



UNIVERSITAT DE
BARCELONA

Genètica i epigenètica de les addiccions i trastorns comòrbids

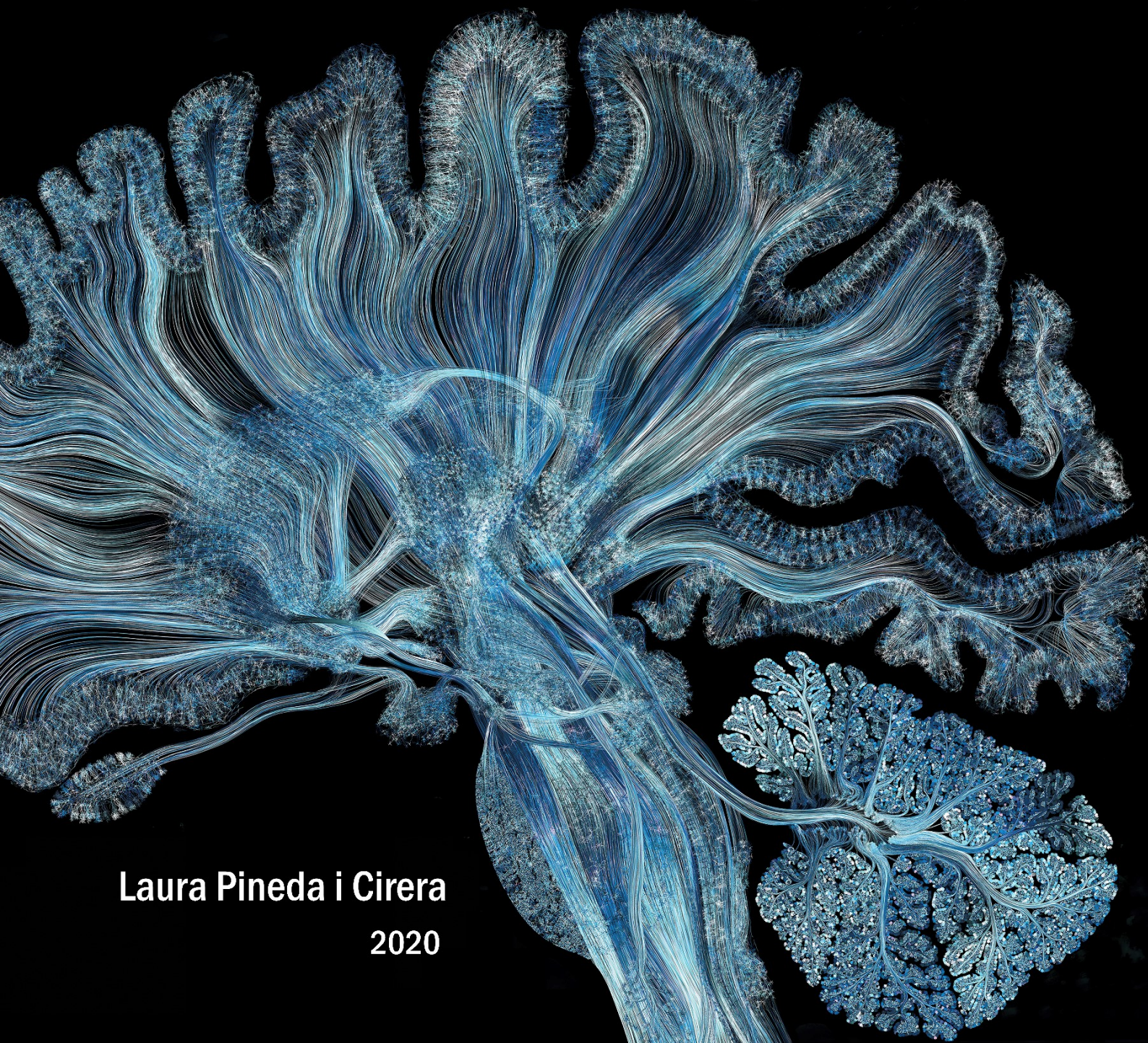
Laura Pineda i Cirera

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Genètica i epigenètica de les addiccions i trastorns comòrbids



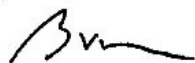
Laura Pineda i Cirera
2020

Genètica i epigenètica de les addiccions i trastorns comòrbids

Memòria presentada per
Laura Pineda i Cirera

Per optar al grau de
Doctora per la Universitat de Barcelona
Programa de Genètica
Departament de Genètica, Microbiologia i Estadística
Facultat de Biologia, Universitat de Barcelona

Tesi doctoral “Genètica i epigenètica de les addiccions i trastorns comòrbids”
amb la direcció i tutoria del **Dr. Bru Cormand Rifà** (codirector i tutor) i la
Dra. Noèlia Fernàndez Castillo (codirectora)



Dr. Bru Cormand Rifà
Codirector/Tutor



Dra. Noèlia Fernàndez Castillo
Codirectora



Laura Pineda i Cirera
Barcelona, Juny del 2020



UNIVERSITAT DE
BARCELONA

AGRAÏMENTS

Aquestes paraules les escric per donar les gràcies a tota aquella gent que directa o indirectament ha posat un granet de sorra en aquesta Tesi, començant pel **Bru** i a la **Noe**, els directors d'aquesta Tesi. Gràcies per donar-me l'oportunitat de realitzar la Tesi en el vostre grup i guiar-me en aquest camí. Tots aquests anys m'heu ajudat a créixer no només com a científica sinó també com a persona, i no tinc cap dubte que no hauria pogut triar un grup millor on anar. A la **Judit**, la meva companya de drogues, amb la que he treballat colze a colze durant tots aquests anys. Sempre he admirat la teva capacitat i estic segura que arribaràs fins on et proposis i tingues clar que aquesta Tesi no seria el que és sense tu, n'estic segura. A la **Bàrbara**, per estar amb mi els primers dies del laboratori i fer que el grup estigui tant unit com està. To **Anu**, thanks for your effort, we finally got it! To **Maja**, you deserve everything, and I hope you finally get the PhD you want here in Barcelona. I finalment, a la **Rous**, una gran científica amb qui sempre pots comptar i qui sempre et donarà bons consells. La teva naturalitat fa que de seguida es pugui confiar amb tu. Tots vosaltres heu fet que aquesta tesi sigui possible, gràcies.

També voldria donar les gràcies a la gent del lab, gràcies a tots aquells que heu passat i passareu durant els pròxims anys per aquest laboratori, us desitjo el millor. A la **Neus**, per moltes més caminades per Barcelona de nit parlant de tot i de res, posant-nos al dia, m'encanten. A la **Noe B**, una noia decidida, amb les idees clares i amb qui vaig començar aquesta aventura del doctorat. A la **Núria**, una peça clau al laboratori, sempre donant alegria. Sembla impossible però al final sempre seguim el mateix camí, mateixa carrera, mateix màster, el doctorat al mateix laboratori i sembla que l'any que ve potser tornarem a estar juntes hehe. A la **Laura C**, em va encantar poder anar amb tu a València al curs durant aquell Sant Jordi. A l'**Ester**, la següent doctora del grup i el granet de sorra que ens faltava als Neurogenetic Angels. A la **Moni**, és una alegria arribar al laboratori cada matí i trobar-t'hi. Sense tu el laboratori no seria el que és. I a l'**Eva**, espero que aconseguis quedar-te al grup per fer la Tesi i que la gaudeixis com ningú. Gràcies també a la **Susanna**, al **Dani**, a la **Lluïsa**, a la **Gemma**, a la **Roser Urreizti** i a la **Kelly**, per aportar sempre idees a tots els projectes. Tenir-vos a tots al laboratori durant aquests anys ha sigut genial.

To **David** and **Aet**, thanks for giving me the opportunity to work with you in Frankfurt. I felt at home there with you and I have had one of my best experiences that I wouldn't mind repeating over and over again. Thanks a lot! **Vanessa**, you are the nicest person that I have never met, I hope you get what you want because you are worth much. **Sofia**, the first person that I met to the lab, we started working together and I will always remember it. Finally, **Jessica**, **Dilhan** and **Anna**, thanks for being there with me, you are also a part of this.

Al grup de la UPF, **Rafa**, **Elena**, **Laura** i **Èric**, gràcies a vosaltres hem pogut arribar molt lluny i encara queda molt camí per recórrer. I també a les noies de Vall d'Hebron, **Marta**, **Cris**, **Paula**, **Iris** i **Laura**, perquè també formeu part d'aquesta Tesi.

Al **Francesc Mestres**, per ensenyar-me com de gran pot ser la Genètica i que, juntament amb la **Conxita** i al **Joan Balanyà**, sou una família per mi dins de la Universitat que sempre heu estat allà quan us he necessitat i que m'heu cuidat com ningú.

A la **Núria Giralte**, al **Jordi Camprubí**, **Lo Pere Franch** i a l'**Imma Sansó**, professors que, durant l'ESO i el Batxillerat, han fet que estimi la Química i la Biologia i els quals van fer que em decidís a estudiar Bioquímica.

També a **Greg Dunn** (Greg Dunn Neuro Art) per proporcionar-me la imatge de la portada amb la que sempre identificaré la meva Tesi.

I ara aquells que no han ajudat des de dins però sí des de fora, començant pel **Josep** i l'**Ale**, per tots els matins esmorçant abans d'entrar al lab, primer a la cafeteria de Biologia amb els sucus, després a farmàcia i finalment al Penyafort, amb el nostre cafè amb el croissant (vinga va lamentable!!). Realment els matins amb vosaltres feien que el dia fos molt més amè, no podria haver tingut uns millors companys. A tu **Josep**, per ser com ets, un sol de persona i apassionat amb tot el que fas. I a tu **Ale**, per venir a veurem cada cop que t'ho demanava, per escoltar-me, per acompanyar-me i per tot el que m'has donat aquests anys, l'inoblidable viatge als EEUU, el teu país somiat, la tranquil·litat que vam viure a Cambodja i escapades de cap de setmana. La nostra amistat seguirà així tota la vida, n'estic completament segura. Desitjo que aconseguis trobar el teu postdoc als EEUU, t'ho mereixes.

Al **Cortés** i a l'**Ibars** (els **Alberts**), els companys de pis perfectes, així que uns fifes per aquí o un sushi per allà, això sí, no t'afaitarem no Cortés? Hehe. Estem apunt d'acabar aquesta experiència però amb ningú hauria pogut ser millor. Gràcies per estar amb mi i formar part d'això, per haver estat aquí durant l'escriptura i òbviament durant el confinament. Tot i que sempre m'estigüeu burxant, us trobaré molt a faltar. I també a la **Neus** i a la **Carme**, per venir a casa a fer pis i compartir temps amb nosaltres o aprofitar i anar a fer alguna caminada dels Gafarrons oi **Neus**?

Als amics de sempre, **LACROSSA&CO**, des dels 3 anys creixent junts i amb mil històries compartides. Els vespres al 100Montaditos o a l'Snack55, els sopars a casa, els concerts, les barbacoes, les visites en els diferents Erasmus, els viatges d'estiu, tot això només és una petita part del que hem viscut i el que viurem. Realment sempre serem una gran família, sou genials i sense vosaltres tot aquest camí no hagués sigut el mateix. A l'**Anna**, tota la vida inseparables encara que agafem camins diferents, si vam aconseguir formar un equip de futbol quan teníem 12 anys penjant cartells per Barcelona, estic convençuda que podem fer tot el que ens proposem. Gràcies per tot el que fas i per deixar-me formar part de la teva família, ja saps que per mi sou tots molt importants, i que tinc un especial "carinyo" als teus avis i a Ca l'Avi amb les veremes. A la **Carlota**, ets de les persones més persistents que conec i també de les més indecises hehe, ets un sol de persona i per sort sé que sempre estaràs aquí. A la **Rosa**, moltes gràcies per preguntar-me com anava amb la Tesi o escriure'm quan feia molt temps que no sabies res de mi, encara que jo no ho fagi tant. Estic segura que encara ens queden molts calendaris Kinder per compartir. A l'**Oriol**, el senyor Brandy haha, gran periodista però millor persona. Em vas salvar aquell dia a Frankfurt al mercat nadalenc, menys mal que estaves amb mi aguantant el tipus. A la **Sabina**, la persona més autèntica que conec, gallines, ratafia, formatge... sempre fent coses sorprenents que poca altra gent intentaria fer pel seu compte. Per quan el Lovaina part IV? A la **Silvia**, la meva veïna, amb tu tornava totes les nits, amb tu he estat estudiant a la biblioteca tota la carrera, amb tu (i la Sabina) ens n'hem anat al cul de Catalunya a concerts de Brams i molts altres grups. Estic molt contenta de tot el que hem fet i espero que segueixi així. A l'**Albert**, sí altre cop tu, perquè estàs fins la sopa ja ho sabem, hehe, però m'encanta. I finalment, als **Marçals**, per donar tanta vida al grup i seguir amb nosaltres des de fa tant de temps.

Als **coronabirres**, tots vosaltres també heu estat presents, i heu fet que tot passi més ràpid, sobretot el confinament, perquè ja és dir acabar una Tesi confinada a casa. Fa molt temps que ens coneixem i estic segura que seguirem així sempre, sent nosaltres mateixos, no es pot desitjar res millor. A l'**Estelita**, una de les persones més importants per mi amb qui he compartit mil experiències. Amb tu vaig començar la Universitat (bé, començar no, que la primera classe et vas equivocar d'aula) i he viscut les primeres experiències en un laboratori. Estic molt contenta que estiguis tant present, per tot el que hem viscut i per tot el que viurem. Al **Pich**, perquè gràcies a tu i a la teva insistència seguim vivint la vida com quan estàvem al Batxillerat. Ets una persona molt especial i amb molt talent, estic impacient per veure on arribes. Al **Pau**, admiro que hagi sigut capaç d'emprendre l'aventura del doctorat a l'estranger i que hagi sigut capaç de

marxar durant 4 anys per fer el que volies, molts del grup serem doctors, però en el teu cas amb molt més mèrit. A l'**Aina** i al **Marc**, sé que sempre podré comptar amb vosaltres pel que sigui tal i com he pogut veure tots aquests anys. I finalment, a la **Judit** i al **Xavi**, dos peces clau del grup, poseu el granet de sorra que falta perquè estiguem del tot complets.

Als **Bioqninges**, als que estan fent el doctorat, i els que han decidit prendre altres camins. A l'**Elena** i al **Pau**, qui ho hauria de dir que formariem aquest magnífic grup, estic molt contenta de que sigueu amics meus. **Pau**, mai ningú m'havia sorprès tant com tu, ets un tot terreny. Sempre recordaré totes les tornades amb bici l'acabar les classes i les tardes al 100Montaditos, estic molt contenta que mai deixéssim de veure'ns. **Elena**, ets una noia espectacular, i encara que facis veure que t'enfades cada cop que juguem, sé que en el fons t'encanta que et burxem. I ja ho tenim això, perquè finalment dipositem a la vegada! Per cert! Prepara't, que el Pau i jo serem els super tiets dels teus nens, els mimarem com ningú i els ensenyarem a jugar als nostres jocs de taula hehe. Al **Dalili**, una escapadeta a Milà per veure't sempre va bé per descansar del doctorat, o a ratejar una mica per Sri Lanka també, perquè enganyar-nos (What is this?, kalakalakelekelekilikili). Al **corqui** i la **Marta**, per tots els pollos, scape rooms i viatgets a Mallorca, per quant el següent? Al **Negre**, per donar-me "vidilla", explicar-me mil historietes i estar sempre amb mi encara que no estiguéssim a la mateixa ciutat. Al **Borja**, per tots els concerts, viatges i experiències. Et desitjo el millor en aquesta nova etapa a Barcelona ara que ja ets doctor. A la **Cris** per ser l'alegria de la huerta. A la **Sam**, amb qui també he compartit molts moments al llarg de la carrera i el doctorat. I també a la **Carol**, per formar part d'aquesta família.

De Frankfurt, agrair a l'**Edu** i la **Montse** que ens fessin sentir com a casa i per ajudar-nos tant. Sempre recordaré el primer dia que ens vaucollir a l'aeroport de Frankfurt i ens vaucollir a casa vostre així com les visites i les sorpreses! To **Mike**, **Manuela** and **Sannie** for trusting us, you have been more than our landlords, we will be happy to show you Barcelona if you decide to come. I finalment a la **Sandra** i a l'**Albert**, espero que gaudiu molt de l'experiència igual que vam fer nosaltres.

Als pandalovers, **Maria** i **Albert**, des del moment que us vaig veure per primer cop, vaig sentir com si us conegués de tota la vida, sou genials. Gràcies per fer més amè el confinament i escriptura de la tesi amb les minivisites i els passejos. Ens esperen molts més plans gourmet i viatges.

A l'**Arantxa**, que tot i veure'ns poc, sé que sempre estaràs amb mi, ets la persona més treballadora i decidida que conec. No podia trobar millor companyia durant els anys al Conservatori.

A les nenes de futbol, les **BELLES GLÒRIES**, un partidet sempre ve bé per descansar i alliberar endorfines, encara que sigui a un quart de dotze de la nit haha.

A la família del Francesc, **Marta**, **Joan** i **Ricard**, per acollir-me com una més a casa vostra, amb les fideuàs els diumenges i el que sigui necessari.

A la família: **als pares**, **germans**, **tiets**, **cosins** i **els nens**. Per tots els dinars, sopars i sortides i els caps de setmana a Prats de Lluçanès. En especial a l'**Anna** i l'**Andreu** per explicar-me com funciona tot aquest món, al **Jordi** per la companyia en els dinars a casa o a Pidiscat quan estava de baixa, discutint de què es pot canviar del plat perquè encara sigui més bo o tastant nous cafès, a l'**Elisenda** per ser la meua companya de ciències de la família i **als pares**, perquè mai heu dubtat que puguem fer el que ens proposem.

I finalment al **Francesc**, per estar amb mi tots aquests mesos i fer que tot sigui fàcil, per recordar-me que és el que importa i recolzar-me en aquest final de Tesis confinada. Sempre dic que Frankfurt és una de les millors experiències de la meua vida, i si això és així, és perquè la vaig compartir amb tu i tota la family hehe.

Moltes gràcies a tots per fer possible aquesta Tesi i per acompanyar-me en tot aquest camí. Aquesta Tesi no és només meua, també és de tots vosaltres.

ABSTRACT

Addiction is a complex neuropsychiatric disorder characterized by compulsive and uncontrolled use of an addictive stimulus such as drugs of abuse or food. Its etiology involves genetic and environmental factors, with epigenetics as the interplay between them. Although family and twin studies have estimated a high contribution of genetics to the disease, the underlying mechanisms remain largely unknown. In addition, multiple studies have shown that addiction is highly comorbid with other psychiatric disorders, which can be explained, at least in part, by shared genetics.

The work presented in this Doctoral Thesis has contributed to the study and identification of genetic and epigenetic factors underlying drug and food addiction, as well as comorbid disorders. First, we have replicated a variant previously identified in a genome wide association study for drug dependence and we have discarded that *NFAT5*, a candidate gene for cocaine dependence, modulates dopamine synthesis by regulating *DDC* expression in dopaminergic neural cell lines. Subsequently, we focused on the identification of genes and epigenetic risk factors for drug and food addiction. We identified three variants associated with drug addiction that alter the methylation of CpG sites or the binding of miRNAs, pinpointing *CTNBL1*, *SCP2*, *ECHDC2* and *PLCB1* as candidate genes for the disorder. Furthermore, we have identified several mRNAs and miRNAs that are differentially expressed in two brain areas of mice with extreme behaviours towards food addiction-like phenotype (vulnerable and resilient), some of these changes explained by promoter methylation alterations. Then, we have identified multiple variants affecting the methylation of CpG sites associated with ADHD. Some of these variants have an impact on the expression of three genes, *ARTN*, *PIDD1* and *C2orf82*, pointing them as candidates to contribute to the disorder. For the *C2orf82* gene, we have explored its expression in the mouse brain, obtaining the highest expression in the nucleus accumbens, and we have also phenotyped a knockout mouse model, observing hypolocomotion, impulsivity and lower cocaine motivation. Finally, we have identified multiple variants that affect brain methylation and expression of eight genes with pleiotropic effects on several psychiatric disorders.

ÍNDEX

ABREVIATURES

INTRODUCCIÓ	1
1. ADDICCIÓ A LES DROGUES I AL MENJAR	3
1.1. EPIDEMIOLOGIA DEL CONSUM DE DROGUES I PREVALENÇA DE LES ADDICCIONS	3
1.2. DIAGNÒSTIC	5
1.3. NEUROBIOLOGIA	7
2. TRASTORNS COMÒRBIDS	10
Trastorn per dèficit d'atenció amb hiperactivitat (TDAH)	12
3. GENÈTICA DE LES ADDICCIONS I TRASTORNS COMÒRBIDS	15
3.1. HERETABILITAT	16
3.2. ESTUDIS D'ASSOCIACIÓ GENÈTICA	17
3.2.1. Estudis d'associació de gens i sistemes genètics candidats	18
3.2.2. Estudis d'associació a escala genòmica	20
3.2.3. Model omnigènec	22
4. EPIGENÈTICA DE LES ADDICCIONS I TRASTORNS COMÒRBIDS	23
4.1. MODIFICACIONS POSTTRADUCCIONALS D'HISTONES	24
4.2. METILACIÓ DEL DNA	26
Metilació específica d'al·lel	29
4.3. MICRORNAS	30
5. MODELS ANIMALS	33
5.1. MODELS PER L'ESTUDI DE LES ADDICCIONS	34
5.2. MODELS PER L'ESTUDI DEL TDAH	37
OBJECTIUS	41
RESULTATS	45
INFORME DELS DIRECTORS	47
CAPÍTOL 1. Estudi de factors genètics de risc a l'addicció a drogues identificats prèviament.	51
ARTICLE 1: "AVALUACIÓ DE VARIANTS PRÈVIAMENT ASSOCIADES A ESCALA GENÒMICA A LA DEPENDÈNCIA DE DROGUES EN UNA MOSTRA ESPANYOLA"	51
ARTICLE 2: "NFAT5 (TONEBP) NO REGULA L'EXPRESSIÓ DE LA DDC EN LÍNIES CEL·LULARS NEURONALS DOPAMINÈRGiques"	63

CAPÍTOL 2. Identificació de factors epigenètics de susceptibilitat a l'addicció a les drogues i al menjar.	75
ARTICLE 3: "ESTUDI DE LA METILACIÓ ESPECÍFICA D'AL·LEL EN LA SUSCEPTIBILITAT A LA DEPENDÈNCIA DE DROGUES"	75
ARTICLE 4: "ASSOCIACIÓ DEL GEN PLCB1 AMB LA DEPENDÈNCIA DE DROGUES"	103
ARTICLE 5: "EXPRESSIÓ GÈNICA I PERFIL EPIGENÈTIC D'UN MODEL DE RATOLÍ D'ADDICCIÓ AL MENJAR"	117
CAPÍTOL 3. Identificació de factors epigenètics de susceptibilitat a trastorns comòrbids a l'addicció a les drogues.	165
ARTICLE 6: "ESTUDI DE VARIANTS GENÈTIQUES QUE TENEN UN EFECTE EN ELS NIVELLS DE METILACIÓ DEL CERVELL EN EL TRASTORN PER DÈFICIT D'ATENCIÓ AMB HIPERACTIVITAT"	165
ANNEX DE L'ARTICLE 6: CONTRIBUTIÓ DEL GEN C2ORF82 AL TDAH MITJANÇANT LA UTILITZACIÓ D'UN MODEL GENOANUL·LAT MURÍ.	207
ARTICLE 7: IDENTIFICATION OF GENETIC VARIANTS INFLUENCING METHYLATION IN BRAIN WITH PLEIOTROPIC EFFECTS ON PSYCHIATRIC DISORDERS	239
DISCUSSIÓ	299
<hr/>	
1. CONSIDERACIONS SOBRE LES METODOLOGIES EMPRADES	302
1.1. ESTUDIS D'ASSOCIACIÓ CAS-CONTROL POBLACIONALS	302
1.2. ESTUDIS D'ASSOCIACIÓ A ESCALA GENÒMICA (GWAS)	308
1.3. ESTUDIS TRANSCRIPTÒMICS I METILÒMICS	309
1.4. MODELS ANIMALS	310
2. IDENTIFICACIÓ DE FACTORS EPIGENÈTICS DE SUSCEPTIBILITAT	312
2.1. VARIANTS GENÈTIQUES QUE TENEN INFLUÈNCIA SOBRE LA METILACIÓ DEL DNA	312
2.2. VARIANTS GENÈTIQUES EN LLOCS D'UNIÓ DE MIRNAS	315
2.3. ESTUDI TRANSCRIPTÒMIC EN UN MODEL D'ADDICCIÓ AL MENJAR	316
2.4. GENS I FACTORS EPIGENÈTICS DE RISC IDENTIFICATS EN ELS NOSTRES ESTUDIS	318
3. FACTORS GENÈTICS DE RISC A L'ADDICCIÓ A DROGUES I TRASTORNS COMÒRBIDS I ESTUDIS FUNCIONALS	320
4. PERSPECTIVES DE FUTUR	325
CONCLUSIONS	329
<hr/>	
BIBLIOGRAFIA	333
ANNEX	359
<hr/>	

ABREVIATURES

3'UTR: *3'-untranslated region*

5'UTR: *5'-untranslated region*

AMY: Amígdala

ASM: Metilació específica d'al·lel, *Allele-Specific Methylation*

CB: Cerebel

CDCV: Malaltia comuna- variant comuna, *common disease - common variant*

CDRV: Malaltia comuna- variant rara, *common disease - rare variant*

COMT: Catecol-O-metil transferasa

CPP: Preferència de lloc condicionada, *conditioned place preference*

CPT: *Continuous Performance Task*

CR: *Correct rejection*

D₁₋₅: Receptor de dopamina 1-5

DA: Dopamina

DAT: Transportador monoaminèrgic de dopamina

DDC (o AADC): Decarboxilasa d'aminoàcids aromàtics

DMR: Regions diferencialment metilades, *Differentially Methylated Regions*

DNMT: Metiltransferasa del DNA

DS: Nucli estriat dorsal

DSM: Manual Diagnòstic i Estadístic de Trastorns Mentals, *Diagnostic and Statistical Manual of Mental Disorders*

EPM: *Elevated Plus Maze*

FAR: *False Alarm Rate*

FC: *Fear Conditioning*

FDR: *False Discovery Rate*

FR: Test de raó fixa, *fixed ratio*

GABA: Àcid γ -aminobutíric

GEC: *Genetic Type I Error Calculator*

GWAS: Estudi d'associació a escala genòmica, *genome-wide association study*

HDAC: Desacetilasa d'histones

HWE: Equilibri Hardy-Weinberg

HPC: Hipocamp

HET: Heterozigot

HT: Hipotàlem

HPC: Hipocamp

HR: *Hit Rate*

ICD: Classificació Internacional de Malalties, *International Classification of Diseases*

ID: Demora en la ingesta, *ingestion delay*

IMC: Índex de massa corporal

ITI: Interval entre intents, *Inter-Trial Interval*

KO: Genoanul·lat, *gene Knock-Out*

LD: Desequilibri de lligament

LDB: *Light Dark box*

LH: Retenció limitada, *Limited Hold*

MAF: Freqüència de l'al·lel menor

MeCP2: Proteïna d'unió 2 a grups metil CpG, *methyl-CpG binding protein 2*

miRNA: microRNA

MBD-seq: *Methyl Binding Domain sequencing*

mPFC: Escorça prefrontal medial

MPT: Modificació posttraduccional

NAC: Nucli accumbens

NOR: *Novel Object Recognition*

OF: *Open field*

OMS: Organització Mundial de la Salut

PFC: Escorça prefrontal

PGC: *Psychiatric Genomics Consortium*

PPI: *Prepulse Inhibition test*

PR: Test de raó progressiva, *progressive ratio*

RNAseq: Seqüenciació massiva de RNA

S-: Estímul negatiu

S+: Estímul positiu

SC: Medul·la espinal

SCID: Entrevista estandarditzada *Structured Clinical Interview*

SD: Duració de l'estímul, *Stimulus Duration*

SHR: Rata espontàniament hipertensa

smallRNAseq: Seqüenciació massiva de smallRNA

SNC: Sistema nerviós central

SNP: Canvi d'un sol nucleòtid, *single nucleotide polymorphism*

TDH: Trastorn per dèficit d'atenció amb hiperactivitat

TO: *Time-out*

TSS: Inici de transcripció, *Transcription Start Site*

UNODC: Oficina de les Nacions Unides Contra la Droga i els Delictes

VTA: Àrea tegmental ventral

YFAS: *Yale Food Addiction Scale*

WT: Soca salvatge, *wild-type*

INTRODUCCIÓ

1. Addicció a les drogues i al menjar

L'addicció és un trastorn neuropsiquiàtric complex caracteritzat per comportaments maladaptatius i destructius, causats per un estímul addictiu, que pot ser una droga o un comportament (com per exemple l'addicció al menjar, al joc, a l'exercici, a internet, a comprar o al sexe). En aquesta Tesi ens centrem en l'addicció a les drogues en general, en l'addicció a la cocaïna en particular i també en l'addicció al menjar.

L'addicció consisteix en l'ús d'una substància o l'execució d'una conducta de forma persistent, compulsiva i incontrolada. L'estímul addictiu produeix una sensació de plaer, reforç i gratificació que desplaça les recompenses naturals de l'organisme, tot afavorint la repetició de la conducta o el consum reiterat de la substància i provocant canvis adaptatius al sistema nerviós central (SNC). Aquests canvis n'alteren el funcionament i donen lloc a la simptomatologia associada a l'addicció: la tolerància (cada cop cal un estímul de major magnitud per aconseguir la mateixa recompensa), la pèrdua de control, l'establiment del deler o *craving* (desig intens d'experimentar els efectes de l'estímul addictiu), l'abstinència i els canvis emocionals i anímics que persisteixen fins molt de temps després que cesi el consum o la conducta. Els canvis neurobiològics es mantenen, són progressius i són responsables de la cronificació de l'addicció i l'aparició de recaigudes.

L'addicció al menjar es caracteritza pel consum compulsiu d'aliments amb cert poder addictiu (aliments saborosos), que solen ser rics en greixos o en sucres, capaços d'activar el sistema de recompensa en humans de forma similar a les drogues d'abús. La idea de menjars amb poder addictiu va sorgir a finals del segle XIX, però no ha estat fins a principis del segle XXI quan s'ha començat a parlar de l'addicció al menjar (Meule, 2019). Tot i així, encara hi ha certa controvèrsia en la seva classificació com a una addicció a un comportament, ja que en ella també hi intervenen propietats addictives intrínseques dels menjars saborosos similars a les de les drogues d'abús (Gordon *et al.*, 2018; Meule, 2019).

1.1. Epidemiologia del consum de drogues i prevalença de les addiccions

El consum i abús de les drogues comporta problemes tant a nivell individual com a escala global. A nivell individual la droga afecta l'entorn familiar, les relacions interpersonals i laborals i la salut del consumidor. A nivell global, a causa de la seva elevada prevalença, el consum de drogues té un gran impacte socio-sanitari, econòmic i polític i fomenta la delinqüència, la inestabilitat i la inseguretat ciutadana.

Gran part de la població consumeix drogues lícites (com l'alcohol o el tabac), però també un percentatge considerable, que ha incrementat en els darrers anys, en consumeix d'il·lícites. L'any 2017, l'Oficina de les Nacions Unides contra les Drogues i el Crim (UNODC) va estimar que 271 milions de persones d'entre 15 i 64 anys (5,5% de la població mundial d'aquest grup d'edat) havien

consumit alguna droga il·lícita almenys una vegada durant l'últim any (UNODC, 2019) (Taula 1). Aquesta xifra suposa un increment del 30% comparat amb el 2009, quan 210 milions de persones havien consumit alguna droga il·lícita almenys una vegada durant l'últim any. Tot i que aquest augment era degut en part a un creixement del 10% de la població global de 15 a 64 anys, les dades actuals mostren una major prevalença del consum a nivell mundial (UNODC, 2019). A més, també va haver-hi un augment del 15% en el nombre d'individus diagnosticats amb trastorns per ús de drogues il·lícites, situant-se en 35,3 milions de persones (0,7% de la població mundial), i un increment de morts a causa de les drogues il·lícites, que s'ha situat en 585.000 l'any 2017 (UNODC, 2019).

Si ens referim a les drogues lícites, en aquesta Tesi ens hem centrat només en l'alcohol. L'any 2016 l'Organització Mundial de la Salut (OMS) va comptabilitzar que més de la meitat de la població mundial de més de 15 anys (57% o 3.100 milions de persones) havia consumit alcohol en els últims dotze mesos (Taula 1). D'aquests, aproximadament 2.300 milions de persones eren bevedors habituals, i si ens fixem en els adolescents (15-19 anys), 155 milions (26,5%) són bevedors habituals. Per altra banda, el consum d'alcohol anual va passar dels 5,5 litres d'alcohol pur el 2005 a 6,4 litres el 2010, nivell que s'ha mantingut estable fins el 2016 (World Health Organization, 2018). Finalment, pel que fa a aquelles persones diagnosticades amb un trastorn per ús d'alcohol, l'any 2016 aquest número se situava en 283 milions (5,1% de la població mundial de més de 15 anys) (World Health Organization, 2018).

Taula 1. Prevalença anual i número de consumidors de drogues il·lícites d'entre 15 i 64 anys a nivell mundial l'any 2017 i de consumidors d'alcohol de més de 15 anys a nivell mundial.

Droga	Prevalença (%)	Número (milions)
Alcohol	57	3.100
Cànnabis	3,8	188
Opioides	1,08	53,4
Opiacis	0,59	29,2
Estimulants de tipus amfetamínic	0,59	28,9
Èxtasi	0,41	21,3
Cocaïna	0,37	18,1

En relació a l'addicció al menjar, la prevalença a la població general o entre els estudiants se situa entre el 5 i el 10% (Gearhardt, Corbin i Brownell, 2009; Meule, 2011; Meule i Gearhardt, 2014), sent per tant més alta que les calculades pels trastorns per ús de drogues il·lícites o d'alcohol. Resulta interessant destacar que s'ha descrit una relació no lineal (similar a una corba en J, on primer trobem un petit descens i a continuació un increment pronunciat) entre l'addicció al menjar i l'índex de massa corporal (IMC), amb prevalences més elevades en les persones amb IMC

extremes (baix pes o obesitat) (Figura 1) (Meule, 2011; Hauck *et al.*, 2017). Aquests estudis doncs, assenyalen la importància d'entendre el paper de l'addicció al menjar tant en trastorns de la conducta alimentària (on incloem la bulímia nerviosa i el trastorn per afartament) com en l'obesitat (Meule, 2011; Hauck *et al.*, 2017).

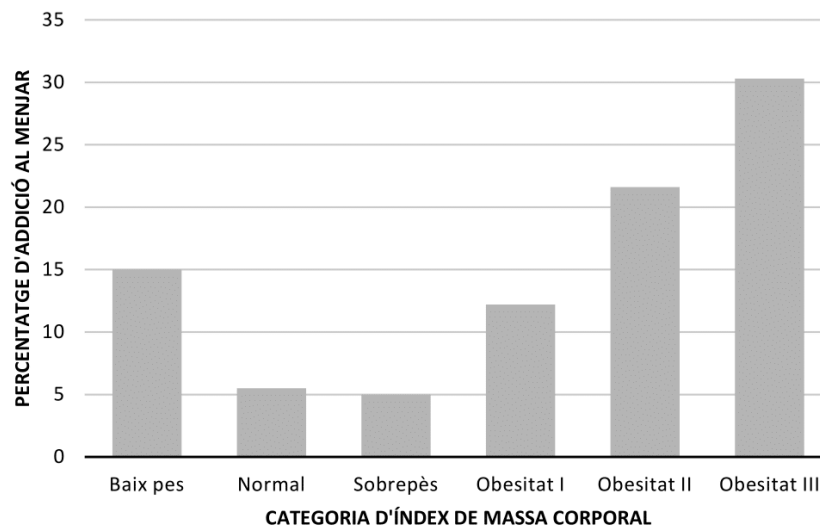


Figura 1. Prevalença de l'addicció al menjar segons l'índex de massa corporal diagnosticat segons YFAS2.0. Adaptada de Hauck et al., 2017.

1.2. Diagnòstic

L'ús o el consum de drogues és un terme que engloba diferents tipus de conductes que van des del consum esporàdic, passant per l'abús i fins la dependència. El diagnòstic clínic diferencial entre aquestes conductes es fa en base als criteris descrits en el Manual Diagnòstic i Estadístic de Trastorns Mentals de l'Associació Americana de Psiquiatria (*Diagnostic and Statistical Manual of Mental Disorders; DSM*), la versió vigent del qual és el DSM V, tot i que en molts centres se segueix utilitzant la quarta edició revisada del manual (DSM-IV-TR).

El DSM conté una classificació dels trastorns mentals i proporciona descripcions clares de les categories diagnòstiques. La versió DSM-V buscava solucionar els problemes identificats a la versió DSM-IV-TR, per poder oferir una aproximació millorada del diagnòstic dels trastorns per ús de substàncies (American Psychiatric Association, 2013). El primer gran canvi que va introduir el DSM-V va ser la unificació dels criteris d'abús i dependència en una única categoria, els trastorns per ús de substàncies, amb dos canvis en els criteris diagnòstics utilitzats fins el moment: eliminació dels problemes legals deguts a l'ús de substàncies i incorporació del deler (o *craving*) (Taula 2). Un altre canvi va ser la incorporació de nous l·lindars indicadors de severitat. En el DSM-IV-TR es considerava abús si es complien un o més criteris durant un any, sempre i quan el

Introducció

síntomes no assolissin els criteris de dependència. La dependència, en canvi, es diagnosticava quan es complien tres o més criteris. En la nova versió (DSM-V) ja no es fa aquesta distinció i es diagnostiquen tres graus de trastorns per ús de substàncies: de dos a tres criteris es considera addicció lleu, de quatre a cinc addicció moderada i a partir de sis criteris addicció severa (Hasin *et al.*, 2013).

Taula 2. Criteris diagnòstics dels trastorns per ús de substàncies inclosos en els manuals de diagnòstic DSM-IV-TR i DSM-V. Adaptada de Hasin *et al.*, 2013.

	DSM-IV-TR (abús) ¹	DSM-IV-TR (dependència) ²	DSM-V ³
Incompliment d'obligacions per l'ús recurrent	✓		✓
Ús recurrent en situacions perilloses	✓		✓
Problemes legals per l'ús recurrent	✓		
Problemes socials i interpersonalment per l'ús recurrent	✓		✓
Tolerància ⁴		✓	✓
Abstinència ⁵		✓	✓
Consum de grans quantitats o al llarg de molt temps		✓	✓
Intents repetits d'aturar o controlar l'ús		✓	✓
Molt temps dedicat a l'ús		✓	✓
Abandonament d'activitats socials o laborals degut al consum		✓	✓
Deler (<i>craving</i>)			✓

¹L'abús es defineix quan es compleixen un o més criteris dins d'un període d'un any i no hi ha diagnòstic de dependència. ²La dependència es defineix quan es compleixen tres o més criteris dins d'un període d'un any. ³L'addicció es defineix si es compleixen dos criteris o més dins d'un període d'un any, amb diferents graus de dependència. ⁴La tolerància és definida per a) la necessitat d'incrementar el consum per aconseguir l'efecte desitjat o b) quan l'efecte d'una mateixa quantitat de droga disminueix amb l'ús continuat. ⁵L'abstinència és definida per a) la síndrome d'abstinència típica de cada substància o b) el consum de la substància amb l'objectiu de reduir o evitar els símptomes de l'abstinència.

Un altre manual utilitzat per al diagnòstic de les addiccions és la Classificació Internacional de Malalties (*International Classification of Diseases; ICD*), desenvolupat per l'OMS, sent la versió ICD-11 la més recent. L'ICD, a diferència del DSM-V, conserva encara la dependència de substàncies com a diagnòstic principal. Dins del diagnòstic de la dependència, hi ha criteris similars al DSM-V, com la tolerància, l'abstinència, la persistència en l'ús malgrat conèixer les conseqüències negatives que això comporta, i altres indicadors de pèrdua de control. Tot i que hi ha diferències entre els dos manuals, ambdós permeten classificar de forma fiable la dependència d'una substància.

L'addicció al menjar té moltes similituds amb l'addicció a les drogues, ja que ambdues mostren patrons de comportament similars, que queden recollits en els criteris diagnòstics de la dependència de drogues: tolerància, abstinència, consum durant més temps o en més quantitat

del previst, intents repetits per controlar o aturar el consum, molt de temps dedicat a l'obtenció o consum, abandonament d'activitats socials o laborals pel consum o el consum malgrat saber-ne les conseqüències negatives. Tot i les semblances, l'addicció al menjar encara no està inclosa en els manuals de diagnòstic oficials, i és per això que s'utilitza un qüestionari anomenat *Yale Food Addiction Scale* (YFAS). La primera versió d'aquest qüestionari fou publicada l'any 2009 per la *Yale University's Rudd Center for Food Policy and Obesity*. Aquest qüestionari constava de 25 preguntes i es basava en els criteris diagnòstics per la dependència de substàncies descrits en el manual DSM-IV (Gearhardt, Corbin i Brownell, 2009). L'any 2016 es va revisar aquest qüestionari degut a la implementació del manual DSM-V incrementant el número de preguntes fins a 35 i desenvolupant així el YFAS 2.0 (Gearhardt, Corbin i Brownell, 2016).

1.3. Neurobiologia

En un inici, l'estudi de la neurobiologia de les addiccions es va centrar en l'efecte agut de les drogues d'abús al sistema nerviós central (SNC). Aquesta tendència, però, ha anat canviant i actualment l'interès central rau en entendre quins són els mecanismes genètics, epigenètics, cel·lulars i moleculars implicats en la transició de l'ús esporàdic fins a la pèrdua de control sobre la cerca i consum de la droga i la recaiguda crònica fins i tot després d'un període d'abstinència llarg (Koob i Volkow, 2010, 2016).

Les drogues d'abús i els aliments saborosos actuen sobre el sistema de recompensa del cervell, un circuit format per diferents tipus de neurones, entre elles neurones dopaminèrgiques, que es projecten principalment des de l'àrea tegmental ventral (VTA) fins al nucli accumbens (NAc), a l'amígdala (AMY), a l'hipocamp (HPC) i a l'escorça prefrontal (PFC) (Figura 2) (Gardner, 1997; Carter *et al.*, 2016; Cooper, Robison i Mazei-Robison, 2017; Walker i Nestler, 2018). Cadascuna de les àrees juga un paper clau en la recompensa obtinguda a partir de les drogues o del menjar. La VTA participa en el plaer, la recompensa, la cognició i la motivació; el NAc juga un paper clau en la presa de decisions i en el sistema de recompensa; l'HPC i l'AMY estan involucrats en la formació de memòries de la relació entre l'estímul i la recompensa, en el cas de l'HPC en la memòria declarativa, orientativa i espacial i en el de l'AMY en la memòria i processament d'emocions; i finalment, la PFC està implicada en la presa de decisions i en el control inhibitori. Així doncs, el sistema de recompensa està involucrat en la recompensa, la motivació, la presa de decisions i el control inhibitori (Volkow *et al.*, 2013a), tot de processos centrals en el desenvolupament de l'addicció i en el consum excessiu de menjar.

Introducció

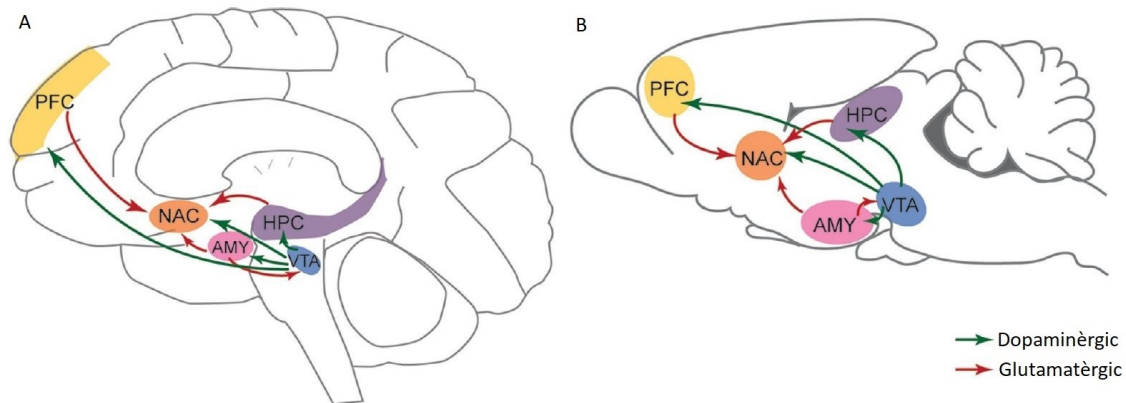


Figura 2. El sistema de recompensa és similar en algunes espècies, com per exemple humans o ratolins. A la imatge es representen les àrees més importants relacionades amb el sistema de recompensa en l'home (A) i en el ratolí (B): les neurones dopaminèrgiques (en verd) de l'àrea tegmental ventral (VTA) tenen projeccions cap al nucli accumbens (NAC), a l'escorça prefrontal (PFC), a l'amígdala (AMY) i a l'hipocamp (HPC). El NAC també rep informació de neurones glutamatèrgiques (en vermell) que venen del PFC, AMY i HPC. Tot i que els mecanismes d'acció són específics per a cada substància, totes actuen a través del sistema de recompensa incrementant els nivells de dopamina. Adaptada de Walker i Nestler, 2018.

En el desenvolupament de l'addicció hi participen molts neurotransmissors, entre ells la dopamina (DA), que té un paper fonamental en molts aspectes del comportament, des del control motor a la regulació d'emocions, cognició, conducta, motivació i recompensa, entre d'altres (Björklund i Dunnett, 2007). La síntesi de DA s'inicia amb la conversió de fenilalanina a tirosina mitjançant l'enzim fenilalanina hidroxilasa (PAH). A continuació, l'enzim tirosina hidroxilasa (TH) oxida la tirosina convertint-la en L-dihidrofenilalanina (L-DOPA), pas limitant de la producció de DA, noradrenalina i adrenalina al cervell. La L-DOPA és llavors convertida en DA per la decarboxilasa d'aminoàcids aromàtics (AADC o DDC), que al seu torn és convertida en noradrenalina per la dopamina beta-hidroxilasa (DBH). Hi ha cinc tipus de receptors de DA amb diferents propietats estructurals i farmacològiques. S'agrupen en dues categories: els similars a D_1 , que activen l'adenilat ciclase (receptors D_1 i D_5) i els similars al receptor D_2 , que inhibeixen la formació d'AMPc (receptors D_2 , D_3 i D_4) (Girault i Greengard, 2004). Finalment, el transportador de dopamina DAT s'encarrega de la recaptació de DA de l'espai sinàptic, tot regulant-ne els nivells. Un cop recaptada, la DA pot ser degradada alternativament pels enzims catecol-O-metil transferasa (COMT) o monoamina oxidasa (MAO).

Les drogues d'abús, i també el menjar, activen de forma directa o indirecta l'activitat dopaminèrgica tot incrementant la concentració de DA a l'espai sinàptic (Koob i Le Moal, 2001; Nestler, 2001, 2005; Di Chiara *et al.*, 2004; Wise, 2008; Volkow *et al.*, 2013a). Els receptors D_2 tenen entre 10 i 100 vegades més afinitat per la DA que els receptors D_1 , i per tant són activats a

concentracions de DA molt més petites. En circumstàncies normals, les neurones dopaminèrgiques de la VTA alliberen una dosi petita i estable de DA (Uhl, Koob i Cable, 2019). En canvi, davant d'una resposta a una recompensa extraordinària, com la que s'obté després del consum de drogues d'abús o de menjar saborosos (Volkow *et al.*, 2013a; Volkow i Morales, 2015), es produeix un increment molt ràpid i pronunciat de DA a l'espai sinàptic, activant així els receptors de dopamina D₁, que serien necessaris per tenir una activació completa del sistema de recompensa (Baik, 2013). Els receptors de dopamina D₁ estimulen la recompensa i l'aprenentatge associatiu o condicionament (Uhl, Koob i Cable, 2019), una altra característica clau de l'addicció, ja que permet als individus associar un estímul amb una recompensa o un càstig. A mesura que s'estableix aquest condicionament, les neurones dopaminèrgiques deixen de respondre a l'estímul en si mateix, la droga, i ho fan de forma anticipada en resposta a un estímul neutre condicionat que inicialment no produïa cap resposta (Schultz, 2002). Aquest tipus de condicionament es dona al NAc i a l'amígdala i és el que s'anomena aprenentatge incentiu- condicionat. A més, en aquest procés hi participen l'aprenentatge d'hàbits (nucli caudat i putamen) i la memòria declarativa o explícita (HPC), de forma que qualsevol estímul associat al consum (per exemple, l'ambient en què es consumeix, les persones amb qui s'acostuma a consumir, l'estat d'ànim, etc.) produeix per si sol la motivació per la cerca de la substància i desig de consum (Volkow *et al.*, 2002, 2013a). L'exposició repetida a menjar saborosos també pot donar lloc a aquest aprenentatge en què l'estímul condicionat podria ser l'olor dels aliments (Berridge, 2009). Aquest condicionament està produït per una sèrie d'adaptacions cel·lulars i moleculars que persisteixen després de llargs períodes d'abstinència i segueixen provocant desig de consum en resposta als estímuls condicionats (Volkow, Koob i McLellan, 2016).

La DA és, doncs, fonamental en la motivació per les drogues o el menjar, però no és l'únic neurotransmissor involucrat. En el procés de transició de l'ús a la dependència hi intervenen gran part dels neurotransmissors i neuromoduladors, com ara els pèptids opioides, l'oxitocina, el glutamat, l'àcid γ -aminobutíric (GABA), la serotonina (5-HT), l'acetilcolina (ACh) i el sistema endocannabinoide. Els estímuls addictius actuen sobre tots aquests sistemes tot produint neuroadaptacions que donen lloc a la dependència. El sucre (i probablement també altres aliments saborosos) incrementen directament els nivells endògens d'opioïdes (Volkow, Wang i Baler, 2011) i agonistes i antagonistes d'opioïdes produeixen un increment o una disminució, respectivament, de la ingesta d'aliments saborosos (Stice *et al.*, 2013). El sistema endocannabinoide, per la seva banda, afavoreix un increment del consum d'aliments a través de la modulació de les entrades excitadores i inhibidores a la VTA i al NAc i, indirectament, activant la transmissió de DA al sistema mesolímbic. A més, el sistema endocannabinoide també regula

accions de mediadors en l'hipotàlem (HT) que regulen la sensació de gana (D'Addario *et al.*, 2014). Finalment, el sistema serotoninèrgic està involucrat en la regulació de processos cognitius, en particular el control d'impulsos i també té un paper modulador en la regulació de l'activitat dopaminèrgica neuronal. A més, ha estat una de les dianes terapèutiques principals en la cerca de fàrmacs per la pèrdua de pes (Higgins, Sellers i Fletcher, 2013).

El sistema endocrí també està involucrat en el procés addictiu ja que els pèptids responsables de la regulació de funcions homeostàtiques, com el consum d'aliments, també influeixen en la recompensa de les drogues (Volkow *et al.*, 2013b; Zallar *et al.*, 2017). La grelina és l'hormona encarregada de generar la sensació de gana, i actua sobre les neurones AgRP en l'HT i sobre les neurones dopaminèrgiques de la VTA, per incrementar la ingesta d'aliments, per incrementar la motivació per obtenir menjar i per regular la glucosa en sang (Kamegai *et al.*, 2000; Zallar *et al.*, 2017). Les hormones que suprimeixen la sensació de gana (per exemple la insulina o la leptina) contraposen les accions de la grelina. El nivell d'aquestes hormones augmenta quan hi ha ingesta d'aliments, i tenen la funció d'atenuar el sistema de recompensa i reduir la motivació pel menjar (Labouèbe *et al.*, 2013). Així doncs, l'estat metabòlic té un efecte en el sistema de recompensa, de manera que es fa patent el vincle existent entre els mecanismes homeostàtics a l'HT i el sistema de recompensa.

2. Trastorns comòrbids

Les diferents addiccions són trastorns que presenten comorbiditat entre elles (múltiples drogues) i també amb un gran ventall de trastorns psiquiàtrics. La comorbiditat és la presència concomitant o concurrent, en un pacient, d'una o més malalties o trastorns amb una malaltia o trastorn primari. Definim, doncs, com a trastorns comòrbids aquells trastorns o malalties addicionals al trastorn primari. El fet que dues malalties siguin comòrbides també implica que les malalties interactuen, tot afectant el curs i el pronòstic de totes dues (Ross i Peselow, 2012; Santucci, 2012). Hi ha múltiples trastorns o condicions psiquiàtriques que són comòrbides a la dependència de drogues, com els trastorns d'ansietat, on incloem el trastorn per ansietat generalitzada, el trastorn de pànic i el trastorn per estrès posttraumàtic (Torrens, Gilchrist i Domingo-Salvany, 2011; Magidson *et al.*, 2012; Brady *et al.*, 2013; Vorspan *et al.*, 2015), la depressió i el trastorn bipolar (Pettinati, O'Brien i Dundon, 2013; Grant *et al.*, 2016), el trastorn per dèficit d'atenció amb hiperactivitat (De Alwis *et al.*, 2014; Harstad *et al.*, 2014; van de Glind *et al.*, 2014; Chen *et al.*, 2018) o l'esquizofrènia (Hunt *et al.*, 2018).

Estudis epidemiològics descriuen que aproximadament la meitat de les persones diagnosticades amb un trastorn psiquiàtric desenvoluparan addicció a les drogues en algun moment de la seva vida i viceversa (Ross i Peselow, 2012). També s'han fet alguns estudis sobre la comorbiditat en joves, que suggereixen que el 60% dels adolescents diagnosticats amb addicció a les drogues també compleixen els criteris diagnòstics per com a mínim una altra malaltia mental (Hser *et al.*, 2001).

Els pacients que pateixen més d'un trastorn psiquiàtric tenen pitjor pronòstic, una major gravetat de la malaltia i també una trajectòria pitjor de la malaltia en diferents aspectes, com per exemple un risc més alt de recaiguda, un compliment deficient de la medicació o majors taxes de suïcidi, dificultats legals, problemes mèdics, recaiguda o estrès familiar, entre d'altres. A més, també s'ha vist que si hi ha dificultats en tractar un dels trastorns, sense importar quin, això afecta negativament el tractament adequat dels altres (Ross i Peselow, 2012).

Les vies principals que poden contribuir a la comorbiditat entre l'addicció a les drogues i altres trastorns psiquiàtrics són, per una banda, factors de risc comuns entre ambdós trastorns, i per l'altra, els mateixos trastorns psiquiàtrics que, per la seva naturalesa, poden contribuir a l'ús i addicció a les drogues i viceversa. Quan parlem de factors de risc, ens referim a factors genètics i epigenètics de vulnerabilitat, problemes en àrees cerebrals similars i factors ambientals com ara l'exposició primerenca a l'estrès o al trauma. Entre els factors de risc genètics hi ha per exemple gens concrets que alteren de forma directa la resposta que té una persona a un fàrmac o el temps que aquest roman en el cos, o també aquells que estan involucrats en diferents sistemes de neurotransmissió, com seria el cas dels gens implicats en l'alliberació de diferents neurotransmissors com la dopamina, la serotonina, el glutamat o la norepinefrina (Tsapakis i Travis, 2002; Cerdá *et al.*, 2010; Müller i Homberg, 2015; Xing, Li i Gao, 2016). També cal considerar aquells gens que actuen de forma indirecta tot alterant la resposta d'una persona a l'estrès (Enoch, 2012) o incrementant els comportaments de presa de riscos o de cerca de novetats (Kelly i Daley, 2013). Finalment, hi ha afeccions en algunes àrees cerebrals que són comunes tant per l'addicció com per altres trastorns psiquiàtrics, com per exemple aquelles implicades en la recompensa, la motivació, la presa de decisions, el control dels impulsos o les emocions (Ross i Peselow, 2012).

D'entre tots els trastorns psiquiàtrics comòrbids a les addiccions, en aquesta Tesi ens hem centrat amb el trastorn per dèficit d'atenció amb hiperactivitat (TDAH), que ha mostrat una alta comorbiditat a les addiccions en múltiples estudis (De Alwis *et al.*, 2014; Harstad *et al.*, 2014; van de Glind *et al.*, 2014; Weissenberger *et al.*, 2017; Chen *et al.*, 2018; Romo *et al.*, 2018).

Trastorn per dèficit d'atenció amb hiperactivitat (TDAH)

Diversos estudis clínics demostren una alta comorbiditat entre el TDAH i l'addicció a les drogues, i també amb l'addicció al menjar (De Alwis *et al.*, 2014; Harstad *et al.*, 2014; van de Glind *et al.*, 2014; Weissenberger *et al.*, 2017; Chen *et al.*, 2018; Romo *et al.*, 2018). Pel que fa a l'addicció a drogues, els nens amb TDAH tenen 2,5 vegades més risc de desenvolupar alguna addicció a les drogues (Harstad *et al.*, 2014). Si analitzem les addiccions a les diferents drogues, observem que els nens diagnosticats amb TDAH tenen tres vegades més risc de patir dependència de la nicotina, el doble de risc de patir dependència de l'alcohol o de la cocaïna i aproximadament 1,5 vegades més de risc d'addicció al cànnabis (Lee *et al.*, 2011). En el cas dels adults amb TDAH, aquests tenen nou vegades més risc de desenvolupar algun trastorn per ús de substàncies (Chen *et al.*, 2018). Tot i això, diferents estudis suggereixen que a través d'un correcte tractament del TDAH amb estimulants es pot reduir el risc de desenvolupar trastorns per ús de substàncies (Biederman *et al.*, 1999; Harstad *et al.*, 2014). Pel que fa a l'addicció al menjar, un estudi ha determinat que estudiants (d'entre 18 i 24 anys) amb TDAH tenien 2,27 vegades més risc de patir addicció al menjar (Romo *et al.*, 2018).

El TDAH és un trastorn neuropsiquiàtric que apareix en la infància amb una prevalença del 5% (Polanczyk *et al.*, 2007) i que es manté en l'edat adulta amb una prevalença del voltant del 2,5% (Simon *et al.*, 2009). El sexe és un factor que també afecta la prevalença, essent el trastorn quatre cops més freqüent en nens que en nenes segons estudis clínics (Polanczyk *et al.*, 2007) i 2,4 vegades més freqüent segons estudis poblacionals. Aquestes diferències en funció del sexe, però, són pràcticament inexistents en l'edat adulta (Matte *et al.*, 2015). Els criteris diagnòstics del TDAH es recullen en el DSM o en l'ICD, les versions actualitzades dels quals són el DSM-V i l'ICD-11. En el diagnòstic del TDAH es tenen en compte els dos símptomes principals del trastorn, el dèficit d'atenció per una banda i la hiperactivitat amb impulsivitat per l'altra, i el DSM-V recull un total de nou criteris diferents per diagnosticar cadascun d'aquests dos trets (Taula 3). Es considera que un pacient presenta un d'aquests dos trets si compleix sis o més dels criteris, o, en individus que tenen més de 16 anys, almenys cinc criteris durant els últims sis mesos amb un grau que no concorda amb l'edat de desenvolupament i/o amb un impacte negatiu en les activitats socials, acadèmiques o professionals (Taula 3). Segons la simptomatologia que presenten els pacients podem diferenciar entre tres subtipus de TDAH: predominantment amb dèficit d'atenció, predominantment hiperactiu/impulsiu, i combinat, que són els casos on es presenta dèficit d'atenció i hiperactivitat/impulsivitat.

Taula 3. Criteris diagnòstics del trastorn per dèficit d'atenció amb hiperactivitat.**CRITERIS PEL DÈFICIT D'ATENCIÓ**

- 1) Sovint no presta atenció suficient als detalls o comet errors deguts a descuits en les tasques escolars, a la feina o en altres activitats.
- 2) Sovint té dificultats per mantenir l'atenció en tasques o en activitats lúdiques.
- 3) Sovint sembla no escoltar quan se li parla directament.
- 4) Sovint no segueix instruccions i no finalitza tasques escolars, encàrrecs, o obligacions en el centre de treball.
- 5) Sovint té dificultats per organitzar tasques o activitats.
- 6) Sovint evita, li desagrada o és reticent pel que fa a dedicar-se a tasques que requereixen un esforç mental sostingut.
- 7) Sovint extravia objectes necessaris per a tasques o activitats.
- 8) Sovint es distreu fàcilment per estímuls irrellevants.
- 9) Sovint és descuidat en les activitats diàries.

CRITERIS PER LA HIPERACTIVITAT I IMPULSIVITAT

- 1) Sovint mou en excés mans o peus, o es mou en el seu seient.
- 2) Sovint abandona el seu seient a classe o en altres situacions en què s'espera que romangui assegut.
- 3) Sovint corre o salta excessivament en situacions en què és inapropiat fer-ho.
- 4) Sovint té dificultats per jugar o dedicar-se tranquil·lament a activitats d'oci.
- 5) Sovint "està en marxa" o sol actuar com si tingués un motor.
- 6) Sovint parla en excés.
- 7) Sovint precipita respostes abans d'haver estat completades les preguntes.
- 8) Sovint té dificultats per guardar torn.
- 9) Sovint interromp o interfereix en les activitats d'altres.

A més a més també s'han de complir les següents condicions:

- A) Alguns dels símptomes de dèficit d'atenció o hiperactivitat i impulsivitat han d'haver estat presents abans dels 12 anys.
- B) Alguns dels símptomes de dèficit d'atenció o hiperactivitat i impulsivitat es presenten en dos o més ambients (per exemple: casa i escola/feina; amb amics o familiars).
- C) Hi ha proves clares que els símptomes interfereixen o redueixen la qualitat de l'activitat social, acadèmica o laboral.
- D) Els símptomes no es presenten durant el transcurs d'un altre trastorn psicòtic, i no s'expliquen millor per la presència d'un altre trastorn mental (trastorns de l'estat d'ànim, ansietat, trastorn dissociatiu o trastorn de la personalitat).

Hi ha una elevada similitud entre els criteris descrits en el DSM-V i l'ICD-11 a l'hora de diagnosticar el TDAH. L'ICD-11 però, té una visió més àmplia pel que fa a la caracterització clínica del TDAH ja que defineix els límits amb la "normalitat" i amb altres diagnòstics, a més de mostrar un panorama evolutiu (canvis en el temps o manifestacions dependents de l'edat) així com diferències pel que fa al gènere.

Introducció

Hi ha múltiples regions corticals del cervell que estan implicades en el TDAH: l'escorça prefrontal dorsolateral està relacionada amb la memòria del treball, l'escorça prefrontal ventromedial està implicada en la presa de decisions i la planificació estratègica, i finalment l'escorça parietal està implicada en l'atenció orientativa (Figura 3). També hi ha regions subcorticals que estan implicades en el TDAH: la circumvolució dorsal anterior del cíngol promou components afectius i cognitius del control executiu. Aquesta àrea, juntament amb el gangli basal (que inclou el NAc, el nucli caudat i el putamen) formen el circuit frontoestriatal (Figura 3) (Faraone *et al.*, 2015). Estudis de neuroimatge en pacients amb TDAH mostren deficiències en totes aquestes estructures i també en l'AMY i el cerebel a nivell d'estructura (Valera *et al.*, 2007; Ellison-Wright, Ellison-Wright i Bullmore, 2008; Nakao *et al.*, 2011; Frodl i Skokauskas, 2012; Norman *et al.*, 2016; Hoogman *et al.*, 2017), funció (Hart *et al.*, 2012; Plichta i Scheres, 2014; Norman *et al.*, 2016) i connectivitat (Chen *et al.*, 2016; Cortese *et al.*, 2016; Aoki, Cortese i Castellanos, 2018). Tot i així, el descobriment més consistent és una reducció en la mida total del cervell amb canvis específics al nucli caudat, a la matèria blanca de l'escorça prefrontal, al cos callós i a la vermis del cerebel (Tripp i Wickens, 2009).

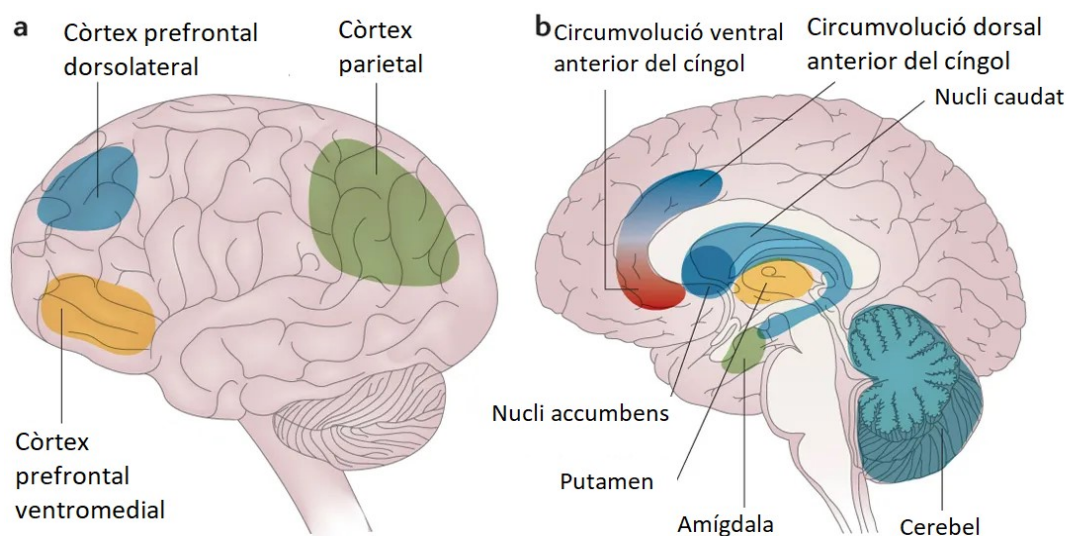


Figura 3. Àrees corticals i subcorticals cerebrals implicades en el TDAH. a) Visió lateral b) Visió medial. Adaptada de Faraone *et al.*, 2015.

Igual que per les addiccions, la dopamina (DA) també és un neurotransmissor clau en el TDAH. De fet, una de les teories proposades per explicar la fisiopatologia del TDAH és la teoria del dèficit de transferència de DA (Figura 4) (Blum *et al.*, 2008). Una transferència normal de la senyalització de DA dona primerament una gran alliberació de DA degut a una recompensa no esperada. Després,

en les primeres fases de l'aprenentatge, part de l'alliberació de DA és transferida a l'estímul condicionat previ a la recompensa. Al final de l'aprenentatge tota la resposta de DA és transferida a l'estímul condicionat (Figura 4). Segons aquesta teoria, la transferència d'alliberació de DA a l'estímul condicionat previ està disminuïda en nens amb TDAH, de manera que és ineficaç. El dèficit de transferència de DA explica doncs els símptomes de falta d'atenció, com la manca d'atenció als detalls o el fet de cometre errors per descuits, ja que hi ha una absència del reforç a l'atenció continuada obtingut a partir de l'alliberació anticipada de DA. De la mateixa manera, també explica la hiperactivitat i la impulsivitat, com per exemple quan el nen s'aixeca de la cadira, ja que l'alteració de la transferència dopaminèrgica fa que no obtinguin la recompensa (Tripp i Wickens, 2009). A més, aquesta teoria també es veu reforçada pel fet que el mecanisme d'acció dels tractaments més utilitzats pel TDAH, com el metilfenidat, les amfetamines o l'atomoxetina, entre d'altres, és l'alliberació de DA i norepinefrina al SNC.



Figura 4. Teoria del dèficit de transferència de dopamina. A) Transferència normal de l'alliberació de DA a l'estímul condicionat previ durant l'aprenentatge. B) Transferència deficiente de l'alliberació de DA a l'estímul condicionat previ durant l'aprenentatge, situació proposada en pacients amb TDAH. Adaptada de Rege, 2019.

3. Genètica de les addiccions i trastorns comòrbids

Els trastorns neuropsiquiàtrics com les addiccions o el TDAH són trastorns multifactorials en què hi intervenen factors genètics i ambientals de predisposició que determinen la susceptibilitat. Durant dècades, els investigadors han treballat per identificar les causes moleculars subjacents als trastorns psiquiàtrics. En un inici, l'estudi de la genètica d'aquests trastorns es va centrar en buscar alteracions en diferents components dels sistemes de neurotransmissió de DA,

noradrenalina i serotonina mitjançant estudis dirigits de gens candidats. Tot i així, l'adveniment dels estudis genòmics a gran escala, que no parteixen d'hipòtesis prèvies, ha permès descriure molts altres gens que també estan implicats en els trastorns.

3.1. Heretabilitat

Tot i que és fàcil identificar quins són els factors ambientals que participen en el desenvolupament de la dependència, estudis familiars, d'adopcions i de bessons demostren l'existència d'una forta component genètica. L'heretabilitat (H^2), en termes estadístics, és la proporció de variància fenotípica total d'una població que pot ser atribuïda a factors genètics. En el cas dels trastorns per ús de drogues, les heretabilitats oscil·len entre el 33 i el 79% segons la droga d'abús, essent la dependència de cocaïna de les més elevades (Goldman, Oroszi i Ducci, 2005; Agrawal *et al.*, 2012; Ducci i Goldman, 2012) (Figura 5). L'inici del consum i l'ús de les drogues d'abús presenten en general una heretabilitat menor (però significativa) que la dependència. Per tant, podem dir que els factors genètics no només predisposen a la dependència, sinó que també tenen un paper en l'inici del consum i de l'ús de la droga o conducta (Goldman, Oroszi i Ducci, 2005; Ducci i Goldman, 2012). A més, també s'ha descrit que els familiars en primer grau de pacients dependents de cocaïna, opiàcics, cànnabis i/o alcohol tenen un risc vuit vegades superior de desenvolupar dependència d'aquestes drogues que els familiars d'individus controls (Merikangas *et al.*, 1998). Finalment, també s'ha observat una agregació familiar pel que fa a la droga d'abús predominant, recolzant la idea de l'existència de factors genètics de risc específics de cada substància a part dels factors de risc comuns per a totes les drogues.

De moment, encara no s'han realitzat estudis de bessons per determinar l'heretabilitat de l'addicció al menjar (Figura 5). Malgrat això, sí que s'ha estudiat l'heretabilitat d'altres conductes que tenen relació amb l'addicció al menjar, com l'obesitat o el trastorn per afartament (Davis i Bonder, 2019). Un estudi de bessons realitzat en 11.000 parelles de bessons va estimar una heretabilitat per a l'índex de massa corporal d'un 37% (Llewellyn *et al.*, 2013). Un altre estudi recent, que incloïa 18.000 bessons adults, va determinar una heretabilitat del trastorn per afartament del 65% (Capusan *et al.*, 2017). Aquests resultats obtinguts per Capusan *et al.*, 2017, són força congruents amb els de dos estudis previs en els quals s'obtenien heretabilitats pel trastorn d'afartament del 41% i del 57% (Reichborn-Kjennerud *et al.*, 2004; Javaras, 2007). Així doncs, tot i que l'heretabilitat de l'addicció al menjar encara no està determinada, estudis en conductes similars suggereixen que gran part de la variabilitat fenotípica en l'addicció al menjar és deguda a la genètica (Davis i Bonder, 2019).

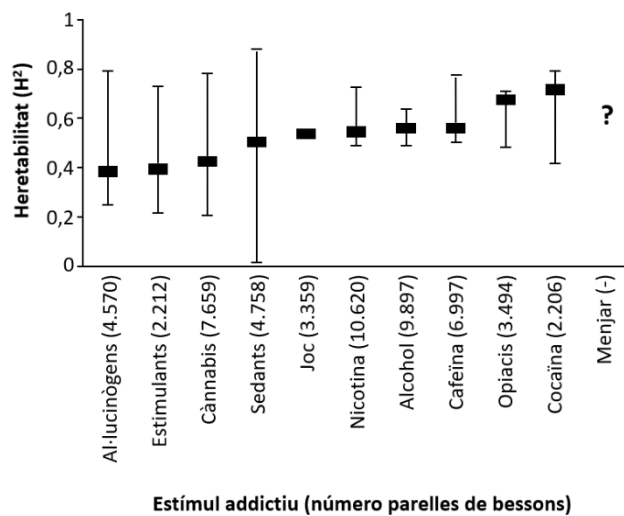


Figura 5. Heretabilitat (mitjana ponderada i rang) de 10 trastorns addictius: dependència o abús d'al·lucinògens, estimulants, cànnabis, sedants, opiacis i cocaïna; dependència de l'alcohol; persistència del tabaquisme; consum de cafeïna i jocs patològics. Les mitjanes ponderades estan calculades a partir de les dades d'estudis de bessons adults. Entre parèntesis: número de parelles de bessons. Adaptada de Ducci i Goldman, 2012.

Finalment, l'últim trastorn amb què hem treballat en aquesta Tesi és el TDAH, que té una heretabilitat del 70-80%, tot mostrant la gran importància de la genètica en aquest trastorn (Faraone *et al.*, 2005). També s'ha descrit que germans de pacients amb TDAH tenen aproximadament nou vegades més risc de desenvolupar el trastorn en comparació amb la població general (Chen *et al.*, 2008).

3.2. Estudis d'associació genètica

Els estudis d'associació de tipus cas-control són una de les metodologies més utilitzades per a la identificació de factors genètics de susceptibilitat a les malalties complexes. Es tracta de mètodes no paramètrics basats en la comparació, mitjançant taules de contingència, de les freqüències genotípiques o al·lèliques d'un determinat marcador entre un grup d'individus afectats i un grup de controls. La mostra control defineix dos tipus d'estudis d'associació, els poblacionals, en què la mostra control consisteix en individus no relacionats amb els casos; i els familiars, en què la mostra control consisteix en els al·lèls no transmesos dels progenitors de famílies amb un únic fill afectat.

La gran majoria d'estudis d'associació realitzats se centren en variants genètiques comunes (polimorfismes) que estan presents a la població general, però que tenen una freqüència superior en els individus afectats. Així doncs, els primers treballs es van centrar en la hipòtesi "malaltia comuna - variant comuna" o "*common disease - common variant (CDCV)*". Recentment, aquesta tendència està canviant gràcies a l'aparició de tecnologies de seqüenciació a gran escala, que

permeten explorar també variants poc comunes o rares en moltes malalties complexes, per tal de contrastar la hipòtesi “malaltia comuna-variant rara” o “*common disease - rare variant (CDRV)*”. Segons aquesta segona hipòtesi, les variants genètiques poc freqüents també podrien contribuir a l’etiologia de les malalties, probablement amb efectes més penetrants sobre el fenotip.

Els estudis d’associació entre fenotips i variants comunes utilitzen dos tipus d’aproximacions: 1) Estudis d’associació de gens o sistemes gènics candidats basats en evidències prèvies; 2) Estudis d’associació a escala genòmica (*Genome-Wide Association Studies, GWAS*), en què s’analitzen polimorfismes (en general de canvis d’un únic nucleòtid, o SNPs) que cobreixen tot el genoma.

3.2.1. Estudis d’associació de gens i sistemes gènics candidats

Aquests estudis preseleccionen un conjunt de gens o sistemes gènics candidats amb una possible implicació en el trastorn en base a evidències prèvies obtingudes en estudis farmacològics, d’expressió gènica, de neuroimatge, d’estudis realitzats en animals o de deduccions derivades dels fenotips.

Des de principis dels 2000 s’han publicat un gran nombre d’estudis dirigits a gens candidats per les addiccions a drogues o a una droga específica, la gran majoria d’ells centrats en gens que codifiquen proteïnes implicades en dels diferents sistemes de neurotransmissió (dopaminèrgic, serotoninèrgic, GABAèrgic, etc.) (Hall *et al.*, 2013). Si considerem aquells estudis d’associació per l’addicció a les drogues en general, s’han trobat associacions amb gens que codifiquen receptors de diversos sistemes de neurotransmissió (*DRD2, DRD3, DRD4, OPRD1, OPRK1, OPRM1* i *GABRA*), així com per un transportador del sistema serotoninèrgic (*SLC6A4*) o altres gens relacionats amb els sistemes de neurotransmissió, com la catecol-O-metiltransferasa (*COMT*) (Comings *et al.*, 1994; Persico *et al.*, 1996; Vandenberg *et al.*, 1997, 2000; Kranzler *et al.*, 1998; Krebs *et al.*, 1998; Hoehe *et al.*, 2000; Luo *et al.*, 2003; Agrawal *et al.*, 2006; Zhang *et al.*, 2008; Al-Eitan *et al.*, 2014). També s’han publicat nombrosos estudis dirigits a gens candidats per la dependència de cocaïna específicament, una de les addiccions considerades en aquesta Tesi. De fet, el nostre grup ha contribuït a aquest coneixement realitzant estudis d’associació que consideren tant gens dels sistemes serotoninèrgic i dopaminèrgic com gens relacionats amb l’alliberament de neurotransmissors (Fernández-Castillo *et al.*, 2010, 2012, 2013; Tristán-Noguero *et al.*, 2013). Quan considerem els estudis d’associació realitzats per la dependència de cocaïna, alguns dels gens destacats són els mateixos que els destacats per la dependència de drogues en general com és d’esperar, ja que tot i que cada droga activa vies específiques (Russo *et al.*, 2009; Zhou *et al.*, 2011; Graziane *et al.*, 2016), totes elles actuen directa o indirectament sobre el sistema de

recompensa (Wise, 1996; MacNicol, 2017). En les associacions obtingudes en els estudis d'associació de gens candidats a la dependència de cocaïna hi trobem altre cop receptors de diversos sistemes de neurotransmissió (*DRD2, DRD3, HTR3A, HTR3B, CHRNA3, CHRNA4, CHRNA5, CNR1* o *NK3R*), així com el transportador *DAT* o altres proteïnes relacionades com *COMT, DBH, PDYN, NPY, TPH2* o *CARTPT* (Noble *et al.*, 1993; Cubells *et al.*, 2000; Chen *et al.*, 2002; Messas *et al.*, 2005; Dahl *et al.*, 2006; Wetherill *et al.*, 2008; Foroud *et al.*, 2008; Grucza *et al.*, 2008; Lohoff, Bloch, *et al.*, 2008; Lohoff, Weller, *et al.*, 2008; Zuo *et al.*, 2009; Sherva *et al.*, 2010; Fernández-Castillo *et al.*, 2010, 2012; Clarke *et al.*, 2013; Yang i Li, 2014).

Altres trastorns psiquiàtrics comòrbids a l'addicció també han estat associats a alteracions en diferents components dels sistemes de neurotransmissió, com el TDAH, un trastorn que ha sigut àmpliament estudiat pel grup mitjançant diverses aproximacions, per exemple estudis d'associació (Sánchez-Mora *et al.*, 2010, 2011; Ribasés, Ramos-Quiroga, *et al.*, 2012; Ribasés, Sánchez-Mora, *et al.*, 2012; Sánchez-Mora *et al.*, 2013; Garcia-Martínez *et al.*, 2016; Pagerols *et al.*, 2017). En el cas del TDAH, els estudis d'associació principalment s'han centrat en els sistemes dopaminèrgic i noradrenèrgic, dianes dels medicaments utilitzats per aquest trastorn, tot i que també s'han considerat altres sistemes de neurotransmissió, com el sistema serotoninèrgic. Algunes de les associacions que s'han trobat han destacat gens que codifiquen receptors, de dopamina (*DRD1, DRD4* i *DRD5*) o de serotonina (*HTR1B*) o que codifiquen transportadors, de dopamina *DAT (SLC6A3)*, de norepinefrina (*SLC6A2*), o de serotonina (*SLC6A4*). Igual que en el cas de l'addicció a drogues i cocaïna, també s'han identificat associacions amb gens que codifiquen productes relacionats de forma no tant directa amb els sistemes de neurotransmissió, com *COMT*, la monoamino oxidasa A (*MAOA*), la dopamina beta-hidroxilasa (*DBH*), la tirosina hidroxilasa (*TH*) o SNAP-25 (amb un rol important en la regulació de l'alliberació de neurotransmissors i en la funció sinàptica) (Faraone *et al.*, 2005; Gizer, Ficks i Waldman, 2009; Franke *et al.*, 2010; Landaas *et al.*, 2010; Sánchez-Mora *et al.*, 2011; Zhang *et al.*, 2012; Gallo i Posner, 2016; Pagerols *et al.*, 2017). Com podem observar, alguns dels gens destacats en aquests estudis són comuns amb els obtinguts per la dependència de drogues, segurament a causa de l'alta comorbiditat existent entre ambdós trastorns.

Tot i que s'han publicat un gran nombre d'estudis d'associació per l'addicció o per trastorns comòrbids, la majoria d'associacions identificades són nominals i en molts pocs casos s'han pogut replicar els resultats obtinguts. Això podria ser degut a diferències en el disseny o a una mida mostral insuficient per detectar associacions.

3.2.2. Estudis d'associació a escala genòmica

Des de l'aparició de les tecnologies de genotipació a gran escala, van sorgir noves aproximacions que permetien realitzar estudis lliures d'hipòtesis, com els estudis d'associació a escala genòmica (GWAS). Aquests estudis no seleccionen gens o regions candidates en base a coneixements previs, sovint molt limitats, de manera que els resultats obtinguts no presenten biaixos. Aquest enfoc requereix la realització d'un elevat nombre de tests estadístics de contrast, i per tant és necessària una mostra molt més gran per tal d'assolir suficient poder estadístic en les associacions. A partir d'aquests estudis també es va començar a calcular l'anomenada heretabilitat deguda a SNPs, un terme molt important en els trets complexos, que es defineix com la proporció de variació fenotípica atribuïble als efectes additius d'un conjunt d'SNPs.

Pel que fa a les addiccions, en els últims anys, els GWAS han permès identificar múltiples regions del genoma o gens concrets que poden contribuir a la dependència de drogues en general o a la dependència d'una droga específica. Alguns d'aquests GWAS han assenyalat gens de risc que ja s'havien explorat en estudis de gens candidats previs, com per exemple la família alcohol deshidrogenasa (ADH) per la dependència de l'alcohol (Samochowiec *et al.*, 2014; Deak, Miller i Gizer, 2019). En canvi, hi ha altres estudis que han assenyalat nous candidats per a la dependència de drogues en general o per una droga en particular, com ara la cocaïna, tot identificant noves funcions que podrien estar relacionades amb l'etiologia de la malaltia (Taula 4). La gran majoria de GWAS sobre drogues han explorat en la dependència de l'alcohol (Tawa, Hall i Lohoff, 2016), però n'hi ha també sobre la dependència de cocaïna (Gelernter *et al.*, 2014; Cabana-Domínguez *et al.*, 2019) o sobre les drogues en general (McGue *et al.*, 2013; Wetherill *et al.*, 2014; Johnson *et al.*, 2015). Alguns d'aquests GWAS descriuen associacions que són biològicament interessants i robustes estadísticament (Taula 4), però calen estudis de rèplica en mostres independents i també estudis funcionals que ajudin a entendre els mecanismes pels quals aquests factors genètics són importants en la predisposició a la dependència.

Taula 4. Associacions significatives a escala genòmica obtingudes en estudis GWAS de dependència de cocaïna o de múltiples drogues. Adaptada de Jensen, 2016.

Variant	Referència	Associat a	Població	P-valor	Gen proper
rs2629540	Gelernter et al., 2014	Dependència de cocaïna	AA, EA	4,28E-08	<i>FAM53B</i>
rs1868152	McGue et al., 2013		EA	4,9E-08	<i>CRYGS</i>
rs9829896	Johnson et al., 2015	Dependència de múltiples drogues	AA	4,63E-08	<i>KAT2B</i>
rs2952621	Wetherill et al., 2014		EA	1,8E-08	<i>LOC151121</i>
rs2567261			EA	3,8E-08	<i>ARHGAP28</i>

EA = americans d'origen europeu; AA = americans d'origen africà

Per altra banda, fins ara només s'ha realitzat un GWAS per l'addicció al menjar, ara fa 4 anys (Cornelis *et al.*, 2016). En aquest estudi hi van participar un total de 9.314 dones amb ascendència europea, de les quals 287 van ser diagnosticades amb addicció al menjar segons el qüestionari YFAS. Malgrat el poder limitat d'aquest estudi a causa del baix nombre de casos, es van identificar associacions amb dos *loci* per sota del llindar de significació a escala genòmica (17q21.31 i 11q13.4), assenyalant *PRKCA* i *NTM* com a possibles gens candidats del trastorn.

En els primers GWAS realitzats per TDAH no es van descriure variants per sota el llindar de significació (Lasky-Su, Anney, *et al.*, 2008; Lasky-Su, Neale, *et al.*, 2008; Lesch *et al.*, 2008; Mick *et al.*, 2010; Neale, S. Medland, *et al.*, 2010; Hinney *et al.*, 2011; Fliers *et al.*, 2012; Ebejer *et al.*, 2013; Yang *et al.*, 2013; Zayats *et al.*, 2015; Sánchez-Mora *et al.*, 2015), malgrat que moltes d'aquestes mostres es van combinar en una metanàlisi (Neale, S. E. Medland, *et al.*, 2010). Probablement la no obtenció d'associacions significatives es devia a que els primers GWAS realitzats tenien mides mostrals petites, amb poder suficient per detectar només riscos relatius grans. Malauradament, els riscos relatius de les variants genètiques comunes associades a les malalties mentals solen ser petits, de manera que, per tal d'obtenir associacions significatives calia augmentar significativament la mida mostral (Sullivan *et al.*, 2018). El *Psychiatric Genomics Consortium* (PGC) és un consorci que es funda l'any 2007 amb el propòsit de coordinar esforços d'investigadors d'arreu del món per tal de poder aconseguir mides mostrals molt més grans. Per exemple, l'any 2019 es va publicar la primera metanàlisi de dades GWAS amb resultats significatius a escala genòmica per al TDAH, utilitzant 12 cohorts diferents que inclouen un total de 20.183 casos i 35.191 controls (Demontis *et al.*, 2019). D'aquesta manera, aquest consorci ha fet grans aportacions per a la comprensió de les bases genètiques de diversos trastorns psiquiàtrics, promovent la realització de GWAS i metanàlisis que han identificat un gran nombre d'associacions significatives a escala genòmica i han calculat l'heretabilitat deguda a SNPs per a cadascun d'aquests trastorns (Taula 5).

Taula 5. Mides mostrals i resultats obtinguts pel *Psychiatric Genomics Consortium* (PGC) fins l'any 2017. Adaptada de Sullivan et al., 2018.

Trastorn psiquiàtric	Casos (N)	Associacions	H ² general (bessons)	H ² SNP
Esquizofrènia	60.955	155	81%	45%
Depressió	130.664	44	30%-40%	8.9%
Trastorn bipolar	20.352	19	80%	21%
Trastorn per dèficit d'atenció amb hiperactivitat	20.183	12	70%-80%	22%
Trastorns de l'espectre autista	18.381	5	75%	12%
Anorexia nerviosa	3.495	8	56%	≈20%
Trastorn per ús de substàncies	12.798	1	50%	≈8%
Síndrome de Tourette	4.232	1	60%-80%	58%
Trastorn obsessiu-compulsiu	2.688	0	45%-65%	37%
Trastorn per estrès posttraumàtic	5.131	0	30%-40%	5%-35%

Associacions: associacions independents per sota del llindar de significació a escala genòmica ($p < 5E-08$); H² general (bessons): Heretabilitat estimada a partir d'estudis de bessons; H² SNP: Heretabilitat deguda a SNPs estimada a partir de resultats de GWAS.

Aquest consorci no només ha realitzat GWAS per a trastorns psiquiàtrics específics, sinó que també ha publicat GWAS que reuneixen diversos trastorns psiquiàtrics, amb l'objectiu de conèixer les bases genètiques comunes (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013). L'últim estudi transversal va incloure un total de vuit trastorns psiquiàtrics: anorèxia, trastorn per dèficit d'atenció amb hiperactivitat, trastorns de l'espectre autista, trastorn bipolar, depressió, trastorn obsessiu-compulsiu, esquizofrènia i síndrome de Tourette. Aquest estudi va incloure un total de 232.964 casos i 494.162 controls i va descriure un total de 109 *loci* associats com a mínim a dues malalties, tot demostrant que hi ha factors genètics de risc comuns entre els diferents trastorns (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019).

3.2.3. Model omnigènic

L'aparició dels GWAS i dels estudis de seqüenciació d'exomes, la fracció codificant del genoma, ha permès desxifrar part de les bases genètiques de les malalties complexes, però tot i així, les associacions significatives obtingudes en aquests estudis només expliquen una part de la variació genètica de les malalties complexes, tal i com s'indica amb el càlcul de l'heretabilitat deguda a SNPs per cada trastorn. L'any 2009, Manolio et al. van introduir un nou concepte, "l'heretabilitat perduda", que definia tota aquella variació genètica que no s'explicava per les senyals d'associació significatives dels estudis a gran escala (Manolio *et al.*, 2009). Més tard, nous estudis van descriure que l'heretabilitat perduda podia explicar-se a través dels efectes grans de variants rares (Marouli *et al.*, 2017) però també pels efectes moderats dels polimorfismes d'un únic nucleòtid situats per sota del llindar de significació dels estudis a gran escala (Yang *et al.*, 2010; Shi, Kichaev i Pasaniuc, 2016). En malalties complexes, moltes de les variants que presenten associacions significatives en

GWAS cauen en regions no codificants del genoma, tot indicant un possible efecte dels elements de regulació de l'expressió gènica. De fet, per diferents trets i malalties complexes, s'ha observat un enriquiment d'associacions significatives en regions de cromatina activa, com ara regions promotores o regions amplificadores (*enhancers*, en anglès) (Maurano *et al.*, 2012; Kundaje *et al.*, 2015; Wang *et al.*, 2016). Totes aquestes evidències fan pensar que les malalties complexes estan causades per l'acumulació de petits efectes en gens clau i en vies reguladores d'aquests gens.

L'any 2017 es va descriure un nou model d'herència per les malalties complexes anomenat model omnigènic (Boyle, Li i Pritchard, 2017), que destacava la importància de les variants de risc comunes amb associacions per sota del llindar de significació. El model omnigènic proposa que la majoria dels trets complexos estan directament afectats per un nombre modest de gens o vies reguladores amb una funció específica per l'etiologia de la malaltia. Aquests són els anomenats gens essencials (*core genes*) i tenen un rol biològic evident en la malaltia. A més, mencionen que l'efecte d'aquests gens és relatiu, ja que gran part de la contribució genètica total ve de gens perifèrics que no tenen un rol directe en la malaltia i que superen en número el grup de gens essencials (Boyle, Li i Pritchard, 2017). Dos anys més tard, els autors van aprofundir més en aquest nou model, tot postulant que la funció/expressió dels gens essencials estaria probablement afectada per un gran nombre de variants situades en *trans* amb un efecte petit (variants perifèriques). Si suposem que un tret pot tenir centenars de gens essencials, això pot ajudar a explicar perquè hi ha tants punts diferents en el genoma que contribueixen a l'herència d'aquests fenotips. A més, aquest model més desenvolupat suggereix que la major part de l'heretabilitat és deguda als efectes en *trans*, sobretot si els gens essencials tendeixen a ser co-regulats (Liu, Li i Pritchard, 2019).

4. Epigenètica de les addiccions i trastorns comòrbids

En el cas de les addiccions i els trastorns comòrbids, en tractar-se de malalties complexes, hi intervenen factors genètics i ambientals, amb l'epigenètica com a vincle entre ambdues. El terme epigenètica s'utilitza per descriure les modificacions moleculars dinàmiques que es generen a la cromatina, ja sigui directament sobre la molècula de DNA o a nivell de la carcassa proteica que l'embolcalla. Aquestes modificacions tenen un efecte funcional en la regulació de diferents processos com la reparació del DNA, l'organització de la cromatina, l'*splicing* o la transcripció de l'RNA entre d'altres (Hamilton i Nestler, 2019). Els mecanismes epigenètics més estudiats són les modificacions d'histones, la metilació del DNA i els microRNAs (miRNAs).

Alguns factors ambientals com l'estrès crònic, els traumes, les drogues o l'exposició a fàrmacs poden induir canvis estables en l'expressió gènica, que poden alterar el funcionament de circuits neuronals i en última instància tenir un impacte sobre el comportament (Peña *et al.*, 2014). L'epigenètica estaria darrere les interaccions entre ambient i genètica. Diferents estudis també comencen a proposar que aquests canvis epigenètics poden ser transmesos a la següent generació (Nestler, 2014), però, aquest és encara un tema controvertit. Per altra banda, alguns estudis suggereixen que alguns factors ambientals poden interaccionar amb la genètica tot generant canvis epigenètics en estadis del desenvolupament concrets que incrementarien el risc de desenvolupar trastorns psiquiàtrics o addicció a algunes substàncies (Nestler, 2014; Peña *et al.*, 2014). Un exemple és el cas del consum de cànnabis en l'adolescència (aquí el cànnabis seria el factor de risc ambiental), que mostra una clara associació amb un risc més elevat de patir episodis psicòtics (Barkus i Murray, 2010; Parakh i Basu, 2013).

Pel que fa a l'addicció, les cèl·lules nervioses del sistema de recompensa pateixen adaptacions a nivell epigenètic a mesura que un individu esdevé addicte degut a l'exposició repetida a una substància. Aquestes adaptacions epigenètiques (per exemple metilació de certs gens) són les que generen canvis en la funcionalitat del cervell i contribueixen a comportaments característics de l'addicció que es mantindran de per vida. Estudiar aquests mecanismes epigenètics pot ajudar a entendre les bases biològiques de l'addicció i també buscar noves teràpies. Malauradament, la complexitat del paisatge neuroepigenètic generat per les substàncies addictives fa difícil saber quins són els canvis concrets que contribueixen als mecanismes patogènics de l'addicció (Hamilton i Nestler, 2019). Entre aquests mecanismes hi ha les modificacions posttraduccionals d'histones, la metilació del DNA i els miRNAs.

4.1. Modificacions posttraduccionals d'histones

Les modificacions posttraduccionals (MPTs) d'histones són el mecanisme epigenètic més ben estudiat. A l'interior del nucli cel·lular hi ha la cromatina, un complex macromolecular format per DNA empaquetat amb histones (octàmers proteics) per formar els nucleosomes (Kornberg, 1974). Les modificacions posttraduccionals regulen el grau de condensació de la cromatina fent que els gens d'una regió concreta siguin més o menys accessibles a la maquinària de transcripció. Les subunitats de les histones poden ser modificades per nombroses MPTs. Les MPTs més freqüents són l'acetilació dels residus de lisina (K), associada a un increment de la transcripció, la metilació dels residus de lisina o arginina (R) i la fosforilació dels residus de serina (S) o treonina (T), que poden promoure o reprimir la transcripció, en funció del residu que es modifiqui. A més, hi ha altres tipus de MPTs com la ubiquitinació, la sumoilació i l'ADP-ribosilació, l'efecte de les quals

l'addicció, però encara hi ha poc coneixement sobre com la regulació d'aquestes proteïnes afectaria el fenotip addictiu (Hamilton i Nestler, 2019).

En el cas del TDAH la implicació de les MPTs d'histones en el trastorn ha estat molt poc estudiat, però tot i així, aquest mecanisme epigenètic sí que s'ha relacionat amb la patogènesi d'altres trastorns psiquiàtrics comòrbids a les addiccions, principalment amb l'esquizofrènia i la depressió. En esquizofrènia, es proposa HDAC2 com a possible diana terapèutica (Kurita *et al.*, 2012) i també se suggereix que la combinació de l'acetilació d'histones i la metilació del DNA en promotors de GABA en neurones pot tenir un paper clau en la fisiopatologia d'aquesta malaltia (Tremolizzo *et al.*, 2002; Dong *et al.*, 2007). En un model de depressió en ratolí s'ha relacionat amb una disminució de l'expressió de *Bdnf* amb un increment dels nivells de metilació de les histones situades en el seu promotor. A més, aquesta disminució de l'expressió s'ha pogut corregir a través de l'administració d'un antidepressiu que actua incrementant els nivells d'acetilació de les histones del promotor de *Bdnf* (Onishchenko *et al.*, 2008). Finalment, també s'han observat patrons de metilació i acetilació d'histones aberrants en pacients amb depressió o esquizofrènia (Chase *et al.*, 2013; Schmitt *et al.*, 2014).

4.2. Metilació del DNA

La metilació del DNA és una marca epigenètica dinàmica i reversible però més estable que les modificacions d'histones. Pot ser modificada per factors ambientals i és catalitzada per metiltransferases del DNA (DNMTs) (Figura 7A), que afegeixen un grup metil al C5 de la citosina (5-mC), majoritàriament en dinucleòtids CpG (Klose i Bird, 2006). En mamífers hi ha tres DNA metiltransferases principals: DNMT1, encarregada de la metilació de manteniment, i DNMT3a i DNMT3b, encarregades de la metilació *de novo* (Figura 7A) (Liu *et al.*, 2018). Les proteïnes TET són les implicades en la demetilació (Tahiliani *et al.*, 2009).

La funció principal de la metilació del DNA és la regulació de l'expressió gènica. La metilació de dinucleòtids CpG prop dels llocs d'inici de transcripció presenta amb freqüència una correlació negativa amb l'expressió gènica i, en canvi, quan la metilació es produeix en el cos dels gens la correlació és positiva. La metilació intergènica, per altra banda, no sembla que tingui tan impacte sobre l'expressió gènica (Liu *et al.*, 2018). En el cas de la repressió de la transcripció, la metilació del DNA en les zones promotores pot actuar de dues maneres: (1) evitant la unió de factors de transcripció a les zones promotores de gens (Figura 7B) (Hope, 1998; Domcke *et al.*, 2015) o (2) promovent l'entrada de proteïnes d'unió a grups metil-CpG, com MeCP2 (methyl-CpG binding protein 2) que recluten co-repressors transcripcionals com HDACs (Figura 7C), que afavoreixen la compactació de la cromatina (Bird i Wolffe, 1999; Kundu i Rao, 1999; Klose i Bird, 2006).

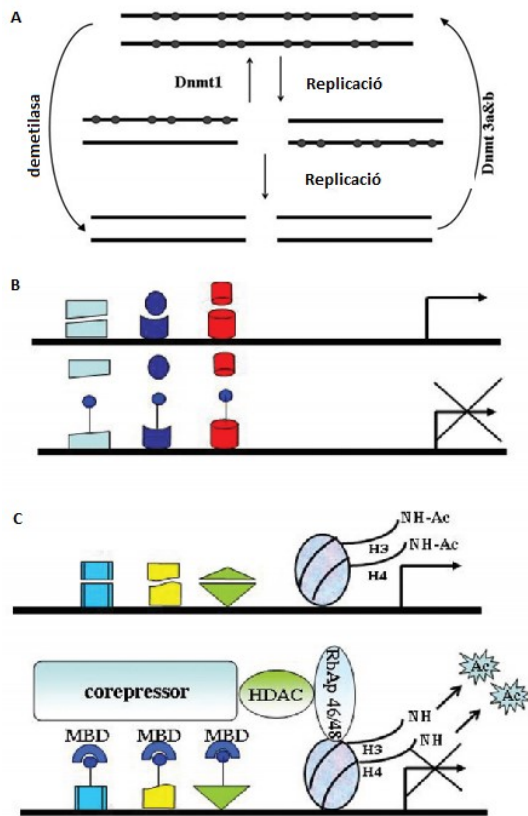


Figura 7. Característiques principals de la metilació del DNA. A) Metabolisme de la metilació del DNA. Dnmt1 s'encarrega de la metilació *de novo*, en canvi DNMT3A i DNMT3B s'encarreguen de la metilació de manteniment. B) Regulació de l'expressió gènica directa. La metilació dels llocs CpG (piruletes blaves) evita que els factors de transcripció s'hi uneixin, de manera que s'evita la iniciació de la transcripció. C) Regulació de l'expressió gènica indirecta. Les citosines metilades a la regió promotora dels gens recluten proteïnes d'unió a grups metil i complexes remodeladors de la cromatina per inhibir la transcripció. Adaptada de Razin i Kantor, 2005.

Gràcies a l'aparició de la seqüenciació a gran escala, durant els darrers anys s'han desenvolupat nombrosos mètodes per avaluar la metilació al llarg de tot el genoma (perfil metilòmic), per exemple MeDIP-seq, MBD-seq, MRE-seq, RRBS-seq o WGBS-seq (Yong, Hsu i Chen, 2016). En els mètodes de MeDIP-seq, MBD-seq i MRE-seq es duu a terme un enriquiment de DNA metilat previ a la seqüenciació. MeDIP-seq precipita el DNA metilat utilitzant un anticòs que reconeix la 5mC, MBD-seq precipita el DNA metilat utilitzant la proteïna MeCP2, que té afinitat pels grups metil-CpG, i finalment, MRE-seq utilitza uns enzims sensibles a la metilació que només tallaran aquells fragments no metilats, enriquint així la mostra amb fragments de DNA metilat. La RRBS-seq utilitza enzims de restricció per dur a terme un enriquiment de regions del genoma amb alt número de dinucleòtids CpG que posteriorment es bisulfiten i finalment es seqüencien. Finalment, la WGBS-seq ens permet avaluar la metilació de tot el genoma mitjançant la bisulfiteació del DNA seguida de seqüenciació. En tots els casos, aquestes tecnologies ens permetran identificar a gran escala regions diferencialment metilades (*Differentially Methylated Regions*, DMRs) quan comparem diferents grups.

Aquests últims anys s'ha suggerit que la metilació del DNA podria jugar un paper important en l'etiologia dels trastorns neuropsiquiàtrics. Estudis de metilació a escala genòmica han determinat

nivells globals de metilació anormals en la majoria de trastorns psiquiàtrics (Shimabukuro *et al.*, 2007; Bönsch *et al.*, 2012; Melas *et al.*, 2012; Nishioka *et al.*, 2013; Tremolizzo *et al.*, 2014; Tseng *et al.*, 2014; Murphy *et al.*, 2015; Liu *et al.*, 2018). A més, s'ha descrit que pacients amb esquizofrènia tractats amb diferents fàrmacs neurolèptics mostren una correcció d'aquest canvi en els nivells de metilació globals per arribar pràcticament als nivells de metilació observats en controls (Bönsch *et al.*, 2012; Melas *et al.*, 2012).

Si ens referim a gens concrets, s'han observat canvis dels nivells de metilació de gens relacionats amb la neurotransmissió en diferents trastorns psiquiàtrics (Taula 6) (Liu *et al.*, 2018). També s'han detectat canvis d'expressió de proteïnes implicades en la metilació del DNA i en la seva funció reguladora. Per exemple, tant el consum de cocaïna i metamfetamina, com l'esquizofrènia o el trastorn bipolar s'han relacionat amb canvis d'expressió d'algunes isoformes de DNMT (Costa *et al.*, 2007; Anier *et al.*, 2010; Laplant *et al.*, 2010; Bodetto *et al.*, 2013; Jayanthi *et al.*, 2014) i en el cas de la proteïna MeCP2, s'han observat canvis d'expressió d'aquesta proteïna relacionats amb el consum de drogues, l'autisme o el TDAH (Nagarajan *et al.*, 2006; Im *et al.*, 2010; Lv *et al.*, 2013; Deng *et al.*, 2014). De fet, els ratolins genoanul·lats per la proteïna MeCP2 (un repressor de la transcripció que reconeix llocs CpG metilats) presenten deficiències locomotores i reduccions en la mida del cervell, característiques similars als símptomes de la síndrome de Rett, una patologia amb certs paral·lelismes amb els trastorns de l'espectre autista (Guy *et al.*, 2001; Collins *et al.*, 2004).

Taula 6. Canvis dels nivells de metilació de gens implicats en sistemes de neurotransmissió en trastorns psiquiàtrics. Adaptada de Liu et al., 2018

Gene	Locus	Funció	Teixit	Efecte en la metilació
Sistema dopaminèrgic				
<i>DRD2</i>	11q23.2	Receptor DA	Cervell, Sang	↓SZ, ↑AN
<i>DRD4</i>	11p15.5	Receptor DA	Cervell, Sang	↓SZ i TDAH
<i>DRD5</i>	4p16.1	Receptor DA	Sang	↓SZ
<i>MB-COMT</i>	22q11.21	Regulador dels nivells homeostàtics de DA en sinapsis	Cervell, Saliva	↓SZ i BP
<i>S-COMT</i>	22q11.21	Regulador dels nivells homeostàtics de DA en sinapsis	Cervell, Sang, Saliva	↑SZ
<i>SLC6A3 (DAT)</i>	5p15.33	Transportador DA	Sang	↑AN i BN, ↓AUD
<i>DBH</i>	9q34.2	DA beta-hidroxilasa	Sang	↓AUD
Sistema serotoninèrgic				
<i>HTR1A</i>	5q12.3	Receptor de serotonina	Sang	↑SZ i BD
<i>HTR2A</i>	13.q14.2	Receptor de serotonina	Cervell	↑SZ
<i>SLC6A4 (5-HTT)</i>	17q11.2	Transportador de serotonina	Cervell, Sang, Saliva	↑SZ, BD, MDD i TDAH
Sistema GABAèrgic				
<i>GAD1 (GAD67)</i>	22q31.1	Sintetitzador de GABA	Cervell, Sang	↓SZ, BD i AUD
<i>RELN</i>	7q22.1	Migració i posicionament neuronal	Cervell	↑SZ
<i>GABRP</i>	-	Receptor de GABA	Sang	↑AUD
Sistema glutamatèrgic				
<i>GRM2</i>	3p21.2	Receptor GLU metabotròpic	Cervell	↓SZ
<i>GRM5</i>	11q14.2.3	Receptor GLU metabotròpic	Cervell	↓SZ
<i>SCG2</i>	2q36.1	Estimulador de la secreció de GLU	Cervell	↓SZ
<i>VGLUT1</i>	19q13.33	Empaquetador de GLU en vesícules sinàptiques	Cervell	↑SZ
<i>VGLUT2</i>	11p14.3	Transportador de glutamat	Cervell	↓SZ
<i>SLC1A2 (EAAT2)</i>	11p13	Proteïna transportadora de soluts	Sang	↑BD
<i>BDNF</i>	11p14.1	Factor neurotròfic	Cervell, Sang	↑BD, SZ i BN

AN: Anorèxia, AUD: trastorn per ús d'alcohol, BD: trastorn bipolar; BN: bulímia nerviosa, DA: dopamina, GABA: àcid gamma-aminobutíric; GLU: glutamat, TDAH: Trastorn per dèficit d'atenció amb hiperactivitat, SZ: esquizofrènia

En aquesta Tesi, hem tingut especial interès amb un mecanisme epigenètic relacionat amb la metilació del DNA, la metilació específica d'al·lel.

Metilació específica d'al·lel

Anteriorment s'assumia que la metilació de dinucleòtids CpG era equivalent en els dos al·lells a les cèl·lules diploides amb l'excepció d'un número petit de gens imprintats (fenomen epigenètic pel qual certs gens s'expressen de manera específica segons l'origen parental) i regions subjectes al fenomen de la inactivació del cromosoma X en femelles. Però la realitat és molt més complexa a causa de l'existència de la metilació específica d'al·lel (*Allele-Specific Methylation, ASM*), característica bastant freqüent en el genoma humà (Meaburn, Schalkwyk i Mill, 2010).

L'ASM és un mecanisme epigenètic pel qual els nivells de metilació d'un lloc CpG depenen del genotip present en una determinada posició polimòrfica situada en *cis* o en *trans*. Per exemple, per a una variant genètica C/T, els genotips CC, CT o TT determinarien diferents nivells de metilació d'un dinucleòtid CpG situat a una certa distància. És un tret quantitatiu i heterogeni entre individus i teixits (Meaburn, Schalkwyk i Mill, 2010) que pot determinar una expressió específica d'al·lel (ASE) tot alterant l'activitat dels promotors de gens de la regió (Gaur *et al.*, 2013). A més, s'ha observat un enriquiment d'ASM en zones reguladores del genoma, remarcant el seu possible paper funcional (Benton *et al.*, 2019). L'ASM en *cis* s'hereta generalment de forma mendeliana, ja que en aquest cas dependria dels al·lells d'un polimorfisme genètic presents en una regió propera al lloc que es metila. Aquest mecanisme ens pot ajudar en la interpretació dels resultats dels GWAS, en què sovint es detecten variants genètiques associades a una patologia concreta que estan localitzades lluny de qualsevol gen, en regions no codificants (Tycko, 2010a, 2010b), i així desxifrar part de l'heretabilitat perduda d'alguns trastorns psiquiàtrics.

Al llarg dels últims anys, hi ha diferents estudis que s'han centrat en la identificació d'ASM en múltiples teixits (Do *et al.*, 2017), alguns d'ells en regions cerebrals (Gibbs *et al.*, 2010; Zhang *et al.*, 2010). A més, s'ha observat un enriquiment de variants ASM en diferents trastorns psiquiàtrics (Gamazon *et al.*, 2013; Hannon *et al.*, 2015), destacant la seva importància funcional. Finalment, alguns estudis recents han explorat la contribució de l'ASM a diferents trastorns psiquiàtrics com ara l'autisme (Andrews *et al.*, 2017), el trastorn bipolar (Chuang *et al.*, 2013) o l'esquizofrènia (Van Eijk *et al.*, 2015; Gagliano *et al.*, 2016; Wu i Pan, 2019), entre d'altres.

4.3. MicroRNAs

Els microRNAs (miRNAs), són molècules de RNA no codificants (ncRNAs) d'uns 22 nucleòtids de longitud. Es tracta de molècules altament conservades en l'evolució que actuen com a reguladors post-transcripcionals de l'expressió gènica. Es transcriuen a partir del DNA nuclear donant lloc a un precursor de diversos centenars de nucleòtids de longitud que pren una estructura de forquilla degut a la seva autocomplementarietat. Aquesta molècula precursora rep el nom de pri-miRNA i pot contenir d'un a sis precursors de miRNAs, és processada seguint diferents passos que tenen lloc tant al nucli com, posteriorment, al citoplasma, per generar una molècula funcional de cadena única d'RNA: el miRNA madur (Figura 8).

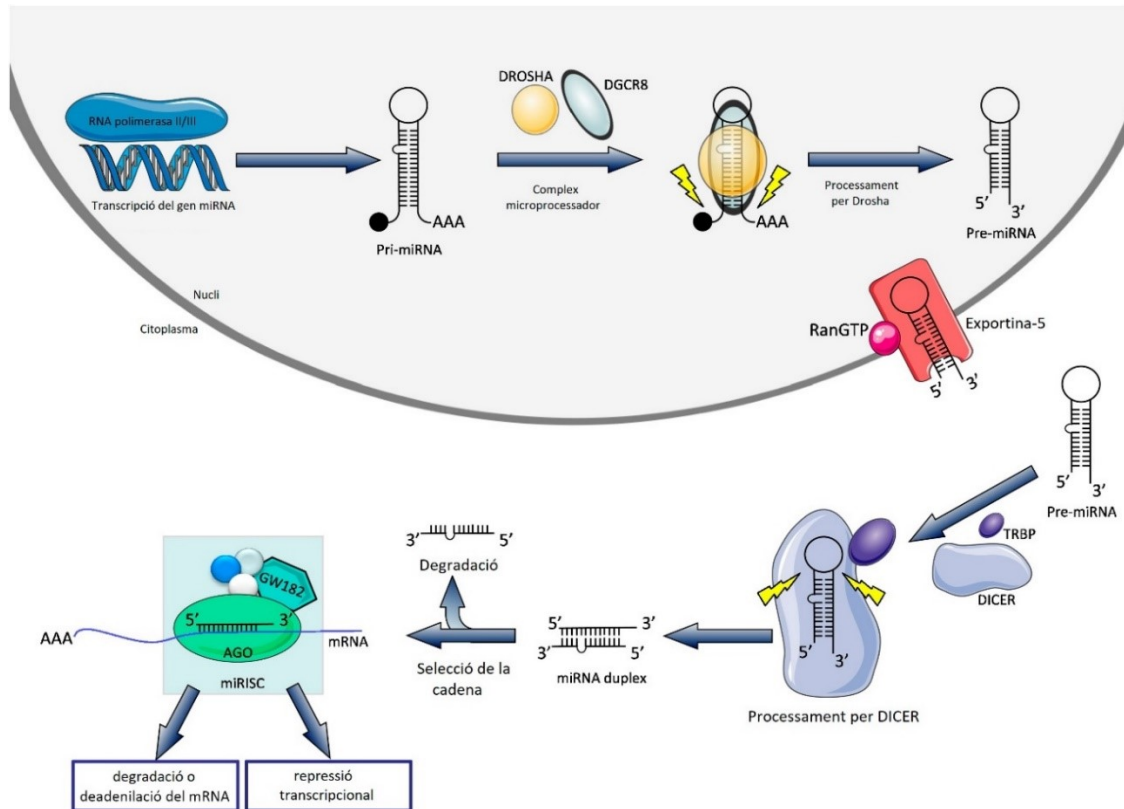


Figura 8. Biogènesi dels miRNAs. La transcripció dels gens que codifiquen miRNAs dona lloc als pri-miRNAs, que seran processats en el nucli cel·lular per un complex microprocessador que inclou les proteïnes DROSHA i DGCR8, tot produint unes estructures d'uns 60-100 nucleòtids anomenades pre-miRNAs. Aquest pre-miRNAs són translocats del nucli al citoplasma amb l'ajuda de l'exportina-5. Un cop al citoplasma, el pre-miRNA és processat per DICER, que talla el miRNA a nivell del *loop*, donant lloc a un dúplex de RNA de 19-24 nucleòtids de llargada. A continuació una de les dues cadenes es degrada i l'altra s'uneix al complex miRISC, que té com a components principals les proteïnes AGO, entre d'altres. El complex miRISC guiarà el miRNA cap al mRNA diana, modificant així l'expressió del mRNA mitjançant la deadenilació /degradació o la repressió traduccional. Adaptada de Luoni i Riva, 2016.

Els miRNAs funcionen per complementaritat de bases amb l'RNA missatger (mRNA), unint-se a la seva regió 3'UTR (*3'-untranslated region*), tot i que en alguns casos també es pot unir a la regió 5'UTR (*5'-untranslated region*) o a la regió codificant del mRNA (Lytle, Yario i Steitz, 2007; Lee *et al.*, 2009; Forman i Coller, 2010). La unió miRNA-mRNA bloqueja l'accés dels ribosomes al mRNA o indueix la degradació del mRNA, donant lloc a una disminució de la síntesi de la proteïna (Figura 8). A més de la funció citoplasmàtica canònica, hi ha nombroses evidències que demostren que els miRNAs poden actuar al nucli regulant l'estabilitat del mRNA al nuclèol i també el *splicing* alternatiu, així com activant o inhibint la transcripció dels gens diana (Catalanotto, Cogoni i Zardo, 2016). Un mateix miRNA pot regular l'expressió de 10 a 100 mRNAs diferents, i a la vegada, un mRNA pot estar regulat per desenes o centenars de miRNAs (Friedman *et al.*, 2009). Aquest fet, sumat al fet que l'acció dels miRNAs no és del tipus "tot o res" (*on-off switch*), sinó que permet la regulació fina de l'expressió del gen diana, fa que els miRNAs funcionin com un mecanisme de

control dinàmic de l'expressió gènica (Kim, Han i Siomi, 2009; Afonso-Grunz i Müller, 2015; Gulyaeva i Kushlinskiy, 2016).

Aquests darrers anys s'han anat acumulant estudis que demostren que els miRNAs són uns mediadors importants de nombrosos mecanismes clau del SNC, des del desenvolupament del cervell (dendritogènesi, formació de les espines o creixement de les neurites), fins la modulació i plasticitat sinàptica i la regulació d'activitats cerebrals superiors, incloses la memòria i l'aprenentatge (Schratt, 2009; Ceman i Saugstad, 2011; Meza-Sosa *et al.*, 2012; Wang, Kwon i Tsai, 2012; Sun i Shi, 2015; Luoni i Riva, 2016). L'any 2014 un estudi a gran escala va caracteritzar sistemàticament les interaccions entre miRNAs i mRNAs al cervell humà, mitjançant la tècnica d'immunoprecipitació contra la proteïna argonauta 2 (AGO2), una de les proteïnes del complex miRISC. En aquest treball es van identificar unes 7.000 regions d'unió a AGO2 que contenien 916 SNPs que alterarien la unió miRNA:mRNA (Boudreau *et al.*, 2014). En els últims anys s'ha demostrat que els miRNAs tenen un paper clau en la remodelació del sistema de recompensa induïda per les drogues, fet que probablement impulsa l'aparició de l'addicció i també alteracions de la plasticitat sinàptica (Smith i Kenny, 2018). A més, s'han descrit múltiples miRNAs específics relacionats amb les addiccions, com ara els miR-212, miR-181a, let7 o miRNA-124, que s'han relacionat àmpliament amb la dependència de cocaïna (Chandrasekar i Dreyer, 2009, 2011; Bali i Kenny, 2013).

En el cas del TDAH, tot i que encara hi ha pocs estudis centrats en investigar la funció dels miRNAs en el trastorn (Wu *et al.*, 2010, 2015, 2017; Németh *et al.*, 2013; Sánchez-Mora *et al.*, 2013; Kandemir *et al.*, 2014; Garcia-Martínez *et al.*, 2016), diverses evidències suggereixen que hi juguen un paper important, similar a l'observat en altres trastorns psiquiàtrics (Srivastav, Walitza i Grünblatt, 2018). Estudis d'expressió de miRNAs a gran escala realitzats en àrees cerebrals de pacients amb trastorns psiquiàtrics demostren l'existència de perfils de miRNAs específics segons l'àrea cerebral i el trastorn (Luoni i Riva, 2016). Per exemple, Smalheiser *et al.* 2012, van observar una disminució dels nivells d'expressió de miRNAs a l'escorça prefrontal de pacients amb depressió suïcida (Smalheiser *et al.*, 2012). La majoria d'estudis de miRNAs en trastorns psiquiàtrics però, s'han centrat en l'esquizofrènia, el trastorn bipolar i la depressió. Per exemple, la neurotrofina BDNF és un element clau en el fenotip depressiu, i s'ha descrit una reducció de la seva expressió per l'acció de diferents miRNAs: miR-30a-5p (Mellios *et al.*, 2008), miR-124a (Bahi, Chandrasekar i Dreyer, 2014), miR-16 (Bai *et al.*, 2012) i miR-182 (Y. Li *et al.*, 2016). Finalment, els sistemes de neurotransmissió també mostren alteracions de la seva expressió per l'acció de miRNAs en diferents trastorns psiquiàtrics. Per exemple, un estudi recent descriu una correlació

entre un increment de l'expressió del miRNA Let-7a i una reducció de l'expressió del receptor de serotonina 4 en l'hipocamp de rates que presenten un fenotip depressiu com a conseqüència de la privació materna (Bai *et al.*, 2014). Un altre estudi, realitzat en un model de ratolí de depressió, va detectar associacions entre el miR-135a i el transportador monoaminèrgic de serotonina SERT i el receptor de serotonina 5-HT-1A en un model de ratolí de depressió (Issler *et al.*, 2014).

Totes aquestes evidències suggereixen que els miRNAs poden influir en el desenvolupament de trastorns psiquiàtrics, una camp d'estudi que ha estat sent explorada durant els últims anys (Hunsberger *et al.*, 2009; Geaghan i Cairns, 2015; Issler i Chen, 2015; Luoni i Riva, 2016; Kuehner *et al.*, 2019).

5. Models animals

Finalment, tant en relació a la genètica com a l'epigenètica, els models animals són una aproximació àmpliament utilitzada. Un organisme model és una espècie no humana que s'utilitza per comprendre fenòmens biològics particulars, amb l'expectativa que els descobriments realitzats en l'organisme model proporcionin una visió del funcionament en l'organisme humà. Els organismes model s'utilitzen per investigar malalties humanes quan l'experimentació humana seria inviable o poc ètica.

El desenvolupament de models animals convincents i útils per estudiar trastorns psiquiàtrics representa un repte important a causa de l'heterogeneïtat que hi ha en els símptomes, la manca de biomarcadors i el coneixement incomplet de la seva neurobiologia i genètica. Al llarg dels anys, s'han seguit diferents metodologies per a l'obtenció de models animals en psiquiatria, que inclouen l'encreuament selectiu, l'enginyeria genètica, les lesions cerebrals i les manipulacions ambientals (Taula 7) (Nestler i Hyman, 2010).

Taula 7. Tipus d'aproximacions existents per a la creació de models animals en trastorns psiquiàtrics. Adaptada de Nestler i Hyman, 2010.

Mètode	Avantatges	Inconvenients
Genètic		
Encreuaments selectius	Focus en el fenotip d'interès	Pot donar lloc a una fenocòpia
Mutació aleatòria	Focus en el fenotip d'interès	Pot donar lloc a una fenocòpia
Models transgènics (<i>knock-out</i> , <i>knock-in</i> , <i>knock-down</i> ...)	Focus en el gen d'interès; pot reproduir alteracions genètiques humanes	Penetrància variable; pot ser difícil d'establir la rellevància del fenotip observat
Administració de gens al cervell mitjançant virus	Focus en el gen d'interès; control temporal i espacial	No reproduceix la causa genètica que es dona en humans
Farmacològic		
Administració d'agonistes i antagonistes de neurotransmissors	Focus en el sistema neurotransmissor d'interès; control temporal i espacial si l'administració és intracranial	No hi ha evidències que els trastorns mentals siguin causats per alteracions a només un sistema neurotransmissor
Ambiental		
Estrès social crònic	Pot reproduir factors de risc que es donen en humans	Falta de validesa d'especificitat per trastorns humans
Estrès físic crònic	Fàcil d'administrar	Falta de validesa del model per la majoria de trastorns humans
Estimulació elèctrica i lesions		
Estimulació del cervell, incloent optogenètica	Control temporal i espacial; pot reproduir efectes descrits en humans	Limitacions del coneixement de les alteracions en els circuits neuronals de cada trastorn
Lesions anatòmiques	Pot produir anomalies de conducta que recorden el trastorn humà	Falta d'evidències de lesions anatòmiques que causin el trastorn en humans

5.1. Models per l'estudi de les addiccions

S'han desenvolupat molts models animals que permeten investigar signes i símptomes específics de les addiccions, tot i que encara no hi ha cap model que mimetitzï a la perfecció la condició humana d'addicció. Els models animals desenvolupats recentment tenen en compte la diversitat de resposta a una droga existent entre individus i soques, permetent així l'estudi dels processos neurobiològics i dels factors ambientals que intervenen en el risc a l'addicció, ja sigui per generar vulnerabilitat o resiliència. També s'està començant a explorar la influència de l'etapa evolutiva de l'individu i del seu sexe en la resposta a les drogues, per entendre millor l'increment en la vulnerabilitat a l'addicció a les drogues quan l'ús s'inicia en l'adolescència i les diferents trajectòries d'ús de les drogues observades entre homes i dones (Koob i Volkow, 2016).

En les addiccions a les drogues, els tests utilitzats en models animals es basen en el condicionament, tant clàssic (o Pavlovià), com operant (Sanchis-Segura i Spanagel, 2006). Aquests models permeten valorar tant l'efecte reforçant de la droga, com altres aspectes relacionats amb l'addicció, entre els quals hi ha la sensibilització, la tolerància, la impulsivitat, la compulsivitat i la recaiguda després d'un període d'abstinència.

El model de preferència de lloc condicionada (CPP) és una forma de condicionament Pavlovià per mesurar els efectes motivacionals i de recompensa de les drogues d'abús. El test es duu a terme en un aparell que conté dos compartiments dissenyats perquè puguin ser clarament diferenciats i consisteix en tres fases: habituació, condicionament i test de preferència (Figura 9). En l'habitució l'animal té lliure accés als dos compartiments. En la fase de condicionament, els animals només tenen accés a un dels dos compartiments. En un dels compartiments sempre se subministrarà la droga d'abús, de manera que aquell compartiment s'associarà als efectes motivacionals; en canvi, en l'altre compartiment sempre se subministrarà una solució salina (Figures 9A i 9B). Finalment, en el test de preferència (CPP) l'animal té lliure moviment per poder escollir entre els dos compartiments (Figura 9C). Els efectes reforçants de l'estímul addictiu es defineixen per l'existència d'una preferència més alta pel compartiment on aquest se subministrava (Cunningham, Gremel i Groblewski, 2006; Sanchis-Segura i Spanagel, 2006; Olive i Kalivas, 2010).

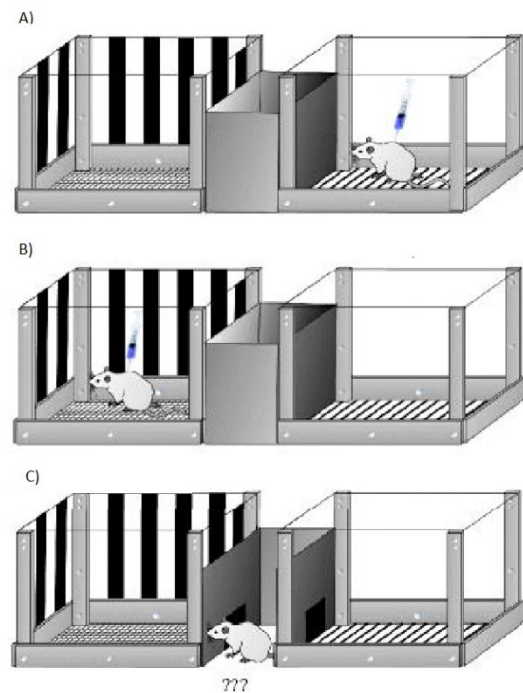


Figura 9. Esquema del model de preferència de lloc condicionada (CPP). A) Condicionament amb salí en un dels compartiments. B) Condicionament amb la droga d'abús en l'altre compartiment. C) Test de preferència. Adaptada de Olive i Kalivas, 2010.

El model d'autoadministració operant es realitza en una cambra adaptada per poder utilitzar catèters per administrar la droga d'abús i que a més, conté dues palanques (o forats), una inactiva i una activa (Figura 10A). La palanca activa administra la droga i va acompanyada d'un estímul condicionat (llum o so) que l'animal associarà a la droga. Primerament, s'estableix una raó fixa (FR) per a obtenir la droga i a continuació una raó progressiva (PR), en què l'animal ha de realitzar un nombre de respostes creixent per a obtenir la substància, per així valorar la motivació de l'animal per la droga. Aquest model és més motivacional que el CPP ja que l'animal ha de fer un esforç per obtenir la droga (Jacobs *et al.*, 2003; Sanchis-Segura i Spanagel, 2006) (Figura 10A).

Hi ha una altra versió d'aquest model anomenat model d'autoadministració operant acoblat a control (Figura 10B). Aquest model funciona igual que l'anterior, però inclou un altre animal

Introducció

acoblat a l'animal que s'autoadministra la droga, el qual rep la mateixa dosi de droga però de forma passiva. A més, es pot connectar un tercer subjecte que rep una infusió salina cada cop que el primer animal s'autoadministra la substància. L'animal acoblat no té control sobre el consum de les drogues, i per tant no genera cap aprenentatge. Així doncs, aquest model permet diferenciar entre els efectes directes de la droga d'abús d'aquells deguts a l'aprenentatge (Jacobs *et al.*, 2003; Sanchis-Segura i Spanagel, 2006) (Figura 10B).

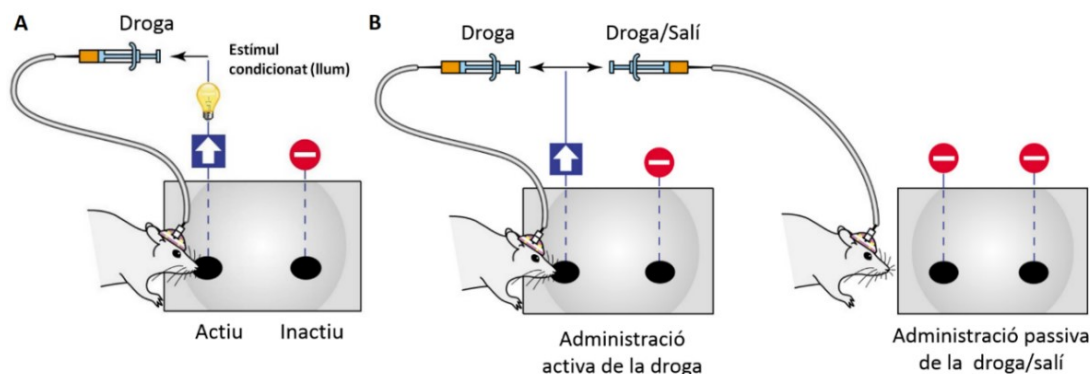


Figura 10. Esquema del model d'autoadministració. A) Model d'autoadministració B) Model d'autoadministració operant acoblat a control. Adaptada de Jacobs *et al.*, 2003.

Al llarg d'aquesta última dècada també s'ha validat un model animal d'addicció al menjar basat en el model d'autoadministració operant (Martín-García *et al.*, 2011; Mancino *et al.*, 2015; Domingo-Rodríguez *et al.*, 2020), en el qual, en comptes de administrar droga a través d'un catèter, s'administren pèl·lets de menjar saborós cada cop que es prem la palanca activa. En aquest model es tenen en compte tres criteris que es mesuren al principi i al final de l'aprenentatge i que definiran la resistència o vulnerabilitat a l'addicció: la persistència del comportament, la motivació i la resistència al càstig. Primerament s'estableixen cinc sessions amb un FR1 i a continuació es comencen les sessions d'autoadministració amb un FR5 que inclou un període de 10 minuts on no hi ha disponibilitat de pèl·lets per avaluar la persistència del comportament (Figura 11). Al principi i al final de l'entrenament es duu a terme un PR per avaluar la motivació pel menjar. Finalment, per avaluar la resistència al càstig, al principi i al final de l'entrenament també es realitza una sessió de FR5 en una cambra amb una ambientació diferent en la qual l'animal cada cop que rep un pèl·let, també rep un xoc elèctric de baixa intensitat als peus (Martín-García *et al.*, 2011; Mancino *et al.*, 2015; Domingo-Rodríguez *et al.*, 2020).

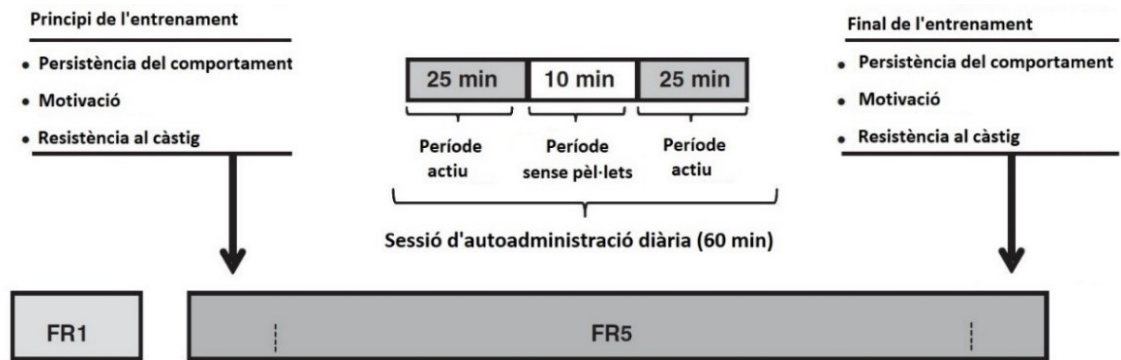


Figura 11. Disseny experimental del model d'autoadministració per l'addicció al menjar. Adaptada de Mancino et al., 2015.

Pel que fa als models transgènics, s'han estudiat molts gens relacionats amb els sistemes de neurotransmissió per la seva relació directa amb el sistema de recompensa i per tant amb l'establiment de la dependència, com els receptors i transportadors de DA, serotonina i noradrenalina. S'ha estudiat la contribució d'aquests gens en diferents addiccions a drogues en ratolins genoanul·lats, en homozigosi o en heterozigosi (falta de dues o una còpia del gen, respectivament) (Sora *et al.*, 2010). Per exemple, en ratolins genoanul·lats pel receptor de serotonina 5HT1B s'observava una major autoadministració de cocaïna i morfina (Crabbe *et al.*, 1996; Rocha *et al.*, 1998). Els ratolins genoanul·lats pel receptor de dopamina D₂ o pel receptor cannabinoid CB₁ presentaven una reducció de la recompensa com a resposta a la morfina (Maldonado *et al.*, 1997; Ledent *et al.*, 1999). Finalment, un últim exemple seria el dels ratolins genoanul·lats per la subunitat β_2 del receptor colinèrgic nicotínic, que mostraven una reducció de la recompensa en resposta a la cocaïna.

5.2. Models per l'estudi del TDAH

De manera semblant a les addiccions, no hi ha cap model animal que mimetizzi a la perfecció el TDAH, degut a la complexitat i heterogeneïtat del trastorn. Tot i així, hi ha molts models i paradigmes que es poden utilitzar pels símptomes principals del TDAH: el dèficit d'atenció, la hiperactivitat i la impulsivitat.

El paradigma de *Delay-discounting* s'utilitza com a mesura de la impulsivitat en la selecció, de manera que la impulsivitat és definida com una major tendència a triar una recompensa més petita abans que una recompensa més gran que s'obté més tard. Per altra banda, la *Stop-Signal Reaction Time task* (SSRT) també serveix com a mesura de la impulsivitat, en aquest cas tenint en compte la capacitat d'inhibir o retenir una resposta motora ja iniciada (Wickens, Hyland i Tripp, 2011).

La *Continuous Performance Task* (CPT) s'ha utilitzat per avaluar diferents tipus d'atenció en humans, i recentment s'ha pogut traslladar als models murins a través de l'ús de cambres amb pantalles tàctils. Igual que en molts CPT humans, és necessari que els ratolins responguin als estímuls del patró visual "objectiu" i retinguin respostes als estímuls "no objectius" durant les sessions en què els dos tipus d'estímuls es presenten seqüencialment en una ubicació única de la pantalla. En funció de la precisió en la realització i l'aprenentatge de la tasca es podrà avaluar l'atenció i la impulsivitat dels ratolins (Kim *et al.*, 2015).

La rata espontàniament hipertensa (SHR) de la soca Wistar-Kyoto es va desenvolupar inicialment com a model d'hipertensió (Okamoto i Aoki, 1963), però casualment també va mostrar una alta activitat motora espontània, suggerint que podria utilitzar-se com a model per al TDAH (Moser *et al.*, 1988). La SHR mostra molts dels símptomes principals del TDAH, com la impulsivitat, dificultats per aprendre o una capacitat d'espera reduïda (Sagvolden, 2000). Els dèficits observats en aquest model sembla que poden estar relacionats alteracions en els diferents sistemes de neurotransmissió. Per exemple, la SHR mostra una deficiència en la secreció de DA a l'escorça prefrontal, al NAc i al nucli caudat-putamen (Deutch i Roth, 1990; Russell, 2000). A més, també s'ha vist que mascles joves d'aquest model mostren un increment de la densitat de receptors D_1 i D_5 en el neocòrtex i al NAc (Carey *et al.*, 1998), i una reducció de l'expressió de D_4 a l'escorça prefrontal (Li *et al.*, 2007). Pel que fa al sistema noradrenèrgic, s'ha observat una concentració més elevada de noradrenalina al locus coeruleus, a la substància negra i a l'escorça prefrontal (de Villiers *et al.*, 1995) i una reducció de receptors beta-adrenoreceptors (Myers, Whitemore i Hendley, 1981). Tot i així, aquest model de TDAH no és perfecte ja que tenim un factor de confusió, la hipertensió.

Pel que fa a models transgènics pel TDAH, al llarg del temps se n'han generat molts, i alguns d'ells han estat validats (Taula 8) (de la Peña *et al.*, 2018). El model de ratolí genoanul·lat pel transportador de dopamina DAT (DAT-KO) és el model transgènic de TDAH més ben estudiat i presenta els tres símptomes principals del TDAH: hiperactivitat, dèficit d'atenció i impulsivitat. Al model de ratolí DAT-KO li manca el gen *Slc6a3*, que codifica el transportador de dopamina DAT, el responsable de la recaptació o eliminació de la DA de l'espai sinàptic. Aquest model es va desenvolupar per tal d'estudiar l'efecte terapèutic del metilfenidat i l'amfetamina, dos compostos utilitzats en el tractament del TDAH (Giros *et al.*, 1996; Leo i Gainetdinov, 2013).

Taula 8. Models de ratolins transgènics validats de TDAH. Adaptada de de la Peña et al., 2018.

Nom	Descripció	Síntomes
DAT-KO	Falta del transportador de dopamina DAT (<i>Slc6a3</i>)	H, DA i I
SNAP-25 mutant	Deleció de 20 gens al cromosoma 2 (incloent <i>Snap25</i>)	H, DA i I
NK1R-KO	Interrupció del receptor NK1R (<i>Tacr1</i>)	H, DA i I
TR β PV-KI	Addició del gen humà de la hormona tiroïdal β	H, DA i I
P35-KO	Eliminació del cofactor p35, activador de Cdk5	H

H: Hiperactivitat; DA: Dèficit d'atenció, I: Impulsivitat.

Malgrat que aquests models no poden mimetitzar a la perfecció el TDAH, ens proporcionen informació sobre el trastorn que no es pot obtenir a partir estudis humans a causa de diverses limitacions, tant ètiques com del coneixement existent de les bases etiològiques, neurobiològiques i genètiques del trastorn.

OBJECTIUS

Aquesta Tesi té com a objectiu principal contribuir a l'estudi i a la identificació de factors genètics i epigenètics de susceptibilitat a les addiccions i a trastorns comòrbids, tant en humans com en models animals. Els objectius concrets d'aquest treball són els següents:

CAPÍTOL 1. Estudi de factors genètics de risc a l'addicció a drogues identificats prèviament.

- 1.1. Realització d'un estudi de rèplica d'associacions a la dependència de drogues, i de cocaïna en particular, prèviament identificades en estudis d'associació genètica a escala genòmica.
- 1.2. Avaluació de l'efecte d'NFAT5, un gen candidat a l'addicció a cocaïna descrit prèviament pel grup, sobre la síntesi de la dopamina a través de la regulació de l'expressió del gen *DDC* en dues línies cel·lulars neurals dopaminèrgiques.

CAPÍTOL 2. Identificació de factors epigenètics de susceptibilitat a l'addicció a les drogues i al menjar.

- 2.1. Estudi de la contribució a la dependència de drogues de variants genètiques amb un efecte sobre la metilació específica d'al·lel (ASM).
- 2.2. Identificació de variants genètiques de susceptibilitat a la dependència de drogues que poden alterar el lloc d'unió de miRNAs, i posterior avaluació funcional de les variants identificades.
- 2.3. Exploració dels perfils transcriptòmics (mRNAs i miRNAs) i metilòmics en un model de ratolí d'addicció al menjar.

CAPÍTOL 3. Identificació de factors epigenètics de susceptibilitat a trastorns comòrbids a l'addicció a les drogues.

- 3.1. Exploració de la contribució de variants genètiques amb un efecte en l'ASM al trastorn per dèficit d'atenció amb hiperactivitat (TDAH).
- 3.2. Anàlisi de la funció de *C2orf82* al cervell i de la seva contribució al TDAH: determinació de l'expressió del gen en diferents àrees cerebrals de ratolí i caracterització fenotípica d'un model murí genoanul·lat.
- 3.3. Avaluació de l'efecte pleotròpic de variants genètiques que presenten ASM en diferents trastorns psiquiàtrics.

RESULTATS

INFORME DELS DIRECTORS DE TESI SOBRE LA CONTRIBUCIÓ DE LA DOCTORANDA A LES PUBLICACIONS D'AQUESTA TESI DOCTORAL

Títol de la Tesi: “Genètica i epigenètica de les addiccions i trastorns comòrbids”

Autora: Laura Pineda i Cirera

Directors: Dr. Bru Cormand Rifà i Dra. Noèlia Fernández Castillo

CAPÍTOL 1. Estudi de factors genètics de risc a l'addicció a drogues identificats prèviament.

Article 1

Pineda-Cirera L, Cabana-Domínguez J, Roncero C, Cozar M, Grau-López L, Abad AC, Martínez-Luna N, Robles-Martínez M, Sánchez-Mora C, Ramos-Quiroga JA, Casas M, Ribasés M, Fernández-Castillo N, Cormand B. *Evaluation of previous substance dependence genome-wide significant findings in a Spanish sample*. Drug and Alcohol Dependence 2018. 187:358-362. doi: 10.1016/j.drugalcdep.2018.03.013.

Índex de qualitat: IF₂₀₁₈: 3,466. 1r quartil de les categories SUBSTANCE ABUSE i PSYCHIATRY.

Aportació de l'autora a l'article: Participació en l'extracció de mostres de DNA genòmic de part dels casos i controls. Preparació de les mostres per genotipar, disseny i selecció de SNPs del panell de genotipació i realització de l'estudi d'associació i de l'estudi de metanàlisi. Redacció del primer esborrany del manuscrit, figures i taules, i participació en l'edició final.

Altres: Aquest article forma també part de la Tesi de la Dra. Judit Cabana Domínguez, titulada “Genètica i dependència de cocaïna: estudis d'associació, transcriptòmica i models animals.”

Article 2

Pineda-Cirera L, Cabana-Domínguez J, Benetó N, Diez H, Arenas C, Cormand B, Fernández-Castillo N. *DDC expression is not regulated by NFAT5 (TonEBP) in dopaminergic neural cell lines*. Gene 2020; 742:144569. doi: 10.1016/j.gene.2020.144569.

Índex de qualitat: IF₂₀₁₈: 2,638. 2n quartil de la categoria GENETICS & HEREDITY.

Aportació de l'autora a l'article: Realització del treball experimental en els dos models cel·lulars testats: cultius cel·lulars, localització subcel·lular per immunocitoquímica, generació i transducció de lentivirus, extracció d'RNA, avaluació de l'expressió de les isoformes de *DDC* i quantificació de l'expressió per RT-qPCR. Participació en l'anàlisi estadística dels resultats. Redacció dels materials i mètodes del primer esborrany del manuscrit, figures i taules, i participació en l'edició final.

CAPÍTOL 2. Identificació de factors epigenètics de susceptibilitat a l'addicció a les drogues i al menjar.

Article 3

Pineda-Cirera L, Cabana-Domínguez J, Grau-López L, Sánchez-Mora C, Daigre C, Palma-Álvarez RF, Ramos-Quiroga JA, Ribasés M, Cormand B, Fernández-Castillo N. *Exploring allele specific methylation in drug dependence susceptibility.*

Índex de qualitat: IF₂₀₁₈: 3,917. 1r quartil de la categoria PSYCHIATRY (sotmès, en segona revisió).

Aportació de l'autora a l'article: Participació en l'extracció de mostres de DNA genòmic de part dels casos i controls. Disseny i selecció de SNPs d'un dels dos panells de genotipació emprats. Preparació de les mostres per genotipar i realització de l'estudi d'associació. Redacció del primer esborrany del manuscrit, figures i taules, i participació en l'edició final.

Article 4

Cabana-Domínguez J, Roncero C, Pineda-Cirera L, R. Palma-Álvarez F, Ros-Cucurull E, Grau-López L, Esojo A, Casas M, Arenas C, Ramos-Quiroga JA, Ribasés M, Fernández-Castillo N, Cormand B. *Association of the PLCB1 gene with drug dependence.* Scientific Reports 2017; 7:10110. doi: 10.1038/s41598-017-10207-2.

Índex de qualitat: IF₂₀₁₇: 4,122. 1r quartil de la categoria MULTIDISCIPLINARY SCIENCES.

Aportació de l'autora a l'article: Participació en l'extracció de mostres de DNA genòmic de part dels casos i controls. Participació en la preparació de les mostres de la rèplica per a la seva genotipació i en l'edició final del manuscrit.

Altres: Aquest article també forma part de la Tesi de la Dra. Judit Cabana Domínguez, titulada "Genètica i dependència de cocaïna: estudis d'associació, transcriptòmica i models animals."

Article 5

Pineda-Cirera L*, Cabana-Domínguez J*, Domingo L*, Martín-García E, Maldonado R#, Fernández-Castillo N#, Cormand B#. *Gene expression and epigenetic profiling of a mouse model of food addiction* (Manuscrit en preparació).

*equally contributed to this work; #equally supervised this work

Aportació de l'autora a l'article: Participació en l'extracció de DNA, RNA i miRNA de les mostres i preparació de les mostres DNA i RNA per seqüenciar. Realització de l'anàlisi de MBDseq i d'anàlisis posteriors: càlcul de la probabilitat de la superposició de gens, recollida de dades funcionals dels gens destacats i identificació de regions d'interès. Redacció del primer esborrany del manuscrit, figures i taules, i participació en l'edició final.

CAPÍTOL 3. Identificació de factors epigenètics de susceptibilitat a trastorns comòrbids a l'addicció a les drogues.

Article 6

Pineda-Cirera L*, Shivalikanjli A*, Cabana-Domínguez J, Demontis D, Rajagopal VM, Børglum AD, Faraone SV, Cormand B, Fernández-Castillo N. *Exploring genetic variation that influences brain methylation in attention-deficit/hyperactivity disorder*. Translational Psychiatry 2019. 3;9(1):242. doi: 10.1038/s41398-019-0574-7.

*equally contributed to this work

Índex de qualitat: IF₂₀₁₈: 5,182. 1r quartil de la categoria PSYCHIATRY.

Aportació de l'autora a l'article: Selecció inicial d'SNPs i filtrat mitjançant dades genètiques i funcionals: obtenció de variants en alt desequilibri de lligament, anàlisi del context genètic de les variants destacades i inspecció de marques d'histones i d'eQTLs. Participació en la redacció del primer esborrany del manuscrit, figures i taules, i participació en l'edició final.

Article 7

Pineda-Cirera L, Cabana-Domínguez J, Fernández-Castillo N, Cormand B. *Identification of genetic variants influencing methylation in brain with pleiotropic effects on psychiatric disorders*. (Manuscrit en preparació).

Aportació de l'autora a l'article: Realització d'anàlisis posteriors a l'estudi d'associació: obtenció de variants en alt desequilibri de lligament, anàlisi del context genètic de les variants destacades i inspecció de marques d'histones i d'eQTLs. Redacció del primer esborrany del manuscrit, figures i taules, i participació en l'edició final.

Barcelona, 19 de Juny de 2020

Signatura dels directors



Dr. Bru Cormand Rifà



Dra. Noèlia Fernández Castillo

CAPÍTOL 1. Estudi de factors genètics de risc a l'addicció a drogues identificats prèviament.

Article 1

“Avaluació de variants prèviament associades a escala genòmica a la dependència de drogues en una mostra espanyola”

Resum:

La dependència de drogues és un trastorn crònic i recurrent produït per la interacció de factors de risc genètics i ambientals. L'objectiu d'aquest treball és replicar a la nostra mostra espanyola diverses associacions a escala genòmica identificades prèviament per altres grups mitjançant estudis GWAS de dependència de drogues o cocaïna. Hem realitzat un estudi d'associació cas-control en una mostra de 1.711 individus dependents de drogues (dels quals 1.011 són dependent de cocaïna) i 1.719 controls. Hem avaluat tres SNPs que presentaven associacions significatives a escala genòmica en tres estudis previs: rs1868152 i rs2952621 (situats prop dels gens *LOC253573* i *LOC151121*, respectivament) associats a la dependència de drogues, i rs2629540 (al primer intró del gen *FAM53B*) associat a la dependència de cocaïna. Hem pogut replicar l'associació entre la variant rs2952621 i la dependència de drogues ($P = 0,02$) sota el model dominant, en què l'al·lel de risc (T) és el mateix que s'havia descrit en l'estudi previ. A continuació hem realitzat una metanàlisi per analitzar aquesta variant conjuntament a les dues mostres de l'estudi original (*Collaborative Studies on Genetics of Alcoholism (COGA)* i *Study of Addiction: Genetics and Environment (SAGE)*) i a la nostra mostra. La metanàlisi inclou en total 3.747 casos i 4.043 controls, i confirma l'associació d'aquesta variant amb la patologia ($OR = 1,26$; $95\% CI = 1,15 - 1,39$). En conclusió, s'ha pogut replicar a la nostra mostra l'associació entre la variant rs2952621, situada prop del gen *LOC151121*, encara no caracteritzat, i la dependència de drogues. Serà necessari realitzar més estudis per poder determinar la contribució d'aquesta variant a la predisposició a la dependència de drogues.

Referència:

Pineda-Cirera L, Cabana-Domínguez J, Roncero C, Cozar M, Grau-López L, Abad AC, Martínez-Luna N, Robles-Martínez M, Sánchez-Mora C, Ramos-Quiroga JA, Casas M, Ribasés M, Fernández-Castillo N, Cormand B. *Evaluation of previous substance dependence genome-wide significant findings in a Spanish sample*. Drug Alcohol Depend. 2018 Jun 1;187:358-362. doi: 10.1016/j.drugalcdep.2018.03.013.



Contents lists available at ScienceDirect

Drug and Alcohol Dependence

journal homepage: www.elsevier.com/locate/drugalcddep

Short communication

Evaluation of previous substance dependence genome-wide significant findings in a Spanish sample



Laura Pineda-Cirera^{a,b,c,d}, Judit Cabana-Domínguez^{a,b,c,d}, Carlos Roncero^{e,f,g,h,i},
 Mònica Cozar^{a,b,c,d}, Lara Grau-López^{e,f,g,h}, Alfonso C. Abad^{f,h}, Nieves Martínez-Luna^{e,f,h},
 María Robles-Martínez^{f,h}, Cristina Sánchez-Mora^{g,h,j}, Josep Antoni Ramos-Quiroga^{g,h,j,k},
 Miquel Casas^{e,g,h}, Marta Ribasés^{g,h,j}, Noèlia Fernández-Castillo^{a,b,c,d,*}, Bru Cormand^{a,b,c,d,*}

^a Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 643, 08028 Barcelona, Catalonia, Spain

^b Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Av. Monforte de Lemos 3-5, Pabellón 11, 280029, Madrid, Spain

^c Institut de Biomedicina de la Universitat de Barcelona (IBUB), Av. Diagonal 643, 08028 Barcelona, Catalonia, Spain

^d Institut de Recerca Sant Joan de Déu (IR-SJD), Passeig de Sant Joan de Déu 2, 08950 Esplugues de Llobregat, Barcelona, Catalonia, Spain

^e Psychiatry Department, Universitat Autònoma de Barcelona, 08193 Bellaterra, Catalonia, Spain

^f Addiction and Dual Diagnosis Unit, Psychiatric Services, Hospital Universitari Vall d'Hebron-ASPB, Passeig Vall d'Hebron 119-129, Edifici Mediterrània, 08035 Barcelona, Catalonia, Spain

^g Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Instituto de Salud Carlos III, Av. Monforte de Lemos 3-5, Pabellón 11, 280029 Madrid, Spain

^h Department of Psychiatry, Hospital Universitari Vall d'Hebron, Passeig Vall d'Hebron 119-129, Edifici Mediterrània, 08035 Barcelona, Catalonia, Spain

ⁱ Psychiatric Service, University of Salamanca Health Care Complex and Institute of Biomedicine of Salamanca, University of Salamanca, Castile and Leon, Spain

^j Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, 08193 Bellaterra, Catalonia, Spain

^k Department of Psychiatry and Legal Medicine, Universitat Autònoma de Barcelona, 08193 Bellaterra, Catalonia, Spain

ARTICLE INFO

Keywords:

Replication study
 Genome-wide association study (GWAS)
 Substance dependence
 Cocaine dependence
 rs2952621
 Meta-analysis

ABSTRACT

Background: Substance dependence is a chronic and relapsing disorder explained by genetic and environmental risk factors. The aim of our study is to replicate previous genome-wide significant (GWS) hits identified in substance dependence in general or in cocaine dependence in particular using an independent sample from Spain.

Methods: We evaluated, in a Spanish sample of 1711 subjects with substance dependence (1011 of them cocaine dependent) and 1719 control individuals, three SNPs identified as GWS in previous studies: rs1868152 and rs2952621 (located near *LINC02052* and *LINC01854*, respectively), associated with substance dependence, and rs2629540 (in the first intron of *FAM53B*), associated with cocaine dependence.

Results: We replicated the association between rs2952621 and substance dependence under the dominant model ($P = 0.020$), with the risk allele (T) being the same in our sample and in those two reported previously. We then performed a meta-analysis of the two samples used in the original study that reported the association of rs2952621 with substance dependence (Collaborative Studies on Genetics of Alcoholism (COGA) and Study of Addiction: Genetics and Environment (SAGE)) together with our Spanish sample. The meta-analysis of 3747 cases and 4043 controls confirmed the association (OR = 1.26, 95% CI = 1.15–1.39).

Conclusions: The rs2952621 variant, located downstream from the yet uncharacterized gene *LINC01854*, is associated with substance dependence in our Spanish sample. Further research is needed to understand its contribution to the susceptibility to substance dependence.

1. Introduction

Substance dependence is a complex psychiatric disorder characterized

by loss of control in drug intake, craving, and withdrawal (Koob and Volkow, 2010). The heritability estimates of substance dependence range from 39% to 72% depending on the drug of abuse (Goldman et al., 2005),

* Corresponding authors at: Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Avinguda Diagonal 643, edifici annex, 3^a planta, 08028 Barcelona, Catalonia, Spain.

E-mail addresses: noefernandez@ub.edu (N. Fernández-Castillo), bcormand@ub.edu (B. Cormand).

<https://doi.org/10.1016/j.drugalcddep.2018.03.013>

Received 5 November 2017; Received in revised form 2 March 2018; Accepted 8 March 2018

Available online 16 April 2018

0376-8716/ © 2018 Elsevier B.V. All rights reserved.

although other factors such as clinical assessment, ethnicity, or age may also have an impact on these scores. Despite the fact that the genetic contribution to the disorder is substantial, the specific underlying genetic risk factors remain largely unknown. Candidate gene association studies assessing susceptibility to dependence to drugs of abuse have focused mainly on genes involved in the neurotransmitter systems of the brain, but only a few consistent associations have been found. These include the genes *CNR1*, *CHRNA5* and *DRD2* (reviewed by Bühler et al., 2015). On the other hand, Genome-Wide Association Studies (GWAS) have highlighted new candidate genes such as *LINC01854*, *ARHGAP28* (Wetherill et al., 2014), or *KAT2B* (Johnson et al., 2015) for substance dependence and *FAM53B* (Gelernter et al., 2014) for cocaine dependence. However, most GWAS have been performed for alcohol dependence (Tawa et al., 2016). These GWAS have highlighted some risk genes that had already emerged through candidate gene studies, like some members of the *ADH* family and *ALDH2*, encoding alcohol and aldehyde dehydrogenases, respectively (Samochowiec et al., 2014). Others have emerged that were not *a priori* candidates for the disorder, pointing at new functions that may be related to its etiology.

Replication of previous GWS findings in independent samples is key to validating association signals, especially when they have been identified in studies with limited sample sizes, as is the case for most studies performed in substance dependence so far. The aim of the present study is to replicate, in a Spanish sample, several GWS findings reported in previous GWAS of substance dependence in general and dependence to cocaine performed in samples with European ancestry (Gelernter et al., 2014; McGue et al., 2013; Wetherill et al., 2014). Our analyses include meta-analytical approaches and functional annotation of replicated findings.

2. Materials and methods

2.1. Samples and DNA isolation

The case sample consists of 1711 substance dependent patients, including 1011 cocaine-dependent subjects (Supplementary Table S1 and Supplementary Fig. 1). All patients were diagnosed under DSM-IV-TR criteria (Diagnostic and Statistical Manual of Mental Disorders, 4th ed. Text revision) and the SCID (Structured Clinical Interview) (First et al., 1997) at the Addiction and Dual Diagnosis Unit of the Psychiatry Department of Hospital Universitari Vall d'Hebron (Barcelona, Spain). They were included in the study if they met the criteria for dependence for at least one drug of abuse (Supplementary Table S1 and Supplementary Fig. 1). The control sample consisted of 1719 individuals (Supplementary Table S1) recruited at the Blood and Tissues Bank of Barcelona. All cases and controls were unrelated, Spanish, Caucasian, and sex-matched. The study was approved by the ethics committee of our institution according to the Helsinki Declaration. Population stratification of our sample was discarded in a previous study of our group (Fernández-Castillo et al., 2013) by genotyping 48 unlinked anonymous SNPs located at least 100 kb from any known gene and analyzing the results using the STRUCTURE software (Pritchard et al., 2000), the Fst coefficient (Goudet, 1995), and the Pritchard and Rosenberg method (Pritchard and Rosenberg, 1999).

Genomic DNA was isolated from peripheral blood lymphocytes using the salting-out method (Miller et al., 1988). DNA concentration of all samples was measured on a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Termo Fisher Scientific Inc., Wilmington, DE, USA).

2.2. SNP selection and genotyping

We selected SNPs that showed a GWS association ($P < 5e-08$) with substance dependence (dependence to at least one drug of abuse) in samples with European ancestry (see Supplementary Table S2 for details). Two SNPs were identified under these criteria and were selected for genotyping: rs1868152, located 5' from *LINC02052* (McGué et al.,

2013), and rs2952621, located 3' from *LINC01854* (Wetherill et al., 2014). Both were associated with substance dependence. We also considered a third SNP that was GWS for cocaine dependence in a meta-analysis of samples of African- and European-American individuals: rs2629540, within the first intron of the *FAM53B* gene (Gelernter et al., 2014). This previous study included two additional SNPs, rs200085570 and rs2456778, showing GWS associations in some of the individual samples but not in the meta-analysis, and they were not considered in our replication study. Genotyping of all SNPs was performed using KASP technology (LGC genomics, UK). Replicates of some samples and negative controls were included.

2.3. Statistical analyses

The statistical power of our samples was calculated *post hoc* using the Genetic Power Calculator software (<http://csg.sph.umich.edu/abecasis/cats/>) (Purcell et al., 2003). We calculated the power under the additive, dominant, and recessive models using a level of significance of 0.05, odds ratio of 1.2, a disease prevalence of 0.026 (Compton et al., 2007), and the minor allele frequency (MAF) calculated in our control sample (MAF = 0.1 for rs1868152, MAF = 0.42 for rs2952621 and MAF = 0.27 for rs2629540).

Departure from Hardy-Weinberg equilibrium (HWE) ($P < 0.05$) was tested for all the SNPs separately in controls and cases using the SNPpass package of the R library (González et al., 2007). The case-control association test was carried out using the same package under four different genetic models: additive, dominant, recessive and codominant. All results were adjusted by age, as significant differences in age were detected between cases and controls using the non-parametric test U Mann-Whitney with SPSS22 (SPSS Inc., Chicago, IL, USA). Bonferroni correction threshold for multiple testing was set at $P < 4.16e-03$ (0.05/(3 SNPs*4 genetic models)).

Finally, for rs2629540, we also performed a symptom count analysis similar to the one reported in the original study (Gelernter et al., 2014) by including the severity of the consumption of opioid, alcohol, and nicotine as covariates.

2.4. Meta-analysis

Meta-analysis for rs2952621 (3747 cases and 4043 controls) was performed with the meta R package (www.cran.r-project.org/web/packages/rmeta/index.html) for three different substance dependence samples: the Collaborative Studies on Genetics of Alcoholism (COGA) (genotype data from Wetherill et al., 2014), the Study of Addiction: Genetics and Environment (SAGE) (data obtained through dbGAP application, accessions phs000092.v1.p1.c1 and c2, under project 15342), and our Spanish sample. The samples gathered by COGA and SAGE have been widely used in substance dependence GWAS either as discovery or as replication samples. The COGA sample included 824 cases and 935 controls. For the SAGE sample, which consisted of 1224 cases and 1390 controls, we used European-American individuals and excluded related individuals and subjects also present in the COGA sample from the analysis. We considered the dominant model, which displayed the best *p*-value in the Spanish sample. The Mantel-Haenszel test (fixed effects model) was applied, as the samples did not show heterogeneity.

2.5. Functional annotation

We searched for other variants in linkage disequilibrium (LD) with rs2952621 using the SNAP tool (<http://archive.broadinstitute.org/mpg/snap/ldsearch.php>) (Johnson et al., 2008), considering $r^2 \geq 0.8$ and the genotype data from the Utah Residents (CEPH) with Northern and Western European Ancestry (CEU) population of 1000 Genomes Pilot 1. The predicted functional effect of all these variants (38 in total) was evaluated using the SNP Function Prediction software ([359](https://</p>
</div>
<div data-bbox=)

Table 1
Replication study of three genome-wide significant associations in a sample of Caucasian Spanish subjects with substance dependence and controls.

SNP	Phenotype	N		Alleles	MAF	Genotypes						p-value			
		Cases	Controls			Cases (%)			Controls (%)			Codominant	Dominant	Recessive	Additive
						11	12	22	11	12	22				
rs1868152	SD	1698	1707	C > T	0.1	79.6	19.3	1.1	79.6	19.4	1.0	0.8039	0.7338	0.6270	0.8401
rs2952621	SD	1699	1718	C > T	0.42	31.3	50.7	18.0	35.0	46.7	18.3	0.0488	0.0196	0.9675	0.1109
rs2629540	CD	1011	1144	C > G	0.27	53.7	38.7	7.6	53.9	38.4	7.7	0.6347	0.4154	0.8258	0.5837

Bold values: significant *p*-values; MAF: Minor Allele Frequency; SD: Substance Dependence; CD: Cocaine Dependence.

snpinfo.nih.gov/snpinfo/snpfunc.html) in the CEU population. These SNPs were tested as potential quantitative trait loci (eQTLs) using GTEx Analysis, Release V6p (GTEx Consortium et al., 2013). Histone modification data for H3K27ac, H3K9ac, H3K9me1 or H3K9me3 of 10 different brain regions (hippocampus middle, substantia nigra, anterior caudate, cingulate gyrus, inferior temporal lobe, angular gyrus, dorso-lateral prefrontal cortex, germinal matrix, and male and female fetal brain) related to enhancer or promoter regions were explored using the Haploreg v4.1 tool (Ward and Kellis, 2016).

3. Results

We assessed three previous genome-wide significant findings identified in GWAS of substance dependence in general or cocaine dependence in particular (rs1868152, rs2952621, and rs2629540, see Supplementary Table S2) in a Spanish Caucasian sample of substance dependence (59% of the subjects being cocaine-dependent) and the corresponding controls (Supplementary Table S1 and Supplementary Fig. 1).

The statistical power of our sample, assuming an OR of 1.2, was 6–68% (rs1868152), 40–97% (rs2952621), and 12–79% (rs2629540), depending on the model of inheritance considered. All three SNPs were in Hardy-Weinberg Equilibrium (HWE) in controls and cases.

We identified an association between rs2952621 and substance dependence under the dominant model (best *p*-value, *P* = 0.020) and the codominant model (*P* = 0.049), neither of which overcame the Bonferroni correction for multiple testing (Table 1). These nominal associations showed the same direction as in the original study (Wetherill et al., 2014), with the T allele as the risk variant. As the association was described under the additive model in the original study, we inspected the genotypes of COGA (discovery) and SAGE (replication) samples and observed that the associations were also positive for the dominant model (*P* = 3.3e-04 and *P* = 9.9e-03, respectively) and the codominant model (*P* = 1.0e-05 and *P* = 0.035,

respectively). Then, we performed the meta-analysis of the COGA, SAGE, and Spanish samples (3747 cases and 4043 controls in total) under the dominant model, which yielded the best association results in most samples. The association was confirmed in the meta-analysis (OR = 1.26, 95% CI = 1.15–1.39). All three datasets showed the same direction of the effect, and T was the risk allele (Fig. 1). We did not observe significant associations for rs1868152 or rs2629540 under any of the tested genetic models (Table 1).

We then assessed the possible functional effect of the rs2952621 variant using bioinformatics tools. This SNP is in high LD ($r^2 \geq 0.8$) with 37 SNPs (Supplementary Table S3). Our SNP of interest, and those located closest to it, overlap with multiple histone marks classically related with enhancer regions in different brain areas (Supplementary Table S3). On the other side, although rs2952621 is not an eQTL and is not predicted to have any functional effect, eight SNPs out of those 37 that are in LD with it are located in predicted binding sites for transcription factors, and two of them overlap with splice sites within the *LINC01854* gene. Finally, 11 out of these 37 SNPs are eQTLs for the pseudogene *FAR2P1* (located 790 kb distal from rs1251179) in the nerve-tibial tissue.

4. Discussion

The aim of this study is to follow-up previous GWAS results on substance dependence in general or on cocaine dependence through a replication study in a Spanish sample. One of the three GWS signals, the rs2952621 SNP associated with substance dependence (Wetherill et al., 2014), was replicated in our study. We did not replicate the associations reported for rs1868152 with substance dependence or rs2629540 with cocaine dependence.

The original study (Wetherill et al., 2014) identified an association between rs2952621 and substance dependence (*P* = 1.77e-08) in a discovery sample (COGA, Collaborative Study on the Genetics of Alcoholism) of around 800 cases and 900 controls. The authors replicated these findings in a second sample of 2647 individuals (SAGE, Study of Addiction: Genetics and Environment) with the T allele as the risk variant in both groups (Wetherill et al., 2014). Our study replicates this finding in the same direction in a Spanish Caucasian sample (*P* = 0.020, dominant model). Finally, the meta-analysis of the three samples showed an effect for the risk allele T with an OR = 1.26 (95% CI = 1.15–1.39) (Fig. 1).

The rs2952621 SNP is located 1.3 Kb downstream from a yet uncharacterized gene, *LINC01854*, identified in two previous studies (Ota et al., 2004; Strausberg et al., 2002). It is expressed in most tissues, including the brain. Unfortunately, no further information is available about the function of this gene to understand its possible role in substance dependence. Rs2952621 seems to be located in an enhancer region, but it is not described as an eQTL and does not have a predicted functional impact. Several SNPs are in high LD with it and were explored for potential functional involvement. Among them, rs1251179 seems to be the most interesting one: the degree of LD with rs2952621 is high ($r^2 = 0.97$ in the CEU population), it is predicted to be located in a transcription factor binding site, and it is described as an eQTL in the

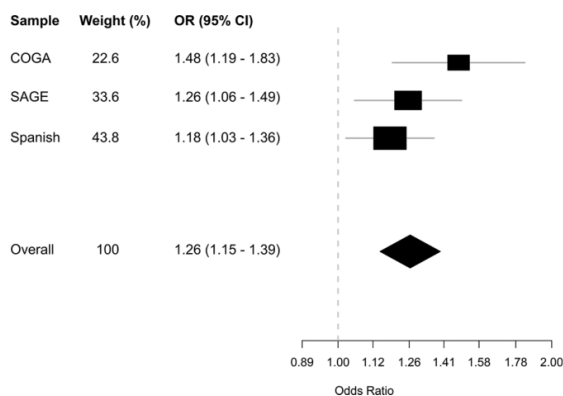


Fig. 1. Meta-analysis of rs2952621. Forest plot of the ORs (95% CI) of marker rs2952621 for our Spanish sample and the COGA and SAGE datasets.

nerve-tibial tissue for the pseudogene *AC018865.9* (located 790 kb distal from rs1251179). Further research is needed to clarify whether it could have a role in the susceptibility to substance dependence.

The lack of replication of the previously associated SNPs rs1868152 and rs2629540 could be explained by phenotypic heterogeneity of the substance dependence samples used in the different studies or by lack of statistical power due to limited sample size and/or modest effect size of the variants. In the case of rs2629540, the reported association (Gelernter et al., 2014) reached GWS in a meta-analysis of European-American and African-American individuals ($P = 4.3e-08$). However, when the two ethnic groups were analyzed separately, the association with rs2629540 was only nominal, with the African sample performing much better ($P = 1.4e-06$) than the European one ($P = 2.6e-03$) with comparable sample sizes. The lack of replication in our sample, which is also of European origin, would make sense in this context.

5. Conclusions

We replicated the reported association between substance dependence and the variant rs2952621, located close to the uncharacterized gene *LINC01854*. A meta-analysis of our sample and the discovery and replication samples of the original study confirmed the association and the direction of the effect. This risk variant for substance dependence (T allele) may contribute to the susceptibility to this chronic relapsing disorder, although further studies are required to understand the molecular basis of this contribution.

Role of funding source

Major financial support for this research was received by BC from the Spanish 'Ministerio de Economía y Competitividad' (SAF2015-68341-R) and AGAUR, 'Generalitat de Catalunya' (2017-SGR-738). CR received financial support from 'Instituto de Salud Carlos III', Spain (PI13/1911) and 'Plan Nacional Sobre Drogas', Spain (2013/044). The research leading to these results also received funding from the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement no 602805 and from the European Union H2020 Program [H2020/2014-2020] under grant agreement no 667302, from 'Instituto de Salud Carlos III', Spain (PI17/00289), the European Regional Development Fund (ERDF), AGAUR, 'Generalitat de Catalunya' (2014SGR1461), 'Departament de Salut, Generalitat de Catalunya' and a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation to MR. LP-C and JC-D were supported by 'Generalitat de Catalunya' (2016 FI_B 00728 and 2015 FI_B 00448, respectively). LP-C was also supported by 'Ministerio de Educación, Cultura y Deporte' (FPU15/03867). NF-C was supported by a contract of the 'Centro de Investigación Biomédica en Red de Enfermedades Raras' (CIBERER). MR and CS-M are recipients of a Miguel de Servet (CP09/00119) and a Sara Borrell (CD15/00199) contract, respectively, from 'Instituto de Salud Carlos III', Spain.

Contributors

LP-C performed the association study and the statistical analysis; LP-C and JC-D performed the meta-analysis; LG-L, CR, ACA, NM-L, MR-M, JAR-Q and MC participated in the recruitment of patients and clinical assessment and coordinated the clinical research; LP-C, JC-D, CS-M and NF-C isolated genomic DNA from samples; LP-C, BC and NF-C designed the study; MR contributed to the genetic analysis. LP-C prepared the first draft of the manuscript and all figures and tables; BC and NF-C coordinated the study and supervised the manuscript preparation. All authors contributed to and approved the final version of the manuscript.

Conflict of interest

JAR-Q was on the speakers' bureau and/or acted as a consultant for Eli-Lilly, Janssen-Cilag, Novartis, Shire, Lundbeck, Almirall, Braingaze, Sincrolab, and Rubió in the last 5 years. He also received travel awards for taking part in psychiatric meetings from Janssen-Cilag, Rubió, Shire, and Eli-Lilly. The Department of Psychiatry chaired by him received unrestricted educational and research support from the following pharmaceutical companies in the last 5 years: Eli-Lilly, Lundbeck, Janssen-Cilag, Actelion, Shire, Ferrer, and Rubió. The rest of the authors declare no conflicts of interest or relevant financial interests.

Acknowledgments

We are thankful to all the Blood and Tissues Bank of Barcelona donors, patients, families and clinical collaborators who contributed to the recruitment of patients (Dr. Elena Ros-Cucurull, Abderraman Esojo, R. Felipe Palma-Álvarez, Joan Alvarós, Ángel Egido and Laia Rodríguez-Cintas). We are also grateful to Leah Wetherill for kindly providing the genotype data of the COGA sample.

Funding support for the Study of Addiction: Genetics and Environment (SAGE) was provided through the NIH Genes, Environment and Health Initiative [GEI] (U01 HG004422). SAGE is one of the genome-wide association studies funded as part of the Gene Environment Association Studies (GENEVA) under GEI. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Support for collection of datasets and samples was provided by the Collaborative Study on the Genetics of Alcoholism (COGA; U10 AA008401), the Collaborative Genetic Study of Nicotine Dependence (COGEND; P01 CA089392), and the Family Study of Cocaine Dependence (FSCD; R01 DA013423). Funding support for genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research, was provided by the NIH GEI (U01HG004438), the National Institute on Alcohol Abuse and Alcoholism, the National Institute on Drug Abuse, and the NIH contract "High throughput genotyping for studying the genetic contributions to human disease" (HHSN268200782096C). The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000092.v1.p1 through dbGaP accession number phs000092.v1.p.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.drugalcdep.2018.03.013>.

References

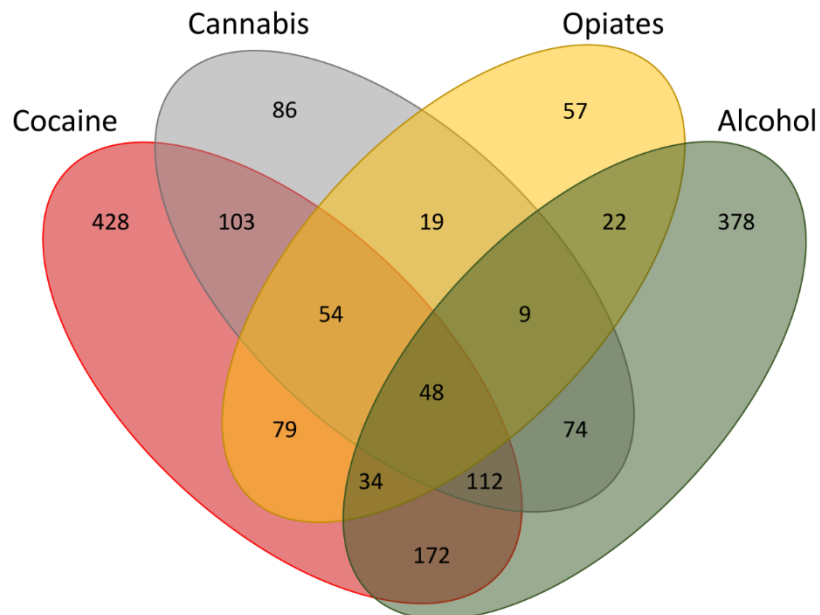
- Bühler, K.-M., Giné, E., Echeverry-Alzate, V., Calleja-Conde, J., de Fonseca, F.R., López-Moreno, J.A., 2015. Common single nucleotide variants underlying drug addiction: more than a decade of research. *Addict. Biol.* 20, 845–871.
- Compton, W.M., Thomas, Y.F., Stinson, F.S., Grant, B.F., 2007. Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Arch. Gen. Psychiatry* 64, 566–576.
- Fernández-Castillo, N., Roncero, C., Grau-Lopez, L., Barral, C., Prat, G., Rodríguez-Cintas, L., Sánchez-Mora, C., Gratacòs, M., Ramos-Quiroga, J.A., Casas, M., Ribasés, M., Cormand, B., 2013. Association study of 37 genes related to serotonin and dopamine neurotransmission and neurotrophic factors in cocaine dependence. *Genes Brain Behav.* 12, 39–46.
- First, M.B., Spitzer, S.R., Gibbon, M., Williams, J.B.M., 1997. Structured Clinical Interview for DSM-IV Axis I Disorders (SCID). American Psychiatric Press, Washington, D.C.
- GTEx Consortium, Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron, J., Fleming, J., Siminoff, L., Traino, H., Mosavel, M., Barker, L., et al., 2013. The genotype-Tissue expression

- (GTEx) project. *Nat. Publ. Gr.* 45, 580–585.
- Gelernter, J., Sherva, R., Koesterer, R., Almasy, L., Zhao, H., Kranzler, H.R., Farrer, L., 2014. Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Mol. Psychiatry* 19, 717–723.
- Goldman, D., Oroszi, G., Ducci, F., 2005. The genetics of addictions: uncovering the genes. *Nat. Rev. Genet.* 6, 521–532.
- González, J.R., Armengol, L., Solé, X., Guinó, E., Mercader, J.M., Estivill, X., Moreno, V., 2007. SNPpass: an R package to perform whole genome association studies. *Bioinformatics* 23, 644–645.
- Goudet, J., 1995. FSTAT (Version 1.2): a computer program to calculate F-statistics. *J. Hered.* 86, 485–486.
- Johnson, A.D., Handsaker, R.E., Pulit, S.L., Nizzari, M.M., O'Donnell, C.J., de Bakker, P.I.W., 2008. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 24, 2938–2939.
- Johnson, E.O., Hancock, D.B., Levy, J.L., Gaddis, N.C., Page, G.P., Glasheen, C., Saccone, N.L., Bierut, L.J., Kral, A.H., 2015. KAT2B polymorphism identified for drug abuse in African Americans with regulatory links to drug abuse pathways in human prefrontal cortex. *Addict. Biol.* 21, 1217–1232.
- Koob, G.F., Volkow, N.D., 2010. Neurocircuitry of addiction. *Neuropsychopharmacology* 35, 217–238.
- McGue, M., Zhang, Y., Miller, M.B., Basu, S., Vrieze, S., Hicks, B., Malone, S., Oetting, W.S., Iacono, W.G., 2013. A genome-wide association study of behavioral disinhibition. *Behav. Genet.* 43, 363–373.
- Miller, S.A., Dykes, D.D., Polesky, H.F., 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16, 1215.
- Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., et al., 2004. Complete sequencing and characterization of 21, 243 full-length human cDNAs. *Nat. Genet.* 36, 40–45.
- Pritchard, J.K., Rosenberg, N.A., 1999. Use of unlinked genetic markers to detect population stratification in association studies. *Am. J. Hum. Genet.* 65, 220–228.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Purcell, S., Cherny, S.S., Sham, P.C., 2003. Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19, 149–150.
- Samochowiec, J., Samochowiec, A., Puls, I., Bienkowski, P., Schott, B.H., 2014. Genetics of alcohol dependence: a review of clinical studies. *Neuropsychobiology* 70, 77–94.
- Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.L., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L., et al., 2002. Generation and initial analysis of more than 15, 000 full-length human and mouse cDNA sequences. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16899–16903.
- Tawa, E.A., Hall, S.D., Lohoff, F.W., 2016. Overview of the genetics of alcohol use disorder: alcohol Alcohol. *Natl. Inst. Health* 51, 507–514.
- Ward, L.D., Kellis, M., 2016. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* 44, D877–D881.
- Wetherill, L., Agrawal, A., Kapoor, M., Bertelsen, S., Bierut, L.J., Brooks, A., Dick, D., Hesselbrock, M., Hesselbrock, V., Koller, D.L., Le, N., Nurnberger, J.I., Salvatore, J.E., Schuckit, M., Tischfield, J.A., Wang, J.C., Xuei, X., Edenberg, H.J., Porjesz, B., Bucholz, K., Goate, A.M., Foroud, T., 2014. Association of substance dependence phenotypes in the COGA sample. *Addict. Biol.* 20, 617–627.

Supplementary material for:

EVALUATION OF PREVIOUS SUBSTANCE DEPENDENCE GENOME-WIDE SIGNIFICANT FINDINGS IN A SPANISH SAMPLE

Laura Pineda-Cirera, Judit Cabana-Domínguez, Carlos Roncero, Mònica Cozar, Lara Grau-López, Alfonso C. Abad, Nieves Martínez-Luna, María Robles-Martínez, Cristina Sánchez-Mora, Josep Antoni Ramos-Quiroga, Miquel Casas, Marta Ribasés, Noèlia Fernàndez-Castillo, Bru Cormand.



Supplementary Figure 1. Venn diagram of the distribution of the main four dependencies (cocaine, cannabis, alcohol and opiates) in the Spanish sample of patients used in this study. Other dependencies with a frequency lower than 10% are not displayed.

Supplementary Table S1. Descriptive characteristics of the Caucasian Spanish individuals used in the present study

	Substance dependence sample		Cocaine dependence sample	
	Cases N=1,711	Controls N=1,719	Cases N=1,011	Controls N=1,144
Gender N (%)				
Male	1,299 (75.8)	1,305 (75.8)	826 (81.7)	935 (81.7)
Female	412 (24.2)	414 (24.2)	185 (18.3)	209 (18.3)
Age (mean and SD)				
	40.7 ± 11.7	54.3 ± 16.6	36.6 ± 8.4	54.9 ± 16.7
Substance Dependence N (%)				
Cocaine	1,030 (60.2)	-	1,011 (100.0)	-
Alcohol	849 (49.6)	-	354 (35.0)	-
Cannabis	505 (29.5)	-	310 (30.6)	-
Opiates	322 (18.8)	-	204 (20.2)	-
Tobacco	621 (36.3)	-	271 (26.8)	-
Benzodiacepines	139 (8.1)	-	65 (6.4)	-
Others	49 (2.9)	-	31 (3.1)	-

Supplementary Table S2. SNPs selected for the replication study

SNP	Reference	Nearby Gene	Distance (Kb)	Alleles	MAF	Assoc. with Sample	p-value	Replication study
rs1868152	McGue et al. 2013	LOC253573	8.6 (5')	C>T	0.12	SD	4.90e-08	No
rs2952621	Wetherill et al. 2015	LOC151121	1.3 (3')	C>T	0.47	SD	1.77e-08	Yes (P = 0.02, SAGE sample)
rs2629540	Gelernter et al. 2014	FAM53B	0 (intron 1)	C>G	EA: 0.07 AA: 0.25	CD	4.28e-08	Not replicated (SAGE sample)

MAF: Minor Allele Frequency; Assoc: Association; SD: substance dependence; CD: cocaine dependence; Ca: cases; Ct: controls; EA: European-American; AA: African-American

Supplementary Table S3. Predicted functional effect of the associated variant rs2952621 and variants in linkage disequilibrium with it.

SNP	eQTL Information ^a				SNP Func. Pred. ^b			HaploReg Information ^c						
	Position ^d	Alleles	R ²	D'	Gene	P-Value	Effect Size	Tissue	TFBS	Splicing	H3K4me1	H3K4me3	H3K27ac	H3K9ac
rs35258248	129700092	C/T	0.846	1										
rs1270481	129700134	C/T	0.84	0.964										
rs735987	129700872	A/G	0.967	1										
rs746631	129701108	T/A	0.967	1										
rs1251214	129701584	A/C	0.967	1	AC018865.9	5.7 e-05	0.31	Nerve - Tibial						
rs1251215	129701624	T/C	1	1										
rs1268546	129701652	T/C	0.934	0.966										
rs1251216	129701680	G/C	0.967	1										
rs2406641	129701895	A/G	0.934	0.966										
rs894109	129702012	C/G	0.967	1	AC018865.9	9.7 e-05	0.30	Nerve - Tibial						
rs1251205	129703990	A/T	1	1	AC018865.9	9.1 e-05	0.30	Nerve - Tibial						
rs4662907	129706667	A/G	0.875	1							2			
rs13033217	129707339	C/T	0.875	1							8	1	1	
rs1078826	129707482	A/G	0.818	1							8	1	1	
rs1251195	129708713	C/T	0.967	1							8			
rs1660895	129710802	A/T	0.87	0.965										
rs35878075	129711971	A/G	0.846	1										
rs879533	129713216	T/A	0.967	1							1			
rs1734330	129713458	A/T	0.967	1							4		4	1
rs3856372	129713553	T/A	0.846	1							5		4	1
rs894110	129714013	C/G	0.967	1							3		5	1
rs935085	129714503	C/T	0.934	0.966							3		5	1
rs2952621	129714913	C/T	1	1							3		3	1
rs1251172	129715660	A/G	0.967	1					Yes	Yes	2			
rs4662640	129716266	A/C	0.846	1					Yes	Yes	1			
rs4662641	129716742	A/G	0.846	1					Yes	Yes				
rs10167178	129718195	C/T	0.846	1					Yes	Yes				
rs1441141	129718543	A/G	0.846	1					Yes	Yes				
rs6708498	129718656	C/T	0.84	0.964	AC018865.9	7.1e-05	0.32	Nerve - Tibial	Yes	Yes				
rs1441142	129719748	A/G	0.846	1					Yes	Yes				
rs1251179	129720205	A/G	0.967	1	AC018865.9	6.3e-05	0.31	Nerve - Tibial	Yes	Yes				
rs837827	129720621	A/G	0.967	1					Yes	Yes				
rs7591211	129722233	A/T	0.84	0.964	AC018865.9	7e-05	0.31	Nerve - Tibial			1			
rs6756671	129723645	A/G	0.811	0.963	AC018865.9	5.4e-05	0.32	Nerve - Tibial						
rs2197022	129724220	G/T	0.84	0.964	AC018865.9	5.7e-05	0.31	Nerve - Tibial						
rs1524190	129724516	T/C	0.84	0.964	AC018865.9	5.9e-05	0.31	Nerve - Tibial						
rs13034230	129724734	A/G	0.811	0.963	AC018865.9	6.1e-05	0.31	Nerve - Tibial						
rs13017147	129724741	C/T	0.84	0.964	AC018865.9	8.6e-05	0.32	Nerve - Tibial						

^aNCBI36/h18 assembly

^bGTEx Analysis, Release V6p. <http://www.gtexportal.org/home/>

^cSNP Function Prediction. <https://snpinfo.nih.gov/snpinfo/snpfunc.html>

^dHaploReg 4.1 tool. <http://archive.broadinstitute.org/mammals/haploreg/haploreg.php>. We list the total number of marks considering 10 different brain areas.

SNP in bold: SNP associated with substance dependence

Article 2

“NFAT5 (TonEBP) no regula l’expressió de la *DDC* en línies cel·lulars neuronals dopaminèrgiques”

Resum:

NFAT5, també conegut com a TonEBP, és un factor de transcripció que regula la resposta osmoadaptativa en múltiples teixits i que presenta una elevada expressió en el sistema nerviós central en desenvolupament. Un estudi previ va descriure que l’activació d’NFAT5 a través d’estrès hipertònic incrementava l’expressió de l’enzim dopa descarboxilasa (*DDC*), també conegut com a AADC, en cèl·lules del tub renal proximal, donant lloc a un augment de la síntesi de dopamina. En un estudi anterior vam identificar *NFAT5* com a gen candidat a la dependència de cocaïna, un trastorn psiquiàtric complex en què la neurotransmissió dopaminèrgica té un paper important. Per tant, hem formulat ara la hipòtesi que NFAT5 també podia afectar els nivells de dopamina al sistema nerviós mitjançant la regulació de l’expressió de *DDC*. Per tal de testar la nostra hipòtesi, hem examinat si es produïa aquesta regulació en dues línies cel·lulars neurals dopaminèrgiques, SH-SY5Y i PC12.

L’efecte d’NFAT5 sobre l’expressió de la isoforma neuronal de la *DDC* s’ha avaluat per qRT-PCR. Com a efecte de l’estrès hipertònic, NFAT5 s’activa i s’acumula al nucli i, posteriorment, es produeix un augment de l’expressió d’NFAT5 i dels seus gens diana *SMIT* i *TAUT*, tal i com esperàvem. No obstant, hem observat una disminució de l’expressió de *DDC*. En silenciar l’expressió d’NFAT5 amb un shRNA específic hem observat que la disminució de l’expressió de la *DDC* és independent d’NFAT5 en les dues línies cel·lulars estudiades i que és deguda a l’estrès hipertònic.

En conclusió, NFAT5 no regula l’expressió de la isoforma neuronal de *DDC* en línies cel·lulars neurals dopaminèrgiques i, en conseqüència, no modula la síntesi de dopamina, almenys a través de la *DDC*.

Referència:

Pineda-Cirera L, Cabana-Domínguez J, Benetó N, Díez H, Arenas C, Cormand B, Fernàndez-Castillo N. *DDC expression is not regulated by NFAT5 (TonEBP) in dopaminergic neural cell lines*. Gene 2020 Mar 9;742:144569. doi: 10.1016/j.gene.2020.144569



Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

DDC expression is not regulated by NFAT5 (TonEBP) in dopaminergic neural cell lines



Laura Pineda-Cirera^{a,b,c,d}, Judit Cabana-Domínguez^{a,b,c,d}, Noelia Benetó^{a,b,c,d}, Hector Diez^{e,f}, Concepció Arenas^a, Bru Cormand^{a,b,c,d,*}, Noèlia Fernández-Castillo^{a,b,c,d,*}

^a Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain

^b Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain

^c Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Catalonia, Spain

^d Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Catalonia, Spain

^e Minoryx Therapeutics, Mataró, Catalonia, Spain

^f qGenomics, Esplugues de Llobregat, Catalonia, Spain

ARTICLE INFO

Keywords:

NFAT5

DDC

TonEBP

Hypertonic stress

Neural dopaminergic cell lines

SH-SY5Y

PC12

AADC

ABSTRACT

The nuclear factor of activated T-cells 5 (NFAT5), also known as tonicity-responsive enhancer-binding protein (TonEBP), is a transcription factor that regulates osmoadaptive response in multiple tissues and is highly expressed in the developing central nervous system. A former study reported that NFAT5 activation through hypertonic stress increases the expression of the dopa decarboxylase enzyme (DDC), also known as aromatic-L-amino-acid decarboxylase (AADC), in human renal proximal tubule cells, leading to an increase of dopamine synthesis. In a previous study, we identified *NFAT5* as a candidate gene for cocaine dependence, a complex psychiatric disorder in which dopaminergic neurotransmission plays an important role. Therefore, to test the hypothesis that NFAT5 may also affect dopamine levels in the nervous system through the regulation of *DDC* expression, we examined this regulation using two neural dopaminergic cell lines, SH-SY5Y and PC12.

The effect of NFAT5 on the expression of the neuronal isoform of *DDC* was evaluated by qRT-PCR. Upon hypertonic stress, NFAT5 was activated and accumulated into the nuclei and, subsequently, the expression of *NFAT5* and its known targets sodium/myo-inositol cotransporter 1 (*SMIT*) and sodium chloride/taurine cotransporter (*TAUT*) increased, as expected. However, the expression of *DDC* decreased. When silencing the expression of *NFAT5* with a specific shRNA we observed that the downregulation of *DDC* is independent from NFAT5 in both cell lines and is due to hypertonic stress.

In conclusion, NFAT5 does not regulate the expression of the neuronal isoform of *DDC* in neural dopaminergic cell lines and, consequently, it does not modulate dopamine synthesis through *DDC*.

1. Introduction

NFAT5 (TonEBP) is a transcription factor of the Rel family activated by hypertonicity that is ubiquitously expressed, it is essential for the regulation of homeostasis under osmotic stress, and it has been especially related to immune system and inflammatory functions (Aramburu and López-Rodríguez, 2019; Ho, 2003; Lee et al., 2019). The activation

of NFAT5, and its translocation to the nucleus, activates the expression of many genes, such as the ones for the osmolyte transporters *SMIT* (sodium/myo-inositol cotransporter) and *TAUT* (sodium chloride/taurine cotransporter), and it also activates its own transcription (Aramburu et al., 2006; Halterman et al., 2012; Loyher et al., 2004). It has been reported that an acute hypertonic injection in rats increases the amount of NFAT5 protein in the nuclei of neurons (Maallem et al.,

Abbreviations: *NFAT5*, nuclear factor of activated T-cells 5; TonEBP, tonicity-responsive enhancer-binding protein; *DDC*, dopa decarboxylase; AADC, aromatic-L-amino-acid decarboxylase; *SMIT*, sodium/myo-inositol cotransporter 1; *TAUT*, sodium chloride/taurine cotransporter; *AQP4*, aquaporin 4; *GAPDH*, glyceraldehyde-3-phosphate dehydrogenase; *PGK1*, phosphoglycerate kinase 1; NaCl, sodium chloride; RNase, ribonuclease; cDNA, DNA complementary to RNA; shRNA, short hairpin RNA; PCR, polymerase chain reaction; qRT-PCR, quantitative polymerase chain reaction; DAPI, 4',6-diamidino-2-phenylindole stain; ANOVA, analysis of variance

* Corresponding authors at: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain.

E-mail addresses: bcormand@ub.edu (B. Cormand), noefernandez@ub.edu (N. Fernández-Castillo).

¹ These authors contributed equally to this work.

<https://doi.org/10.1016/j.gene.2020.144569>

Received 17 July 2019; Received in revised form 21 January 2020; Accepted 8 March 2020

Available online 09 March 2020

0378-1119/ © 2020 Elsevier B.V. All rights reserved.

2006a,b). However, NFAT5 is not only regulated by tonicity, and it can also be activated through other stimuli in hypertonic and isotonic tissues (Halterman et al., 2012). NFAT5 is expressed in the adult brain, but it is also highly and specifically expressed in the developing brain at embryonic stages, suggesting a role in embryogenesis and cellular homeostasis (Loyher et al., 2004; Yang et al., 2018). NFAT5 function in the brain varies among cell types. For example it protects neurons against ischemic damage, may participate in inflammation in microglia, regulates the expression of *AQP4* in astrocytes and could influence dopaminergic neurotransmission (Jeong et al., 2016; Mak et al., 2012; Yang et al., 2018; Yi et al., 2013).

In a previous study of our group, the *NFAT5* gene was found upregulated in a dopaminergic neuronal model after cocaine exposure, and it was also found to carry genetic risk variants predisposing to cocaine dependence (Fernández-Castillo et al., 2015). It has been suggested that NFAT5 may participate in modulating dopamine levels in peripheral tissues through the direct regulation of DDC, an enzyme involved in dopamine synthesis (Hsin et al., 2011). The authors observed that the expression of *DDC* was upregulated by NFAT5 (TonEBP) in cells of the human renal proximal tubule (HK-2) and this resulted in increased dopamine levels (Hsin et al., 2011). This connection between NFAT5 and dopamine is interesting, since addiction is a neuropsychiatric disorder in which this neurotransmitter (and also serotonin) plays an important role by mediating the effects of the drug on the reward system (Volkow et al., 2017).

DDC (also known as AADC), the aromatic L-amino acid decarboxylase, catalyses the synthesis of dopamine and serotonin, which act as neurotransmitters and hormones in neural and endocrine tissues. Dopaminergic (and also serotonergic) neurotransmission in the brain plays a key role in several neuropsychiatric disorders, including addiction (Volkow et al., 2017). The *DDC* gene has two tissue-specific isoforms that differ in their 5'UTR, one expressed in neuronal cell types and the other one in non-neuronal tissues (Ichinose et al., 1989).

Since NFAT5 has been reported to be a transcription factor for *DDC* in human renal proximal tubule cells (Hsin et al., 2011), it has been suggested that NFAT5 may also have a role in dopaminergic neurotransmission in brain (Yang et al., 2018). This would substantiate its possible contribution to cocaine addiction.

Therefore, the aim of the present study is to explore whether NFAT5 regulates the expression of the *DDC* neuronal isoform in dopaminergic neural cell lines, modulating in consequence the production of dopamine in the brain.

2. Material and methods

2.1. Cell culture and osmotic shock

PC12 cells from rat adrenal gland (ATCC) provided by Celltec UB were grown in RPMI medium supplemented with 10% horse serum, 5% fetal bovine serum, 1% penicillin/streptomycin and 1% glutamine (Life Technologies), and SH-SY5Y cells from human neuroblastoma (ATCC) were grown in DMEM:F12 (50:50) medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin (Life Technologies), both at 37 °C in a humidified 5% CO₂ atmosphere incubator.

To activate and upregulate NFAT5 we used a hypertonic stress procedure based on previous studies performed in brain (with sucrose) (Maallem et al., 2006a,b) or in neural cell lines (with NaCl) (Bitoun and Tappaz, 2000; Isaacks et al., 1994). Osmotic shock was performed at the same osmolality by adding a NaCl or sucrose solution to the medium up to 550 mOsm/Kg, examined by VAPRO 5520 osmometer (Wescor) to reach a hypertonic condition. Cells were exposed to the hypertonic medium for 6, 24, 36 or 48 h before RNA isolation. A total of five replicates per condition were performed.

2.2. PCR and quantitative real-time PCR

RNA isolation was performed with the High Pure RNA Isolation Kit (Roche) and quantified with Nanodrop (NanoDrop Technologies). RNA was retrotranscribed to cDNA using the High-Capacity cDNA Reverse kit (Life Technologies) and RNase inhibitor (Applied Biosystems).

Since *DDC* has a neuronal and a non-neuronal isoform, we designed specific primers for the two isoforms to assess which ones are expressed in PC12 and SH-SY5Y by PCR. The primers used are indicated in Table S1. The expression of both isoforms was tested in PC12 and SH-SY5Y cells either not exposed to osmotic shock or exposed to 6 h of hypertonic stress with NaCl up to 550 mOsm/Kg. We used cDNA from rat liver and Hep3B cells as a positive control for the non-neuronal isoform. PCR products were resolved by electrophoresis on 2% agarose gels followed by staining with RedSafe (iNtRON Biotechnology).

The expression levels of *NFAT5* and *DDC* were assessed by quantitative PCR (qRT-PCR), as well as those of the *SMIT* and *TAUT* genes, known NFAT5 targets. QRT-PCR experiments were performed using LightCycler® 480 SYBR Green I Master reagent (Roche) with the LightCycler® 480II system (Roche), and data were analyzed with the LightCycler® 480 Software, Version 1.5. For normalization, *GAPDH* and *PGK1* expression was used as a reference, which were stable across conditions. The primers used for this study are listed in Table S1.

2.3. Immunocytochemistry

PC12 or SH-SY5Y cells were stained during 15 min with Wheat germ agglutinin Alexa Fluor® 488 conjugate (Life Technologies) at 1:2000 in cold medium to stain the cellular membrane, fixed in 4% paraformaldehyde for 30 min and permeabilized with 0.1 M glycine and 0.5% Tween solution. After blocking non-specific binding with 10% normal donkey serum (blocking solution, BS) for 1 h, we stained the samples during 1 h at 37 °C with the anti-NFAT5 rabbit polyclonal primary antibody (Abcam) at 1:100 dilution in BS, 45 min at 37 °C with the donkey anti-rabbit IgG-Cy3 secondary antibody (Jackson ImmunoResearch) at 1:400 in BS and finally, 10 min at RT with DAPI (Invitrogen) at 1:10,000 diluted in BS. Images were acquired with a Leica SP2 AOBs Confocal Microscope (Leica Microsystems) with 63×/1.32–0.6 oil immersion objective and they were analyzed using ImageJ (Schindelin et al., 2012).

2.4. NFAT5 silencing, lentiviral production and infection

To silence *NFAT5* expression, we used an shRNA against *NFAT5* (shNFAT5 5'-GGTCAAACGACGAGATTGTA-3') previously used by others (Drews-Elger et al., 2009). The shRNA was inserted into a psi-LVRU6P plasmid that includes the U6 promoter and a puromycin selection cassette (Genecopoeia). A scrambled shRNA was produced in the same plasmid backbone (Genecopoeia). We also used Tet-O-FUW-EGFP plasmid (Addgene) that contains the EGFP cassette under a tetracycline promoter and FUW-M2rtTA (Addgene) to quantify infection efficiency.

Third generation lentivirus particles were produced essentially as previously described (Ding and Kilpatrick, 2013). Briefly, the lentiviral expression vector (shRNAs or EGFP vectors) and three helper plasmids (pMD2.G, pMDLg/pRRE and pRSV-Rev (Addgene)) were cotransfected into HEK293T plated in T175 flasks using calcium phosphate transfection method, considering a 10:3:4:2 plasmid ratio respectively with a total amount of 142 µg. The medium was replaced 24 h after transfection and viral supernatants were collected 24–30 h later and filtered through a 0.45 µm filter. Finally, supernatants were ultracentrifuged and resuspended in 100 µl of DMEM medium.

PC12 cells or SH-SY5Y cells were plated in six-wells plates and infected with 0.8 µl of concentrated lentiviral particles in fresh medium with polybrene (6.5 µg/mL) (Sigma). Viral medium was removed after 24 h and replaced with fresh medium or fresh medium containing

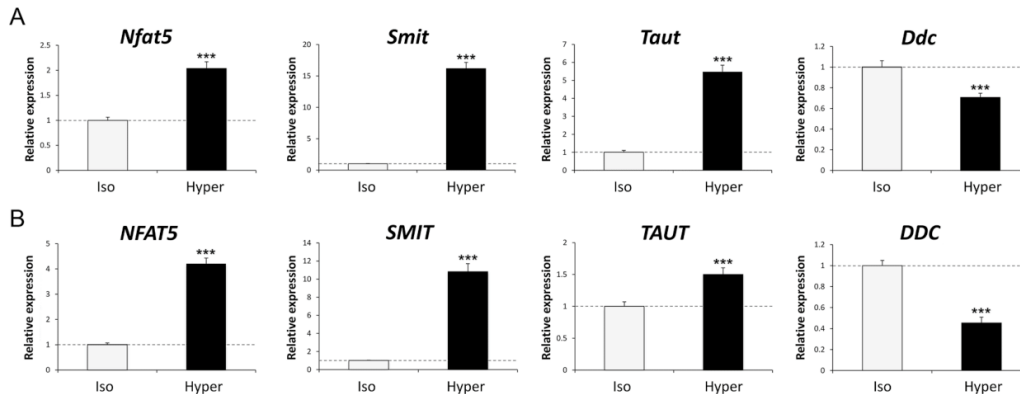


Fig. 1. Gene expression response to hypertonic stress. Cells were exposed to an osmotic shock with salt (NaCl) at 550 mOsm/kg during 6 h. A) PC12 cell line B) SH-SY5Y cell line. Mean and SD is shown. *** $P < 0.001$ when compared to isotonic condition. Expression levels were normalized using *GAPDH* expression. Similar results were obtained when normalizing with *PGK1*.

doxycycline (0.5 $\mu\text{g}/\text{mL}$) (Merck) for genes under tetracycline promoter. Twenty-four hours later, the osmotic shock was performed during 6, 24, 36 or 48 h with NaCl as described above. A total of five replicates per condition were performed.

2.5. Statistical analysis

The data in all figures were presented as mean \pm SD. The analyses of the effect of NFAT5 activation under hypertonic stress on *SMIT*, *TAUT* and *DDC* expression were performed with one-way ANOVA. Gaussian Generalized Linear Models were used to investigate the possible effect of *NFAT5* on the *SMIT*, *TAUT* and *DDC* expression in the experiments silencing *NFAT5*. In all the analyses, when necessary for normality assumption, log transformation was made. All computations were performed in R version 3.4.1 (R Core Team, 2014) and a value of $P < 0.05$ was considered to be statistically significant.

3. Results

The aim of the present study is to explore whether NFAT5, previously shown to modulate dopamine synthesis in renal tubule cells

through activation of the DDC enzyme, also regulates the expression of the *DDC* neuronal isoform in two dopaminergic neural cell lines (PC12 and SH-SY5Y cells), therefore regulating dopamine levels in the brain. To reach our objective, we activated NFAT5 by inducing hypertonic stress with salt or sucrose up to 550 mOsm/Kg during 6, 24, 36 or 48 h, as previously used to activate NFAT5 in neural cell lines (Bitoun and Tappaz, 2000; Isaacks et al., 1994).

First, we inspected which *DDC* isoforms were expressed in PC12 and SH-SY5Y and we determined that both cell lines express only the neuronal isoform in both conditions, with or without hypertonic stress (Fig. S1).

When activation of NFAT5 was induced under hypertonic stress (550 mOsm/kg) with salt (NaCl) we observed an expected significant upregulation of *NFAT5* expression, both in PC12 and in SH-SY5Y cells (Fig. 1). Similar results were observed when hypertonic stress was induced with sucrose at the same osmolality (Fig. S2). Immunocytochemistry demonstrated that the NFAT5 protein was translocated into the nuclei and therefore activated upon hypertonic stress with salt in both cell lines (Fig. 2). The genes *SMIT* and *TAUT*, known NFAT5 targets, were significantly upregulated under hypertonic stress with salt as expected, and the expression of *NFAT5* also increased

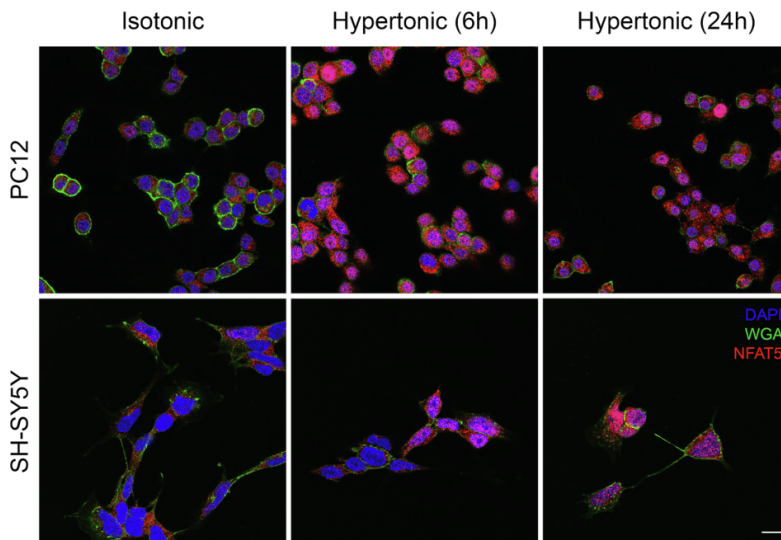


Fig. 2. Immunocytochemistry images showing the translocation of NFAT5 (in red) from the cytoplasm to the nucleus of the cells after a hypertonic stress. Cells were exposed to an osmotic shock with salt (NaCl) at 550 mOsm/kg during 6 h or 24 h. NFAT5 (TonEBP) was visualized using an antibody targeting its C-terminal segment (red), DAPI (blue) was used to stain the nuclei and WGA (green) was used to stain the membrane. Scale bar: 10 μm . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

significantly, which confirms the actual activation of the transcription factor NFAT5 in both cell lines (Fig. 1). However, the expression of *DDC* decreased significantly in PC12 and in SH-SY5Y (Fig. 1), contrarily to the effect reported in renal proximal tubule cells (Hsin et al., 2011). Similar results were observed when both cell lines were exposed to hypertonic stress with sucrose (Fig. S2). So, for subsequent experiments we only induced hypertonic stress with salt, more used in neural cell culture studies (Bitoun and Tappaz, 2000; Isaacks et al., 1994).

In order to confirm that NFAT5 does not regulate the expression of *DDC* in these cell lines, we silenced *NFAT5* using a specific shRNA that inhibited its expression about 50% (Figs. 3 and 4). Hypertonic stress

increased the expression (through NFAT5) of the targets of *NFAT5*-activated transcriptional activity *SMIT* and *TAUT*. When *NFAT5* was silenced, the transcription levels of both genes decreased compared to the control in the two cell lines, indicating an actual downregulation of *NFAT5* activity by the shRNA (Figs. 3 and 4). In contrast, the combination of *NFAT5* inhibition and hypertonic stress resulted in a reduction of *DDC* expression similar to that observed when *NFAT5* is not silenced (using a control shRNA) (Figs. 3 and 4). We observed a 70–80% decrease of *DDC* expression upon 24 h of hypertonic shock in PC12 (Fig. 3) and 60–80% upon 6 h of hypertonic shock in SH-SY5Y (Fig. 4). These results suggest that, in contrast to *SMIT* and *TAUT*, the expression

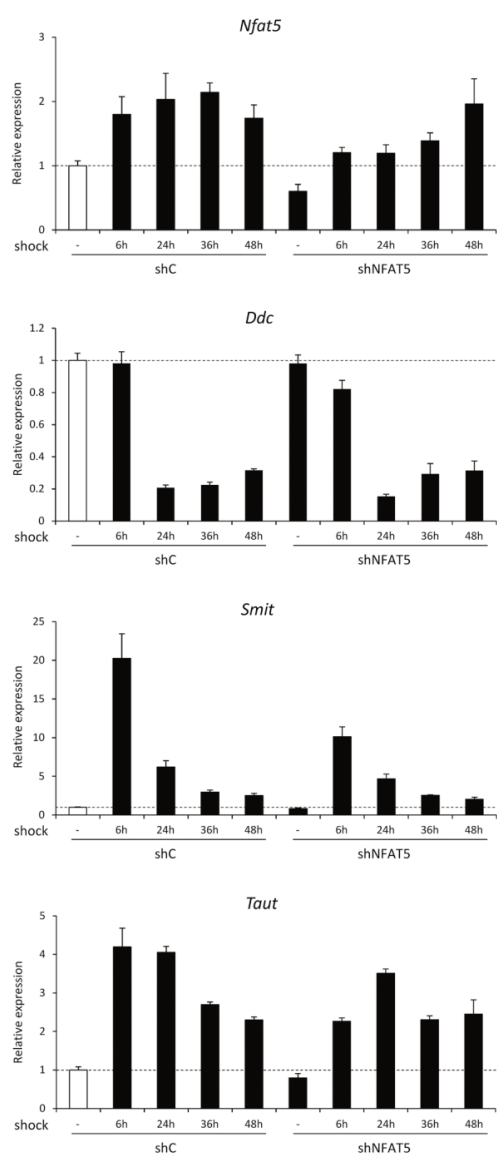


Fig. 3. Gene expression response to hypertonic stress in PC12 cells. Cells were exposed to an osmotic shock at 550 mOsm/kg during a 6, 24, 36 or 48 h treatment with salt (NaCl), with or without *NFAT5* silencing. Two different shRNA were used, one against *Nfat5* (shNFAT5) and a scrambled shRNA as a control (shC). Mean and SD is shown. Expression levels were normalized using *Gapdh* expression.

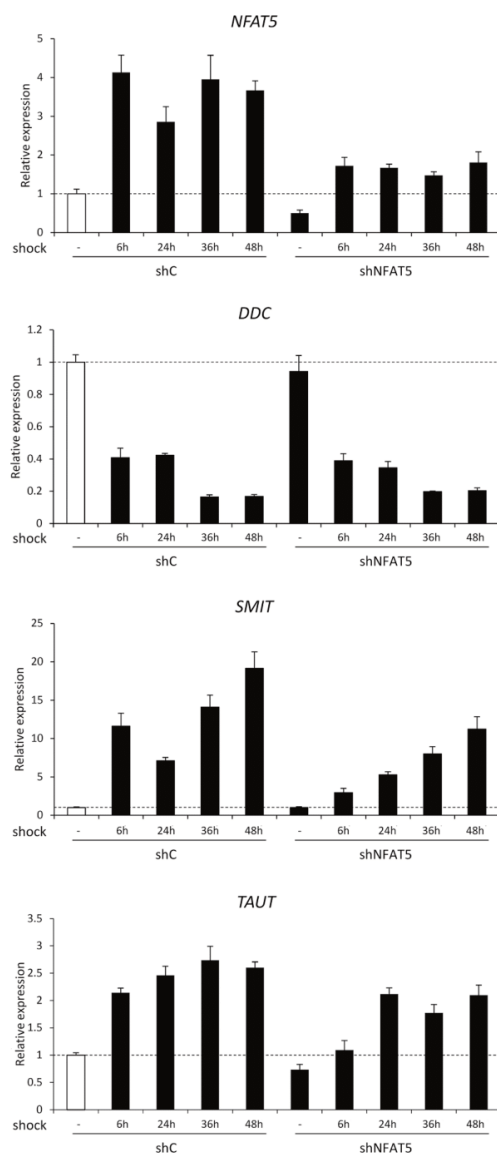


Fig. 4. Gene expression response to hypertonic stress in SH-SY5Y cells. Cells were exposed to an osmotic shock at 550 mOsm/kg during a 6, 24, 36 or 48 h treatment with salt (NaCl), with or without *NFAT5* silencing. Two different shRNA were used, one against *NFAT5* (shNFAT5) and a scrambled shRNA as a control (shC). Mean and SD is shown. Expression levels were normalized using *GAPDH* expression.

of the neuronal isoform of *DDC* is regulated through a mechanism that is independent from NFAT5 in PC12 and in SH-SY5Y cells.

Then, we assessed the effect of NFAT5 on the expression of these three genes using a generalized linear model. As reported previously (Aramburu et al., 2006; Halterman et al., 2012; Loyher et al., 2004), NFAT5 showed a significant effect on *SMIT* and *TAUT* expression in PC12 ($P = 2.23E-04$ and $P = 6.05E-04$, respectively) and SH-SY5Y ($P = 1.34E-03$ and $P = 4.62E-02$, respectively), both increasing their expression when NFAT5 is upregulated, reinforcing their value as targets of NFAT5-regulated transcriptional activity (Figs. 3 and 4 and Table S2). On the other hand, NFAT5 had no significant effect on the expression of *DDC* in PC12 ($P = 0.13$) nor in SH-SY5Y ($P = 0.48$) (Figs. 3 and 4 and Table S2). Furthermore, we observed that the expression of *DDC* decreased significantly by the effect of the osmotic shock upon 24 h in PC12 ($P < 2E-16$), and at all time points in SH-SY5Y ($2E-16 < P < 1.27E-08$) (Figs. 3 and 4 and Table S2).

4. Discussion

The present study suggests that NFAT5 is not involved in dopamine neurotransmission in the nervous system (a hypothesis that had previously been formulated (Fernández-Castillo et al., 2015; Yang et al., 2018)), at least through the regulation of the expression of *DDC*. Our work in dopaminergic neural-like cells shows NFAT5 does not regulate the neuronal isoform of *DDC*, contrarily to what had been reported in renal tubular cells, where hypertonicity activates NFAT5, which in turn upregulates the non-neuronal isoform of *DDC*, increasing the levels of dopamine (Hsin et al., 2011). Our group previously described that NFAT5 was upregulated by cocaine and that genetic risk variants in this gene were associated with cocaine dependence, a complex psychiatric disorder in which dopaminergic neurotransmission plays an important role (Volkow et al., 2017). We therefore hypothesized that NFAT5 could also regulate *DDC* in the brain, activating its expression and leading to an increase of dopamine, a neurotransmitter involved in reward and motivation, important for addiction, but also related to other neuropsychiatric disorders. However, rather than observing an upregulation of *DDC*, we detected a decrease of *DDC* expression after the activation of NFAT5 through hypertonic stress in dopaminergic neural-like cell lines. Subsequently, we demonstrated that this decrease in the expression of the neuronal isoform of *DDC* was due to hypertonic stress and independent from NFAT5 regulation.

Our results indicate that NFAT5 was expressed and translocated into the cell nuclei, where it becomes an active transcription factor, as shown by immunocytochemistry and by the increased expression of two known target genes and itself (Figs. 1 and 2), consistent with previous studies (Aramburu et al., 2006; Halterman et al., 2012; Loyher et al., 2004). But NFAT5 does not seem to regulate *DDC* expression in dopaminergic neural cell lines, in contrast to the results reported by Hsin et al. in renal proximal tubule cells (Hsin et al., 2011). Thus, NFAT5 regulates the expression of the non-neuronal isoform of *DDC* but not that of the neuronal isoform. However, we cannot discard that NFAT5 may play a role in the regulation of dopamine levels in the brain not through *DDC*. We also show that the neuronal isoform of *DDC* is downregulated by osmotic shock in the two investigated neural cell lines. This observation goes in the opposite direction to what is seen in renal tubule cells, where dopamine has a role in regulating natriuresis. One possible explanation may be that the neuronal isoform of *DDC* is not involved in the osmoadaptive response mediated by dopamine and *DDC* would be downregulated as part of the transcription repression generated by a cell adaptation process (Burg et al., 2007) in neural cell lines.

There are some strengths and limitations in our study that should be discussed. Strengths: (i) Comparable results have been obtained for the two dopaminergic neural cell lines studied (PC12 and SH-SY5Y), (ii) expression changes are maintained up to 48 h, (iii) similar expression changes were obtained when hypertonic stress was induced with

sucrose or with NaCl at the same osmolality, (iv) we confirmed that both cell lines (PC12 and SH-SY5Y) express only the neuronal isoform of the *DDC*. Limitations: (i) The mechanism by which hypertonic stress modulates the expression of the neuronal isoform of *DDC* in PC12 and in SH-SY5Y cells has not been studied by us, (ii) NFAT5 may be involved in the regulation of dopamine levels through another mechanism not considered in the present study and (iii) what we see in these two neural cell lines may differ from actual mechanisms in the brain.

To sum up, we observed that the expression of the neuronal isoform of *DDC* is independent from NFAT5 regulation in two dopaminergic neural cell lines (PC12 and SH-SY5Y) upon hypertonic stress.

CRedit authorship contribution statement

Laura Pineda-Cirera: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Judit Cabana-Domínguez:** Investigation, Writing - review & editing. **Noelia Benetó:** Methodology, Investigation, Writing - review & editing. **Héctor Díez:** Methodology, Writing - review & editing. **Concepció Arenas:** Formal analysis, Funding acquisition, Writing - review & editing. **Bru Cormand:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Noelia Fernández-Castillo:** Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The funder Minoryx Therapeutics provided support in the form of salaries for author HD, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. All other authors declare no competing interests.

Acknowledgements

We are thankful to Jose Francisco Aramburu (Universitat Pompeu Fabra, Barcelona) for helpful advice on several experiments and also for the use of the VAPRO 5520 osmometer of his laboratory. We are also grateful to Isaac Canals for help with the lentivirus production and for kindly providing the plasmids needed.

This work was supported mainly by the Spanish 'Ministerio de Economía y Competitividad' [grant number SAF2015-68341-R and RTI2018-100968-B-100] and AGAUR, 'Generalitat de Catalunya' [grant number 2017-SGR-738]. BC also received funding from the European Union H2020 Program [H2020/2014-2020, grant agreement 667302 and 728018]. CA has received funding from and AGAUR, 'Generalitat de Catalunya' [grant number 2017-SGR-622]. LP-C and JC-D were supported by 'Generalitat de Catalunya' [grant numbers 2016 FI_B 00728 and 2015 FI_B 00448, respectively]. LP-C was also supported by 'Ministerio de Educación, Cultura y Deporte' [grant number FPU15/03867]. NF-C is supported by a contract of the 'Centro de Investigación Biomédica en Red de Enfermedades Raras' (CIBERER). The funder Minoryx Therapeutics provided support in the form of salaries for author HD.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.gene.2020.144569>.

References

- Aramburu, J., Drews-Elger, K., Estrada-Gelonch, A., Minguillón, J., Morancho, B., Santiago, V., López-Rodríguez, C., 2006. Regulation of the hypertonic stress response and other cellular functions by the Rel-like transcription factor NFAT5. *Biochem. Pharmacol.* 72, 1597–1604.
- Aramburu, J., López-Rodríguez, C., 2019. Regulation of inflammatory functions of macrophages and T lymphocytes by NFAT5. *Front. Immunol.* 10, 535.
- Bitoun, M., Tappaz, M., 2000. Gene expression of the transporters and biosynthetic enzymes of the osmolytes in astrocyte primary cultures exposed to hyperosmotic conditions. *Glia* 32, 165–176.
- Burg, M.B., Ferraris, J.D., Dmitrieva, N.I., 2007. Cellular response to hyperosmotic stresses. *Physiol. Rev.* 87, 1441–1474.
- Ding, B., Kilpatrick, D.L., 2013. Lentiviral vector production, titration, and transduction of primary neurons. In: Zhou, R., Mei, L. (Eds.), *Neural Development. Methods in Molecular Biology (Methods and Protocols)*. Humana Press, Totowa, NJ, pp. 119–131.
- Drews-Elger, K., Ortells, M.C., Rao, A., López-Rodríguez, C., Aramburu, J., 2009. The transcription factor NFAT5 is required for cyclin expression and cell cycle progression in cells exposed to hypertonic stress. *PLoS One* 4, e5245.
- Fernández-Castillo, N., Cabana-Domínguez, J., Soriano, J., Sánchez-Mora, C., Roncero, C., Grau-López, L., Ros-Cucurull, E., Daigre, C., van Donkelaar, M.M.J., Franke, B., Casas, M., Ribasés, M., Cormand, B., 2015. Transcriptomic and genetic studies identify NFAT5 as a candidate gene for cocaine dependence. *Transl. Psychiatry* 5, e667.
- Halterman, J.A., Kwon, H.M., Wamhoff, B.R., 2012. Tonicity-independent regulation of the osmosensitive transcription factor TonEBP (NFAT5). *Am. J. Physiol. Cell Physiol.* 302, C1–C8.
- Ho, S.N., 2003. The role of NFAT5/TonEBP in establishing an optimal intracellular environment. *Arch. Biochem. Biophys.* 413, 151–157.
- Hsin, Y.H., Tang, C.H., Lai, H.T., Lee, T.H., 2011. The role of TonEBP in regulation of AAD expression and dopamine production in renal proximal tubule cells upon hypertonic challenge. *Biochem. Biophys. Res. Commun.* 414, 598–603.
- Ichinose, H., Kurosawa, Y., Titani, K., Fujita, K., Nagatsu, T., 1989. Isolation and characterization of a cDNA clone encoding human aromatic L-amino acid decarboxylase. *Biochem. Biophys. Res. Commun.* 164, 1024–1030.
- Isaacs, R.E., Bender, A.S., Kim, C.Y., Prieto, N.M., Norenberg, M.D., 1994. Osmotic regulation of myo-inositol uptake in primary astrocyte cultures. *Neurochem. Res.* 19, 331–338.
- Jeong, G.R., Im, S.-K., Bae, Y.-H., Park, E.S., Jin, B.K., Kwon, H.M., Lee, B.-J., Bu, Y., Hur, E.-M., Lee, B.D., 2016. Inflammatory signals induce the expression of tonicity-responsive enhancer binding protein (TonEBP) in microglia. *J. Neuroimmunol.* 295, 21–29.
- Lee, N., Kim, D., Kim, W.-U., 2019. Role of NFAT5 in the Immune System and Pathogenesis of Autoimmune Diseases. *Front. Immunol.* 10, 270.
- Loyer, M.L., Mutin, M., Woo, S.K., Kwon, H.M., Tappaz, M.L., 2004. Transcription factor tonicity-responsive enhancer-binding protein (TonEBP) which transactivates osmoprotective genes is expressed and upregulated following acute systemic hypertonicity in neurons in brain. *Neuroscience* 124, 89–104.
- Maallem, S., Berod, A., Mutin, M., Kwon, H.M., Tappaz, M.L., 2006a. Large discrepancies in cellular distribution of the tonicity-induced expression of osmoprotective genes and their regulatory transcription factor TonEBP in rat brain. *Neuroscience* 142, 355–368.
- Maallem, S., Mutin, M., Kwon, H.M., Tappaz, M.L., 2006b. Differential cellular distribution of tonicity-induced expression of transcription factor TonEBP in the rat brain following prolonged systemic hypertonicity. *Neuroscience* 137, 51–71.
- Mak, K.M.C., Lo, A.C.Y., Lam, A.K.M., Yeung, P.K.K., Ko, B.C.B., Chung, S.S.M., Chung, S.K., 2012. Nuclear Factor of Activated T Cells 5 Deficiency Increases the Severity of Neuronal Cell Death in Ischemic Injury. *Neurosignals* 20, 237–251.
- R Core Team, 2014. *R: A Language and Environment for Statistical Computing*.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682.
- Volkow, N.D., Wise, R.A., Baler, R., 2017. The dopamine motive system: Implications for drug and food addiction. *Nat. Rev. Neurosci.* 18, 741–752.
- Yang, X.-L., Wang, X., Peng, B.-W., 2018. NFAT5 has a job in the brain. *Rev. Dev. Neurosci.* 40, 289–300.
- Yi, M.-H., Lee, Y.S., Kang, J.W., Kim, S.J., Oh, S.-H., Kim, Y.M., Lee, Y.H., Do Lee, S., Kim, D.W., 2013. NFAT5-dependent expression of AQP4 in astrocytes. *Cell. Mol. Neurobiol.* 33, 223–232.

SUPPLEMENTARY MATERIAL

Table S1. Primers used for the *DDC* isoform analysis and for the real time quantitative PCR.

Human		
Gene	Forward	Reverse
Non-neuronal <i>DDC</i>	AAGGAATTCGAATTTCCAGCAT	GCGGATCCTCAATGCCTTCCATGTAGTT
Neuronal <i>DDC</i>	TTCTGTGCCTCTTAACTGTCACTG	GCGGATCCTCAATGCCTTCCATGTAGTT
<i>NFAT5</i>	TATTTGATGCCGACAGTGCC	GCTCCTTTCCTCACTTTTAAC
<i>DDC</i>	ACCACAACATGCTGCTCCTTTG	CATTCAGAAGGTGCCGGAACTC
<i>SMIT</i>	CACTCGCCGATCCTCCAG	ACTCTCCACAAGACCATCAGC
<i>TAUT</i>	CACCCAGGCTCTCTGAAATG	GACCAGCAGCACCAGGAG
<i>GAPDH</i>	AGCCACATCGCTCAGACAC	GCCCAATACGACCAAATCC
<i>PGK1</i>	ACCGAATCACCGACCTCTCT	TCTCTCCAGCCACTAAGCCA
Rat		
Non-neuronal <i>Ddc</i>	TTGCAGAGCTGGACTGAGTG	ACACTGGACGTCCCTCAATG
Neuronal <i>Ddc</i>	GCAAGAGAGCGAATAGAGAGGA	ACACTGGACGTCCCTCAATG
<i>Nfat5</i>	CGAGTAAAGCCACACGGATTC	ATTCTGGCCTCAACATCAGC
<i>Ddc</i>	GTTGTCACCCTAGGAACCACATC	CGCCATTCAGAAGATACCGGAAC
<i>Smit</i>	GCTCATGCCAAAAGTTCTACTC	CTGCTTCCACACACTTGCATG
<i>Taut</i>	GCCTACCCAAAAGCTGTGACC	CCCGACGATAACCCTTCCTTAG
<i>Gapdh</i>	ATGGTGAAGGTCGGTGTGA	CTTGCCGTGGGTAGAGTCAT
<i>Pgk1</i>	GCAAAGACTGGCCAAGCTAC	GCCTCAGCATATTTCTTACTGCT

Table S2. Generalized linear model *P* values and effects for the different factors that play a role in gene expression of *SMIT*, *TAUT* and *DDC*.

PC12				SH-SY5Y			
	<i>Smit</i>				<i>SMIT</i>		
	Effect	<i>P</i> value	Sig.		Effect	<i>P</i> value	Sig.
Intercept	-0.259	6.70E-04	***	Intercept	-0.194	4.97E-03	**
<i>Nfat5</i> expression	0.258	2.23E-04	***	<i>NFAT5</i> expression	0.188	1.34E-03	**
Shock 6h	2.792	< 2E-16	***	Shock 6h	1.866	1.28E-13	***
Shock 24h	1.552	< 2E-16	***	Shock 24h	1.620	< 2E-16	***
Shock 36h	0.785	1.76E-11	***	Shock 36h	2.095	3.78E-16	***
Shock 48h	0.732	5.95E-14	***	Shock 48h	2.454	< 2E-16	***
shNFAT5	-0.090	7.47E-02		shNFAT5	0.106	6.97E-02	
Shock 6h and shNFAT5	-0.446	2.51E-07	***	Shock 6h and shNFAT5	-1.041	6.46E-10	***
Shock 24h and shNFAT5	0.025	7.45E-01		Shock 24h and shNFAT5	-0.183	5.66E-02	
Shock 36h and shNFAT5	0.133	8.92E-02		Shock 36h and shNFAT5	-0.207	1.41E-01	
Shock 48h and shNFAT5	-0.195	2.40E-02	*	Shock 48h and shNFAT5	-0.294	1.29E-02	
	<i>Taut</i>				<i>TAUT</i>		
	Effect	<i>P</i> value	Sig.		Effect	<i>P</i> value	Sig.
Intercept	-0.207	1.43E-03	**	Intercept	-0.099	7.73E-02	
<i>Nfat5</i> expression	0.204	6.05E-04	***	<i>NFAT5</i> expression	0.094	4.62E-02	*
Shock 6h	1.269	< 2E-16	***	Shock 6h	0.472	3.24E-03	**
Shock 24h	1.192	< 2E-16	***	Shock 24h	0.729	2.17E-09	***
Shock 36h	0.762	3.57E-13	***	Shock 36h	0.731	6.75E-06	***
Shock 48h	0.684	3.23E-15	***	Shock 48h	0.711	2.34E-06	***
shNFAT5	-0.149	9.80E-04	***	shNFAT5	-0.266	1.08E-06	***
Shock 6h and shNFAT5	-0.340	2.65E-06	***	Shock 6h and shNFAT5	-0.192	9.57E-02	
Shock 24h and shNFAT5	0.176	1.24E-02	*	Shock 24h and shNFAT5	0.229	5.09E-03	**
Shock 36h and shNFAT5	0.145	3.24E-02	*	Shock 36h and shNFAT5	0.066	5.70E-01	
Shock 48h and shNFAT5	0.160	3.03E-02	*	Shock 48h and shNFAT5	0.223	2.30E-02	*
	<i>Ddc</i>				<i>DDC</i>		
	Effect	<i>P</i> value	Sig.		Effect	<i>P</i> value	Sig.
Intercept	0.952	< 2E-16	***	Intercept	-0.029	5.88E-01	
<i>Nfat5</i> expression	0.048	1.34E-01		<i>NFAT5</i> expression	0.032	4.76E-01	
Shock 6h	-0.057	1.17E-01		Shock 6h	-1.006	1.27E-08	***
Shock 24h	-0.842	< 2E-16	***	Shock 24h	-0.921	8.90E-13	***
Shock 36h	-0.829	< 2E-16	***	Shock 36h	-1.903	< 2E-16	***
Shock 48h	-0.720	< 2E-16	***	Shock 48h	-1.871	< 2E-16	***
shNFAT5	0.000	9.95E-01		shNFAT5	-0.051	2.70E-01	
Shock 6h and shNFAT5	-0.131	7.04E-04	***	Shock 6h and shNFAT5	0.084	4.46E-01	
Shock 24h and shNFAT5	-0.013	7.29E-01		Shock 24h and shNFAT5	-0.119	1.21E-01	
Shock 36h and shNFAT5	0.105	7.15E-03	**	Shock 36h and shNFAT5	0.316	6.67E-03	**
Shock 48h and shNFAT5	-0.011	7.94E-01		Shock 48h and shNFAT5	0.302	1.83E-03	**

Sig: Significancy; * *P* < 0.05; ** *P* < 0.01; *** *P* < 0.001.

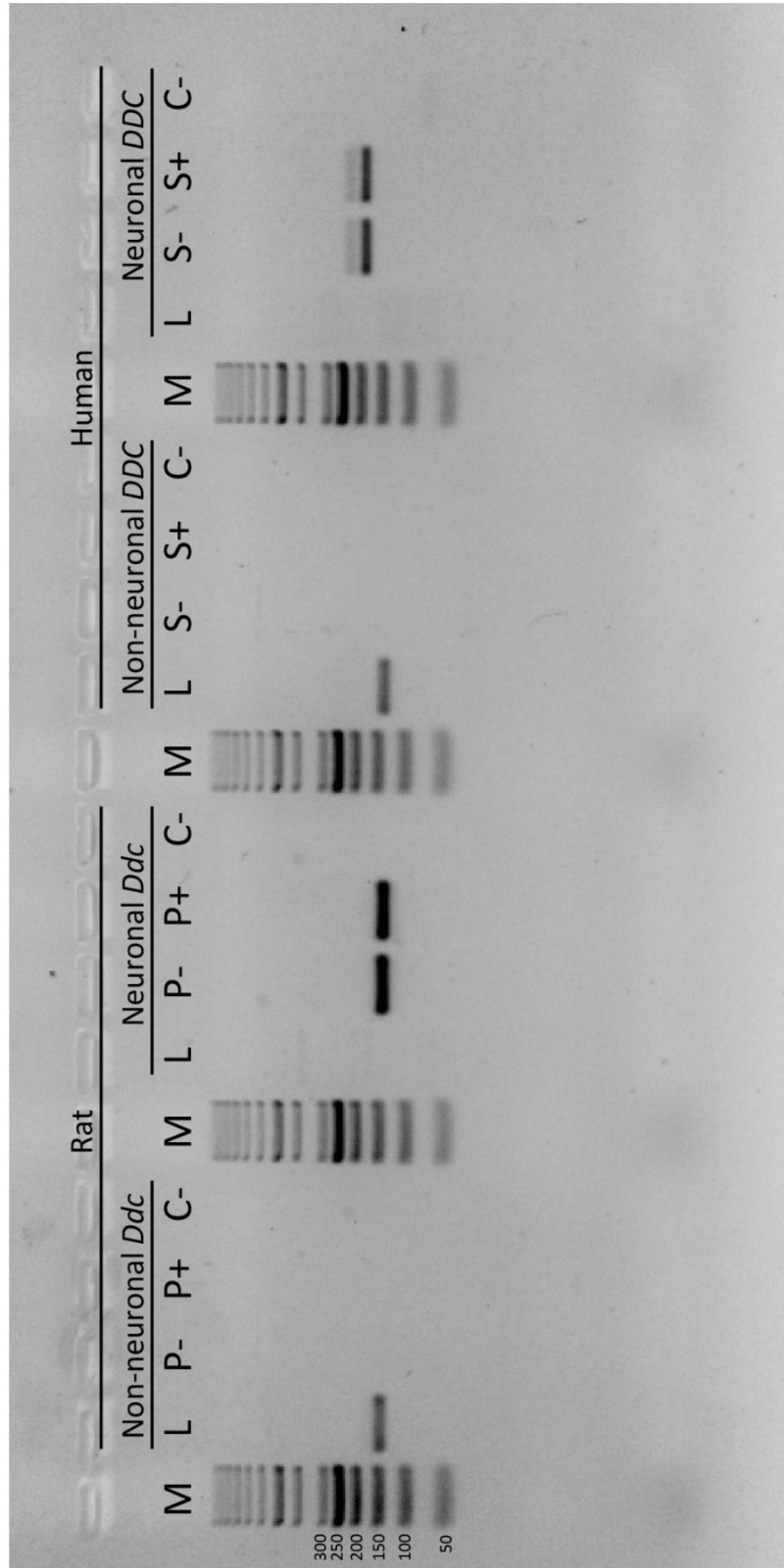


Figure S1. Analysis of the DDC isoforms expressed in rat and human. M: 50 bp marker; L: liver; P-: PC12 cells not exposed to osmotic shock; P+: PC12 cells exposed to 6h osmotic shock at 550mOsm/kg; S-: SH-SY5Y cells not exposed to osmotic shock; S+: SH-SY5Y cells exposed to 6h osmotic shock at 550mOsm/Kg; C-: negative control.

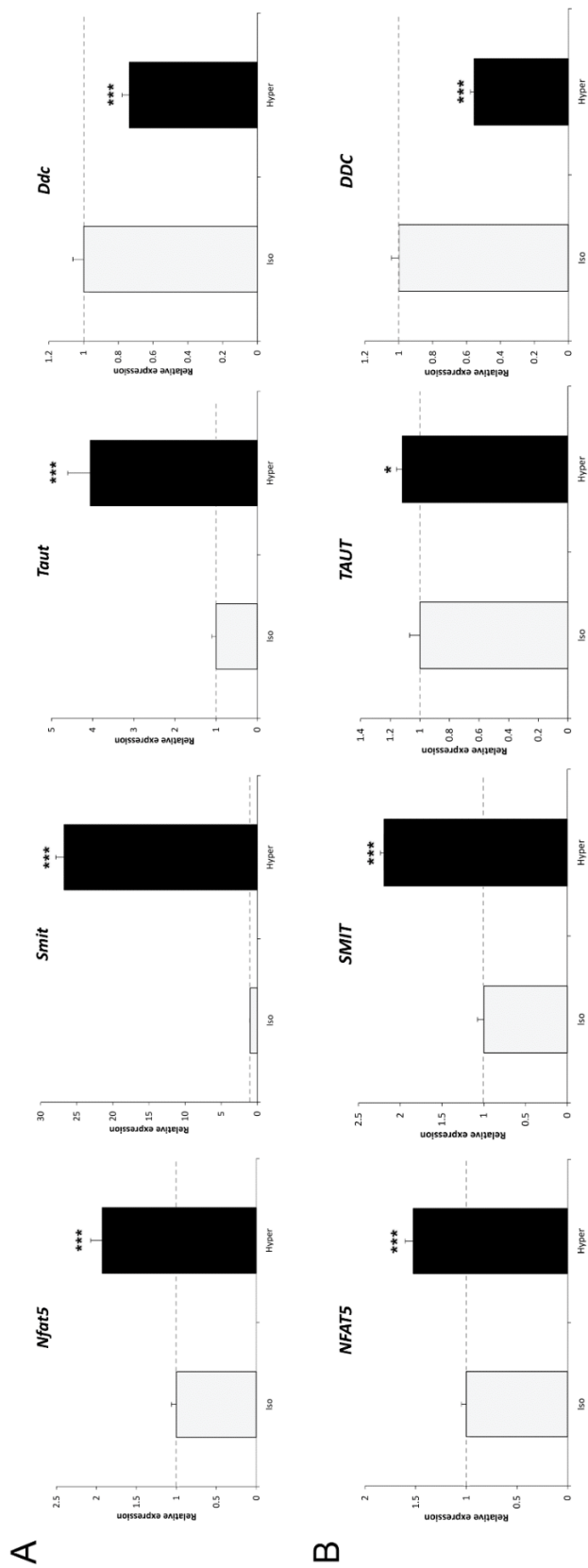


Figure S2. Gene expression response to hypertonic stress. Cells were exposed to an osmotic shock at 550mOsm/kg during a 6 hours treatment with sucrose compared to non-treated controls. A) PC12 cell line B) SH-SY5Y cell line. Iso: cells exposed to an isotonic medium; Hyper: cells exposed to 6h osmotic shock at 550mOsm/Kg. Mean and SD is shown * $P < 0.05$, *** $P < 0.001$. Expression levels were normalized using *GAPDH* expression. Similar results were obtained when normalizing with *PGK1*.

CAPÍTOL 2. Identificació de factors epigenètics de susceptibilitat a l'addicció a les drogues i al menjar.

Article 3

“Estudi de la metilació específica d'al·lel en la susceptibilitat a la dependència de drogues”

Resum:

La dependència de drogues és una malaltia neuropsiquiàtrica causada per factors genètics, epigenètics i ambientals. La metilació específica d'al·lel (ASM) és un mecanisme epigenètic comú i estable pel qual els al·lells d'una variant genètica presenten correlació amb la variació dels nivells de metilació d'un lloc CpG. En aquest estudi hem seleccionat 182 polimorfismes d'un sol nucleòtid (SNPs) que sabem que tenen influència sobre la metilació en *cis* en diferents regions del cervell humà per avaluar la seva possible contribució a la dependència de drogues. Hem dut a terme un estudi d'associació cas-control en una mostra de descoberta de 578 pacients dependents de drogues (incloent 428 dependents de cocaïna) i 656 controls espanyols i a continuació, hem avaluat aquelles variants que mostraven una associació significativa amb el fenotip en una mostra independent de 1.119 casos (incloent 589 dependents de cocaïna) i 1.092 controls també espanyols. A la mostra de descoberta hem identificat cinc associacions nominals de les quals una s'ha replicat a la mostra de rèplica (rs6020251). L'anàlisi conjunta de les dues mostres ha revelat una associació entre la dependència de drogues i la variant rs6020251, però també amb la variant rs11585570. Hem fet la mateixa anàlisi tenint en compte només pacients dependents de cocaïna i hem obtingut resultats similars. La variant rs6020251 presenta correlació amb els nivells de metilació diferencials del lloc CpG cg17974185 i està situada al primer intró del gen *CTNNB1*, en una regió genòmica amb múltiples marques d'histones relacionades amb regions potenciadores i promotores en cervell. La variant rs11585570 és un eQTL en el cervell i en sang pels gens *SCP2* i *ECHDC2*, i presenta correlació amb els nivells de metilació dels llocs CpG cg27535305 i cg13461509, situats a la regió promotora dels dos gens. En resum, gràcies a l'aproximació utilitzada, que combina dades genètiques i epigenètiques, hem pogut identificar els gens *CTNNB1*, *SCP2* i *ECHDC2* com a possibles factors de risc a la dependència de drogues.

Referència: Pineda-Cirera L, Cabana-Domínguez J, Grau-López L, Sánchez-Mora C, Daigre C, Palma-Álvarez RF, Ramos-Quiroga JA, Ribasés M, Cormand B, Fernàndez-Castillo N. *Exploring allele specific methylation in drug dependence susceptibility*. *Journal of Psychiatric Research* (sotmès, en segona revisió).

Exploring allele specific methylation in drug dependence susceptibility

Laura Pineda-Cirera^{a,b,c,d}, Judit Cabana-Domínguez^{a,b,c,d}, Lara Grau-López^{e,f,g,h}, Constanza Daire^{e,f,g,h}, Cristina Sánchez-Mora^{a,e,g,i}, Raul Felipe Palma-Álvarez^{e,f,g,h}, Josep Antoni Ramos-Quiroga^{e,f,g,h}, Marta Ribasés^{a,e,g,i}, Bru Cormand^{a,b,c,d} & Noèlia Fernández-Castillo^{a,b,c,d} &

^aDepartament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain

^bInstitut de Biomedicina de la Universitat de Barcelona (IBUB), Catalonia, Spain

^cCentro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Spain

^dInstitut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Catalonia, Spain

^eDepartment of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Catalonia, Spain

^fGroup of Psychiatry, Mental Health and Addictions, Catalonia, Spain

^gBiomedical Network Research Centre on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain

^hDepartment of Psychiatry and Legal Medicine, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain

ⁱPsychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Spain

[&]Senior authors that equally contributed

CORRESPONDING AUTHORS:

Noèlia Fernández Castillo: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain, Avinguda Diagonal 643, edifici Prevosti, 1^a planta, sala 134. 08028, Barcelona, Catalonia, Spain. Tel.: (+34) 93 4037082. Fax: (+34) 93 4034420; Email:

noefernandez@ub.edu

Bru Cormand: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain, Avinguda Diagonal 643, edifici Prevosti, 3^a planta, 08028, Barcelona, Catalonia, Spain. Tel.: (+34) 93 4021013. Fax: (+34) 93 4034420; Email: bcormand@ub.edu

ABSTRACT

Drug dependence is a neuropsychiatric condition that involves genetic, epigenetic and environmental factors. Allele-specific methylation (ASM) is a common and stable epigenetic mechanism that involves genetic variants correlating with differential levels of methylation at CpG sites. We selected 182 single-nucleotide polymorphisms (SNPs) described to influence *cis* ASM in human brain regions to evaluate their possible contribution to drug dependence susceptibility. We performed a case-control association study in a discovery sample of 578 drug-dependent patients (including 428 cocaine-dependent subjects) and 656 controls from Spain, and then, we followed-up the significant associations in an independent sample of 1,119 cases (including 589 cocaine-dependent subjects) and 1,092 controls. In the discovery sample, we identified five nominal associations, one of them replicated in the follow-up sample (*rs6020251*). The pooled analysis revealed an association between drug dependence and *rs6020251* but also *rs11585570*.

We performed the same analysis considering only cocaine-dependent patients and obtained similar results. The rs6020251 variant correlates with differential methylation levels of cg17974185 and lies in the first intron of the *CTNNB1* gene, in a genomic region with multiple histone marks related to enhancer and promoter regions in brain. Rs11585570 is an eQTL in brain and blood for the *SCP2* and *ECHDC2* genes and correlates with differential methylation of cg27535305 and cg13461509, located in the promoter regions of both genes. To conclude, using an approach that combines genetic and epigenetic data, we highlighted the *CTNNB1*, *SCP2* and *ECHDC2* genes as potential contributors to drug dependence susceptibility.

KEYWORDS: *CTNNB1*, *SCP2*, *ECHDC2*, brain DNA methylation, SNP, association

INTRODUCTION

Drug use is a major health problem worldwide, with about 35 million people, or 0.7% of the global adult population being affected by illegal substance use disorders (UNODC, 2019) and 5.1% affected by alcohol use disorders (World Health Organization, 2018). Drug dependence is a complex psychiatric disorder in which genetic and environmental risk factors are involved. Heritability estimates of addiction range from 39 to 72%, depending on the drug of abuse (Ducci and Goldman, 2012). Although multiple studies have been performed to understand the genetic basis of drug dependence (reviewed by Bühler et al., 2015; Edenberg et al., 2019; Jensen, 2016; Mulligan, 2019), the underlying molecular mechanisms are still largely unknown. It has been widely suggested that epigenetics, and particularly DNA methylation, may have a role in addiction (Cadet et al., 2016; Hamilton and Nestler, 2019), as the interplay between genetic and environmental factors. Furthermore, DNA methylation is implicated in the regulation of multiple brain events like neurogenesis, differentiation and brain development (Ladd-Acosta et al., 2007) in addition to synaptic plasticity, memory formation and acquired behavior (Bashtrykov and Jeltsch, 2017), many of them involved in the transition from drug use to dependence.

DNA methylation is a common and stable epigenetic mechanism catalyzed by DNA methyltransferases (DNMT) that add a methyl group usually to the cytosine of a CpG dinucleotide (Klose and Bird, 2006). DNMT expression alterations have been linked to drug dependence and to other psychiatric disorders like schizophrenia (Ajonijebu et al., 2017). In terms of specific drugs, methamphetamine and cocaine administration are related to expression changes of some DNMT isoforms that lead to changes in DNA methylation in the brain (Cadet et al., 2016). Furthermore,

cocaine is also related to an increased expression of methyl-CpG binding protein 2 (Cadet et al., 2016). Based on growing evidence that links DNA methylation with addiction, a pretreatment with methionine, a methyl donor, has been proposed as a possible therapy for this disorder (Brown and Feng, 2017).

Allele-specific methylation (ASM) is an epigenetic mechanism in which the genotype of a genetic variation correlates with DNA methylation (usually of a CpG dinucleotide) either in *cis* or in *trans* (Meaburn et al., 2010). It is quantitative and heterogeneous across tissues and individuals (Meaburn et al., 2010) and can lead to allele-specific expression (ASE) through the alteration of the activity of gene promoters (Gaur et al., 2013). During the last decade, several studies have focused on the identification of ASM SNPs in multiple human tissues (Do et al., 2017), some of them in brain regions (Gibbs et al., 2010; Zhang et al., 2010). Remarkably, ASM has previously been associated with psychiatric disorders, such as autism (Andrews et al., 2017), bipolar disorder (Chuang et al., 2013), schizophrenia (Gagliano et al., 2016; Wu and Pan, 2019) or ADHD (Pineda-Cirera et al., 2019).

The present study aims at elucidating the contribution of ASM to drug dependence susceptibility through a case-control study in a Spanish sample with European ancestry.

MATERIAL AND METHODS

Subjects

The discovery sample consisted of 578 substance-dependent patients (including 428 cocaine-dependent subjects) and 656 control individuals and the follow-up sample included 1,119 substance-dependent patients (with 843 cocaine-dependent subjects) and 1,092 control individuals (Supplementary Table 1 and Figure 1). Both samples had the same sex ratio in the case and control groups and all individuals were Spanish, with European ancestry and unrelated. Patients were diagnosed under DSM-IV-TR criteria (Diagnostic and Statistical Manual of Mental Disorders, 4th ed., text revision) and the SCID-I (Structured Clinical Interview) (First et al., 1997) at the Addiction and Dual Diagnosis Section of the Department of Psychiatry at Hospital Universitari Vall d'Hebron (Barcelona, Spain). Patients were included in the study if they met the criteria for dependence to at least one drug of abuse (Figure 1). The control individuals were blood donors recruited at the Blood and Tissues Bank of Barcelona, who had not injected drugs intravenously. The study was approved by the ethics committee of our institution according to

the Helsinki Declaration, the study design was reviewed by an appropriate ethical committee and informed consent was obtained from all participants. Population stratification was previously discarded in part of our sample (Fernández-Castillo et al., 2013).

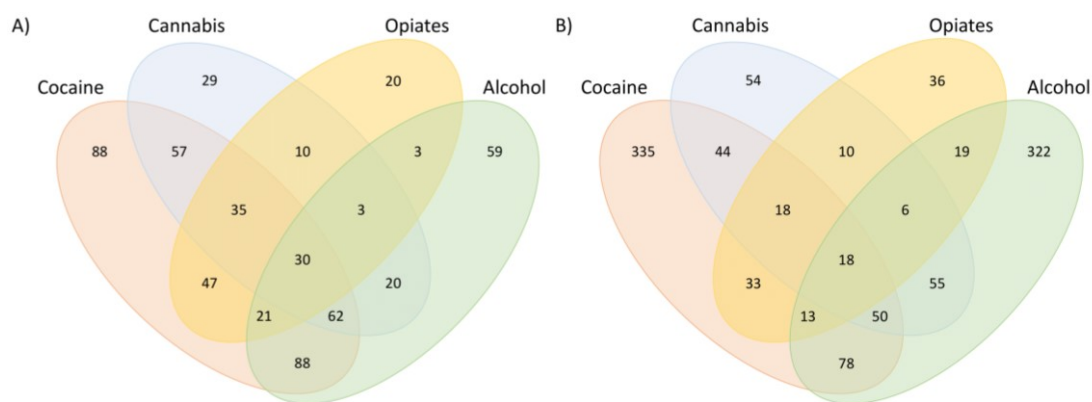


Figure 1. Distribution of the four main dependencies in the patients' sample (cocaine, cannabis, alcohol and opiates). Other dependencies with a frequency lower than 10% are not displayed. A) Discovery sample. B) Replication sample. The percentage of cases addicted to more than one drug is 65.7% (A) and 31.5% (B).

DNA isolation and quantification

Genomic DNA was isolated from peripheral blood using the salting-out method (Miller et al., 1988). DNA concentration of all samples was measured on a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Thermo Fisher Scientific Inc., Wilmington, DE, USA). DNA concentration for the discovery sample was also quantified using Picogreen (Life Technologies, Carlsbad, CA, USA) with a Glomax[®]-Multi Detection System (Promega, Madison, WI, USA).

Selection of Allele-Specific Methylation SNPs and genotyping

The SNP selection was made based on two previous studies that listed ASM variants for multiple brain regions of *postmortem* human samples (Gibbs et al., 2010; Zhang et al., 2010). Gibbs et al., (2010) considered four brain regions (cerebellum, frontal cortex, caudal pons and temporal cortex) of 150 subjects and Zhang et al., (2010) used only the cerebellum of another group of 153 subjects. Genomic variants located within the probe sequences used in methylation arrays may affect their binding affinity and cause inaccurate methylation measurements. To avoid potential artefacts in our results, we confirmed that the probes used to detect the CpG sites reported by us do not target genomic regions containing common genetic variants.

In the study by Zhang *et al.*, (2010) a total of 12,117 SNP-CpG pairs associations were reported in cerebellum and Gibbs *et al.*, (2010) listed a total of 12,135 SNP-CpG pairs in frontal cortex, 11,374 in caudal pons, 16,734 in temporal cortex and 12,102 in cerebellum (Figure 2). We combined the two studies and obtained a total of 43,132 SNP-CpG pairs involving 33,944 different SNPs and 5,306 CpG sites (Figure 2).

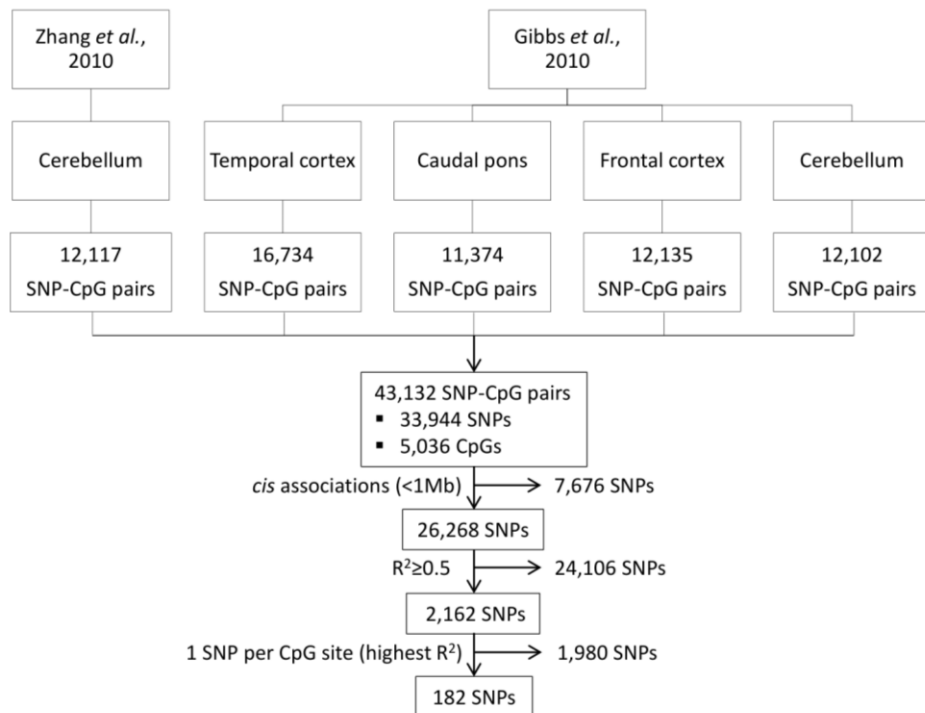


Figure 2. Selection of allele-specific methylation (ASM) SNPs from two previous studies.

We selected ASM SNPs showing high *cis* correlations between the SNP and the methylation levels of a CpG site (threshold set at $R^2 \geq 0.5$). Per each CpG site, we selected the SNP showing the highest R^2 and, in case of a tie, the SNP closest to the CpG site (Figure 2). In almost all cases the SNP showing the highest R^2 matched across tissues. However, occasionally this did not occur, so that for one of the tissues a variant different from the selected SNP displayed a higher correlation (R^2). In those cases, we calculated the linkage disequilibrium (LD) between the selected variant and the variant of highest R^2 in the particular tissue to take into account information redundancy. Thus, if the two SNPs were in high LD ($r^2 \geq 0.75$), we did not consider the second SNP, but in case of lower LD values ($r^2 < 0.75$), we also selected the SNP with higher R^2 for this particular tissue. Using these criteria, we ended up with 182 SNPs (Figure 2, Supplementary Table 2).

The linkage disequilibrium (LD) analysis was carried out using Haploview software (Barrett, 2009) with the Central European (CEU) reference panel from HapMap (www.hapmap.org; release 23).

The genotyping of the discovery sample was performed using the BioMark HD System (Fluidigm, CA, USA) by the “Servicio General de Genómica: Unidad de Secuenciación y Genotipado de la Universidad del País Vasco”, Spain. The genotyping of the replication sample was done with the KASP technology (LGC Biosearch Technologies, CA, USA) and was performed by LGC genomics, UK. Both assays are based on competitive allele-specific PCR that enable bi-allelic scoring of single nucleotide polymorphisms (SNPs) through a fluorescence-based detection. Primers used under the two genotyping technologies were designed by the corresponding company. Duplicates of some samples and negative controls were included in the genotyping of both the discovery and replication samples.

Statistical analysis

First, we assessed if there were significant age differences between cases and controls since means differed considerably (Supplementary Table 1). The Saphiro-Wilk test determined the data was not fitting a normal distribution and hence, we used the non-parametric test U Mann-Whitney with SPSS22 (SPSS Inc., Chicago, IL, USA) and obtained significant differences in age between cases and controls both in the discovery and in the replication samples.

We excluded from the association study those SNPs with more than 5% missing genotypes. Hardy-Weinberg equilibrium (HWE) was tested for all the SNPs separately in controls and cases using the SNPAssoc package of the R library (González et al., 2007), and those showing departure from HWE ($p < 0.01$) were excluded. The case-control association test was carried out using the same package, which considers four different models: additive, dominant, recessive and codominant. As we detected differences in age between cases and controls, we also calculated and presented all the p-values adjusted by age.

For the discovery sample we only considered the additive model to limit the number of tests. The SNPs that showed nominal associations in the discovery sample, and remained significant after adjusting by age, were subsequently assessed in the replica sample under the different genetic models. Finally, we conducted a pooled analysis of all the individuals from the discovery and replica samples under the model showing the best performance.

Functional annotation

We investigated the possible functional role of the associated ASM SNPs using different tools. First, we evaluated the predicted effect of the variant on the binding of transcription factors using the FeatSNP tool (Ma et al., 2019). Second, we evaluated whether the ASM SNPs are located in a region with histone marks related to enhancer regions (H3K4me1 and H3K27ac) and promoter regions (H3K4me3 and H3K9ac) using the Haploreg v4.1 tool (Ward and Kellis, 2016) and considering all the available brain regions. Finally, we inspected the possible effect of the associated variants on gene expression checking if they were described to be eQTLs using GTEx data (Release V8) (Lonsdale et al., 2013). We considered eQTL information for blood and from all available brain tissues.

RESULTS

The aim of the present study was to assess the contribution of SNPs showing allele-specific methylation in brain regions to drug dependence susceptibility. Starting from two previous studies (Gibbs et al., 2010; Zhang et al., 2010) that report more than 43,000 SNP-CpG pairs in several brain tissues, we selected 182 SNPs (Figure 2 and Supplementary Table 2). The selection was restricted to variants showing a high correlation in *cis* with methylation levels ($R^2 \geq 0.5$), taking only one SNP per CpG site to avoid redundancies. We further excluded 8 SNPs from the analysis that showed more than 5% of missing genotypes and one SNP that deviated from Hardy-Weinberg equilibrium (Supplementary Table 2). Finally, a total of 173 SNPs were investigated through a case-control association study in a Spanish sample of drug-dependent individuals and the corresponding controls, all with European ancestry (Supplementary Table 2).

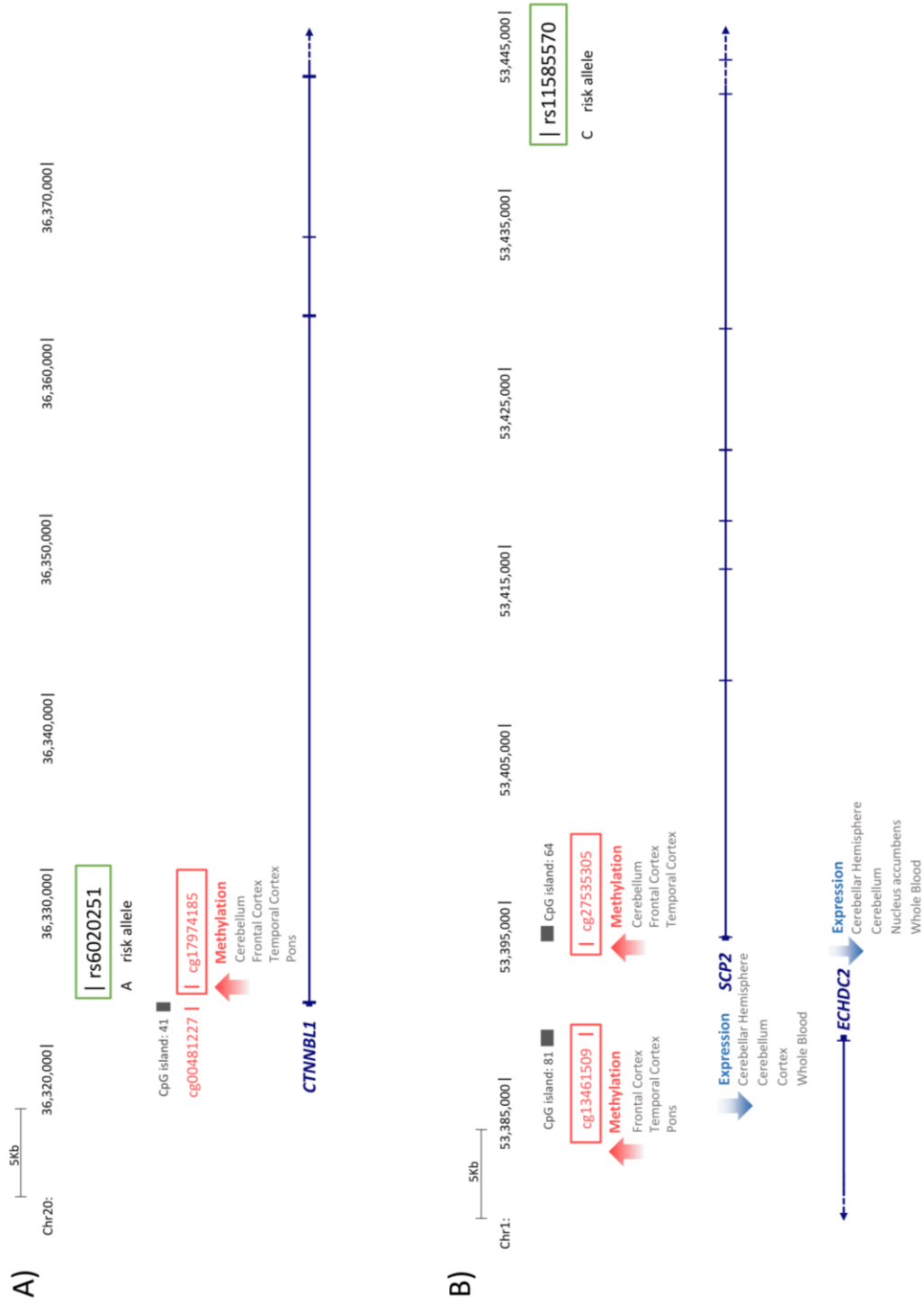
In the discovery sample (578 cases and 656 controls) we identified five SNPs nominally associated with drug dependence under the additive model: rs823080, rs3766612, rs6020251, rs9891018 and rs11585570 (Supplementary Table 3). We followed up these five associations in an independent sample (1,119 cases and 1,092 controls) and the association between rs6020251 and drug dependence was replicated under the recessive model ($P = 4.79E-02$; $Padj_{age} = 3.86E-02$) (Table 1A, Supplementary Table 4). We also performed a pooled analysis under the recessive model including the discovery and replication samples where rs6020251 ($P = 3.70E-03$; $Padj_{age} = 2.35E-03$) and also rs11585570 ($P = 1.54E-03$; $Padj_{age} = 7.65E-03$) remained associated with drug dependence (Table 1A).

We then performed the association analysis for the subset of cocaine-dependent patients (about 60% of our patients' sample) under the recessive model. Both variants showed the same direction of association as in the analysis for drug dependence (Table 1B).

Rs6020251 correlates with differential levels of methylation of cg17974185 ($R^2 \geq 0.6$) in all the four brain areas studied (Gibbs et al., 2010; Zhang et al., 2010), with the risk allele A associated with higher levels of methylation (Figure 3A and Table 2). The SNP is located 26 bp upstream from the CpG site (cg17974185) and both lie within the first intron of the *CTNNB1* reference transcript isoform, about 900 bp downstream from the transcription start site (TSS) and at the 5' end of other isoforms of the same gene (49-82 kb 5' from the TSS of the other isoforms) (Figure 3A). Remarkably, this variant lies in a region with multiple brain histone marks related to enhancer and promoter regions (Table 2 and Supplementary Table 5). Finally, it is also predicted to increase the binding affinity of four transcription factors: NKX3-2, NKX3-1, ZNF354C and ISL2 (Table 2 and Supplementary Table 6). It is important to highlight that the SNP rs6020251 is in high LD ($r^2 \geq 0.85$) with other SNPs that also correlated with differences in methylation of cg17974185 and were not genotyped in this study (Supplementary Table 7).

On the other hand, rs11585570 correlates with differential levels of methylation of cg27535305 ($R^2 > 0.5$) in three brain areas, but also with cg13461509 ($R^2 \geq 0.27$) in three regions (Figure 3B and Table 2). In both cases, the risk allele C is associated with higher levels of methylation (Figure 3B and Table 2). This variant is an eQTL for *SCP2* and *ECHDC2* in different brain areas and blood, with the risk allele C associated with less gene expression (Figure 3B and Table 2). Interestingly, both CpG sites, cg27535305 and cg13461509, lie in the possible promoter regions of the *SCP2* and *ECHDC2* genes (Figure 3B). The cg27535305 lies about 400 bp and 5.3 kb upstream from the TSS of the *SCP2* and the *ECHDC2* reference transcripts, respectively. The cg13461509 lies about 5.4 kb and 200 bp upstream from the TSS of the *SCP2* and the *ECHDC2* reference transcripts, respectively. Furthermore, rs11585570 is predicted to affect the binding of the transcription factor HOXA5, with an increase in the affinity when the risk allele C is present (Table 2 and Supplementary Table 6). The SNP rs11585570 is also in high LD with other SNPs affecting methylation of these two CpG sites and with a possible functional effect (Supplementary Table 7).

Figure 3. Genomic context of A) rs6020251 and the associated CpG site cg17974185. B) rs11585570 and the associated CpG sites cg27535305 and cg13461509. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG islands are represented in gray; CpG sites inspected in the reference studies appear in pink; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs; arrows indicate the effect on methylation/expression of the risk variants, with indication of the brain regions where it is described. GRCh37/hg19 assembly was used.



DISCUSSION

Epigenetic factors are known to mediate the link between genetics and environment, which makes them ideal candidates as risk factors for psychiatric conditions. DNA methylation, and particularly ASM, has been investigated as a potential etiological factor in multiple psychiatric disorders. An enrichment of ASM SNPs in GWAS has been reported in psychiatric disorders like schizophrenia and ADHD (Gagliano et al., 2016; Pineda-Cirera et al., 2019). In the present study, we have explored the possible contribution to substance use disorders of ASM variants previously described in the human brain and studied their possible impact on gene expression. Thus, we have been able to identify two variants (rs6020251 and rs11585570), that influence methylation at nearby CpG sites, associated with drug dependence and also with cocaine dependence through a case-control association study in a Spanish sample with European ancestry.

In general, ASM is associated with genetic variation in *cis* (up to 1 Mb), although there is also a proportion of variation in *trans*. Thus, SNPs located far away from CpG sites can also have an impact on methylation (Meaburn et al., 2010). There are multiple mechanisms that can lead to ASM, for instance SNPs that affect the binding of regulatory proteins like transcription factors or the transcriptional repressor CTCF (Do et al., 2016; Tycko, 2010). Any alteration of these protein-DNA interactions by the effect of SNP variants can lead to changes in the 3D structure of the DNA and therefore, produce effects on distant CpG sites. In the two studies that we used to select ASM in brain, most of the reported effects were in *cis* (up to 1Mb). Thus, Gibbs calculated an average distance of 81 kb between the SNP and the CpG site in associations in *cis*, which implies that SNPs that are distant from CpG sites can also influence their methylation (Gibbs et al., 2010). In general, in these studies, multiple SNPs (most of them in high linkage disequilibrium, LD) do correlate with methylation levels of a single CpG site. Considering this and in order to minimize redundancy, we did not inspect all the SNP variants reported to influence methylation at every CpG site. Instead, we selected the SNP showing the best correlation with methylation for each site. Therefore, there are other SNPs displaying ASM for the same CpG sites in high LD with the reported ones that could be the ones involved in the effect on DNA methylation (Supplementary Table 7). Importantly, the two variants identified in our study highlight CpG sites that could have a role in dependence.

The rs6020251 variant was associated with drug dependence and also specifically with cocaine dependence in two independent samples and correlates with the methylation levels of cg17974185, that lies in the first intron of the *CTNBL1* gene (Figure 3A). Noteworthy, this variant is in a region with multiple histone marks related to enhancer and promoter regions in multiple brain areas and is predicted to affect the binding of four TFBS (Table 2 and Supplementary Tables

5 and 6), reinforcing its role in regulation. Furthermore, this SNP has been reported to be an mQTL in human fetal brain (Hannon et al., 2015). The *CTNBL1* is expressed in most tissues, including brain, and encodes a component of the PRP19-CDC5L protein complex, an integral part of the spliceosome involved in the activation of pre-mRNA splicing (van Maldegem et al., 2015). Interestingly, *CTNBL1* has been associated with obesity (Liu et al., 2008) and with memory (Papassotiropoulos et al., 2013), both related to drug dependence (Goodman and Packard, 2016; Kutlu and Gould, 2016; Volkow et al., 2013). The effect of food in obesity resembles that of the drug in addiction, as both result in powerful reinforcing effects characterized by an increase of dopamine in brain (Volkow et al., 2013). Additionally, it has been suggested that individuals presenting obesity or drug dependence have impairments in the dopaminergic pathways that mediate reward (Volkow et al., 2013). About memory, it is known that drug use is related to maladaptive memories that support drug addiction and relapse. Moreover, drugs of abuse are often related to disrupted learning (Goodman and Packard, 2016; Kutlu and Gould, 2016).

The rs11585570 variant was found associated with drug and with cocaine dependence in the discovery sample but not in the replication dataset, although the pooled analysis yielded significant results. This variant correlates with differential levels of methylation of two CpG sites (cg27535305 and cg13461509) that lie in the promoter regions of the *SCP2* and *ECHDC2* genes (Figure 3B) and is also predicted to affect the binding of a transcription factor, HOXA5 (Table 2 and Supplementary Table 6). In general, DNA methylation in promoter regions inversely correlates with levels of gene expression (Li and Zhang, 2014), and this is observed with variant rs11585570, where the risk allele C is associated with hypermethylation of cg27535305 and cg13461509 in brain and to downregulation of *SCP2* and *ECHDC2* in brain and blood (Figure 3B and Table 2). Furthermore, cg13461509 methylation also correlates with the expression of *SCP2* in peripheral blood (Wu et al., 2018), pointing it out as a possible biomarker for the disorder. Remarkably, the variant rs11585570 is an eQTL for the *ECHDC2* gene in the nucleus accumbens, a key area in addiction. Addictive drugs increase the extracellular concentration of dopamine in the nucleus accumbens, a neurotransmitter involved in reward and motivation, two main features of addiction (Di Chiara, 2002; Di Chiara et al., 2004; Koob and Volkow, 2016). This effect on gene expression in the nucleus accumbens is highlighting a possible window to be studied in depth to understand the specific role of this variant to drug dependence susceptibility. The *SCP2* and *ECHDC2* genes are expressed in most tissues, including brain. *ECHDC2* is involved in the metabolism of fatty acids and lipids and *SCP2* plays a role in the lipid transfer between membranes and interestingly, was found increased in the brain of mice exposed to alcohol (Myers-Payne et al., 1996).

All these evidences and predictions together with the results obtained in the present work support the functional role of these variants at an epigenetic level, involving methylation of the CpG sites cg17974185, cg27535305 and cg13461509 in addition. Furthermore, two of the highlighted genes (*SCP2* and *CTNBL1*) were previously linked to other disorders or conditions related to drug dependence (Liu et al., 2008; Myers-Payne et al., 1996; Papassotiropoulos et al., 2013).

There are several strengths and limitations of this study that should be discussed. Multiple factors may influence the statistical power of an association study, for instance sample size, homogeneity of the sample used or prevalence of the disease in the general population (Nsengimana and Bishop, 2017). Sample size is one of the most important elements that contribute to statistical power and it directly correlates with it (Hong and Park, 2012). The sample size of this study is substantially larger than the ones used in previous studies on drug dependence; nonetheless, it is still limited taking into account that most of the variants contributing to complex disorders like addiction have modest effects. Homogeneity of the sample used is also relevant, since heterogeneity could lead to false positives (Nsengimana and Bishop, 2017). In order to control this, all subjects included in our study were Spanish, with European ancestry, recruited in Barcelona (Spain), sex ratios were similar in cases and controls, and all patients were evaluated in the same hospital and under the same clinical assessment. Although we could not assess drug dependence in control individuals, the impact of the contamination of cases within the control sample is expected to be low, given that only 0.7% of the global population aged 15-64 suffer from illicit drug use disorders (UNODC, 2019) and 5.1% of the global population aged 15 years and older suffer from alcohol use disorders (World Health Organization, 2018). This design could eventually dilute positive findings (false negative results), but importantly, it is not a reason for the occurrence of false positive results. In association studies it is also important to minimize marker redundancies to improve statistical power. Lastly, replication studies in independent samples are widely used to confirm the association and direction of the effect of genetic variants identified in a discovery sample (Chanock et al., 2007). When a variant replicates in a follow-up sample, its association with the evaluated phenotype is reinforced (Chanock et al., 2007). In this work, although the associations found in the discovery sample were nominal, we were able to replicate one of them in a second sample that doubled the size of the discovery. Moreover, this hit and a second one showed association in a pooled analysis combining the discovery and replication samples, strengthening the results obtained.

In conclusion, we report an association of rs6020251 and rs11585570 with drug dependence and with cocaine dependence. Interestingly, these SNPs correlate with methylation levels of nearby CpG sites (Gibbs et al., 2010; Zhang et al., 2010) that are located in possible regulatory regions of the *CTNBL1* or the *SCP2* and *ECHDC2* genes, respectively. In addition, the risk allele of rs11585570 is associated with a lower expression of *SCP2* and *ECHDC2* in brain and blood. Our results highlight these genes as new candidates for involvement in the susceptibility to drug dependence, although further functional studies are needed to ascertain their possible role in the disorder.

CONFLICT OF INTEREST

JAR-Q was on the speakers' bureau and/or acted as consultant for Eli-Lilly, Janssen-Cilag, Novartis, Shire, Takeda, Bial, Shionogui, Lundbeck, Almirall, Braingaze, Sincrolab, Medice and Rubió in the last 5 years. He also received travel awards (air tickets + hotel) for taking part in psychiatric meetings from Janssen-Cilag, Rubió, Shire, Takeda, Shionogui, Bial, Medice and Eli-Lilly. The Department of Psychiatry chaired by him received unrestricted educational and research support from the following companies in the last 5 years: Eli-Lilly, Lundbeck, Janssen-Cilag, Actelion, Shire, Ferrer, Oryzon, Roche, Psious, and Rubió. RFP-A has received fees to give talks for Exeltis, Lundbeck, and Takeda. LG-L has received fees to give talks for Janssen and Lundbeck. The other authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The authors thank technical and human support provided by SGIker of UPV/EHU and European funding (ERDF and ESF). The authors are also grateful to "Agència de Salut Pública de Barcelona", the Blood and Tissues Bank of Barcelona donors, patients, families and clinical collaborators who contributed to the recruitment of patients (Dra. Nieves Martinez-Luna, Dra. Elena Ros-Cucurull, Eduardo Castrillo and Miguel Ángel Cantillo).

BIBLIOGRAPHY

- Ajonijebu, D.C., Abboussi, O., Russell, V.A., Mabandla, M. V., Daniels, W.M.U., 2017. Epigenetics: a link between addiction and social environment. *Cell. Mol. Life Sci.* 74, 2735–2747.
- Andrews, S. V., Ellis, S.E., Bakulski, K.M., Sheppard, B., Croen, L.A., Hertz-Picciotto, I., Newschaffer, C.J., Feinberg, A.P., Arking, D.E., Ladd-Acosta, C., Fallin, M.D., 2017. Cross-tissue integration of genetic and epigenetic data offers insight into autism spectrum disorder. *Nat. Commun.* 8, 1011.
- Barrett, J.C., 2009. Haploview: Visualization and Analysis of SNP Genotype Data. *Cold Spring Harb. Protoc.* 4, 1–5.
- Bashtrykov, P., Jeltsch, A., 2017. Epigenome Editing in the Brain, in: Delgado-Morales R. (Ed.), *Neuroepigenomics in Aging and Disease. Advances in Experimental Medicine and Biology.* Springer, Cham, pp. 409–424. <https://doi.org/10.1007/978-3-319-53889-1>
- Brown, A.N., Feng, J., 2017. Drug addiction and DNA modifications, in: Delgado-Morales R. (Ed.), *Neuroepigenomics in Aging and Disease. Advances in Experimental Medicine and Biology.* Springer, Cham, pp. 105–125.
- Bühler, K.-M., Giné, E., Echeverry-Alzate, V., Calleja-Conde, J., de Fonseca, F.R., López-Moreno, J.A., 2015. Common single nucleotide variants underlying drug addiction: more than a decade of research. *Addict. Biol.* 20, 845–871.
- Cadet, J.L., McCoy, M.T., Jayanthi, S., 2016. Epigenetics and addiction. *Clin. Pharmacol. Ther.* 99, 502–511.
- Chanock, S.J., Manolio, T., Boehnke, M., Boerwinkle, E., Hunter, D.J., Thomas, G., Hirschhorn, J.N., Abecasis, G., Altshuler, D., Bailey-Wilson, J.E., Brooks, L.D., Cardon, L.R., Daly, M., Donnelly, P., Fraumeni, J.F., Freimer, N.B., Gerhard, D.S., Gunter, C., Guttmacher, A.E., Guyer, M.S., Harris, E.L., Hoh, J., Hoover, R., Kong, C.A., Merikangas, K.R., Morton, C.C., Palmer, L.J., Phimister, E.G., Rice, J.P., Roberts, J., Rotimi, C., Tucker, M.A., Vogan, K.J., Wacholder, S., Wijsman, E.M., Winn, D.M., Collins, F.S., 2007. Replicating genotype-phenotype associations. *Nature* 447, 655–660. <https://doi.org/10.1038/447655a>
- Chuang, L.-C., Kao, C.-F., Shih, W.-L., Kuo, P.-H., 2013. Pathway analysis using information from allele-specific gene methylation in genome-wide association studies for bipolar disorder. *PLoS ONE* 8, e53092.
- Di Chiara, G., 2002. Nucleus accumbens shell and core dopamine: Differential role in behavior and addiction. *Behav. Brain Res.* 137, 75–114.
- Di Chiara, G., Bassareo, V., Fenu, S., De Luca, M.A., Spina, L., Cadoni, C., Acquas, E., Carboni, E., Valentini, V., Lecca, D., 2004. Dopamine and drug addiction: The nucleus accumbens shell connection. *Neuropharmacology* 47, 227–241.
- Do, C., Lang, C.F., Lin, J., Darbary, H., Krupska, I., Gaba, A., Petukhova, L., Vonsattel, J.P., Gallagher, M.P., Goland, R.S., Clynes, R.A., Dwork, A., Kral, J.G., Monk, C., Christiano, A.M., Tycko, B., 2016. Mechanisms and Disease Associations of Haplotype-Dependent Allele-Specific DNA Methylation. *Am. J. Hum. Genet.* 98, 934–955. <https://doi.org/10.1016/j.ajhg.2016.03.027>
- Do, C., Shearer, A., Suzuki, M., Terry, M.B., Gelernter, J., Grealley, J.M., Tycko, B., 2017. Genetic-epigenetic interactions in cis: a major focus in the post-GWAS era. *Genome Biol.* 18, 120.
- Ducci, F., Goldman, D., 2012. The genetic basis of addictive disorders. *Psychiatr. Clin. North Am.* 35, 495–519.
- Edenberg, H.J., Gelernter, J., Agrawal, A., 2019. Genetics of Alcoholism. *Curr. Psychiatry Rep.* 21, 26. <https://doi.org/10.1007/s11920-019-1008-1>
- Fernández-Castillo, N., Roncero, C., Grau-Lopez, L., Barral, C., Prat, G., Rodriguez-Cintas, L., Sánchez-Mora, C., Gratacòs, M., Ramos-Quiroga, J.A., Casas, M., Ribasés, M., Cormand, B., 2013. Association study of 37 genes related to serotonin and dopamine neurotransmission and neurotrophic factors in cocaine dependence. *Genes. Brain. Behav.* 12, 39–46.
- First, M.B., Spitzer, S.R., Gibbon, M., Williams, J.B.M., 1997. Structured Clinical Interview for DSM-IV Axis I Disorders (SCID). *Am. Psychiatr. Press.*
- Gagliano, S.A., Ptak, C., Mak, D.Y.F., Shamsi, M., Oh, G., Knight, J., Boutros, P.C., Petronis, A., 2016. Allele-Skewed DNA Modification in the Brain: Relevance to a Schizophrenia GWAS. *Am. J. Hum. Genet.* 98, 956–962.

- Gaur, U., Li, K., Mei, S., Liu, G., 2013. Research progress in allele-specific expression and its regulatory mechanisms. *J. Appl. Genet.* 54, 271–283.
- Gibbs, J.R., van der Brug, M.P., Hernandez, D.G., Traynor, B.J., Nalls, M.A., Lai, S.L., Arepalli, S., Dillman, A., Rafferty, I.P., Troncoso, J., Johnson, R., Zielke, H.R., Ferrucci, L., Longo, D.L., Cookson, M.R., Singleton, A.B., 2010. Abundant quantitative trait loci exist for DNA methylation and gene expression in Human Brain. *PLoS Genet.* 6, e1000952.
- González, J.R., Armengol, L., Solé, X., Guinó, E., Mercader, J.M., Estivill, X., Moreno, V., 2007. SNPAssoc: An R package to perform whole genome association studies. *Bioinformatics* 23, 644–645.
- Goodman, J., Packard, M.G., 2016. Memory systems and the addicted brain. *Front. Psychiatry* 7, 24.
- Hamilton, P.J., Nestler, E.J., 2019. Epigenetics and addiction. *Curr. Opin. Neurobiol.* 59, 128–136. <https://doi.org/10.1016/j.conb.2019.05.005>
- Hannon, E., Spiers, H., Viana, J., Pidsley, R., Burrage, J., Murphy, T.M., Troakes, C., Turecki, G., O'Donovan, M.C., Schalkwyk, L.C., Bray, N.J., Mill, J., 2015. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nat. Neurosci.* 19, 48–54.
- Hong, E.P., Park, J.W., 2012. Sample Size and Statistical Power Calculation in Genetic Association Studies. *Genomics Inform.* 10, 117. <https://doi.org/10.5808/gi.2012.10.2.117>
- Jensen, K.P., 2016. A Review of Genome-Wide Association Studies of Stimulant and Opioid Use Disorders. *Mol. neuropsychiatry* 2, 37–45.
- Klose, R.J., Bird, A.P., 2006. Genomic DNA methylation: The mark and its mediators. *Trends Biochem. Sci.* 31, 89–97. <https://doi.org/10.1016/j.tibs.2005.12.008>
- Koob, G.F., Volkow, N.D., 2016. Neurobiology of addiction: a neurocircuitry analysis. *The Lancet Psychiatry* 3, 760–773. [https://doi.org/10.1016/S2215-0366\(16\)00104-8](https://doi.org/10.1016/S2215-0366(16)00104-8)
- Kutlu, M.G., Gould, T.J., 2016. Effects of drugs of abuse on hippocampal plasticity and hippocampus-dependent learning and memory: contributions to development and maintenance of addiction. *Learn. Mem.* 23, 515–533.
- Ladd-Acosta, C., Pevsner, J., Sabunciyan, S., Yolken, R.H., Webster, M.J., Dinkins, T., Callinan, P.A., Fan, J.-B., Potash, J.B., Feinberg, A.P., 2007. DNA Methylation Signatures within the Human Brain. *Am. J. Hum. Genet.* 81, 1304–1315.
- Li, E., Zhang, Y., 2014. DNA methylation in mammals. *Cold Spring Harb. Perspect. Biol.* 6, a019133. <https://doi.org/10.1101/cshperspect.a019133>
- Liu, Y.J., Liu, X.G., Wang, L., Dina, C., Yan, H., Liu, J.F., Levy, S., Papasian, C.J., Drees, B.M., Hamilton, J.J., Meyre, D., Delplanque, J., Pei, Y.F., Zhang, L., Recker, R.R., Froguel, P., Deng, H.W., 2008. Genome-wide association scans identified CTNBL1 as a novel gene for obesity. *Hum. Mol. Genet.* 17, 1803–1813. <https://doi.org/10.1093/hmg/ddn072>
- Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., Foster, B., Moser, M., Karasik, E., Gillard, B., Ramsey, K., Sullivan, S., Bridge, J., Magazine, H., Syron, J., Fleming, J., Siminoff, L., Traino, H., Mosavel, M., Barker, L., Jewell, S., Rohrer, D., Maxim, D., Filkins, D., Harbach, P., Cortadillo, E., Berghuis, B., Turner, L., Hudson, E., Feenstra, K., Sobin, L., Robb, J., Branton, P., Korzeniewski, G., Shive, C., Tabor, D., Qi, L., Groch, K., Nampally, S., Buia, S., Zimmerman, A., Smith, A., Burges, R., Robinson, K., Valentino, K., Bradbury, D., Cosentino, M., Diaz-Mayoral, N., Kennedy, M., Engel, T., Williams, P., Erickson, K., Ardlie, K., Winckler, W., Getz, G., DeLuca, D., Daniel MacArthur, Kellis, M., Thomson, A., Young, T., Gelfand, E., Donovan, M., Meng, Y., Grant, G., Mash, D., Marcus, Y., Basile, M., Liu, J., Zhu, J., Tu, Z., Cox, N.J., Nicolae, D.L., Gamazon, E.R., Im, H.K., Konkashbaev, A., Pritchard, J., Stevens, M., Flutre, T., Wen, X., Dermitzakis, E.T., Lappalainen, T., Guigo, R., Monlong, J., Sammeth, M., Koller, D., Battle, A., Mostafavi, S., McCarthy, M., Rivas, M., Maller, J., Rusyn, I., Nobel, A., Wright, F., Shabalina, A., Feolo, M., Sharopova, N., Sturcke, A., Paschal, J., Anderson, J.M., Wilder, E.L., Derr, L.K., Green, E.D., Struwing, J.P., Temple, G., Volpi, S., Boyer, J.T., Thomson, E.J., Guyer, M.S., Ng, C., Abdallah, A., Colantuoni, D., Insel, T.R., Koester, S.E., A Roger Little, Bender, P.K., Lehner, T., Yao, Y., Compton, C.C., Vaught, J.B., Sawyer, S., Lockhart, N.C., Demchok, J., Moore, H.F., 2013. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* 45, 580–585. <https://doi.org/10.1038/ng.2653>
- Ma, C., Madden, P., Gontarz, P., Wang, T., Zhang, B., 2019. FeatSNP : An Interactive Database for Brain-Specific Epigenetic Annotation of Human SNPs. *Front. Genet.* 10, 262.

- Meaburn, E.L., Schalkwyk, L.C., Mill, J., 2010. Allele-specific methylation in the human genome: Implications for genetic studies of complex disease. *Epigenetics* 5, 578–582.
- Miller, S.A., Dykes, D.D., Polesky, H.F., 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16, 1215.
- Mulligan, M.K., 2019. Genetic Factors in Cannabinoid Use and Dependence, in: Bukiya A. (Ed.), *Recent Advances in Cannabinoid Physiology and Pathology. Advances in Experimental Medicine and Biology.* Springer, Cham, pp. 129–150. https://doi.org/10.1007/978-3-030-21737-2_7
- Myers-Payne, S.C., Fontaine, R.N., Loeffler, A., Pu, L., Rao, A.M., Kier, A.B., Wood, W.G., Schroeder, F., 1996. Effects of chronic ethanol consumption on sterol transfer proteins in mouse brain. *J. Neurochem.* 66, 313–320.
- Nsengimana, J., Bishop, D.T., 2017. Design considerations for genetic linkage and association studies, in: Elston, R. (Ed.), *Statistical Human Genetics. Methods in Molecular Biology.* Humana Press, New York, NY, pp. 257–281. https://doi.org/10.1007/978-1-4939-7274-6_13
- Papassotiropoulos, A., Stefanova, E., Vogler, C., Gschwind, L., Ackermann, S., Spalek, K., Rasch, B., Heck, A., Aerni, A., Hanser, E., Demougin, P., Huynh, K.D., Luechinger, R., Klarhöfer, M., Novakovic, I., Kostic, V., Boesiger, P., Scheffler, K., De Quervain, D.J.F., 2013. A genome-wide survey and functional brain imaging study identify CTNBL1 as a memory-related gene. *Mol. Psychiatry* 18, 255–263.
- Pineda-Cirera, L., Shivalikanjli, A., Cabana-Domínguez, J., Demontis, D., Rajagopal, V.M., Børglum, A.D., Faraone, S. V., Cormand, B., Fernández-Castillo, N., 2019. Exploring genetic variation that influences brain methylation in attention-deficit/hyperactivity disorder. *Transl. Psychiatry* 9, 242. <https://doi.org/10.1038/s41398-019-0574-7>
- Tycko, B., 2010. Allele-specific DNA methylation: Beyond imprinting. *Hum. Mol. Genet.* 19, 210–220. <https://doi.org/10.1093/hmg/ddq376>
- UNODC, 2019. World Drug Report Sales No. E.19.XI.9.
- van Maldegem, F., Maslen, S., Johnson, C.M., Chandra, A., Ganesh, K., Skehel, M., Rada, C., 2015. CTNBL1 facilitates the association of CWC15 with CDC5L and is required to maintain the abundance of the Prp19 spliceosomal complex. *Nucleic Acids Res.* 43, 7058–7069.
- Volkow, N.D., Wang, G.J., Tomasi, D., Baler, R.D., 2013. Obesity and addiction: Neurobiological overlaps. *Obes. Rev.* 14, 2–18. <https://doi.org/10.1111/j.1467-789X.2012.01031.x>
- Ward, L.D., Kellis, M., 2016. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* 44, D877–D881.
- World Health Organization, 2018. Global status report on alcohol and health. Geneva. Licence: BY-NC-SA 3.0 IGO.
- Wu, C., Pan, W., 2019. Integration of methylation QTL and enhancer–target gene maps with schizophrenia GWAS summary results identifies novel genes. *Bioinformatics* 35, 3576–3583.
- Wu, Y., Zeng, J., Zhang, F., Zhu, Z., Qi, T., Zheng, Z., Lloyd-Jones, L.R., Marioni, R.E., Martin, N.G., Montgomery, G.W., Deary, I.J., Wray, N.R., Visscher, P.M., McRae, A.F., Yang, J., 2018. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat. Commun.* 9, 918.
- Zhang, D., Cheng, L., Badner, J. a., Chen, C., Chen, Q., Luo, W., Craig, D.W., Redman, M., Gershon, E.S., Liu, C., 2010. Genetic Control of Individual Differences in Gene-Specific Methylation in Human Brain. *Am. J. Hum. Genet.* 86, 411–419.

Supplementary data for:

Exploring allele specific methylation in drug dependence susceptibility

Laura Pineda-Cirera^{a,b,c,d}, Judit Cabana-Domínguez^{a,b,c,d}, Lara Grau-López^{e,f,g,h}, Constanza Daigre^{e,f,g,h}, Cristina Sánchez-Mora^{a,e,g,i}, Raul Felipe Palma-Álvarez^{e,f,g,h}, Josep Antoni Ramos-Quiroga^{e,f,g,h}, Marta Ribasés^{a,e,g,i}, Bru Cormand^{a,b,c,d} & Noèlia Fernández-Castillo^{a,b,c,d} &

Supplementary Table 1. Descriptive characteristics of the Spanish individuals with European ancestry with substance dependence and controls.

	Discovery sample		Replication sample	
	Cases N = 578	Controls N = 656	Cases N = 1119	Controls N = 1092
Gender N (%)				
Male	448 (77.4)	508 (77.5)	843 (75.3)	824 (75.5)
Female	130 (22.6)	148 (22.5)	275 (24.7)	268 (24.5)
Age (mean and SD)				
	38.2 ± 9.6	55 ± 16.2	44.1 ± 12.8	53.3 ± 16.9
Substance dependence N (%)				
Cocaine	428 (74.0)	-	589 (52.6)	-
Alcohol	286 (49.5)	-	561 (50.1)	-
Cannabis	246 (42.6)	-	255 (22.8)	-
Opiates	169 (29.2)	-	153 (13.7)	-
Benzodiazepines	50 (8.7)	-	88 (7.9)	-
Others	18 (3.1)	-	31 (2.8)	-

Resultats capítol 2. Article 3

Supplementary Table 2. ASM SNP acting in *cis* selected.

SNP	Chr	SNP Position	p-value (additive model)		Comments	IlmnID	CpG position	CrbI [#]	CrbI*	Fctx*	Pons*	Tctx*
			Not corrected by age	Corrected by age								
rs2268175	1	9223195	8.59E-01	7.45E-01		cg22601917	9216836		x	x	x	x
rs6429769	1	15841197	5.36E-01	3.00E-01		cg11706911	15783576	x	x	x	x	x
						cg17385448	15784264	x	x			
rs666371	1	20914142	3.49E-01	7.54E-01		cg15613048	20916593			x		x
rs926438	1	25626225	2.49E-01	6.69E-01		cg06388544	25471465	x				
						cg24857721	25471642	x		x	x	x
rs7520588	1	39987112	8.44E-01	8.41E-01		cg24920358	39976897		x	x	x	x
rs663818	1	44351146	8.92E-01	6.60E-01		cg08679985	44356912	x	x	x	x	x
rs11585570	1	53210718	<u>1.22E-02</u>	<u>1.60E-02</u>		cg13461509	53160189		x	x		x
						cg27535305	53165213			x	x	x
rs7536099	1	70594345	7.30E-01	9.47E-01		cg19342782	70594369			x	x	x
rs2994948	1	84824430	9.30E-01	6.16E-01		cg08380539	84812265			x	x	x
rs6687387	1	94118915	1.47E-01	2.44E-01		cg22807700	94118285	x	x			
rs3754446	1	110054764	5.86E-01	6.44E-01		cg04987894	110056114		x	x	x	x
						cg23719124	110056467			x	x	x
rs1888962	1	151019420	6.40E-02	1.19E-01		cg14696820	151066817					x
						cg15531099	151037151		x	x	x	x
rs1048535	1	151044089	8.09E-01	7.64E-01		cg03960217	150914376					x
						cg14696820	151066817					x
						cg15531099	151037151		x	x	x	x
						cg21065959	151025404					x
						cg21846488	150947442					x
						cg23413307	151014121					x
						cg24884084	151269797					x
rs4515830	1	151903096	6.84E-01	6.24E-01		cg07576541	151897532			x	x	x
						cg23851011	151852564		x			
rs6670721	1	158803185	7.40E-01	9.99E-01		cg02945019	158817137		x			
rs7410867	1	159099423	8.15E-01	4.32E-01		cg10106388	159099286	x				
rs17400053	1	159629445	9.39E-01	7.68E-01		cg10431340	159545757		x	x	x	x
rs4656311	1	159765474	1.27E-01	1.83E-01		cg04384208	159785995		x	x	x	x
						cg04567009	159867368					x
rs7549675	1	160747192	1.49E-01	3.82E-01		cg07175007	160733834	x	x	x	x	x
						cg19451466	160734090					x
rs2281962	1	165326384	2.36E-01	2.12E-01		cg06665322	165325964	x	x	x	x	x
rs1063412	1	170677590	1.39E-01	1.38E-01		cg08587864	170680210		x			
rs1951627	1	172158440	1.43E-01	2.77E-01		cg14047008	172153659	x	x		x	x
rs16840522	1	194977539	3.67E-01	1.56E-01		cg04614339	195010270		x	x	x	x
						cg12687463	195054448		x			x
rs4915215	1	199343835	1.18E-01	3.49E-01		cg22815214	199349743	x	x	x	x	x
rs823121	1	203990925			Genotyping problems	cg14159672	204085827	x	x	x	x	x
						cg14893161	204085849	x	x	x	x	
rs823080	1	204055905	<u>1.98E-02</u>	<u>3.15E-02</u>		cg14159672	204085827		x	x	x	x
						cg14893161	204085849		x	x	x	
rs6685182	1	208034942	4.60E-01	1.83E-01		cg23283495	208046427	x	x	x	x	x
rs3766612	1	208086636	<u>2.10E-04</u>	<u>4.00E-04</u>		cg23166289	208067705	x				
rs4659664	1	234621933	8.39E-01	5.89E-01		cg09809672	234624280	x	x			
rs3844080	1	245373456	6.88E-01	5.26E-01		cg15590526	245401267		x	x	x	x
						cg18354594	245343270	x				
						cg21918500	245402069		x	x		x
rs4854183	2	3348056	4.17E-01	9.82E-01		cg12447832	3362239	x	x	x	x	x
rs13006081	2	24198418	7.49E-01	6.32E-01		cg04809136	24153687					x
						cg08965143	24161775		x			
rs1662955	2	30524477	8.33E-01	6.89E-01		cg17749961	30523342		x	x	x	x
rs829584	2	30530779			Genotyping problems	cg17749961	30523342	x	x	x	x	x
rs920391	2	42872851	1.60E-01	3.47E-01		cg01561916	42873648	x	x			x
						cg20289949	42873118	x				x
rs7583085	2	42876472			Genotyping problems	cg01561916	42873648	x	x			x
						cg20289949	42873118	x				x
rs2357691	2	160433385	8.67E-01	8.89E-01		cg10107725	160469777	x	x			
						cg15352315	160363282	x				
						cg23995753	160469003	x	x	x	x	x
rs6744553	2	160469045	2.67E-01	3.38E-01		cg23995753	160469003		x			x
rs6735923	2	160477834	4.27E-01	8.29E-01		cg10107725	160469777	x	x			
						cg23995753	160469003	x	x	x	x	x
						cg10107725	160469777		x			
						cg23995753	160469003		x	x	x	x
rs836722	2	168383434	1.46E-01	1.33E-01		cg19723473	168383261		x			
						cg25427580	168383578		x			
rs854572	7	94792632			Genotyping problems	cg07404485	94791614	x	x	x	x	x
						cg24750391	94864147	x				
rs10932241	2	208685200	1.81E-01	1.79E-01		cg25356886	208697187			x	x	x
rs4627502	2	209020883	6.59E-01	7.83E-01		cg07915542	208828578		x			
rs10197460	2	234253929	4.55E-01	6.49E-01		cg10432859	234255304	x	x			

Resultats capítol 2. Article 3

					cg27418851	54201659	x					
rs3849969	10	75196005	6.60E-01	5.53E-01	cg16540259	75242201	x	x	x			x
rs2180968	10	89408710	1.56E-01	4.23E-01	cg13926569	89408903	x	x				
rs1061437	10	100166144	9.18E-01	7.19E-01	cg03852570	100165121	x	x				
					cg08397758	100164868	x	x			x	
rs908365	10	126134955	3.43E-01	5.16E-01	cg07596401	126140051	x	x	x	x	x	x
rs10829605	10	131234403	7.22E-01	7.17E-01	cg07638938	131238614			x	x	x	x
					cg20537325	131238436			x			
rs7087388	10	134607844	1.79E-01	6.17E-01	cg01757745	134605566				x	x	
rs16928297	11	2442696	4.43E-01	6.22E-01	cg01734338	2447832	x	x	x	x	x	x
rs3758949	11	17457639	6.15E-01	2.03E-01	cg13185308	17455420		x				
rs2403279	11	18373191	5.54E-01	3.20E-01	cg05740244	18390566			x	x		
rs10833312	11	20342877	6.73E-01	4.75E-01	cg15576195	20342112		x	x	x	x	x
rs4930461	11	67151471	9.33E-01	7.04E-01	cg05244766	67107075	x					
					cg06235429	67129715	x	x	x			x
					cg09701102	67129634	x	x	x			
					cg06521852	36471365	x					
rs360718	11	111539949	6.36E-01	8.94E-01	cg11304234	111539986	x	x				
rs11216858	11	117742615	4.44E-01	5.80E-01	cg15880738	117720297	x	x	x			x
rs10772679	12	7769867	2.30E-01	1.80E-01	cg01120761	7794412		x	x			
rs11054075	12	10963569	5.64E-02	1.17E-01	cg04520391	11399184	x			x	x	
					cg14076161	11354671	x	x		x	x	x
					cg21874193	11066165	x					
					cg27345534	11400850	x					x
rs7134036	12	11076851	3.50E-01	5.47E-01	cg25677688	11067628	x	x	x	x	x	x
rs3759129	12	48640704	<u>2.80E-02</u>	5.39E-02	cg23855989	48642113		x	x	x	x	x
rs12824366	12	51271560	7.42E-01	5.73E-01	cg06494782	51281688	x	x				
rs10747675	12	52286298	7.58E-01	8.94E-02	cg08995424	52357158	x					
					cg22546318	52356620	x	x				
rs3204635	12	55923860	8.09E-01	9.45E-01	cg14747072	55931400			x	x	x	x
					cg21878918	55921054			x			x
rs2082529	12	74713746	3.28E-01	6.16E-01	cg27182761	74712885			x	x	x	x
rs1880867	12	78121579	8.57E-01	5.81E-01	cg22333868	78135486	x					
rs8181778	12	78146034	9.32E-01	5.92E-01	cg22333868	78135486	x					
rs1153051	12	79599649	8.29E-01	7.18E-01	cg21126707	79635143	x					
					cg26207503	79634855	x	x				
rs10859990	12	94860170	6.53E-01	4.10E-01	cg25229172	94860227		x				
rs11068050	12	108660313	2.44E-01	2.48E-01	cg01335367	108656616						x
					cg02351381	108656754						x
rs8001444	13	94750438	8.21E-01	8.78E-01	cg24476569	94752408		x	x	x	x	x
rs2277424	13	110320421	6.81E-02	5.79E-02	cg04513422	110320340	x	x				
					cg22131172	110320891		x				
					cg22131172	110320866	x					
rs2273629	14	23968591	7.91E-01	6.43E-01	cg13105904	23969859			x	x	x	x
rs10151959	14	34171345	2.25E-01	3.25E-01	cg17346022	34169846	x		x	x	x	x
					cg24276491	34079218	x	x				
rs6574232	14	74967628	9.98E-01	8.48E-01	cg22143352	74967569		x				
rs17103483	14	75115983	7.42E-01	5.17E-01	cg04001333	75115101	x					
rs7152947	14	76361846	5.37E-01	4.63E-01	cg08675585	76362511		x				
					cg08675585	76362486	x					
					cg24887211	76362378	x					
rs2284235	14	77233176	9.76E-01	6.70E-01	cg13507326	77295317	x	x	x	x	x	x
					cg20385229	77245223	x		x	x	x	x
					cg21787280	77295304	x					
rs2402074	14	91363271	5.21E-01	7.04E-01	cg25729068	91373013			x	x	x	x
rs8007280	14	93452715	5.36E-01	8.17E-01	cg20022541	93455173		x	x	x	x	x
rs2602141	15	41511938	<u>4.77E-02</u>	3.73E-01	cg00410921	41672430	x					
					cg03679734	41672382	x					
					cg12784172	41449224	x	x				
rs6494519	15	63455148	6.99E-01	3.04E-01	cg16003238	63457929	x					
					cg25172835	63458780	x					
rs3866543	15	74416664	7.60E-02	3.68E-01	cg02831604	74415805		x	x			x
rs770144	15	77887772	5.90E-01	6.41E-01	cg02196730	77975807			x	x	x	x
					cg23855392	77976862			x	x	x	x
rs4786500	16	4465266	7.37E-01	3.19E-01	cg14951292	4466012	x	x	x	x	x	x
rs8062405	16	28745407			Genotyping problems	cg00000292	28797626	x	x	x	x	x
						cg00931491	28515789	x				
						cg05590982	28457672	x				
						cg15149645	28458120	x				
rs3888190	16	28796987	1.43E-01	<u>2.93E-02</u>	cg00000292	28797626		x	x	x	x	x
rs747973	16	29782742	9.67E-01	6.03E-01	cg24235633	29782653		x	x	x	x	x
rs4783962	16	55552539	7.19E-01	1.53E-01	cg12564453	55553341	x					
					cg26916607	55552296	x	x				
rs7200919	16	65874101	1.89E-01	2.79E-01	cg19514469	65790958		x				
					cg25341653	65790803		x				
rs10492813	16	70679964	6.48E-01	6.89E-01	cg23815491	70646148	x	x	x	x	x	x

rs9899621	17	36728272	8.62E-01	3.23E-01	cg11710560	36642695		x				
rs11079035	17	37542538	9.51E-01	6.73E-01	cg02968557	37528627	x	x				
					cg05346140	37528558	x	x				
rs11079772	17	42748738	1.07E-01	7.21E-02	cg08085267	42756807	x	x	x	x	x	
rs606911	17	44263227	4.14E-01	1.94E-01	cg08137716	44226929	x	x				
rs1005376	17	46200634	5.53E-02	<u>2.64E-02</u>	cg18490846	46200746				x		x
rs6504178	17	59166465	4.08E-01	2.70E-01	cg02988947	59132545	x					
					cg06873352	59173722	x	x	x	x	x	
					cg07117700	59257059	x					
					cg20131968	59203744	x					
					cg24498554	59327969	x					
rs9891018	17	59181588	<u>4.53E-02</u>	<u>4.84E-02</u>	cg02988947	59132570	x	x				
					cg06873352	59173747	x					
					cg24498554	59327969	x					
rs3809700	17	72094821	3.51E-01	9.30E-01	cg08666623	72094303	x	x	x			x
rs4789817	17	78201278	6.57E-01	8.77E-01	cg10281770	78200261	x	x	x	x	x	
rs4519393	18	13905410	7.39E-01	8.01E-01	cg00335286	13906330	x	x	x	x	x	
rs8094212	18	27288302	9.22E-01	7.31E-01	cg14871544	27281858	x	x	x	x		
rs8092040	18	75890106			Genotyping problems	cg04727522	75894170	x		x	x	x
						cg07851192	75725501	x				
rs3744872	18	75906844	3.53E-01	4.19E-01	cg04727522	75894195		x	x	x	x	x
rs2301742	19	626513	8.50E-01	5.11E-01	cg10503234	627739	x					
					cg23260026	626390	x					
rs8104762	19	701173			Genotyping problems	cg17376813	701049	x	x	x	x	x
rs873289	19	1126396	8.44E-01	9.09E-01	cg14401897	1125642		x	x	x	x	x
rs4804802	19	7713610	8.14E-01	3.22E-01	cg01532771	7733560	x					
					cg01618851	7718265	x					
					cg03221619	7673348	x					
					cg04816348	7702382	x					
					cg12387247	7672974	x					
					cg21372914	7734144	x					
rs371671	19	13731153	8.10E-01	8.04E-01	cg16474696	13735989		x	x	x	x	x
rs3810427	19	15907650	8.95E-01	9.08E-01	cg03190825	15906788	x					
					cg24655310	15906094	x	x	x	x	x	x
rs4806171	19	40714481	8.75E-01	5.64E-01	cg25681177	40716178	x	x	x	x	x	
rs4806187	19	40841601	2.58E-01	3.40E-01	cg22802439	40849640			x	x	x	x
rs7255146	19	46223966	3.03E-01	4.11E-01	cg25427638	46081222		x	x	x	x	x
rs7253209	19	47755334	2.45E-01	4.56E-01	cg10237469	46825249	x					
					cg11500727	47586265	x					
					cg12973651	47586293	x					
					cg25839766	48076917	x					
rs7246342	19	55115538	2.52E-01	3.99E-01	cg04315771	55125761	x					
					cg23878206	55123262	x	x				
rs12459644	19	56731658	9.40E-01	4.58E-01	cg16293105	56726807		x	x	x	x	x
rs11672222	19	56919728	4.09E-01	2.03E-01	cg13300756	56919850	x					
					cg14141399	56919860	x					
rs10421319	19	60903743	5.29E-01	6.40E-01	cg26209676	60846219		x	x	x		
rs159872	19	63239261	9.66E-01	8.39E-01	cg07942995	63238144	x					
rs6051669	20	3164203	1.07E-01	2.00E-01	cg08284598	3166120	x					
rs6020251	20	35756728	<u>2.44E-02</u>	<u>1.19E-02</u>	cg17974185	35756779		x	x	x	x	x
rs386155	20	43892590	2.40E-01	<u>3.89E-03</u>	cg00690280	43766147	x					
					cg08201421	43889558	x					
					cg09874776	43764255	x					
rs11702277	21	18008191	2.78E-01	8.66E-02	cg01284306	18112475		x				
rs2243521	21	29353601	<u>3.48E-02</u>	1.09E-01	cg05406101	29314115	x	x	x	x	x	
rs8127496	21	37364352	6.83E-01	8.04E-01	cg21832243	37366118			x	x	x	
rs1571737	21	42403740	<u>2.94E-02</u>	1.03E-01	cg19766460	42401249	x	x	x	x	x	
rs3810597	22	17660403	7.83E-01	5.21E-01	cg07251788	17660190	x	x	x	x	x	
rs1002286	22	22587337	2.20E-01	1.10E-01	cg04394102	22651838	x	x	x	x	x	
					cg24330042	22714159	x					
rs713835	22	36394596	4.86E-01	7.67E-01	cg00292662	36401089		x	x			x
					cg19853760	36401648		x				
rs5756813	22	36505423	8.30E-01	3.84E-01	cg25404088	36473089	x	x				
rs7291782	22	38745563	3.75E-01	7.80E-01	cg21771250	38735970	x	x	x	x	x	x

ASM: Allele-Specific Methylation; SNP and CpG positions: NCBI36/hg18 assembly; IlmnID: CpG site name; #: ASM from Zhang *et al.*, 2010 study. *: ASM from Gibbs *et al.*, 2010; Crbl: Cerebellum; Tctx: Temporal cortex; Fctx: Frontal cortex; Underlined p-values: significant p-values (p < 0.05); HWE: Hardy-Weinberg equilibrium.

Supplementary Table 3. SNPs with nominal associations in the discovery sample under the additive model.

SNP	Chr	Position	Alleles (1/2)	Discovery						Additive model	
				Controls N=656 (%)			Cases N=578 (%)			p-value	Adj. p-value
				11	12	22	11	12	22		
rs11585570	1	53,438,130	T/C	276 (42.1)	321 (48.9)	59 (9.0)	233 (40.4)	246 (42.6)	98 (17)	1.22E-02	1.60E-02
rs823080	1	205,789,282	G/A	233 (35.8)	318 (48.8)	100 (15.4)	178 (31)	282 (49.1)	114 (19.9)	1.98E-02	3.15E-02
rs3766612	1	210,020,013	G/T	391 (59.7)	232 (35.4)	32 (4.9)	403 (69.8)	156 (27)	18 (3.1)	2.10E-04	4.00E-04
rs9891018	17	61,827,856	T/C	355 (54.1)	255 (38.9)	46 (7.0)	343 (59.5)	202 (35.1)	31 (5.4)	4.53E-02	4.84E-02
rs6020251	20	36,323,314	G/A	409 (62.4)	226 (34.5)	20 (3.1)	332 (57.5)	213 (36.9)	32 (5.5)	2.44E-02	1.19E-02

SNP: Single Nucleotide Polymorphism; Chr: Chromosome; Position: build hg19; Adj. p-value: p-value adjusted by age; All alleles are reported in the forward strand.

Supplementary Table 4. Drug dependence case-control study in a sample of Caucasian Spanish subjects. Four different genetic models are considered: additive, dominant, codominant and recessive.

Discovery									
SNP	Additive model		Dominant model		Codominant model		Recessive model		
	p-value	Adj. p-value	p-value	Adj. p-value	p-value	Adj. p-value	p-value	Adj. p-value	
rs823080	<u>1.98E-02</u>	<u>3.15E-02</u>	7.67E-02	5.72E-02	6.03E-02	9.82E-02	<u>3.87E-02</u>	1.14E-01	
rs3766612	<u>2.10E-04</u>	<u>4.00E-04</u>	<u>2.00E-04</u>	<u>2.40E-04</u>	<u>8.30E-04</u>	<u>1.14E-03</u>	1.14E-01	2.28E-01	
rs6020251	<u>2.44E-02</u>	<u>1.19E-02</u>	7.94E-02	<u>4.38E-02</u>	<u>4.39E-02</u>	<u>2.38E-02</u>	<u>2.98E-02</u>	<u>2.15E-02</u>	
rs9891018	<u>4.53E-02</u>	<u>4.84E-02</u>	5.78E-02	<u>3.22E-02</u>	1.34E-01	1.01E-01	2.39E-01	5.55E-01	
rs11585570	<u>1.22E-02</u>	<u>1.60E-02</u>	5.47E-01	3.83E-01	<u>1.10E-04</u>	<u>1.26E-03</u>	<u>3.00E-05</u>	<u>2.70E-04</u>	

Replication									
SNP	Additive model		Dominant model		Codominant model		Recessive model		
	p-value	Adj. p-value	p-value	Adj. p-value	p-value	Adj. p-value	p-value	Adj. p-value	
rs823080	7.11E-01	8.84E-01	6.31E-01	8.37E-01	8.88E-01	9.77E-01	9.30E-01	9.89E-01	
rs3766612	7.16E-01	7.02E-01	6.33E-01	6.38E-01	8.63E-01	8.81E-01	9.09E-01	9.66E-01	
rs6020251	5.17E-01	4.70E-01	9.12E-01	9.50E-01	1.10E-01	9.30E-02	<u>4.79E-02</u>	<u>3.86E-02</u>	
rs9891018	5.74E-01	6.45E-01	9.86E-01	9.72E-01	3.02E-01	5.22E-01	1.39E-01	2.71E-01	
rs11585570	9.43E-02	1.33E-01	7.93E-02	1.03E-01	2.08E-01	2.63E-01	4.33E-01	5.35E-01	

Pooled analysis									
SNP	Additive model		Dominant model		Codominant model		Recessive model		
	p-value	Adj. p-value	p-value	Adj. p-value	p-value	Adj. p-value	p-value	Adj. p-value	
rs823080	9.59E-02	1.88E-01	1.60E-01	2.43E-01	2.50E-01	4.19E-01	1.88E-01	3.31E-01	
rs3766612	5.61E-02	7.94E-02	6.64E-02	7.89E-02	1.60E-01	2.05E-01	3.05E-01	4.40E-01	
rs6020251	6.28E-02	<u>4.48E-02</u>	3.37E-01	2.84E-01	<u>1.46E-02</u>	<u>9.49E-03</u>	<u>3.70E-03</u>	<u>2.35E-03</u>	
rs9891018	4.56E-01	4.39E-01	2.56E-01	2.51E-01	3.76E-01	3.84E-01	6.50E-01	6.87E-01	
rs11585570	<u>5.39E-03</u>	<u>1.08E-02</u>	8.73E-02	8.54E-02	<u>5.06E-03</u>	<u>1.88E-02</u>	<u>1.54E-03</u>	<u>7.65E-03</u>	

SNP: Single Nucleotide Polymorphism; Adj. p-value: p-value adjusted by age; Underlined p-values: nominally significant p-values

Supplementary Table 5. Histone marks for rs6020251 correlating with differential methylation of cg17974185.

Brain region	Chromatin states model		Enhancer histone marks		Promoter histone marks	
	Core 15-state	Core 25-state	H3K4me1	H3K27ac	H3K4me3	H3K9ac
	Hippocampus Middle	TssAFlnk	PromD1	H3K4me1	H3K27ac	H3K4me3
Substantia Nigra	TssAFlnk	PromD1	H3K4me1		H3K4me3	H3K9ac
Anterior Caudate	TssAFlnk	PromD1	H3K4me1	H3K27ac	H3K4me3	H3K9ac
Cingulate Gyrus	TssAFlnk	PromD1	H3K4me1	H3K27ac	H3K4me3	H3K9ac
Inferior Temporal Lobe	TssAFlnk	PromD1	H3K4me1	H3K27ac	H3K4me3	H3K9ac
Angular Gyrus	TssAFlnk	PromD1	H3K4me1		H3K4me3	H3K9ac
Dorsolateral Prefrontal Cortex	TssAFlnk	PromD1	H3K4me1	H3K27ac	H3K4me3	H3K9ac
Germinal Matrix	TssAFlnk	PromD1	H3K4me1		H3K4me3	
Fetal Female	TssAFlnk	PromD1	H3K4me1		H3K4me3	
Fetal Male	TssAFlnk	PromD2	H3K4me1		H3K4me3	

H3K4me1 and H3K27ac: Histone marks related to enhancer regions; H3K4me3 and H3K9ac: Histone marks related to promoter regions; TssAFlnk: Flanking Active TSS; PromD1: Promoter Downstream TSS 1; PromD2: Promoter Downstream TSS 2; Gray cells: experiments from Roadmap are not available.

Supplementary Table 6. Transcription Factor Binding Sites (TFBS) motif scores predictions for A) rs11585570 and B) rs6020251

Transcription Factor	Alleles (1/2)	Strand of the TFBS	Allele 1			Allele 2				
			TFBS score*	p-value	Qvalue	TFBS motif	TFBS score*	p-value	Qvalue	TFBS motif
A) HOXA5	T/C	-	2.28	8.12E-03	0.227	CAATAATG	10.04	1.69E-04	0.00474	CAGTAATG
B) NKX3-2	G/A	+	2.42	5.99E-03	0.16	CCCGCTTTG	9.02	3.08E-04	0.01	CCCACTTTG
NKX3-1	G/A	+	0.27	6.40E-03	0.17	CCCGCTTTG	9.05	2.88E-04	0.01	CCCACTTTG
ZNF354C	G/A	+	NA	NA	NA	NA	8.08	1.22E-03	0.04	CCCCAC
ISL2	G/A	+	NA	NA	NA	NA	8.20	5.08E-04	0.01	CCCACTTTG

*The TFBS score is an estimation of fitness between TFBS motif position weight matrix (PWM) and target sequence by using log-likelihood calculation method. Higher score indicates a higher similarity of target sequence to TFBS motif PWM. All alleles are reported in the forward strand; Risk allele underlined; In bold: Allele 1 or 2 in the TFBS motif.

Supplementary Table 7. SNPs in high LD ($r^2 > 0.85$) with rs6020251 and rs11585570 influencing ASM of cg13461509, cg27535305 and cg17974185.

CpG site (position)	SNP	SNP position	Distance (bp)	Tissue	pvalue	R ²
cg13461509 (chr1:53387576)	rs11206043	chr1:53175140	14976	crblm	2.19E-10	0.3174
				tctx	1.03E-11	0.3104
				fctx	6.36E-11	0.2791
	rs1775540	chr1:53180481	20317	crblm	6.39E-10	0.3037
				tctx	2.76E-11	0.2996
				fctx	2.96E-10	0.2623
	rs1672913	chr1:53180914	20750	crblm	6.39E-10	0.3037
				tctx	2.76E-11	0.2996
				fctx	2.96E-10	0.2623
	rs576891	chr1:53137564	22600	tctx	1.83E-11	0.3041
				crblm	6.96E-10	0.3026
				fctx	1.14E-10	0.2728
	rs511599	chr1:53135406	24758	tctx	1.83E-11	0.3041
				crblm	6.96E-10	0.3026
				fctx	1.14E-10	0.2728
	rs567880	chr1:53121152	39012	tctx	1.83E-11	0.3041
				crblm	6.96E-10	0.3026
				fctx	1.14E-10	0.2728
	rs481440	chr1:53117564	42600	tctx	1.83E-11	0.3041
				crblm	6.96E-10	0.3026
				fctx	1.14E-10	0.2728
	rs11585570	chr1:53210718	50554	crblm	1.91E-10	0.3191
				tctx	2.97E-11	0.2988
				fctx	2.10E-10	0.2661
rs514881	chr1:53109325	50839	crblm	8.37E-10	0.3002	
			fctx	9.70E-11	0.2745	
			tctx	3.53E-08	0.2166	
rs480299	chr1:53092862	67302	crblm	8.37E-10	0.3002	
			fctx	9.70E-11	0.2745	
			tctx	3.53E-08	0.2166	
rs11206043	chr1:53175140	9901	tctx	4.59E-32	0.6722	
			fctx	1.40E-25	0.5671	
			pons	2.63E-19	0.4824	
rs1775540	chr1:53180481	15242	tctx	1.17E-31	0.6673	
			fctx	1.10E-25	0.5687	
			pons	2.39E-19	0.4832	
rs1672913	chr1:53180914	15675	tctx	1.17E-31	0.6673	
			fctx	1.10E-25	0.5687	
			pons	2.39E-19	0.4832	
rs576891	chr1:53137564	27675	tctx	5.47E-31	0.659	
			fctx	4.50E-25	0.5594	
			pons	1.26E-17	0.4491	
rs511599	chr1:53135406	29833	tctx	5.47E-31	0.659	
			fctx	4.50E-25	0.5594	
			pons	1.26E-17	0.4491	
cg27535305 (chr1:53392651)	rs567880	chr1:53121152	44087	tctx	5.47E-31	0.659
				fctx	4.50E-25	0.5594

			pons	1.26E-17	0.4491	
			tctx	6.04E-33	0.6826	
rs11585570	chr1:53210718	45479	fctx	2.08E-26	0.5795	
			pons	1.41E-20	0.5063	
			tctx	5.47E-31	0.659	
rs481440	chr1:53117564	47675	fctx	4.50E-25	0.5594	
			pons	1.26E-17	0.4491	
			tctx	1.38E-23	0.5527	
rs514881	chr1:53109325	55914	fctx	1.59E-21	0.5012	
			pons	5.98E-16	0.4138	
			tctx	1.38E-23	0.5527	
rs480299	chr1:53092862	72377	fctx	1.59E-21	0.5012	
			pons	5.98E-16	0.4138	
			pons	2.55E-28	0.6299	
rs6020251	chr20:36323314	26	crblm	1.29E-24	0.6298	
			fctx	6.22E-29	0.615	
			tctx	2.85E-27	0.6092	
			crblm	2.02E-20	0.5564	
rs4811030	chr20:36347834	24494	pons	6.36E-22	0.5304	
			fctx	1.17E-22	0.5206	
			tctx	1.63E-19	0.4809	
			pons	1.66E-27	0.6185	
rs6096026	chr20:36375397	52057	crblm	1.32E-23	0.6133	
			fctx	9.31E-27	0.5846	
			tctx	6.47E-24	0.5581	
			crblm	7.15E-22	0.5833	
cg17974185 (chr20:36323340)	rs6063608	chr20:36430407	107067	pons	4.35E-25	0.5825
			fctx	4.02E-20	0.4762	
			tctx	2.09E-18	0.4595	
			pons	3.53E-27	0.6138	
rs2281148	chr20:36433288	109948	crblm	6.20E-22	0.5844	
			tctx	3.20E-21	0.5123	
			fctx	4.61E-22	0.5105	
			pons	2.95E-25	0.5852	
rs6063714	chr20:36492927	169587	crblm	1.50E-19	0.5394	
			tctx	2.14E-19	0.4786	
			fctx	7.08E-20	0.4717	
			pons	3.32E-19	0.4805	
rs6068029	chr20:36511453	188113	fctx	3.06E-17	0.4209	
			tctx	1.61E-15	0.3994	
			crblm	6.01E-13	0.3882	

All the information presented in this table is described in Zhang et al., 2010 and Gibbs et al., 2010; SNP: Single Nucleotide Polymorphism; Distance (bp): Distance between the SNP and the CpG site; Positions in build hg19; tctx: Temporal cortex; fctx: Frontal cortex; crblm: cerebellum; In bold: SNP reported in our study.

Article 4

“Associació del gen *PLCB1* amb la dependència de drogues”

Resum:

Els factors genètics implicats en la susceptibilitat a la dependència de drogues són avui dia encara força desconeguts. Diversos estudis donen suport al paper dels microRNAs (miRNAs) en la plasticitat neuronal induïda per les drogues que dona lloc a l'addicció. En aquest treball hem investigat el paper dels miRNAs en l'establiment de la dependència de drogues. Amb aquest objectiu hem seleccionat 62 variants d'un sol nucleòtid (SNPs) situades a la regió 3'UTR de gens diana i que potencialment podrien alterar la unió de molècules de miRNAs. Hem realitzat un estudi cas-control en un mostra espanyola de 735 casos (majoritàriament individus dependents de cocaïna que a la vegada són dependents d'altres drogues d'abús) i 739 controls, i hem identificat associació amb la variant rs1047383 del gen *PLCB1*, que s'ha replicat en una mostra independent (663 casos i 667 controls). A continuació hem testat 9 miRNAs que potencialment s'unirien a la regió del SNP rs1047383, però no s'ha pogut demostrar que cap d'ells alteri els nivells d'expressió del gen *PLCB1*. A més, hem testat dos miRNAs que podrien unir-se a una altra regió del gen *PLCB1* que conté un SNP, rs708019, en elevat desequilibri de lligament amb la variant associada. Tot i que el miRNA miR-582 podia unir-se a aquesta regió tot disminuint l'expressió del gen *PLCB1*, no hem observat diferències entre els dos al·lels del SNP. Finalment, hem explorat la possibilitat que la cocaïna alteri l'expressió del gen *PLCB1*. Tant en nucli accumbens de subjectes que abusen de la cocaïna com en neurones dopaminèrgiques humanes en cultiu tractades amb cocaïna s'ha detectat un increment de l'expressió del gen. Aquests resultats, conjuntament amb d'altres publicats en estudis previs, suggereixen que el gen *PLCB1* podria participar en la susceptibilitat a la dependència de drogues en general, i a la cocaïna en particular.

Referència:

Cabana-Domínguez J, Roncero C, Pineda-Cirera L, R. Palma-Álvarez F, Ros-Cucurull E, Grau-López L, Esojo A, Casas M, Arenas C, Ramos-Quiroga JA, Ribasés M, Fernández-Castillo N, Cormand B. *Association of the PLCB1 gene with drug dependence*. Scientific Reports 2017; 7:10110. doi: 10.1038/s41598-017-10207-2.

SCIENTIFIC REPORTS

OPEN

Association of the *PLCB1* gene with drug dependence

Judit Cabana-Domínguez^{1,2,3,4}, Carlos Roncero^{5,6,7,8}, Laura Pineda-Cirera^{1,2,3,4}, R. Felipe Palma-Álvarez^{5,8}, Elena Ros-Cucurull^{5,6,8}, Lara Grau-López^{5,6,7,8}, Abderaman Esojo^{6,8}, Miquel Casas^{5,7,8}, Concepció Arenas¹, Josep Antoni Ramos-Quiroga^{5,7,8,9}, Marta Ribasés^{7,8,9}, Noèlia Fernández-Castillo^{1,2,3,4} & Bru Cormand^{1,2,3,4}

1d: 4 May 2017

2d: 4 August 2017

3d online: 31 August 2017

Genetic factors involved in the susceptibility to drug addiction still remain largely unknown. MiRNAs seem to play key roles in the drug-induced plasticity of the brain that likely drives the emergence of addiction. In this work we explored the role of miRNAs in drug addiction. With this aim, we selected 62 SNPs located in the 3'UTR of target genes that are predicted to alter the binding of miRNA molecules and performed a case-control association study in a Spanish sample of 735 cases (mainly cocaine-dependent subjects with multiple drug dependencies) and 739 controls. We found an association between rs1047383 in the *PLCB1* gene and drug dependence that was replicated in an independent sample (663 cases and 667 controls). Then we selected 9 miRNAs predicted to bind the rs1047383 region, but none of them showed any effect on *PLCB1* expression. We also assessed two miRNAs binding a region that contains a SNP in linkage disequilibrium with rs1047383, but although one of them, hsa-miR-582, was found to downregulate *PLCB1*, no differences were observed between alleles. Finally, we explored the possibility that *PLCB1* expression is altered by cocaine and we observed a significant upregulation of the gene in the nucleus accumbens of cocaine abusers and in human dopaminergic-like neurons after cocaine treatment. Our results, together with previous studies, suggest that *PLCB1* participates in the susceptibility to drug dependence.

Drug dependence is one of the major health problems worldwide. In Europe, about 25% of adults are estimated to have tried illicit drugs at some point in their lives¹. Usually drug consumers use more than one drug at the same time: for example, within the group of European individuals who consumed a psychoactive substance in the last 12 months, 33% had consumed two different substances and 10% had used three². This high prevalence of poly-drug abuse is due to common and drug-specific genetic and environmental factors^{3–5}. It is well known that addictions are moderately to highly heritable (from 0.39 in the case of hallucinogens to 0.72 for cocaine), although the specific genetic risk factors involved in its predisposition remain largely unknown^{6–9}. Transcriptomic studies in animal and cellular models, as well as human studies in postmortem brain samples from addicted individuals, have revealed that both acute and chronic drug exposure produce epigenetic adaptations and changes in gene expression¹⁰. Furthermore, recent studies have shown that some genes whose expression is altered by cocaine also contribute to cocaine dependence susceptibility^{11,12}.

MicroRNAs (miRNAs) are small regulatory noncoding RNA molecules (about 18–25 nucleotides in length) that control gene expression through direct binding to 3'untranslated regions (3'UTRs) of target mRNAs causing translational repression or mRNA degradation. One single miRNA can target and regulate hundreds of mRNAs

¹Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain. ²Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain. ³Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Catalonia, Spain. ⁴Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Catalonia, Spain. ⁵Department of Psychiatry and Legal Medicine, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain. ⁶Addiction and Dual Diagnosis Unit Vall Hebron, Psychiatric Services, Hospital Universitari Vall d'Hebron-ASPB, Barcelona, Catalonia, Spain. ⁷Biomedical Network Research Centre on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain. ⁸Department of Psychiatry, Hospital Universitari Vall d'Hebron, Barcelona, Catalonia, Spain. ⁹Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain. Noèlia Fernández-Castillo and Bru Cormand contributed equally to this work. Correspondence and requests for materials should be addressed to N.F.-C. (email: noefernandez@ub.edu) or B.C. (email: bcormand@ub.edu)

Gender N (%)	Discovery sample		Replication sample	
	Drug dependence N = 735	Control N = 739	Drug dependence N = 663	Control N = 667
Male	574 (78.1)	577 (78.1)	507 (76.5)	510 (76.5)
Female	161 (21.9)	162 (21.9)	156 (23.5)	157 (23.5)
Age (mean and SD)				
	37.3 ± 9.6	55.4 ± 15.9	37.3 ± 9.9	55.8 ± 16.3

Table 1. Descriptive characteristics of the Spanish Caucasian individuals with drug dependence and controls used in the case-control association study. SD: Standard deviation.

and, conversely, one mRNA can be regulated by several miRNAs. This is a complex and dynamic system that allows the cells to fine-tune gene expression^{13–15}.

MiRNAs are very abundant in the central nervous system and play important roles in neuronal development, differentiation and survival^{16, 17}. Many studies have shown their contribution to several psychiatric disorders such as schizophrenia, bipolar disorder, autism or drug dependence^{18–21}. In human prefrontal cortex of alcoholic patients 35 miRNA were found up-regulated as compared to controls²². Animal model studies have demonstrated that drugs of abuse induce robust alterations in the expression of a wide range of miRNAs. Cocaine administration in rats alters miR-124, miR-181 and let-7 in mesolimbic dopaminergic system^{23, 24} and miR-212 in dorsal striatum^{25, 26}. On the other hand, alcohol regulates miR-9 increasing alcohol tolerance²⁷. Also, miRNAs have been shown to play an important role in different processes related to addiction such as reward, synaptic plasticity, learning, memory, withdrawal and relapse²⁸.

Some studies suggest that single nucleotide polymorphisms (SNPs) located in miRNAs or in their target sites can alter the miRNA-mediated regulation of gene expression that underlies disease and non-pathological phenotypes^{29–31}. A recent study generated a transcriptome-wide map of the miRNA binding sites in human brain. Based on the interaction between argonaute 2 protein (AGO2) and miRNAs, they identify target regions in mRNAs. These regions contain 916 common SNPs that could potentially alter miRNA:mRNA binding³².

In this study, we aimed at examining the contribution to drug dependence susceptibility of SNPs that alter the binding of miRNAs to their target mRNAs. For that purpose we selected SNPs located in the 3'UTR identified in the study mentioned above and performed a case-control association study in drug addiction in a discovery and a replication samples from Spain. The identified variants were subjected to functional testing. Finally, we assessed the impact of cocaine on the expression of those genes where the associated SNPs are located.

Material and Methods

Association study. *Subjects.* Patients were recruited and evaluated at the Addiction and Dual Diagnosis Unit of the Psychiatry Department of the Hospital Universitari Vall d'Hebron (Barcelona, Spain) according to DSM-IV-TR criteria (Diagnostic and Statistical Manual of Mental Disorders, 4th ed. Text revision). The Structured Clinical Interview (SCID)³³ was administered and volunteers with DSM-IV lifetime diagnosis for substance dependence were included in the study. About 73.5% of our patient sample consists of cocaine-dependent patients, most of which are dependent to other drugs of abuse. Controls were recruited at the Blood and Tissues Bank of Barcelona, and both patients and controls were Spanish and Caucasian, with the two last names (one from each parent) of Spanish origin. Other ethnicities such as Moroccan, Gypsies or South American individuals, among others, were discarded. Patients and controls were divided randomly into a discovery sample, which consisted of 735 patients and 739 controls, and a replication sample of 663 patients and 667 controls (Table 1 and Fig. 1). Population stratification was previously discarded in our sample³⁴. The study was approved by the ethics committee of our institution, the Institutional Review Board of the University of Barcelona (IRB00003099), and informed consent was obtained from all participants, in accordance with the Helsinki Declaration. All experiments were performed in accordance with relevant guidelines and regulations.

DNA isolation and quantification. Genomic DNA samples were obtained from peripheral blood lymphocytes using the salting-out method³⁵ and were quantified using Nanodrop ND-1000 Spectrophotometer (Nanodrop Technologies, Thermo Fisher Scientific Inc., Wilmington, DE, USA).

SNP selection and genotyping. We selected SNPs in target genes that might alter the binding of regulator miRNAs using a previously described list of 916 SNPs located in AGO2 binding sites³². From this list we selected SNPs within the 3'UTR of genes and with a minimum allele frequency (MAF) of 0.15 in CEU individuals from HapMap project (www.hapmap.org; release 23). A total of 62 SNPs were selected under these criteria and genotyped in the discovery sample. We considered a maximum percentage of 5% of missing genotypes, and the SNPs finally evaluated had an average call rate of 99.6%. Nominally significant associated SNPs were subsequently assessed in a replication sample. Genotyping of both discovery and replication samples was performed by KASP technology (LGC, Teddington, Middlesex, UK).

Statistical analyses. The minimal statistical power for discovery (23–93%) and replication samples (18–81%) were calculated *post hoc* using the software Power Calculator for Genetic Studies (<http://sph.umich.edu/csg/abecasis/CaTS/>), under the multiplicative model (equivalent to log-additive) and assuming an odds ratio (OR) between 1.1 and 1.3, a significance threshold of 0.05, the lowest MAF value in our study (0.119), and a prevalence for substance dependence of 0.026³⁶. The R library *SNPassoc*³⁷ was used to assess Hardy-Weinberg equilibrium

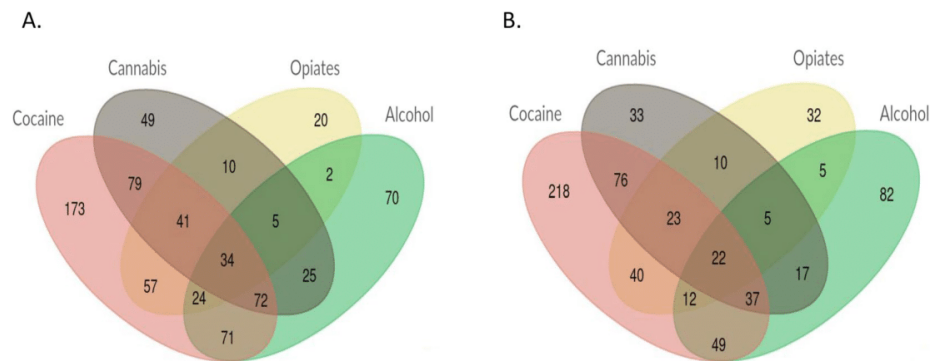


Figure 1. Distribution of the four main dependencies (cocaine, cannabis, alcohol and opiates) in the sample of patients included in the case-control association study, depicted in Venn diagrams. Other dependencies with a frequency lower than 10% are not displayed. **(A)** Discovery sample. **(B)** Replication sample.

Marker	Locus	Discovery (735 cases–739 controls)				Replication (663 cases–667 controls)				Pooled analysis (1393 cases–1406 controls)		
		Alleles	p-value ¹	Adj. p-value ²	OR [95% CI] ²	Alleles	p-value ¹	Adj. p-value ²	OR [95% CI] ²	p-value ¹	Adj. p-value ²	OR [95% CI] ²
rs6840	<i>SCD5</i>	<u>C</u> /T	0.011	0.025	1.24 [1.03–1.50]	—	—	—	—	—	—	—
rs1285	<i>ID11</i>	<u>C</u> /T	0.029	—	—	C/T	0.010	0.034	1.32 [1.02–1.71]	—	—	—
rs1872353	<i>FBXO45</i>	<u>C</u> /T	0.044	—	—	—	—	—	—	—	—	—
rs6855973	<i>GRIA2</i>	<u>A</u> /T	0.014	—	—	—	—	—	—	—	—	—
rs1047383	<i>PLCB1</i>	<u>C</u> /T	0.039	9.6e-03	1.27 [1.06–1.53]	<u>C</u> /T	3.8e-03	1.6e-03	1.37 [1.13–1.67]	4.8e-04	3.7e-04	1.29 [1.12–1.49]
rs1057377	<i>SPOCK3</i>	<u>A</u> /G	0.026	0.033	1.23 [1.02–1.49]*	—	—	—	—	—	—	—
rs2597775	<i>QDPR</i>	<u>C</u> /T	0.043	—	—	—	—	—	—	—	—	—

Table 2. SNPs located in miRNA binding sites associated with drug dependence. ¹Log-additive model; ²Adjusted by age; Risk allele underlined.*When OR < 1 the inverted score is shown; SNP, Single Nucleotide Polymorphism.

(HWE, threshold set at $P < 0.01$) and to compare genotypic frequencies between cases and controls for each marker considering the log-additive model and a significance threshold of 0.05. As age differed significantly between cases and controls, we considered it as a covariate in all tests. All the p-values shown in the different association analyses are the ones adjusted for age, except in Table 2 and Supplementary Tables where both p-values are shown. The Bonferroni correction threshold for multiple testing was set at $P < 8.5e-04$ (0.05/59 SNPs) in the discovery sample, and at $P < 7.2e-03$ (0.05/7 SNPs) in the replication sample and in the pooled analysis.

Evaluation of functional effect of associated variants. *Linkage disequilibrium analysis.* Genotype data for the *PLCB1* gene plus 10 kb flanking sequences upstream and downstream were available for 554 individuals from our control sample³⁸. The analysis of linkage disequilibrium (LD) was performed using Haploview software³⁹ setting a maximum r^2 threshold at 0.85.

Functional evaluation of SNPs effect on microRNA regulation. We assessed the possible functional effect of rs1047383 and two other variants found in LD with it (rs708910 and rs1047381), all located in the 3'UTR of the *PLCB1* gene. To do that, we used a luciferase reporter system to test the possible impact of these SNPs on the regulation of gene expression mediated by miRNAs, as previously described¹². The prediction tools FuncPred, mirWalk, mirSNP, mrSNP, mirdSNP, miRNASNP and RNAhybrid were used to select miRNAs which binding sites in *PLCB1* is potentially affected by these SNPs. We chose the best predictions for each SNP: for rs1047383, hsa-miR-124-1, hsa-miR-139, hsa-miR-140, hsa-miR-144, hsa-miR-377, hsa-miR-506, hsa-miR-548h, hsa-miR-1324 and hsa-miR-3148; and for rs708910, hsa-miR-582 and hsa-miR-140. All miRNAs were cloned into a pCMV-MIR vector (OriGene, Rockville, MD, USA) and expression was confirmed after transfection into HeLa or HEK293 cells by qRT-PCR using the miScript PCR System (Qiagen, Hilden, Germany). Two regions from the 3'UTR of the *PLCB1* gene were cloned in the pmirGLO Dual-Luciferase miRNA Target Expression Vector (Promega, Madison, WI, USA): one fragment of the 3'UTR of the *PLCB1* gene containing SNP rs1047383 (hg19/chr20:8,864,935–8,865,116; 182 bp) and another one containing SNPs rs708910 and rs1047381 (hg19/chr20:8,864,108–8,864,373; 267 bp). All the constructs were used to test the effect of the selected miRNAs in HeLa or in HEK293 cells. Luciferase expression was assessed using the Dual-luciferase Reporter Assay System (Promega). As our data did not follow a normal distribution (tested using Shapiro-Wilk test), differences between

the two conditions were evaluated with the nonparametric Mann-Whitney U-test using the SPSS statistics software version 22.0 (IBM, Armonk, NY, USA), and $P < 0.05$ was considered significant.

Effect of cocaine on *PLCB1* expression. To assess the possible impact of cocaine on *PLCB1* expression we used data available from previous studies. Data from nucleus accumbens samples of human cocaine abusers (10 cases and 10 controls matched by age, race, sex and brain pH) were kindly provided by the authors⁴⁰. When comparing differences between the expression levels of *PLCB1* between cases and controls, we considered the ratio between each case with its matched control. Normality of ratio was confirmed by using Shapiro-Wilk test ($P = 0.91$). Thus, the null hypothesis that the ratio is one, was tested with the parametric Student's t test, using the SPSS statistics software version 22.0 and considering $P < 0.05$ as significant. Furthermore, we also used RNA samples obtained in a previous study from our group on the effect of cocaine on gene expression¹². In this study we generated a human dopaminergic neuron-like model (differentiated SH-SY5Y cells) and RNA samples were obtained at different time points after an acute cocaine exposure (30 min). Samples were retrotranscribed using High-Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific). For the present study, we assessed *PLCB1* expression by quantitative Real-Time PCR (qRT-PCR) using LightCycler[®] 480 SYBR Green I Master (Roche Life Sciences, Branford, CT, USA), and relative quantification was performed as previously described, using *GAPDH* and *ACTB* as reference genes¹². As we only have three replicates for each condition, differences between conditions were evaluated with a t-student test⁴¹ using SPSS statistics software version 22.0.

Ethics statement. This study was approved by the local Ethics Committee and informed consent was obtained from all adult subjects, children and their parents according to the Helsinki declaration.

Results

In this study we evaluated the contribution to drug dependence predisposition of SNPs located in the 3'UTR of genes expressed in the brain that are predicted to alter the binding of miRNA molecules.

We performed a case-control association study in a sample of 735 drug-dependent patients and 739 sex-matched controls from Spain. A total of 59 SNPs in 56 genes were finally evaluated (from the 62 SNPs that were genotyped, one showed poor genotyping and two were not in HWE). The comparison of genotype frequencies between cases and controls under the log-additive model showed nominally significant differences before adjusting by age for seven SNPs located in the genes *SCD5*, *ID11*, *FBXO45*, *GRIA2*, *PLCB1*, *SPOCK3* and *QDPR* (Table 2 and Supplementary Table S1). These associated SNPs were subsequently evaluated in an independent Spanish sample of 663 drug-dependent patients and 667 sex-matched controls and the association remained significant for rs1285 in the *ID11* gene ($P = 0.034$; OR = 1.32, CI = [1.02–1.71]) and rs1047383 in the *PLCB1* gene ($P = 1.6e-03$; OR = 1.37, CI = [1.13–1.67]). In the first one the direction of the effect was not the same in the discovery and replication samples, but it was so for the rs1047383 association, which also survived the Bonferroni correction (Table 2 and Supplementary Table S2). In the pooled analysis of both the discovery and replication samples, only rs1047383 in the *PLCB1* gene remained associated with the disorder ($P = 3.7e-04$; OR = 1.29, CI = [1.12–1.49], Table 2 and Supplementary Table S3), with a higher frequency of subjects carrying the C allele in the group of drug-dependent subjects (39%) as compared to controls (35%). This association withstood the Bonferroni correction for multiple testing. Finally, we also explored these results in the subgroup of cocaine-dependent patients, including patients where cocaine is one of multiple drugs of abuse or the only one (about 75% of our patients' sample), and the association remained significant both in the discovery sample (551 cases; $P = 0.03$; OR = 1.25, CI = [1.02–1.53]) and in the replication sample (478 cases; $P = 0.01$; OR = 1.35, CI = [1.05–1.73]), and also in the pooled analysis (1,029 cases; $P = 5.7e-03$; OR = 1.26, CI = [1.07–1.48]).

We selected nine different miRNAs that potentially target the 3'UTR of *PLCB1*, with their binding being predicted to be affected by variation in rs1047383. No regulatory effect on *PLCB1* expression was observed for any of the assayed miRNAs using a gene reporter system in HeLa cells. We then investigated the existence of SNPs in LD with rs1047383 and found two polymorphisms, rs6056229 and rs708910 ($r^2 = 1$ and 0.88, respectively). The first one is located outside the *PLCB1* gene and we could not find any functional prediction for it. The latter is located within the 3'UTR of *PLCB1* and is predicted to alter a binding site for several miRNAs by the FuncPred software. This SNP is only 25 bp distant from another SNP, rs1047381, which is in moderate LD with the associated variant in our Spanish sample ($r^2 = 0.61$). To test this, we used a construct that includes both SNPs and assessed the possible effect of the two most frequent haplotypes (rs1047381C-rs708910G and rs1047381T-rs708910A) on the regulation mediated by hsa-miR-140 and hsa-miR-582. No differences were observed in HeLa cells, but in this cell line hsa-miR-582 could not be successfully overexpressed upon transfection. For this reason we repeated the experiment in HEK293 and observed that hsa-miR-582 decreases *PLCB1* expression (25%, $P = 3.96e-03$, Supplementary Figure 2), but without differences between the two rs1047381-rs708910 haplotypes.

Previous studies had reported that two genes (*NFAT5* and *NTNG1*) with altered expression under cocaine also bear variants that confer susceptibility to cocaine dependence as shown by association studies^{11,12}. Since we have shown that *PLCB1* is involved in the vulnerability to cocaine dependence, we explored the possibility that its expression is altered by the action of cocaine. For that purpose we used data from a previous study using samples from nucleus accumbens of human cocaine abusers⁴⁰. The sample's mean expression levels of *PLCB1* gene were increased by 1.19-fold ($P = 0.012$) when we compared each case with its matched control. Furthermore, we investigated the effect of cocaine on the regulation of *PLCB1* in a human neuron-like dopaminergic model (differentiated SH-SY5Y cells) using samples previously produced in our lab¹². We have now performed qRT-PCR experiments in this model and, interestingly, a significant upregulation of *PLCB1* after cocaine treatment was observed, showing a maximum of 1.65-fold increase ($P = 5.8e-03$) as compared to untreated cells (Fig. 2).

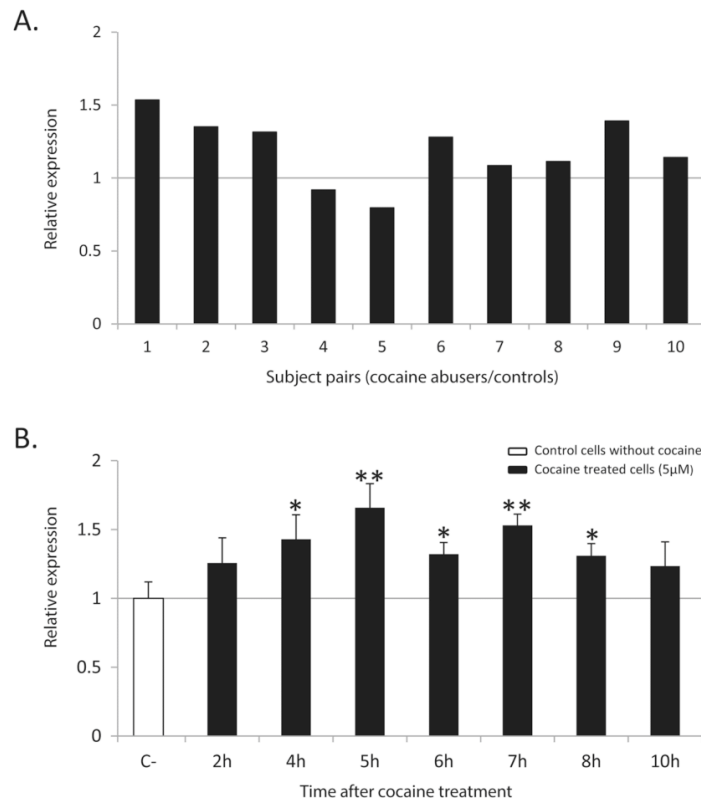


Figure 2. Expression of the *PLCB1* gene after treatment with cocaine. **(A)** Expression levels in the nucleus accumbens of cocaine abusers compared with their matched controls. **(B)** Transcription levels in human dopaminergic neuron-like cells (differentiated SH-SY5Y) at different time points after a 30-min exposure to 5 μM cocaine. Significant differences compared to control cells (not exposed to cocaine) normalized to *GAPDH* are indicated. Error bars indicate s.d. **p*-value < 0.05, ***p*-value < 0.01.

Discussion

In this study we explored, for the first time, the possible role in drug dependence of SNPs located in the 3'UTR of genes potentially altering the binding of the corresponding mRNA to miRNA molecules. These SNPs were selected on the basis of a previous study³² which listed variants located in binding sites for the AGO2 protein, one of the molecules of the RNA-induced silencing complex (RISC) that interacts with both the miRNA and the mRNA. We found that the rs1047383 variant in the *PLCB1* gene is associated with drug dependence in two independent samples. Then we investigated a possible effect of this SNP and rs1047381-rs708910 (in LD with rs1047383) on miRNA binding, with negative results for all the tested miRNAs (a total of 10 out of 35 predictions identified in the databases we used). The only exception was hsa-miR-582, which reduced gene expression but with no differences between the two rs1047381-rs708910 haploalleles. It is important to note that, although many prediction tools are available, the degree of overlap of the different outputs is often limited¹⁴, so ranking the predictions is not straightforward.

The *PLCB1* gene encodes the Phospholipase C beta 1 protein expressed in the brain, mainly in cortex, hippocampus and amygdala. It is considered a molecular mediator of synaptic plasticity and it plays an important role in modulating cognitive behavior and emotions^{42,43}. Many neurotransmitters such as dopamine, serotonin and glutamate activate PLCβ1 by a G-protein-coupled receptor that signals through G_{q/11}⁴⁴⁻⁴⁷. Furthermore, *PLCB1* has previously been related to other psychiatric and neurological disorders such as schizophrenia, autism and epilepsy⁴⁸⁻⁵².

Several lines of evidence support a role for *PLCB1* in drug dependence. A region of overlapping clusters of SNPs in the *PLCB1* gene were identified in a previous study that assessed common genomic regions in two GWAS of illegal substance dependence and cocaine dependence⁵³. However, in another GWAS reported by Gelernter *et al.* in 2014⁵⁴ none of the *PLCB1* SNPs showed a suggestive association ($P < 1e-05$) with cocaine dependence. These discordant results might be explained by differences in the case-control designs, as the controls used in their study⁵⁴ were individuals not dependent to cocaine who had taken this drug at least once in their lives and, in our study, the controls were individuals from the general population. Also, increased *Plcb1* expression was

previously described in the nucleus accumbens of mice after administration of cocaine during 7 days and also during withdrawal⁵⁵. This is consistent with our results, in which we identified increased expression in both the nucleus accumbens of human cocaine abusers and in cultured dopaminergic-like human neurons treated with cocaine. Furthermore, a previous study of our group identified changes in the expression of *PLCB1* in a mouse model of frustrated expected reward⁵⁶.

In conclusion, although we could not prove that the SNP found associated with the phenotype alters miRNA-mediated regulation of gene expression, our data provide evidence for the contribution of the *PLCB1* gene to cocaine dependence, identifying an associated variant that was replicated in a second sample, as well as alterations in the expression of *PLCB1* induced by cocaine.

Several strengths and limitations of the present study should be discussed. Strengths: i) To minimize sample heterogeneity, both patients and controls in the discovery and replication samples were recruited in the same small geographical area around Barcelona (Spain); cases and controls were Spanish, Caucasian and sex-matched; ii) cases were evaluated by following a unique clinical assessment; iii) the associated variant was replicated in a second independent sample; iv) altered expression of *PLCB1* induced by cocaine was seen in human brain post-mortem samples and also in a human neuron-like model, with the same direction; v) previous studies are in agreement with our findings at *PLCB1* both in the association and expression studies. Limitations: i) The disease phenotype was not excluded in the control samples, which may have potentially diluted positive findings in the association study; ii) the sample size, 1398 cases and 1406 controls, is relatively small and may have led to false negative results; iii) the associated variant could not be proven to have a functional effect on the binding of any of the tested miRNAs, although it could have an effect on other non-investigated miRNAs; iv) changes in *PLCB1* expression levels could be proven at mRNA level but we could not test the protein as samples were not available for this type of analysis.

In conclusion, a variant in the *PLCB1* gene was found associated with drug dependence in two independent samples. Although the sample size is altogether relatively small, the results obtained in the discovery sample have been replicated. Also, the expression of *PLCB1* was found to be altered by cocaine. Together with previous results, this study highlights *PLCB1* as a gene that may contribute to drug dependence. Finally, these findings and previous results from our and other groups on *NFAT5* and *NTNG1*^{11,12} suggest that genes which expression is altered by the effect of drugs of abuse may play an important role in the susceptibility to drug dependence.

References

1. European Monitoring Centre for Drugs and Drug Addiction(EMCDDA). *European Drug Report. Trends and Developments* (2016).
2. European Monitoring Centre for Drugs and Drug Addiction(EMCDDA). *National report to the EMCDDA by the Reitox National Focal Point 2014. Health (San Francisco)* (2014).
3. Palmer, R. H. C. *et al.* Genetic Etiology of the Common Liability to Drug Dependence: Evidence of Common and Specific Mechanisms for DSM-IV Dependence Symptoms. *Drug Alcohol Depend* **123**, S24–S32 (2012).
4. Xian, H. *et al.* Genetic and environmental contributions to nicotine, alcohol and cannabis dependence in male twins. *Addiction* **103**, 1391–1398 (2008).
5. Ozburn, A. R., Janowsky, A. J. & Crabbe, J. C. Commonalities and Distinctions Among Mechanisms of Addiction to Alcohol and Other Drugs. *Alcohol. Clin. Exp. Res.* **39**, 1863–1877 (2015).
6. Bevilacqua, L. & Goldman, D. Genes and addictions. *Clin. Pharmacol. Ther.* **85**, 359–361 (2009).
7. Bühler, K. M. *et al.* Common single nucleotide variants underlying drug addiction: More than a decade of research. *Addict. Biol.* **20**, 845–871 (2015).
8. Goldman, D., Oroszi, G. & Ducci, F. The genetics of addictions: uncovering the genes. *Nat. Rev. Genet.* **6**, 521–532 (2005).
9. Hall, F. S., Drgonova, J., Jain, S. & Uhl, G. R. Implications of genome wide association studies for addiction: Are our a priori assumptions all wrong? *Pharmacol. Ther.* **140**, 267–279 (2013).
10. Zhou, Z., Enoch, M. A. & Goldman, D. Gene Expression in the Addicted Brain. *Int Rev Neurobiol* **116**, 251–273 (2014).
11. Kelai, S. *et al.* Netrin g1: Its downregulation in the nucleus accumbens of cocaine-conditioned mice and genetic association in human cocaine dependence. *Addict. Biol.* 1–13, doi:10.1111/adb.12485 (2017).
12. Fernández-Castillo, N. *et al.* Transcriptomic and genetic studies identify NFAT5 as a candidate gene for cocaine dependence. *Transl. Psychiatry* **5**, e667 (2015).
13. Kim, V. N., Han, J. & Siomi, M. C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **10**, 126–39 (2009).
14. Afonso-Grunz, F. & Müller, S. Principles of miRNA–mRNA interactions: beyond sequence complementarity. *Cell. Mol. Life Sci.* **72**, 3127–3141 (2015).
15. Gulyaeva, L. F. & Kushlinskiy, N. E. Regulatory mechanisms of microRNA expression. *J. Transl. Med.* **14**, 143 (2016).
16. Nampoothiri, S. S. & Rajanikant, G. K. Decoding the ubiquitous role of microRNAs in neurogenesis. *Mol. Neurobiol.* doi:10.1007/s12035-016-9797-2 (2016).
17. Stappert, L., Roese-Koerner, B. & Brüstle, O. The role of microRNAs in human neural stem cells, neuronal differentiation and subtype specification. *Cell Tissue Res* **359**, 47–64 (2015).
18. Kolshus, E., Dalton, V. S., Ryan, K. M. & McLoughlin, D. M. When less is more—microRNAs and psychiatric disorders. *Acta Psychiatr. Scand.* **129**, 241–56 (2014).
19. Im, H.-I. & Kenny, P. J. MicroRNAs in neuronal function and dysfunction. *Trends Neurosci.* **35**, 325–34 (2012).
20. Luoni, A. & Riva, M. A. MicroRNAs and psychiatric disorders: From aetiology to treatment. *Pharmacol. Ther.* **167**, 13–27 (2016).
21. Hollins, S. L. & Cairns, M. J. MicroRNA: Small RNA mediators of the brains genomic response to environmental stress. *Prog. Neurobiol.* **143**, 61–81 (2016).
22. Lewohl, J. M. *et al.* Up-regulation of microRNAs in brain of human alcoholics. *Alcohol. Clin. Exp. Res.* **35**, 1928–1937 (2011).
23. Chandrasekar, V. & Dreyer, J.-L. microRNAs miR-124, let-7d and miR-181a regulate Cocaine-induced Plasticity. *Mol. Cell. Neurosci.* **42**, 350–362 (2009).
24. Chandrasekar, V. & Dreyer, J.-L. Regulation of MiR-124, Let-7d, and MiR-181a in the Accumbens Affects the Expression, Extinction, and Reinstatement of Cocaine-Induced Conditioned Place Preference. *Neuropsychopharmacology* **36**, 1149–1164 (2011).
25. Hollander, J. A. *et al.* Striatal microRNA controls cocaine intake through CREB signalling. *Nature* **466**, 197–202 (2010).
26. Im, H.-I., Hollander, J. A., Bali, P. & Kenny, P. J. MeCP2 controls BDNF expression and cocaine intake through homeostatic interactions with microRNA-212. *Nat. Neurosci.* **13**, 1120–1127 (2010).
27. Pietrzykowski, A. Z. *et al.* Posttranscriptional Regulation of BK Channel Splice Variant Stability by miR-9 Underlies Neuroadaptation to Alcohol. *Neuron* **59**, 274–287 (2008).

28. Doura, M. B. & Unterwald, E. M. MicroRNAs Modulate Interactions between Stress and Risk for Cocaine Addiction. *Front. Cell. Neurosci* **10**, 125 (2016).
29. Vosa, U., Esko, T., Kasela, S. & Annilo, T. Altered gene expression associated with microRNA binding site polymorphisms. *PLoS One* **10**, 1–24 (2015).
30. Saunders, M. A., Liang, H. & Li, W.-H. Human polymorphism at microRNAs and microRNA target sites. *Proc. Natl. Acad. Sci* **104**, 3300–3305 (2007).
31. Borel, C. & Antonarakis, S. E. Functional genetic variation of human miRNAs and phenotypic consequences. *Mamm. Genome* **19**, 503–509 (2008).
32. Boudreau, R. L. *et al.* Transcriptome-wide discovery of microRNA binding sites in human brain. *Neuron* **81**, 294–305 (2014).
33. First, M. B., Spitzer, S. R., Gibbon, M. & Williams, J. B. M. *Structured Clinical Interview for DSM-IV Axis I Disorders (SCID)*. (American Psychiatric Press, 1997).
34. Fernández-Castillo, N. *et al.* Association study of 37 genes related to serotonin and dopamine neurotransmission and neurotrophic factors in cocaine dependence. *Genes. Brain. Behav* **12**, 39–46 (2013).
35. Miller, S. A., Dykes, D. D. & Polesky, H. F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16**, 1215 (1988).
36. Compton, W. M., Thomas, Y. F., Stinson, F. S. & Grant, B. F. Prevalence, Correlates, Disability, and Comorbidity of DSM-IV Drug Abuse and Dependence in the United States. *Arch. Gen. Psychiatry* **64**, 566 (2007).
37. Gonzalez, J. R. *et al.* SNPassoc: an R package to perform whole genome association studies. *Bioinformatics* **23**, 654–655 (2007).
38. Sánchez-Mora, C. *et al.* Case–Control Genome-Wide Association Study of Persistent Attention-Deficit Hyperactivity Disorder Identifies FBXO33 as a Novel Susceptibility Gene for the Disorder. *Neuropsychopharmacology* **40**, 915–926 (2015).
39. Barrett, J. C. Haploview: Visualization and analysis of SNP genotype data. *Cold Spring Harb. Protoc.* **2009**, pdb.ip71 (2009).
40. Albertson, D. N. *et al.* Gene expression profile of the nucleus accumbens of human cocaine abusers: Evidence for dysregulation of myelin. *J. Neurochem* **88**, 1211–1219 (2004).
41. Winter, J. Using the Student's t-test with extremely small sample sizes. *Pr. Assessment, Res. Evaluation* **18**, 1–12 (2013).
42. McOmish, C. E., Burrows, E. L., Howard, M. & Hannan, A. J. PLC- β 1 knockout mice as a model of disrupted cortical development and plasticity: Behavioral endophenotypes and dysregulation of RGS4 gene expression. *Hippocampus* **18**, 824–834 (2008).
43. Caricasole, A., Sala, C., Roncarati, R., Terstappen, G. C. & Formenti, E. Cloning and characterization of the human phosphoinositide-specific phospholipase C-beta 1 (PLC β 1). *Biochim. Biophys. Acta - Mol. Cell Res* **1517**, 63–72 (2000).
44. Choi, S.-Y. *et al.* Multiple Receptors Coupled to Phospholipase C Gate Long-Term Depression in Visual Cortex. *J. Neurosci.* **25**, (2005).
45. Wallace, M. A. & Claro, E. Comparison of serotonergic to muscarinic cholinergic stimulation of phosphoinositide-specific phospholipase C in rat brain cortical membranes. *J. Pharmacol. Exp. Ther.* **255** (1990).
46. Wallace, M. A. & Claro, E. A novel role for dopamine: inhibition of muscarinic cholinergic-stimulated phosphoinositide hydrolysis in rat brain cortical membranes. *Neurosci. Lett.* **110**, 155–161 (1990).
47. Kim, D. *et al.* Phospholipase C isozymes selectively couple to specific neurotransmitter receptors. *Nature* **389**, 290–293 (1997).
48. St Pourcain, B. *et al.* Variability in the common genetic architecture of social-communication spectrum phenotypes during childhood and adolescence. *Mol. Autism* **5**, 18 (2014).
49. Girirajan, S. *et al.* Refinement and discovery of new hotspots of copy-number variation associated with autism spectrum disorder. *Am. J. Hum. Genet.* **92**, 221–37 (2013).
50. Schoonjans, A.-S. *et al.* PLCB1 epileptic encephalopathies; Review and expansion of the phenotypic spectrum. *Eur. J. Paediatr. Neurol.* **20**, 474–9 (2016).
51. Udawela, M., Scarr, E., Hannan, A. J., Thomas, E. A. & Dean, B. Phospholipase C beta 1 expression in the dorsolateral prefrontal cortex from patients with schizophrenia at different stages of illness. *Aust. N. Z. J. Psychiatry* **45**, 140–147 (2011).
52. Lo Vasco, V. R., Cardinale, G. & Polonia, P. Deletion of PLCB1 gene in schizophrenia-affected patients. *J. Cell. Mol. Med.* **16**, 844–51 (2012).
53. Drgon, T. *et al.* 'Replicated' genome wide association for dependence on illegal substances: genomic regions identified by overlapping clusters of nominally positive SNPs. **156** (2), 125–138 (2012).
54. Gelernter, J. *et al.* Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene. *Mol. Psychiatry* **19**, 717–723 (2014).
55. Eipper-Mains, J. E. *et al.* Effects of cocaine and withdrawal on the mouse nucleus accumbens transcriptome. *Genes, Brain Behav* **12**, 21–33 (2013).
56. Martín-García, E. *et al.* Frustrated expected reward induces differential transcriptional changes in the mouse brain. *Addict. Biol.* **20**, 22–37 (2015).

Acknowledgements

We are grateful to all the patients, their families and clinical collaborators who contributed to the recruitment of patients (Drs. N. Martínez-Luna, A.C. Abad, J. Alvarós, A. Egido, and M. Robles). We are very grateful to all the healthy participants and the Blood and Tissues Bank of Barcelona who contributed to their recruitment. We thank Michael Bannon for kindly providing us with the transcriptomic data from the study in human post-mortem samples. J.C.-D. and N.F.-C. were supported by contracts from the 'Centro de Investigación Biomédica en Red de Enfermedades Raras' (CIBERER) and L.P.-C. with a contract of the "Ministerio de Educación, Cultura y Deporte" (FPU15/03867) of the Spanish Government. J.C.-D. and L.P.-C. were also supported by 'Generalitat de Catalunya' (2015 FI_B 00448 and 2016 FI_B00728, respectively) and N.F.-C. was awarded with an EMBO short-term fellowship (ASTF 573-2016). Major financial support for this research was received by B.C. from the Spanish 'Ministerio de Economía y Competitividad' (SAF2015-68341-R) and AGAUR, 'Generalitat de Catalunya' (2014SGR932). M.R. is a recipient of a Miguel de Servet contract from the 'Instituto de Salud Carlos III', Spain (CP09/00119 and CPII15/00023) and received financial support from 'Instituto de Salud Carlos III' (PI12/01139, PI14/01700, PI15/01789 and PI16/01505), by the European Regional Development Fund (ERDF) and AGAUR, 'Generalitat de Catalunya' (2014SGR1357 and 2014SGR0932) and the NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation. C.R. received financial support by 'Instituto de Salud Carlos III' (PI13/1911) and 'Plan Nacional Sobre Drogas' (2013/044). We were also supported by the European Community's Seventh Framework Program (under grant agreement number 602805, Aggressotype), the European Community's H2020 Program (under grant agreement number 667302, CoCA) and the ECNP network "ADHD across the lifespan".

Author Contributions

J.C.-D. performed the association study, the miRNA functional studies, the expression study of *PLCB1* and the statistical analysis; C.R., R.F.P.-A., E.R.-C., L.G.-L., A.E. and J.A.R.-Q. and M.C. participated in the recruitment of patients and clinical assessment and coordinated the clinical research; J.C.-D., L.P.-C. and N.F.-C. isolated genomic DNA from samples; C.A. participated in the statistical analysis; J.C.-D., B.C. and N.F.-C. designed the study; J.C.-D. prepared the first draft of the manuscript and all figures and tables; B.C., N.F.-C. and M.R. coordinated the study and supervised the manuscript preparation. All authors contributed to and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-10207-2](https://doi.org/10.1038/s41598-017-10207-2)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



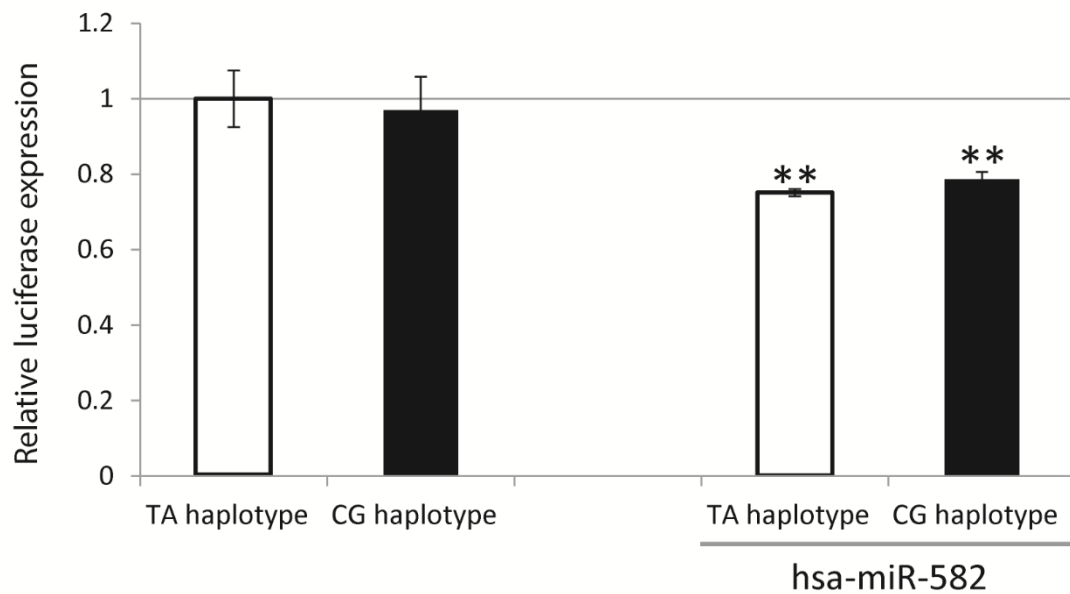
Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017

Supplementary data for:

ASSOCIATION OF THE *PLCB1* GENE WITH DRUG DEPENDENCE

Judit Cabana-Domínguez, Carlos Roncero, Laura Pineda-Cirera, R. Felipe Palma-Álvarez, Elena Ros-Cucurull, Lara Grau-López, Abderaman Esojo, Miquel Casas, Concepció Arenas, Josep Antoni Ramos-Quiroga, Marta Ribasés, Noèlia Fernández-Castillo, Bru Cormand



Supplementary Figure 1. Effect of rs708910 on miRNA-based regulation of gene expression. Relative luciferase expression of the pmirGlo vector with the 3'UTR region of *PLCB1* containing the rs1047381-rs708910 T-A and C-G haplotypes in the presence and absence of hsa-miR-582. **p-value < 0.01

Table S1. SNPs associated with drug dependence in the discovery sample: 735 cases and 739 controls.

Marker	Locus	Genotypes (11-12-22)						Sum	p-value ¹	Adj p-value ²	Adj.(OR -95%CI) ²	
		Controls N (%)			Cases N (%)							
		11	12	22	11	12	22					
rs6840	SCD5	341 (46.5)	317 (43.2)	76 (10.3)	734	284 (39.2)	355 (49.0)	86 (11.8)	725	0.011	0.025	1.24 (1.03-1.50)
rs1285	ID1	491 (66.6)	231 (31.4)	15 (2.0)	737	535 (72.9)	181 (24.7)	18 (2.4)	734	0.029	0.116	1.22 (0.95-1.54)*
rs1872353	FBXO45	455 (61.8)	248 (33.7)	33 (4.5)	736	484 (66.4)	222 (30.4)	23 (3.2)	729	0.044	0.177	1.16 (0.93-1.45)*
rs6855973	GRIA2	483 (65.4)	228 (30.8)	28 (3.8)	739	514 (70.0)	208 (28.4)	12 (1.6)	734	0.014	0.114	1.20 (0.95-1.54)*
rs1047383	PLCB1	307 (41.6)	339 (45.9)	92 (12.5)	738	269 (36.6)	356 (48.5)	109 (14.9)	734	0.039	9.6e-03	1.27 (1.06-1.53)
rs1057377	SPOCK3	343 (46.5)	329 (44.6)	66 (8.9)	738	393 (53.6)	277 (37.8)	63 (8.6)	733	0.026	0.033	1.23 (1.02-1.49)*
rs2597775	QDPR	341 (46.2)	322 (43.6)	75 (10.2)	738	307 (41.9)	330 (45.1)	95 (13.0)	732	0.043	0.439	1.08 (0.89-1.30)

¹Log-additive model; ²Adjusted by age; * When OR<1 the inverted score is shown. In bold: nominally significant p-values

Table S2. SNPs associated with drug dependence in the replication sample: 663 cases and 667 controls.

Marker	Locus	Genotypes (11-12-22)						Sum	p-value ¹	Adj p-value ²	Adj.(OR -95%CI) ²	
		Controls N (%)			Cases N (%)							
		11	12	22	11	12	22					
rs6840	SCD5	295 (44.7)	291 (44.1)	74 (11.2)	660	295 (44.8)	283 (42.9)	81 (12.3)	659	0.785	0.420	-
rs1285	ID1	492 (74.4)	157 (23.8)	12 (1.8)	661	453 (68.8)	182 (27.7)	23 (3.5)	658	0.010	0.034	1.32 (1.02-1.71)
rs1872353	FBXO45	407 (61.4)	229 (34.5)	27 (4.1)	663	447 (67.6)	186 (28.2)	28 (4.2)	661	0.051	3.8e-03	-
rs6855973	GRIA2	458 (69.3)	186 (28.1)	17 (2.6)	661	452 (68.5)	191 (28.9)	17 (2.6)	660	0.779	0.486	-
rs1047383	PLCB1	284 (42.7)	289 (43.5)	92 (13.8)	665	213 (32.4)	351 (53.3)	94 (14.3)	658	3.8e-03	1.5e-03	1.37 (1.13-1.67)
rs1057377	SPOCK3	349 (52.6)	253 (38.1)	62 (9.3)	664	331 (50.3)	265 (40.3)	62 (9.4)	658	0.516	0.945	-
rs2597775	QDPR	293 (44.3)	294 (44.4)	75 (11.3)	662	294 (44.8)	267 (40.6)	96 (14.6)	657	0.461	0.794	-

¹Log-additive model; ²Adjusted by age; In bold: nominally significant p-values

Table S3. SNPs associated with drug dependence in the pooled sample: 1393 cases and 1406 controls.

Marker	Locus	Genotypes (11-12-22)										Sum	Sum	p-value ¹	Adj p-value ²	Adj.(OR -95%CI) ²
		Controls N (%)			Cases N (%)			22	22	22	Sum					
		11	12	Sum	11	12	Sum									
rs1285	<i>ID1</i>	983 (70.3)	388 (27.8)	27 (1.9)	1398	988 (71)	363 (26.1)	41 (2.9)	1392	0.857	0.754	-				
rs1047383	<i>PLCB1</i>	591 (42.1)	628 (44.8)	184 (13.1)	1403	482 (34.6)	707 (50.8)	203 (14.6)	1392	4.8e-04	3.7e-04	1.26 (1.07-1.48)				

¹Log-additive model; ²Adjusted by age; In bold: nominally significant p-values.

Article 5

“Expressió gènica i perfil epigenètic d’un model de ratolí d’addicció al menjar”

Resum:

Hi ha força evidències que els aliments saborosos (rics en sucres refinats i greixos) són addictius i condueixen a algunes formes d’obesitat. Diversos processos, entre els que s’inclouen la recompensa i la plasticitat neuronal, estan implicats en la transició de l’ús a la dependència mitjançant canvis persistents en l’expressió de gens en regions cerebrals que juguen un paper important en l’addicció. S’han assenyalat els mecanismes epigenètics com a candidats ideals per impulsar aquests canvis en trastorns complexos com l’addicció. La metilació del DNA i l’acció dels miRNAs són importants per a la regulació de l’expressió gènica i juguen un paper clau en la plasticitat neuronal, la memòria i l’aprenentatge, factors centrals en l’addicció. En aquest treball ens hem proposat avaluar els canvis d’expressió de gens que codifiquen proteïnes i de gens que codifiquen miRNAs al nucli accumbens i a l’escorça prefrontal medial de dues subpoblacions extremes de ratolins vulnerables o resistents a l’addicció al menjar. A més, també hem explorat la metilació del DNA com un possible mecanisme regulador que provocaria aquests canvis en l’expressió gènica. Utilitzant aquesta aproximació, hem identificat 154 gens que codifiquen proteïnes i nou que codifiquen miRNAs que mostren una expressió diferencial en els dos grups de ratolins. Els resultats s’han replicat en una mostra independent, i molts dels gens s’havien associat prèviament amb altres trastorns psiquiàtrics, fenotips cognitius, obesitat o trets metabòlics. També hem identificat regions diferencialment metilades situades al promotor de 265 gens que codifiquen proteïnes i en dos gens de miRNAs, tot revelant una possible explicació de la regulació dels canvis d’expressió observats. A més, hem observat un enriquiment de gens expressats diferencialment en una via prèviament relacionada amb l’addicció i amb trastorns alimentaris, la ruta de senyalització d’oxitocina. En resum, en aquest estudi hem identificat diversos mRNAs i miRNAs com a possibles factors que contribueixen a la susceptibilitat a l’addicció al menjar. Cal destacar que, en alguns d’ells, hem trobat alteracions en la metilació de la seva regió promotora que podrien explicar els canvis observats en l’expressió gènica. Cal fer més estudis per entendre la seva contribució específica al trastorn.

Referència : Pineda-Cirera L*, Cabana-Domínguez J*, Domingo L*, Martín-García E, Maldonado R#, Fernández-Castillo N#, Cormand B#. Gene expression and epigenetic profiling of a mouse model of food addiction. (Manuscrit en preparació).

*equally contributed to this work; #equally supervised this work

Gene expression and epigenetic profiling of a mouse model of food addiction

Laura Pineda-Cirera^{1,2,3,4,*}, Judit Cabana-Domínguez^{1,2,3,4,*}, Laura Domingo^{5*}, Elena Martín-García⁵, Rafael Maldonado^{5,#}, Noèlia Fernández-Castillo^{1,2,3,4#}, Bru Cormand^{1,2,3,4#}

¹Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain

²Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Spain

³Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Catalonia, Spain

⁴Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Catalonia, Spain

⁵Laboratori de Neurofarmacologia, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, PRBB, Barcelona, Spain

*Equally contributed

#Equally supervised this work

CORRESPONDING AUTHORS:

Bru Cormand. Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Avinguda Diagonal 643, edifici Prevosti, 3^a planta, 08028, Barcelona, Catalonia, Spain. Tel.: (+34) 93 4021013. Fax: (+34) 93 4034420; Email: bcormand@ub.edu

Noèlia Fernández-Castillo. Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Avinguda Diagonal 643, edifici Prevosti, 3^a planta, 08028, Barcelona, Catalonia, Spain. Tel.: (+34) 93 4037082. Fax: (+34) 93 4034420; Email: noefernandez@ub.edu

ABSTRACT

Increasing evidence supports the idea that hyperpalatable food (high in refined sugars and fats) is addictive and leads to some forms of obesity. Several processes, that include reward and neuroadaptation, mediate the transition from use to dependence via long-lasting changes in gene expression within regions of the brain which play an important role in addiction. Epigenetic mechanisms have been pointed out as ideal candidates in driving these long-lasting changes in complex disorders like addiction. DNA methylation and miRNAs are important regulators of gene expression, and they play a remarkable role in neuronal plasticity, memory and learning that underlie addiction. Here, we aimed to inspect expression changes of protein-coding genes and miRNAs in the nucleus accumbens and medial prefrontal cortex of two extreme subpopulations of mice in the phenotype continuum from vulnerability to resilience to food addiction. Furthermore, we explored DNA methylation as a possible regulatory mechanism driving these changes in gene expression. Using this approach, we highlighted 154 protein-coding genes and nine miRNAs showing differential expression that were replicated in an independent sample, most of them previously associated with other psychiatric conditions, obesity or metabolic traits. We

also identified differentially methylated regions in the promoter region of 26 protein-coding genes and two miRNA genes, revealing a possible regulatory explanation for the expression changes observed. Furthermore, we described an enrichment of differentially expressed genes in a pathway previously related to addiction and eating disorders, the “Oxytocin signalling pathway”. To sum up, we identified several protein-coding genes and miRNAs as possible contributors to food addiction susceptibility. Noteworthy, in some of them, we found alterations in the methylation of their promoter region that could drive the changes observed in gene expression. Further studies are required to understand their specific contribution to the disorder.

INTRODUCTION

Obesity is a worldwide problem with 650 million affected individuals in 2016 (body mass index, BMI ≥ 30 kg/m²) with a threefold increase since 1975 (World Health Organization, 2020). The availability of low-cost high calorie foods that are hyperpalatable (HP) together with the promotion of sedentary lifestyles have contributed to the increase of obesity prevalence. HP food has reinforcing effects that are similar to the ones obtained from drugs of abuse, both acting through an increase of dopamine release in the nucleus accumbens (Volkow et al., 2013; Volkow, Wise and Baler, 2017). Over the past few years, food addiction has been pointed as a new behavioural addiction, although it is still not included as a clinical disorder in the Diagnostic and statistical manual of mental disorders (DSM-V) (American Psychiatric Association, 2013). It is diagnosed through the YFAS2.0 questionnaire (Gearhardt, Corbin and Brownell, 2016) based on the DSM-V criteria for addiction (American Psychiatric Association, 2013). The prevalence of food addiction in the general population ranges from 5 to 10% (Gearhardt, Corbin and Brownell, 2009, 2016) and it presents higher values within extreme weight subpopulations (underweight, overweight and obesity), reaching the maximum prevalence of 30% in the obesity III group (BMI ≥ 40 kg/m²) (Hauck et al., 2017). This pattern of prevalence proves its involvement in obesity but also in eating disorders (including bulimia nervosa and binge eating disorder) (Meule, 2011; Hauck et al., 2017).

Food and drug addiction have similarities with regard to the underlying neurobiological mechanisms and some behaviours that are shared among them: loss of control, persistent desire, fail in controlling the amount of food/drug consumed, high motivation for food/drug seeking and repeated consumption in spite of negative consequences (Hebebrand et al., 2014; Lindgren et al., 2018). In the last years, animal models for drug addiction have been widely used to understand the genetic, epigenetic and environmental factors underlying these neurobiological mechanisms

and behaviours (Koob and Volkow, 2016) and interestingly, a mouse model for food addiction based on classical operant conditioning has been recently validated (Mancino et al., 2015; Domingo-Rodriguez et al., 2020) and can be used for the same purposes.

The transition from drug use to dependence is linked to neuroadaptive changes in the central nervous system, including the regulation of gene expression, and it is known that epigenetic mechanisms play a crucial role in driving these changes in gene expression (Zhou, Enoch and Goldman, 2014; Hamilton and Nestler, 2019; Nestler and Lüscher, 2019). Epigenetics is defined as changes in chromatin structure that regulate gene expression without affecting the DNA sequence of the genes themselves. DNA methylation and microRNA are two epigenetic mechanisms involved in gene expression regulation. DNA methylation is the most stable one and is catalysed by DNA methyltransferases (DNMTs) that add a methyl group mainly to cytosines of CpG dinucleotides, affecting the chromatin conformation (Klose and Bird, 2006). When affects a gene promoter, methylation typically acts repressing gene transcription. MiRNAs are small molecules of around 22 nucleotides that work as posttranscriptional regulators by binding mainly to the 3' UTR of target genes (Forman and Collier, 2010). There are multiple evidences showing the involvement of these epigenetic mechanisms to drug addiction and obesity (Feng and Nestler, 2013; Nestler, 2014; Cadet, McCoy and Jayanthi, 2016; Albuquerque et al., 2017; Sayols-Baixeras et al., 2017; Walker and Nestler, 2018; Davis and Bonder, 2019; Hamilton and Nestler, 2019), and therefore, given the close relationship between these two disorders and food addiction, we may expect a similar involvement of these epigenetic mechanisms to the susceptibility to food addiction.

The aim of the present study is to assess expression changes of protein coding-genes and miRNA genes in a mouse model of food addiction, with focus on two key brain areas of the reward system, the nucleus accumbens (NAc) and the medial prefrontal cortex (mPFC). For this purpose, we compare two extreme subpopulations of mice that are vulnerable or resilient to food addiction. Finally, we ascertained if these expression changes can be driven by DNA methylation changes lying in the promoter region of our candidate genes.

MATERIAL AND METHODS

Animals

C57BL/6J inbred male mice (N = 51, two months old) were used for the experiments. Mice were housed individually, with food and water provided ad libitum and under controlled laboratory conditions (Temperature: $21 \pm 1^\circ\text{C}$; Humidity: $55 \pm 10\%$). Tests were carried out during the dark phase of a reverse light cycle (lights off at 8 a.m and on at 8 p.m). All procedures were approved by the local ethics committee (Comitè Ètic d'Experimentació Animal-Parc de Recerca Biomèdica de Barcelona, CEEA-PRBB, agreement N°9687) and performed in compliance with the guidelines of the European Communities Council Directive 2010/63/EU regulating animal research. All efforts were made to minimize the number of mice used and the suffering.

Chocolate-flavoured self-administration experimental design

Mice were trained in operant chambers (Model ENV-307A-CT, Med Associates, Georgia, VT, USA) with chocolate-flavoured pellets (20 mg, highly isocaloric pellet, TestDiet, Richmond, IN, USA) using a protocol previously described (Mancino et al., 2015; Domingo-Rodriguez et al., 2020). A daily self-administration session for 98 days is performed, it comprises two active periods with access to pellets of 25 min separated by a 10 min period with no access to pellets. Each session ends when 100 rewards are delivered or after 1h, whichever occurs first. The protocol consists of six days of self-administration sessions under a FR1 (fixed ratio) schedule of reinforcement followed by a FR5 schedule for all the other sessions. A detailed timeline of similar protocols can be found in two previous studies (Mancino et al., 2015; Domingo-Rodriguez et al., 2020).

Three addiction criteria defined previously (Mancino et al., 2015; Domingo-Rodriguez et al., 2020) were evaluated during the late period (days 82-92): i) persistence to response, ii) motivation, iii) compulsivity. An animal was considered positive for any of the criteria when the score of each criterion was above the 75th percentile of the normal distribution of the population of mice used (Mancino et al., 2015; Domingo-Rodriguez et al., 2020). Mice that achieved two or three criteria were considered vulnerable to food addiction-like phenotype. Finally, from the total population of mice, we selected the two extreme subpopulations of six animals related to the phenotypes of vulnerability and resilience to food addiction-like phenotype for the discovery sample. We repeated the same process with the remaining animals trained with chocolate flavoured pellets to obtain the replica sample (Supplementary Figure 1).

Sample collection and DNA, RNA and miRNA isolation

Tissue collection was performed after the last FR5 session. After decapitation, brains were removed from the skulls and processed rapidly on ice. The mPFC and NAc were collected and kept in individual tubes, frozen on dry ice and stored at -80°C. DNA, RNA and miRNA were isolated from mPFC and NAc of all the animals from the discovery and replica samples using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen, Melbourne, VIC, Australia). DNA was stored at -20°C and RNA and miRNA were stored at -80°C.

mRNA sequencing and analysis

Expression changes in mRNAs was assessed by NGS sequencing of mRNAs (mRNAseq), which was carried out by the “Centre Nacional d’Anàlisi Genòmica” (CNAG-CRG, Barcelona, Spain). RNA integrity was previously assessed by Bioanalyzer and quantified with a Qubit fluorometer. The KAPA Stranded mRNA-Seq Kit (Roche, Basel, Switzerland) was used to construct the libraries and sequenced on an Illumina HiSeq 3000/4000 with 75 bp paired end (PE) reads for the discovery sample and on an Illumina NovaSeq 6000 SP with 50 PE reads for the replica sample. RNAseq reads were cleaned of adapters and low quality reads and mapped against the *Mus musculus* genome of reference (GRCm38/mm10) using STAR version 2.5.3a (Dobin et al., 2013) with ENCODE parameters. Quantification of annotated genes (Gencode version M18 and M24 for the discovery and replica sample, respectively) was done using RSEM version 1.3.0 (Li and Dewey, 2011) with default options. Differential expression between conditions was performed with DESeq2 version 1.18 (Love, Huber and Anders, 2014) with default parameters. This analysis was carried out by the CNAG-CRG (Barcelona, Spain).

SmallRNA sequencing and analysis

For assessing expression changes in miRNAs, smallRNAseq was performed also by the Centre de Regulació Genòmica (CRG, Barcelona, Spain). RNA integrity was previously assessed and quantified in the same samples. The NEBNext® Ultra™ Small RNA Library Prep kit (New England Biolabs, Ipswich, MA, USA) was used for the library preparation and the Illumina HiSeq 2500 System was used for the sequencing, with 50 bp single end (SE) reads for the discovery and replica samples. The analysis of smallRNAseq was carried out through the OASIS2 pipeline (<http://oasis.dzne.de/>) (Rahman et al., 2018). Briefly, FastQC v0.10.1 (Andrews S., 2010) was used to inspect the reads quality and CutAdapt 1.7.1 (Martin, 2011) was used to clean the data of adapters and low quality reads. After the trimming of the adaptor, only the reads between 15-32 nucleotides were kept. Next, reads were mapped against the *Mus musculus* genome of reference

(GRCm38/mm10) with STAR 2.4.1d (Dobin et al., 2013), reads of length 15-19 nucleotides were aligned allowing for 0 mismatches whereas reads of length 20-32 nucleotides were mapped allowing for one mismatch. Finally, the differential expression analysis was done by DESeq2 (Love, Huber and Anders, 2014). In order to discard those miRNAs with very low expression or sequencing errors, we only considered the mature miRNAs showing at least five reads across all the samples.

Selection of mRNAs and miRNAs

We selected those genes encoding protein coding mRNAs and miRNAs that showed nominal differential expression in NAc or mPFC, with concordant effects on gene expression in both the discovery and the replica samples. We then used the two-tailed Fisher Exact test to determine the statistical significance of the overlap observed between the discovery and replica samples for each brain area considering the upregulated and downregulated genes separately. For further analysis, we only used those genes showing a statistical significance of the overlap observed between the discovery and replica samples.

Differences in methylation levels were assessed by Methyl Binding Domain sequencing (MBDseq), which was carried out by Macrogen Inc. (Seoul, South Korea), only in the discovery sample. DNA integrity was confirmed by agarose gel electrophoresis and quantified by Picogreen (Invitrogen, Carlsbad, CA, USA) using Victor 3 fluorometry. MBD enrichment was performed using the MethylMiner Methylated DNA Enrichment Kit. Library preparation was carried out with Truseq DNA nano (350 bp) kit (Illumina) and sequencing was performed on an Illumina Novaseq 6000 with 100 bp PE reads. Raw data read quality was analysed with FastQC (Andrews S., 2010) and cleaned of adapters or low quality reads using Trimmomatic 0.36 (Bolger, Lohse and Usadel, 2014) under the following parameters: LEADING:3, TRAILING:3, SLIDINGWINDOW:4:20 and MINLEN:36. Subsequently, reads were mapped against the *Mus musculus* genome of reference (GRCm38/mm10) with Bowtie2 (Langmead and Salzberg, 2012) using the default parameters. MEDIPS R package (Lienhard et al., 2014) was used to identify differential methylated regions (DMRs) between mice vulnerable or resilient to food addiction-like phenotype under the following parameters: $uniq = 1e-3$, $extend = 100$, $shift = 0$, $ws = 100$, $minRowSum = 10$. All quality controls included in MEDIPS software were performed to ensure reliable and reproducible results: saturation analysis, correlation between samples, sequence pattern coverage and CpG enrichment.

We retrieved those DMRs showing nominal associations with the phenotype that were located in the putative promoter region (5 kb upstream and 1 kb downstream from the transcription start site (TSS)) of all transcript variants of those DE mRNA and miRNA genes showing a statistical significance of the overlap observed between the discovery and replica samples. Every transcript could have one or more DMRs. When it contained more than one DMR, the DMRs were considered only if they all had the same direction of the effect. Finally, we only selected those DMRs with an effect on methylation opposite to that observed for gene expression.

Functional annotation

We compiled functional annotations for the highlighted mRNAs, miRNAs and DMRs using multiple databases. For the selected mRNAs, we tested enrichment in Gene Ontology categories (considering “Biological Process”, “Cellular Component” and “Molecular Function”) or KEGG pathways (Kyoto Encyclopedia of Genes and Genomes) using the WebGESTALT software (Wang et al., 2017). We also inspected whether the selected protein-coding genes had previously been associated with psychiatric conditions, cognitive phenotypes, body mass index (BMI) or metabolic traits using the GWAS Catalog (<https://www.ebi.ac.uk/gwas/>) (Buniello et al., 2019). For those DE miRNAs with DMRs in their putative promoter region, we assessed if there was an enrichment of their target protein-coding genes among the differentially expressed genes in the discovery and replica samples. We used two lists of genes targeted by miRNAs: experimentally validated (miRTarBase; <http://mirtarbase.mbc.nctu.edu.tw/php/index.php>) (Chou et al., 2018) and predicted (miRSystem; <http://mirsystem.cgm.ntu.edu.tw/>) (Lu et al., 2012). Finally, we assessed whether the highlighted DMRs overlap with chromatin state marks related to transcriptional regions, enhancers, promoters or heterochromatin regions in the developing mouse brain (Gorkin et al., 2017). This was done by using the subtrack “Chromatin state of embryonic tissue from ENCODE 3” included in the “ENCODE Regulation Track” in the UCSC Genome Browser (<https://genome.ucsc.edu/>) (Kent et al., 2002).

RESULTS

Acquisition of the operant training maintained by food

Mice were subjected to operant training with chocolate-flavoured pellets to identify animals showing vulnerability or resilience to a food addiction-like phenotype. To do so, we calculated the individual scores for each criterion (motivation, persistence to response and compulsivity) to

classify the animals in vulnerable (2-3 criteria) or resilient (0 criterion) to a food addiction-like phenotype (Figure 1 and Supplementary Figure 1). From the total of 51 animals, 23.5% were classified as vulnerable to a food addiction-like phenotype, demonstrating that highly palatable food can lead to an operant-seeking behaviour and a loss of control over food intake. We then selected the discovery sample by choosing the most extreme subpopulations of mice related to the vulnerability and resilience to a food addiction-like phenotype (N = 6 per group) and repeated the procedure with the remaining animals for the replica sample (N = 6 per group) (Figure 1 and Supplementary Figure 1). We did not observe differences in pellet intake among the groups of mice selected for the discovery and replica samples.

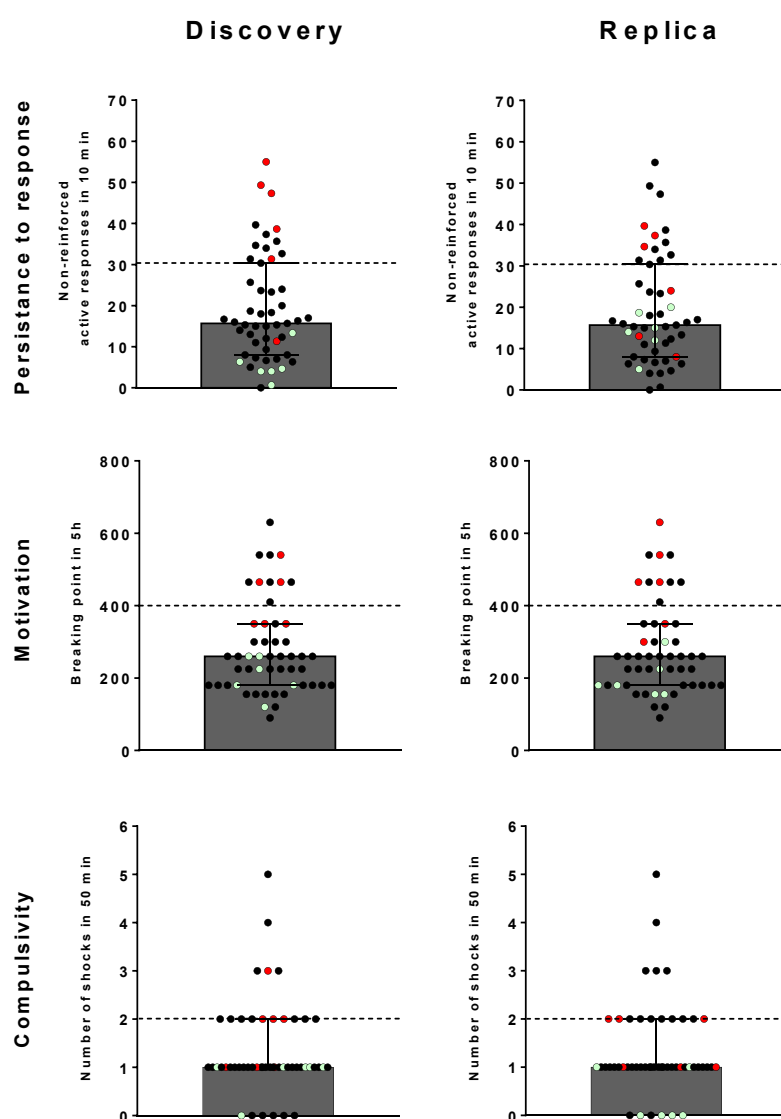


Figure 1. Individual scores of the three addiction criteria (persistence to response, motivation and compulsivity) obtained in the behavioural tests (N = 51). The animals selected for the discovery and replica samples are highlighted in colour, red (vulnerable to food addiction-like phenotype, N = 6) or green (resilient to food addiction-like phenotype, N = 6). The dashed line determines the 75th percentile of the normal distribution of the mice population, used as a threshold to consider a mouse positive for the criterion.

Differences in gene expression of protein coding genes and its regulation by methylation

Differential expressed protein-coding genes (DEGs) were identified in two brain regions, mPFC and NAc. Nominal DEGs ($p < 0.05$) with concordant effects in the discovery and in the replica samples were selected for subsequent analyses. In mPFC we identified differential expression of 45 protein-coding genes, 12 of them upregulated and 33 downregulated in the subpopulation of mice that were vulnerable to the food addiction-like phenotype compared to the resilient group. In NAc we detected altered expression of 121 protein-coding genes, 68 upregulated and 53 downregulated in the subpopulation vulnerable to the food addiction-like phenotype (Supplementary Table 1). We found significant gene overlaps among the discovery and replica samples in all comparisons, except for the upregulated genes in mPFC, that were discarded for further analysis (Table 1). Remarkably, most of the remaining genes had previously been associated with body mass index (BMI), metabolic traits, psychiatric conditions like anorexia and schizophrenia, or cognitive phenotypes like intelligence, among others (Supplementary Table 2).

Table 1. Comparative analyses examining overlap of differentially expressed genes between the discovery and replica samples.

	Nº DEG discovery	Nº DEG replica	Nº DEG in common	p-value	OR (95% CI)
mPFC					
Downregulated	770	428	33	<u>2.42E-05</u>	2.35 (1.68 - 3.20)
Upregulated	1452	530	12	1.00E+00	
NAc					
Downregulated	531	793	53	<u>2.94E-11</u>	3.08 (2.37 - 3.97)
Upregulated	598	943	68	<u>2.92E-13</u>	2.99 (2.37 - 3.74)

DEG: Differentially expressed genes; OR: Odds ratio; Underlined: significant p-values ($p < 0.05$)

The analysis of GO categories and KEGG pathways was performed by considering all the differentially expressed genes for NAc (121 genes) and only the downregulated genes for mPFC (33). We detected an enrichment in several pathways and categories for NAc but not for mPFC, probably because of the low number of DEGs. Interestingly, among the pathways identified in NAc (Table 2), we reported the “Oxytocin signalling”, that was also nominally enriched in mPFC (Nº of genes: 2 (-); Expect: 0.3; Ratio: 7.5; p-value: 2.80E-02), although it did not survive corrections for multiple testing.

Table 2. Enrichment of differentially expressed genes in NAC in GO categories and KEGG Pathways.

Gene Set	Description	Nº of genes		Size	Expect	Ratio	p-value	p-value _{adj}
		Upreg.	Downreg.					
Cellular component								
GO:0044463	Cell projection part	14	7	1559	7.0	3.0	4.12E-06	2.38E-03
GO:0045202	Synapse	9	9	1316	5.9	3.0	1.89E-05	4.38E-03
GO:0043005	Neuron projection	10	9	1471	6.6	2.9	2.37E-05	4.46E-03
GO:0098590	Plasma membrane region	10	6	1238	5.6	2.9	1.16E-04	1.49E-02
GO:0030054	Cell junction	8	6	1080	4.9	2.9	3.18E-04	3.66E-02
Biological process								
GO:0010629	Negative regulation of gene expression	17	7	1632	8.9	2.7	5.16E-06	4.65E-02
KEGG pathway								
mmu05165	Human papillomavirus infection	5	4	370	2.2	4.2	2.34E-04	3.81E-02
mmu04921	Oxytocin signalling pathway	3	3	153	0.9	6.7	2.37E-04	3.81E-02

GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; DEG: Differentially Expressed Genes; Nº of genes: Number of genes upregulated (Upreg.) and downregulated (Downreg.); Size: Number of genes included in the category; p-value_{adj}: p-value adjusted for multiple testing correction via 5 % FDR.

In order to identify possible explanations for the differential expression of protein-coding genes observed, we explored methylation profiles in the promoter region of DEGs. We focused on differentially methylated regions (DMRs) with an effect on methylation that is opposite to that observed for gene expression. We found nine significant DMRs in the possible promoter regions of three differentially expressed protein-coding genes in mPFC, and 60 DMRs for 23 genes in NAC, all of them with opposite effects on methylation and gene expression (Supplementary Table 3). Remarkably, 66 out of the 69 DMRs lie in regions with at least one chromatin state mark described in developing mouse brain (Supplementary Table 3). Among the genes with DMRs in their possible promoter region we highlight *Adams2*, *Arsj* and *Klhl40*, because they presented a remarkable gene expression fold change ($FC > 1.3$) and because the DMRs located in their putative promoter region lie in regions with multiple chromatin state marks related to heterochromatin, enhancer or promoter elements (Supplementary Figures 2, 3 and 4 and Supplementary Tables 1 and 3).

Differences in miRNA expression and regulation by methylation

Following the same strategy as in the DEG analysis, we have identified eleven miRNAs that show differential expression in mPFC both in the discovery and in the replica sample, nine of them downregulated (mmu-miR-29c-3p, -124-3p, -137-3p, -211-5p, -544-3p, -665-3p, -876-5p, -3072-3p and -3085-3p) and two upregulated (mmu-miR-100-5p and -192-5p) in the subpopulation of

mice that are vulnerable to food addiction compared to the resilient group (Supplementary Table 4). We were not able to replicate any of the miRNAs found in the discovery sample for NAc. We found a significant overlap for the downregulated miRNAs in mPFC but not for the upregulated genes, that were excluded from further analyses (Table 3).

Table 3. Comparative analyses examining overlap of DE miRNAs in mPFC between the discovery and replica samples.

	Nº DE miRNAs discovery	Nº DE miRNAs replica	Common	p-value	OR (95% CI)
Downregulated	44	61	9	<u>5.06E-03</u>	3.34 (1.55 - 6.78)
Upregulated	28	39	2	4.18E-01	

DE: Differentially expressed; OR: Odds ratio; Underlined: significant p-values ($p < 0.05$)

The analysis of differential methylation in DE miRNA genes identified four significant DMRs in the possible promoter region of mmu-miR-124-3p and five significant DMRs for mmu-miR-3085-3p with an opposite effect on methylation to that observed for gene expression (Supplementary Table 5). Noteworthy, all the DMRs for mmu-miR-124-3p lie in a region predicted to be an active promoter region in the mouse developing brain, supporting that these DMRs can influence gene expression (Supplementary Figure 5).

We also explored whether the mRNAs that are targeted by mmu-miR-124-3p or mmu-miR-3085-3p were enriched in DEGs. This analysis included both experimentally validated and predicted targets of downregulated miRNAs from mPFC. We observed an enrichment of DEGs among the mmu-miR-124-3p target mRNAs in the discovery and replica samples in mPFC (Table 4), highlighting its potential role in gene expression regulation.

Table 4. Enrichment analysis of mmu-miR-124-3p and mmu-miR-3085-3p target genes among DEG in mPFC.

Database	Sample	Nº DEG	Nº target genes	Nº of DE target		p-value	OR (95% CI)
				genes	genes		
mmu-miR-124-3p							
miRTarBase	Discovery	2222	450	96		<u>1.30E-12</u>	2.46 (2.01 - 3.00)
	Replica	958	450	32		<u>5.16E-03</u>	1.69 (1.21 - 2.32)
miRSystem	Discovery	2222	1028	168		<u>2.50E-10</u>	1.79 (1.54 - 2.07)
	Replica	958	1028	74		<u>1.90E-05</u>	1.75 (1.41 - 2.16)
mmu-miR-3085-3p							
miRTarBase	Discovery	2222	197	26		1.00E-01	
	Replica	958	197	14		5.20E-02	
miRSystem	Discovery	2222	-	-		-	-
	Replica	958	-	-		-	-

DEG: Differentially expressed genes; DE: Differentially expressed; OR: Odds ratio; Underlined: significant p-values ($p < 0.05$); “-“: data not available.

DISCUSSION

Highly palatable food and drugs of abuse act on the reward system, a circuit composed by dopaminergic neuron bodies located in the ventral tegmental area (VTA) that project their axons mainly towards the nucleus accumbens (NAc), the hippocampus, the amygdala and the prefrontal cortex (PFC) (Carter et al., 2016; Walker and Nestler, 2018). The reward circuitry of the brain is similar across humans and mice (Walker and Nestler, 2018) and the gene expression profile is also highly conserved among these two species, especially in tissues with a predominance of gene-specific expression like the brain (Breschi, Gingeras and Guigó, 2017). We took advantage of this and explored the transcriptomic profile in the Nac and mPFC of two subpopulations of mice with extreme food addiction-like phenotypes.

NAc is involved in decision-making and rewarding effects of natural rewards, highly palatable food and drugs of abuse (Volkow et al., 2013; Volkow, Wise and Baler, 2017). In this area, we reported differential expression of 121 protein-coding genes, 23 of them with DMRs in their possible promoter regions with an opposite direction on methylation and gene expression. Among them, we highlight *Adamts2*, *Arsj* and *Khlh40*, because they display an important fold change on gene expression and have DMRs in their promoter regions with chromatin state marks that support a possible regulatory effect. *Adamts2* is involved in the cleavage of type I and II collagen propeptides, *Arsj* belongs to the arylsulfatase family and finally, *Khlh40* is a part of a E3 ubiquitin ligase complex that acts as a key regulator of skeletal muscle development. Moreover, previous studies have suggested a possible involvement of *Adamts2* and *Arsj* in psychiatric traits and conditions linked to food addiction. Specifically, *Adamts2* was previously related to schizophrenia (Crespo-Facorro, Prieto and Sainz, 2015; Ruso-Julve et al., 2019) and GWAS studies showed suggestive associations with sugar intake (Hwang et al., 2019) and attention deficit/hyperactivity disorder (Lasky-Su et al., 2008). On the other hand, *Arsj* showed nominal associations with alcohol dependence (Kalsi et al., 2010) and a GWAS study pointed at variants in this gene with suggestive associations with motor impulsivity (Sanchez-Roige et al., 2019). We observed an enrichment of differentially expressed genes in NAc in GO categories and pathways of interest for food addiction like “Negative regulation of gene expression” or the “Oxytocin signalling pathway”. Oxytocin is a neuropeptide involved in the regulation of many reward-related behaviours (Carter, 2014) and has been previously related to drug addiction and also to eating disorders (Sarnyai and Kovács, 1994, 2014; Kovács, Sarnyai and Szabó, 1998; Giel, Zipfel and Hallschmid, 2018; Leong et al., 2018; Leslie et al., 2019). Furthermore, low levels of this neuropeptide have been associated with increased food consumption, especially of sweet carbohydrates. Taking into account all these evidences, oxytocin deficiency has been proposed as a risk factor for compulsive and addictive

overeating (Davis and Moghimi, 2017), making oxytocin signalling a good candidate pathway to be pursued for understanding food addiction susceptibility or even as a target for a possible treatment.

The other studied region, mPFC, is involved in the executive control over choices made, including self-control (Volkow et al., 2013). In this area, we detected altered expression of 33 protein-coding genes and nine miRNA genes. From those, three protein-coding genes and two miRNA genes (mmu-miR-124-3p and -3085-3p) have DMRs in their possible promoter regions with an opposite effect on methylation than the one observed for gene expression. Noteworthy, we observed an enrichment of validated and predicted target genes of mmu-miR-124-3p that are also differentially expressed, highlighting its potential involvement in gene expression regulation. Furthermore, miR-124 was previously related to cocaine addiction (Chandrasekar and Dreyer, 2009, 2011; Cabana-Domínguez et al., 2018; Viola et al., 2019) as well as to different neuronal processes (Svahn et al., 2016; Kutsche et al., 2018; Veremeyko et al., 2019) and to other psychiatric disorders (Dwivedi, 2017; Gu, Pan and Chen, 2019), supporting the idea of its implication in food addiction susceptibility.

Strengths and limitations of the present study should be discussed. Although the results obtained on altered gene expression are nominal, we have observed a significant overlap of genes between the discovery and the replica samples that strengthen the reliability of our findings. Furthermore, the replica sample supports the results obtained (Chanock et al., 2007) as well as a considerable sample size for this type of experiment. Remarkably, most of the highlighted genes had previously been associated with BMI, metabolic traits, cognitive traits or psychiatric conditions and we obtained an enrichment of differential expression in the “Oxytocin signalling pathway”, extensively related to addiction and to eating disorders (Sarnyai and Kovács, 1994, 2014; Kovács, Sarnyai and Szabó, 1998; Giel, Zipfel and Hallschmid, 2018; Leong et al., 2018; Leslie et al., 2019; King, Gano and Becker, 2020) due to its involvement in the adaptive processes related to reward, tolerance and memory.

To sum up, we assessed differences in the expression of protein-coding genes and miRNA genes between two subpopulations of mice with extreme behaviours towards food addiction-like phenotype (vulnerable and resilient). This allowed us to highlight multiple candidate genes to contribute to the susceptibility to food addiction. Furthermore, we provided a possible explanation for some of the gene expression changes observed by assessing the methylation profiles on their possible promoter regions, known to inversely correlate with gene expression.

FUNDINGS

This work was supported by the Spanish 'Ministerio de Economía y Competitividad-MINECO' (#SAF2017-84060-R-AEI/FEDER-UE), the Spanish 'Instituto de Salud Carlos III, RETICS-RTA' (#RD12/0028/0023), the 'Generalitat de Catalunya, AGAUR' (#2017 SGR-669), 'ICREA-Acadèmia' (#2015) and the Spanish 'Ministerio de Sanidad, Servicios Sociales e Igualdad, 'Plan Nacional Sobre Drogas of the Spanish Ministry of Health' (#PNSD-2017I068) to RM, 'Fundació La Marató-TV3' (#2016/20-30) and 'Plan Nacional Sobre Drogas of the Spanish Ministry of Health' (#PNSD-2019I006) to E.M-G., Spanish 'Ministerio de Ciencia, Innovación y Universidades' (#RTI2018-100968-B-100), Spanish 'Ministerio de Economía y Competitividad' (#SAF2015-68341-R), 'AGAUR-Generalitat de Catalunya' (#2017-SGR-738) and Spanish 'Ministerio de Sanidad, Servicios Sociales e Igualdad, 'Plan Nacional Sobre Drogas of the Spanish Ministry of Health' (#PNSD-2017I050) to BC. The research leading to these results has also received funding from the European Union H2020 Program [H2020/2014-2020] under grant agreements n° 667302 (CoCA), 643051 (MiND) and 728018 (Eat2beNICE) to BC. LP-C was supported by 'Ministerio de Educación, Cultura y Deporte' [grant number FPU15/03867]. JC-D was supported by the H2020 CoCA and Eat2beNICE projects and NF-C by 'Centro de Investigación Biomédica en Red de Enfermedades Raras' (CIBERER).

ACKNOWLEDGMENTS

We are grateful to the Genomics Unit at the CRG for assistance with the mRNAseq and smallRNAseq.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest or relevant financial interests.

BIBLIOGRAPHY

- Albuquerque, D. et al. (2017) 'The contribution of genetics and environment to obesity', *British Medical Bulletin*, 123(1), pp. 159–173. doi: 10.1093/bmb/ldx022.
- American Psychiatric Association (2013) *Diagnostic and statistical manual of mental disorders* (5th ed.), Arlington, VA: American Psychiatric Publishing.
- Andrews S. (2010) 'FastQC: A quality control tool for high throughput sequence data'. Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Bolger, A. M., Lohse, M. and Usadel, B. (2014) 'Trimmomatic: a flexible trimmer for Illumina sequence data', *Bioinformatics*, 30(15), pp. 2114–2120.
- Breschi, A., Gingeras, T. R. and Guigó, R. (2017) 'Comparative transcriptomics in human and mouse', *Nature Reviews Genetics*, 18(7), pp. 425–440. doi: 10.1038/nrg.2017.19.
- Buniello, A. et al. (2019) 'The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019', *Nucleic Acids Research*, 47(D1), pp. D1005–D1012. doi: 10.1093/nar/gky1120.
- Cabana-Domínguez, J. et al. (2018) 'MiR-9, miR-153 and miR-124 are down-regulated by acute exposure to cocaine in a dopaminergic cell model and may contribute to cocaine dependence', *Translational Psychiatry*, 8(1), p. 173. doi: 10.1038/s41398-018-0224-5.
- Cadet, J. L., McCoy, M. T. and Jayanthi, S. (2016) 'Epigenetics and addiction', *Clinical Pharmacology and Therapeutics*, 99(5), pp. 502–511.
- Carter, A. et al. (2016) 'The Neurobiology of "Food Addiction" and Its Implications for Obesity Treatment and Policy', *Annual Review of Nutrition*, 36(1), pp. 105–128. doi: 10.1146/annurev-nutr-071715-050909.
- Carter, C. S. (2014) 'Oxytocin Pathways and the Evolution of Human Behavior', *Annual Review of Psychology*, 65(1), pp. 17–39. doi: 10.1146/annurev-psych-010213-115110.
- Chandrasekar, V. and Dreyer, J. L. (2009) 'microRNAs miR-124, let-7d and miR-181a regulate Cocaine-induced Plasticity', *Molecular and Cellular Neuroscience*, 42(4), pp. 350–362. doi: 10.1016/j.mcn.2009.08.009.
- Chandrasekar, V. and Dreyer, J. L. (2011) 'Regulation of MiR-124, Let-7d, and MiR-181a in the accumbens affects the expression, extinction, and reinstatement of cocaine-induced conditioned place preference', *Neuropsychopharmacology*, 36(6), pp. 1149–1164. doi: 10.1038/npp.2010.250.
- Chanock, S. J. et al. (2007) 'Replicating genotype-phenotype associations', *Nature*, 447(7145), pp. 655–660. doi: 10.1038/447655a.
- Chou, C.-H. et al. (2018) 'miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions', *Nucleic Acids Research*, 46(D1), pp. D296–D302. doi: 10.1093/nar/gkx1067.
- Crespo-Facorro, B., Prieto, C. and Sainz, J. (2015) 'Schizophrenia Gene Expression Profile Reverted to Normal Levels by Antipsychotics', *International Journal of Neuropsychopharmacology*, 18, p. pyu066.
- Davis, C. and Bonder, R. (2019) 'Genetics and epigenetics of food addiction', in Cottone, P. et al. (eds) *Compulsive Eating Behavior and Food Addiction: Emerging Pathological Constructs*. York University, Toronto, ON, Canada: Academic Press, pp. 293–327. doi: 10.1016/B978-0-12-816207-1.00010-X.
- Davis, C. and Moghimi, E. (2017) "'Oxytocin Deficiency": Implications for the Food-Addiction Construct', *Current Addiction Reports*, 4(2), pp. 158–164. doi: 10.1007/s40429-017-0135-1.
- Dobin, A. et al. (2013) 'STAR: Ultrafast universal RNA-seq aligner', *Bioinformatics*, 29(1), pp. 15–21. doi: 10.1093/bioinformatics/bts635.
- Domingo-Rodríguez, L. et al. (2020) 'A specific prelimbic-nucleus accumbens pathway controls resilience versus vulnerability to food addiction', *Nature Communications*, 11(1), p. 782. doi: 10.1038/s41467-020-14458-y.

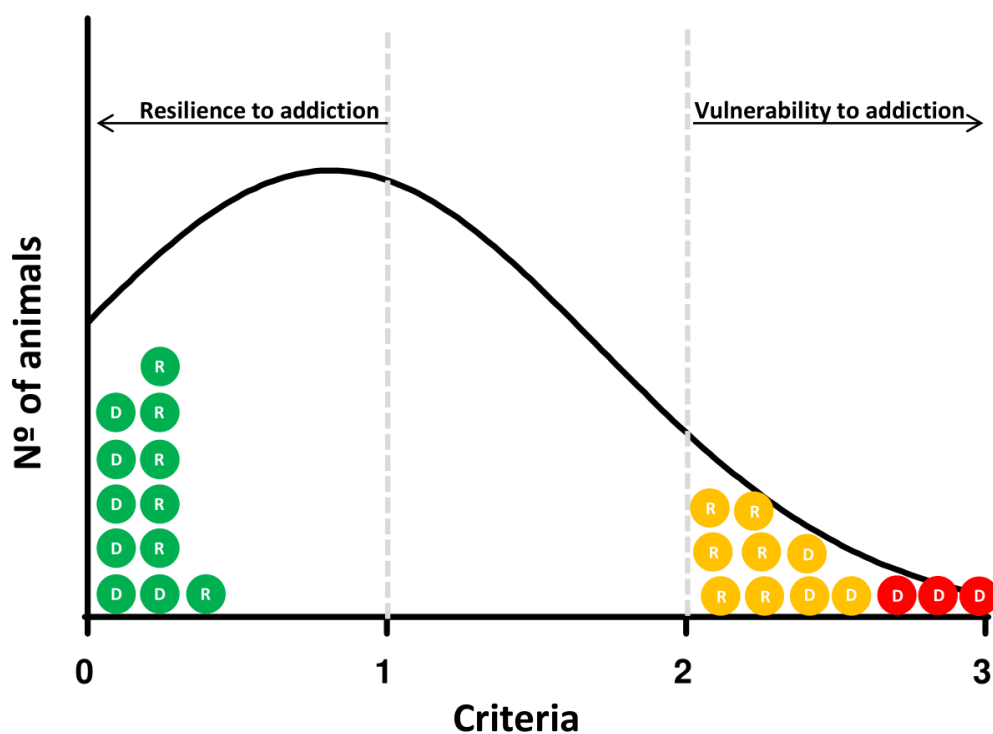
- Dwivedi, Y. (2017) 'microRNA-124: a putative therapeutic target and biomarker for major depression', *Expert Opinion on Therapeutic Targets*, 21(7), pp. 653–656. doi: 10.1080/14728222.2017.1328501.
- Feng, J. and Nestler, E. J. (2013) 'Epigenetic mechanisms of drug addiction', *Current Opinion in Neurobiology*, 23(4), pp. 521–528. doi: 10.1016/j.conb.2013.01.001.
- Forman, J. J. and Coller, H. A. (2010) 'The code within the code: MicroRNAs target coding regions', *Cell Cycle*, 9(8), pp. 1533–1541. doi: 10.4161/cc.9.8.11202.
- Gearhardt, A. N., Corbin, W. R. and Brownell, K. D. (2009) 'Preliminary validation of the Yale Food Addiction Scale', *Appetite*, 52(2), pp. 430–436. doi: 10.1016/j.appet.2008.12.003.
- Gearhardt, A. N., Corbin, W. R. and Brownell, K. D. (2016) 'Development of the Yale Food Addiction Scale Version 2.0', *Psychology of Addictive Behaviors*, 30(1), pp. 113–121. doi: 10.1037/adb0000136.
- Giel, K., Zipfel, S. and Hallschmid, M. (2018) 'Oxytocin and Eating Disorders: A Narrative Review on Emerging Findings and Perspectives', *Current Neuropharmacology*, 16(8), pp. 1111–1121. doi: 10.2174/1570159x15666171128143158.
- Gorkin, D. U. et al. (2017) 'Systematic mapping of chromatin state landscapes during mouse development', *bioRxiv*, p. 166652. doi: 10.1101/166652.
- Gu, Z., Pan, J. and Chen, L. (2019) 'MIR-124 suppression in the prefrontal cortex reduces depression-like behavior in mice', *Bioscience Reports*, 39(9), p. BSR20190186. doi: 10.1042/BSR20190186.
- Hamilton, P. J. and Nestler, E. J. (2019) 'Epigenetics and addiction', *Current Opinion in Neurobiology*, 59, pp. 128–136. doi: 10.1016/j.conb.2019.05.005.
- Hauck, C. et al. (2017) 'Prevalence of "Food Addiction" as Measured with the Yale Food Addiction Scale 2.0 in a Representative German Sample and Its Association with Sex, Age and Weight Categories', *Obesity Facts*, 10(1), pp. 12–24. doi: 10.1159/000456013.
- Hebebrand, J. et al. (2014) "'Eating addiction", rather than "food addiction", better captures addictive-like eating behavior', *Neuroscience and Biobehavioral Reviews*, 47, pp. 295–306. doi: 10.1016/j.neubiorev.2014.08.016.
- Hwang, L. D. et al. (2019) 'New insight into human sweet taste: A genome-wide association study of the perception and intake of sweet substances', *American Journal of Clinical Nutrition*, 109(6), pp. 1724–1737. doi: 10.1093/ajcn/nqz043.
- Kalsi, G. et al. (2010) 'A systematic gene-based screen of chr4q22-q32 identifies association of a novel susceptibility gene, DKK2, with the quantitative trait of alcohol dependence symptom counts', *Human Molecular Genetics*, 19(12), pp. 2497–2506. doi: 10.1093/hmg/ddq112.
- Kent, W. J. et al. (2002) 'The Human Genome Browser at UCSC', *Genome Research*, 12(6), pp. 996–1006. doi: 10.1101/gr.229102.
- King, C. E., Gano, A. and Becker, H. C. (2020) 'The role of oxytocin in alcohol and drug abuse', *Brain Research*, 1736, p. 146761. doi: 10.1016/j.brainres.2020.146761.
- Klose, R. J. and Bird, A. P. (2006) 'Genomic DNA methylation: The mark and its mediators', *Trends in Biochemical Sciences*, 31(2), pp. 89–97. doi: 10.1016/j.tibs.2005.12.008.
- Koob, G. F. and Volkow, N. D. (2016) 'Neurobiology of addiction: a neurocircuitry analysis', *The Lancet Psychiatry*, 3(8), pp. 760–773. doi: 10.1016/S2215-0366(16)00104-8.
- Kovács, G. L., Sarnyai, Z. and Szabó, G. (1998) 'Oxytocin and addiction: A review', *Psychoneuroendocrinology*, 23(8), pp. 945–962. doi: 10.1016/S0306-4530(98)00064-X.
- Kutsche, L. K. et al. (2018) 'Combined Experimental and System-Level Analyses Reveal the Complex Regulatory Network of miR-124 during Human Neurogenesis', *Cell systems*, 7(4), pp. 438–452.e8. doi: 10.1016/j.cels.2018.08.011.

- Langmead, B. and Salzberg, S. L. (2012) 'Fast gapped-read alignment with Bowtie 2', *Nature Methods*, 9(4), pp. 357–359.
- Lasky-Su, J. et al. (2008) 'Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations', *American Journal of Medical Genetics Part B*, 147B(8), pp. 1345–1354. doi: 10.1002/ajmg.b.30867.
- Leong, K. C. et al. (2018) 'Oxytocin and Rodent Models of Addiction', in Olive, M. F. and Tomek, S. E. (eds) *International Review of Neurobiology. Animal Models for Examining Social Influences on Drug Addiction*, pp. 201–247. doi: 10.1016/bs.irn.2018.07.007.
- Leslie, M. et al. (2019) 'The influence of oxytocin on eating behaviours and stress in women with bulimia nervosa and binge eating disorder', *Molecular and Cellular Endocrinology*, 497, p. 110354. doi: 10.1016/j.mce.2018.12.014.
- Li, B. and Dewey, C. N. (2011) 'RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome', *BMC Bioinformatics*, 12, p. 323. doi: 10.1186/1471-2105-12-323.
- Lienhard, M. et al. (2014) 'MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments', *Bioinformatics*, 30(2), pp. 284–286. doi: doi: 10.1093/bioinformatics/btt650.
- Lindgren, E. et al. (2018) 'Food addiction: A common neurobiological mechanism with drug abuse', *Frontiers in Bioscience*, 23(5), pp. 811–836. doi: 10.2741/4618.
- Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biology*, 15(12), p. 550. doi: 10.1186/s13059-014-0550-8.
- Lu, T.-P. et al. (2012) 'miRSystem: An Integrated System for Characterizing Enriched Functions and Pathways of MicroRNA Targets', *PLoS ONE*, 7(8), p. e42390. doi: 10.1371/journal.pone.0042390.
- Mancino, S. et al. (2015) 'Epigenetic and proteomic expression changes promoted by eating addictive-like behavior', *Neuropsychopharmacology*, 40(12), pp. 2788–2800. doi: 10.1038/npp.2015.129.
- Martin, M. (2011) 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet.J.*, 17(1), pp. 10–12. doi: 10.14806/ej.17.1.200.
- Meule, A. (2011) 'How prevalent is "food addiction"?', *Frontiers in Psychiatry*, 2, p. 61. doi: 10.3389/fpsy.2011.00061.
- Nestler, E. J. (2014) 'Epigenetic mechanisms of drug addiction', *Neuropharmacology*, 76, pp. 259–268. doi: 10.1016/j.neuropharm.2013.04.004.
- Nestler, E. J. and Lüscher, C. (2019) 'The Molecular Basis of Drug Addiction: Linking Epigenetic to Synaptic and Circuit Mechanisms', *Neuron*, 102(1), pp. 48–59. doi: 10.1016/j.neuron.2019.01.016.
- Rahman, R.-U. et al. (2018) 'Oasis 2: improved online analysis of small RNA-seq data', *BMC Bioinformatics*, 19(1), p. 54. doi: 10.1186/s12859-018-2047-z.
- Ruso-Julve, F. et al. (2019) 'Dopaminergic control of ADAMTS2 expression through cAMP/CREB and ERK: molecular effects of antipsychotics', *Translational Psychiatry*, 9(1), p. 306. doi: 10.1038/s41398-019-0647-7.
- Sanchez-Roige, S. et al. (2019) 'Genome-wide association studies of impulsive personality traits (BIS-11 and UPPS-P) and drug experimentation in up to 22,861 adult research participants identify loci in the CACNA11 and CADM2 genes', *Journal of Neuroscience*, 39(13), pp. 2562–2572. doi: 10.1523/JNEUROSCI.2662-18.2019.
- Sarnyai, Z. and Kovács, G. L. (1994) 'Role of oxytocin in the neuroadaptation to drugs of abuse', *Psychoneuroendocrinology*, 19(1), pp. 85–117. doi: 10.1016/0306-4530(94)90062-0.
- Sarnyai, Z. and Kovács, G. L. (2014) 'Oxytocin in learning and addiction: From early discoveries to the present', *Pharmacology Biochemistry and Behavior*, 119, pp. 3–9. doi: 10.1016/j.pbb.2013.11.019.

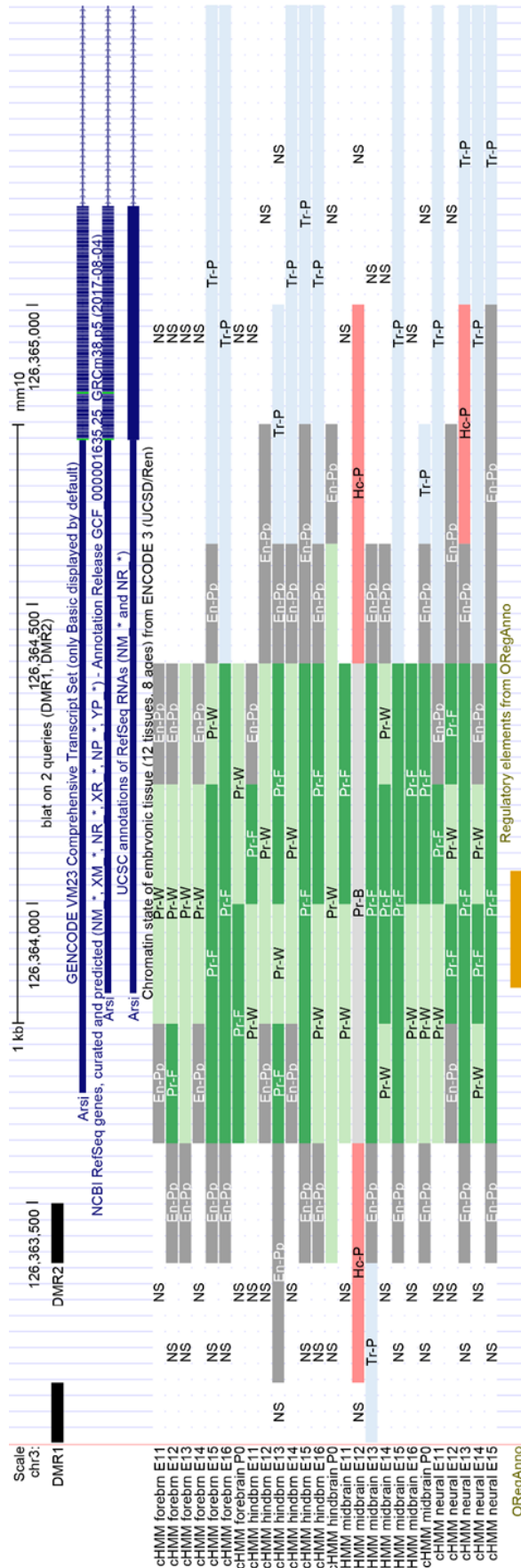
- Sayols-Baixeras, S. et al. (2017) 'DNA methylation and obesity traits: An epigenome-wide association study. The REGICOR study', *Epigenetics*, 12(10), pp. 909–916. doi: 10.1080/15592294.2017.1363951.
- Svahn, A. J. et al. (2016) 'miR-124 Contributes to the functional maturity of microglia', *Developmental neurobiology*, 76(5), pp. 507–518. doi: 10.1002/dneu.22328.
- Veremeyko, T. et al. (2019) 'Neuronal extracellular microRNAs miR-124 and miR-9 mediate cell–cell communication between neurons and microglia', *Journal of Neuroscience Research*, 97(2), pp. 162–184. doi: 10.1002/jnr.24344.
- Viola, T. W. et al. (2019) 'Peripheral blood microRNA levels in females with cocaine use disorder', *Journal of Psychiatric Research*, 114, pp. 48–54. doi: 10.1016/j.jpsychires.2019.03.028.
- Volkow, N. D. et al. (2013) 'Obesity and addiction: Neurobiological overlaps', *Obesity Reviews*, 14(1), pp. 2–18. doi: 10.1111/j.1467-789X.2012.01031.x.
- Volkow, N. D., Wise, R. A. and Baler, R. (2017) 'The dopamine motive system: Implications for drug and food addiction', *Nature Reviews Neuroscience*, 18(12), pp. 741–752. doi: 10.1038/nrn.2017.130gvf.
- Walker, D. M. and Nestler, E. J. (2018) 'Neuroepigenetics and addiction', in Geschwind, D. H., Paulson, H. L., and Klein, C. (eds) *Handbook of Clinical Neurology. Neurogenetics, Part II*, pp. 747–765. doi: 10.1016/B978-0-444-64076-5.00048-X.
- Wang, J. et al. (2017) 'WebGestalt 2017: A more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit', *Nucleic Acids Research*, 45(W1), pp. W130–W137. doi: 10.1093/nar/gkx356.
- World Health Organization (2020) Obesity and overweight Fact Sheet. Available at: <https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight> (Accessed: 12 June 2020).
- Zhou, Z., Enoch, M. A. and Goldman, D. (2014) 'Gene expression in the addicted brain', in Hitzemann, R. and Mcweaney, S. (eds) *International Review of Neurobiology. Brain Transcriptome*, pp. 251–273. doi: 10.1016/B978-0-12-801105-8.00010-2.

Supplementary Material

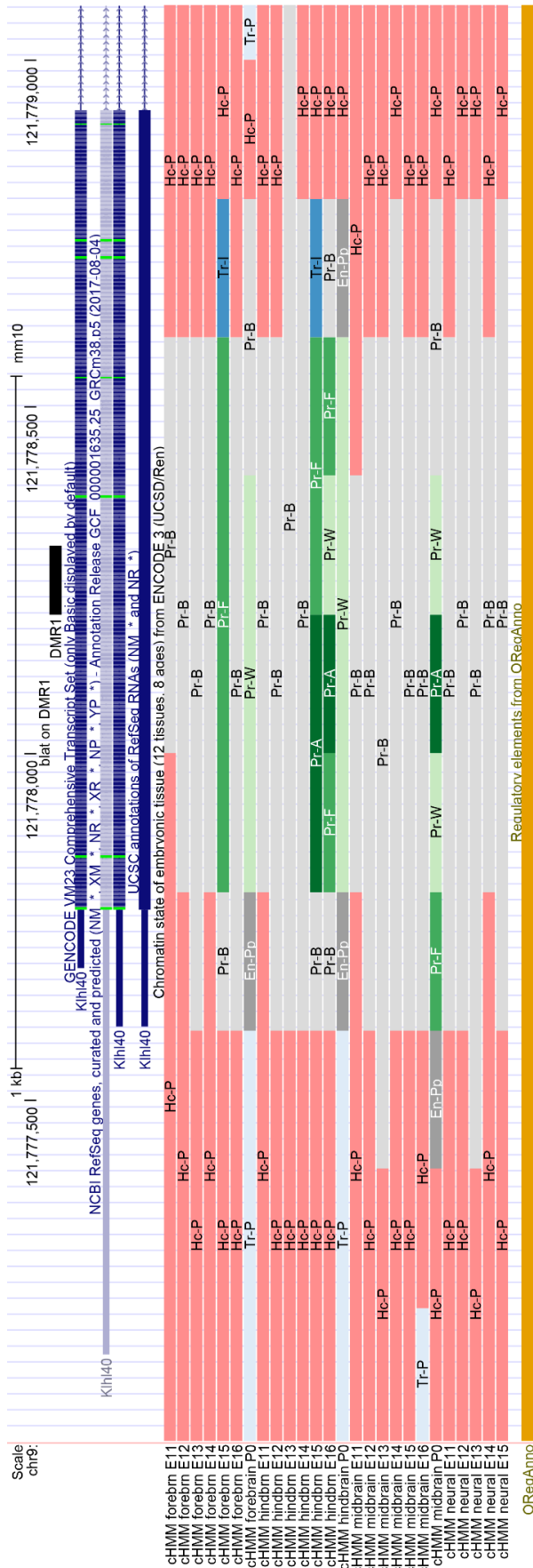
GENE EXPRESSION AND EPIGENETIC PROFILING OF A MOUSE MODEL OF FOOD ADDICTION



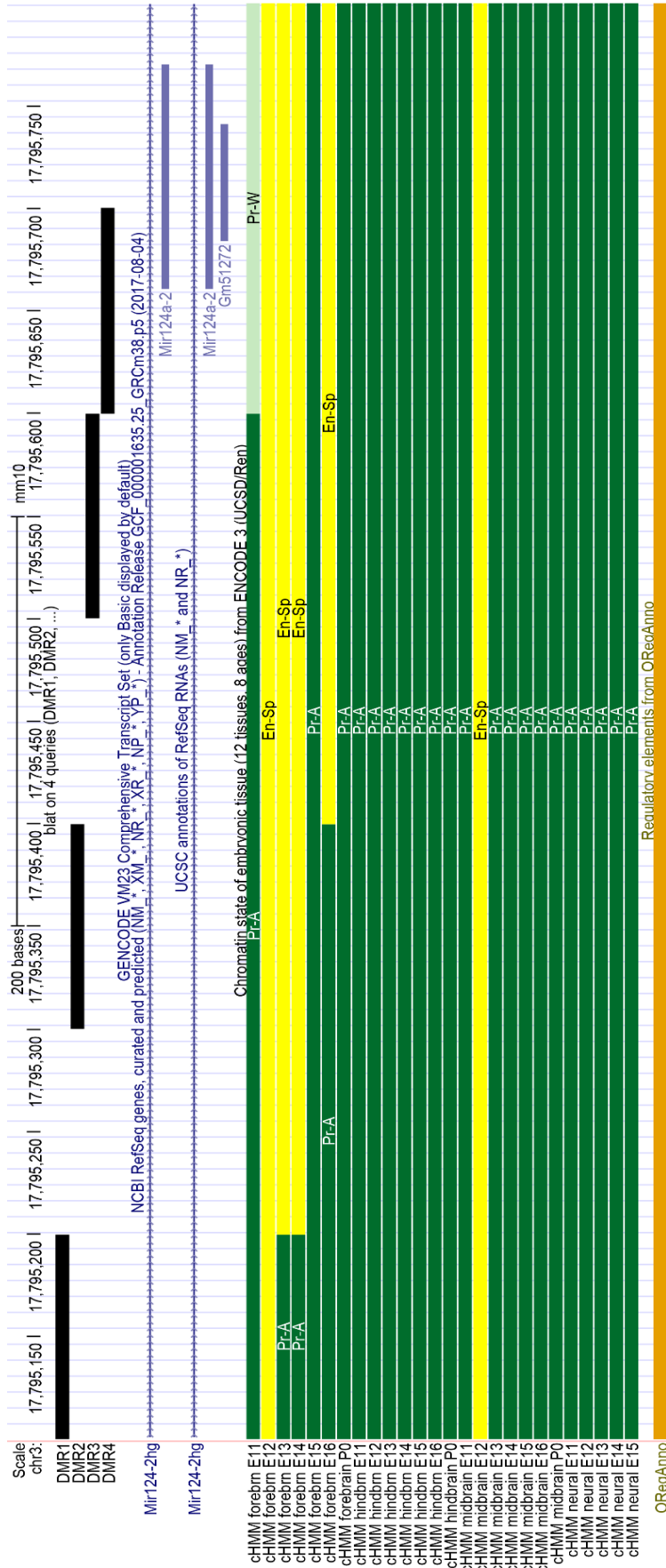
Supplementary Figure 1. Distribution scheme of the selected mice for the discovery (D) and replica (R) samples. The Gaussian curve represent the population of animals distributed depending on the score obtained for the three criteria of food addiction (motivation, persistence to response and compulsivity). Each animal selected is represented by a circle, which correspond to the extremes of the population. The dashed lines indicate the threshold for resilience or vulnerability to food addiction-like phenotype. Green circles: animals resilient to addiction-like phenotype (0 criterion); Orange circles: animals vulnerable to addiction-like phenotype that achieved two or more addiction criteria; Red circles: animals vulnerable to addiction-like phenotype that achieved three addiction criteria.



Supplementary Figure 3. *Arsj* genomic context from UCSC genome browser (mouse GRCh38/mm10). Black bars indicate the differentially methylated regions (DMR1: chr3:126363101-126363200 and DMR2: chr3:126363401-126363500). The tracks displayed are the GENCODE VM23, NCBI RefSeq (RefSeq All and UCSC RefSeq), Chromatin state of embryonic tissue from Encode Regulation (only brain regions) and Regulatory elements from ORegAnno. Pr-W: Promoter Weak; Pr-B: Promoter Bivalent; Pr-F: Promoter Flanking Region; En-Pp: Enhancer Poised TSS-proximal; Tr-P: Transcription Permissive; Hc-P: Heterochromatin Polycomb-associated; Ns: no significant signal. The orange line from the ORegAnno track indicates a transcription factor binding site.



Supplementary Figure 4. *Klfh40* genomic context from UCSC genome browser (mouse GRCh38/mm10). Black bars indicate the differentially methylated regions (DMR1: chr9:121778201-121778300). The tracks displayed are the GENCODE VM23, NCBI RefSeq (RefSeq All and UCSC RefSeq), Chromatin state of embryonic tissue from Encode Regulation (only brain regions) and Regulatory elements from ORegAnno. Pr-A: Promoter active; Pr-W: Promoter Weak; Pr-B: Promoter Bivalent; Pr-F: Promoter Flanking Region; En-Pp: Enhancer Poised TSS-proximal; Tr-P: Transcription Permissive; Tr-I: Transcription Initiation; Hc-P: Heterochromatin Polycomb-associated; Ns: no significant signal. The orange line from the ORegAnno track indicates a transcription factor binding site.



Supplementary Figure 5. miRNA124 genomic context from UCSC genome browser (mouse GRCm38/mm10). Black bars indicate the differentially methylated regions (DMR1: chr3: 17795101-17795200, DMR2: chr3: 17795301-17795400, DMR3: chr3: 17795501-17795600 and DMR4: chr3: 17795601-17795700). The tracks displayed are the GENCODE VM23, NCBI RefSeq (RefSeq All and UCSC RefSeq), Chromatin state of embryonic tissue from Encode Regulation (only brain regions) and Regulatory elements from ORegAnno. Pr-A: Promoter active; En-Sp: Enhancer Strong TSS-proximal. The orange line from the ORegAnno track indicates a transcription factor binding site.

Supplementary Table 1. Protein coding genes differentially expressed in mPFC and NAc

ID	Gene symbol	Discovery			Replica		
		Expression	FC	p-value	Expression	FC	p-value
mPFC							
ENSMUSG00000109588	<i>Lnp1</i>	2.88E+01	-2.98	2.43E-03	1.02E+01	-3.25	1.08E-02
ENSMUSG00000027971	<i>Ndst4</i>	1.95E+03	-2.32	3.62E-02	7.34E+02	-3.11	1.00E-02
ENSMUSG00000019865	<i>Nmbr</i>	2.80E+02	-1.79	2.14E-02	9.29E+01	-1.95	2.12E-03
ENSMUSG00000043659	<i>Npsr1</i>	3.26E+02	-2.39	2.07E-04	9.95E+01	-1.72	2.14E-04
ENSMUSG00000049001	<i>Ndnf</i>	4.16E+03	-1.25	3.56E-02	1.43E+03	-1.63	4.12E-03
ENSMUSG00000028017	<i>Egf</i>	8.75E+01	-1.28	3.55E-02	3.24E+01	-1.55	5.87E-03
ENSMUSG00000032776	<i>Mctp2</i>	1.16E+02	-1.40	4.83E-02	4.52E+01	-1.53	5.58E-03
ENSMUSG00000038352	<i>Arl5c</i>	8.21E+01	-1.36	9.28E-03	3.22E+01	-1.50	1.17E-02
ENSMUSG00000042514	<i>Klhl14</i>	2.32E+02	-1.85	1.71E-02	6.28E+01	-1.48	2.89E-02
ENSMUSG00000019874	<i>Fabp7</i>	8.23E+02	-1.25	4.43E-02	2.57E+02	-1.47	2.25E-02
ENSMUSG00000049112	<i>Oxtr</i>	8.82E+02	-1.45	2.74E-02	2.70E+02	-1.47	2.68E-02
ENSMUSG00000037469	<i>Papl</i>	9.92E+01	-1.52	1.01E-02	3.86E+01	-1.41	1.66E-02
ENSMUSG00000044071	<i>Fam19a2</i>	3.50E+03	-1.27	7.36E-03	1.16E+03	-1.36	3.44E-03
ENSMUSG00000043668	<i>Tox3</i>	1.53E+03	-1.25	4.24E-03	4.94E+02	-1.31	3.03E-03
ENSMUSG00000044816	<i>AC101869.3</i>	3.42E+02	-1.22	2.52E-03	1.21E+02	-1.30	1.29E-03
ENSMUSG00000018126	<i>Baiap2l2</i>	2.45E+02	-1.22	1.49E-02	7.06E+01	-1.29	2.94E-02
ENSMUSG00000035131	<i>Brinp3</i>	2.07E+03	-1.22	3.66E-02	6.17E+02	-1.28	2.44E-02
ENSMUSG00000031119	<i>Gpc4</i>	1.79E+03	-1.22	2.95E-02	4.98E+02	-1.27	2.26E-02
ENSMUSG00000040061	<i>Plcb2</i>	3.42E+02	-1.17	4.24E-02	1.09E+02	-1.24	1.71E-02
ENSMUSG00000055633	<i>Zfp580</i>	1.53E+03	-1.11	1.99E-02	3.82E+02	-1.20	4.03E-04
ENSMUSG00000030092	<i>Cntn6</i>	1.73E+03	-1.25	1.43E-02	5.69E+02	-1.19	2.05E-02
ENSMUSG00000039179	<i>Tekt5</i>	8.65E+02	-1.16	3.38E-02	2.62E+02	-1.17	3.87E-02
ENSMUSG00000093803	<i>Ppp2r3d</i>	1.61E+03	-1.10	5.14E-03	4.69E+02	-1.12	9.34E-03
ENSMUSG00000038181	<i>Chpf2</i>	5.57E+03	-1.14	2.94E-02	1.70E+03	-1.12	3.84E-03
ENSMUSG00000029344	<i>Tpst2</i>	2.15E+03	-1.10	3.90E-02	5.72E+02	-1.11	3.39E-02
ENSMUSG00000000823	<i>Zfp512b</i>	6.36E+03	-1.13	1.78E-02	2.19E+03	-1.11	4.65E-02
ENSMUSG00000069631	<i>Strada</i>	1.81E+03	-1.12	1.04E-02	6.19E+02	-1.10	2.12E-02
ENSMUSG00000022540	<i>Rogdi</i>	8.51E+03	-1.09	2.39E-02	2.73E+03	-1.09	1.11E-03
ENSMUSG00000048503	<i>Tmem136</i>	1.52E+03	-1.10	2.80E-02	5.35E+02	-1.09	2.87E-02
ENSMUSG00000037570	<i>Mcrs1</i>	5.52E+03	-1.11	1.22E-02	1.61E+03	-1.08	2.64E-02
ENSMUSG00000004933	<i>Matk</i>	1.01E+04	-1.09	3.20E-02	3.70E+03	-1.08	5.11E-03
ENSMUSG00000040415	<i>Dtx3</i>	1.39E+04	-1.06	1.65E-02	4.46E+03	-1.07	2.39E-03
ENSMUSG00000025158	<i>Rfng</i>	6.19E+03	-1.07	4.93E-02	1.79E+03	-1.06	2.01E-02
ENSMUSG00000018199	<i>Trove2</i>	9.45E+03	1.08	1.26E-02	3.04E+03	1.06	3.82E-02
ENSMUSG00000021188	<i>Trip11</i>	6.00E+03	1.07	1.43E-02	1.89E+03	1.08	1.32E-02
ENSMUSG00000031666	<i>Rbl2</i>	5.93E+03	1.09	3.01E-02	1.76E+03	1.08	3.74E-02
ENSMUSG00000038375	<i>Trp53inp2</i>	1.77E+04	1.13	3.91E-03	4.90E+03	1.11	3.69E-02
ENSMUSG00000075232	<i>Amd1</i>	1.37E+04	1.10	1.57E-02	4.30E+03	1.12	2.01E-03
ENSMUSG00000054619	<i>Mettl7a1</i>	2.77E+03	1.23	1.96E-03	8.33E+02	1.16	3.29E-02
ENSMUSG00000078866	<i>RP23-67E6.5</i>	8.84E+02	1.15	4.48E-02	3.13E+02	1.17	1.92E-02
ENSMUSG00000022425	<i>Enpp2</i>	1.85E+04	1.19	1.73E-02	5.42E+03	1.20	3.24E-02
ENSMUSG00000042834	<i>Nrep</i>	3.60E+03	1.23	3.71E-02	1.11E+03	1.23	3.00E-02
ENSMUSG00000026475	<i>Rgs16</i>	8.86E+02	1.18	2.93E-02	2.67E+02	1.24	1.79E-02
ENSMUSG00000072568	<i>Fam84b</i>	1.50E+03	1.18	4.28E-02	4.34E+02	1.29	4.13E-02
ENSMUSG00000032265	<i>Fam46a</i>	6.43E+02	1.34	1.10E-02	2.09E+02	1.30	4.95E-02

NAC							
ENSMUSG00000046561	<i>Arsj</i>	1.21E+02	-1.86	4.55E-02	4.26E+01	-1.85	3.80E-02
ENSMUSG00000049744	<i>Arhgap15</i>	2.17E+02	-2.33	3.57E-02	8.04E+01	-1.78	4.87E-02
ENSMUSG00000035258	<i>Abi3bp</i>	2.29E+03	-3.49	2.99E-03	7.06E+02	-1.71	4.53E-02
ENSMUSG00000034164	<i>Emid1</i>	3.66E+02	-1.76	3.00E-02	1.23E+02	-1.68	3.90E-04
ENSMUSG00000022468	<i>Endou</i>	7.40E+01	-1.55	3.03E-02	2.29E+01	-1.63	2.41E-02
ENSMUSG00000039405	<i>Prss23</i>	9.90E+02	-1.83	4.70E-02	3.47E+02	-1.60	1.46E-02
ENSMUSG00000074001	<i>Klhl40</i>	1.67E+02	-1.78	7.55E-03	5.13E+01	-1.49	2.31E-02
ENSMUSG00000035513	<i>Ntng2</i>	1.94E+03	-1.60	4.47E-02	5.93E+02	-1.48	1.63E-02
ENSMUSG00000071714	<i>Csf2rb2</i>	7.34E+01	-1.29	2.37E-02	2.41E+01	-1.47	2.56E-02
ENSMUSG00000022372	<i>Sla</i>	4.99E+02	-1.74	4.23E-02	1.76E+02	-1.45	1.78E-02
ENSMUSG00000068037	<i>Mas1</i>	6.38E+02	-1.88	4.26E-02	2.21E+02	-1.39	2.64E-02
ENSMUSG00000029055	<i>Plch2</i>	4.84E+03	-1.52	4.60E-02	1.63E+03	-1.36	4.42E-02
ENSMUSG00000035357	<i>Pdzrn3</i>	8.22E+02	-2.36	1.47E-02	2.37E+02	-1.35	4.71E-02
ENSMUSG00000050860	<i>Phospho1</i>	2.56E+03	-1.54	4.17E-02	7.61E+02	-1.34	5.87E-03
ENSMUSG00000050211	<i>Pla2g4e</i>	3.93E+02	-1.45	3.86E-03	1.35E+02	-1.34	2.32E-02
ENSMUSG00000029544	<i>Cabp1</i>	6.94E+03	-1.64	3.86E-02	2.03E+03	-1.32	3.01E-02
ENSMUSG00000039116	<i>Adgrg6</i>	4.95E+02	-1.85	1.57E-02	1.70E+02	-1.32	4.62E-02
ENSMUSG00000036545	<i>Adamts2</i>	6.68E+02	-1.35	7.41E-03	2.30E+02	-1.30	4.87E-02
ENSMUSG00000033327	<i>Tnxb</i>	4.71E+02	-1.18	4.51E-02	1.88E+02	-1.29	4.89E-02
ENSMUSG00000040258	<i>Nxph4</i>	5.61E+02	-1.75	1.88E-02	1.63E+02	-1.28	4.93E-02
ENSMUSG00000028125	<i>Abca4</i>	2.99E+02	-1.33	3.21E-02	1.17E+02	-1.28	4.90E-02
ENSMUSG00000043822	<i>Adamts15</i>	5.26E+02	-1.33	7.30E-03	1.52E+02	-1.28	6.91E-04
ENSMUSG00000023192	<i>Grm2</i>	7.87E+02	-1.40	1.75E-02	2.49E+02	-1.27	2.50E-02
ENSMUSG00000040998	<i>Npnt</i>	7.92E+02	-1.89	1.25E-02	2.53E+02	-1.25	4.78E-02
ENSMUSG00000021257	<i>Angel1</i>	7.92E+02	-1.30	2.84E-02	2.52E+02	-1.25	1.85E-02
ENSMUSG00000042532	<i>Golga7b</i>	5.36E+03	-1.41	4.82E-02	1.60E+03	-1.23	1.96E-02
ENSMUSG00000031883	<i>Car7</i>	5.36E+02	-1.35	1.53E-02	1.89E+02	-1.23	1.57E-02
ENSMUSG00000069171	<i>Nr2f1</i>	2.33E+03	-1.45	2.74E-02	6.82E+02	-1.22	4.73E-02
ENSMUSG00000046417	<i>Lrrc75a</i>	8.23E+02	-1.28	2.52E-02	2.67E+02	-1.22	1.00E-02
ENSMUSG00000026688	<i>Mgst3</i>	2.98E+03	-1.36	1.17E-02	9.37E+02	-1.21	4.17E-03
ENSMUSG00000024968	<i>Rcor2</i>	9.01E+02	-1.31	3.20E-02	2.91E+02	-1.21	3.30E-02
ENSMUSG00000002068	<i>Ccne1</i>	9.63E+02	-1.38	1.48E-02	3.31E+02	-1.21	1.43E-02
ENSMUSG00000091735	<i>Gpr62</i>	1.05E+03	-1.11	4.13E-02	3.16E+02	-1.19	1.37E-02
ENSMUSG00000053279	<i>Aldh1a1</i>	2.00E+03	-1.30	2.56E-02	7.60E+02	-1.19	4.66E-02
ENSMUSG00000096847	<i>Tmem151b</i>	1.07E+04	-1.23	2.04E-02	3.24E+03	-1.18	1.27E-02
ENSMUSG00000095407	<i>Tmem200c</i>	1.35E+03	-1.20	3.86E-02	4.17E+02	-1.17	2.71E-02
ENSMUSG00000015745	<i>Plekho1</i>	1.27E+03	-1.31	2.05E-02	3.89E+02	-1.17	2.83E-02
ENSMUSG00000042797	<i>Aqp11</i>	1.03E+03	-1.22	3.56E-02	3.37E+02	-1.17	4.72E-02
ENSMUSG00000043456	<i>Zfp536</i>	4.05E+03	-1.38	2.72E-02	1.50E+03	-1.16	3.33E-02
ENSMUSG00000032172	<i>Olfm2</i>	5.01E+03	-1.24	2.52E-02	1.58E+03	-1.16	4.32E-02
ENSMUSG00000069806	<i>Cacng7</i>	1.48E+04	-1.23	3.03E-02	4.56E+03	-1.15	3.12E-03
ENSMUSG00000058441	<i>Panx2</i>	9.87E+03	-1.13	4.90E-02	2.92E+03	-1.14	5.45E-03
ENSMUSG00000022537	<i>Tmem44</i>	4.64E+03	-1.17	4.40E-02	1.61E+03	-1.14	9.01E-03
ENSMUSG00000034891	<i>Sncb</i>	2.66E+04	-1.22	3.65E-02	7.90E+03	-1.13	4.82E-02
ENSMUSG00000045994	<i>B3gat1</i>	1.68E+04	-1.23	1.23E-02	5.56E+03	-1.13	2.60E-02
ENSMUSG00000055435	<i>Maf</i>	2.28E+03	-1.14	4.16E-02	7.26E+02	-1.13	1.28E-02
ENSMUSG00000029608	<i>Rph3a</i>	2.39E+04	-1.23	3.04E-02	8.12E+03	-1.13	3.71E-02
ENSMUSG00000038453	<i>Srcin1</i>	3.79E+04	-1.16	3.93E-02	1.17E+04	-1.12	1.39E-02
ENSMUSG00000032175	<i>Tyk2</i>	1.34E+03	-1.11	1.28E-02	4.68E+02	-1.12	2.74E-02

Resultats capítol 2. Article 5

ENSMUSG0000001034	<i>Mapk7</i>	1.34E+03	-1.09	3.41E-02	4.48E+02	-1.12	9.87E-03
ENSMUSG00000039936	<i>Pik3cd</i>	2.04E+03	-1.16	2.41E-02	6.71E+02	-1.11	3.59E-02
ENSMUSG00000021219	<i>Rgs6</i>	2.38E+03	-1.17	3.83E-02	7.96E+02	-1.11	1.38E-02
ENSMUSG00000027674	<i>Pex5l</i>	1.35E+04	-1.23	1.11E-02	4.44E+03	-1.09	3.35E-02
ENSMUSG00000032329	<i>Hmg20a</i>	8.27E+03	1.05	4.34E-02	2.91E+03	1.06	4.00E-02
ENSMUSG00000038984	<i>Tspyl5</i>	5.48E+03	1.08	2.12E-02	1.91E+03	1.07	3.29E-02
ENSMUSG00000019777	<i>Hdac2</i>	6.86E+03	1.05	4.86E-02	2.47E+03	1.07	4.36E-02
ENSMUSG00000039879	<i>Heca</i>	3.49E+03	1.04	4.87E-02	1.20E+03	1.07	3.92E-02
ENSMUSG00000029571	<i>Tmem106b</i>	2.41E+04	1.06	2.27E-02	8.86E+03	1.07	3.36E-02
ENSMUSG00000040570	<i>Rundc3b</i>	5.25E+03	1.09	3.62E-02	1.86E+03	1.07	3.83E-02
ENSMUSG00000066643	<i>Wdr35</i>	4.56E+03	1.07	9.85E-03	1.60E+03	1.07	1.98E-02
ENSMUSG00000032216	<i>Nedd4</i>	6.05E+04	1.09	1.96E-02	2.14E+04	1.08	3.66E-03
ENSMUSG00000040550	<i>Otud6b</i>	6.80E+03	1.09	2.74E-03	2.45E+03	1.08	1.59E-02
ENSMUSG00000022016	<i>Akap11</i>	4.52E+04	1.05	4.33E-02	1.60E+04	1.08	1.40E-02
ENSMUSG00000027708	<i>Dcun1d1</i>	5.32E+03	1.05	3.72E-02	1.94E+03	1.08	3.08E-02
ENSMUSG00000060373	<i>Hnrnpc</i>	1.51E+04	1.09	4.05E-02	5.63E+03	1.09	4.68E-02
ENSMUSG00000021840	<i>Mapk1ip1l</i>	5.41E+03	1.06	2.96E-02	1.87E+03	1.09	8.24E-03
ENSMUSG00000040928	<i>S100pbp</i>	7.41E+03	1.09	3.98E-02	2.63E+03	1.09	2.20E-02
ENSMUSG00000009406	<i>Elk1</i>	6.22E+03	1.08	3.67E-02	2.09E+03	1.09	3.63E-02
ENSMUSG00000048799	<i>Cep120</i>	6.76E+03	1.07	4.03E-02	2.43E+03	1.09	1.04E-02
ENSMUSG00000022403	<i>St13</i>	1.58E+04	1.07	3.75E-02	5.62E+03	1.09	2.41E-02
ENSMUSG00000071359	<i>Tbpl1</i>	5.46E+03	1.06	3.25E-02	1.89E+03	1.09	4.63E-03
ENSMUSG00000024182	<i>Axin1</i>	2.78E+03	1.07	6.41E-03	9.43E+02	1.09	9.22E-03
ENSMUSG00000027122	<i>Arl14ep</i>	4.03E+03	1.07	4.35E-02	1.48E+03	1.09	3.06E-02
ENSMUSG00000059263	<i>Usp47</i>	1.03E+04	1.06	3.35E-02	3.75E+03	1.09	3.88E-04
ENSMUSG00000019979	<i>Apaf1</i>	1.75E+03	1.07	3.56E-02	6.21E+02	1.09	4.10E-02
ENSMUSG00000032745	<i>Gbbp1</i>	9.86E+03	1.05	2.83E-02	3.61E+03	1.09	6.75E-03
ENSMUSG00000042426	<i>Dhx29</i>	3.32E+03	1.08	4.61E-02	1.15E+03	1.10	1.47E-02
ENSMUSG00000000266	<i>Mid2</i>	5.89E+03	1.08	4.98E-02	2.10E+03	1.10	5.89E-04
ENSMUSG00000059482	<i>RP23-144K18.5</i>	3.01E+03	1.08	4.78E-02	1.12E+03	1.10	4.78E-03
ENSMUSG00000037172	<i>RP23-83N17.1</i>	1.15E+04	1.08	1.62E-02	4.25E+03	1.10	4.85E-02
ENSMUSG00000049281	<i>Scn3b</i>	1.41E+04	1.08	2.50E-02	5.18E+03	1.10	1.46E-02
ENSMUSG00000022814	<i>Umps</i>	1.46E+03	1.09	4.19E-02	5.12E+02	1.10	3.06E-02
ENSMUSG00000050587	<i>Lrrc4c</i>	1.33E+04	1.08	7.74E-03	4.57E+03	1.10	2.29E-02
ENSMUSG00000026771	<i>Spopl</i>	2.18E+03	1.07	5.37E-03	8.62E+02	1.11	2.29E-02
ENSMUSG00000005034	<i>Prkacb</i>	6.01E+04	1.12	4.72E-02	2.13E+04	1.11	5.42E-03
ENSMUSG00000026361	<i>Cdc73</i>	6.63E+03	1.05	3.96E-02	2.49E+03	1.11	3.98E-03
ENSMUSG00000035620	<i>Ric8b</i>	1.03E+04	1.13	4.25E-02	3.73E+03	1.11	3.20E-02
ENSMUSG00000072964	<i>Bhlhb9</i>	7.78E+03	1.10	1.84E-02	2.95E+03	1.11	1.95E-02
ENSMUSG00000027763	<i>Mbnl1</i>	1.45E+04	1.09	2.91E-02	5.44E+03	1.11	2.27E-02
ENSMUSG00000034021	<i>Pds5b</i>	1.68E+04	1.09	2.70E-02	6.19E+03	1.11	2.98E-03
ENSMUSG00000021779	<i>Thrb</i>	6.27E+03	1.13	1.53E-02	2.11E+03	1.11	4.30E-02
ENSMUSG00000026787	<i>Gad2</i>	8.99E+04	1.20	2.07E-02	3.22E+04	1.11	2.21E-02
ENSMUSG00000039680	<i>Mrps6</i>	1.62E+03	1.13	4.23E-02	5.47E+02	1.12	4.91E-02
ENSMUSG00000048701	<i>Ccdc6</i>	8.23E+03	1.14	1.64E-02	2.86E+03	1.12	4.09E-03
ENSMUSG00000055026	<i>Gabrg3</i>	8.25E+03	1.13	2.66E-02	3.00E+03	1.12	1.69E-02
ENSMUSG00000041216	<i>Clvs1</i>	7.00E+03	1.10	2.90E-02	2.48E+03	1.13	2.46E-03
ENSMUSG00000032475	<i>Nck1</i>	1.74E+03	1.09	2.74E-02	6.63E+02	1.13	1.65E-02
ENSMUSG00000036712	<i>Cyld</i>	3.18E+04	1.16	7.03E-03	1.14E+04	1.13	3.51E-02
ENSMUSG00000058897	<i>Col25a1</i>	8.73E+03	1.14	4.63E-02	3.26E+03	1.13	6.60E-03

ENSMUSG00000036006	<i>Fam65b</i>	1.28E+04	1.15	2.91E-02	4.26E+03	1.13	2.60E-02
ENSMUSG00000025812	<i>Pard3</i>	5.23E+03	1.17	1.52E-02	1.75E+03	1.13	3.43E-02
ENSMUSG00000027351	<i>Spred1</i>	2.30E+04	1.09	4.09E-02	8.39E+03	1.14	8.67E-03
ENSMUSG00000043241	<i>Upf2</i>	6.41E+03	1.10	4.21E-02	2.47E+03	1.14	1.14E-02
ENSMUSG00000038718	<i>Pbx3</i>	6.14E+03	1.18	4.38E-02	2.05E+03	1.14	3.79E-02
ENSMUSG00000039706	<i>Ldb2</i>	2.32E+03	1.20	4.60E-02	8.19E+02	1.14	1.74E-02
ENSMUSG00000034480	<i>Diaph2</i>	6.66E+03	1.10	4.04E-02	2.55E+03	1.14	4.25E-02
ENSMUSG00000024642	<i>Tle4</i>	1.12E+04	1.15	1.74E-02	3.74E+03	1.14	1.42E-02
ENSMUSG00000022722	<i>Arl6</i>	2.06E+03	1.10	2.51E-02	7.81E+02	1.15	3.25E-03
ENSMUSG00000024491	<i>Rbm27</i>	5.84E+03	1.07	2.14E-02	2.10E+03	1.15	2.82E-03
ENSMUSG00000020134	<i>Peli1</i>	5.61E+03	1.13	1.99E-02	2.04E+03	1.16	1.73E-02
ENSMUSG00000056014	<i>WI1-955B2.1</i>	1.74E+03	1.10	1.09E-02	6.32E+02	1.16	8.86E-04
ENSMUSG00000032009	<i>Sesn3</i>	4.78E+04	1.13	3.67E-02	1.68E+04	1.16	5.75E-03
ENSMUSG00000026360	<i>Rgs2</i>	8.19E+03	1.17	2.49E-02	2.77E+03	1.16	6.37E-03
ENSMUSG00000028035	<i>Dnajb4</i>	6.95E+03	1.11	4.51E-02	2.64E+03	1.17	1.35E-03
ENSMUSG00000036510	<i>Cdh8</i>	6.05E+03	1.19	2.57E-02	2.14E+03	1.18	1.21E-03
ENSMUSG00000042604	<i>Kcna4</i>	6.28E+03	1.11	3.43E-02	2.17E+03	1.18	4.20E-02
ENSMUSG00000089756	<i>RP23-355G16.6</i>	1.12E+03	1.22	2.44E-02	5.57E+02	1.20	2.26E-03
ENSMUSG00000035021	<i>Baz1a</i>	4.67E+02	1.26	3.13E-02	1.90E+02	1.28	3.17E-02
ENSMUSG00000049511	<i>Htr1b</i>	5.29E+03	1.27	3.36E-02	1.92E+03	1.31	2.88E-02
ENSMUSG00000035578	<i>lqcg</i>	3.48E+02	1.22	3.52E-02	1.17E+02	1.34	2.28E-02
ENSMUSG00000022805	<i>Maats1</i>	1.56E+02	1.54	3.33E-02	5.27E+01	1.44	3.93E-02

FC: Fold change

Supplementary Table 2. GWAS catalog information for the differentially expressed genes in mPFC and NAc considering psychiatric conditions, cognitive phenotypes, body mass index (BMI) or metabolic traits.

Gene	Variant and risk allele	Location	P-value	Mapped gene	Reported trait	Trait(s)	Study accession	
mPFC	NDST4	rs11725345-C	5E-09	NDST4	Self-reported math ability	mathematical ability	GCST006573	
		rs35134647-A	4E-10	MRRPS33P3, NDST4	Highest math class taken (MITAG)	mathematical ability	GCST006568	
		rs6818100-A	2E-09	MRRPS33P3, NDST4	Self-reported math ability (MITAG)	mathematical ability	GCST006569	
	NMBR		rs561736395-A	7E-06	NMBR, RPS3AP23	Waist-hip ratio	waist-hip ratio	GCST008155
			rs7770810-?	2E-08	RPS3AP23, NMBR	Body mass index	body mass index	GCST006369
			rs1577948-?	2E-08	RPS3AP23, NMBR	Body mass index	body mass index	GCST006369
			rs1572604-?	2E-08	NMBR, RPS3AP23	Body mass index	body mass index	GCST006369
		rs1572605-?	2E-08	RPS3AP23, NMBR	Body mass index	body mass index	GCST006369	
		rs17054002-?	1E-07	NMBR, RPS3AP23	Body mass index	body mass index	GCST006369	
	NDNF		rs17071124-G	3E-06	NMBR, RPS3AP23	Sleep quality	sleep quality	GCST002049
		rs10018902-A	4E-09	PRDM5, NDNF	Obesity-related traits	body weight	GCST001762	
		rs40045992-T	4E-06	ELOVL6, EGF	Waist-to-hip ratio adjusted for BMI	BMI-adjusted waist-hip ratio	GCST008159	
		rs10012953-G	3E-06	EGF, ELOVL6	Obesity-related traits	adiponectin measurement	GCST001762	
MCTP2		rs8093195-A	5E-07	MCTP2	Hippocampal volume	hippocampal volume	GCST002761	
		rs11718289-T	3E-07	CAV3, OXTR	Caudate activity during reward	caudate nucleus measurement, functional brain measurement, response to reward	GCST004970	
PAPL		rs237899-?	3E-06	OXTR, CAV3	Depression and alcohol dependence	alcohol dependence	GCST001313	
		rs472265-G	9E-06	ACP7	Type 2 diabetes	Type II diabetes mellitus	GCST001033	
FAM19A2		rs144026674-T	19E-09	ACP7	Intelligence	intelligence	GCST006250	
		rs2198776-A	12E-06	TAF42	Obesity-related traits	body mass index	GCST001762	
		rs2198776-A	12E-06	TAF42	Obesity-related traits	body weight	GCST001762	
		rs11174202-A	12E-06	TAF42	Anorexia nervosa	anorexia nervosa	GCST004367	
		rs11174202-?	12E-06	TAF42	Anorexia nervosa or obsessive-compulsive disorder	anorexia nervosa, obsessive-compulsive disorder	GCST006576	
		rs3759289-?	12E-06	TAF42	Body mass index	body mass index	GCST007039	
		rs117359675-A	4E-06	TAF42	Body mass index	body mass index	GCST008158	
		rs117359675-A	2E-06	TAF42	Weight	body weight	GCST008152	
		rs349885-?	12E-06	TAF42	Reaction time	reaction time measurement	GCST006268	
		rs1146079-C	12E-06	TAF42	Educational attainment (years of education)	self reported educational attainment	GCST006442	
TOX3		rs141480969-T	12E-06	TAF42	Hippocampal volume	hippocampal volume	GCST002761	
		rs1348526-T	12E-06	ACO78789.1, TAF42	Educational attainment (years of education)	self reported educational attainment	GCST006571	
		rs11174399-A	12E-06	TAF42	Educational attainment (years of education)	self reported educational attainment	GCST006442	
		rs12921005-?	16E-05	CASC22, TOX3	Educational attainment (years of education)	self reported educational attainment	GCST007037	
		rs12921005-A	16E-05	CASC22, TOX3	Educational attainment (years of education)	self reported educational attainment	GCST006442	
		rs993638-A	16E-05	TOX3	Neuroticism	neuroticism measurement	GCST006476	
		rs1362546-?	16E-05	TOX3	Neuroticism	neuroticism measurement	GCST007084	
		rs4784224-A	16E-05	TOX3	Neuroticism	neuroticism measurement	GCST007339	
		rs142739979-?	22E-08	BAIAP2L2	White matter microstructure (radial diffusivities)	white matter microstructure measurement	GCST009540	
		rs142739979-?	22E-08	BAIAP2L2	White matter microstructure (mean diffusivities)	white matter microstructure measurement	GCST009538	
BAIAP2L2		rs142739979-?	22E-08	BAIAP2L2	White matter microstructure (fractional anisotropy)	white matter microstructure measurement	GCST009539	
		rs62235125-?	22E-08	BAIAP2L2	White matter microstructure (fractional anisotropy)	white matter microstructure measurement	GCST009539	
		rs132930-?	22E-08	BAIAP2L2, AL022322.1	White matter microstructure (fractional anisotropy)	white matter microstructure measurement	GCST009539	
		rs72220602-?	22E-08	BAIAP2L2	White matter microstructure (fractional anisotropy)	white matter microstructure measurement	GCST009539	

rs5750538-?	22:38105400	5E-08 BAIAP2L2	White matter microstructure (fractional anisotropy)	white matter microstructure measurement	GCST009539
rs12484232-?	22:38107472	3E-09 BAIAP2L2	White matter microstructure (fractional anisotropy)	white matter microstructure measurement	GCST009539
rs12484232-?	22:38107472	4E-08 BAIAP2L2	White matter microstructure (fractional anisotropy)	white matter microstructure measurement	GCST009539
rs815753-C	1:189997267	4E-09 BRINP3, AL359976.1	Subjective well-being (MTAG)	wellbeing measurement	GCST005325
rs34580810-A	1:190038476	6E-10 AL359976.1, BRINP3	Chromotype	chromotype measurement	GCST007576
rs2419374-A	1:190088550	1E-06 BRINP3, AL359976.1	Suicide ideation score in major depressive disorder	suicide ideation measurement, unipolar depression	GCST002974
rs11803128-G	1:190090965	5E-09 BRINP3, AL359976.1	Depressive symptoms (MTAG)	depressive symptom measurement	GCST005323
rs11803128-G	1:190090965	7E-11 BRINP3, AL359976.1	Insomnia	insomnia measurement	GCST007988
rs1935881-T	1:190097256	1E-06 BRINP3, AL359976.1	Major depressive disorder	unipolar depression	GCST006041
rs146820337-I	1:190125997	2E-10 BRINP3	Chromotype	chromotype measurement	GCST007576
rs2247300-C	1:190162440	5E-06 BRINP3	Body mass index	body mass index	GCST008158
rs10920678-A	1:190270777	7E-20 BRINP3, AL354771.1	Body mass index	body mass index	GCST009003
rs10920678-A	1:190270777	9E-07 BRINP3, AL354771.1	Body mass index	body mass index	GCST002783
rs10920678-A	1:190270777	7E-20 BRINP3, AL354771.1	Body mass index	body mass index	GCST009004
rs10920678-A	1:190270777	4E-10 BRINP3, AL354771.1	Body mass index	body mass index	GCST005951
rs10920678-A	1:190270777	1E-06 BRINP3, AL354771.1	Body mass index	body mass index	GCST002783
rs10920678-A	1:190270777	7E-11 BRINP3, AL354771.1	Body mass index	body mass index	GCST006368
rs10920678-G	1:190270777	3E-13 BRINP3, AL354771.1	Body mass index	body mass index	GCST004904
rs10920678-G	1:190270777	6E-10 BRINP3, AL354771.1	Subjective well-being	wellbeing measurement	GCST005328
rs10920678-G	1:190270777	2E-11 BRINP3, AL354771.1	Subjective well-being (MTAG)	wellbeing measurement	GCST005325
rs815163-?	1:190325596	1E-15 AL354771.1, BRINP3	Body mass index	body mass index	GCST007039
rs491055-G	1:190339704	3E-10 BRINP3, AL354771.1	Body mass index	body mass index	GCST004904
rs1171044-T	1:190416174	4E-09 BRINP3	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs1171040-A	1:190431854	5E-11 BRINP3	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs1171040-A	1:190431854	9E-13 BRINP3	Educational attainment (MTAG)	self reported educational attainment	GCST006571
rs1869901-?	15:40303426	3E-06 PLCB2	Schizophrenia	schizophrenia	GCST001242
rs139701434-A	3:1100097	2E-06 CNTN6	Caudate activity during reward	caudate nucleus measurement, functional brain measurement, response to reward	GCST004970
rs17193334-A	3:1355689	6E-06 CNTN6	Depressive symptoms (stressful life events interaction)	depressive symptom measurement, stressful life event measurement	GCST003835
rs116256240-?	3:1385090	2E-06 CNTN6	Loneliness	loneliness measurement	GCST003771
rs192399095-?	3:1526228	4E-08 CNTN6, RPL23AP38	DNA methylation variation (age effect)	DNA methylation	GCST006660
rs13906-T	12:49558611	2E-09 MCRS1	Smoking initiation (ever regular vs never regular)	smoking status measurement	GCST007474
rs12104221-A	19:3797102	3E-08 MATK	Obesity-related traits	energy expenditure	GCST001762
rs11729278-T	4:113934816	9E-07 ARSJ	Impulsivity (motor)	behavioural disinhibition measurement	GCST007628
rs62171698-?	2:143201527	1E-10 ARHGAP15	Body mass index	body mass index	GCST007039
rs6710871-A	2:143203024	3E-19 ARHGAP15	Body mass index	body mass index	GCST006368
rs6710871-A	2:143203024	3E-16 ARHGAP15	Body mass index	body mass index	GCST009004
rs13002158-A	2:143264143	1E-12 ARHGAP15	Body mass index	body mass index	GCST009003
rs7601485-A	2:143295869	4E-08 ARHGAP15, AC079793.1	Educational attainment (MTAG)	self reported educational attainment	GCST006571
rs1343763-A	2:143361183	5E-06 AC079793.1, ARHGAP15	Perceived sweetness of sucrose	sensory perception of sweet taste	GCST008760
rs17824247-T	2:143394970	2E-12 AC079793.1, ARHGAP15	Educational attainment (years of education)	self reported educational attainment	GCST003676
rs62172064-T	2:143397319	9E-11 AC079793.1, ARHGAP15	Intelligence (MTAG)	intelligence	GCST005316
rs1320139-C	2:143400608	8E-26 ARHGAP15, AC079793.1	Educational attainment (years of education)	self reported educational attainment	GCST006442

rs10189912-A	2:143405040	1E-11 ARHGAP15, AC079793.1	Intelligence	intelligence	GCT006250
rs62172117-A	2:143411098	1E-17 AC079793.1, ARHGAP15	Morningness	chronotype measurement	GCT007983
rs10200533-A	2:143425433	4E-12 AC079793.1, ARHGAP15	Intelligence (MTAG)	intelligence	GCT005316
rs12991555-A	rs12991555-A	7E-10 AC079793.1, ARHGAP15	Highest math class taken	mathematical ability	GCT006574
rs13024996-A	2:143467646	6E-13 ARHGAP15, AC079793.1	Alcohol consumption (drinks per week)	alcohol consumption measurement	GCT007461
rs13024996-A	2:143467646	2E-15 ARHGAP15, AC079793.1	Alcohol consumption (drinks per week) (MTAG)	alcohol consumption measurement	GCT008757
rs4273169-?	2:143473740	4E-13 ARHGAP15, AC079793.1	Alcohol consumption	alcohol consumption measurement	GCT008595
		3E-12 AC079793.1, ARHGAP15	Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)	schizophrenia, intelligence, self reported educational attainment	
rs28380327-?	2:143474922	1E-10 ARHGAP15, AC079793.1	Morning person	chronotype measurement	GCT007083
rs28380327-A	2:143474922	5E-25 ARHGAP15, AC079793.1	Morning person	chronotype measurement	GCT007565
rs28380327-A	2:143474922	5E-25 ARHGAP15, AC079793.1	Chronotype	chronotype measurement	GCT007576
rs34112283-C	2:143475682	5E-06 ARHGAP15, AC079793.1	Neuroticism	neuroticism measurement	GCT003770
rs13413953-T	2:143479467	3E-09 AC079793.1, ARHGAP15	Alcohol consumption (drinks per week)	alcohol consumption measurement	GCT007328
rs13413953-T	2:143479467	8E-09 AC079793.1, ARHGAP15	Bitter alcoholic beverage consumption	bitter alcoholic beverage consumption measurement	GCT008522
rs13413953-G	2:143479467	3E-09 AC079793.1, ARHGAP15	Depressed affect	unipolar depression, mood disorder	GCT006475
rs10048736-A	2:143481734	1E-09 ARHGAP15, AC079793.1	Neuroticism	neuroticism measurement	GCT006476
rs35789697-A	2:143491149	3E-09 ARHGAP15, AC079793.1	Experiencing mood swings	mood instability measurement	GCT006944
rs35825582-A	2:143491193	2E-10 AC079793.1, ARHGAP15	Intelligence (MTAG)	intelligence	GCT005316
rs35825582-A	2:143491193	8E-08 AC079793.1, ARHGAP15	Smoking initiation (ever regular vs never regular) (MTAG)	smoking status measurement	GCT007468
rs35825582-?	2:143491193	3E-08 AC079793.1, ARHGAP15	Depressed affect	unipolar depression, mood disorder	GCT006476
rs34030812-T	2:143491336	2E-08 ARHGAP15, AC079793.1	Neuroticism	neuroticism measurement	GCT006475
rs34030812-?	2:143491336	3E-08 ARHGAP15, AC079793.1	Depressed affect	unipolar depression, mood disorder	GCT006475
rs13428598-T	2:143492918	1E-11 AC079793.1, ARHGAP15	Cognitive performance	cognitive function measurement	GCT006572
rs13428598-T	2:143492918	1E-18 AC079793.1, ARHGAP15	Cognitive performance (MTAG)	cognitive function measurement	GCT006570
rs13428598-T	2:143492918	4E-13 AC079793.1, ARHGAP15	Intelligence (MTAG)	intelligence	GCT005316
rs13428598-T	2:143492918	3E-09 AC079793.1, ARHGAP15	General cognitive ability	intelligence	GCT006269
rs13428598-T	2:143492918	5E-09 AC079793.1, ARHGAP15	Extremely high intelligence	intelligence	GCT007044
rs13428598-T	2:143492918	2E-19 AC079793.1, ARHGAP15	Highest math class taken (MTAG)	mathematical ability	GCT006568
rs13428598-?	2:143492918	4E-16 AC079793.1, ARHGAP15	Educational attainment (years of education)	self reported educational attainment	GCT007037
rs13428598-T	2:143492918	1E-29 AC079793.1, ARHGAP15	Educational attainment (MTAG)	self reported educational attainment	GCT006571
rs12465778-T	2:143497632	3E-10 AC079793.1, ARHGAP15	Intelligence (MTAG)	intelligence	GCT005316
rs35125029-?	2:143504088	3E-07 AC079793.1, ARHGAP15	Morning vs. evening chronotype	circadian rhythm	GCT003429
rs10191758-?	2:143505711	4E-08 AC079793.1, ARHGAP15	Cognitive ability	intelligence	GCT005142
rs10191758-A	2:143505711	3E-09 AC079793.1, ARHGAP15	Intelligence	intelligence	GCT004364
rs10191758-A	2:143505711	3E-09 AC079793.1, ARHGAP15	Intelligence	intelligence	GCT004364
rs10191758-?	2:143505711	2E-14 AC079793.1, ARHGAP15	Cognitive ability (MTAG)	intelligence, self reported educational attainment	GCT005141
rs10204230-T	2:143512879	6E-09 ARHGAP15, AC079793.1	Intelligence	intelligence	GCT006250
rs4233567-T	2:143514807	9E-10 AC079793.1, ARHGAP15	Alcohol consumption (drinks per week)	alcohol consumption measurement	GCT008811
rs4233567-T	2:143514807	1E-08 AC079793.1, ARHGAP15	Intelligence (MTAG)	intelligence	GCT005316
rs2731553-A	2:143616372	1E-10 ARHGAP15	Morningness	chronotype measurement	GCT007983
rs2731553-C	2:143616372	2E-14 ARHGAP15	Morning person	chronotype measurement	GCT007565
rs10928190-T	2:143618464	8E-10 ARHGAP15	Educational attainment (years of education)	self reported educational attainment	GCT006442
rs10928190-T	2:143618464	2E-10 ARHGAP15	Educational attainment (MTAG)	self reported educational attainment	GCT006571

rs10928195-C	2:143671039	4E-06 ARHGAP15	Quantitative traits	body weight	GCST000337
rs73961845-A	2:143721414	3E-15 AC092652.2, ARHGAP15	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs73961845-A	2:143721414	3E-17 AC092652.2, ARHGAP15	Educational attainment (MITAG)	self reported educational attainment	GCST006571
rs6721148-T	2:143745646	2E-13 ARHGAP15	Highest math class taken (MITAG)	mathematical ability	GCST006568
rs971230-C	3:100759471	4E-08 ABI3BP	Educational attainment (MITAG)	self reported educational attainment	GCST006571
rs12632235-T	3:100817440	5E-06 ABI3BP	Alcohol dependence	alcohol dependence	GCST003121
rs10898553-?	11:86914854	2E-06 PRSS23, APO01528.1	Suicide attempts	attempted suicide	GCST006428
rs123509-T	3:42691976	1E-06 KLHL40	Body fat distribution (trunk fat ratio)	body fat distribution	GCST007294
rs123509-T	3:42691976	3E-08 KLHL40	Body fat distribution (trunk fat ratio)	body fat distribution	GCST007294
rs123509-T	3:42691976	3E-06 KLHL40	Body fat distribution (arm fat ratio)	body fat distribution	GCST007293
rs10793903-A	9:132101265	1E-08 NTNG2, MED27	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs11243638-T	9:132146221	5E-10 MED27, NTNG2	Highest math class taken (MITAG)	mathematical ability	GCST006568
rs11243638-T	9:132146221	2E-09 MED27, NTNG2	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs12005151-A	9:132155262	2E-08 MED27, NTNG2	Cognitive performance (MITAG)	cognitive function measurement	GCST006570
rs12005151-A	9:132155262	4E-09 MED27, NTNG2	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs12005151-A	9:132155262	2E-10 MED27, NTNG2	Educational attainment (MITAG)	self reported educational attainment	GCST006571
rs6597547-?	9:132173502	4E-06 NTNG2	General cognitive ability	intelligence	GCST006269
rs942820-?	1:2448266	3E-09 PLCH2	Schizophrenia	schizophrenia	GCST009337
rs4648845-T	1:2455662	9E-10 PLCH2	Schizophrenia	schizophrenia	GCST002539
rs4648845-?	1:2455662	6E-09 PLCH2	Schizophrenia	schizophrenia	GCST008459
rs4648845-T	1:2455662	7E-09 PLCH2	Schizophrenia	schizophrenia	GCST003048
rs4648845-?	1:2455662	2E-10 PLCH2	Schizophrenia	schizophrenia	GCST006803
rs3762444-?	1:2496273	2E-14 PLCH2	Body mass index	body mass index	GCST007039
rs1109480-A	12:120645476	9E-10 CABP1	Smoking status (ever vs never smokers)	smoking behavior	GCST007327
rs1109480-A	12:120645476	6E-10 CABP1	Age of smoking initiation (MITAG)	smoking initiation	GCST007462
rs1109480-?	12:120645476	5E-10 CABP1	Smoking status	smoking status measurement	GCST007085
rs1109480-A	12:120645476	2E-10 CABP1	Smoking initiation (ever regular vs never regular)	smoking status measurement	GCST007474
rs1109480-A	12:120645476	2E-14 CABP1	Smoking initiation (ever regular vs never regular) (MITAG)	smoking status measurement	GCST007468
rs2152090-?	6:142284734	2E-08 ADGRG6, AL360007.1	White matter microstructure (mean diffusivities)	white matter microstructure measurement	GCST009538
rs6570507-A	6:142358435	5E-07 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference	GCST004063
rs6570507-A	6:142358435	6E-11 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference	GCST004063
rs6570507-A	6:142358435	2E-06 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference	GCST004063
rs6570507-A	6:142358435	4E-06 ADGRG6	Waist circumference adjusted for BMI in non-smokers	BMI-adjusted waist circumference	GCST004504
rs6570507-A	6:142358435	4E-08 ADGRG6	Waist circumference adjusted for BMI in non-smokers	BMI-adjusted waist circumference	GCST004504
rs6570507-A	6:142358435	3E-33 ADGRG6	Body fat distribution (trunk fat ratio)	body fat distribution	GCST007294
rs6570507-A	6:142358435	2E-31 ADGRG6	Body fat distribution (trunk fat ratio)	body fat distribution	GCST007294
rs6570507-A	6:142358435	1E-23 ADGRG6	Body fat distribution (leg fat ratio)	body fat distribution	GCST007295
rs6570507-A	6:142358435	8E-24 ADGRG6	Body fat distribution (leg fat ratio)	body fat distribution	GCST007295
rs6570507-A	6:142358435	1E-06 ADGRG6	Body fat distribution (arm fat ratio)	body fat distribution	GCST007293
rs6570507-A	6:142358435	8E-06 ADGRG6	Waist circumference adjusted for BMI in smokers	smoking behavior, BMI-adjusted waist circumference	GCST004503
rs6570507-A	6:142358435	2E-08 ADGRG6	Waist circumference adjusted for BMI (adjusted for smoking behaviour)	smoking behavior, BMI-adjusted waist circumference	GCST004500
rs6570507-A	6:142358435	2E-10 ADGRG6	Waist circumference adjusted for BMI (adjusted for smoking behaviour)	smoking behavior, BMI-adjusted waist circumference	GCST004500

rs6570507-A	6:142358435	2E-11 ADGRG6	Waist circumference adjusted for BMI (joint analysis main effects and smoking interaction)	smoking behavior, BMI-adjusted waist circumference	GCST004501
rs6570507-A	6:142358435	2E-08 ADGRG6	Waist circumference adjusted for BMI (joint analysis main effects and smoking interaction)	smoking behavior, BMI-adjusted waist circumference	GCST004501
rs6922607-A	6:142382346	2E-09 ADGRG6	Body mass index	body mass index	GCST009004
rs6922607-?	6:142382346	1E-08 ADGRG6	Body mass index	body mass index	GCST007039
rs4896582-?	6:142382740	7E-14 ADGRG6	Waist circumference adjusted for BMI (joint analysis main effects and physical activity interaction)	BMI-adjusted waist circumference, physical activity measurement	GCST004563
rs4896582-?	6:142382740	2E-08 ADGRG6	Waist circumference adjusted for BMI (joint analysis main effects and physical activity interaction)	BMI-adjusted waist circumference, physical activity measurement	GCST004563
rs4896582-G	6:142382740	9E-10 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference, physical activity measurement	GCST004562
rs4896582-?	6:142382740	1E-08 ADGRG6	Waist circumference adjusted for BMI (joint analysis main effects and physical activity interaction)	BMI-adjusted waist circumference, physical activity measurement	GCST004563
rs4896582-?	6:142382740	2E-09 ADGRG6	Waist circumference adjusted for BMI (joint analysis main effects and physical activity interaction)	BMI-adjusted waist circumference, physical activity measurement	GCST004563
rs4896582-?	6:142382740	2E-06 ADGRG6	Waist circumference adjusted for BMI (joint analysis main effects and physical activity interaction)	BMI-adjusted waist circumference, physical activity measurement	GCST004563
rs4896582-G	6:142382740	1E-13 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference, physical activity measurement	GCST004562
rs4896582-G	6:142382740	9E-09 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference, physical activity measurement	GCST004562
rs4896582-?	6:142382740	3E-13 ADGRG6	Waist circumference adjusted for BMI (joint analysis main effects and physical activity interaction)	BMI-adjusted waist circumference, physical activity measurement	GCST004563
rs4896582-G	6:142382740	9E-06 ADGRG6	Waist circumference adjusted for BMI in active individuals	BMI-adjusted waist circumference, physical activity measurement	GCST004564
rs4896582-G	6:142382740	2E-08 ADGRG6	Waist circumference adjusted for BMI in active individuals	BMI-adjusted waist circumference, physical activity measurement	GCST004564
rs4896582-G	6:142382740	2E-06 ADGRG6	Waist circumference adjusted for BMI in active individuals	BMI-adjusted waist circumference, physical activity measurement	GCST004564
rs4896582-G	6:142382740	1E-08 ADGRG6	Waist circumference adjusted for BMI in active individuals	BMI-adjusted waist circumference, physical activity measurement	GCST004564
rs4896582-G	6:142382740	4E-07 ADGRG6	Waist circumference adjusted for BMI in inactive individuals	BMI-adjusted waist circumference, physical activity measurement	GCST004565
rs4896582-G	6:142382740	7E-06 ADGRG6	Waist circumference adjusted for BMI in inactive individuals	BMI-adjusted waist circumference, physical activity measurement	GCST004565
rs4896582-G	6:142382740	9E-08 ADGRG6	Waist circumference adjusted for BMI in inactive individuals	BMI-adjusted waist circumference, physical activity measurement	GCST004565
rs4896582-G	6:142382740	1E-12 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference, physical activity measurement	GCST004562
rs4896582-G	6:142382740	2E-06 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference, physical activity measurement	GCST004562
rs6570508-A	6:142392705	2E-07 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference	GCST008161
rs6570508-G	6:142392705	2E-08 ADGRG6	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference, physical activity measurement	GCST004562

rs6570509-T	6:142395149	1E-29 ADGRG6	Hip circumference adjusted for BMI	BMI-adjusted hip circumference	GCST004067
rs6570509-T	6:142395149	3E-16 ADGRG6	Hip circumference adjusted for BMI	BMI-adjusted hip circumference	GCST004067
rs6570509-T	6:142395149	2E-16 ADGRG6	Hip circumference adjusted for BMI	BMI-adjusted hip circumference	GCST004067
rs931619-A	5:179087728	1E-07 ZNF354C, ADAMTS2	Waist-to-hip ratio adjusted for BMI	BMI-adjusted waist-hip ratio	GCST008994
rs10464079-?	5:179098778	5E-08 ZNF354C, ADAMTS2	Waist-hip ratio	waist-hip ratio	GCST007067
rs888762-C	5:179120312	1E-17 ADAMTS2	Body fat distribution (trunk fat ratio)	body fat distribution	GCST007294
rs888762-C	5:179120312	2E-14 ADAMTS2	Body fat distribution (leg fat ratio)	body fat distribution	GCST007295
rs888762-C	5:179120312	3E-14 ADAMTS2	Body fat distribution (leg fat ratio)	body fat distribution	GCST007295
rs888762-C	5:179120312	7E-18 ADAMTS2	Body fat distribution (trunk fat ratio)	body fat distribution	GCST007294
rs10039254-A	5:179143913	8E-06 ADAMTS2	Attention deficit hyperactivity disorder (time to onset)	attention deficit hyperactivity disorder	GCST000280
rs340122-?	5:179260682	6E-06 ADAMTS2	Reaction time	reaction time measurement	GCST006268
rs4700798-C	5:179292363	5E-06 ADAMTS2	Intake of total sugars	sugar consumption measurement	GCST008759
rs575377383-T	6:32041710	8E-11 TNXB	Automobile speeding propensity	risk-taking behaviour	GCST007329
rs34214527-C	6:32046679	4E-10 TNXB	Highest math class taken	mathematical ability	GCST006574
rs433061-A	6:32047051	4E-17 TNXB	Well-being spectrum (multivariate analysis)	wellbeing measurement	GCST007341
rs41316748-C	6:32051735	2E-08 TNXB	Depressive symptoms	depressive symptom measurement	GCST007340
rs41316748-C	6:32051735	2E-08 TNXB	Neuroticism	neuroticism measurement	GCST007339
rs41316748-C	6:32051735	1E-08 TNXB	Well-being spectrum (multivariate analysis)	wellbeing measurement	GCST007341
rs369637-?	6:32058086	2E-10 TNXB	Autism spectrum disorder or schizophrenia	schizophrenia, autism spectrum disorder	GCST004521
rs1150757-A	6:32061428	1E-14 TNXB	Depressive symptoms	depressive symptom measurement	GCST007340
rs1150757-A	6:32061428	3E-14 TNXB	Neuroticism	neuroticism measurement	GCST007339
rs1150757-A	6:32061428	1E-14 TNXB	Positive affect	wellbeing measurement	GCST007338
rs204883-A	6:32064966	7E-09 TNXB	Cognitive performance	cognitive function measurement	GCST006572
rs204883-A	6:32064966	1E-10 TNXB	Cognitive performance (MITAG)	cognitive function measurement	GCST006570
rs204883-A	6:32064966	2E-09 TNXB	Highest math class taken (MITAG)	wellbeing measurement	GCST007341
rs12211410-T	6:32081646	2E-13 TNXB	Well-being spectrum (multivariate analysis)	wellbeing measurement	GCST006442
rs185819-T	6:32082290	2E-12 TNXB	Educational attainment (years of education)	self reported educational attainment	GCST007329
rs1150754-C	6:32082981	3E-12 TNXB	Automobile speeding propensity	risk-taking behaviour	GCST007201
rs1150752-C	6:32096949	6E-13 TNXB	Schizophrenia	schizophrenia	GCST007337
rs41268902-A	6:32103231	5E-06 TNXB, AL662884.2	Life satisfaction	wellbeing measurement	GCST003048
rs429150-T	6:32107786	2E-09 AL662884.2, TNXB	Schizophrenia	schizophrenia	GCST006442
rs429150-T	6:32107786	1E-11 AL662884.2, TNXB	Educational attainment (years of education)	self reported educational attainment	GCST006571
rs2269426-A	6:32108722	4E-11 AL662884.2, TNXB, AL662884.5	Educational attainment (years of education)	self reported educational attainment	GCST007340
rs2269426-A	6:32108722	3E-11 AL662884.2, TNXB, AL662884.5	Depressive symptoms	depressive symptom measurement	GCST007339
rs2269426-A	6:32108722	1E-14 AL662884.2, TNXB, AL662884.5	Neuroticism	neuroticism measurement	GCST006940
rs2269426-A	6:32108722	4E-15 AL662884.2, TNXB, AL662884.5	Feeling worry	worry measurement	GCST006950
rs113397893-?	12:57217502	5E-12 NXP4	Body mass index	body mass index	GCST007039
rs113397893-A	12:57217502	9E-10 NXP4	Body mass index	body mass index	GCST009004
rs61736007-A	12:57224836	7E-06 NXP4	Trunk fat mass	fat body mass	GCST008154

rs12826178-?	12:57228588	6E-10 NXP4, AC137834.2	anorexia nervosa, attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive-compulsive disorder, schizophrenia, or Tourette syndrome (pleiotropy)	anorexia nervosa, obsessive-compulsive disorder, attention deficit hyperactivity disorder, Tourette syndrome, unipolar depression, schizophrenia, autism spectrum disorder, bipolar disorder	GCST009600
rs12826178-G	12:57228588	1E-11 NXP4, AC137834.2	Schizophrenia	schizophrenia	GCST003048
rs12826178-G	12:57228588	2E-12 NXP4, AC137834.2	Schizophrenia	schizophrenia	GCST002539
rs12826178-?	12:57228588	4E-10 NXP4, AC137834.2	Autism spectrum disorder or schizophrenia	schizophrenia, autism spectrum disorder	GCST004521
rs12826178-?	12:57228588	9E-11 NXP4, AC137834.2	Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)	schizophrenia, intelligence, self reported educational attainment	GCST008595
rs6686643-?	1:165647351	7E-06 MGS73	Total ventricular volume	brain volume measurement	GCST000898
rs11231640-?	11:63923577	8E-09 RCOR2, NAA40	anorexia nervosa, attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive-compulsive disorder, schizophrenia, or Tourette syndrome (pleiotropy)	anorexia nervosa, obsessive-compulsive disorder, attention deficit hyperactivity disorder, Tourette syndrome, unipolar depression, schizophrenia, autism spectrum disorder, bipolar disorder	GCST009600
rs12462975-A	19:29781295	1E-25 CCNE1, AC008798.3	Body mass index	body mass index	GCST009003
rs12462975-A	19:29781295	1E-25 CCNE1, AC008798.3	Body mass index	body mass index	GCST009004
rs113701136-?	19:29786822	2E-24 CCNE1, AC008798.3	Body mass index	body mass index	GCST007039
rs17513613-T	19:29795915	6E-25 CCNE1, AC008798.3	Body mass index	body mass index	GCST009001
rs17513613-T	19:29795915	2E-07 CCNE1, AC008798.3	Body mass index	body mass index	GCST006368
rs17513613-C	19:29795915	3E-06 CCNE1, AC008798.3	Body mass index	body mass index	GCST002783
rs17513613-T	19:29795915	4E-09 CCNE1, AC008798.3	Body mass index	body mass index	GCST005951
rs17513613-C	19:29795915	5E-06 CCNE1, AC008798.3	Body mass index	body mass index	GCST002783
rs17513613-T	19:29795915	5E-12 CCNE1, AC008798.3	Waist-hip ratio	waist-hip ratio	GCST008996
rs62104477-?	19:29804084	7E-06 ALDH1A1	Reaction time	reaction time measurement	GCST006268
rs76004885-G	9:73024947	9E-06 ALDH1A1	Sensorimotor dexterity	cognitive function measurement	GCST009310
rs3758354-C	9:73149649	3E-06 ALDH1A1, ANXA1	Schizophrenia, bipolar disorder and depression (combined)	unipolar depression, mental or behavioural disorder, bipolar disorder	GCST000774
rs1940597-A	18:5888042	7E-09 TMEM200C, AP001021.1	Highest math class taken	mathematical ability	GCST006574
rs7190-A	18:5889766	3E-10 TMEM200C, AP001021.1	Cognitive performance (MTAG)	cognitive function measurement	GCST006570
rs7190-A	18:5889766	1E-08 TMEM200C, AP001021.1	Self-reported math ability (MTAG)	mathematical ability	GCST006569
rs7190-A	18:5889766	5E-14 TMEM200C, AP001021.1	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs7190-A	18:5889766	2E-10 TMEM200C, AP001021.1	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs7190-A	18:5889766	2E-14 TMEM200C, AP001021.1	Educational attainment (MTAG)	self reported educational attainment	GCST006571
rs12457278-T	18:5893150	7E-09 AP001021.1, TMEM200C	Age of smoking initiation (MTAG)	smoking initiation	GCST007462
rs12457278-T	18:5893150	5E-12 AP001021.1, TMEM200C	Smoking initiation (ever regular vs never regular) (MTAG)	smoking status measurement	GCST007468
rs72694951-?	1:150146987	7E-09 VPS45, PLEKHO1	Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)	schizophrenia, intelligence, self reported educational attainment	GCST008595
rs7544145-T	1:150166501	5E-08 AC242988.2, PLEKHO1	Bipolar disorder	educational attainment	GCST008103
rs55802315-T	1:150166511	2E-14 AC242988.2, PLEKHO1	Self-reported math ability (MTAG)	bipolar disorder	GCST006569
rs55802315-C	1:150166511	4E-14 AC242988.2, PLEKHO1	Self-reported math ability	mathematical ability	GCST006573
rs2303100-C	19:9857758	1E-10 OLFM2	Sleep duration	mathematical ability	GCST006914
rs80076900-?	19:9945644	4E-10 OLFM2, COL5A3	Waist-hip ratio	sleep duration	GCST007067
rs678916-T	19:53937946	3E-09 CACNG7	Cognitive performance (MTAG)	waist-hip ratio	GCST007067
				cognitive function measurement	GCST006570

rs678916-?	19:53937946	3E-06 CACNG7	General cognitive ability	intelligence	GCST006269
rs307914-A	19:53938536	4E-12 CACNG7	Self-reported math ability (MTAG)	mathematical ability	GCST006569
rs307914-G	19:53938536	1E-09 CACNG7	Self-reported math ability	mathematical ability	GCST006573
rs307913-A	19:53939797	6E-12 CACNG7	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs307913-A	19:53939797	1E-08 CACNG7	Educational attainment (MTAG)	self reported educational attainment	GCST006571
rs112915006-G	22:50166267	1E-06 PANX2, MOV101	Type 2 diabetes	type II diabetes mellitus	GCST009379
rs5771206-?	22:50178377	5E-09 PANX2	Brain region volumes	brain volume measurement	GCST009518
rs5771206-?	22:50178377	2E-08 PANX2	Brain region volumes	brain volume measurement	GCST009518
rs58475265-C	11:134394248	8E-09 B3GAT1	Insomnia	insomnia measurement	GCST007988
rs7928758-?	11:134396073	3E-06 B3GAT1	Alcohol consumption (max-drinks)	alcohol drinking	GCST002294
rs232918-?	12:112645027	9E-07 RPH3A	Reaction time	reaction time measurement	GCST006268
rs2891403-?	12:112699767	4E-08 RPH3A	Body mass index	body mass index	GCST007039
rs7294822-?	12:112767698	4E-08 RPH3A	Smoking status	smoking status measurement	GCST007085
rs34536443-G	19:10352442	4E-15 TYK2	Type 1 diabetes	type I diabetes mellitus	GCST005536
rs12720356-A	19:10359299	4E-07 TYK2	Type 1 diabetes	type I diabetes mellitus	GCST005536
rs2304256-?	19:10364976	1E-10 TYK2	Type 1 diabetes	type I diabetes mellitus	GCST001191
rs2304256-C	19:10364976	4E-09 TYK2	Type 1 diabetes	type I diabetes mellitus	GCST005539
rs35607894-A	14:71941921	5E-11 RGS6	Schizophrenia	schizophrenia	GCST004946
rs35607894-A	14:71941921	6E-10 RGS6	Schizophrenia	schizophrenia	GCST007201
rs2190864-?	14:71949502	1E-12 RGS6	Schizophrenia	schizophrenia	GCST009337
rs2332700-?	14:71950609	5E-10 RGS6	anorexia nervosa, attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive-compulsive disorder, schizophrenia, or Tourette syndrome (pleiotropy)	anorexia nervosa, obsessive-compulsive disorder, attention deficit hyperactivity disorder, Tourette syndrome, unipolar depression, schizophrenia, autism spectrum disorder, bipolar disorder	GCST009600
rs2332700-?	14:71950609	1E-09 RGS6	Schizophrenia	schizophrenia	GCST008459
rs2332700-C	14:71950609	1E-09 RGS6	Schizophrenia	schizophrenia	GCST003048
rs2332700-?	14:71950609	2E-10 RGS6	Schizophrenia	schizophrenia	GCST006803
rs2332700-C	14:71950609	7E-10 RGS6	Schizophrenia	schizophrenia	GCST007201
rs2332700-C	14:71950609	5E-09 RGS6	Schizophrenia	schizophrenia	GCST002539
rs2332700-C	14:71950609	2E-09 RGS6	Autism spectrum disorder or schizophrenia	schizophrenia, autism spectrum disorder	GCST004521
rs2190872-?	14:71962347	2E-08 RGS6	Educational attainment (years of education)	self reported educational attainment	GCST007037
rs12888615-T	14:71968988	2E-13 RGS6	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs12888615-T	14:71968988	9E-12 RGS6	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs4902961-A	14:71970560	1E-10 RGS6	Self-reported math ability (MTAG)	mathematical ability	GCST006569
rs4902961-G	14:71970560	3E-09 RGS6	Self-reported math ability	mathematical ability	GCST006573
rs34328930-?	14:71971007	5E-11 RGS6	Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)	schizophrenia, intelligence, self reported educational attainment	GCST008595
rs12897542-A	14:71982874	3E-12 RGS6	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs12897542-A	14:71982874	2E-13 RGS6	Educational attainment (MTAG)	self reported educational attainment	GCST006571
rs36032616-G	14:72017439	1E-06 RGS6	Night sleep phenotypes	sleep quality	GCST003542
rs981975-T	14:72119843	9E-07 RGS6	antipsychotic drug dosage in schizophrenia or schizoaffective disorder	schizophrenia, antipsychotic drug use measurement, schizoaffective disorder	GCST003341
rs699363-G	14:72225785	5E-06 RGS6	Obesity (extreme)	obesity	GCST001240
rs2238199-A	14:72332295	3E-10 RGS6	Adventurousness	risk-taking behaviour	GCST007324
rs2238199-A	14:72332295	4E-09 RGS6	General risk tolerance (MTAG)	risk-taking behaviour	GCST007325

PEX5L	rs7159300-?	14:72417166	4E-06	RG56	Smoking cessation	smoking cessation	GCST000778
	rs7630877-A	3:179943530	7E-06	PEX5L	Type 2 diabetes	type II diabetes mellitus	GCST001033
	rs191580074-A	3:179974963	6E-06	PEX5L, AC092939.1	General cognitive ability	Intelligence	GCST006269
	rs62007304-A	3:180084176	2E-06	PEX5L, RN45SP149	Bipolar disorder	bipolar disorder	GCST008103
HMG20A	rs7178572-G	15:77430853	2E-08	HMG20A	Educational attainment (years of education)	self reported educational attainment	GCST006442
	rs7178572-G	15:77454848	6E-07	HMG20A	Type 2 diabetes	type II diabetes mellitus	GCST004894
	rs7178572-G	15:77454848	2E-08	HMG20A	Type 2 diabetes	type II diabetes mellitus	GCST004894
	rs7178572-G	15:77454848	1E-11	HMG20A	Type 2 diabetes	type II diabetes mellitus	GCST004894
	rs7178572-G	15:77454848	2E-11	HMG20A	Type 2 diabetes	type II diabetes mellitus	GCST002352
	rs7178572-G	15:77454848	7E-11	HMG20A	Type 2 diabetes	type II diabetes mellitus	GCST001213
	rs7119-T	15:77485290	1E-08	HMG20A	Type 2 diabetes	type II diabetes mellitus	GCST001550
	rs12910361-?	15:77489993	5E-07	AC090984.1, HMG20A	Type 2 diabetes	type II diabetes mellitus	GCST001033
	rs12910361-G	15:77489993	1E-11	HMG20A, AC046168.1	Type 2 diabetes	type II diabetes mellitus	GCST007077
	rs12898370-?	15:77516256	4E-08	AC046168.1, HMG20A	Type 2 diabetes	type II diabetes mellitus	GCST007847
	rs12234369-A	7:11853649	1E-08	THSD7A, TMEM106B	Risky sexual behaviors (alcohol dependence interaction)	alcohol dependence, risky sexual behaviour measurement	GCST004956
TMEM106B	rs12234369-A	7:11853649	6E-11	THSD7A, TMEM106B	Educational attainment (years of education)	self reported educational attainment	GCST006442
	rs12531804-T	7:11856428	2E-10	THSD7A, TMEM106B	Educational attainment (MTAG)	self reported educational attainment	GCST006571
	rs57107575-T	7:11857885	5E-09	TMEM106B, THSD7A	Highest math class taken (MTAG)	mathematical ability	GCST006568
	rs2883598-?	7:11867594	1E-07	TMEM106B, THSD7A	Educational attainment (years of education)	self reported educational attainment	GCST006442
	rs114945094-G	7:12153806	4E-06	TMEM106B, THSD7A	General cognitive ability	Intelligence	GCST006269
	rs74826345-T	7:12158819	4E-06	THSD7A, TMEM106B	Anorexia nervosa	anorexia nervosa	GCST001961
	rs62435692-T	7:12203113	1E-10	THSD7A, TMEM106B	Plasma kynurenine levels in major depressive disorder	unipolar depression, kynurenine measurement	GCST005342
	rs57506017-A	7:12205599	6E-09	TMEM106B, THSD7A	Neuroticism	neuroticism measurement	GCST007339
	rs57506017-?	7:12205599	2E-09	TMEM106B, THSD7A	Neuroticism	neuroticism measurement	GCST005232
	rs3807866-A	7:12210752	5E-08	TMEM106B, THSD7A	Depressed affect	unipolar depression, mood disorder	GCST006475
	rs3807865-A	7:12210776	7E-12	TMEM106B, THSD7A	Lifetime anxiety disorder	anxiety disorder	GCST009575
	rs3807864-?	7:12210947	3E-06	THSD7A, TMEM106B	Depression (broad)	unipolar depression	GCST005902
	rs6460895-G	7:12212914	2E-14	TMEM106B	General cognitive ability	Intelligence	GCST006269
	rs2043539-A	7:12214254	8E-24	TMEM106B	Depressive symptoms	depressive symptom measurement	GCST007340
	rs3823612-C	7:12219129	5E-11	TMEM106B	Depression	unipolar depression	GCST007342
	rs11509880-A	7:12222285	1E-09	TMEM106B	Depression	unipolar depression	GCST006477
	rs11509880-A	7:12222285	4E-11	TMEM106B	Feeling tense	feeling tense measurement	GCST006952
	rs11509880-A	7:12222285	6E-11	TMEM106B	Experiencing mood swings	mood instability measurement	GCST006944
	rs11509880-A	7:12222285	9E-11	TMEM106B	Neuroticism	neuroticism measurement	GCST006476
	rs10950393-C	7:12223920	2E-11	TMEM106B	Neuroticism	neuroticism measurement	GCST006940
	rs10950398-A	7:12225245	3E-08	TMEM106B	Depressed affect	unipolar depression, mood disorder	GCST006475
	rs5011432-C	7:12229042	2E-08	TMEM106B	Depression	unipolar depression	GCST005839
	rs14978-?	7:12233526	7E-06	TMEM106B	Major depressive disorder (probable)	unipolar depression	GCST005904
	rs7805419-T	7:12242825	1E-08	TMEM106B	Response to amphetamines	response to drug	GCST001651
	rs7805419-C	7:12242825	9E-13	TMEM106B	General factor of neuroticism	neuroticism measurement	GCST007709
WDR35	rs113386058-?	2:19944070	1E-09	WDR35	Well-being spectrum (multivariate analysis)	wellbeing measurement	GCST007341
NEDD4	rs17819300-?	15:55918731	1E-10	NEDD4	Waist-hip ratio	waist-hip ratio	GCST007067
	rs574687503-A	15:55990260	7E-06	NEDD4	Brain connectivity	brain connectivity measurement	GCST001889
					Hip circumference adjusted for BMI	BMI-adjusted hip circumference	GCST008156

MAPK1IP1L	rs67652508-?	14:55088650	3E-08 LGALS3, MAPK1IP1L	Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)	schizophrenia, intelligence, self reported educational attainment	GCST008595
CEP120	rs4308481-T	5:123316412	1E-08 CEP120, ACO10493.1	Body mass index	body mass index	GCST004904
	rs1582931-A	5:123321505	2E-13 ACO10493.1, CEP120	Body mass index	body mass index	GCST009004
	rs1582931-?	5:123321505	6E-13 ACO10493.1, CEP120	Body mass index	body mass index	GCST007039
	rs1582931-G	5:123321505	9E-08 ACO10493.1, CEP120	Body mass index	body mass index	GCST006368
	rs1582931-A	5:123321505	5E-09 ACO10493.1, CEP120	Body mass index	body mass index	GCST004904
	rs112446794-T	5:123329771	3E-10 CEP120, ACO10493.1	Obesity (extreme)	obese body mass index status	GCST007241
	rs2115172-A	5:123349434	2E-06 CEP120	Weight	body weight	GCST000299
	rs10051787-T	5:123396542	4E-06 CEP120	Waist circumference	waist circumference	GCST003335
	rs10051787-T	5:123396542	5E-08 CEP120	Waist circumference	waist circumference	GCST003335
	rs10051787-T	5:123396542	7E-12 CEP120	Waist circumference	waist circumference	GCST003335
ST13	rs13054099-T	22:40819668	7E-14 SLC25A17, ST13	Neuroticism	neuroticism measurement	GCST006476
	rs13054099-T	22:40819668	3E-10 SLC25A17, ST13	Neuroticism	neuroticism measurement	GCST006940
	rs13054099-T	22:40819668	1E-08 SLC25A17, ST13	Feeling worry	worry measurement	GCST006950
TBP1L	rs13201465-G	6:133912955	1E-06 TBP1L, TCF21	Night sleep phenotypes	sleep quality	GCST003542
AXIN1	rs7200589-A	16:2999331	4E-13 AXIN1	Body mass index	body mass index	GCST009003
	rs11866815-C	16:337867	2E-07 AXIN1	Body mass index	body mass index	GCST006368
	rs11866815-C	16:337867	2E-06 AXIN1	Body mass index	body mass index	GCST002783
	rs11866815-T	16:337867	2E-16 AXIN1	Body mass index	body mass index	GCST009004
	rs11866815-T	16:337867	5E-11 AXIN1	Body mass index	body mass index	GCST005951
	rs11866815-C	16:337867	1E-07 AXIN1	Body mass index	body mass index	GCST002783
	rs11866815-C	16:337867	2E-07 AXIN1	Body mass index	body mass index	GCST002783
	rs7201895-?	16:357723	1E-13 AXIN1, MRPL28	Body mass index	body mass index	GCST007039
ARL14EP	rs1765139-T	11:30332178	5E-14 ARL14EP	Body mass index	body mass index	GCST009001
	rs1765142-A	11:30357012	2E-06 ARL14EP, RPL12P30	Schizophrenia	schizophrenia	GCST003048
USP47	rs529308902-A	11:11826030	1E-06 ACO104383.2, USP47	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference	GCST008161
APAF1	rs116928965-A	12:98735094	3E-08 APAF1, ANKS1B	Body fat percentage	body fat percentage	GCST009531
SCN3B	rs3851101-G	11:123653388	4E-07 SCN3B	Schizoaffective disorder-bipolar type	Schizoaffective disorder-bipolar type	GCST009817
LRR4C	rs180682174-?	11:40203478	2E-08 LRR4C	DNA methylation variation (age effect)	DNA methylation	GCST006660
	rs4611189-?	11:40816163	6E-07 LRR4C	Response to antipsychotic treatment in schizophrenia (reasoning)	reasoning	GCST000885
	rs1462231-?	11:41011553	6E-06 LRR4C	Cerebellum cortex volume	cerebellum cortex volume measurement	GCST009203
	rs11511923-T	11:41069380	7E-10 LRR4C	General risk tolerance (MTAG)	risk-taking behaviour	GCST007325
	rs7124946-?	11:41116149	5E-08 LRR4C	Smoking status	smoking status measurement	GCST007085
	rs10444280-T	11:41134347	6E-09 LRR4C	Highest math class taken (MTAG)	mathematical ability	GCST006568
	rs10444280-T	11:41134347	1E-08 LRR4C	Educational attainment (years of education)	self reported educational attainment	GCST006442
	rs10444280-T	11:41134347	1E-09 LRR4C	Educational attainment (MTAG)	self reported educational attainment	GCST006571
	rs10768673-?	11:41235344	1E-07 LRR4C	Educational attainment (years of education)	self reported educational attainment	GCST007037
	rs1506719-T	11:41396711	2E-08 LRR4C, ACO90138.1	Risk-taking tendency (4-domain principal component model)	risk-taking behaviour	GCST007323
	rs2939756-A	11:41414747	3E-09 ACO90138.1, LRR4C	Smoking status (ever vs never smokers)	smoking behavior	GCST007327
	rs2939756-?	11:41414747	4E-12 ACO90138.1, LRR4C	Smoking status	smoking status measurement	GCST007085
	rs2939756-A	11:41414747	7E-10 ACO90138.1, LRR4C	Smoking initiation (ever regular vs never regular)	smoking status measurement	GCST007474
	rs2939755-?	11:41412922	7E-06 LRR4C, ACO90138.1	General cognitive ability	intelligence	GCST006269
	rs1506724-T	11:41462843	8E-11 LINC02741, LRR4C	Educational attainment (MTAG)	self reported educational attainment	GCST006571

rs11036413-T	11:41480329	8E-09 LINC02741, LRRRC4C	Age of smoking initiation (MTAG)	smoking initiation	GCST007462
rs11036413-T	11:41480329	1E-14 LINC02741, LRRRC4C	Smoking initiation (ever regular vs never regular) (MTAG)	smoking status measurement	GCST007468
rs12273007-A	11:41492154	1E-08 LRRRC4C, LINC02741	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs10170310-G	2:138521352	1E-07 SPOPL	Response to antipsychotic treatment	response to olanzapine	GCST001773
rs1216006-T	12:106794002	1E-13 RIC8B	Waist-hip ratio	waist-hip ratio	GCST007067
rs12369214-A	12:106804833	1E-12 RIC8B	Waist-to-hip ratio adjusted for BMI	BMI-adjusted waist-hip ratio	GCST008994
rs10444533-C	12:106835927	4E-06 RIC8B	Social autistic-like traits	autism spectrum disorder symptom	GCST002228
rs4472028-T	3:152335461	2E-10 MBNL1	Type 2 diabetes	Type II diabetes mellitus	GCST006867
rs7633673-A	3:152366454	4E-08 MBNL1	Depressive symptoms	depressive symptom measurement	GCST007340
rs7633673-A	3:152366454	4E-08 MBNL1	Neuroticism	neuroticism measurement	GCST007339
rs7633673-A	3:152366454	4E-08 MBNL1	Well-being spectrum (multivariate analysis)	wellbeing measurement	GCST007341
rs13065698-A	3:152368744	8E-13 MBNL1	Type 2 diabetes	Type II diabetes mellitus	GCST009379
rs9595893-T	13:32601656	4E-20 PDS5B	Body mass index	body mass index	GCST009001
rs9595908-T	13:32610151	5E-23 PDS5B	Body mass index	body mass index	GCST006368
rs9595908-T	13:32610151	6E-06 PDS5B	Body mass index	body mass index	GCST002783
rs9595908-T	13:32610151	6E-06 PDS5B	Body mass index	body mass index	GCST002783
rs9595908-T	13:32610151	4E-20 PDS5B	Body mass index	body mass index	GCST009004
rs9595908-T	13:32610151	4E-20 PDS5B	Body mass index	body mass index	GCST009003
rs4942792-G	13:32663326	1E-06 PDS5B	Logical memory (delayed recall)	memory performance	GCST007010
rs7428850-T	3:24116906	2E-06 THRB, LINC00691	antipsychotic drug dosage in schizophrenia or schizoaffective disorder	schizophrenia, antipsychotic drug use measurement, schizoaffective disorder	GCST003341
rs1349265-A	3:24117896	2E-13 THRB	Intelligence	intelligence	GCST006250
rs9310732-A	3:24174066	3E-06 THRB	Schizophrenia	schizophrenia	GCST003048
rs4858582-?	3:24178030	9E-06 THRB	General cognitive ability	intelligence	GCST006269
rs1171832-C	10:59908059	3E-09 CCDC6, LINC01553	Adventurousness	risk-taking behaviour	GCST007324
rs17561924-T	15:26985669	2E-12 GABRG3	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs11635966-?	15:26989121	4E-10 GABRG3	Educational attainment (years of education)	self reported educational attainment	GCST007037
rs891793-C	15:27010874	2E-12 GABRG3	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs891792-A	15:27010970	3E-12 GABRG3	Educational attainment (years of education)	self reported educational attainment	GCST006571
rs2195817-A	15:27016500	2E-11 GABRG3	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs112401219-G	15:27065401	8E-06 GABRG3, GABRG3-AS1	Cognitive performance (MTAG)	cognitive function measurement	GCST006570
rs184405367-?	15:27153352	4E-09 GABRG3-AS1, GABRG3	Bitter beverage consumption	bitter beverage consumption measurement	GCST008521
rs140780521-T	15:27393141	6E-06 GABRG3	Anthropometric traits (multi-trait analysis)	body weights and measures	GCST007490
rs4633669-T	15:27432822	3E-08 GABRG3	Weight	body weight	GCST008152
rs62003775-A	15:27450563	1E-08 GABRG3	Cognitive performance (MTAG)	cognitive function measurement	GCST006570
rs4595752-G	15:27456432	5E-06 GABRG3	General risk tolerance (MTAG)	risk-taking measurement	GCST007325
rs34078361-A	15:27457801	2E-10 GABRG3	Eudaimonic well-being	wellbeing measurement	GCST006635
rs7174138-A	15:27458569	6E-11 GABRG3	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs7174138-A	15:27458569	1E-09 GABRG3	Self-reported math ability (MTAG)	mathematical ability	GCST006569
rs13279512-?	8:61045943	9E-07 AC022182.2, CLVS1	Self-reported math ability	mathematical ability	GCST006573
rs10097298-?	8:61079697	3E-06 CLVS1	Opioid dependence	opioid dependence	GCST005368
rs6471915-?	8:61082222	2E-06 CLVS1	Prudent dietary pattern	diet measurement	GCST004639
rs12681792-A	8:61141904	3E-12 CLVS1	Body mass index	body mass index	GCST009003
rs12681792-A	8:61141904	3E-12 CLVS1	Body mass index	body mass index	GCST009004
rs12681792-?	8:61141904	4E-09 CLVS1	Body mass index	body mass index	GCST007039

rs12681792-?	8:61141904	9E-07 CLVS1	Migraine	migraine disorder	GCST002081
rs6471932-T	8:61166345	1E-14 CLVS1	Body mass index	body mass index	GCST006368
rs4738873-A	8:61173617	2E-06 CLVS1	Body mass index	body mass index	GCST002783
rs4738873-A	8:61173617	5E-07 CLVS1	Body mass index	body mass index	GCST002783
rs6985602-A	8:61223177	3E-06 CLVS1	Waist circumference adjusted for body mass index	BMI-adjusted waist circumference	GCST008161
rs903027-?	8:61496869	6E-06 CLVS1	Whole-brain volume	whole-brain volume	GCST000891
rs1052618-G	3:136855659	8E-09 NCK1-DT, SLC35G2	Body mass index	body mass index	GCST008129
rs1609391-A	3:136896426	3E-11 NCK1	Neuroticism	neuroticism measurement	GCST005232
rs1609391-A	3:136896426	4E-08 NCK1	Neuroticism	neuroticism measurement	GCST006940
rs9867325-C	3:136900067	1E-06 NCK1, RAD51AP1P1	Body mass index	body mass index	GCST002783
rs9867325-C	3:136900067	1E-06 NCK1, RAD51AP1P1	Body mass index	body mass index	GCST002783
rs9867325-C	3:136900067	1E-06 NCK1, RAD51AP1P1	Body mass index	body mass index	GCST002783
rs57537217-T	4:108819029	8E-06 COL25A1	Maximum habitual alcohol consumption	alcohol consumption measurement	GCST008675
rs17039583-?	4:108959574	4E-06 COL25A1	Schizophrenia (age at onset)	schizophrenia, age at onset	GCST001115
rs3096490-G	4:109057975	5E-07 COL25A1	Body mass index	body mass index	GCST002571
rs9991319-?	4:109249480	4E-06 COL25A1	Reaction time	reaction time measurement	GCST006268
rs1780437-?	10:33967875	4E-08 PARD3, LINC02629	DNA methylation variation (age effect)	DNA methylation	GCST006660
rs1780436-?	10:33968684	7E-06 LINC02629, PARD3	Major depressive disorder	unipolar depression	GCST006582
rs1780436-?	10:33968684	3E-06 LINC02629, PARD3	Major depressive disorder	unipolar depression	GCST001083
rs61846062-?	10:33971707	3E-08 LINC02629, PARD3	DNA methylation variation (age effect)	DNA methylation	GCST006660
rs148822459-?	10:33973918	5E-09 LINC02629, PARD3	DNA methylation variation (age effect)	DNA methylation	GCST006660
rs3004929-?	10:34125937	2E-08 PARD3	Body mass index	body mass index	GCST007039
rs758480150-CCT	10:34212305	2E-06 PARD3	Hip circumference adjusted for BMI	BMI-adjusted hip circumference	GCST008156
rs641162-A	10:34277992	5E-10 PARD3	Self-reported math ability (MTAG)	mathematical ability	GCST006569
rs641162-A	10:34277992	8E-14 PARD3	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs1765174-C	10:34279104	1E-09 PARD3	Highest math class taken	mathematical ability	GCST006574
rs2197012-A	10:34294707	2E-12 PARD3	Educational attainment (MTAG)	self reported educational attainment	GCST006571
rs2007655-T	10:34297761	9E-10 PARD3	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs142712198-?	10:34449119	6E-06 PARD3	General cognitive ability	intelligence	GCST006269
rs73267375-?	10:34600928	8E-06 PARD3	General cognitive ability	intelligence	GCST006269
rs3938236-?	15:38253524	2E-08 SPRED1	HDL cholesterol levels x long total sleep time interaction (2df test)	sleep duration, high density lipoprotein cholesterol measurement	GCST009368
rs6495962-T	15:38359831	3E-08 SPRED1, FAM98B	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs7040561-A	9:125766699	7E-10 PBX3	Caudate nucleus volume	caudate nucleus volume	GCST009666
rs10987006-C	9:125889260	7E-09 PBX3	Sleep duration	sleep duration	GCST006914
rs1238134-A	9:125984333	8E-10 AL589923.1, PBX3	Medication use (opioids)	Opioid use measurement	GCST007936
rs6818414-T	4:16599041	5E-09 LDB2, AC104656.1	Body mass index	body mass index	GCST009004
rs158271-C	4:16826126	5E-06 LDB2	Waist circumference	waist circumference	GCST008151
rs158271-C	4:16826126	5E-06 LDB2	Waist circumference	waist circumference	GCST008160
rs49969-T	4:16830189	1E-08 LDB2	Number of sexual partners	risky sexual behaviour measurement	GCST007326
rs2035442-?	4:16891547	7E-08 LDB2	Body mass index	body mass index	GCST007039
rs11138313-?	9:79626053	2E-08 TLE4	Body mass index	body mass index	GCST007039
rs11138313-A	9:79626053	6E-10 TLE4	Body mass index	body mass index	GCST009004
rs9886703-A	9:79631436	2E-09 TLE4	Highest math class taken (MTAG)	mathematical ability	GCST006568
rs9886703-A	9:79631436	4E-13 TLE4	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs9886703-A	9:79631436	9E-14 TLE4	Educational attainment (MTAG)	self reported educational attainment	GCST006571

rs914715-?	9:79695983	1E-06 TLE4	Bipolar disorder	bipolar disorder	GCST000961
rs7039819-G	9:79815503	5E-10 TLE4, AL161912.3	Lifetime smoking index	smoking behaviour measurement	GCST0009096
rs4877152-T	9:79823670	3E-11 AL161912.3, TLE4	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs7031698-?	9:79826571	1E-08 AL161912.3, TLE4	Educational attainment (years of education)	self reported educational attainment	GCST007037
rs4877516-A	9:79827816	8E-13 AL161912.3, TLE4	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs10780329-T	9:79828183	2E-13 AL161912.3, TLE4	Educational attainment (MTAG)	self reported educational attainment	GCST006571
rs12401738-A	1:77981077	2E-10 DNAJB4, GIPC2	Body mass index	body mass index	GCST002783
rs12401738-A	1:77981077	8E-08 DNAJB4, GIPC2	BMI in non-smokers	body mass index	GCST004499
rs12401738-A	1:77981077	2E-06 DNAJB4, GIPC2	BMI in non-smokers	body mass index	GCST004499
rs12401738-A	1:77981077	4E-14 DNAJB4, GIPC2	Body mass index	body mass index	GCST005951
rs12401738-A	1:77981077	6E-10 DNAJB4, GIPC2	Body mass index	body mass index	GCST002783
rs12401738-A	1:77981077	1E-10 DNAJB4, GIPC2	Body mass index	body mass index	GCST002783
rs12401738-A	1:77981077	9E-08 DNAJB4, GIPC2	BMI (adjusted for smoking behaviour)	smoking behavior, body mass index	GCST004495
rs12401738-A	1:77981077	4E-06 DNAJB4, GIPC2	BMI (adjusted for smoking behaviour)	smoking behavior, body mass index	GCST004495
rs12401738-A	1:77981077	3E-07 DNAJB4, GIPC2	Body mass index (joint analysis main effects and smoking interaction)	smoking behavior, body mass index	GCST004497
rs12401738-A	1:77981077	6E-06 DNAJB4, GIPC2	Body mass index (joint analysis main effects and smoking interaction)	smoking behavior, body mass index	GCST004497
rs34517439-?	1:77984833	2E-40 GIPC2, DNAJB4	Body mass index	body mass index	GCST007039
rs34517439-A	1:77984833	9E-09 GIPC2, DNAJB4	Smoking initiation (ever regular vs never regular) (MTAG)	smoking status measurement	GCST007468
rs4130548-T	1:77998184	2E-16 GIPC2, DNAJB4	Body mass index	body mass index	GCST006368
rs4130548-?	1:77998184	5E-09 GIPC2, DNAJB4	Body mass index (joint analysis main effects and physical activity interaction)	physical activity measurement, body mass index	GCST004558
rs4130548-?	1:77998184	1E-09 GIPC2, DNAJB4	Body mass index (joint analysis main effects and physical activity interaction)	physical activity measurement, body mass index	GCST004558
rs4130548-?	1:77998184	6E-10 GIPC2, DNAJB4	Body mass index (joint analysis main effects and physical activity interaction)	physical activity measurement, body mass index	GCST004558
rs4130548-?	1:77998184	3E-08 GIPC2, DNAJB4	Body mass index (joint analysis main effects and physical activity interaction)	physical activity measurement, body mass index	GCST004558
rs4130548-C	1:77998184	1E-09 GIPC2, DNAJB4	Body mass index	physical activity measurement, body mass index	GCST004557
rs4130548-C	1:77998184	2E-10 GIPC2, DNAJB4	Body mass index	physical activity measurement, body mass index	GCST004557
rs4130548-C	1:77998184	1E-09 GIPC2, DNAJB4	Body mass index	physical activity measurement, body mass index	GCST004557
rs4130548-C	1:77998184	5E-10 GIPC2, DNAJB4	Body mass index	physical activity measurement, body mass index	GCST004557
rs4130548-C	1:77998184	9E-07 GIPC2, DNAJB4	Body mass index in physically active individuals	physical activity measurement, body mass index	GCST004559
rs4130548-C	1:77998184	2E-06 GIPC2, DNAJB4	Body mass index in physically active individuals	physical activity measurement, body mass index	GCST004559
rs4130548-C	1:77998184	2E-06 GIPC2, DNAJB4	Body mass index in physically active individuals	physical activity measurement, body mass index	GCST004559
rs4130548-C	1:77998184	8E-07 GIPC2, DNAJB4	Body mass index in physically active individuals	physical activity measurement, body mass index	GCST004559

rs4130548-?	1:77998184	2E-09 <i>GIPC2</i> , <i>DNAJB4</i>	Cognitive ability, years of educational attainment or schizophrenia (pleiotropy)	schizophrenia, intelligence, self reported educational attainment	GCST008595
rs1130548-T	1:77998184	1E-06 <i>GIPC2</i> , <i>DNAJB4</i>	Waist circumference	waist circumference	GCST004065
rs1130548-T	1:77998184	3E-10 <i>GIPC2</i> , <i>DNAJB4</i>	Waist circumference	waist circumference	GCST004065
rs12446623-?	16:61713765	3E-06 <i>CDH8</i>	General cognitive ability	intelligence	GCST006269
rs143812851-A	16:61739945	1E-09 <i>CDH8</i>	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs143812851-A	16:61739945	9E-11 <i>CDH8</i>	Educational attainment (MTAG)	self reported educational attainment	GCST006570
rs140663152-A	16:61740153	4E-08 <i>CDH8</i>	Cognitive performance (MTAG)	cognitive function measurement	GCST006568
rs78513878-T	16:61755851	6E-10 <i>CDH8</i>	Highest math class taken (MTAG)	mathematical ability	GCST008522
rs78423589-G	16:61794878	4E-06 <i>CDH8</i>	Bitter alcoholic beverage consumption	bitter alcoholic beverage consumption measurement	GCST007324
rs117167357-A	16:61801391	8E-09 <i>CDH8</i>	Adventurousness	risk-taking behaviour	GCST007468
rs72798770-A	16:61887196	1E-10 <i>CDH8</i>	Smoking initiation (ever regular vs never regular) (MTAG)	smoking status measurement	GCST006569
rs968627-A	16:62019156	3E-11 <i>CDH8</i>	Self-reported math ability (MTAG)	mathematical ability	GCST006573
rs968627-A	16:62019156	4E-08 <i>CDH8</i>	Self-reported math ability	mathematical ability	GCST007329
rs7201179-T	16:62021842	6E-10 <i>CDH8</i>	Automobile speeding propensity	risk-taking behaviour	GCST006570
rs12325429-A	16:62025344	3E-14 <i>CDH8</i>	Cognitive performance (MTAG)	cognitive function measurement	GCST003813
rs200312707-?	16:62031776	1E-06 <i>CDH8</i>	Response to antidepressants and depression	unipolar depression, mood disorder, response to antidepressant	GCST006572
rs12446238-A	16:62041234	3E-10 <i>CDH8</i> , <i>RNU6-21P</i>	Cognitive performance	cognitive function measurement	GCST006250
rs12446238-A	16:62041234	5E-09 <i>CDH8</i> , <i>RNU6-21P</i>	Intelligence	intelligence	GCST006568
rs12162090-A	16:62047130	5E-13 <i>RNU6-21P</i> , <i>CDH8</i>	Highest math class taken (MTAG)	mathematical ability	GCST007037
rs74710412-?	16:62052134	6E-09 <i>RNU6-21P</i> , <i>CDH8</i>	Educational attainment (years of education)	self reported educational attainment	GCST006442
rs7202829-T	16:62097799	6E-11 <i>RNU6-21P</i> , <i>CDH8</i>	Educational attainment (years of education)	self reported educational attainment	GCST006571
rs7202829-T	16:62097799	7E-14 <i>RNU6-21P</i> , <i>CDH8</i>	Educational attainment (MTAG)	self reported educational attainment	GCST006660
rs62049757-?	16:62133113	5E-09 <i>RNU6-21P</i> , <i>CDH8</i>	DNA methylation variation (age effect)	DNA methylation	GCST006442
rs929556-T	16:62159736	2E-12 <i>RNU6-21P</i> , <i>CDH8</i>	Educational attainment (years of education)	self reported educational attainment	GCST003800
rs77945277-?	14:34840969	2E-07 <i>BAZ1A</i>	Response to bupropion in depression	unipolar depression, response to bupropion, mood disorder	GCST006570
rs13212041-T	6:77461407	1E-09 <i>HTR1B</i> , <i>AL365222.1</i>	Cognitive performance (MTAG)	cognitive function measurement	GCST006442
rs13212041-T	6:77461407	7E-09 <i>HTR1B</i> , <i>AL365222.1</i>	Educational attainment (years of education)	self reported educational attainment	GCST006571
rs13212041-T	6:77461407	4E-12 <i>HTR1B</i> , <i>AL365222.1</i>	Educational attainment (MTAG)	self reported educational attainment	GCST006569
rs6297-T	6:77462224	6E-13 <i>HTR1B</i>	Self-reported math ability (MTAG)	mathematical ability	GCST006568
rs6297-T	6:77462224	5E-15 <i>HTR1B</i>	Highest math class taken (MTAG)	mathematical ability	GCST006573
rs12173930-C	6:77476717	4E-10 <i>RPS6P7</i> , <i>HTR1B</i>	Self-reported math ability	mathematical ability	GCST006574
rs7748725-G	6:77493665	5E-09 <i>RPS6P7</i> , <i>HTR1B</i>	Highest math class taken	mathematical ability	GCST008156
rs76793066-T	6:24962027	8E-06 <i>RIPOR2</i>	Hip circumference adjusted for BMI	BMI-adjusted hip circumference	

Trait: Standardised trait term from the Experimental Factor Ontology to represent the trait investigated; Reported trait: trait definition based on the authors' description of the phenotype analysed.

Supplementary Table 3. Differentially expressed protein-coding genes with DMRs in their possible promoter region (5Kb upstream and 1 Kb downstream from the TSS).

Gene Symbol	Discovery				RNA-seq data				DMR	MBD-seq data				Chromatin state of mouse developing brain																
	Expression	FC	p-value	Replica	Expression	FC	p-value	Transcript/s ID (transcript type)		CF	VALP	RALP	FC	p-value	Pr-A	Pr-W	Pr-B	Pr-F	En-Sd	En-Sp	En-W	En-Pd	En-Pp	Tr-S	Tr-P	Tr-I	Hc-P	Hc-H	Ns	
<i>Egfr</i>	87.5	-1.28	3.55E-02	32.4	-1.55	5.87E-03	ENSMUST000000197250 (R)	chr3:129681201-129681300	2	10.7	6.6	1.57	2.53E-02																17	
							ENSMUST000000197250 (R)	chr3:129681301-129681400	2	9.8	6.1	1.58	3.34E-02															9		
							ENSMUST000000197250 (R)	chr3:129681501-129681600	1	3.5	1.4	2.26	2.01E-02											1	8			17		
<i>Chpf2</i>	5574.0	-1.14	2.94E-02	1695.7	-1.12	3.84E-03	ENSMUST000000197933 (P)	chr5:24590801-24590900	3	4.4	2.1	1.93	4.89E-02									4	1	12	8			1		
							ENSMUST000000197933 (P)	chr5:24591001-24591100	3	8.3	4.5	1.71	1.20E-02										5	1	10	5			1	
							ENSMUST000000197933 (P)	chr5:24591101-24591200	4	10.0	6.3	1.51	3.77E-02										5	1	10	5			1	
							ENSMUST000000197933 (P)	chr5:24591201-24591300	3	11.2	7.1	1.52	2.71E-02										4	1	6	2			13	
							ENSMUST000000197933 (P)	chr5:24591301-24591400	9	11.6	7.3	1.55	2.59E-02										4	1	6	2			13	
<i>Mctp2</i>	116.4	-1.40	4.83E-02	45.2	-1.53	5.58E-03	ENSMUST000000205410 (M)	chr7:72166401-72166500	1	3.8	1.6	2.24	1.68E-02									4	1	6	2			23		
NAC																														
<i>Adgrg6</i>	495.3	-1.85	1.57E-02	170.2	-1.32	4.62E-02	ENSMUST000000208429 (P)	chr10:14550501-14550600	2	4.7	1.8	2.48	9.11E-03																26	
							ENSMUST000000208429 (P)	chr10:14550601-14550700	3	4.4	2.0	2.04	3.51E-02																26	
<i>Adamts2</i>	668.4	-1.35	7.41E-03	229.6	-1.30	4.87E-02	ENSMUST00000040523	chr11:5059701-5059800	3	4.2	1.8	2.30	1.69E-02																2	
							ENSMUST00000040523	chr11:5059801-5059900	2	4.7	2.0	2.28	8.54E-03																2	
							ENSMUST00000040523	chr12:83047901-83048000	3	1.6	0.2	4.58	3.88E-02										2	3					6	
<i>Rgs6</i>	2383.6	-1.17	3.83E-02	796.4	-1.11	1.38E-02	ENSMUST000000190964 (R)	chr13:24501101-24501200	1	3.4	6.5	-1.88	1.35E-02																12	
<i>Ripor2</i>	12838.7	1.15	2.91E-02	4258.5	1.13	2.60E-02	ENSMUST000000238974 (P)	chr13:24501201-24501300	3	4.5	8.2	-1.78	2.75E-02																9	
							ENSMUST000000238974 (P)	chr13:24501301-24501400	1	2.3	4.7	-2.00	2.43E-02																7	
							ENSMUST000000238974 (P)	chr13:24582101-24582200	2	2.3	4.9	-2.06	1.77E-02																7	
							ENSMUST000000238974 (P)	chr13:24614301-24614400	1	0.3	1.5	-4.04	2.25E-02																1	
							ENSMUST000000238974 (P)	chr13:24636601-24636700	2	1.1	2.8	-2.34	4.92E-02																21	
							ENSMUST000000238974 (P)	chr13:24674701-24674800	2	0.7	2.4	-3.07	1.70E-02																21	
							ENSMUST000000238974 (P)	chr13:24674901-24675000	2	0.4	1.6	-3.21	3.52E-02																3	
							ENSMUST000000238974 (P)	chr13:24686001-24686100	1	1.1	3.0	-2.75	2.60E-02																2	
							ENSMUST000000238974 (P)	chr13:24686201-24686300	2	1.5	4.2	-2.77	4.38E-03																2	
							ENSMUST000000238974 (P)	chr13:24686401-24686500	5	2.0	6.0	-3.11	2.15E-04																2	
							ENSMUST000000238974 (P)	chr13:24686601-24686700	1	1.5	3.4	-2.25	4.12E-02																2	
<i>Mapk1ip1l</i>	5408.1	1.06	2.96E-02	1873.0	1.09	8.24E-03	ENSMUST000000164235 (P)	chr13:24688001-24688100	2	1.8	3.6	-1.97	4.73E-02																2	
<i>Csf2rb2</i>	73.4	-1.29	2.37E-02	24.1	-1.47	2.56E-02	ENSMUST00000096356 (P)	chr14:47309001-47309100	1	0.1	1.2	-5.30	2.15E-02																9	
							ENSMUST00000096356 (P)	chr15:78308601-78308700	1	3.0	0.7	4.15	5.79E-03																17	
<i>Pamx2</i>	9867.0	-1.13	4.90E-02	2916.3	-1.14	5.45E-03	ENSMUST000000161372 (P)	chr15:78308701-78308800	2	3.4	1.2	2.70	7.86E-03																	17
							ENSMUST000000161372 (P)	chr15:89057801-89057900	1	2.8	0.9	2.71	1.93E-02																	4
							ENSMUST000000161372 (P)	chr15:89057901-89058000	3	4.0	0.9	3.81	1.98E-04																	4
							ENSMUST000000161372 (P)	chr15:89058001-89058100	1	4.2	1.1	3.50	3.18E-04																	4
							ENSMUST000000161372 (P)	chr15:89058101-89058200	2	4.4	1.2	3.36	2.99E-04																	4
							ENSMUST000000161372 (P)	chr15:89058201-89058300	1	3.4	1.5	2.16	2.45E-02																	4
<i>lacc</i>	348.5	1.22	3.52E-02	117.0	1.34	2.28E-02	ENSMUST000000133219 (P)	chr16:33023301-33023400	1	4.1	7.7	-1.85	1.61E-02																8	
							ENSMUST000000133219 (P)	chr16:33023401-33023500	1	3.7	6.3	-1.70	3.54E-02																	18
<i>Arlgapp15</i>	217.5	-2.33	3.57E-02	80.4	-1.78	4.87E-02	ENSMUST000000055776 (P)	chr2:43749501-43749600	1	1.7	0.3	4.56	1.86E-02																	26
<i>Arsf</i>	120.6	-1.86	4.55E-02	42.6	-1.85	3.80E-02	ENSMUST000000112822 (M)	chr3:126363101-126363200	2	4.9	2.3	2.03	1.55E-02																	25
							ENSMUST000000112822 (M)	chr3:126363401-126363500	1	4.1	2.2	1.83	4.91E-02																	12
<i>Npnt</i>	792.1	-1.89	1.25E-02	253.5	-1.25	4.78E-02	ENSMUST000000132732 (R)	chr3:132911101-132911200	1	3.4	1.2	2.50	2.55E-02																	18

<i>S100bpb</i>	7406.3	1.09	3.98E-02	2626.6	1.09	2.20E-02	ENSMUST00000072431 (PC); ENSMUST00000106059 (PC); ENSMUST00000106061 (PC); ENSMUST00000117350 (PC); ENSMUST00000117497 (PC); ENSMUST00000117965 (PC);	chr4:129191601-129191700	1	0.2	1.2	-5.29	3.21E-02	7	1	5	13
<i>Plk3cd</i>	2039.5	-1.16	2.41E-02	670.6	-1.11	3.59E-02	ENSMUST0000012059 (PC); ENSMUST00000124510 (PT);	chr4:149705801-149705900	1	4.6	2.2	2.02	2.15E-02				26
<i>Rund3b</i>	5249.4	1.09	3.62E-02	1856.6	1.07	3.83E-02	ENSMUST00000047485 (PC);	chr5:8622801-8622900	17	0.1	1.2	-5.30	2.15E-02	22	4		
<i>Ldb2</i>	2316.3	1.20	4.60E-02	818.7	1.14	1.74E-02	ENSMUST00000199471 (PT);	chr5:4459701-44597300	1	0.1	1.4	-6.37	7.91E-03				11
<i>Pds5b</i>	16764.3	1.09	2.70E-02	6189.3	1.11	2.98E-03	ENSMUST00000016569 (PC); ENSMUST00000038900 (PC);	chr5:150673401-150673500	2	0.4	1.6	-3.21	3.70E-02	5	16	5	
<i>Zfp536</i>	4053.2	-1.38	2.72E-02	1500.5	-1.16	3.33E-02	ENSMUST00000206763 (PC)	chr7:37571901-37572000	2	4.2	1.6	2.55	1.20E-02				25
							chr7:37572001-37572100	1	6.7	3.1	2.13	9.88E-03					24
							chr7:37572101-37572200	1	7.9	4.3	1.81	1.65E-02					24
							chr7:37572201-37572300	2	8.8	4.9	1.74	2.01E-02					24
							chr7:37572301-37572400	3	8.8	5.2	1.63	2.63E-02					24
<i>Gabrg3</i>	8249.1	1.13	2.66E-02	2997.4	1.12	1.69E-02	ENSMUST0000068394 (PC); ENSMUST0000068911 (PC);	chr7:5739201-5739300	1	0.1	1.4	-5.84	1.17E-02	3			7
<i>Aqp11</i>	1031.8	-1.22	3.56E-02	337.2	-1.17	4.72E-02	ENSMUST0000084986 (PC);	chr7:9774091-97741000	2	8.1	4.8	1.69	2.32E-02				24
<i>Par3d</i>	5227.3	1.17	1.52E-02	1750.3	1.13	3.43E-02	ENSMUST0000026921 (PC); ENSMUST00000079777 (PC); ENSMUST00000159141 (NMID); ENSMUST00000159537 (PC); ENSMUST00000159818 (NMID); ENSMUST00000160272 (PC); ENSMUST00000160581 (PC); ENSMUST00000160593 (RI); ENSMUST00000162035 (RI); ENSMUST00000159940 (PT)	chr8:127353201-127353300	1	2.8	5.2	-1.84	4.60E-02				16
							chr8:127353301-127353400	2	0.3	1.5	-4.04	2.86E-02					21
							chr8:127577801-127577900	1	0.2	1.2	-5.29	2.98E-02					14
							chr8:127577901-127578000	3	0.2	1.6	-6.92	3.43E-03					14
<i>Scrb3b</i>	14062.2	1.08	2.50E-02	5175.3	1.10	1.46E-02	ENSMUST00000176547 (PC);	chr9:40277201-40277300	2	1.0	2.5	-2.39	4.76E-02				10
							chr9:40277301-40277400	1	0.4	1.7	-3.47	3.93E-02					12
							chr9:40277401-40277500	3	0.4	1.6	-3.21	3.59E-02					13
							chr9:40280001-40280100	2	1.8	4.0	-2.26	1.97E-02					13
<i>Neeld4</i>	60519.3	1.09	1.96E-02	21418.4	1.08	3.66E-03	ENSMUST00000238315 (PC); ENSMUST00000184287 (RI)	chr9:72693001-72693100	3	0.3	1.8	-5.07	7.41E-03	7	6	2	10
							chr9:72745201-72745300	1	1.7	4.3	-2.41	5.82E-03					26
							chr9:72745301-72745400	3	2.2	4.6	-2.07	2.15E-02					26
							chr9:72745401-72745500	3	2.2	4.2	-1.89	3.60E-02					26
<i>Klfh4d</i>	166.9	-1.78	7.55E-03	51.3	-1.49	2.31E-02	ENSMUST0000098272 (PC);	chr9:72745501-72745600	1	1.4	3.3	-2.16	3.53E-02				26
							chr9:121778201-121778300	16	4.9	2.3	2.08	2.96E-02	4	20	2		

DMR: Differentially methylated region (build hg19); FC: Fold Change; CF: Coupling factor; number of CpG in the DMR; VALP: Vulnerable to addition-like phenotype; RALP: Resilient to addition-like phenotype; Chromatine state of mouse developing brain: number of chromatin state found considering all brain areas available; PC: protein coding; RI: retained intron; PT: Processed transcript; NMID: Nonsense mediated decay; Pr-A: Promoter active; Pr-B: Promoter Bivalent; Pr-F: Promoter Flanking Region; En-Sd: Enhancer Strong TSS-distal; En-Sp: Enhancer Strong TSS-proximal; En-W: Enhancer Weak; En-Pd: Enhancer Poised TSS-distal; En-Pp: Enhancer Poised TSS-proximal; Tr-S: Transcription Strong; Tr-P: Transcription Permissive; Tr-I: Transcription Initiation; Hc-P: Heterochromatin Polycomb-associated; Hc-H: Heterochromatin H3K9me3-associated; Ns: no significant-signal;

Supplementary Table 4. Differentially expressed miRNAs in mPFC

miRNA	Discovery			Replica		
	Expression	FC	p-value	Expression	FC	p-value
mmu-miR-876-5p	18.51	-1.85	2.50E-04	18.05	-1.36	3.93E-02
mmu-miR-211-5p	74.78	-2.02	2.10E-03	93.08	-1.4	3.81E-02
mmu-miR-3085-3p	232.02	-1.35	7.30E-03	292.99	-1.46	1.90E-03
mmu-miR-665-3p	337.76	-1.27	8.90E-03	356.86	-1.39	4.90E-03
mmu-miR-3072-3p	144.61	-1.41	1.74E-02	135.65	-1.32	3.46E-02
mmu-miR-124-3p	81006.63	-1.26	2.65E-02	83266.03	-1.16	3.26E-02
mmu-miR-29c-3p	3002.98	-1.33	2.93E-02	2336.75	-1.28	1.38E-02
mmu-miR-544-3p	77.83	-1.44	3.66E-02	53.65	-1.32	3.74E-02
mmu-miR-137-3p	9233.46	-1.28	3.83E-02	818.38	-1.28	3.15E-02
mmu-miR-100-5p	91517.51	1.17	4.04E-02	52784.51	1.22	2.62E-02
mmu-miR-192-5p	4450.81	1.16	4.58E-02	2282.71	1.19	2.82E-02

FC: Fold change

Supplementary Table 5. Differentially expressed miRNA genes with DMRs in their possible promoter region (5Kb upstream and 1 Kb downstream from the TSS).

miRNA	RNA-seq data			MBD-seq data			Rpkms mean			Chromatin state of mouse developing brain																
	Expression	FC	p-value	Expression	FC	p-value	CF	VALP	RALP	FC	p-value	Pr-A	Pr-W	Pr-B	Pr-F	En-Sp	En-W	En-Pd	En-Pp	Tr-S	Tr-P	Tr-I	Hc-P	Hc-H	Ns	
mmu-miR-124-3p	8.10E+04	-1.26	2.65E-02	8.10E+04	-1.16	3.26E-02	chr3:17795101-17795200	5	8.28	4.82	1.62	3.77E-02	24													
							chr3:17795301-17795400	3	5.48	3.02	1.73	4.42E-02	22													
							chr3:17795501-17795600	1	6.06	3.06	1.92	2.92E-02	21													
							chr3:17795601-17795700	2	5.93	2.84	1.96	1.65E-02	20	1												
mmu-miR-3085-3p	2.32E+02	-1.35	7.30E-03	2.32E+02	-1.46	1.90E-03	chr19:42282101-42282200	3	13.45	6.98	1.87	6.45E-04	24													
							chr19:422821801-422821900	3	13.58	7.31	1.81	1.00E-03	24													
							chr19:42282001-42282100	1	11.17	6.22	1.76	7.65E-03	24													
							chr19:42282101-42282200	3	10.58	5.49	1.89	2.58E-03	24													
						chr19:42282201-42282300	1	9.26	5.48	1.65	3.49E-02	24														

DMR: Differentially methylated region (build hg19); FC: Fold Change; CF: Coupling factor, number of CpG in the DMR; VALP: Vulnerable to addition-like phenotype; RALP: Resilient to addition-like phenotype; Chromatin state of mouse developing brain: number of chromatin state found considering all brain areas available; Pr-A: Promoter active; Pr-W: Promoter Weak; Pr-B: Promoter Bivalent; Pr-F: Promoter Flanking Region; En-Sd: Enhancer Strong TSS-distal; En-Sp: Enhancer Strong TSS-proximal; En-W: Enhancer Weak; En-Pd: Enhancer Poised TSS-distal; En-Pp: Enhancer Poised TSS-proximal; Tr-S: Transcription Strong; Tr-P: Transcription Permissive; Tr-I: Transcription Initiation; Hc-P: Heterochromatin Polycomb-associated; Hc-H: Heterochromatin H3K9me3-associated; Ns: no significant signal.

CAPÍTOL 3. Identificació de factors epigenètics de susceptibilitat a trastorns comòrbids a l'addicció a les drogues.

Article 6

“Estudi de variants genètiques que tenen un efecte en els nivells de metilació del cervell en el trastorn per dèficit d'atenció amb hiperactivitat”

Resum:

El trastorn per dèficit d'atenció amb hiperactivitat (TDAH) és un trastorn del neurodesenvolupament causat per una interacció de factors genètics i ambientals. L'epigenètica és crucial en la generació de canvis persistents que afecten l'expressió gènica al cervell. Estudis recents suggereixen que la metilació del DNA té un paper clau en el TDAH. En aquest estudi hem investigat la contribució al TDAH de la metilació específica d'al·lel (ASM), un mecanisme epigenètic pel qual els al·lells d'un SNP presenten correlació amb els nivells de metilació de llocs CpG. Hem seleccionat 3.896 tagSNPs que sabem que tenen influència sobre la metilació en el cervell i hem dut a terme un estudi d'associació cas-control utilitzant la metanàlisi GWAS més gran realitzada per TDAH fins ara, que inclou 20.183 casos i 35.191 controls. Hem observat un enriquiment de variants genètiques de risc pel TDAH en el conjunt d'ASM SNPs i hem identificat vuit ASM tagSNPs associats al TDAH (considerant un *false discovery rate* del 5%). Aquests SNPs presenten correlació amb la metilació de llocs CpG situats a les regions promotores de sis gens. Com que la metilació pot afectar l'expressió gènica, hem explorat si aquests vuit ASM tagSNPs, juntament amb 52 ASM SNPs que presenten valors de desequilibri de lligament elevats amb els anteriors, són eQTLs al cervell, i hem observat que l'expressió de tres dels gens identificats prèviament està modificada per ells. Així, els al·lells de risc per TDAH presenten correlació amb un increment de l'expressió (i una disminució de la metilació) d'*ARTN* i *PIDD1* i amb una disminució de l'expressió (i un increment de la metilació) de *C2orf82*. A més, prediccions fetes mitjançant diverses eines bioinformàtiques mostren una expressió alterada d'aquests tres gens en el TDAH, i les variants genètiques de *C2orf82* presenten correlació amb el volum de certes àrees cerebrals. En resum, hem identificat tres gens nous que poden contribuir al trastorn, seguint una aproximació sistemàtica per identificar variants de risc per TDAH que presenten correlació amb una metilació en *cis* diferencial.










Referència:

Pineda-Cirera L, Shivalikanjli A, Cabana-Domínguez J, Demontis D, Rajagopal VM, Børghlum AD, Faraone SV, Cormand B, Fernández-Castillo N. Exploring genetic variation that influences brain methylation in attention-deficit/hyperactivity disorder. *Transl Psychiatry*. 2019 Oct 3;9(1):242. doi: 10.1038/s41398-019-0574-7.

ARTICLE

Open Access

Exploring genetic variation that influences brain methylation in attention-deficit/hyperactivity disorder

Laura Pineda-Cirera ^{1,2,3}, Anu Shivalikanjli ^{1,3,4}, Judit Cabana-Domínguez ^{1,2,3,4}, Ditte Demontis ^{5,6,7}, Veera M. Rajagopal ^{5,6,7}, Anders D. Børghlum ^{5,6,7}, Stephen V. Faraone ⁸, Bru Cormand ^{1,2,3,4} and Noèlia Fernández-Castillo ^{1,2,3,4}

Abstract

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder caused by an interplay of genetic and environmental factors. Epigenetics is crucial to lasting changes in gene expression in the brain. Recent studies suggest a role for DNA methylation in ADHD. We explored the contribution to ADHD of allele-specific methylation (ASM), an epigenetic mechanism that involves SNPs correlating with differential levels of DNA methylation at CpG sites. We selected 3896 tagSNPs reported to influence methylation in human brain regions and performed a case-control association study using the summary statistics from the largest GWAS meta-analysis of ADHD, comprising 20,183 cases and 35,191 controls. We observed that genetic risk variants for ADHD are enriched in ASM SNPs and identified associations with eight tagSNPs that were significant at a 5% false discovery rate (FDR). These SNPs correlated with methylation of CpG sites lying in the promoter regions of six genes. Since methylation may affect gene expression, we inspected these ASM SNPs together with 52 ASM SNPs in high LD with them for eQTLs in brain tissues and observed that the expression of three of those genes was affected by them. ADHD risk alleles correlated with increased expression (and decreased methylation) of *ARTN* and *PIDD1* and with a decreased expression (and increased methylation) of *C2orf82*. Furthermore, these three genes were predicted to have altered expression in ADHD, and genetic variants in *C2orf82* correlated with brain volumes. In summary, we followed a systematic approach to identify risk variants for ADHD that correlated with differential *cis*-methylation, identifying three novel genes contributing to the disorder.

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a common neurodevelopmental disorder with a worldwide prevalence of around 5%¹. Its main symptoms include

inattention and/or hyperactivity-impulsivity (DSM-V)². ADHD is among the most heritable psychiatric disorders, with about 76% of its etiology accounted by genetic risk factors³ and with single-nucleotide polymorphisms (SNPs) explaining around 22% of the phenotypic variance⁴. Furthermore, there is molecular evidence of shared genetic risk factors across many psychiatric disorders⁵. In ADHD, a recent genome-wide association study (GWAS) meta-analysis of 12 sample groups unraveled some of the specific genetic underpinnings of this polygenic disorder for the first time⁴. One of the challenges of GWAS is to establish the causal relationship between the associated genetic variants, especially those

Correspondence: Bru Cormand (bcormand@ub.edu) or Noèlia Fernández-Castillo (noefernandez@ub.edu)

¹Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain

²Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Madrid, Spain

Full list of author information is available at the end of the article

These authors contributed equally: Laura Pineda-Cirera, Anu Shivalikanjli

These authors jointly supervised this work: Bru Cormand, Noèlia Fernández-Castillo

© The Author(s) 2019



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

located outside genes, and the disorder. In this regard, the use of epigenetic information can improve the interpretation of functionality of non-coding genetic variation⁶. In addition, some studies have hypothesized the importance of sub-threshold variants derived from GWAS^{7,8}, particularly those located in enhancer regions, with a potential impact on gene regulation^{9,10}.

DNA methylation is one of the most stable epigenetic mechanisms, involving mainly cytosines of CpG dinucleotides. This mechanism plays an important role in the regulation of neurogenesis, differentiation, and brain development¹¹. Furthermore, epigenetic alterations have been hypothesized to contribute to neurodevelopmental disorders¹², including ADHD¹³, autism spectrum disorders (ASD)^{14,15}, or borderline personality disorder¹⁶.

DNA methylation can be complementary if it involves both alleles, or non-complementary when it affects only one allele, as in chromosome X inactivation in females or allele-specific methylation (ASM)⁶. ASM is a common mechanism by which single nucleotide variants determine differential methylation levels of CpG sites. ASM can alter promoter activity, leading to allele-specific expression¹⁷ in combination with other still quite unknown factors, such as environmental effects⁶. It is quantitative and heterogeneous across tissues and individuals⁶. The environment affects DNA methylation leading to changes in gene regulation, although the underlying mechanism is still not well understood¹⁸. It has been suggested that, during embryonic development, ASM regions could be especially sensitive to environmental effects⁶. Investigating SNPs that display ASM could help to identify risk variants for common diseases, including neuropsychiatric disorders¹⁹, as shown by recent studies of bipolar disorder (BD) and schizophrenia^{10,20}.

The present study investigated the possible contribution of ASM to ADHD using data from the largest GWAS meta-analysis performed to date in ADHD⁴. We also assessed its possible effect on gene expression and on brain volumes to identify new genes contributing to the disorder.

Materials and methods

Selection of ASM SNPs

SNP selection was made based on the results of two previous studies^{21,22}, which identified ASM variants in multiple brain regions of post-mortem human samples. Gibbs et al.²¹, considered four brain regions (cerebellum, frontal cortex, caudal pons, and temporal cortex) of 150 subjects and Zhang et al.²², used only the cerebellum of 153 subjects. Gibbs et al.²¹, unlike Zhang et al.²², excluded those sequences of probes with significant correlation with methylation that contained polymorphisms. To discard possible artifacts in our results, we checked and confirmed that none of the probes used to detect the

six highlighted CpG sites target genomic regions with SNP variants. The genotyping platforms used in the two studies were different (Gibbs et al.²¹ used Infinium HumanHap550 Beadchips and Zhang et al.²² used Affymetrix GeneChip Mapping 5.0K Array). Both studies evaluated DNA methylation using the HumanMethylation27 Beadchips, and performed linear regression analyses by PLINK²³ to determine the correlation between each SNP and methylation of any CpG site^{21,22}. Zhang et al.²², unlike Gibbs et al.²¹ applied quantile normalization to the residuals prior to the linear regression analyses.

In the study by Zhang et al.²², a total of 12,117 SNP–CpG pairs associations were reported in cerebellum, and Gibbs et al.²¹ listed a total of 12,135 SNP–CpG pairs in frontal cortex, 11,374 in caudal pons, 16,734 in temporal cortex, and 12,102 in cerebellum (Fig. 1). We combined the information from both studies and obtained a total of 43,132 SNP–CpG pairs involving 33,944 different SNPs and 5306 CpG sites (Fig. 1). We considered all the ASM SNPs from all the tissues in the two studies, as there are multiple SNP–CpG pairs in common between them (Fig. S1).

We subsequently applied different filters to generate a sub-list of 3896 SNPs (Figs. 1 and S2) out of these 33,944 variants to minimize redundancy: associations in *cis* between the SNP and the CpG site, correlation of the SNP with methylation levels of the CpG ($R^2 \geq 0.2$), as performed in Gibbs et al. (2010)²¹. We considered only autosomal SNPs and selected tagSNPs for each CpG site ($r^2 \geq 0.85$), by assessing linkage disequilibrium (LD) with Haploview software²⁴ using the Central European (CEU) reference panel from 1000 Genomes Project Phase 3²⁵.

Case-control GWAS datasets

We explored the selected ASM SNPs in the summary statistics from a meta-analysis of 11 independent GWAS of ADHD conducted by the Psychiatric Genomics Consortium (PGC) and iPSYCH. This case-control study investigated 8,047,420 markers in 20,183 cases and 35,191 controls from Europe, USA, Canada, and China, with patients diagnosed according to the criteria detailed in Demontis et al. (2019)⁴.

Statistical analysis

To test whether risk variants for ADHD are enriched in ASM SNPs, we carried out an enrichment analysis using the Fisher's exact test in R²⁶ at p -value thresholds ranging from $5E-02$ to $5E-08$ considering the total number of ASM SNPs available from the ADHD GWAS meta-analysis⁴ (32,884 out of 33,944 SNPs).

From our selection of 3896 ASM tagSNPs, we could retrieve information on the association with ADHD of 3771 SNPs (96.8%) that were present in the summary

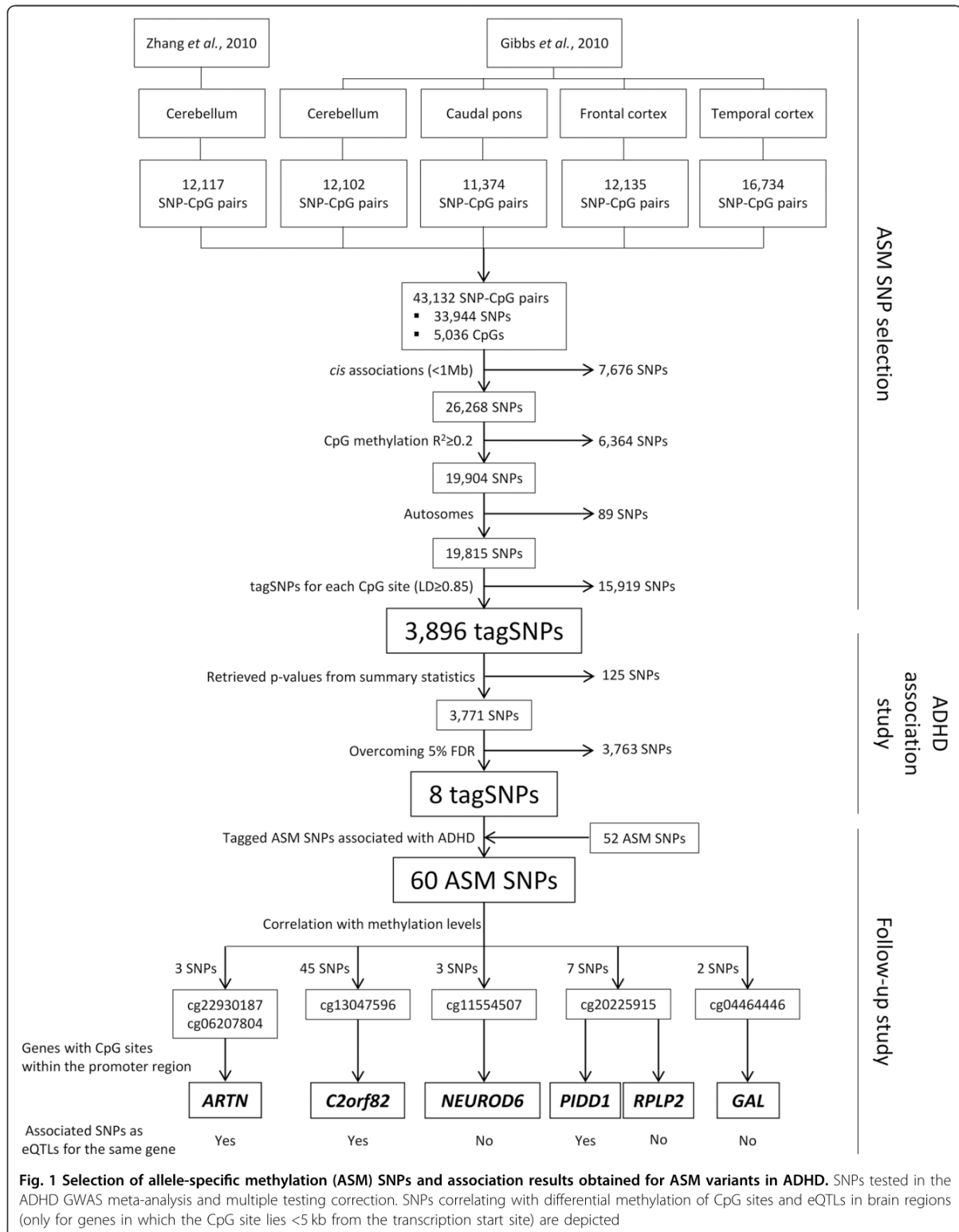


Fig. 1 Selection of allele-specific methylation (ASM) SNPs and association results obtained for ASM variants in ADHD. SNPs tested in the ADHD GWAS meta-analysis and multiple testing correction. SNPs correlating with differential methylation of CpG sites and eQTLs in brain regions (only for genes in which the CpG site lies <5 kb from the transcription start site) are depicted

statistics of the ADHD GWAS meta-analysis (Fig. 1)⁴. False discovery rate (FDR) was applied to correct for multiple testing. We used the *q*-value package for R²⁷ and obtained a threshold *p*-value of 6.78E−05 corresponding to a 5% FDR. CpG sites highlighted by SNPs that were significant at this FDR threshold were followed-up in further analyses (Fig. 1). Additionally, we performed corrections for multiple testing, using Bonferroni and Genetic type 1 Error Calculator (GEC) methods (<http://grass.cgs.hku.hk/gec/>)²⁸. The Bonferroni-corrected threshold was set at $p \leq 1.32E-05$, which considered all the SNPs and tests to be independent (0.05/3771 SNPs). The GEC established the significance threshold at 1.98E−05, which addressed multiple testing for the set of 3771 dependent SNPs by estimating the independent number of tests. The LD between SNPs was calculated according to the 1000 Genomes EUR reference data²⁵.

Finally, we considered and retrieved *p*-values of those tagged ASM SNPs in high LD ($r^2 \geq 0.85$) with the previous ones that also correlated in *cis* with the methylation levels of the same CpG sites ($R^2 \geq 0.2$) (Fig. 1).

Functional annotation of associated ASM SNPs

We applied four methods to obtain information about the possible functional impact of the ASM SNPs that were associated with ADHD. First, we evaluated the presence of possible enhancer or promoter regions using the Haploreg v4.1 tool²⁹. To do this, we considered histone modifications related to enhancer regions (H3K4me1 and H3K27ac) and promoters (H3K4me3 and H3K9ac) of 10 different brain regions (hippocampus middle, substantia nigra, anterior caudate, cingulate gyrus, inferior temporal lobe, angular gyrus, dorsolateral prefrontal cortex, germinal matrix, and male and female fetal brain). Second, we evaluated the effect on gene expression through an eQTL analysis using GTEx data (Release V7)³⁰. We considered eQTL information for all available brain tissues: amygdala, anterior cingulate cortex (BA24), caudate basal ganglia, cerebellar hemisphere, cerebellum, cortex, frontal cortex (BA9), hippocampus, hypothalamus, nucleus accumbens basal ganglia, putamen basal ganglia, spinal cord cervical c-1, and substantia nigra. Third, we considered all the SNPs, not only ASMs, located within ± 1 Mb from the transcription start site (TSS) of each gene to infer if the genetically determined expressions of genes of interest correlated with ADHD. This analysis was carried out using MetaXcan³¹, the input being the summary statistics of the ADHD GWAS meta-analysis⁴ and prediction models trained with RNA-Seq data of 10 GTEx³⁰ brain tissues and CommonMind³² dorsolateral prefrontal cortex. The SNP covariance matrices were generated using the 1000 Genomes Project Phase 3²⁵ EUR genotypes of the prediction model SNPs. Bonferroni correction for multiple testing was considered ($p \leq 2.27E-03$; 0.05/22

tests). Finally, we examined the possible influence of the identified variants on subcortical brain structures. We obtained the summary statistics of a GWAS meta-analysis of eight MRI volumetric measures (nucleus accumbens, amygdala, caudate nucleus, hippocampus, pallidum, putamen, and thalamus) produced by the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) consortium³³. This ENIGMA2 discovery sample included 13,171 subjects of European ancestry and contained association results between seven million markers and variance in the volumes of the mentioned structures³³; we applied the Bonferroni correction ($p \leq 1E-03$; 0.05/50 SNPs).

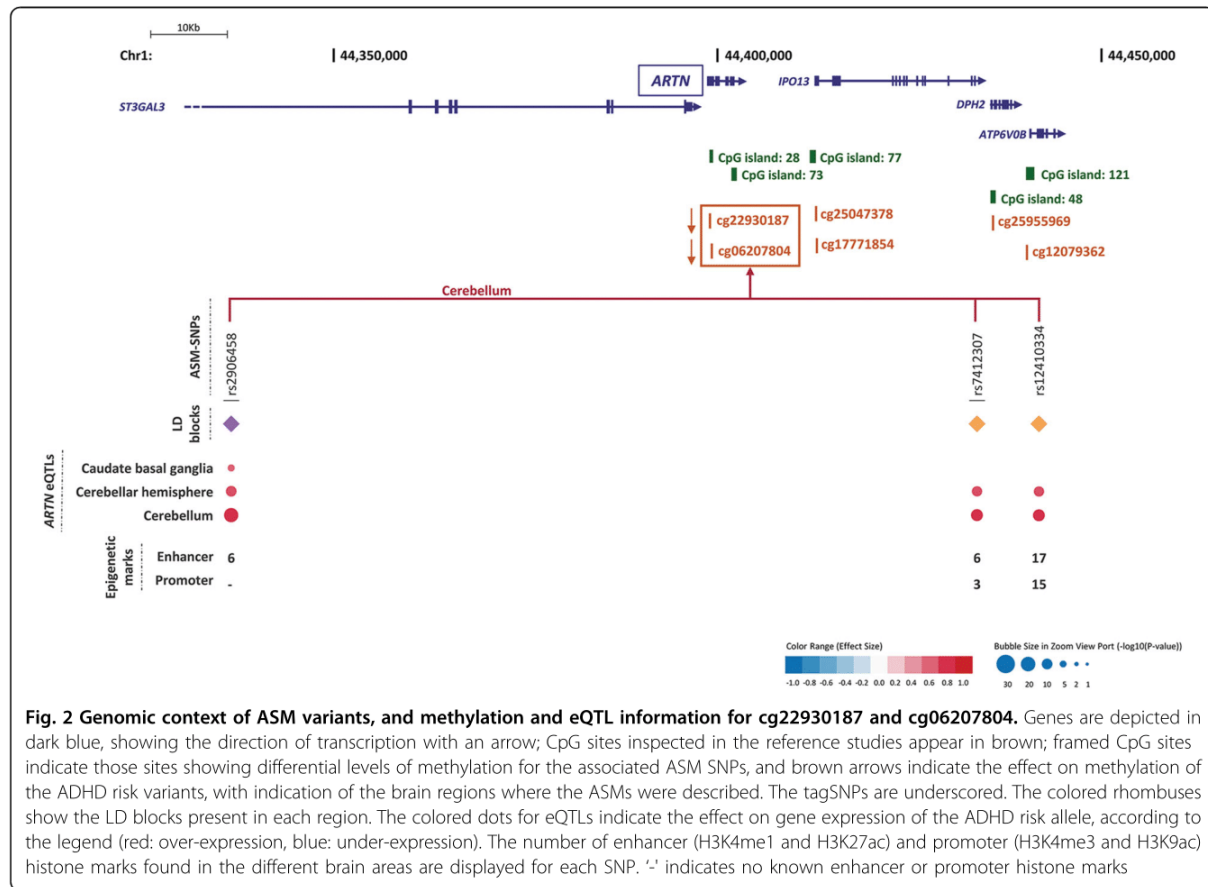
Results

We investigated the possible association with ADHD of SNPs that show ASM in brain regions. Starting from two previous studies^{21,22} that describe ASM in brain tissues we obtained 43,132 SNP–CpG pairs involving 33,944 SNPs and 5306 CpG sites (Figs. 1 and S1). Genetic risk variants for ADHD are enriched in those ASM SNPs, as observed through enrichment analysis at different association *p*-value thresholds (Table S1).

We detected some overlaps and redundancies between studies and tissues (Fig. S1), so we performed a selection process ending up with a list of 3896 ASM tagSNPs (Figs. 1 and S2). Eight ASM tagSNPs were significantly associated with ADHD after correcting for multiple comparisons (5% FDR, $p \leq 6.78E-05$) (Fig. 1 and Table S2). These eight SNPs correlated with differential methylation at six CpG sites in *cis* (three for cg20225915, two for both cg22930187 and cg06207804, and one for each of cg13047596, cg11554507, and cg04464446) in different brain areas (Figs. 2–4 and regional associational plots Figs. S3–S10, Table S2). Three of the eight ASM tagSNPs remained associated with ADHD after applying the Bonferroni and GEC corrections, all correlating with differential methylation at the cg20225915 site (Table S2).

As considering only tagSNPs may overlook true causal variants, we retrieved association results from all the 52 ASM SNPs tagged by the previous ones (LD; $r^2 \geq 0.85$), ending up with 60 variants in eight LD blocks that show association with ADHD and correlate with methylation levels at six CpG sites (Figs. 2–4 and S11–S15 and Table S3). We also selected, for each LD block, the SNP showing the highest number of functional annotations (Table 1), as a putative causal SNP.

Consistently, the direction of the effect of the risk alleles on methylation levels is the same for all the SNPs correlating with the same CpG site. Thus, the risk alleles correlate with decreased methylation of cg22930187, cg06207804, cg11554507 and cg20225915 and with increased methylation of cg13047596 and cg04464446^{21,22} (Figs. 2–4 and Tables 1, S2, and S3).



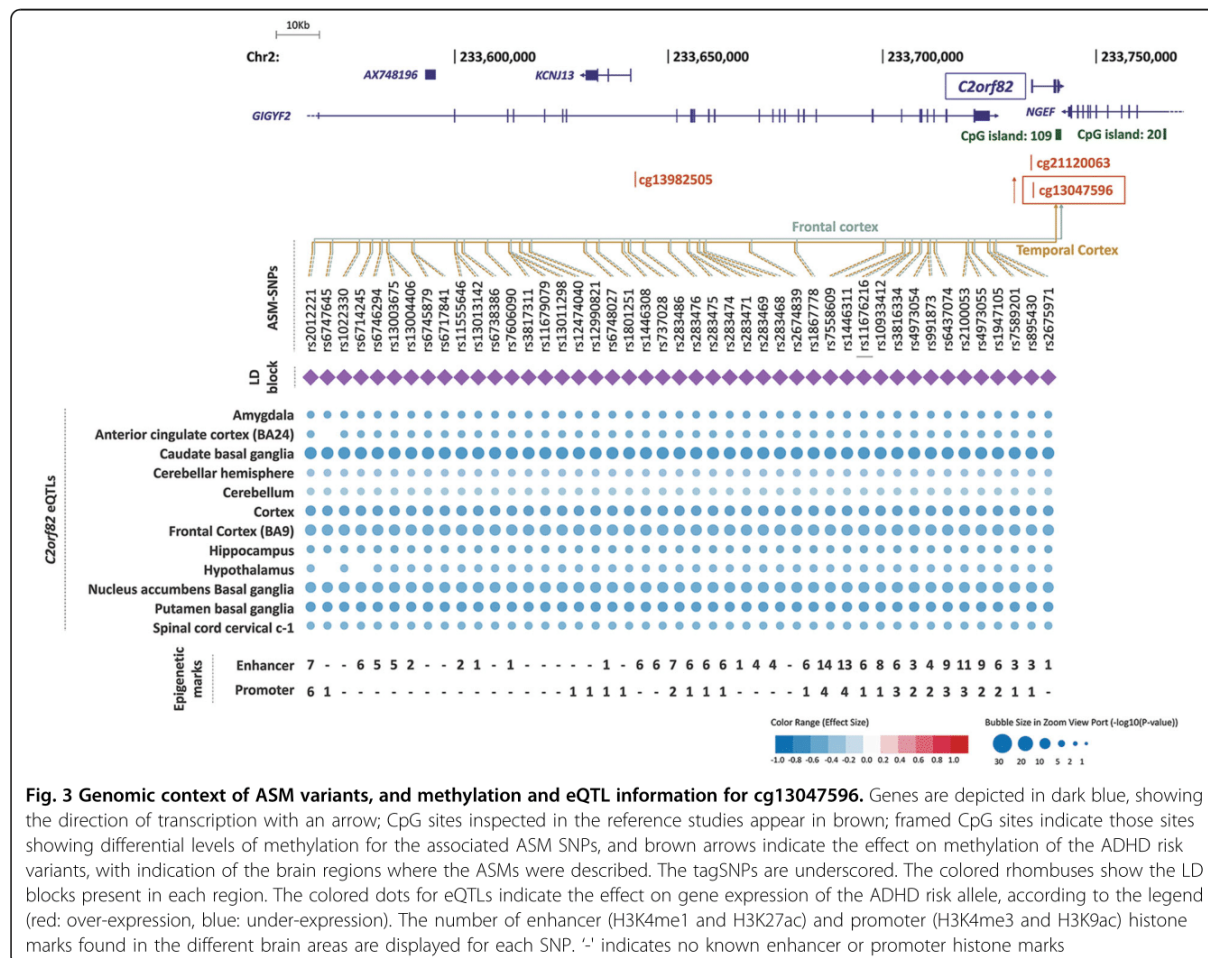
All six CpG sites are located in possible promoter regions (<5000bp upstream from a TSS) of six genes (Table 1), all of them expressed in brain: *ARTN* (cg22930187 and cg06207804), *C2orf82* (cg13047596), *NEUROD6* (cg11554507), *PIDDI* (cg20225915), *RPLP2* (cg20225915), and *GAL* (cg04464446) (Figs. 2–4). Furthermore, 85% of the 60 ASM SNPs are located within a region with enhancer or promoter histone marks in at least one brain area (Figs. 2–4 and Tables S4–S8). All putative causal SNPs selected from each LD block lie within a region with histone marks, ranging from 3 to 17 in enhancer regions and from 4 to 16 in promoter regions (Table 1).

We subsequently assessed the possible effect of those 60 SNPs on gene expression and observed that 57 of them are eQTLs for different genes in brain regions (Table S3). Seven out of the eight putative causal SNPs are eQTLs in brain for at least one gene (Table 1). We focused on methylation in promoter regions, which is well established to inversely correlate with gene expression. The eQTLs for *ARTN*, *C2orf82*, and *PIDDI* correlated with methylation of CpG sites lying on their possible promoter regions, showing opposite directions for methylation and gene

expression levels (Figs. 2–4 and Tables 1 and S3). The ADHD risk alleles are associated with increased expression of *ARTN* (in cerebellum and a subcortical region) and *PIDDI* (in cerebellum and cortex) and with decreased expression of *C2orf82* (in cortical, subcortical, and cerebellar regions) (Figs. 2–4 and Tables 1 and S3).

Consistently, the predicted direction of the effect on gene expression for these three genes is the same when we consider all variants within ±1 MB from the TSS (and not only the ASM SNPs). We found significant associations of gene expression with ADHD for the same three genes in multiple brain tissues using MetaXcan: *ARTN*, *PIDDI* showed increased expression (3.57 < Z-score < 4.19 and 3.57 < Z-score < 5.37, respectively) and *C2orf82* with a decreased expression (−3.64 < Z-score < −3.07) (Table S7), all of them surviving the Bonferroni correction.

We also evaluated the correlation of the 60 ADHD-associated SNPs with subcortical brain volume changes in ENIGMA2 data. SNPs correlating with methylation at cg13047596 and at cg04464446 correlate with nucleus accumbens and/or caudate nucleus volumes, while the only SNP correlating with cg11554507, which is present in ENIGMA2, correlates with thalamus volume (Table S10).



Three of the putative causal SNPs showed correlation with brain volumes (Table 1).

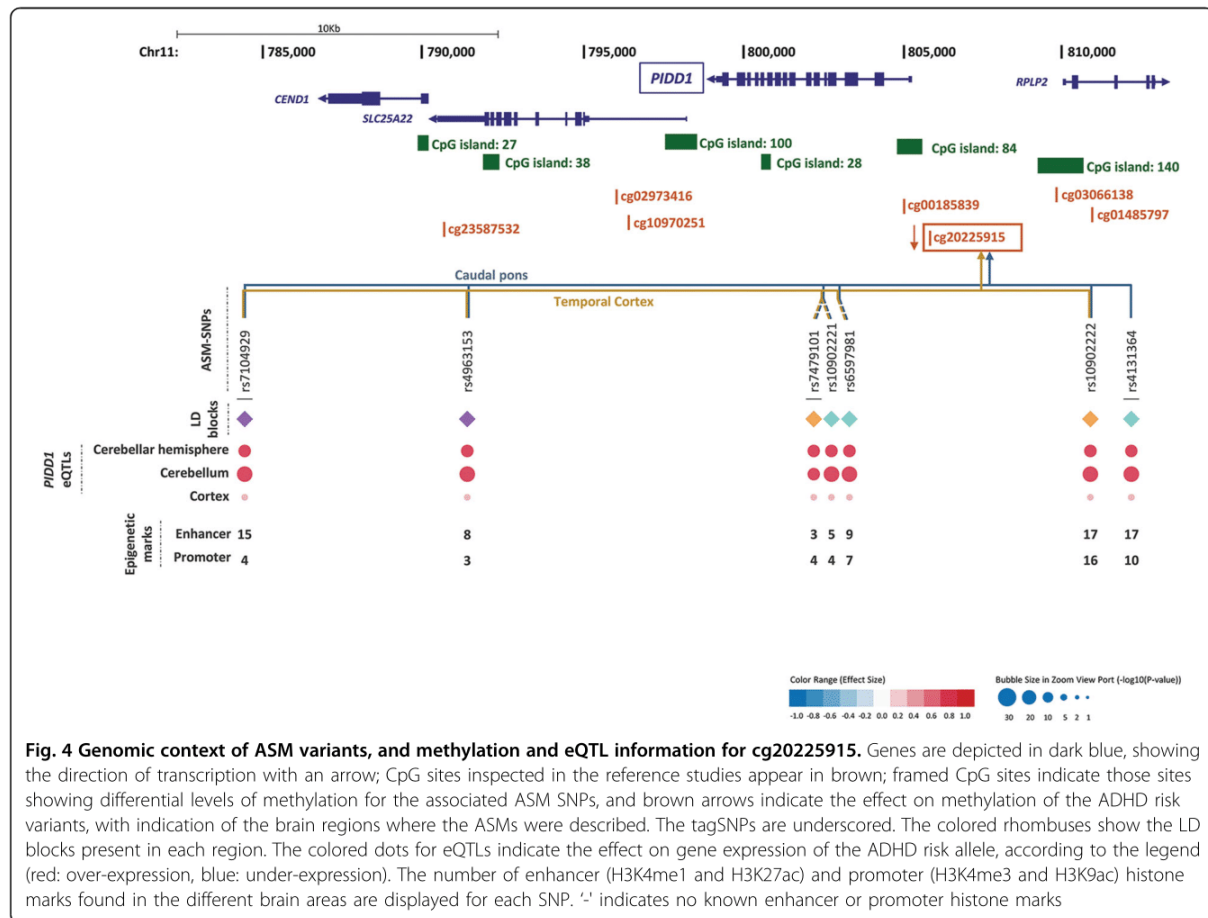
Interestingly, the majority of ASM SNPs that correlate with methylation levels of cg13047596, located in the promoter region of *C2orf82*, are eQTLs in brain for this gene, lie in a region with histone marks and correlate with volume changes of nucleus accumbens and caudate nucleus (Figs. 2–4, Tables 1 and S3–S8 and S10). All this functional evidence highlights the *C2orf82* gene as a good candidate for contributing to ADHD.

Discussion

This study is the first comprehensive assessment of the contribution to ADHD of genetic variants altering methylation in the brain. We identified a total of 60 variants from eight LD blocks associated with ADHD that correlate with differential levels of methylation at six different CpG sites^{21,22} (Tables 1 and S3). All the variants from six out of the eight LD blocks alter the methylation of CpG sites lying at potential promoter regions and are also eQTLs for one of the following three genes in

multiple brain regions: *ARTN*, *C2orf82*, and *PID1* (Figs. 2–4 and Tables 1 and S3). It is well known that DNA methylation in promoter regions inversely correlates with levels of gene expression¹⁸, and all these ASM variants associated with ADHD in our study are concordant with this statement.

The *ARTN* gene, highlighted by two tagSNPs, encodes Artemin, a ligand of the *GDNF* family (glial cell line-derived neurotrophic factor). Artemin supports the survival of sensory and sympathetic peripheral neurons in culture by interacting with GFRα3-RET and possibly also of dopaminergic neurons of the ventral mid-brain through activation of GFRα1-RET complex³⁴. Gene Ontology (GO) pathways link it to key neurodevelopmental functions: axon guidance (GO:0007411), neuroblast proliferation (GO:0007405), and peripheral nervous system development (GO:0007422). Risk alleles for ADHD lead to an overexpression of *ARTN*. Previously, overexpression of *ARTN* has been studied in transgenic mice and been linked to an increase of neuron excitability that leads to hypersensitivity^{35,36}. Another study in *ARTN*



knockout mice reported aberrations in the sympathetic nervous system related to migration and axonal projection³⁷. The *C2orf82* gene (also known as *SNORC*) was highlighted by one tagSNP and it encodes a proteoglycan transmembrane protein that is expressed in brain more than in other tissues³⁰. Little is known about its function. Finally, *PIDD1* was highlighted by three tagSNPs. It is a cell life regulator gene and it has been linked to apoptotic and anti-apoptotic pathways. The PIDD protein initiates apoptosis as a component of the PIDDosome together with RAIDD (RIP-associated ICH-1/ECD3-homologous protein with a death domain) and procaspase-2³⁸ and it also activates an anti-apoptotic pathway involving the transcription factor NF-κB in response to genotoxic stress³⁹.

Alterations in the expression of these three genes (upregulation of *ARTN* and *PIDD1* and downregulation of *C2orf82*) in different brain regions seem to be related to ADHD. Interestingly, most of these regions are relevant for this disorder. Neuroimaging studies have implicated the cerebellum, subcortical and prefrontal regions in ADHD, suggesting a link to problems in the

processing of temporal information⁴⁰. Structural anomalies in the cerebellum have been reported in ADHD individuals through neuroimaging studies^{41–43}. Cerebellar developmental trajectories and hippocampal volumes are linked to the severity of ADHD symptoms^{44–46}. Structural and functional abnormalities in cerebellum and basal ganglia have been associated with motor impairments⁴⁷, which are frequent in nearly half of ADHD cases⁴⁸. Subcortical regions identified through our expression analyses have also been related to ADHD, for instance: (i) remarkably different shapes of caudate-putamen basal ganglia and smaller volumes have been reported in ADHD boys^{49–52}; (ii) in adult males with ADHD, right caudate volume correlates with poor accuracy on sensory selection tasks⁵³ and also with hyperactivity/impulsivity^{54,55}; (iii) nucleus accumbens, caudate nucleus, putamen, amygdala, and hippocampus are structurally altered in the brains of ADHD patients⁵⁶. Remarkably, all the ASM SNPs in the LD block for *C2orf82* with available information nominally correlate with increased volumes of nucleus accumbens and caudate nucleus subcortical regions. Also, the eQTL

Table 1 Selection of putative causal ASM SNPs associated with ADHD according to functional annotations

SNP	Association with ADHD ^a		Effect on methylation ^b	Epigenetic marks ^c		Effect on expression (GTEx data) ^d	Effect on brain volumes ^e
	Risk allele	p-value		Enhancer	Promoter		
rs2906458	G	3.01E-05	↓ cg22930187, ↓ cg06207804	6	0	↑ ARTN	–
rs12410334	A	2.87E-05		17	15		–
rs7558609	A	7.06E-05	↑ cg13047596	14	4	↓ C2orf82	↑ NAc ↑ CN
rs4140961	A	6.05E-05	↓ cg11554507	3	0	–	↑ T
rs7104929	G	<u>7.89E-06</u>	↓ cg20225915	15	4	↑ PIDD1 ↓ <i>PNPLA2</i>	?
rs10902222	T	<u>2.03E-06</u>		17	16		–
rs4131364	A	<u>1.60E-06</u>		17	10		–
rs1054252	G	3.86E-05	↑ cg04464446	4	0	↑ <i>MRPL21</i> , ↑ <i>MRGPRD</i> ↓ <i>IGHMBP2</i>	↓ NAc ↓ CN

ASM: Allele-specific methylation, SNP: single nucleotide polymorphism, NAc: nucleus accumbens, CN: caudate nucleus, T: thalamus. Risk allele: all alleles are reported in the forward strand; Underlined: significant associations between ASM tagSNPs and ADHD overcoming Bonferroni correction for multiple testing and p-value threshold determined using independent number of tests (GEC); ↑: Hypermethylation/overexpression/increased brain volume; ↓: Hypomethylation/underexpression/decreased brain volume; “–”: No significant data for the SNP; “?”: No information available for the SNP; Enhancer: Number of H3K4me1 and H3K27ac marks; Promoter: Number of H3K4me3 and H3K9ac marks; In bold: genes with the reported CpG sites lying in their possible promoter region

^aData obtained from the PGC+IPSYCH ADHD GWAS meta-analysis⁴

^bDescribed in Zhang et al.²² and Gibbs et al.²¹

^cHistone marks found in brain areas

^deQTL information for brain tissues

^eData from the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) consortium³³

effect sizes of these SNPs are the largest for caudate basal ganglia, which volume correlates with the SNP genotype variation. There is evidence about the role in ADHD of cortical thickness, cortical volume and functional connectivity in the anterior cingulate cortex, a region involved in cognitive control, attention, affect and drive^{57–63}. Furthermore, delayed cortical development, e.g. in prefrontal regions has been reported in ADHD patients^{64,65} and this appears to be stronger in ADHD children with below median intelligence quotient⁶⁶. All the above mentioned fronto-subcortical structures and pathways are rich in catecholamines, the molecular targets in pharmacological treatments for ADHD^{48,52,64,67}.

Interestingly, the methylation of cg20225915 has also been associated with *PIDD1* expression in peripheral blood⁶⁸, turning it into a good candidate as a biomarker. The expression of *ARTN* was found to be altered in blood of major depressive disorder (MDD) patients⁶⁹ and the *C2orf82* gene has been associated to schizophrenia^{70,71}. Furthermore, *C2orf82* was highlighted in a cross-disorder GWAS of eight psychiatric conditions, including ADHD and schizophrenia⁷², with the rs778353 lead SNP, located 47 kb downstream from the gene, showing a genome-wide significant association with the phenotype. All three genes overlap with several CNVs that contribute to autism, intellectual disability or aggressive behavior, conditions often comorbid with ADHD (Table S11). It is noteworthy

that some of the CNVs reported in *ARTN*, *C2orf82*, and *PIDD1* are related to brain-specific and overall developmental delay at both fetal and postnatal stages. Thus, it is reasonable to assume that altered expression of these genes might affect brain volumes and cognition. Overall, the fact that these genes have previously been related to neuropsychiatric disorders that are often comorbid with ADHD⁷³ make them appealing candidates to be pursued.

ARTN is the only gene highlighted in our study that is present in one of the top regions reported in the ADHD GWAS meta-analysis⁴, although it did not contain SNPs surviving genome-wide significance. The GWAS findings in the region could be accounted for by one of several genes: *ST3GAL3*, *PTPRF*, *KDM4A*, *RP11-184I16.4*, *XR_246316.1*, *KDM4A-AS1*, and *SLC6A9*. *ST3GAL3* had the most signals. Although two of the reported ASM variants associated with ADHD are intronic to *ST3GAL3*, this gene was not highlighted in our study as none of the associated variants correlated with differential methylation of CpG sites near the *ST3GAL3* TSS (distance from the nearest CpG site: 197 kb) or were eQTLs for the gene in brain tissues. Instead, these SNPs correlated with a nearby gene, *ARTN*, both in terms of methylation and gene expression. This suggests the importance of finding functional connections between disease-associated SNPs and genes, besides considering the genes in the physical vicinity of variants. Furthermore, another of the highlighted genes, *PIDD1*, although not being among the top

findings in the ADHD GWAS meta-analysis⁴, it is pointed out by the gene-based association analysis performed in the same study.

Genetic variants surpassing genome-wide significance in GWAS explain only a small part of the SNP-based heritability and associations not reaching the significance threshold also contribute to disease susceptibility^{4,9}. An omnigenic model has been recently proposed suggesting that the sub-threshold variants could point at regulatory elements of core genes^{7,8}. Indeed, a previous study on a cardiovascular cardiac phenotype reported that nominally significant associations are enriched in enhancer regions⁹, consistent with our findings. Therefore, although none of the variants that we identified in our study display genome-wide significant association with ADHD, they may contribute to the susceptibility to ADHD, as they do have a functional impact (methylation, expression, and in some cases brain structure) via genes that are expressed in brain.

Brain-specific ASM information has also been utilized to detect key genes and pathways in BD²⁰. Also, a higher enrichment of brain ASM was observed in a schizophrenia GWAS in comparison to non-psychiatric GWAS¹⁰. This, together with the enrichment of ASM in ADHD-associated variants found in the present study, reinforces the rationale of utilizing ASM SNPs to highlight genes that are relevant to psychiatric disorders from GWAS data.

There are some strengths and limitations in our study that should be discussed. Strengths: (i) We used the largest GWAS meta-analysis of ADHD performed so far, including around 20,000 cases and 35,000 controls. (ii) The genetic variants identified as associated with ADHD have a functional impact on epigenetic regulation, expression or brain volumes. (iii) Two of the highlighted genes in this study, *ARTN* and *C2orf82*, have previously been associated with other psychiatric disorders. (iv) For two of the genes there is more than one LD block showing the same effect on CpG site methylation. (v) Our results are concordant with eQTL information that had been assessed in an independent sample, with all the SNPs showing the opposite effect on methylation of the promoter region and on the expression of a given gene in brain (more promoter methylation and less gene expression or vice versa), even for the different LD blocks from each region. Limitations: (i) We did not perform a follow-up study to replicate the association findings in an independent sample. (ii) The previous studies that we used for the selection of ASM SNPs were performed on different genotyping platforms that do not include all the existing SNPs in the genome, and therefore we could not test all possible ASMs. (iii) We only considered *cis*-associated ASM variants, which are the vast majority, although non-*cis* ASM also occurs. (iv) There is an overrepresentation of

ASM SNPs from cerebellum compared to the other studied tissues.

To conclude, the present study points to the *ARTN*, *C2orf82*, and *PIDDI* genes as potential contributors to ADHD susceptibility. The identified risk variants have an impact on the methylation levels of different CpG sites located in promoter regions and they inversely correlate with expression of the corresponding genes in brain. This finding is supported by a prediction of increased expression of *ARTN* and *PIDDI*, and a decreased expression of *C2orf82* in ADHD. Moreover, variants correlating with methylation at cg13047596 (near *C2orf82*) influence the volumes of nucleus accumbens and/or caudate nucleus. Further studies are required to elucidate the mechanisms by which these genes contribute to ADHD.

Acknowledgements

Major financial support for this research was received by B.C. from the Spanish 'Ministerio de Economía y Competitividad' (SAF2015-68341-R, RTI2018-100968-B-I00) and AGAUR, 'Generalitat de Catalunya' (2017-SGR-738). The research leading to these results has also received funding from the European Union Seventh Framework Program [FP7/2007–2013] under grant agreement no. 602805 and from the European Union H2020 Program [H2020/2014-2020] under grant agreements nos. 667302 and 643051, the latter supporting the contract of A.S. L.P.-C. and J.C.-D. were supported by 'Generalitat de Catalunya' (2016 FI_B 00728 and 2015 FI_B 00448, respectively). L.P.-C. was also supported by 'Ministerio de Educación, Cultura y Deporte' (FPU15/03867). N.F.-C. was supported by contracts of the 'Centro de Investigación Biomédica en Red de Enfermedades Raras' (CIBERER). V.R. was supported by the Graduate School of Health from the University of Aarhus. The iPSYCH team acknowledges support from the Lundbeck Foundation. Finally, S.F. was supported by the European Union's Seventh Framework Program for research, technological development and demonstration under grant agreement no. 602805, the European Union's Horizon 2020 research and innovation program under grant agreements nos. 667302 and 728018 and NIMH grants 5R01MH101519 and U01 MH109536-01. We are thankful to Roser Corominas (Universitat de Barcelona) for helpful advice. We are also grateful to the ADHD Working Group of the Psychiatric Genomics Consortium (PGC) and the iPSYCH team for distributing the summary statistics of the ADHD GWAS meta-analysis. This study is part of the International Multicentre persistent ADHD Collaboration (ImpACT); www.impactadhdgenomics.com). ImpACT unites major research centres working on the genetics of ADHD persistence across the lifespan and has participants in The Netherlands, Germany, Spain, Norway, the United Kingdom, the United States, Brazil and Sweden. Principal investigators of ImpACT are: Barbara Franke (chair), Andreas Reif (co-chair), Stephen V. Faraone, Jan Haavik, Bru Cormand, J. Antoni Ramos-Quiroga, Marta Ribasés, Philip Asherson, Klaus-Peter Lesch, Jonna Kuntsi, Claiton H.D. Bau, Jan Buitelaar, Alejandro Arias-Vásquez, Tetyana Zayats, Henrik Larsson, Alysya Doyle, and Eugenio H. Grevet. Access to the PGC ADHD data was obtained through dbGaP project number 10608 that includes the following datasets: phs000016.v2.p2, phs000407.v1.p1, phs000358.v1.p1, and phs000490.v1.p1. We thank the ENIGMA consortium for sharing the summary statistics of genome-wide association meta-analyses of MRI phenotypes.

Author details

¹Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain. ²Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Madrid, Spain. ³Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Catalonia, Spain. ⁴Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Barcelona, Catalonia, Spain. ⁵The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus, Denmark. ⁶Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark. ⁷Department of Biomedicine - Human Genetics, Aarhus University, Aarhus, Denmark. ⁸Departments of Psychiatry and Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41398-019-0574-7>).

Received: 15 August 2018 Revised: 3 July 2019 Accepted: 30 July 2019

Published online: 03 October 2019

References

- Polanczyk, G., De Lima, M. S., Horta, B. L., Biederman, J. & Rohde, L. A. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am. J. Psychiatry* **164**, 942–948 (2007).
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Arlington, VA, US: American Psychiatric Publishing, Inc. (2013).
- Faraone, S. V. et al. Molecular genetics of attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **57**, 1313–1323 (2005).
- Demontis, D. et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
- Anttila, V. et al. Analysis of shared heritability in common disorders of the brain. *Science* **360**, eaap8757 (2018).
- Meaburn, E. L., Schalkwyk, L. C. & Mill, J. Allele-specific methylation in the human genome: implications for genetic studies of complex disease. *Epigenetics* **5**, 578–582 (2010).
- Boyle, E. A., Li, Y. I. & Pritchard, J. K. An expanded view of complex traits: from polygenic to omnigenic. *Cell* **169**, 1177–1186 (2017).
- Liu, X., Li, Y. I. & Pritchard, J. K. Trans effects on gene expression can drive omnigenic inheritance. *Cell* **177**, 1022–1034 (2019).
- Wang, X. et al. Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures. *Elife* **5**, e10557 (2016).
- Gagliano, S. A. et al. Allele-skewed DNA modification in the brain: relevance to a Schizophrenia GWAS. *Am. J. Hum. Genet.* **98**, 956–962 (2016).
- Ladd-Acosta, C. et al. DNA methylation signatures within the human brain. *Am. J. Hum. Genet.* **81**, 1304–1315 (2007).
- Jang, H. S., Shin, W. J., Lee, J. E. & Do, J. T. CpG and non-CpG methylation in epigenetic gene regulation and brain function. *Genes* **8**, E148 (2017).
- Xu, Y. et al. Multiple epigenetic factors predict the attention deficit/hyperactivity disorder among the Chinese Han children. *J. Psychiatr. Res.* **64**, 40–50 (2015).
- Schanen, N. C. Epigenetics of autism spectrum disorders. *Hum. Mol. Genet.* **15**, R138–R150 (2006).
- Grayson, D. R. & Guidotti, A. Merging data from genetic and epigenetic approaches to better understand autistic spectrum disorder. *Epigenomics* **8**, 85–104 (2016).
- Dammann, G. et al. Increased DNA methylation of neuropsychiatric genes occurs in borderline personality disorder. *Epigenetics* **6**, 1454–1462 (2011).
- Gaur, U., Li, K., Mei, S. & Liu, G. Research progress in allele-specific expression and its regulatory mechanisms. *J. Appl. Genet.* **54**, 271–283 (2013).
- Li, E. & Zhang, Y. DNA methylation in mammals. *Cold Spring Harb. Perspect. Biol.* **6**, a019133 (2014).
- Do, C. et al. Genetic-epigenetic interactions in cis: a major focus in the post-GWAS era. *Genome Biol.* **18**, 120 (2017).
- Chuang, L.-C., Kao, C.-F., Shih, W.-L. & Kuo, P.-H. Pathway analysis using information from allele-specific gene methylation in genome-wide association studies for bipolar disorder. *PLoS ONE* **8**, e53092 (2013).
- Gibbs, J. R. et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6**, e1000952 (2010).
- Zhang, D. et al. Genetic control of individual differences in gene-specific methylation in human brain. *Am. J. Hum. Genet.* **86**, 411–419 (2010).
- Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- Barrett, J. C. Haploview: visualization and analysis of snp genotype. *Data. Cold Spring Harb. Protoc.* **4**, 1–5 (2009).
- The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
- R Core Team. *R: A Language and Environment for Statistical Computing*. <https://www-project.org/> (2014).
- Storey, John D. with contributions from Bass, Andrew J. AD and, Robinson D. qvalue: Q-value estimation for false discovery rate control. R Package Version 2.6.0. <http://github.com/jdstorey/qvalue> (2015).
- Li, M.-X., Yeung, J. M. Y., Cherny, S. S. & Sham, P. C. Evaluating the effective numbers of independent tests and significant *p*-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum. Genet.* **131**, 747–756 (2012).
- Ward, L. D. & Kellis, M. HaploRegv4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* **44**, D877–D881 (2016).
- GTEx Consortium et al. The Genotype-Tissue Expression (GTEx) project. *Nat. Publ. Gr.* **45**, 580–585 (2013).
- Barbeira, A. N. et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).
- Fromer, M. et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* **19**, 1442–1453 (2016).
- Hibar, D. P. et al. Common genetic variants influence human subcortical brain structures. *Nature* **520**, 224–229 (2015).
- Baloh, R. H. et al. Artemin, a novel member of the GDNF ligand family, supports peripheral and central neurons and signals through the GFRα3-RET receptor complex. *Neuron* **21**, 1291–1302 (1998).
- Elitt, C. M. et al. Artemin overexpression in skin enhances expression of TRPV1 and TRPA1 in cutaneous sensory neurons and leads to behavioral sensitivity to heat and cold. *J. Neurosci.* **26**, 8578–8587 (2006).
- Elitt, C. M., Malin, S. A., Koerber, H. R., Davis, B. M. & Albers, K. M. Overexpression of artemin in the tongue increases expression of TRPV1 and TRPA1 in trigeminal afferents and causes oral sensitivity to capsaicin and mustard oil. *Brain Res.* **1230**, 80–90 (2008).
- Honma, Y. et al. Artemin is a vascular-derived neurotrophic factor for developing sympathetic neurons. *Neuron* **35**, 267–282 (2002).
- Tinel, A. & Tschopp, J. The PIDDosome, a protein complex implicated in activation of caspase-2 in response to genotoxic stress. *Science* **304**, 843–846 (2004).
- Janssens, S., Tinel, A., Lippens, S. & Tschopp, J. PIDD mediates NF-kappaB activation in response to DNA damage. *Cell* **123**, 1079–1092 (2005).
- Toplak, M. E., Dockstader, C. & Tannock, R. Temporal information processing in ADHD: findings to date and new methods. *J. Neurosci. Methods* **151**, 15–29 (2006).
- Berquin, P. C. et al. Cerebellum in attention-deficit hyperactivity disorder: a morphometric MRI study. *Neurology* **50**, 1087–1093 (1998).
- Castellanos, F. X. et al. Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. *Arch. Gen. Psychiatry* **53**, 607–616 (1996).
- Castellanos, F. X. et al. Quantitative brain magnetic resonance imaging in girls with attention-deficit/hyperactivity disorder. *Arch. Gen. Psychiatry* **58**, 289–295 (2001).
- Mackie, S. et al. Cerebellar development and clinical outcome in attention deficit hyperactivity disorder. *Am. J. Psychiatry* **164**, 647–655 (2007).
- Plessen, K. J. et al. Hippocampus and amygdala morphology in attention-deficit/hyperactivity disorder. *Arch. Gen. Psychiatry* **63**, 795–807 (2006).
- Friedman, L. A. & Rapoport, J. L. Brain development in ADHD. *Curr. Opin. Neurobiol.* **30**, 106–111 (2015).
- Pasini, A. & D'Agati, E. Pathophysiology of NSS in ADHD. *World J. Biol. Psychiatry* **10**, 495–502 (2009).
- Curatolo, P., D'Agati, E. & Moavero, R. The neurobiological basis of ADHD. *Ital. J. Pediatr.* **36**, 79 (2010).
- Qiu, A. et al. Basal ganglia volume and shape in children with attention deficit hyperactivity disorder. *Am. J. Psychiatry* **166**, 74–82 (2009).
- Greven, C. U. et al. Developmentally stable whole-brain volume reductions and developmentally sensitive caudate and putamen volume alterations in those with attention-deficit/hyperactivity disorder and their unaffected siblings. *JAMA Psychiatry* **72**, 490–499 (2015).
- Sobel, L. J. et al. Basal ganglia surface morphology and the effects of stimulant medications in youth with attention deficit hyperactivity disorder. *Am. J. Psychiatry* **167**, 977–986 (2010).

52. Rubia, K., Alegria, A. A. & Brinson, H. Brain abnormalities in attention-deficit hyperactivity disorder: a review. *Rev. Neurol.* **58**(Suppl. 1), S3–S16 (2014).
53. Casey, B. J. et al. Implication of right frontostriatal circuitry in response inhibition and attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **36**, 374–383 (1997).
54. Singh, A., Yeh, C. J., Verma, N. & Das, A. K. Overview of attention deficit hyperactivity disorder in young children. *Heal Psychol. Res.* **3**, 23–35 (2015).
55. Onnink, A. M. H. et al. Brain alterations in adult ADHD: Effects of gender, treatment and comorbid depression. *Eur. Neuropsychopharmacol.* **24**, 397–409 (2014).
56. Hoogman, M. et al. Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *Lancet Psychiatry* **4**, 310–319 (2017).
57. Shaw, P. et al. Longitudinal mapping of cortical thickness and clinical outcome in children and adolescents with attention-deficit/hyperactivity disorder. *Arch. Gen. Psychiatry* **63**, 540–549 (2006).
58. Ambrosino, S., de Zeeuw, P., Wierenga, L. M., van Dijk, S. & Durston, S. What can cortical development in attention-deficit/hyperactivity disorder teach us about the early developmental mechanisms involved? *Cereb. Cortex* **27**, 4624–4634 (2017).
59. Castellanos, F. X. & Aoki, Y. Intrinsic functional connectivity in attention-deficit/hyperactivity disorder: a science in development. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **1**, 253–261 (2016).
60. Makris, N. et al. Cortical thinning of the attention and executive function networks in adults with attention-deficit/hyperactivity disorder. *Cereb. Cortex* **17**, 1364–1375 (2007).
61. Makris, N. et al. Anterior cingulate volumetric alterations in treatment-naïve adults with ADHD: a pilot study. *J. Atten. Disord.* **13**, 407–413 (2010).
62. Yang, X.-R., Carrey, N., Bernier, D. & MacMaster, F. P. Cortical thickness in young treatment-naïve children with ADHD. *J. Atten. Disord.* **19**, 925–930 (2015).
63. Qiu, M. et al. Changes of brain structure and function in ADHD children. *Brain Topogr.* **24**, 243–252 (2011).
64. Cortese, S. The neurobiology and genetics of Attention-Deficit/Hyperactivity Disorder (ADHD): what every clinician should know. *Eur. J. Paediatr. Neurol.* **16**, 422–433 (2012).
65. Shaw, P. et al. Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proc. Natl Acad. Sci. USA* **104**, 19649–19654 (2007).
66. de Zeeuw, P. et al. Differential brain development with low and high IQ in attention-deficit/hyperactivity disorder. *PLoS ONE* **7**, e35770 (2012).
67. Pliszka, S. R., Lancaster, J., Liotti, M. & Semrud-Clikeman, M. Volumetric MRI differences in treatment-naïve vs chronically treated children with ADHD. *Neurology* **67**, 1023–1027 (2006).
68. Wu, Y. et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat. Commun.* **9**, 918 (2018).
69. Otsuki, K. et al. Altered expression of neurotrophic factors in patients with major depression. *J. Psychiatr. Res.* **42**, 1145–1153 (2008).
70. Schizophrenia Working Group of the Psychiatric Genomics. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
71. Ripke, S. et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat. Genet.* **45**, 1150–1159 (2013).
72. Lee, P. H. et al. Genome wide meta-analysis identifies genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *bioRxiv*. 528117 (2019).
73. Brown T. E. *ADHD Comorbidities: Handbook for ADHD Complications in Children and Adults*. Washington, DC, US: American Psychiatric Press (2009).

Supplementary Material

EXPLORING GENETIC VARIATION THAT INFLUENCES BRAIN

METHYLATION IN ATTENTION-DEFICIT/HYPERACTIVITY

DISORDER

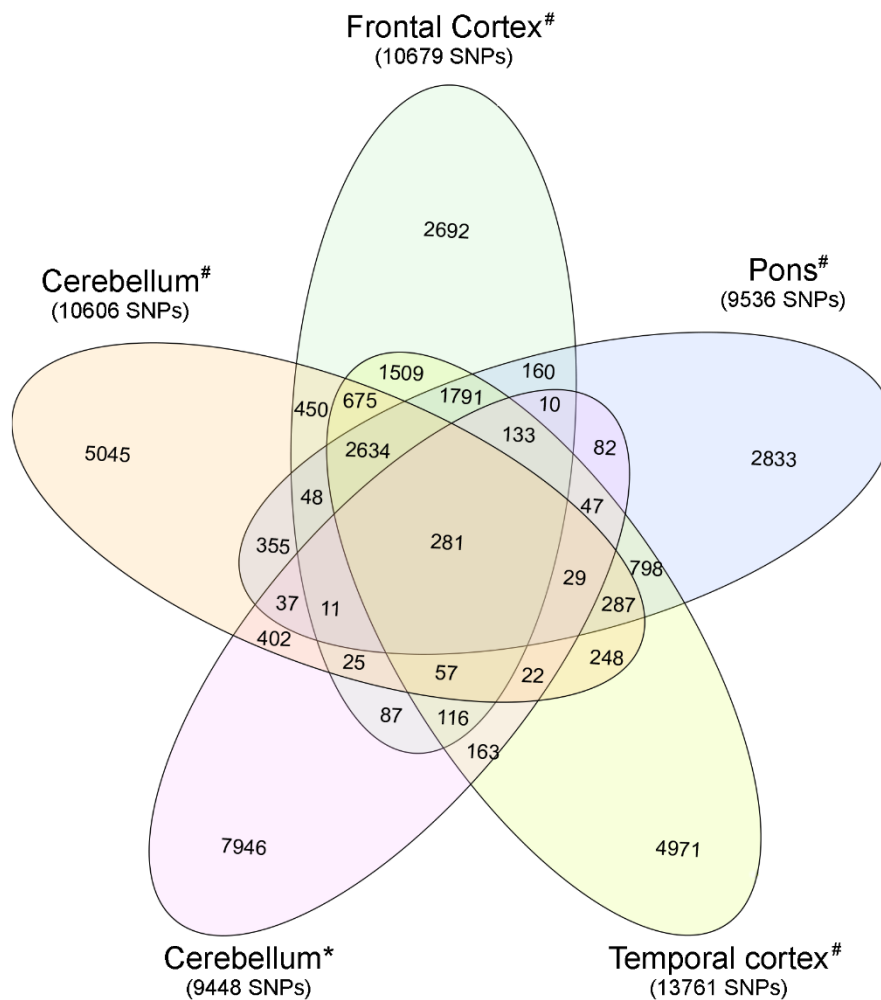


Figure S1. Venn diagram showing the overlaps among the initial 33,944 ASM SNPs selected in the different brain areas. *Cerebellum from the study by Zhang *et al.* 2010. #Brain areas from the study by Gibbs *et al.* 2010. The overlap accounts for 31% of the ASM SNPs; 12.5% ASM SNPs are shared between two tissues, 9% are shared between three tissues, 8.5% are shared between four tissues and only 1% are shared between all the tissues. Sixteen percent of the 9,448 ASM SNPs identified by Zhang *et al.*, 2010 overlap with the ASM SNPs identified in the Gibbs *et al.*, 2010 study.

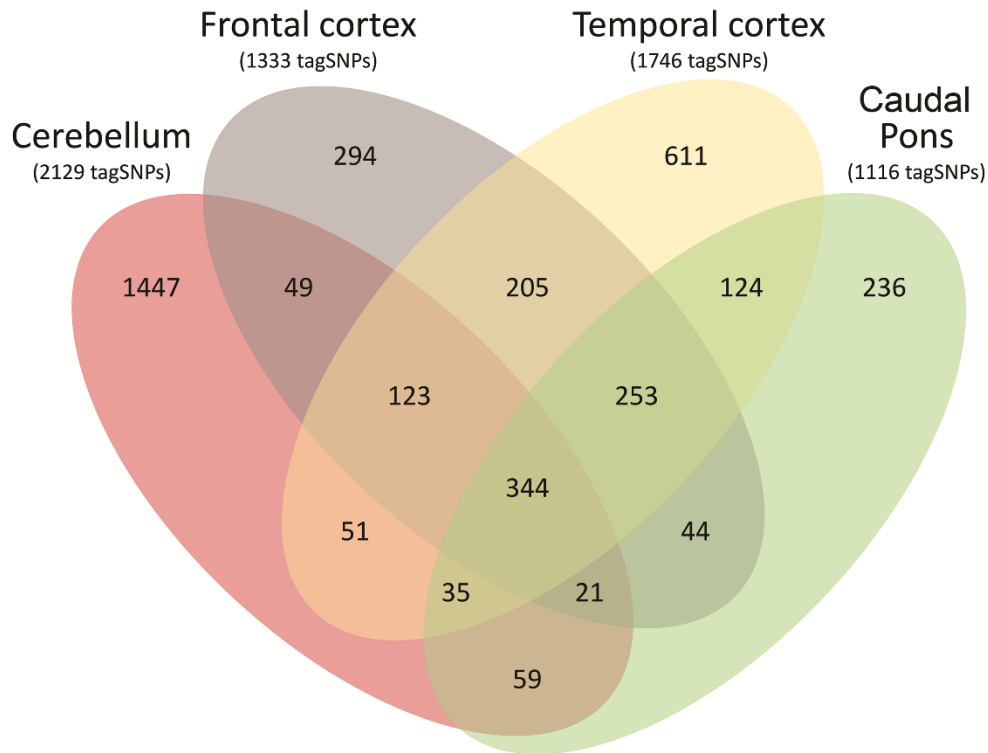


Figure S2. Venn diagram showing the overlaps among the 3,896 ASM tagSNPs selected in the different brain areas. The overlap accounts for 33.5% of tagSNPs; 13.5% tagSNPs are shared between two tissues, 11% are shared between three tissues and 9% are shared between all the tissues.

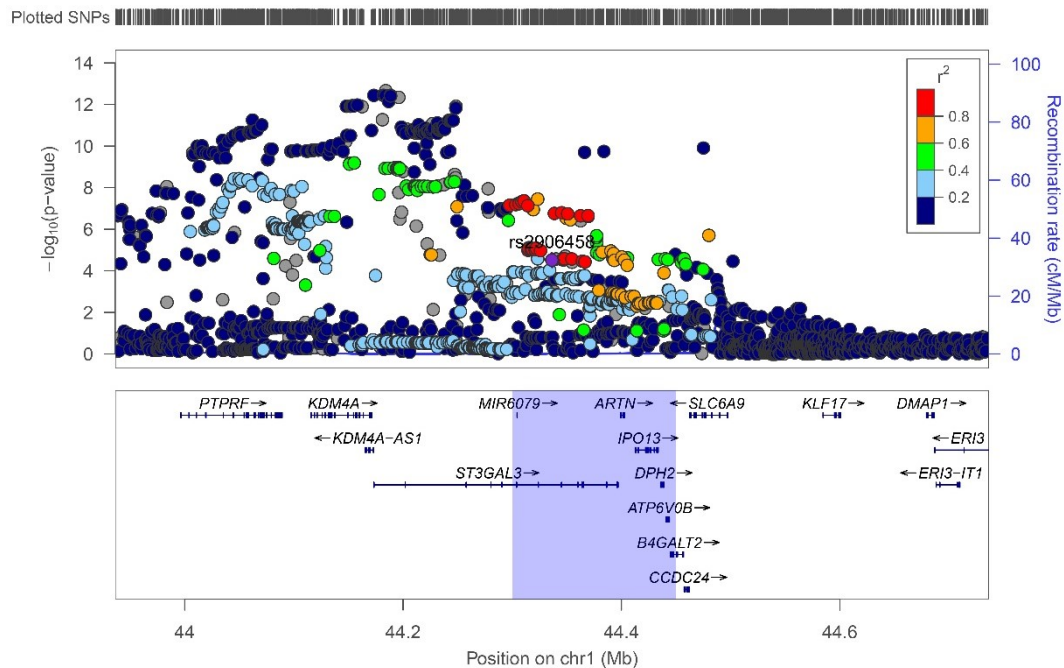


Figure S3. Regional association plot for rs2906458. The SNP represented in the regional plot is depicted in purple. Highlighted in blue: Region represented in Figure 2.

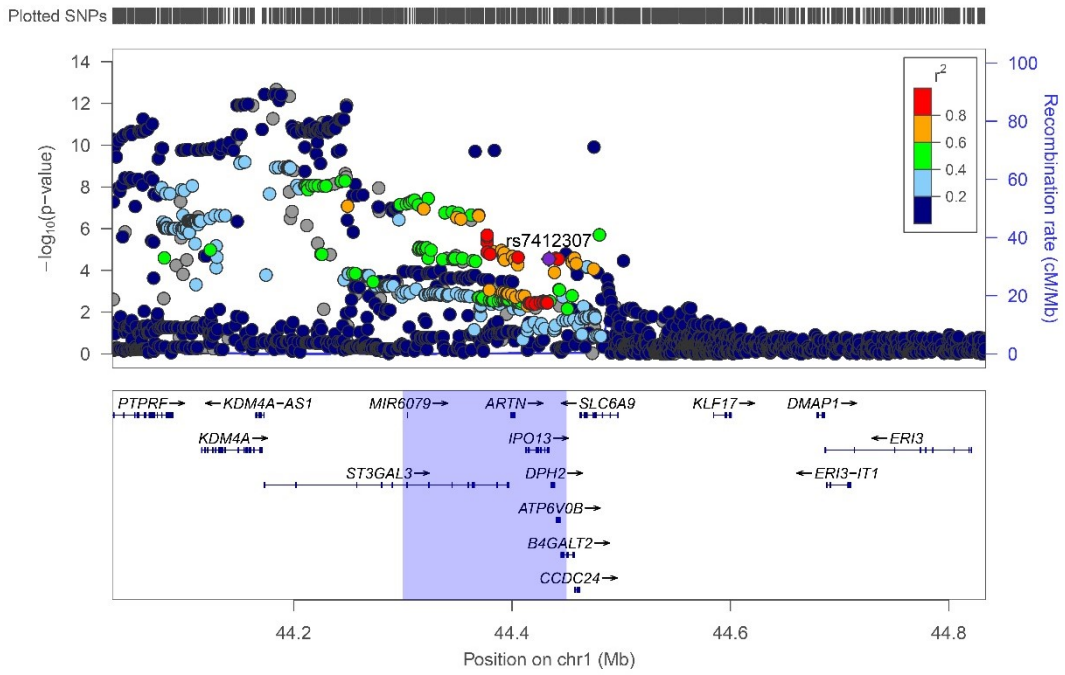


Figure S4. Regional association plot for rs7412307. The SNP represented in the regional plot is depicted in purple. Highlighted in blue: Region represented in Figure 2.

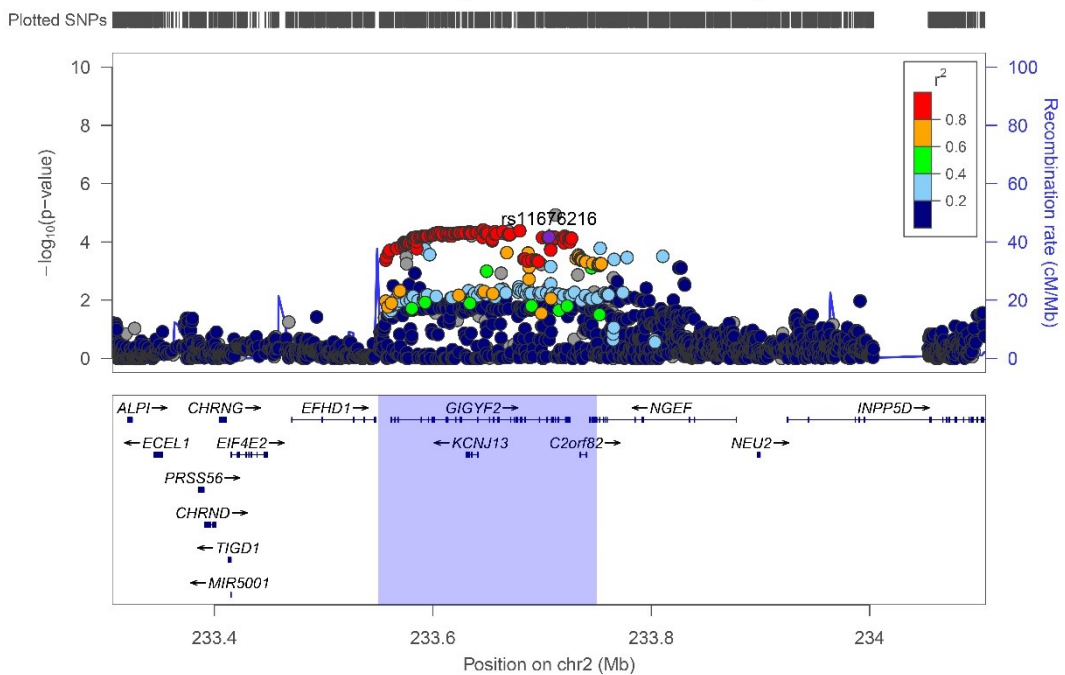


Figure S5. Regional association plot for rs11676216. The SNP represented in the regional plot is depicted in purple. Highlighted in blue: Region represented in Figure 3.

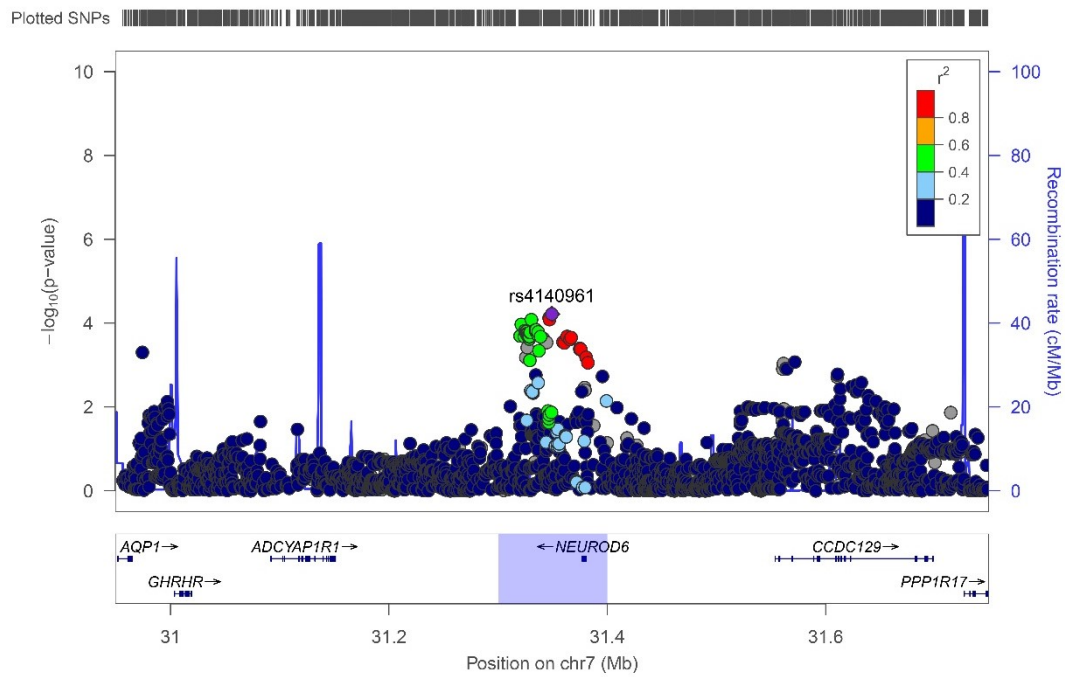


Figure S6. Regional association plot for rs4140961. The SNP represented in the regional plot is depicted in purple. Highlighted in blue: region containing the CpG site and the ASM-SNPs of interest.

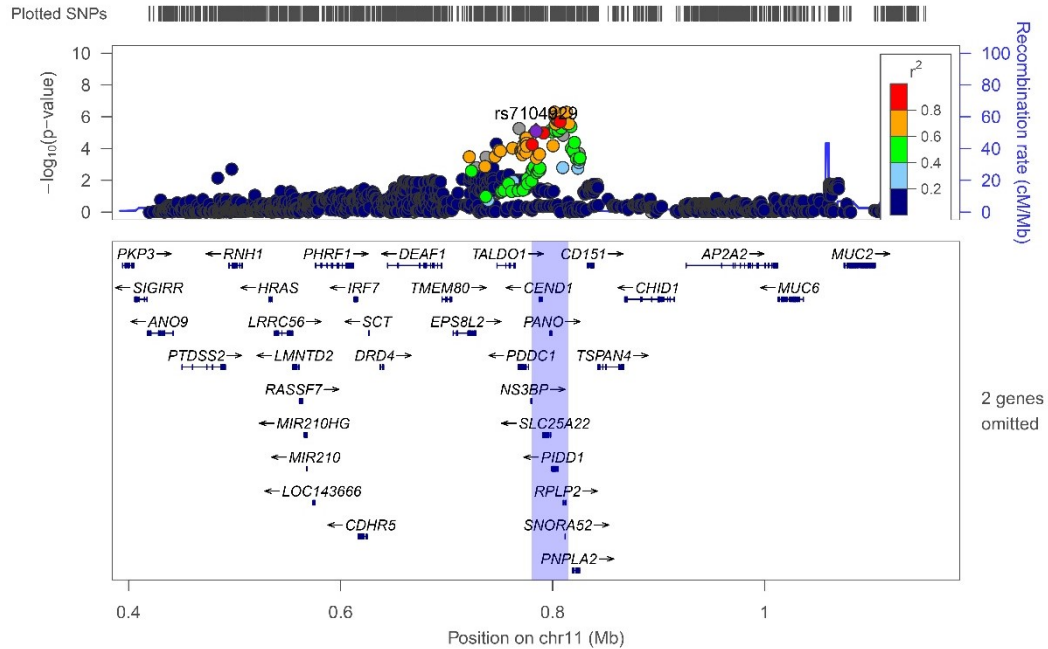


Figure S7. Regional association plot for rs7104929. The SNP represented in the regional plot is depicted in purple. Highlighted in blue: Region represented in Figure 4.

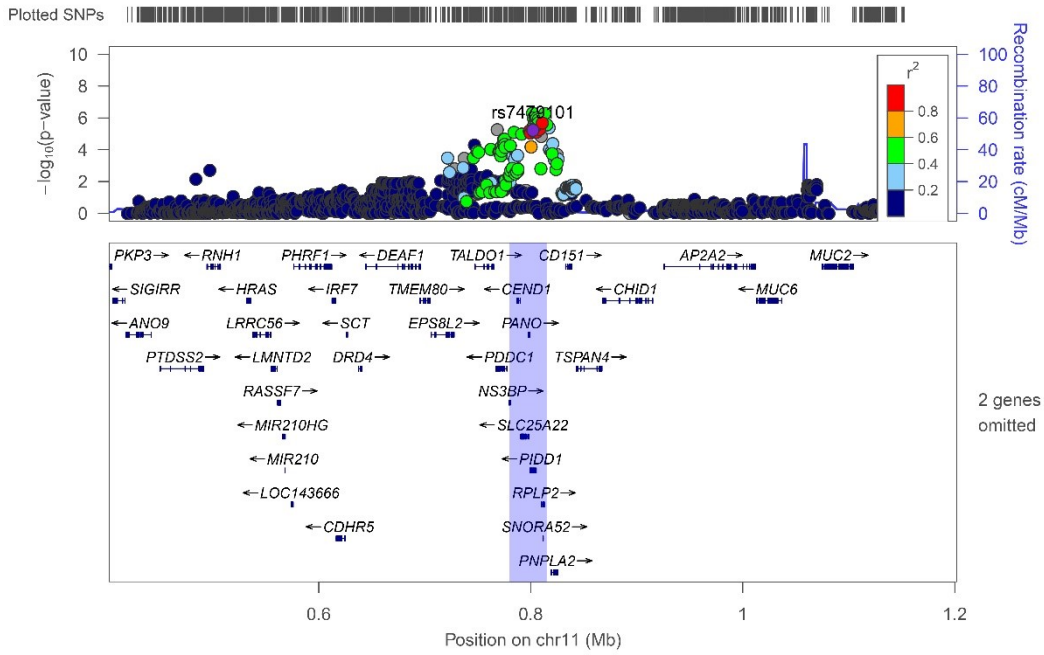


Figure S8. Regional association plot for rs7479101. The SNP represented in the regional plot is depicted in purple. Highlighted in blue: Region represented in Figure 4.

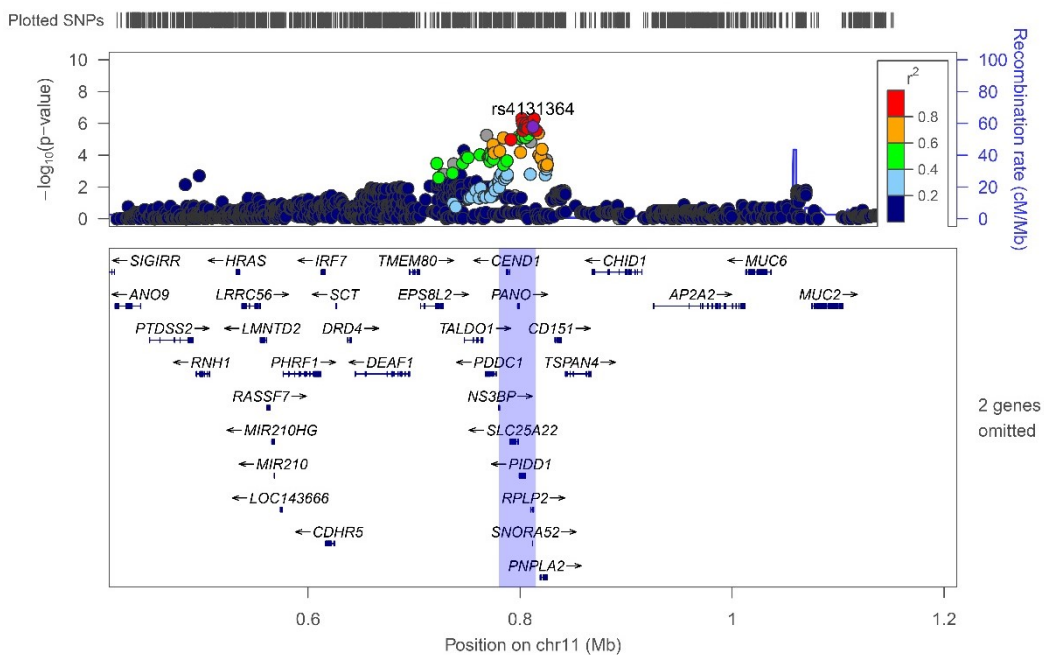


Figure S9. Regional association plot for rs4131364. The SNP represented in the regional plot is depicted in purple. Highlighted in blue: Region represented in Figure 4.

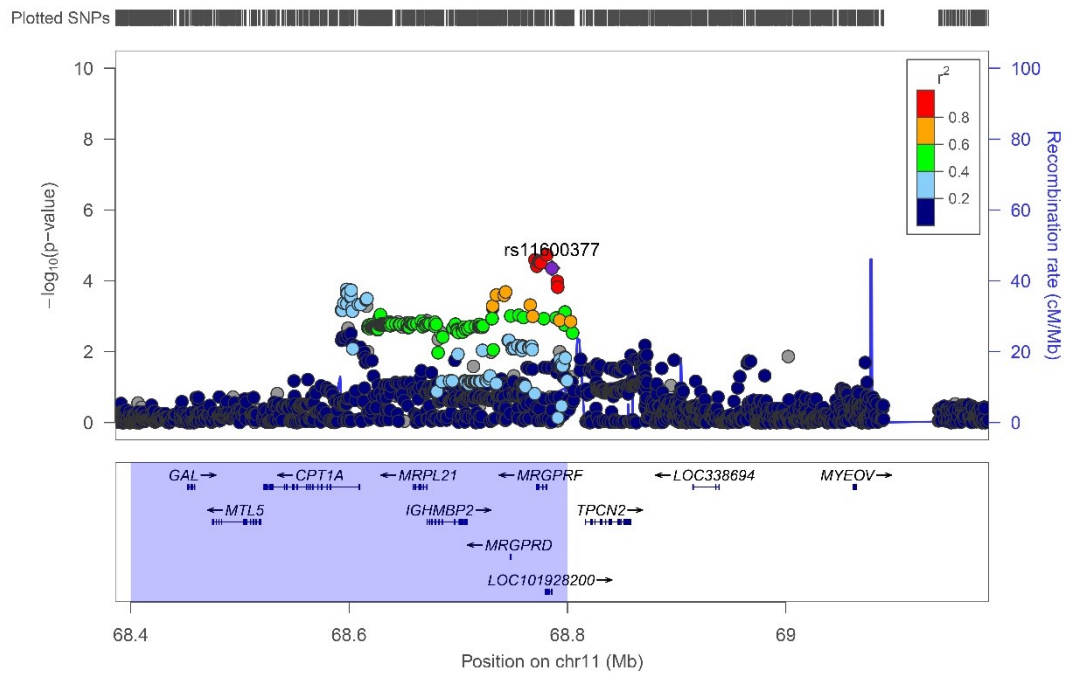


Figure S10. Regional association plot for rs11600377. The SNP represented in the regional plot is depicted in purple. Highlighted in blue: region containing the CpG site and the ASM-SNPs of interest.

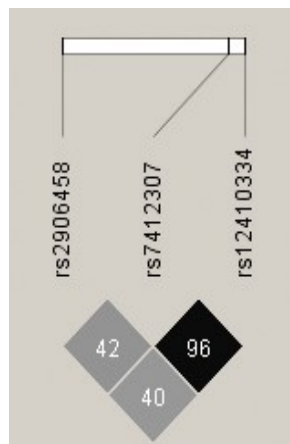


Figure S11. Linkage disequilibrium (LD, r^2) patterns between the three SNPs associated with ADHD that also correlate with differential methylation at two CpG sites, cg22930187 and cg06207804, located in the possible promoter region of *ARTN*.

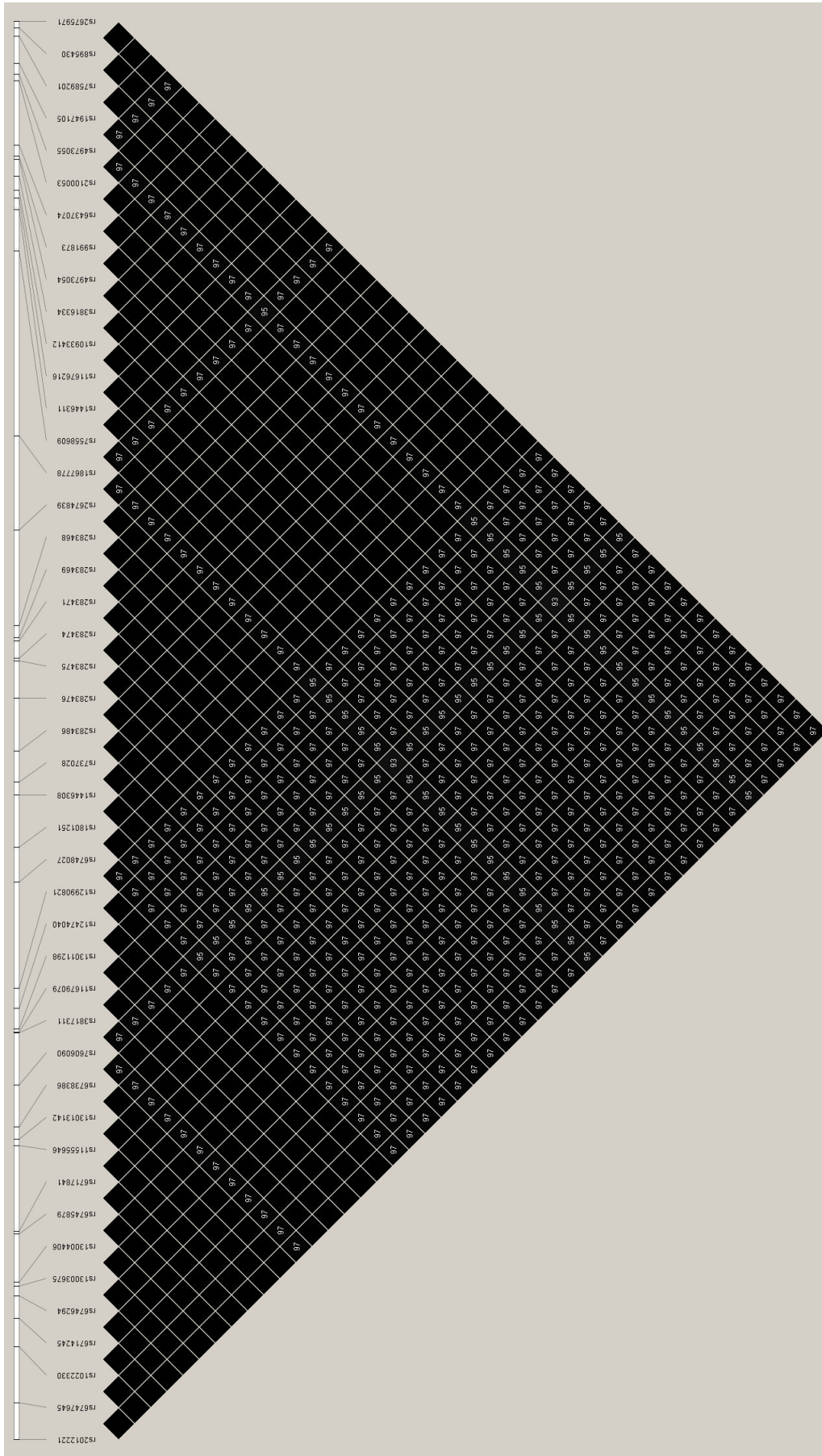


Figure S12. Linkage disequilibrium (LD, r^2) patterns between the 45 SNPs associated with ADHD that also correlate with differential methylation at the CpG site cg13047596, located in the possible promoter region of *C2orf82*.

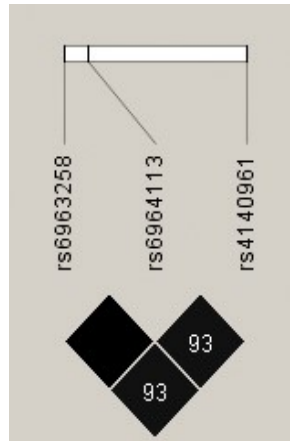


Figure S13. Linkage disequilibrium (LD, r^2) patterns between the three SNPs associated with ADHD that also correlate with differential methylation at the CpG site cg1154507, located in the possible promoter region of *NEUROD6*.

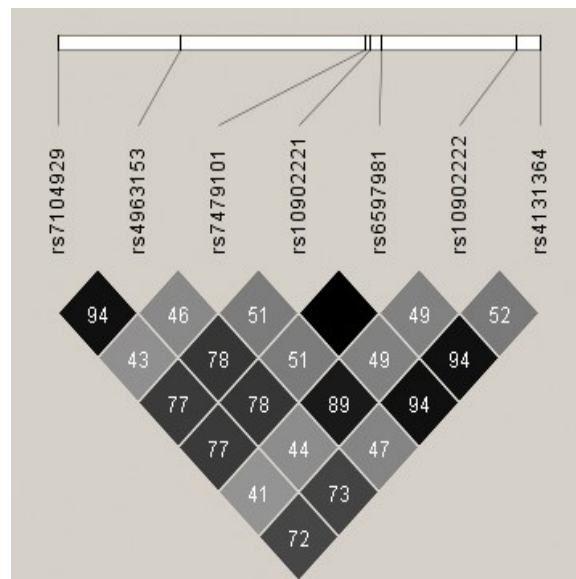


Figure S14. Linkage disequilibrium (LD, r^2) patterns between the 7 SNPs associated with ADHD that also correlate with differential methylation at the CpG site cg20225915, located in the possible promoter region of *PIDD1*.

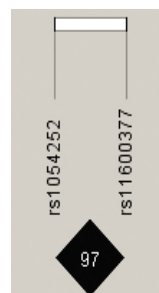


Figure S15. Linkage disequilibrium (LD, r^2) patterns between the two SNPs associated with ADHD that also correlate with differential methylation at the CpG site cg04464446, located in the possible promoter region of *GAL*.

Table S1. Enrichment analysis of ASM SNPs at different significance thresholds in the ADHD GWAS meta-analysis by Demontis et al. (2019).

Significance				
Threshold	N SNPs	N ASM SNPs	p-value	OR
5.00E-08	303	6	<u>1.70E-03</u>	4.92
5.00E-07	945	8	<u>4.30E-02</u>	2.08
5.00E-06	2,122	15	<u>3.15E-02</u>	1.74
5.00E-05	6,970	35	1.31E-01	1.23
5.00E-04	25,288	139	<u>4.58E-04</u>	1.35
5.00E-03	115,681	527	<u>6.94E-03</u>	1.12
5.00E-02	651,772	2790	<u>5.54E-03</u>	1.05

ASM: Allele-specific methylation; N SNPs: Significant SNPs in the ADHD GWAS meta-analysis for the corresponding significance threshold; N ASM SNPs: Significant ASM SNPs in the GWAS meta-analysis; Underlined: Significant enrichment of ASM SNPs in the list of ADHD-associated SNPs; OR: Odds ratio.

Table S2. ASM tagSNPs associated with ADHD.

SNP	SNP information			¹ Association with ADHD				² Correlation with methylation				
	Chr	Pos	Alleles		Freq A1		OR	p-value	CpG site	Tissue	Effect	p-value
			A1	A2	Cases	Controls						
rs2906458	1	44336389	A	<u>G</u>	0.74	0.756	0.94	3.01E-05	cg22930187	Crbl	↓	5.55E-10
									cg06207804	Crbl	↓	1.10E-12
rs7412307	1	44433864	<u>C</u>	G	0.185	0.172	1.07	2.82E-05	cg22930187	Crbl	↓	3.59E-11
									cg06207804	Crbl	↓	1.76E-15
rs11676216	2	233706368	T	<u>C</u>	0.646	0.654	0.95	6.78E-05	cg13047596	Tctx	↑	1.09E-10
										Fctx	↑	2.35E-10
rs4140961	7	31349352	<u>A</u>	G	0.597	0.592	1.06	6.05E-05	cg11554507	Pons	↓	8.45E-25
										Tctx	↓	4.72E-23
rs7104929	11	784340	C	<u>G</u>	0.512	0.526	0.94	7.89E-06	cg20225915	Pons	↓	8.17E-10
										Tctx	↓	1.51E-08
rs7479101	11	802115	A	<u>G</u>	0.317	0.33	0.93	5.90E-06	cg20225915	Pons	↓	2.15E-14
										Tctx	↓	3.37E-14
rs4131364	11	812188	<u>A</u>	G	0.517	0.502	1.07	1.60E-06	cg20225915	Pons	↓	2.49E-10
rs11600377	11	68785803	<u>A</u>	G	0.731	0.72	1.06	4.38E-05	cg04464446	Crbl	↑	3.14E-08

ASM: Allele-specific methylation; ¹Data obtained from the PGC+iPSYCH ADHD GWAS meta-analysis (Demontis et al., 2019); ²Described in Zhang et al., 2010 and Gibbs et al., 2010; SNP: Single Nucleotide Polymorphism; Chr: Chromosome; Pos: Position (build hg19); A1: Allele 1; A2: Allele 2; All alleles are reported in the forward strand; Freq A1: Frequency of allele 1; OR: Odds Ratio (calculated on A1); Effect: Direction of the risk allele effect on DNA methylation levels; Underlined allele: Risk allele for ADHD; In bold: Significant associations for the association between ASM tagSNPs and ADHD p-values overcoming Bonferroni correction for multiple testing and p-value threshold determined using independent number of tests (GEC); Crbl: Cerebellum; Tctx: Temporal cortex; Fctx: Frontal cortex.

Resultats capítol 3. Article 6

Table S3. ASM SNPs associated with ADHD.

SNP	TagSNP	LD (R ²)	Chr	Pos	Alleles		Association with ADHD				² ADHD risk allele effect on CpG methylation (ASM studies)	³ ADHD risk allele effect on gene expression (GTEx data)
					A1	A2	Freq A1 cases	Freq A1 controls	OR	p-value		
rs2906458*	rs2906458	1	1	44336389	A	<u>G</u>	0.74	0.756	0.94	3.01E-05	↓ cg22930187, cg06207804	↑ ARTN
rs7412307*	rs7412307	1	1	44433864	<u>C</u>	G	0.185	0.172	1.07	2.82E-05		
rs12410334		0.96	1	44442521	<u>A</u>	C	0.184	0.171	1.07	2.87E-05		
rs2012221		0.977	2	233566848	T	<u>C</u>	0.642	0.651	0.95	1.67E-04		
rs6747645		0.977	2	233571033	A	<u>G</u>	0.637	0.647	0.95	1.39E-04		
rs1022330		0.977	2	233577330	<u>A</u>	G	0.357	0.348	1.05	1.15E-04		
rs6714245		0.977	2	233580505	T	<u>C</u>	0.643	0.652	0.95	1.12E-04		
rs6746294		0.977	2	233583050	C	<u>G</u>	0.645	0.653	0.95	9.06E-05		
rs13003675		0.977	2	233584109	<u>I</u>	C	0.355	0.346	1.06	6.88E-05		
rs13004406		0.977	2	233584557	A	<u>I</u>	0.642	0.651	0.95	1.49E-04		
rs6745879		0.977	2	233590007	A	<u>G</u>	0.645	0.654	0.95	6.66E-05		
rs6717841		0.977	2	233590255	<u>I</u>	C	0.355	0.346	1.06	6.42E-05		
rs11555646		0.977	2	233599904	A	<u>C</u>	0.645	0.654	0.95	5.80E-05		
rs13013142		0.977	2	233600606	A	<u>G</u>	0.645	0.654	0.95	5.73E-05		
rs6738386		0.977	2	233602028	T	<u>C</u>	0.645	0.654	0.95	5.06E-05		
rs7606090		0.954	2	233606740	T	<u>C</u>	0.645	0.654	0.95	4.90E-05		
rs3817311		0.977	2	233612557	<u>I</u>	C	0.355	0.346	1.06	5.13E-05		
rs11679079		0.977	2	233612656	<u>I</u>	C	0.355	0.346	1.06	5.29E-05		
rs13011298		0.977	2	233612996	A	<u>G</u>	0.645	0.654	0.95	5.30E-05		
rs12474040		0.977	2	233615345	T	<u>G</u>	0.645	0.654	0.95	5.28E-05		
rs12990821		0.977	2	233617585	C	<u>G</u>	0.645	0.654	0.95	5.09E-05		
rs6748027		1.0	2	233629552	T	<u>C</u>	0.645	0.654	0.95	4.86E-05		
rs1801251		1.0	2	233633460	<u>A</u>	G	0.355	0.346	1.06	4.42E-05		
rs1446308		1.0	2	233639309	T	<u>C</u>	0.645	0.654	0.95	4.67E-05		
rs737028		1.0	2	233640750	T	<u>C</u>	0.644	0.653	0.95	4.74E-05		
rs283486	rs11676216	1.0	2	233644223	A	<u>G</u>	0.645	0.653	0.95	4.73E-05	↑ cg13047596	↓ C2orf82
rs283476		1.0	2	233650168	A	<u>G</u>	0.657	0.666	0.95	5.09E-05		
rs283475		1.0	2	233654381	<u>I</u>	G	0.342	0.334	1.06	9.06E-05		
rs283474		1.0	2	233654627	A	<u>G</u>	0.656	0.665	0.95	7.96E-05		
rs283471		1.0	2	233656627	<u>A</u>	G	0.342	0.333	1.06	5.30E-05		
rs283469		1.0	2	233656997	<u>I</u>	G	0.353	0.345	1.06	4.45E-05		
rs283468		1.0	2	233658309	<u>I</u>	C	0.342	0.333	1.06	5.26E-05		
rs2674839		1.0	2	233669040	<u>C</u>	G	0.354	0.346	1.06	5.56E-05		
rs1867778		0.977	2	233679644	T	<u>C</u>	0.646	0.655	0.94	4.11E-05		
rs7558609		1.0	2	233700379	<u>A</u>	G	0.354	0.346	1.06	7.06E-05		
rs1446311		1.0	2	233705071	A	<u>G</u>	0.646	0.654	0.95	6.97E-05		
rs11676216*		1.0	2	233706368	T	<u>C</u>	0.646	0.654	0.95	6.78E-05		
rs10933412		1.0	2	233707226	C	<u>G</u>	0.646	0.654	0.95	6.76E-05		
rs3816334		1.0	2	233708806	<u>A</u>	G	0.354	0.346	1.06	6.66E-05		
rs4973054		1.0	2	233710713	C	<u>G</u>	0.646	0.654	0.95	6.82E-05		
rs991873		1.0	2	233711046	<u>A</u>	G	0.354	0.346	1.06	6.89E-05		
rs6437074		1.0	2	233712296	A	<u>G</u>	0.646	0.654	0.95	7.05E-05		
rs2100053		1.0	2	233719516	T	<u>G</u>	0.645	0.653	0.95	8.18E-05		
rs4973055		0.977	2	233720283	T	<u>G</u>	0.639	0.647	0.95	1.07E-04		
rs1947105		1.0	2	233721455	A	<u>G</u>	0.645	0.654	0.95	7.31E-05		
rs7589201		1.0	2	233724536	<u>A</u>	G	0.355	0.346	1.06	7.49E-05		
rs895430		1.0	2	233725483	A	<u>C</u>	0.646	0.655	0.95	9.84E-05		
rs2675971		1.0	2	233726154	<u>A</u>	G	0.355	0.346	1.06	7.31E-05		
rs6963258		0.937	7	31346832	A	<u>I</u>	0.401	0.405	0.95	7.43E-05		
rs6964113	rs4140961	0.937	7	31347163	<u>C</u>	G	0.598	0.593	1.05	8.16E-05	↓ cg11554507	-
rs4140961*		1	7	31349352	<u>A</u>	G	0.597	0.592	1.06	6.05E-05		
rs7104929*	rs7104929	1.0	11	784340	C	<u>G</u>	0.512	0.526	0.94	7.89E-06		
rs4963153		0.941	11	791462	<u>A</u>	G	0.492	0.478	1.06	1.04E-05		
rs7479101*	rs7479101	1.0	11	802115	A	<u>G</u>	0.317	0.33	0.93	5.90E-06	↓ cg20225915	↑ PIDD1 ↓ PNPLA2
rs10902222		0.896	11	810882	T	G	0.697	0.683	1.07	2.03E-06		
rs10902221		0.941	11	802379	T	<u>C</u>	0.478	0.492	0.93	9.70E-07		
rs6597981	rs4131364	0.941	11	803017	A	<u>G</u>	0.471	0.485	0.94	2.77E-06		
rs4131364*		1.0	11	812188	<u>A</u>	G	0.517	0.502	1.07	1.60E-06		
rs1054252	rs11600377	0.973	11	68772072	A	<u>G</u>	0.27	0.281	0.94	3.86E-05	↑ cg04464446	↑ MRPL21, MRGPRD ↓ IGHMBP2
rs11600377*		1.0	11	68785803	A	G	0.731	0.72	1.06	4.38E-05		

ASM: Allele-specific methylation; ¹ Data obtained from the PGC+iPSYCH ADHD GWAS meta-analysis (Demontis et al., 2019); ² Described in Zhang et al., 2010 and Gibbs et al., 2010; ³ eQTL information for brain tissues; SNP: Single Nucleotide Polymorphism; LD: linkage disequilibrium; Chr: Chromosome; Pos: Position (build hg19); A1: Allele 1; A2: Allele 2; All alleles are reported in the forward strand; Freq A1: Frequency of allele 1; OR: Odds Ratio (calculated on A1); Underlined allele: Risk allele for ADHD; ↑: Hypermethylation/Overexpression; ↓: Hypomethylation/Downexpression; "-": No significant data for the SNP; *: Significant tagSNPs overcoming 5% FDR, the other SNPs are ASM SNPs in LD with these significant tagSNPs; In bold: Significant associations for the association between ASM tagSNPs and ADHD p-values overcoming Bonferroni correction for multiple testing and p-value

Table S4. Histone marks for the ASM SNPs correlating with differential methylation of cg22930187 and cg06207804.

SNP		Hippocampus Middle	Substantia Nigra	Anterior Caudate	Cingulate Gyrus	Inferior Temporal Lobe	Angular Gyrus	Dorsolateral Prefrontal Cortex	Germinal Matrix	Fetal Brain Female	Fetal Brain Male
rs2906458	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K4me1			H3K4me1		H3K4me1	H3K4me1		H3K4me1	
	Promoter histone marks	H3K27ac						H3K27ac			
rs7412307	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K27ac			H3K27ac		H3K27ac				
	Promoter histone marks	H3K4me3					H3K27ac				
							H3K4me3				
rs12410334	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K4me1			H3K4me1		H3K4me1	H3K4me1		H3K4me1	H3K4me1
	Promoter histone marks	H3K27ac			H3K27ac		H3K27ac	H3K27ac			

ASM: Allele-specific methylation; H3K4me1 and H3K27ac: Histone marks related to enhancer regions; H3K4me3 and H3K9ac: Histone marks related to promoter regions; EnhG: Genic enhancer; TxFlnk: Transcription at gene 5' and 3'; TxReg: Transcribed and regulatory (Promoter/Enhancer); TxEnh5: Transcribed 5' preferential and enhancers; Gray cells: Experiments from Roadmap are not available.

Table S5. Histone marks for the ASM SNPs correlating with differential methylation of cg13047596.

SNP		Hippocampus Middle	Substantia Nigra	Anterior Caudate	Cingulate Gyrus	Inferior Temporal Lobe	Angular Gyrus	Dorsolateral Prefrontal Cortex	Germinal Matrix	Fetal Brain Female	Fetal Brain Male
rs2012221	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K4me1 H3K27ac	H3K4me1	H3K4me1	H3K4me1	H3K27ac	H3K4me1	H3K4me1			
	Promoter histone marks	H3K4me3 H3K9ac	H3K9ac	H3K9ac		H3K9ac	H3K4me3 H3K9ac	H3K4me3			
rs6747645	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K4me1 H3K27ac									
	Promoter histone marks	H3K4me3 H3K9ac					H3K9ac				
rs1022330	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K4me1 H3K27ac									
	Promoter histone marks	H3K4me3 H3K9ac									
rs6714245	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K4me1 H3K27ac	H3K4me1	H3K4me1		H3K4me1	H3K4me1	H3K4me1			
	Promoter histone marks	H3K4me3 H3K9ac									
rs6746294	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K4me1 H3K27ac	H3K4me1	H3K4me1							
	Promoter histone marks	H3K4me3 H3K9ac									
rs13003675	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K4me1 H3K27ac	H3K4me1	H3K4me1			H3K4me1	H3K4me1			
	Promoter histone marks	H3K4me3 H3K9ac									
	Chromatin state	15-state model 25-state model									
	Enhancer histone marks	H3K4me1	H3K4me1	H3K4me1			H3K4me1	H3K4me1			

rs13004406	marks	H3K27ac							
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							
	Enhancer histone marks	H3K4me1 H3K27ac		H3K4me1					
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							
rs6745879	Enhancer histone marks	H3K4me1 H3K27ac							
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							
rs6717841	Enhancer histone marks	H3K4me1 H3K27ac							
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							
rs11555646	Enhancer histone marks	H3K4me1 H3K27ac							
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							
	Enhancer histone marks	H3K4me1 H3K27ac							
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							
rs13013142	Enhancer histone marks	H3K4me1 H3K27ac							
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							
	Enhancer histone marks	H3K4me1 H3K27ac							
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							
rs6738386	Enhancer histone marks	H3K4me1 H3K27ac							
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							
	Enhancer histone marks	H3K4me1 H3K27ac							
	Promoter histone marks	H3K4me3 H3K9ac							
	Chromatin state	15-state model 25-state model							

	marks	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac		
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac		
rs4973054	Chromatin state	15-state model	25-state model										
	Enhancer histone marks	H3K4me1	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac		
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac		
rs991873	Chromatin state	15-state model	25-state model										
	Enhancer histone marks	H3K4me1	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac		
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac		
rs6437074	Chromatin state	15-state model	25-state model										
	Enhancer histone marks	H3K4me1	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K4me1	
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	
rs2100053	Chromatin state	15-state model	25-state model										
	Enhancer histone marks	H3K4me1	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	
rs4973055	Chromatin state	15-state model	25-state model										
	Enhancer histone marks	H3K4me1	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	
rs1947105	Chromatin state	15-state model	25-state model										
	Enhancer histone marks	H3K4me1	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	

rs7589201	marks	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac
	Chromatin state	15-state model	25-state model							
rs895430	Enhancer histone marks	H3K4me1	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac
	Chromatin state	15-state model	25-state model							
rs2675971	Enhancer histone marks	H3K4me1	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac
	Promoter histone marks	H3K4me3	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac
	Chromatin state	15-state model	25-state model							

ASM: Allele-specific methylation; H3K4me1 and H3K27ac: Histone marks related to enhancer regions; H3K4me3 and H3K9ac: Histone marks related to promoter regions; Enh: Enhancers; EnhW2: Weak Enhancer 2; EnhG: Genic enhancer; EnhAF: Active enhancer Flank; TxEnh3: Transcribed 3' preferential and enhancers; Gray cells: Experiments from Roadmap are not available.

Table S6. Histone marks for the ASM SNPs correlating with differential methylation of cg11554507.

SNP		Hippocampus Middle	Substantia Nigra	Anterior Caudate	Cingulate Gyrus	Inferior Temporal Lobe	Angular Gyrus	Dorsolateral Prefrontal Cortex	Germinal Matrix	Fetal Brain Female	Fetal Brain Male
rs6963258	Chromatin state	15-state model									
	25-state model										
	Enhancer histone marks	H3K4me1								H3K4me1	H3K4me1
rs6964113	Promoter histone marks	H3K27ac									
		H3K4me3									
		H3K9ac									
rs4140961	Chromatin state	15-state model									
	25-state model										
	Enhancer histone marks	H3K4me1								H3K4me1	H3K4me1
rs4140961	Promoter histone marks	H3K27ac									
		H3K4me3									
		H3K9ac									
rs4140961	Chromatin state	15-state model							Enh	Enh	Enh
	25-state model									EnhAc	EnhAc
	Enhancer histone marks	H3K4me1							H3K4me1	H3K4me1	H3K4me1
rs4140961	Promoter histone marks	H3K27ac									
		H3K4me3									
		H3K9ac									

ASM: Allele-specific methylation; H3K4me1 and H3K27ac: Histone marks related to enhancer regions; H3K4me3 and H3K9ac: Histone marks related to promoter regions; Enh: Enhancers; EnhAc: Primary H3K27ac possible enhancer; Gray cells: Experiments from Roadmap are not available.

Table S7. Histone marks for the ASM SNPs correlating with differential methylation of cg20225915.

SNP		Hippocampus Middle	Substantia Nigra	Anterior Caudate	Cingulate Gyrus	Inferior Temporal Lobe	Angular Gyrus	Dorsolateral Prefrontal Cortex	Germinal Matrix	Fetal Brain Female	Fetal Brain Male
rs7104929	Chromatin state	15-state model								Enh	Enh
		25-state model								EnhW2	EnhAc
	Enhancer histone marks	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1		H3K4me1	H3K4me1
	Promoter histone marks	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac			
rs4963153	Chromatin state			H3K9ac		H3K9ac					
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs7479101	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			
rs10902221	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs6597981	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			H3K4me1
rs10902222	Chromatin state										
		15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3			

Enhancer histone marks	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1
	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac
Promoter histone marks	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3
	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac
Chromatin state	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG	EnhG
	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5
Enhancer histone marks	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1
	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac
Promoter histone marks	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3	H3K4me3
	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac	H3K9ac

rs411364

ASM: Allele-specific methylation; H3K4me1 and H3K27ac: Histone marks related to enhancer regions; H3K4me3 and H3K9ac: Histone marks related to promoter regions; Enh: Enhancers; EnhW2: Weak Enhancer 2; EnhAc: Primary H3K27ac possible enhancer; EnhG: Genic enhancer; TxEnh5: Transcribed 5' preferential and enhancers; Gray cells: Experiments from Roadmap are not available.

Table S8. Histone marks for the ASM SNPs correlating with differential methylation of cg04464446.

SNP		Hippocampus Middle	Substantia Nigra	Anterior Caudate	Cingulate Gyrus	Inferior Temporal Lobe	Angular Gyrus	Dorsolateral Prefrontal Cortex	Germinal Matrix	Fetal Brain Female	Fetal Brain Male
rs1054252	Chromatin state	15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1						H3K4me1			
	Promoter histone marks	H3K27ac	H3K27ac		H3K27ac			H3K27ac			
rs11600377	Chromatin state	15-state model									
		25-state model									
	Enhancer histone marks	H3K4me1									
	Promoter histone marks	H3K27ac									

ASM: Allele-specific methylation; H3K4me1 and H3K27ac: Histone marks related to enhancer regions; H3K4me3 and H3K9ac: Histone marks related to promoter regions; Gray cells: Experiments from Roadmap are not available.

Table S9. MetaXcan prediction of gene expression effects on ADHD for multiple brain tissues.

Gene	Brain tissue*	Z-score	p-value	N SNPs in model	N SNPs used	N ASM SNPs	Predicted R ²
<i>ARTN</i>	Caudate basal ganglia	1.54	1.20E-01	19	16	0	0.10
	Cerebellar hemisphere	4.19	2.50E-05	15	15	4	0.37
	Cerebellum	3.57	3.50E-04	31	31	5	0.36
	Cortex	3.94	8.10E-05	5	4	2	0.14
	Frontal cortex	1.42	1.50E-01	29	24	2	0.09
	Hippocampus	1.55	1.10E-01	140	129	2	0.16
<i>C2orf82</i>	Dorsolateral prefrontal cortex	-3.50	4.50E-04	58	41	9	0.43
	Amygdala	-3.07	2.00E-03	92	89	42	0.40
	Anterior cingulate cortex	-3.55	3.00E-04	11	11	7	0.19
	Caudate basal ganglia	-3.54	3.00E-04	11	11	8	0.44
	Cerebellar hemisphere	-3.64	2.00E-04	37	35	6	0.19
	Cerebellum	-3.50	4.00E-04	30	30	7	0.28
	Cortex	-3.46	5.00E-04	59	54	7	0.27
	Frontal cortex	-3.59	3.00E-04	29	28	6	0.45
	Hippocampus	-3.40	6.00E-04	25	23	8	0.24
	Hypothalamus	-3.37	7.00E-04	30	28	13	0.17
	Nucleus accumbens basal ganglia	-3.59	3.00E-04	29	29	11	0.33
Putamen basal ganglia	-3.20	1.00E-03	52	43	9	0.52	
<i>PIDD1</i>	Dorsolateral prefrontal cortex	4.71	2.41E-06	4	2	1	0.06
	Cerebellar hemisphere	5.00	4.20E-07	32	27	4	0.53
	Cerebellum	5.37	7.60E-08	36	27	5	0.49
	Cortex	3.57	3.40E-04	64	47	2	0.03

ASM: Allele-specific methylation; *ADHD prediction models were only available for some tissues and genes; Z-score: Number of standard deviations change in gene expression in ADHD; p-value: Significance of the association between predicted expression levels and ADHD; N SNPs in model: Number of SNPs used in the training of prediction models for each gene; N SNPs used: Number of SNPs used from the ADHD GWAS meta-analysis summary statistics; N ASM SNPs: Number of ASM SNPs included in the model; Predicted R²: Correlation between the predicted and observed gene expression during prediction model training; In bold: Significant p-values overcoming Bonferroni correction for multiple testing.

Table S10. Correlations with sub-cortical brain volumes of the ASM SNPs associated with ADHD.

SNP	SM for CpG site	Nucleus		Amygdala		Caudate nucleus		Hippocampus		Pallidum		Putamen		Thalamus			
		A1	A2	Effect	p-value	Effect	p-value	Effect	p-value	Effect	p-value	Effect	p-value	Effect	p-value		
rs2906458*		A	G	0.2	0.885	0.7	0.784	0.8	0.881	-4.1	0.442	3.9	0.073	8.9	0.191	3.5	0.628
rs7412307*	cg22930187	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs12410334	cg06207804	A	C	-0.6	0.651	2.9	0.341	-1.6	0.795	-3.0	0.624	3.9	0.119	3.9	0.619	-4.0	0.623
rs2012221		T	C	2.5	0.041	1.7	0.494	10.8	0.040	-2.5	0.618	0.2	0.923	8.2	0.202	-4.9	0.471
rs6747645		A	G	2.5	0.030	1.8	0.473	11.1	0.035	-2.7	0.589	0.5	0.812	8.5	0.187	-4.6	0.496
rs1022330		A	G	2.5	0.029	1.8	0.479	10.8	0.038	-2.6	0.609	0.2	0.908	8.2	0.200	-4.8	0.471
rs6714245		T	C	2.5	0.038	1.8	0.470	10.7	0.040	-2.8	0.577	0.3	0.880	8.1	0.202	-4.9	0.467
rs6746294		C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs13003675		I	C	2.5	0.036	1.8	0.465	10.5	0.044	-2.2	0.663	0.3	0.897	7.9	0.220	-4.9	0.467
rs13004406		A	I	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs6745879		A	G	2.5	0.031	1.8	0.468	10.6	0.040	-2.7	0.585	0.3	0.881	8.1	0.201	-4.9	0.456
rs6717841		I	C	2.4	0.044	2.2	0.377	10.1	0.052	-2.4	0.628	0.2	0.941	7.1	0.268	-4.8	0.471
rs11555646		A	C	2.5	0.028	1.7	0.478	10.6	0.039	-2.7	0.586	0.2	0.923	7.9	0.207	-5.1	0.435
rs13013142		A	G	2.5	0.028	1.7	0.478	10.6	0.039	-2.7	0.587	0.2	0.922	7.9	0.207	-5.1	0.435
rs6738386		T	C	2.5	0.028	1.8	0.476	10.6	0.039	-2.7	0.587	0.2	0.920	7.9	0.208	-5.1	0.435
rs7606090		T	C	2.5	0.027	1.8	0.470	10.7	0.037	-2.7	0.580	0.2	0.912	8.0	0.202	-5.2	0.428
rs3817311		I	C	2.5	0.029	1.8	0.472	10.6	0.038	-2.7	0.577	0.2	0.915	7.8	0.212	-5.5	0.405
rs11679079		I	C	2.5	0.029	1.8	0.472	10.6	0.038	-2.7	0.576	0.2	0.916	7.8	0.213	-5.5	0.405
rs13011298		A	G	2.5	0.029	1.8	0.472	10.6	0.037	-2.7	0.576	0.2	0.916	7.8	0.213	-5.5	0.404
rs12474040		T	G	2.5	0.029	1.8	0.473	10.6	0.038	-2.8	0.567	0.2	0.918	7.7	0.217	-5.6	0.395
rs12990821		C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs6748027		T	C	2.5	0.025	1.6	0.512	10.5	0.040	-3.0	0.536	0.2	0.938	7.6	0.224	-5.5	0.396
rs1801251		A	G	2.5	0.025	1.6	0.516	10.5	0.040	-3.1	0.531	0.2	0.940	7.6	0.224	-5.6	0.394
rs1446308		T	C	2.5	0.025	1.6	0.518	10.5	0.040	-3.1	0.532	0.2	0.937	7.6	0.222	-5.5	0.398
rs737028		T	C	2.5	0.024	1.6	0.519	10.5	0.040	-3.1	0.530	0.2	0.933	7.7	0.218	-5.4	0.404
rs283486	cg13047596	A	G	2.6	0.023	1.6	0.517	10.5	0.039	-3.1	0.525	0.1	0.944	8.0	0.201	-5.2	0.427
rs283476		A	G	2.1	0.068	2.6	0.301	10.5	0.041	-2.1	0.671	-0.7	0.725	6.5	0.307	-3.2	0.634
rs283475		I	G	2.2	0.065	2.6	0.289	10.4	0.044	-2.1	0.665	-0.7	0.716	6.4	0.311	-3.2	0.626
rs283474		A	G	2.2	0.065	2.6	0.301	10.5	0.042	-2.1	0.664	-0.7	0.714	6.5	0.304	-3.2	0.633
rs283471		A	G	2.2	0.062	2.9	0.252	10.7	0.039	-2.1	0.676	-0.7	0.736	6.3	0.318	-2.9	0.667
rs283469		I	G	2.6	0.025	1.6	0.506	10.3	0.044	-3.0	0.541	0.1	0.974	8.4	0.179	-5.0	0.449

rs283468	I	C	2.2	0.064	2.6	0.299	10.5	0.042	-2.2	0.659	-0.8	0.707	6.5	0.308	-3.2	0.630
rs2674839	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs1867778	T	C	2.5	0.030	1.6	0.516	10.3	0.044	-3.2	0.508	0.1	0.973	7.9	0.203	-5.8	0.373
rs7558609	A	G	2.6	0.019	1.5	0.537	10.6	0.037	-3.4	0.484	0.1	0.974	8.0	0.202	-6.4	0.326
rs1446311	A	G	2.6	0.020	1.5	0.526	10.3	0.043	-3.1	0.522	0.1	0.967	7.9	0.202	-6.1	0.348
rs11676216*	T	C	2.6	0.020	1.5	0.527	10.5	0.039	-2.9	0.555	0.0	0.992	8.1	0.194	-6.1	0.348
rs10933412	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs3816334	A	G	2.6	0.019	1.6	0.508	10.4	0.042	-3.0	0.536	0.1	0.959	8.1	0.194	-6.2	0.339
rs4973054	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs991873	A	G	2.6	0.019	1.6	0.507	10.4	0.042	-3.0	0.533	0.1	0.960	8.1	0.194	-6.3	0.337
rs6437074	A	G	2.6	0.020	1.7	0.483	9.9	0.052	-3.0	0.536	0.1	0.948	7.8	0.210	-6.6	0.316
rs2100053	T	G	2.6	0.020	1.6	0.524	10.1	0.047	-3.2	0.514	0.1	0.952	7.9	0.204	-6.0	0.358
rs4973055	T	G	2.6	0.024	2.1	0.398	9.6	0.063	-3.1	0.526	0.3	0.881	7.8	0.214	-5.0	0.444
rs1947105	A	G	2.6	0.021	1.6	0.499	10.1	0.048	-3.2	0.508	0.1	0.959	7.7	0.213	-6.0	0.361
rs7589201	A	G	2.6	0.020	1.6	0.519	10.2	0.046	-3.2	0.507	0.1	0.956	7.8	0.209	-5.9	0.363
rs895430	A	C	2.6	0.020	1.6	0.517	10.2	0.045	-3.2	0.510	0.1	0.951	7.8	0.208	-5.9	0.367
rs2675971	A	G	2.6	0.019	1.6	0.516	10.2	0.045	-3.1	0.522	0.2	0.937	7.9	0.206	-5.8	0.375
rs6963258	A	T	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs6964113	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs4140961*	A	G	1.6	0.147	1.7	0.467	1.9	0.700	7.0	0.147	2.2	0.273	5.8	0.341	13.5	0.035
rs7104929*	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs4963153	A	G	0.1	0.931	1.7	0.514	-2.4	0.643	-0.6	0.899	0.5	0.795	-3.5	0.588	8.0	0.232
rs7479101*	A	G	0.4	0.766	2.5	0.365	-2.0	0.733	-1.4	0.808	0.6	0.798	-1.7	0.814	-2.1	0.777
rs10902222	I	G	0.5	0.715	1.6	0.557	-2.0	0.727	-2.4	0.668	0.5	0.840	-2.1	0.761	-2.6	0.726
rs10902221	T	C	0.1	0.940	1.4	0.574	-1.2	0.825	-2.1	0.681	0.6	0.773	2.8	0.663	2.9	0.662
rs6597981	A	G	0.2	0.881	1.6	0.530	-0.7	0.890	-1.9	0.714	0.7	0.734	3.4	0.596	3.6	0.593
rs4131364*	A	G	0.1	0.941	1.2	0.622	-0.9	0.862	-3.3	0.511	0.6	0.749	0.0	0.998	0.7	0.915
rs1054252	A	G	-2.5	0.037	-1.2	0.651	-11.2	0.038	-2.8	0.587	-1.5	0.477	-3.6	0.583	-1.7	0.804
rs11600377*	A	G	-2.5	0.038	-0.9	0.741	-11.7	0.030	-2.9	0.579	-1.9	0.375	-3.7	0.573	-2.2	0.751

ASM: Allele-specific methylation; A1: Allele 1; A2: Allele 2; All alleles are reported in the forward strand; Effect: Effect sizes are given in units of mm3 per risk allele; *: Significant tagSNPs overcoming 5% FDR, the other SNPs are ASM SNPs in LD with these significant tagSNPs; Underlined allele: Risk allele for ADHD; In bold: Nominally significant p-values; "?: No values retrieved for these SNPs.

Table S11. Copy number variants overlapping with the highlighted genes for ADHD.

Gene	Variant	Type	Sex	Inheritance	Phenotype(s)
ARTN	1:36831723-61417768	Deletion	46XX	Unknown	Hypoplasia of the corpus callosum, Intellectual disability
	1:41126365-45133712	Deletion	46XY	De novo constitutive	Congenital microcephaly
	1:43827703-45107379	Deletion	46XX	De novo constitutive	Autism, Intellectual disability
	1:43870211-45107379	Deletion	46XX	De novo constitutive	Autism, Delayed speech and language development, Intellectual disability
	2:216527137-242783384	Duplication	46XX	De novo constitutive	Intellectual disability
	2:217860965-242783384	Deletion	46XX	Unknown	Hypodysplasia of the corpus callosum
C2orf82	2:228207994-236247812	Deletion	46XY	Unknown	Intellectual disability
	2:231683468-235003677	Duplication	unknown	De novo constitutive	Aggressive behavior, ADHD, Dysgraphia, Self-injurious behavior
	2:231869040-243014630	Duplication	other	De novo constitutive	Aggressive behavior, Dysarthria, Intellectual disability, Microcephaly, Small for gestational age
	2:232756751-238245191	Duplication	46XY	De novo constitutive	Intellectual disability
	2:232763439-242717043	Duplication	46XY	De novo constitutive	Intellectual disability
	2:232831650-239972530	Deletion	46XY	De novo constitutive	Macrocephaly, Mild global developmental delay
P/IDD1	2:232981279-243007359	Deletion	46XX	De novo constitutive	Aggressive behavior, Behavioral abnormality, Failure to thrive, Global developmental delay, Growth delay, Intrauterine growth retardation, Obsessive-compulsive behavior
	2:233508667-241070353	Duplication	46XX	De novo constitutive	Intellectual disability, Seizures.
P/IDD1	2:233527432-242852625	Duplication	46XY	De novo mosaic	Delayed speech and language development, Global developmental delay, Motor delay
	2:233613670-233742556	Duplication	46XY	Unknown	Cognitive impairment
	11:1-18716078	Duplication	46XX	De novo constitutive	Intellectual disability
	11:1-4061418	Duplication	46XX	Maternally inherited, constitutive in mother	Macrocephaly, Postnatal growth retardation
	11:192372-3653923	Triplication	46XX	De novo constitutive	Intellectual disability
	11:196966-3254236	Duplication	46XY	De novo constitutive	Intellectual disability, Small for gestational age
P/IDD1	11:198510-2261562	Duplication	46XX	De novo constitutive	Global developmental delay
	11:203788-3131418	Duplication	46XX	Unknown	Delayed speech and language development, Global developmental delay
	11:210100-2425271	Duplication	46XY	Imbalance arising from a balanced parental rearrangement	Delayed speech and language development, Muscular hypotonia
	11:210300-3363804	Duplication	46XX	De novo constitutive	Autism
P/IDD1	11:353347-872723	Duplication	46XY	Maternally inherited, constitutive in mother	Delayed speech and language development
	11:662317-1157708	Duplication	46XX	Maternally inherited, constitutive in mother	Global developmental delay

Variant: CNV position (build hg19); Source: DECIPHER (v9.27) database; The reported phenotypes have been restricted to those comorbid to ADHD or neurodevelopmental processes.

Annex de l'article 6

Contribució del gen *C2orf82* al TDAH mitjançant la utilització d'un model genoanul·lat murí.**1. ANTECEDENTS**

L'avaluació de la contribució de les variants amb metilació específica d'al·lel en el TDAH ens va permetre identificar el gen *C2orf82*, juntament amb dos altres gens (*ARTN* i *PIDD1*), com a nous gens candidats a contribuir al trastorn (Article 6, Capítol 3). En aquest estudi previ, els al·lells de risc pel TDAH presenten correlació amb un augment de la metilació d'un lloc CpG situat al promotor del gen *C2orf82* i amb una disminució de l'expressió d'aquest mateix gen. A més, els al·lells de risc identificats tenen una influència en el volum del nucli accumbens i del nucli caudat i es va predir una expressió disminuïda d'aquest gen en pacients amb TDAH (Article 6, Capítol 3). El gen *C2orf82*, també conegut com a *SNORC*, mostra una expressió més alta en cervell que en altres teixits. A més, estudis de GWAS previs mostren associacions significatives a escala genòmica entre variants que cauen al gen, o a una regió proximal, i diferents trastorns psiquiàtrics, entre ells l'esquizofrènia (Ripke *et al.*, 2013, 2014; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019).

La funció de *C2orf82* encara està poc estudiada, sobretot a nivell del sistema nerviós central. Els únics treballs realitzats fins ara s'han centrat en la seva funció en cartílag, on mostra una elevada expressió en ratolí (Heinonen *et al.*, 2011, 2017; Jaiswal *et al.*, 2020). Un estudi previ a gran escala tenia com a objectiu crear una llibreria de ratolins genoanul·lats de gens que codificaven proteïnes transmembrana o secretades i realitzar una caracterització fenotípica bàsica per identificar alteracions en el desenvolupament, en el metabolisme, en el sistema nerviós o en el sistema cardiovascular (Tang *et al.*, 2010). Aquest estudi va incloure un total de 472 ratolins genoanul·lats, un dels quals corresponia al gen *C2orf82*. Els ratolins mascles homozigots per la mutació mostraven una disminució dels nivells de triglicèrids en sang en comparar-los amb mascles salvatges de la mateixa camada, i els ratolins femella homozigots per la mutació mostraven una disminució de l'activitat (Tang *et al.*, 2010).

L'objectiu del nostre estudi és caracteritzar l'expressió i la funció de *C2orf82* al cervell així com la seva contribució al TDAH. Per aconseguir els nostres objectius, hem avaluat l'expressió del gen en diferents àrees cerebrals de ratolí mitjançant la tècnica de qRT-PCR, hem dut a terme un estudi de predicció de l'estructura de la proteïna i hem caracteritzat a nivell fenotípic el model animal genoanul·lat per aquest gen, utilitzant tests per una fenotipació general però també tests

específics de comportament per avaluar trets i trastorns psiquiàtrics. Aquest projecte s'ha dut a terme en col·laboració amb el grup de Neurofarmacologia de la Universitat Pompeu Fabra (UPF) liderat pel Dr. Rafael Maldonado i el grup de *Translational Psychiatry* de la *Universitätsklinikum Frankfurt* liderat pel Dr. David Slattery.

2. METODOLOGIA

2.1. Anàlisi del genotip

El ratolí genoanul·lat per la mutació es va obtenir del repositori de ratolins *Mutant Mouse Resource & Research Centers* (MMRRC), on hi havia emmagatzemada la soca generada per Tang et al., 2010. Els ratolins genoanul·lats es van produir per recombinació homòloga reemplaçant els exons 2 i 3 del gen per un casset amb resistència a neomicina. La mutació es va generar en cèl·lules mare embrionàries derivades de la soca 129/SvEvBrd, i els ratolins quimèrics resultants es van creuar amb ratolins albins de la soca C57BL/6-Tyr^{c-Brd} per generar la F1, de manera que el fons genètic dels ratolins amb aquesta mutació és una barreja de les dues soques abans mencionades.

Per genotipar els ratolins vam extreure DNA genòmic a partir d'un fragment petit de l'orella o de la cua. Les mostres es van digerir en 300 µl de NaOH 50mM a 98°C durant 30 minuts. A continuació es van barrejar les mostres amb l'ajuda d'un vòrtex i es va afegir 30 µl de Tris-HCL 1M, pH 8 per neutralitzar la digestió. Finalment, es van centrifugar les mostres a 13.200 rpm per separar la fracció no digerida i es van utilitzar 2 µl del sobrenedant per realitzar la PCR de la genotipació (Gómez-Grau et al., 2017).

Per a la PCR vam utilitzar tres encebadors: dos *forward* (5' – ACTCTGCATCTTGCTCTCC i 5' – GCAGCGCATCGCCTTCTATC) i un *reverse* (5' – GACACAGTGAGTGGTACTTAGC). Utilitzant aquests encebadors obteníem dos productes de PCR, un de 242 pb, corresponent a l'al·lel salvatge, i un de 271 pb, corresponent a l'al·lel mutat. Així doncs, els ratolins salvatges presenten només el producte de 242 pb, els ratolins heterozigots presenten els dos productes de PCR i els ratolins homozigots per la mutació només presenten el producte de 271 pb. Els fragments de PCR es van separar en un gel d'agarosa al 2,5% tenyit amb RedSafe (iNtRON Biotechnology). En totes les PCRs es van utilitzar controls positius per cadascun dels genotips estudiats i un control negatiu. Es va realitzar una genotipació al naixement i una altra al sacrifici de tots els ratolins utilitzats pels diferents experiments.

2.2. Estudi d'expressió gènica en el cervell de ratolí

Per l'estudi d'expressió en el cervell de ratolí es van seleccionar cinc ratolins salvatges (WT) mascles naïf de la soca que presenta la mutació pel gen *C2orf82* de 20 setmanes d'edat. Aquests animals es van sacrificar per decapitació i se'ls va extreure vuit àrees cerebrals (amígdala (AMY), àrea tegmental ventral (VTA), cerebel (CB), escorça prefrontal medial (mPFC), nucli estriat dorsal (DS), hipocamp (HPC), hipotàlem (HT) i nucli accumbens (NAc)) i la part de la medul·la espinal (SC) més proximal al cervell. Totes aquestes àrees es van emmagatzemar a -80°C fins que es va realitzar l'extracció d'RNA total amb el *RNeasy Lipid Tissue Mini Kit* (Qiagen). L'RNA es va quantificar mitjançant la tecnologia Nanodrop (NanoDrop Technologies) i es emmagatzemar a -80°C . Posteriorment, es va retrotranscriure utilitzant el *High-Capacity cDNA Reverse kit* (LifeTechnologies) i *RNAase inhibitor* (Applied Biosystems). Finalment, es va realitzar una quantificació absoluta de l'expressió del gen *C2orf82* a les diferents àrees cerebrals per qRT-PCR utilitzant el reactiu *LightCycler® 480 SYBR Green I Master mix* (Roche), el *LightCycler® 480II system* (Roche) i realitzant les anàlisis amb el *LightCycler® 480 Software* versió 1.5.1.62 (Roche).

2.3. Predicció de l'estructura 3D de la proteïna

Per l'estudi de predicció de l'estructura secundària i terciària de la proteïna es va obtenir la seqüència d'aminoàcids corresponent al transcrit de referència del gen *C2orf82* de ratolí (ENSMUST00000027476.5), de 121 aminoàcids, i vam utilitzar sis servidors diferents: 1) *Protein DisOrder prediction System* (PrDOS): predicció de regions desordenades (regions sense una estructura secundària) (Ishida i Kinoshita, 2007); 2) *A Protein Secondary Structure Prediction Server* (Jpred4): predicció d'estructures secundàries (hèlix alfa o làmines beta) (Drozdetskiy et al., 2015); 3) *TMHMM server v2.0*: predicció d'hèlix transmembrana (Krogh et al., 2001); 4) *Signal P5.0 server*: predicció de pèptids senyal (Almagro Armenteros et al., 2019); 5) *Net-O-Glyc4.0 server*: predicció de llocs de glicosilació (Steenfot et al., 2013); 6) *Robetta*: Predicció d'un possible model 3D de la proteïna (Raman et al., 2009; Song et al., 2013).

2.4. Caracterització fenotípica

Tots els tests s'han realitzat amb ratolins mascles i/o femelles homozigots per una mutació inactivadora del gen (KO), heterozigots per la mutació (HET) i salvatges (WT) de 8-10 setmanes d'edat. Els animals es van mantenir en condicions d'estabulació estables: temperatura de $21\pm 1^{\circ}\text{C}$, humitat del $55\pm 10\%$, cicles de llum/fosc de 12h i menjar i aigua disponibles *ad libitum*. Tots els protocols experimentals realitzats a Barcelona es van dur a terme d'acord amb les recomanacions de la *European Communities Directive 86/609/EEC* que regulen la investigació amb animals i van

ser aprovats pel comitè d'ètica local (CEEA-IMAS-UPF). Els protocols realitzats a Frankfurt es van dur a terme d'acord amb la "Directive of the European Communities Council" del 24 de Novembre del 1986 (86/609/EEC), les lleis de benestar animal alemanyes (TierSchG and TSchV) i van ser aprovats per les autoritats locals. A les instal·lacions de Barcelona vam realitzar els experiments amb una cohort de 10 ratolins mascles per cadascun dels genotips considerats (WT, HET i KO) i a les instal·lacions de Frankfurt vam fer servir dues cohorts de ratolins, una de mascles (9 WT, 8 HET i 7 KO) i una de femelles (4 WT, 2 HET i 6 KO). La cohort d'animals de Barcelona es va utilitzar per als següents tests: Test d'Irwin, caixes PHECOMP, *Novel Object Recognition*, test d'activitat o actimetria, *Elevated Plus Maze*, *Rotarod*, test d'interacció social, *Open Field* (50 i 500 lux) i autoadministració de cocaïna (tots aquests tests els va realitzar un integrant del grup liderat pel Dr. Rafael Maldonado). Les cohorts de Frankfurt van ser utilitzades pels tests de *Open Field* (250 lux), *Light Dark Box* i *Prepulse Inhibition test*. A més, cinc mascles per cada genotip es van utilitzar per la *Continuous Performance Task* i totes les femelles pel test de *Fear Conditioning*. Aquests dos últims test no es van realitzar en els mateixos animals degut al nivell d'estrès que suposen. L'ordre en el qual estan descrits els tests correspon a l'ordre en el que es va executar cada test, de manera que els tests amb més nivell d'estrès són els últims que es van realitzar.

2.4.1. Test d'Irwin

El Test d'Irwin consisteix a introduir el ratolí en un cilindre de plexiglàs obert per dalt per avaluar signes subjectius de disfunció del comportament, que ens donen informació de l'estat de salut de l'animal: posició del cos, tremolors, convulsions, alteracions en la respiració, comportament, por, excitació i reactivitat en tocar-los, entre altres. Aquest estudi es va realitzar amb 10 ratolins mascles per cada genotip a les instal·lacions de Barcelona.

2.4.2. Caixes PHECOMP

Aquest test permet mesurar l'activitat i l'anhedonia (disminució o desaparició de la capacitat d'experimentar plaer). El test es realitza en unes caixes (*PHECOMP boxes*, Panlab and Harvard Apparatus) equipades amb un sistema que monitoritza i enregistra la ingesta de menjar i aigua de l'animal, així com l'activitat horitzontal i vertical, durant les 24h del dia al llarg de les tres setmanes que dura l'experiment. El sistema inclou 12 caixes (197 cm de llarg, 206 cm d'amplada i 212 cm d'alt) amb un terra reixat i una tapa amb filtre. Dins de cada caixa hi ha dos dispensadors de menjar i dos d'aigua. Primerament, els animals es mantenen durant 14 dies (període d'habitució) en aquestes gàbies per habitar-se al sistema, i els dos dispensadors de menjar i els dos d'aigua contenen el mateix producte. A continuació es realitza el test d'anhedonia (test de preferència

per la sucrosa) al llarg de 7 dies, durant els quals se substitueix un dels dispensadors d'aigua per aigua amb sacarosa al 2%. Al llarg d'aquesta setmana cada dia s'intercanvia la posició de les ampolles i es mesura el consum de cada ampolla després d'un interval de 24h. La preferència per la sacarosa es calcula com la quantitat relativa d'aigua amb sacarosa al 2% versus el total de líquid consumit. Passats 7 dies es retira el dosificador amb sacarosa i se substitueix novament per aigua. Els ratolins prefereixen l'aigua amb sacarosa, però si aquests presenten anhedonia, no hi ha cap preferència. Aquest estudi es va realitzar amb 8 ratolins mascles per cada genotip a les instal·lacions de Barcelona.

2.4.3. Novel Object Recognition

El test de *Novel Object Recognition* (NOR) ens permet estudiar la memòria a curt termini. Aquest test es realitza en un laberint de plexiglàs amb dos passadissos (30 cm de llarg x 4,5 cm d'amplada x 15 cm d'alt, Panlab) situats en un angle de 90°. El primer dia els ratolins s'habituen al laberint buit durant 9 minuts. Al cap de 24h els ratolins tornen a ser introduïts al laberint durant 9 minuts, però aquesta vegada amb dos objectes idèntics situats al final de cadascun dels passadissos, i s'enregistra el temps que l'animal passa amb cada objecte. Al cap de 3h els ratolins tornen a ser introduïts al laberint, però un dels objectes familiars ha estat substituït per un objecte nou, i s'enregistra el temps que passa l'animal amb cada objecte. Es considera que el ratolí explora l'objecte si el musell està orientat cap a l'objecte i la distància entre ells és inferior a 2 cm. Finalment es calcula l'índex de discriminació entre l'objecte familiar i l'objecte nou, calculat com el temps d'exploració de l'objecte familiar dividit pel temps total d'exploració dels dos objectes. Un índex de discriminació més alt indica una major retenció de la memòria dependent de l'hipocamp i aprenentatge (Lueptow, 2017). Aquest estudi es va realitzar amb 10 ratolins mascles per cada genotip a les instal·lacions de Barcelona.

2.4.4. Test d'activitat o actimetria

Aquest test s'utilitza per determinar l'activitat locomotora. L'aparell utilitzat consisteix en unes caixes de plexiglàs (9 cm x 20 cm x 11 cm; Imetronic) que estan il·luminades a 20-25 lux en les quals s'enregistra l'activitat locomotora horitzontal i vertical del ratolí durant dues hores. Aquest estudi es va realitzar amb 10 ratolins mascles per cada genotip a les instal·lacions de Barcelona.

2.4.5. Elevated Plus Maze

L'*Elevated Plus Maze* (EPM) ens permet avaluar el conflicte entre la tendència natural dels ratolins a evitar una superfície il·luminada i elevada i la tendència natural a explorar nous entorns. Quan

els animals presenten nivells d'ansietat elevats es redueix aquesta tendència exploratòria (Komada, Takao i Miyakawa, 2008). El test es realitza en un laberint alçat 50 cm respecte el terra amb quatre braços (29 cm de llarg x 6 cm d'ample) situats en creu des d'un quadrat central (6 cm x 6 cm). Dos braços oposats estan delimitats per parets verticals de 15 cm (braços parcialment tancats) i estan exposats a una intensitat lumínica baixa (10 lux). Els altres dos no tenen proteccions a les vores (braços oberts) i estan exposats a una il·luminació més alta (200 lux). A l'inici del test s'introdueix el ratolí al final d'un dels braços tancats mirant cap a la paret. Durant el test (5 minuts de durada), s'enregistra el temps dedicat, el número d'entrades, la velocitat mitja i la distància recorreguda a cada tipus de braç. Es considera que l'animal entra en un dels braços quan les quatre potes de l'animal estan dins del braç. Aquest estudi es va realitzar amb 10 ratolins mascles per cada genotip a les instal·lacions de Barcelona.

2.4.6. Rotarod

El test de *Rotarod* s'utilitza per mesurar la coordinació motora, l'equilibri i la força de les extremitats (Deacon, 2013). El gen *C2orf82* té una elevada expressió en cartílag (Heinonen *et al.*, 2011; Jaiswal *et al.*, 2020), de manera que els resultats d'aquest test són importants per tal d'assegurar que els ratolins genoanul·lats no presenten problemes motors que podrien emmascarar els resultats obtinguts en altres tests. El *rotarod* es realitza en un cilindre que rota elevat del terra i comença amb una sessió d'entrenament a velocitat constant (4 rpm) durant un màxim de 60 segons. A continuació, es torna a situar l'animal en el cilindre però aquesta vegada amb velocitat creixent (4 rpm inicials amb un increment d'1 rpm cada 3s fins a 20 rpm) fins a un màxim de 90 segons o fins que l'animal caigui de l'estructura. Es repeteix aquest protocol cinc vegades i s'enregistra la velocitat màxima a què arriba i el temps màxim que passa el ratolí en l'estructura sense caure en cadascun dels cinc intents. Aquest estudi es va realitzar amb 10 ratolins mascles per cada genotip i es va dur a terme a les instal·lacions de Barcelona.

2.4.7. Test d'interacció social

El test d'interacció social s'utilitza per avaluar el comportament social dels ratolins així com la preferència per iniciar interaccions socials noves amb ratolins desconeguts (Moy *et al.*, 2004). En condicions normals els ratolins són una espècie molt social que presenta una preferència per les interaccions socials noves (Moy *et al.*, 2004). Aquest test es realitza en una caixa amb dues gàbies

en el seu interior. Inicialment, hi ha un període d'habitució durant el qual l'animal es familiaritza amb l'ambient (Figura 1A). A continuació, s'introdueix un ratolí amb el qual mai ha tingut contacte en una de les gàbies (Figura 1B) per valorar el comportament social segons la puntuació obtinguda en diferents paràmetres d'interacció entre ambdós ratolins (temps que passa amb el ratolí, temps que ensuma el ratolí i vegades que va cap al ratolí). Finalment, aquest ratolí familiar s'introdueix en l'altra gàbia i se situa un ratolí nou desconegut a la gàbia que queda (Figura 1C) per mesurar la preferència per les relacions socials noves i es tornen a avaluar les interaccions que té el ratolí d'estudi amb els dos ratolins situats a les gàbies. Així doncs, en aquest test el comportament social es defineix com la predilecció a passar més temps prop d'una gàbia que conté un altre ratolí en comparació amb el temps passat prop d'una gàbia idèntica però buida (Figura 1B), i la preferència per les interaccions socials noves es defineix com la predilecció a passar més temps prop d'una gàbia amb un ratolí desconegut en comparació amb el temps passat prop del ratolí familiar (Figura 1C). Aquest estudi es va realitzar amb 10 ratolins mascles per cada genotip a les instal·lacions de Barcelona.

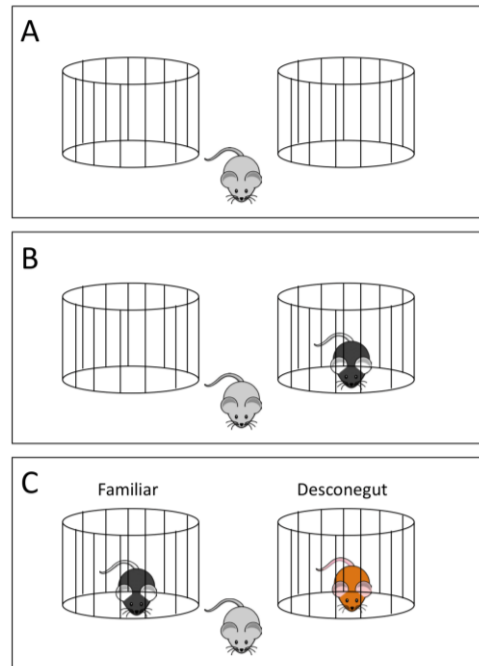


Figura 1. Esquema de l'estructura on es realitza el test d'interacció social. A) Habitució; B) Avaluació del comportament social; C) Avaluació de la preferència per les relacions socials noves.

2.4.8. *Open Field*

El test d'*Open Field* (OF) s'utilitza per mesurar l'activitat locomotora, l'ansietat i la curiositat exploratòria (Seibenhener i Wooten, 2015). Aquest test, igual que el test de EPM es basa en la tendència natural d'un animal a explorar i protegir-se, de manera que en un espai quadrangular o rectangular, un ratolí normal passarà més temps a la zona perifèrica que a la central (la zona més ansiògena), però que, tot i així explorarà la zona central en algun moment. Segons la il·luminació aplicada, els objectius del test difereixen lleument. L'OF a intensitats lumíniques baixes s'utilitza per avaluar la locomoció i la curiositat exploratòria principalment, mentre que a intensitats lumíniques altes (utilitzat com a estímul advers) s'incentiva la mesura de l'ansietat, tot i que també serveix per mesurar la curiositat exploratòria i la locomoció (Valle, 1970). Aquest test

el vam realitzar a les instal·lacions de Barcelona i Frankfurt, per confirmar que podíem observar el mateix fenotip en ambdues instal·lacions, sota diferents condicions lumíniques.

A Barcelona es van utilitzar dos aparells diferents per l'OF en funció de la intensitat lumínica utilitzada (50 lux i 500 lux). L'aparell utilitzat a baixa intensitat lumínica (50 lux) consisteix en una caixa de base quadrada (50 cm x 50 cm x 35 cm) equipada amb un sistema de monitorització (SMART, Letica). En aquest espai es defineix una zona central de 22,5 cm x 22,5 cm. A l'inici del test s'introdueix el ratolí al centre de la caixa i s'enregistra el seu comportament durant 5 minuts. Els paràmetres que es calculen per a cadascuna de les zones (perifèrica i central) són: la distància, les entrades, el temps descansant, el temps dedicat i la velocitat mitja. L'aparell utilitzat a alta intensitat lumínica (500 lux) consisteix en una caixa de base rectangular (70 cm x 90 cm x 60 cm). Aquesta caixa té marcats 63 quadrats (10 cm x 10 cm) en tot el terra. En aquest test la zona perifèrica correspon a la fila de quadrats més externs i tota la resta corresponen a la zona central. L'animal és introduït en el quadrat central i s'enregistra la seva activitat durant 5 minuts. Els paràmetres que es calculen per a cada zona (central i perifèrica) són: el nombre de quadrats travessats, les entrades, el temps dedicat a cada zona, l'activitat vertical (número de vegades que el ratolí se sustenta sobre les potes posteriors), el nombre de defecacions i el nombre de miccions. Aquests tests es van realitzar amb 10 ratolins mascles per cada genotip.

L'aparell utilitzat a Frankfurt consistia en una gàbia amb una base quadrada amb parets fosques verticals (40 cm x 40 cm x 35 cm) il·luminada a 250 lux i equipada amb una càmera a sobre que enregistra tots els moviments dels ratolins. La zona central tenia una mida de 15 cm d'ample i 15 cm de llarg. El test té una durada de 5 minuts durant els quals es mesuren múltiples paràmetres com la distància recorreguda, el temps passat a cada zona (perifèrica i central), el número d'entrades a cada zona o el temps de congelació (*freezing*, llindar establert a partir de 2 segons), definit com l'absència de moviment amb excepció de la respiració. Tots els comportaments són mesurats per un programa de seguiment de vídeo anomenat ANY-MAZE (v6.0). Aquest test es va realitzar amb ratolins mascles (9 WT, 8 HET i 7 KO) i femelles (4 WT, 2 HET i 6 KO).

2.4.9. Light Dark Box

El test de *Light Dark Box* (LDB) ens permet mesurar la locomoció i ansietat dels ratolins i, igual que l'EPM i l'OF, es basa en el conflicte existent entre la tendència innata que tenen els ratolins a buscar refugi en àrees fosques i petites i la curiositat exploratòria dels ambients nous (Takao i Miyakawa, 2006). L'aparell utilitzat en aquest test és una gàbia amb una base quadrada dividida en dos compartiments iguals (20 cm x 40 cm x 35 cm), un il·luminat i un fosc. Els dos

compartiments es comuniquen amb una petita porta (7,5 cm x 7,5 cm) per permetre la transició entre ambdós espais. El test comença col·locant el ratolí al compartiment il·luminat de la gàbia. El test té una durada de 5 minuts durant els quals es mesuren diversos paràmetres similars als obtinguts en l'OF, com la distància recorreguda, el temps passat a cada zona (il·luminada i fosca), el número de transicions entre les dues zones o el temps de congelació (*freezing*, l'indar establert als 2 segons). Tots els comportaments són mesurats per un programa de seguiment de vídeo anomenat ANY-MAZE (v6.0). Aquest test es va realitzar amb ratolins mascles (9 WT, 8 HET i 7 KO) i femelles (4 WT, 2 HET i 6 KO) a les instal·lacions de Frankfurt.

2.4.10. *Prepulse Inhibition test*

En els animals, inclosos els humans, la *startle response*, o resposta de sobresalt, és una resposta defensiva inconscient a estímuls sobtats o amenaçadors com per exemple un soroll. El *Prepulse Inhibition test* (PPI) ens permet calcular la magnitud d'aquesta resposta, així com el percentatge d'inhibició d'aquesta resposta davant la presència d'un preimpuls previ a l'estímul que genera el sobresalt. Aquest test s'ha utilitzat àmpliament en models animals d'esquizofrènia, ja que nombrosos articles han descrit que hi ha un dèficit de la inhibició de la resposta de sobresalt en pacients amb esquizofrènia (Swerdlow *et al.*, 2008). Aquest test es realitza mitjançant el *SRLab™ startle response System* (San Diego Instruments), que consisteix en un cilindre de plexiglàs (4 cm de diàmetre i 13 cm de longitud), on s'introdueix el ratolí, muntat sobre una plataforma de plexiglàs que conté un element piezoelèctric just a sota que és sensible al moviment del ratolí. Aquesta estructura es col·loca en una cambra ventilada i aïllada del so. Es realitzen estímuls sonors mitjançant un altaveu situat a 33 cm per sobre l'estructura de plexiglàs.

El test comença amb la fase d'habitució (5 minuts), en què el ratolí és exposat únicament al soroll de fons (65 dB) que continuarà al llarg de tot el test. Després d'aquest temps d'habitució, venen tres fases en què s'avalua resposta de sobresalt. En la primera fase, el ratolí és exposat a 6 impulsos sonors a 120 dB amb un interval entre intents (ITI) de 10 segons. En la segona fase el ratolí és exposat a 70 combinacions de l'impuls sonor (120 dB) amb preimpulsos (4 dB, 8 dB, 12 dB i 16dB) amb duracions de l'ITI variables (de 20 a 30 segons). Les combinacions possibles són: només impuls, només preimpuls, preimpuls amb impuls i sense impuls. En l'última fase el ratolí és exposat novament a 6 impulsos de 120 dB amb un ITI variable. A partir de les mesures enregistrades es pot calcular la resposta de sobresalt en cadascuna de les tres fases així com el percentatge d'inhibició d'aquesta resposta donada per cada preimpuls segons s'ha descrit prèviament (Geyer i Swerdlow, 1998). Aquest test es va realitzar amb ratolins mascles (9 WT, 8 HET i 7 KO) i femelles (4 WT, 2 HET i 6 KO) a les instal·lacions de Frankfurt.

2.4.11. Fear Conditioning

El test de *Fear Conditioning*, o condicionament de la por, es basa en el condicionament Pavlovià de la por, i ens permet avaluar la capacitat dels ratolins d'aprendre i recordar una associació entre un estímul ambiental i una experiència adversa. Estudis previs han demostrat que la incapacitat d'extingir els records de por comporta un increment de l'activació de l'amígdala (Rauch *et al.*, 2000; Stein *et al.*, 2002; Dilger *et al.*, 2003) i, a més, és un dels símptomes centrals en diversos trastorns psiquiàtrics, com fòbies específiques, trastorn d'ansietat generalitzada o social, trastorn de pànic i trastorn d'estrès posttraumàtic. L'extinció del record es veu reflectit en la fase d'extinció del test, mentre que en la fase d'evocació comprovem si realment han extingit el record de la por.

Aquest test es va realitzar en el *UGO BASILE ANY-maze controlled Fear Conditioning System* (Ugo Basile), una caixa transparent (23 cm x 23 cm x 36 cm) amb un terra amb barres electrificades situades en paral·lel. Aquest sistema permet construir dos ambients visuals, tàctils i olfactoris completament diferents (Context A i B) (Figures 2A i 2B). El test es duu a terme en tres dies consecutius. El primer dia es realitza el condicionament, el segon dia l'extinció i el tercer dia l'evocació (*recall*). El condicionament es realitza en el context A (Figura 2A) i, en canvi, l'extinció i l'evocació es realitzen en el context B (Figura 2B), que consisteix en la mateixa caixa transparent però amb un paper amb una quadrícula negra i blanca a les parets i un terra suau de color gris. L'estructura se situa en una caixa atenuant del so que conté un ventilador, una llum dual (visible/infraroja), un altaveu i una càmera (Figura 2C). A més, també s'introdueix una fragància d'extracte de vainilla dins la caixa per canviar l'ambient olfactiu. En aquest test s'enregistra el temps de congelació (*freezing*, lllindar establert als 2 segons), mitjançant el *software ANY-maze* (v6.0).

En la fase de condicionament el ratolí és col·locat a la cambra amb el context A i es comença amb 3 minuts d'habituaació. A continuació el ratolí és exposat a un to que persisteix durant 30 segons (estímul condicionat de 80 dB). Durant els dos últims segons d'aquests 30 segons, el ratolí experimenta un lleu electroxoc als peus (estímul no condicionat de 0,4 mA). Després, el ratolí segueix a la cambra durant dos minuts més fins que es repeteix el mateix procediment (to durant 30 segons amb electroxoc als últims dos segons). El ratolí se sotmet a aquest procediment tres vegades. L'endemà es realitza la fase d'extinció en el context B, que comença amb 3 minuts d'habituaació, i després es realitzen 16 tons de 30 segons sense electroxoc amb 5 segons de pausa entre cada to. Finalment, al tercer dia es realitza l'evocació en el context B, en què es realitzen tres exposicions al to de 30 segons sense l'electroxoc separades 5 segons entre elles. Aquest va

ser l'últim test que es va realitzar amb la cohort de ratolins femelles (4 WT, 2 HET i 6 KO) a les instal·lacions de Frankfurt.

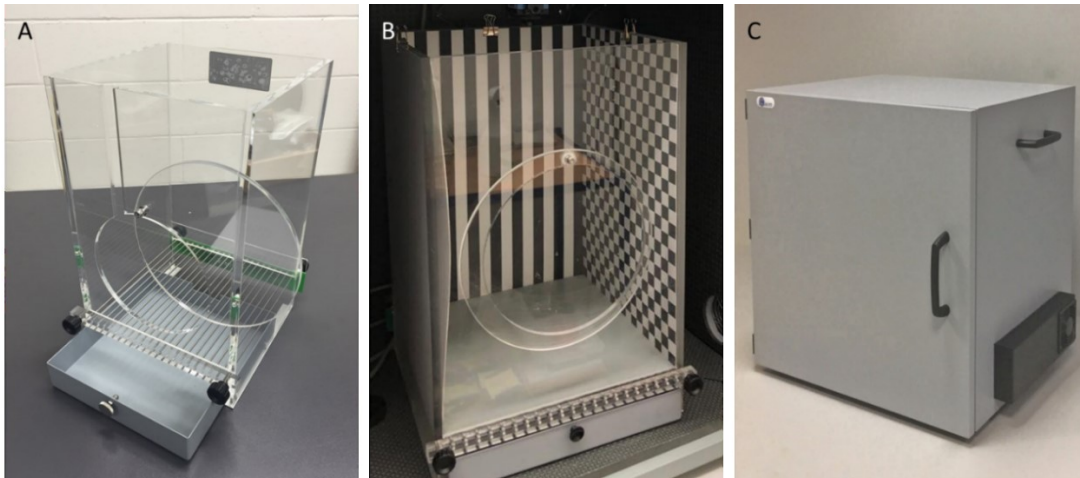


Figura 2. UGO BASILE ANY-maze controlled Fear Conditioning System. A) Context A: Terra reixat i parets transparents; B) Context B: el terra reixat s'ha tapat amb una cartolina grisa clara llisa i les parets s'han recobert amb cartolines amb línies o quadrícules blanques i negres; C) Caixa on se situen les estructures dels Contexts A i B.

2.4.12. Continuous Performance Task

El test de *Continuous Performance Task* (CPT) permet el reconeixement d'impulsivitat i alteracions de diferents tipus d'atenció (Kim *et al.*, 2015) en un paradigma que s'assembla a la prova que es realitza en humans (Caballero-Puntiverio *et al.*, 2019).

Aquest test es realitza en cambres equipades amb pantalles tàctils, un terra reixat, un dispensador de recompenses associat a una llum, unes bombes per subministrar líquid a la zona del dispensador i un altaveu (*Bussey-Saksida Mouse Touch Screen Chambers*, Campden Instruments Ltd.) (Figura 3). Cada cambra es troba dins d'una caixa que atenua el so exterior. El programari utilitzat per dissenyar les diferents sessions del test és el *Animal Behavior Environment Test II (ABET II) software*, i aquest s'acoba a les pantalles tàctils de les cambres mitjançant el controlador WhiskerServer®. La recompensa utilitzada al llarg de totes les tasques és llet amb gust de maduixa (Müllersmilch).

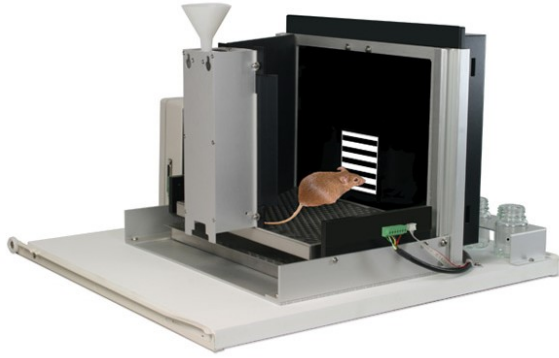


Figura 3. Esquema de la *Bussey-Saksida Mouse Touch Screen Chambers*. El terra de l'estructura és reixat. A la dreta de la imatge veiem la pantalla tàctil on apareixeran els estímuls S+ i S-. A darrera de la pantalla hi ha la bomba encarregada de dispensar el líquid a més de dos pots de vidre que es poden omplir amb la recompensa líquida que s'ha de bombejar. A la part esquerra de la imatge hi ha la zona del dispensador on s'allibera la recompensa en forma de líquid o en forma de pèl·let.

Abans de començar l'experiment es restringeix l'accés al menjar dels ratolins a una hora al dia durant almenys dos dies. Aquesta restricció és continuada al llarg de tot l'experiment de manera que els animals tindran accés al menjar només durant una hora en acabar la sessió. Al llarg d'una setmana es realitzen 6 sessions, i quan s'acaba la sisena sessió els ratolins tindran accés al menjar durant 24h, per així aconseguir que no perdin més del 10% del seu pes al llarg de tot el test.

El primer dia del test els ratolins s'habituen a la cambra i a la recompensa utilitzada tot introduint-los a la cambra durant 20 minuts amb la recompensa en el dispensador. Aquesta sessió es repeteix en dies consecutius fins que el ratolí tasta la recompensa per primera vegada.

A continuació entrem en la fase d'entrenament, que consta de quatre etapes diferents (Taula 1A) que es realitzen en sessions diàries de 45 minuts que es repeteixen fins que els ratolins assoleixen els criteris requerits per a cada etapa. Totes les sessions tenen una durada màxima de 45 minuts, que finalitzarà abans si l'animal aconsegueix 100 recompenses.

En la **primera etapa** (Taula 1A), apareix durant 10 segons (*Stimulus Duration, SD*) un quadrat blanc (3,5 cm x 3,5 cm) emmarcat per una línia blanca situat al centre i 2 cm per sobre de la part baixa de la pantalla. Per poder tenir en compte aquelles respostes que s'han iniciat quan la imatge encara estava present, es defineix la retenció limitada (*Limited Hold, LH*), que correspon a la duració de l'estímul (SD), que és de 10 segons en aquesta etapa, sumat a 0,5 segons extres durant els quals la pantalla encara és activa tot i que no presenta l'estímul. Si el ratolí toca el quadrat blanc, aquest desapareix de la pantalla i, immediatament, es produeix un to d'un segon i s'encén la llum del dispensador de recompenses on la bomba de líquid haurà bombejat la recompensa. Quan el musell del ratolí entra dins del dispensador, s'apaga la llum i entrem dins interval entre intents (*inter-trial interval, ITI*) que dura dos segons. Quan finalitza l'ITI torna a començar el següent intent en què es torna a exposar l'animal al quadrat blanc. Si el ratolí no toca el quadrat







blanc durant els 10 segons d'exposició, s'inicia un ITI i es produeix el següent estímul. Si el ratolí toca la zona on hi havia el quadrat blanc durant l'ITI, aquest interval es repeteix i l'exposició del següent quadrat blanc s'endarrereix. Es considera que l'animal supera aquesta primera etapa si aconseguix més de 60 recompenses en una sessió.

En la **segona etapa** (Taula 1A) substituïm el quadrat blanc per l'estímul positiu (S+) que utilitzarem durant la resta de etapes, a més, la SD es redueix a dos segons, i per tant, el LH passa a ser de 2,5s. L'estímul S+ consisteix en una superfície quadrada de la mateixa mida que l'estímul utilitzat a l'anterior etapa però amb ratlles horitzontals o verticals (un ratolí sempre serà exposat al mateix estímul S+ al llarg de totes les sessions, és a dir, sempre serà exposat a ratlles horitzontals o ratlles verticals). En aquesta etapa i les següents, després de recollir la recompensa hi ha una fase de demora per la ingesta (*Ingestion Delay*, ID) de cinc segons abans de començar l'ITI. Tota la resta dels paràmetres i criteris són idèntics als de la etapa anterior.

La **tercera etapa** (Taula 1A) de la fase d'entrenament consisteix en la introducció d'un estímul negatiu (S-). L'S- és un floc de neu i es presenta juntament amb l'S+ aleatòriament un 50% dels intents. En aquest cas, l'ITI s'incrementa a 5 segons. Si el ratolí toca l'estímul S- durant el LH, aquest desapareix, es dona l'ITI i a continuació comença l'intent de correcció. En l'intent de correcció es presenta l'estímul S- i aquest es repeteix fins que l'animal deixa de reaccionar-hi durant el LH. Tota la resta de paràmetres són els mateixos que els de la segona etapa. Considerem que l'animal adquireix el coneixement si és capaç de discernir entre els estímuls S+ i S- ($d' > 0,6$) durant dues sessions consecutives. Tots els ratolins han de fer aquesta sessió com a mínim 6 vegades per poder establir la corba d'adquisició.

En la **quarta etapa** (Taula 1A) de la fase d'entrenament, el floc de neu ja no s'utilitza com a S-, i aquest és substituït per quatre S- nous. Igual que en l'etapa 2, la probabilitat que aparegui un estímul S- o un S+ segueix sent del 50%. Tota la resta de paràmetres i criteris són els mateixos que els utilitzats en l'etapa 3, amb l'única diferència que en aquesta etapa els animals han de fer com a mínim set sessions enlloc de sis per establir la corba d'adquisició.

Taula 1. Característiques de les diferents etapes de la fase d'entrenament i dels tests realitzats en la *Continuous Performance Task*. Adaptada de Kim et al., 2015.

A) Fase d'entrenament			SD	ITI	Criteri
Primera	Conjunt d'estímuls Quadrat blanc 		10 s	2 s	Aconseguir 60 recompenses
Segona	S+ (100%) 		2 s	2 s	Aconseguir 60 recompenses
Tercera	S+ (50%) 	Un S- (50%) 	2 s	5 s	$d' > 0,6$ en dues sessions consecutives; com a mínim 6 sessions
Quarta	S+ (50%) 	Quatre S- (50%) 	2 s	5 s	$d' > 0,6$ en dues sessions consecutives; com a mínim 7 sessions

B) Tests		
Test	Rang considerat	Número de sessions
<i>Stimulus Duration</i>	0,5 - 2s	2
<i>Stimulus Contrast</i>	12,5 - 100%	1
<i>Flanker distractors</i>	Congruent, incongruent o sense distractors	1
<i>Long Inter Trial Interval</i>	5 i 10 segons	1

SD: Duració de l'estímul; ITI: Interval entre intents.

Finalment, entrem en l'etapa dels tests (Taula 1B), que es realitzen just després de la tercera etapa i de la quarta etapa de la fase d'entrenament. En els tests, es canvien alguns dels paràmetres utilitzats en l'última sessió de l'etapa de l'entrenament corresponent, per avaluar com canvia l'execució dels ratolins a l'introduir aquests canvis. Entre test i test es restableix l'execució de la tasca dels ratolins fent una sessió amb les mateixes característiques que l'últim entrenament realitzat (característiques de la tercera etapa o quarta etapa) per assegurar-nos que segueix discriminant els dos estímuls ($d' > 0,6$). Si s'obté $d' < 0,6$, es repeteix la sessió d'entrenament fins que l'animal torna a ser capaç de discriminar els dos estímuls abans de procedir al següent test.

Stimulus Duration test (Taula 1B). En aquest test es redueix la duració de l'estímul (SD) per incrementar l'atenció necessària per a la tasca. Aquest canvi és una modificació clàssica en estudis en humans per valorar l'atenció. Es realitzen dues sessions: en la primera sessió l'animal és exposat a estímuls S+ i S- que duren 1, 1,5 o 2 s de forma aleatòria, i en la segona els estímuls duren 0,5, 0,75 o 1 s. En les dues sessions però, tot i que desaparegui l'estímul, el LH segueix sent de 2,5 s, de manera que el ratolí pot aconseguir la recompensa (si hi havia l'estímul S+) o entrar dins de l'intent de correcció (si hi havia l'estímul S-).

Stimulus Contrast test (Taula 1B). Aquest test es fa en una sessió i s'utilitza per avaluar l'execució de la tasca quan hi ha condicions de percepció més difícils (reducció del contrast). La il·luminació

de l'estímul segueix sent la mateixa però canvia el contrast en disminuir la diferència d'il·luminació entre les línies fosques i les clares. El contrast es canvia al 12,5, 25, 50 i 100%.

Flanker distractors (Taula 1B). Aquest test es fa en una sessió i s'utilitza per determinar si la presència de distractors flanquejant l'estímul central redueixen l'atenció en l'execució. Uns nivells d'atenció correctes requereixen la inhibició d'informació que pot distreure. S'afegeixen dos requadres de la mateixa mida que el quadrat inicial a cada costat del quadrat inicial on apareixeran distractors. Dels tres requadres, l'únic amb el qual el ratolí obtindrà resposta quan el toqui serà el central. Al llarg de la sessió, es donaran tres situacions: a) sense distractor: no apareix cap imatge en els requadres laterals; b) amb distractors congruents: en els requadres laterals apareixen dos estímuls idèntics al que apareix al mig (tots S+ o tots S-); o c) amb distractors no congruents: els estímuls que apareixen als requadres laterals no són iguals que l'estímul central (al centre S+ i als laterals dos S- iguals o al centre S- i als laterals dos S+).

Long Inter Trial Interval (Taula 1B). Aquest test es fa en una sessió i es canvia la duració de l'interval que hi ha entre la presentació dels estímuls. S'ha demostrat que taxes altes i baixes de presentació de l'estímul redueixen l'atenció sostinguda per a l'execució de la tasca. S'afegeixen dues noves condicions, 5 i 10 segons d'ITI.

Al llarg de totes les sessions el *software* enregistra múltiples paràmetres, entre els quals destaquem: *Hit*: resposta a S+; *Miss*: sense resposta a S+; *False Alarm* (FA): resposta a S- i *Correct Rejection* (CR): sense resposta a S-. Aquests paràmetres seran utilitzats posteriorment per calcular la taxa d'èxit (*Hit Rate*, HR), la taxa d'error (*False Alarm Rate*, FAR), l'índex de discriminació (d') i el biaix en la resposta o voluntat de respondre a l'estímul (c), entre altres valors que ens poden ajudar a conèixer la capacitat d'execució de la tasca dels ratolins. Es considera que l'animal és capaç de discernir entre els dos estímuls si $d' > 0,6$. Valors alts de c indiquen una resposta conservadora i valors baixos es relacionen amb una resposta liberal.

$$\text{Hit rate (HR)} = \frac{\text{Hit}}{\text{Hit} + \text{Miss}}$$

$$\text{False alarm rate (FAR)} = \frac{\text{FA}}{\text{FA} + \text{CR}}$$

$$d' = z(\text{HR}) - z(\text{FAR})$$

$$c = -\frac{z(\text{HR}) - z(\text{FAR})}{2}$$

Aquest va ser l'últim test que es va realitzar amb cinc ratolins mascles per genotip a les instal·lacions de Frankfurt.

2.4.13. Model d'autoadministració de cocaïna

Aquest test ens permet avaluar el comportament que presenten els ratolins envers la cocaïna. Es realitza en cambres operants (*Model ENV-307A-CT, Med Associates, Inc.*) equipades amb dos forats, un dels quals serà seleccionat de forma aleatòria com a forat actiu i l'altre com a forat inactiu. Cada forat té associat un llum i un altaveu que funcionen com a estímuls condicionats. Les cambres són d'alumini, tenen un terra reixat i se situen dins de caixes equipades amb ventiladors que atenuen el so exterior.

Per començar, els ratolins són intervinguts quirúrgicament per introduir-los un catèter silàstic flexible de 6 cm de llarg (0,3 mm de diàmetre interior i 0,6 mm de diàmetre exterior) com a mínim tres dies abans de començar el protocol. El catèter s'insereix 1,3 cm a la vena jugular dreta per la qual s'administrarà la cocaïna (Soria *et al.*, 2005). El catèter es renta diàriament amb una solució salina heparinitzada (30 UI/ml). L'hidroclorur de cocaïna es va obtenir del Ministeri de Sanitat, Política Social i Igualtat (Madrid, Espanya) i es va dissoldre en sèrum fisiològic estèril al 0,9 %. La dosi administrada en cada infusió és 0,5 mg/kg.

El protocol d'autoadministració consisteix en quatre etapes: **adquisició**, **motivació**, **extinció** i **recaiguda**, que es realitzen al llarg de una sessió diària de dues hores de duració (si no s'indica el contrari) a les caixes d'autoadministració (Figura 4). La **fase d'adquisició** comença amb 5 sessions de raó fixa 1 (FR1), en què cada cop que el ratolí fica el musell al forat actiu rep una dosi de cocaïna juntament amb un estímulo lumínic, que s'associa a l'efecte de la droga (estímulo condicionat). A continuació es fan 5 sessions més de raó fixa 3 (FR3) en què el ratolí ha de ficar el musell al forat actiu 3 vegades per obtenir una sola dosi de cocaïna juntament amb l'estímulo lumínic (Figura 4). Després de cada administració d'una dosi de cocaïna hi ha un període de temps anomenat *Time-out* (TO) en què el ratolí no obté cap recompensa quan fica el musell al forat actiu.

A continuació es testa la **motivació** del ratolí per la cocaïna amb un test de raó progressiva (PR; una sessió de 4 hores), en què augmenta progressivament el nombre de vegades que ha de posar el musell al forat actiu per obtenir una infusió de cocaïna (1, 3, 5, 7, 9, 15, 20, 25, 33, 43, 55, 70, 80, 90, 100) (Figura 4). La sessió acaba al cap de 4 hores o fins que el ratolí deixa de respondre a qualsevol forat durant una hora. Un cop acabat el PR, es comprova la funcionalitat del catèter mitjançant l'administració de 0,1 ml de l'anestèsic tiopental sòdic (5mg/ml). Aquells animals que

no presenten símptomes d'anestèsia als 3 segons posteriors a la infusió, es considera que són tiopental negatius i es descarten de l'estudi.

Els animals tiopental positiu continuen amb l'**etapa d'extinció**, que consisteix en múltiples sessions en les mateixes condicions que l'adquisició però sense recompensa ni estímul lumínic associat per simular un període d'abstinència durant 20 dies (Figura 4). Per últim, es va testar la possibilitat de **recaiguda** mitjançant una sessió en què se sotmet l'animal a un estímul lumínic associat a la recompensa sense administració de cocaïna cada vegada que respon al forat actiu, i s'observa el seu comportament (Figura 4). Es considera que l'animal restableix el comportament de cerca de la cocaïna si aquest estímul fa que l'animal torni a ficar el musell de forma selectiva i reiterada al forat actiu, tot i que no hi hagi recompensa associada.



Figura 4. Esquema de les sessions d'autoadministració de cocaïna. FR: Raó fixa. PR: Raó progressiva.

Aquest test va ser l'últim que es va realitzar amb la cohort de ratolins de Barcelona, partint de 10 ratolins mascles per cada genotip, dels quals 8 WT, 9 HET i 7 KO van sobreviure a la cirurgia i van donar positiu al test de tiopental. Totes les anàlisis s'han dut a terme considerant únicament els animals tiopental positius, ja que els resultats obtinguts en l'autoadministració en els tiopental negatius poden estar esbiaixats per la falta de viabilitat del catèter.

2.4.14. Anàlisi estadística

Les dades del *Prepulse Inhibition Test* van ser processades amb Matlab abans de poder utilitzar-les en les anàlisis estadístiques posteriors. Tots els resultats que es mostren es van analitzar estadísticament mitjançant ANOVA, amb el programa SPSS 19.0 o amb el programa GraphPad Prism8, amb un llindar de significació de $p < 0,05$.

3. RESULTATS I DISCUSSIÓ

3.1. Estudi d'expressió en el cervell de ratolí

En l'estudi d'expressió en el cervell de ratolí vam poder confirmar l'expressió del gen *C2orf82* en totes les àrees cerebrals avaluades, mostrant una expressió més elevada en el nucli accumbens i en el nucli estriat dorsal i expressió més baixa en l'amígdala (Figura 5).

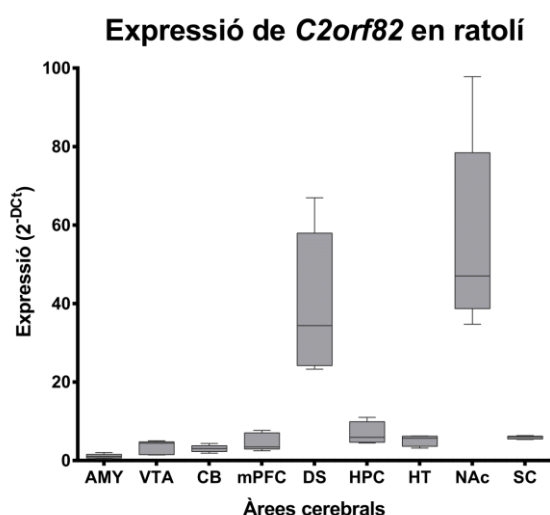


Figura 5. Quantificació absoluta ($2^{-\Delta Ct}$) de l'expressió del gen *C2orf82* en diferents àrees cerebrals de ratolí. Ct: *Cycle threshold*; AMY: Amígdala; VTA: Àrea tegmental ventral; CB: Cerebel; mPFC: Escorça prefrontal medial; DS: Nucli estriat dorsal; HPC: Hipocamp; HT: Hipotàlem; NAc: Nucli accumbens; SC: medulla espinal.

3.2. Predicció de l'estructura 3D de la proteïna

Un estudi previ publicat l'any 2011 va descriure una possible predicció de l'estructura de la proteïna *C2orf82* en ratolí (Heinonen *et al.*, 2011). En la seva predicció la proteïna estava formada per un pèptid senyal del residu 1-24, una regió extracel·lular del residu 25-92 que contenia nou serines/treonines que podien ser glicosilades, una hèlix transmembrana del residu 93 al 115 i una regió intracel·lular del residu 116 al 121. Vam decidir realitzar de nou un estudi de predicció de l'estructura de la proteïna, ja que des del 2011 fins l'actualitat s'han desenvolupat noves eines bioinformàtiques per a la predicció de l'estructura secundària i terciària de polipèptids.

Vam començar avaluant si la proteïna contenia regions desordenades, és a dir, regions sense cap estructura secundària concreta. Mitjançant el servidor PrDOS (Ishida i Kinoshita, 2007) vam poder observar una alta probabilitat de tenir residus desordenats a l'extrem N-terminal (residus 1-6), al mig (44-73 i 81-87) i a la regió C-terminal de la proteïna (117-121) (Figura 6), tot indicant que les altres dues regions podien adoptar estructures secundàries definides.

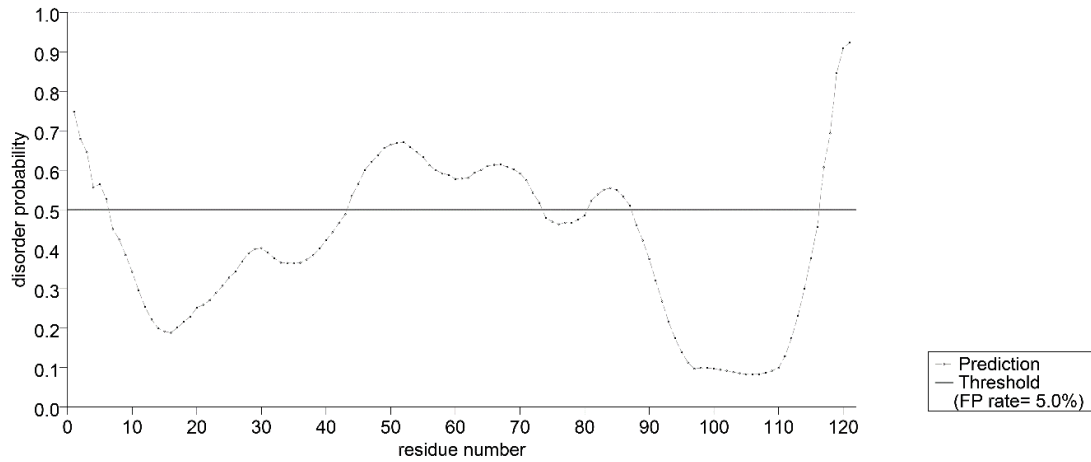


Figura 6. Predicció de regions desordenades en la proteïna (PrDOS). El llindar establert per predir estructures desordenades és 0,5.

Per tal de saber quin tipus d'estructura secundària es podria formar en aquestes dues regions vam utilitzar el programa JPred4 (Drozdetskiy *et al.*, 2015), que ens va predir dues hèlix alfa dels residus 4 al 19 i dels residus 92 al 116, amb una alta probabilitat associada (Figura 7).

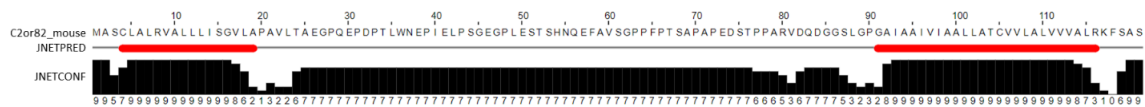


Figura 7. Predicció d'estructures secundàries a la proteïna (JPred4). Les hèlix alfa s'indiquen amb tubs vermells i les làmines beta amb fletxes verdes. JNETPRED: predicció consens; JNETCONF: estimació de la confiança de la predicció, valors alts indiquen una confiança alta.

A continuació vam avaluar si les hèlix alfa tenien una disposició transmembrana amb l'eina TMHMM (Krogh *et al.*, 2001), amb la qual vam predir que ambdues hèlixs alfa tenien una alta probabilitat de ser transmembrana i que la part central desordenada semblava estar situada a la part extracel·lular (Figura 8).

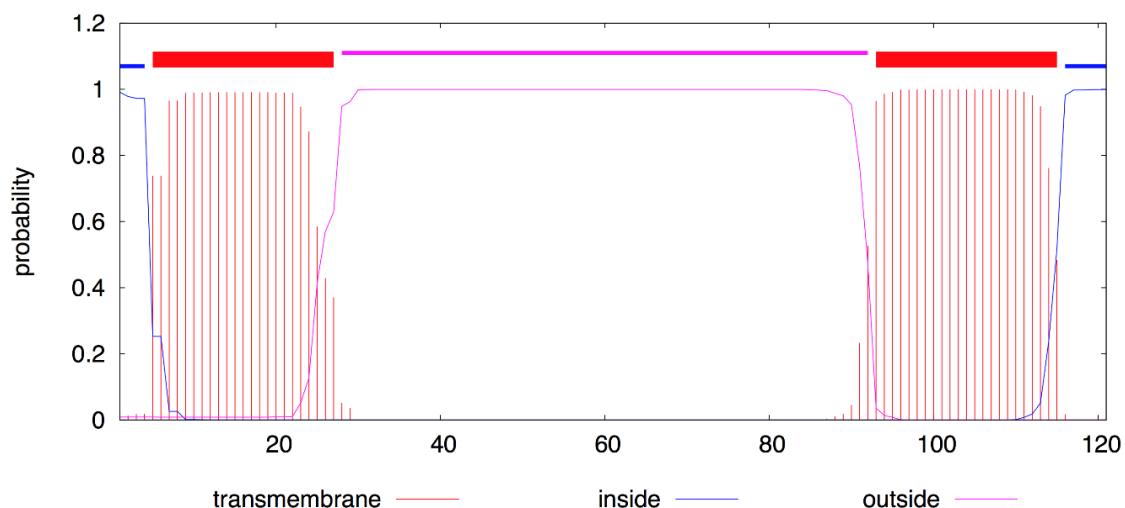


Figura 8. Predicció d'hèlix transmembrana (TMHMM v.2.0). Els números de l'eix de les X indiquen l'aminoàcid de la proteïna. Les hèlix transmembrana s'indiquen amb tubs horitzontals vermells. Les línies vermelles verticals indiquen la probabilitat de cada residu de formar part d'una hèlix transmembrana. En rosa s'indica la probabilitat de cada residu d'estar situat a la part extracel·lular. En blau s'indica la probabilitat de cada residu d'estar situat a la part intracel·lular.

En tractar-se d'una possible proteïna transmembrana, vam explorar si aquesta contenia algun pèptid senyal utilitzant el servidor Signal P5.0 (Almagro Armenteros *et al.*, 2019). Aquest servidor ens va predir un pèptid senyal del residu 1 al 24, solapant-se amb la predicció de l'hèlix alfa situada a l'extrem N-terminal (Figura 9).

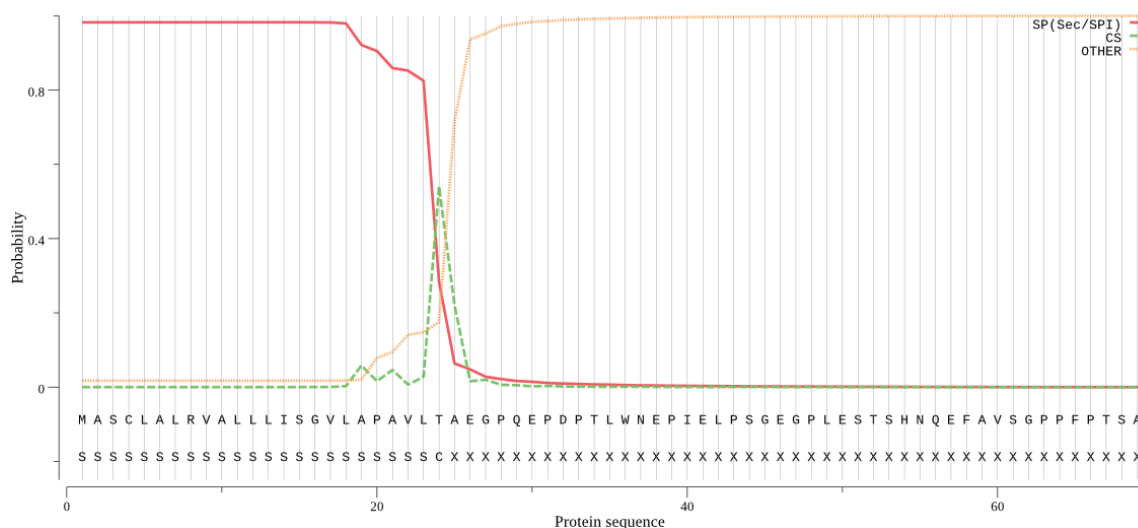


Figura 9. Predicció de pèptids senyals (Signal P5.0). SP: Pèptid senyal; Sec/SPI: substrats secretors tallats per SPase I; CS: lloc d'escissió; Other: els aminoàcids no pertanyen a un pèptid senyal.

Finalment vam explorar la presència de llocs de glicosilació amb el servidor Net-O-Glyc4.0 (Steentoft *et al.*, 2013). Aquesta eina bioinformàtica ens va predir nou llocs de glicosilació als residus 34, 44, 51, 53, 61, 67, 68, 75 i 76. Finalment, el servidor Robetta, un servei de predicció de l'estructura de proteïnes (Raman *et al.*, 2009; Song *et al.*, 2013) que combina diferents eines per al càlcul de possibles models 3D, ens va proposar cinc models possibles per a l'estructura de la proteïna (Figura 10), i en tots ells la proteïna contenia dues hèlixes alfa, amb una probabilitat més alta per la hèlix alfa situada a l'extrem C-terminal.



Figura 10. Predicció de l'estructura 3D de la proteïna C2orf82 de ratolí (Robetta). El degradat de colors indica la confiança de la predicció, on el color groc indica una menor confiança en la predicció i el lila una confiança més elevada. En la imatge es mostra un dels cinc models possibles proposats pel servidor Robetta.

En resum, totes les prediccions donen suport a dues hèlixes transmembrana que anirien des del residu 4 al 19 i des del residu 92 al 116. A més, també vam predir la presència d'un pèptid senyal a l'extrem N-terminal de la proteïna que arribaria fins el residu 24, on se situaria el lloc d'escissió, i nou llocs de glicosilació situats entre els residus 34 i 76. Tots aquests resultats van en la mateixa línia que els resultats obtinguts per Heinonen *et al.* 2011, amb l'única diferència que les noves prediccions no deixen clar si hi ha realment una segona hèlix transmembrana a l'extrem N-terminal. Recentment, un estudi va avaluar la conformació topològica de la proteïna *C2orf82*, tot confirmant que es tracta d'un proteoglicà de transmembrana curt de tipus I amb un únic domini transmembrana situat a la regió C-terminal (Jaiswal *et al.*, 2020), coincidint amb la predicció feta per Heinonen *et al.*, 2011.

3.3. Caracterització fenotípica

No es van observar alteracions de l'estat de salut general, ni problemes de coordinació motora, ni anhedonia en cap dels genotips estudiats (*C2orf82* *-/-*, *C2orf82* *+/-* i *C2orf82* *+/+*). Tampoc vam observar diferències entre genotips pel que fa al comportament social ni en el condicionament de la por.

En les caixes PHECOMP no hi ha diferències entre els genotips durant el període d'anhedonia però sí durant el període d'habitució. S'observa hipolocomoció en els ratolins KO en comparar-los amb els ratolins WT, definida per una menor distància recorreguda (Figura 11A) així com una tendència a una menor activitat (Figura 11B), a una menor velocitat mitja (Figura 11C), a menys estereotípies (Figura 11D) i a un augment del temps de descans (Figura 11E). A més, també s'observa una menor activitat vertical (Figura 11F), que ens indica una disminució del comportament exploratori per part dels ratolins KO, comportament que estudis previs relacionen amb un augment de l'anxietat (Lever, Burton i O'Keefe, 2006).

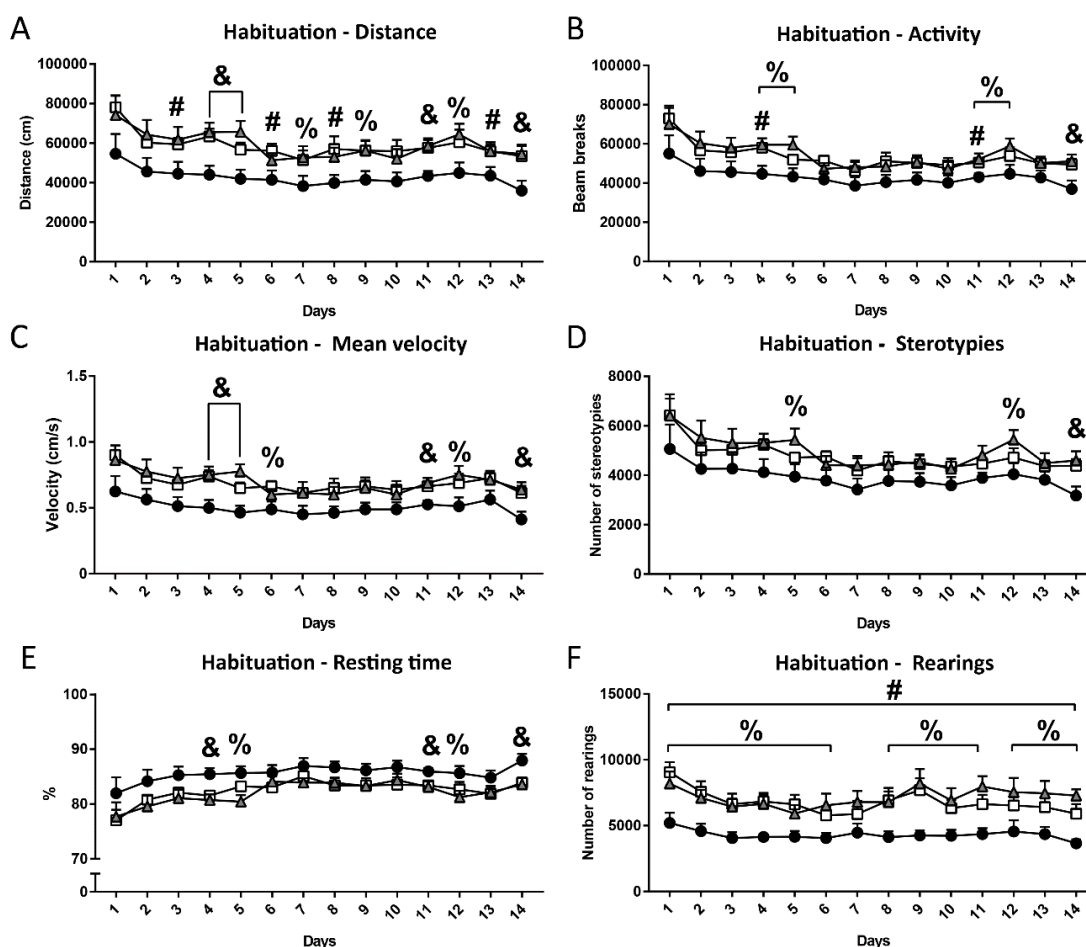


Figura 11. Resultats de les caixes PHECOMP en el període d'habitució. A) Distància total (en cm) recorreguda en períodes de 24h al llarg de 14 dies. B) Activitat en base als *beam breaks* en períodes de 24 h al llarg de 14 dies. C) Velocitat mitjana (cm/s) mesurada en períodes de 24 h al llarg de 14 dies. D) Número d'estereotípies en períodes de 24 h al llarg de 14 dies. E) Percentatge de temps descansant en períodes de 24 h al llarg de 14 dies. F) Activitat vertical en base al nombre de vegades que l'animal s'ha posat dret sobre les dues potes posteriors en períodes de 24 h al llarg de 14 dies. Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT (\square ; n = 10), HET (\triangle ; n = 10) i KO (\bullet ; n = 10). #/%/&p < 0,05. #KO Vs. WT; %KO Vs. HET; &KO Vs. WT i HET.

En el NOR no s'observen diferències pel que fa al temps d'exploració dedicat a cada objecte. Tot i així, s'observa una tendència a un menor índex de discriminació per part dels ratolins KO i HET en comparar-los amb els ratolins WT (p -valor = 0,054 i 0,064 respectivament) (Figura 12). Aquesta tendència a una discriminació disminuïda podria ser indicativa d'una menor memòria dependent de l'hipocamp i aprenentatge (Lueptow, 2017). No obstant, caldria ampliar el número d'animals per confirmar aquest resultat.

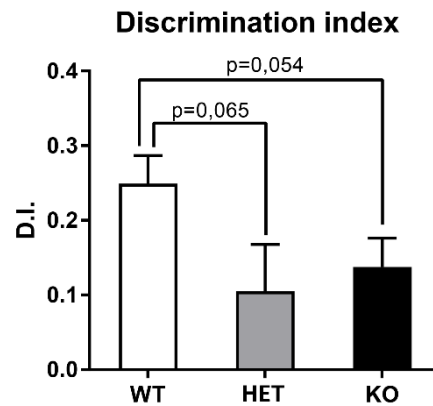


Figura 12. Resultats de l'índex de discriminació calculat entre l'objecte nou i el familiar en el test de *Novel Object Recognition* (NOR). Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT ($n = 10$), HET ($n = 10$) i KO ($n = 10$).

En l'EPM els ratolins KO dediquen menys temps i entren menys vegades als braços oberts (zona ansiògena). En canvi, tot i que també entren menys cops als braços tancats, hi passen més temps (Figura 13). Aquests resultats podrien indicar que els ratolins KO mostren una menor curiositat exploratòria així com possiblement també un fenotip d'ansietat degut a la preferència pels braços tancats (zona menys ansiògena). No obstant, aquests fenotips no es poden confirmar ja que les diferències en el número d'entrades totals ens indiquen que els ratolins presenten una certa hipolocomoció, fenotip que pot afectar la correcta interpretació de la resta de resultats.

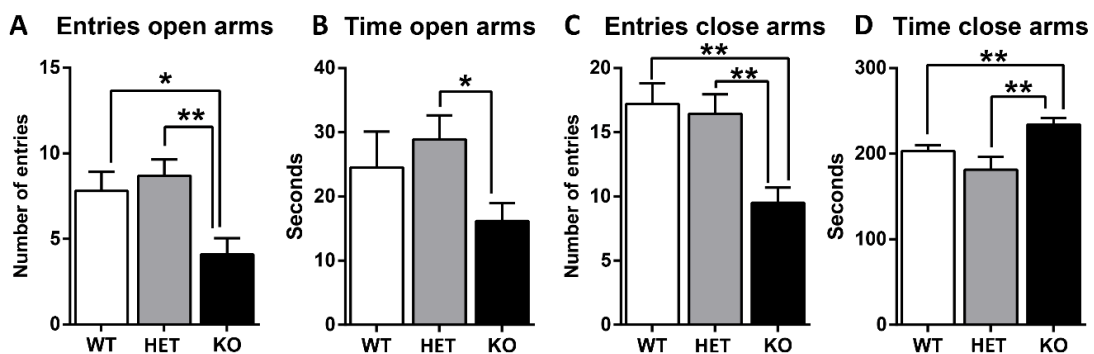


Figura 13. Resultats de l'*Elevated Plus Maze* (EPM). A) Nombre d'entrades als braços oberts. B) Temps (en segons) que els ratolins han estat als braços oberts. C) Nombre d'entrades als braços tancats. D) Temps (en segons) que els ratolins han estat als braços tancats. Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT ($n = 10$), HET ($n = 10$) i KO ($n = 10$). * $p < 0,05$; ** $p < 0,01$.

En l'OF també vam observar el fenotip d'hipolocomoció per part dels ratolins KO en comparar-los amb els ratolins WT així com un possible fenotip d'ansietat i a més, vam confirmar que obteníem els mateixos resultats en les dues instal·lacions on s'han realitzat els experiments. En l'OF a 50 i a 250 lux (realitzats a Barcelona i a Frankfurt, respectivament), vam observar el fenotip d'hipolocomoció, en aquest cas definit per una menor distància recorreguda total (Figures 14A i 15A), així com en cadascuna de les àrees definides (Figures 14B i 15B) i un menor número de transicions entre ambdues zones (Figures 14C i 15C). Per altra banda, també vam observar un possible fenotip d'ansietat ja que els ratolins KO passaven menys temps a la zona interna (més ansiògena) i per tant més temps a la zona perifèrica (Figures 14D i 15D). A més, a les instal·lacions de Frankfurt vam observar un temps de congelació total més elevat per part dels ratolins KO (Figura 15E), fenotip que també es relaciona amb ansietat. A Barcelona, tot i que no vam poder avaluar el temps de congelació, sí que vam observar un increment de temps de descans a la zona perifèrica (Figura 14E). Aquests fenotips van ser observats en mascles, però no es van poder observar en la cohort de femelles utilitzada a Frankfurt, probablement degut al baix nombre de femelles inclòs.

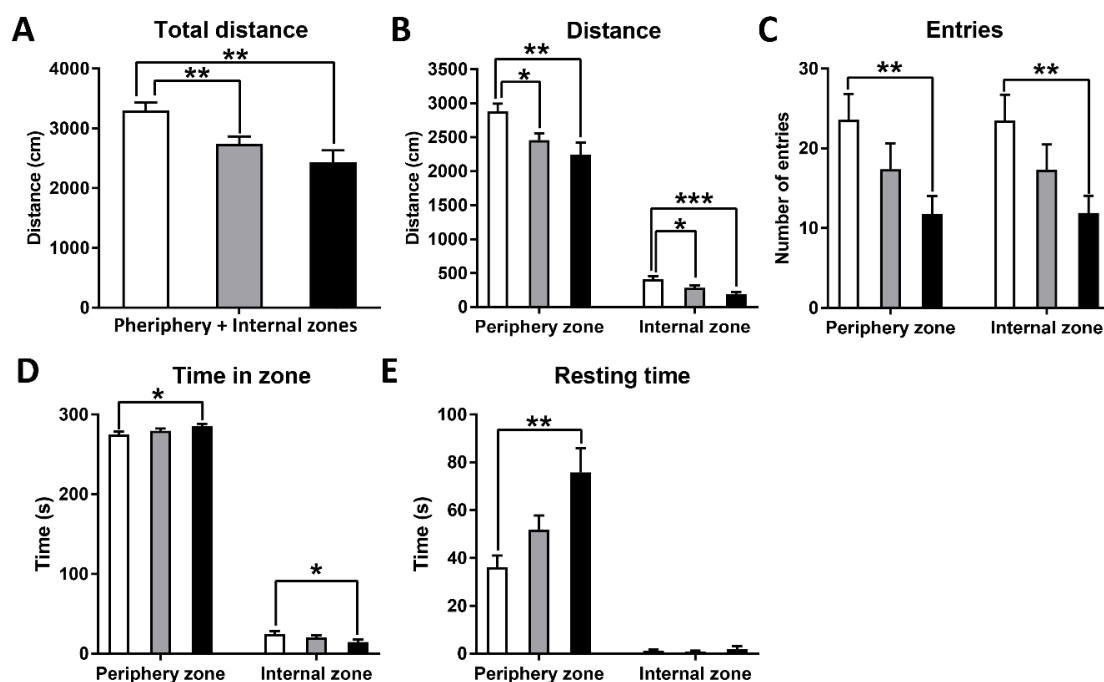


Figura 14. Resultats de l'Open Field (OF) a 50 lux d'intensitat al llarg del temps total del test. A) Distància total (en cm) recorreguda. B) Distància (en cm) recorreguda a cada zona. C) Nombre d'entrades a cada zona. D) Temps (en segons) passat a cada zona. E) Temps (en segons) de descans a cada zona. Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT (□; n = 10), HET (▒; n = 10) i KO (■; n = 10). *p < 0,05; **p < 0,01; ***p < 0,001.

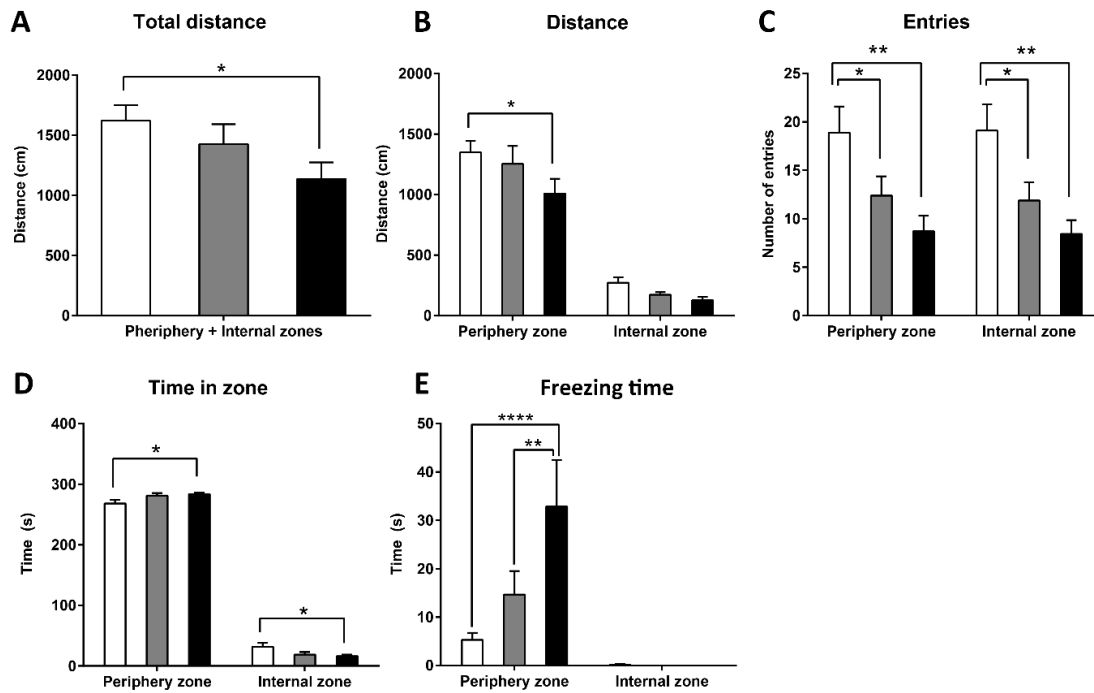


Figura 15. Resultats de l'Open Field (OF) a 250 lux d'intensitat. A) Distància total (en cm) recorreguda. B) Distància (en cm) recorreguda a cada zona. C) Nombre d'entrades a cada zona. D) Temps (en segons) passat a cada zona. E) Temps (en segons) de congelació total. Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT (□; n = 9), HET (▒; n = 8) i KO (■; n = 7). *p < 0,05; **p < 0,01.

A les instal·lacions de Barcelona també es va realitzar el test d'OF a 500 lux, una intensitat lumínica que incentiva les mesures d'ansietat i de curiositat exploratòria. En aquestes condicions lumíniques vam replicar els fenotips d'hipolocomoció (Figures 16A, 16B i 16C) però a més, també vam poder observar una menor curiositat exploratòria dels ratolins KO definida per una menor activitat vertical (Figura 16D), que també es relaciona amb un fenotip d'ansietat.

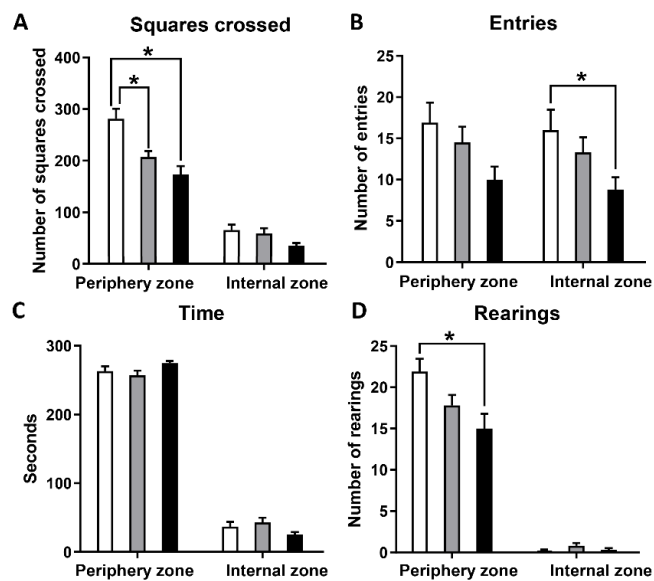


Figura 16. Resultats de l'Open Field (OF) a 500 lux d'intensitat.

A) Distància recorreguda (quadrats travessats) a cada zona. B) Nombre d'entrades a cada zona. C) Temps (en segons) passat a cada zona. D) Activitat vertical en base al nombre de vegades que l'animal s'ha posat dret sobre les dues potes posteriors. Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT (□; n = 10), HET (▒; n = 10) i KO (■; n = 10). *p < 0,05; **p < 0,01; ***p < 0,001

Els resultats obtinguts en el LDB van en la mateixa línia que els obtinguts en l'OF, ja que també observem pels ratolins KO una menor distància recorreguda total (Figura 17A), un menor nombre de transicions (Figura 17B) i un temps de congelació més elevat (Figura 17D), indicatius d'hipolocomoció i possible ansietat en els ratolins mutants. En aquest test però, no observem diferències en el temps passat a cada zona (Figura 17C). Tot i així, igual que en l'EPM, hem de tenir en compte que la hipolocomoció observada en l'OF (50, 250 i 500 lux) i en el LDB pot afectar la interpretació dels resultats pel que fa al fenotip d'ansietat. Igual que en el test d'OF, no observem aquests fenotips en la cohort de femelles, probablement pel baix nombre de femelles inclòs.

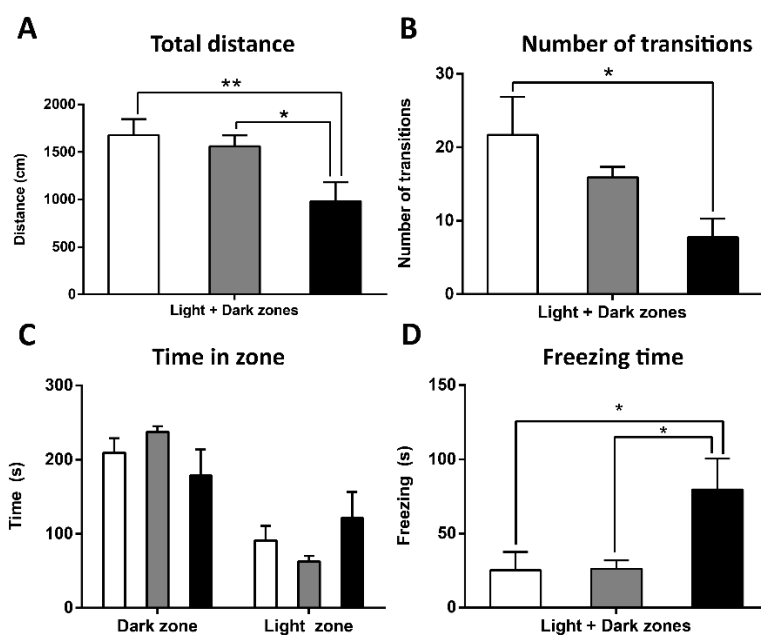


Figura 17. Resultats del Light Dark Box (LDB).

A) Distància (en cm) recorreguda total. B) Nombre de transicions entre ambdues zones. C) Temps (en segons) passat a cada zona. D) Temps (en segons) total de congelació. Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT (□; n = 9), HET (▒; n = 8) i KO (■; n = 7). *p < 0,05; **p < 0,01.

En el PPI hem observat una tendència a un augment del percentatge d'inhibició de de la resposta de sobresalt (*startle response*) en la cohort de femelles (Figura 18), tot i que aquestes diferències no arriben a ser significatives, probablement a causa del baix nombre de femelles utilitzat.

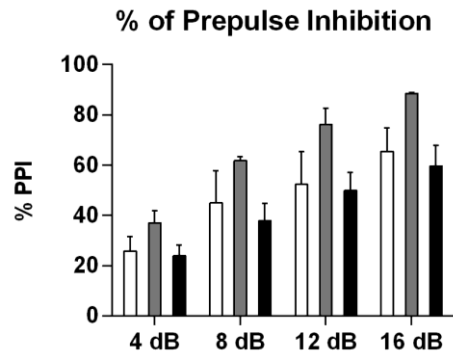


Figura 18. Resultats del percentatge d'inhibició de la resposta de sobresalt (*startle response*) en el test de del *Prepulse Inhibition* (PPI). Es representa la mitjana \pm error estàndard (SEM). Animals femelles WT (\square ; n = 4), HET (\blacksquare ; n = 2) i KO (\blacksquare ; n = 6).

En la CPT, les corbes d'adquisició de l'aprenentatge de les tasques corresponents a la tercera i quarta etapa d'entrenament (Figures 19 i 20), mostren que tot els ratolins han après les tasques i que no hi ha diferències d'aprenentatge segons el genotip.

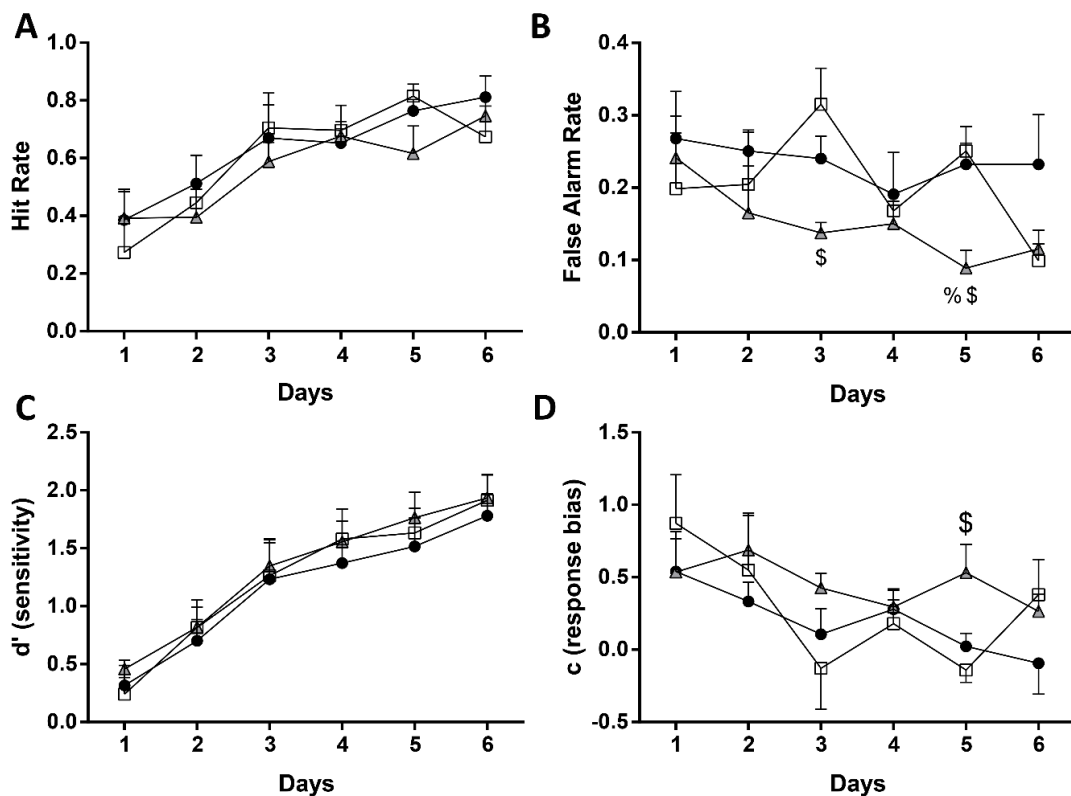


Figura 19. Corbes d'aprenentatge obtingudes després de l'aprenentatge de la tercera etapa d'entrenament (un estímul S+ i un S-) en la *Continuous Performance Task* (CPT). A) Taxa d'èxit (HR). B) Taxa d'error (FAR). C) Índex de discriminació (d'). D) Biaix en la resposta o voluntat de respondre a l'estímul (c). Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT (\square ; n = 5), HET (\triangle ; n = 5) i KO (\bullet ; n = 5). %/\$p < 0,05. %KO Vs. HET; \$HET Vs. WT.

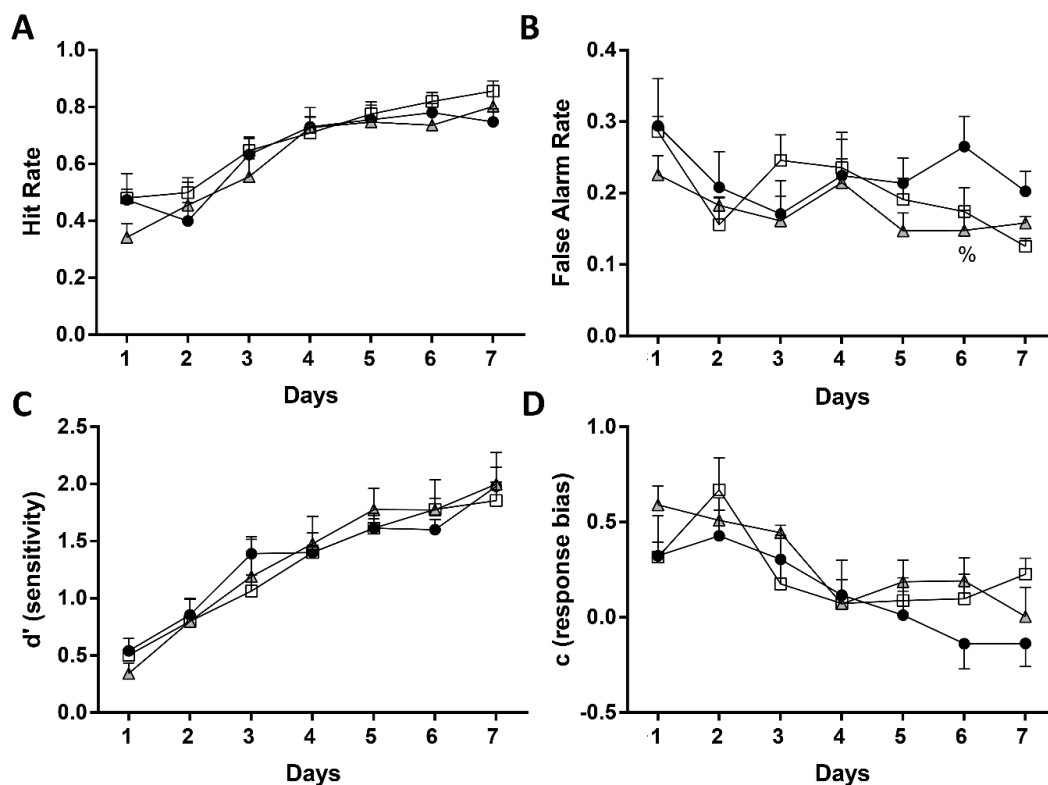


Figura 20. Corbes d'aprenentatge obtingudes després de l'aprenentatge de la quarta etapa d'entrenament (un estímul S+ i quatre S-) en la *Continuous Performance Task*. A) Taxa d'èxit (HR). B) Taxa d'error (FAR). C) Índex de discriminació (d'). D) Biaix en la resposta o voluntat de respondre a l'estímul (c). Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT (\square ; n = 5), HET (\blacktriangle ; n = 5) i KO (\bullet ; n = 5). %p < 0,05. %KO Vs. HET.

En la CPT, només vam observar diferències entre els genotips en el test de *Stimulus Duration* (SD). Vam observar una taxa d'error (FAR) més elevada per part dels ratolins KO quan la duració de l'estímul era més curta en el test realitzat després de la tercera etapa d'entrenament (Figura 21A). Aquestes diferències s'observen també en el test SD realitzat després de la quarta etapa d'entrenament però, en aquest cas, les diferències no són tan acusades i per tant només mostren significació en la SD de 0,5 s (Figura 21B). En aquesta gràfica a més, en tractar-se d'una tasca de més complexitat (quatre estímuls S- enlloc d'un), fins i tot els ratolins HET mostren un valor de FAR més elevat que els ratolins WT. Aquest resultat pot indicar una major impulsivitat dels ratolins KO en comparació amb els ratolins WT, ja que toquen la pantalla sense importar quina és la imatge que apareix quan la duració de l'estímul és més curta. Aquestes diferències també es veuen reflectides en ratolins HET quan la tasca a realitzar és més complicada (Figura 21B).

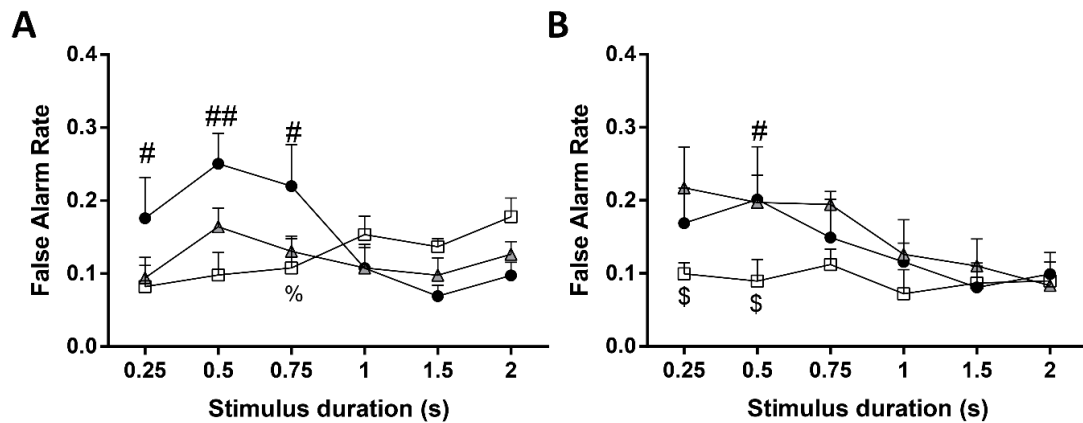


Figura 21. Resultats dels tests realitzats en la *Continuous Performance Task*. A) *Stimulus Duration test* realitzat després de la tercera etapa d'entrenament (un estímul S+ i un estímul S-). B) *Stimulus Duration test* realitzat després de la quarta etapa d'entrenament (un estímul S+ i quatre S-). Es representa la mitjana \pm error estàndard (SEM). Animals mascles WT (□; n = 5), HET (△; n = 5) i KO (●; n = 5). #/%/§p < 0,05; ##p < 0,01. #KO Vs. WT; %KO Vs. HET; §HET Vs. WT.

Finalment, en el test d'autoadministració de cocaïna, els resultats obtinguts en els períodes de raó fixa FR1 i FR3 suggereixen que no hi ha diferències entre genotips en la capacitat d'aprendre el funcionament del sistema d'autoadministració. En la primera sessió (FR1), els animals KO mostren una tendència a una major impulsivitat definida per un nombre més alt de respostes al forat actiu durant el període de *Time-out* (Figura 22A). No obstant això, aquesta tendència desapareix a partir de la segona sessió (Figura 22A), la qual cosa suggereix un efecte preexistent de la impulsivitat degut al genotip, que amb el consum continuat de cocaïna en FR1 i FR3 va desaparèixer. Per altra banda, els ratolins KO presenten una menor motivació per la cocaïna després de 10 dies de consum continuat segons el paradigma de la raó progressiva. A més, els HET presenten un fenotip intermedi (Figura 22B). Finalment, després de l'extinció, hi ha la fase de restabliment de cerca de la droga, que simula la recaiguda en humans. En aquesta fase vam poder observar que tots els ratolins (WT, HET i KO) restablien el comportament, però que tant els ratolins HET com els KO cessaven abans el comportament (Figures 22C i 22D), i que per tant aquests resultats podien indicar que el gen *C2orf82* està implicat en la recaiguda en el consum de cocaïna.

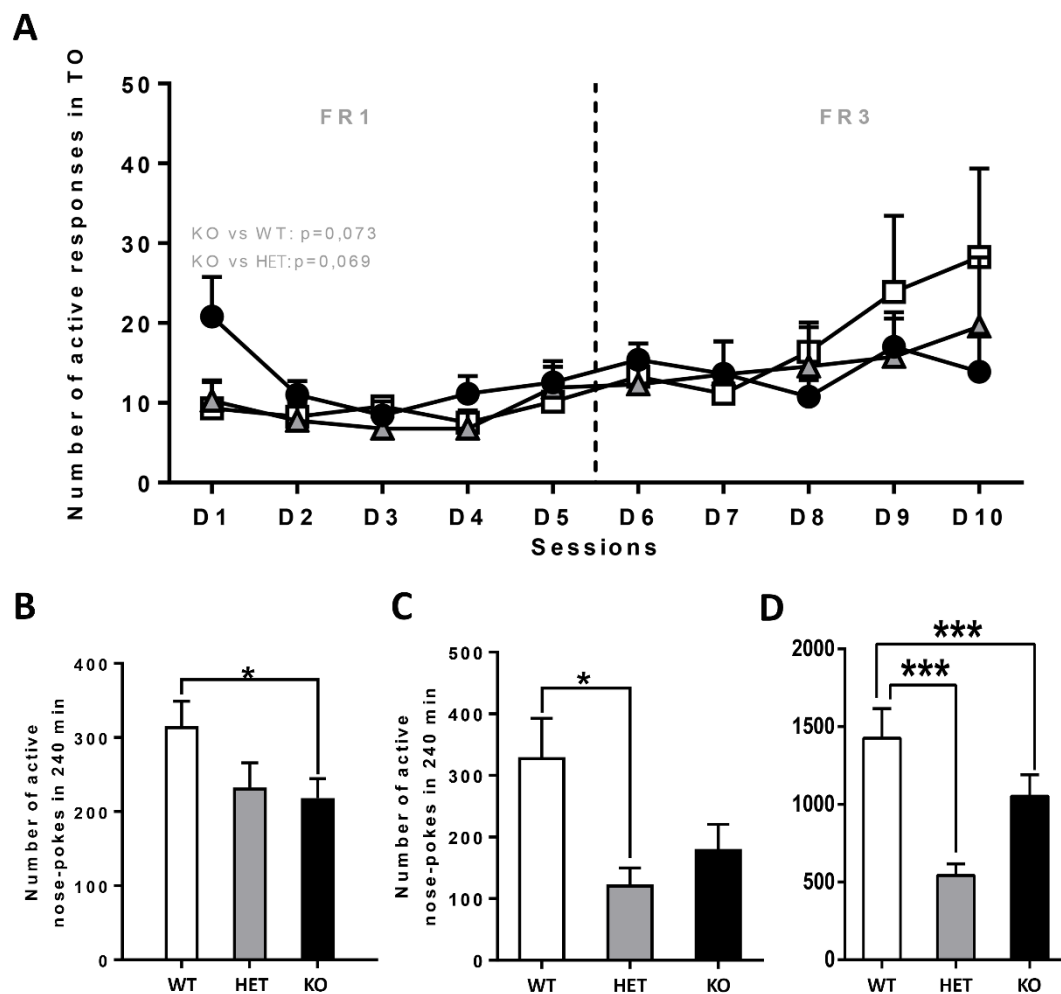


Figura 22. Resultats del test d'autoadministració de cocaïna. A) Mitjana del nombre de respostes al forat actiu en el període de *time-out* durant l'adquisició de l'addicció amb raó fixa 1 (FR1) i raó fixa 3 (FR3). B) Nombre de respostes al forat actiu durant el test de raó progressiva (PR). C) Nombre de respostes al forat actiu en la fase de recaiguda. D) Percentatge de canvi existent entre els períodes d'extinció i recaiguda induïda per l'estímul lumínic. Es representa la mitjana \pm error estàndard (SEM). * $p < 0,05$; *** $p < 0,001$. Animals mascles WT (\square o \square ; $n = 8$), HET (\triangle o \triangle ; $n = 9$) i KO (\bullet o \bullet ; $n = 7$).

4. CONCLUSIONS I PERSPECTIVES DE FUTUR

Les diferents aproximacions utilitzades ens han permès caracteritzar l'expressió de *C2orf82* al cervell així com aprofundir en la seva possible contribució a la susceptibilitat al TDAH. En primer lloc, en l'estudi d'expressió en el cervell hem observat una major expressió en el nucli estriat dorsal i el nucli accumbens. El nucli estriat dorsal està implicat en la memòria i l'aprenentatge i el nucli accumbens és àrea central del sistema de recompensa, un sistema essencial en la majoria de trastorns psiquiàtrics, entre ells el TDAH i les addiccions, objecte d'aquesta Tesi.

En els tests de caracterització fenotípica general no hem observat problemes de salut ni de coordinació motora, confirmant que el ratolí homozigot per la mutació no presenta cap fenotip incapacitant que pugui emascarar els resultats obtinguts en els tests de comportament (Taula 2). Pel que fa als tests de la fenotipació dirigida a avaluar trets i trastorns psiquiàtrics hem observat hipolocomoció i una menor curiositat exploratòria en els ratolins KO en diferents tests realitzats (Taula 2). En alguns dels tests també hem observat un possible fenotip d'ansietat, però aquest fenotip no s'ha pogut confirmar ja que en aquests mateixos tests veiem també el fenotip d'hipolocomoció, que pot afectar a la correcta interpretació de la resta de resultats (Taula 2). Per tal de confirmar el fenotip d'ansietat podríem realitzar el test de *Fear Conditioning*, aquest cop amb una cohort de mascles, ja que en aquest test la hipolocomoció presentada pels ratolins KO no afectaria la interpretació dels resultats. Per altra banda, també és interessant destacar que hem observat certs indicis d'impulsivitat, un dels símptomes principals del TDAH, tant en el test de *Stimulus Duration* de la *Continuous Performance Task* com en la primera sessió de raó fixa del test d'autoadministració de cocaïna (Taula 2). Finalment, en el test d'autoadministració de cocaïna també hem observat que els ratolins KO presenten una menor motivació per la droga i que malgrat que tots els ratolins recorden l'estímul associat amb els afectes reforçants de la droga en la fase de recaiguda, els ratolins HET i KO cessen abans el comportament de cerca de la droga, comportament que podem interpretar com una menor recaiguda en el consum.

Taula 2. Fenotips observats en els ratolins mascles KO en comparació amb els ratolins WT.

	IT	CP	NOR	Act	EPM	R	TIS	OF	LDB	PPI	CPT	CSA
Sense diferències entre genotips	✓			✓		✓				✓		
Menor curiositat exploratòria		✓			✓			✓				
Menor índex de discriminació			✓									
Hipolocomoció		✓			✓			✓	✓			
Ansietat					✓			✓	✓			
Impulsivitat											✓	✓
Motivació per la droga												✓
Cessament primerenc de la cerca de cocaïna												✓

IT: Test d'Irwin; CP: Caixes PHECOMP; NOR: *Novel Object Recognition*; Act: Test d'activitat o actimetria; EPM: *Elevated Plus Maze*; R: *Rotarod*; TIS: Test Interacció social; OF: *Open Field*; LDB: *Light Dark Box*; PPI: *Prepulse Inhibition test*; FC: *Fear Conditioning*; CPT: *Continuous Performance Task*; CSA: *Cocaine self-administration*.

Part dels resultats obtinguts en aquest treball es poden interpretar en el context de la teoria hipodopaminèrgica, una teoria formulada per explicar la simptomatologia dels pacients amb TDAH (Tripp i Wickens, 2009). En condicions normals, l'alliberació de dopamina en resposta a una recompensa es transfereix a l'estímul condicionat previ, i és aquesta transferència la que produeix

una recompensa immediata a nivell cel·lular quan hi ha una recompensa del comportament tardana. En els pacients amb TDAH en canvi, hi ha un dèficit d'aquesta transferència a l'estímul condicionat previ, de manera que això comporta una demora de la recompensa a nivell cel·lular si la recompensa del comportament és tardana. En els resultats obtinguts en aquest treball hem observat una major expressió del gen *C2orf82* en el nucli accumbens, una regió central en el sistema de recompensa, i per tant una de les àrees que pot resultar més afectada per la no expressió del gen en els ratolins KO, afectació es podria relacionar amb les deficiències en aquest sistema. Per altra banda, també hem observat una menor motivació per aconseguir la droga (cocaïna) per part dels ratolins KO quan això suposa més esforç, i també un cessament primerenc de la cerca de la droga. Aquests dos fenotips es podrien explicar amb el dèficit de transferència de dopamina a l'estímul condicionat ja que en ambdós casos els ratolins KO no rebrien la recompensa esperada a l'estímul condicionat previ i per tant cessarien abans el comportament. Tot i així, el número d'animals utilitzats en els tests de comportament encara és molt limitat, i hauríem d'incrementar la mida mostral per tal de poder confirmar els resultats obtinguts, així com algunes tendències observades.

Article 7

“Identificació de variants genètiques que afecten la metilació en cervell amb efectes pleotròpics en trastorns psiquiàtrics”

Resum:

Els trastorns psiquiàtrics afecten un 29,2% de la població general almenys una vegada a la seva vida. Estudis de bessons, familiars i GWAS han demostrat que hi ha una genètica compartida entre molts dels trastorns psiquiàtrics, però hi ha a més mecanismes subjacents específics encara no ben estudiats. Els estudis realitzats al llarg dels últims anys han demostrat que la metilació del DNA té un paper rellevant en els trastorns complexos, i, a més, s’ha descrit un enriquiment de variants que tenen influència sobre la metilació entre les variants que s’han associat amb determinats trastorns psiquiàtrics mitjançant GWAS. En aquest estudi hem explorat la possible contribució de la metilació específica d’al·lel (ASM) a un conjunt de vuit trastorns psiquiàtrics combinant dades genètiques (una metanàlisi GWAS que inclou les vuit patologies), epigenètiques (ASM en cervell) i d’expressió (eQTLs en cervell). Com a resultat, hem identificat 80 SNPs amb efectes pleotròpics en trastorns psiquiàtrics que mostren un efecte contrari en la metilació i l’expressió gènica, tot assenyalant vuit gens candidats: *ZSCAN29*, *ZSCAN31*, *BTN3A2*, *DDAH2*, *HAPLN4*, *ARTN*, *FAM109B* i *NAGA*. *ZSCAN29* és el gen amb l’efecte pleotròpic més ampli, ja que mostra associació amb cinc dels vuit trastorns psiquiàtrics considerats, seguit de *ZSCAN31* i *BTN3A2*, associats amb tres trastorns. Curiosament, tots aquests gens se superposen amb variacions del número de còpies (CNV) relacionades amb fenotips cognitius i trets o trastorns psiquiàtrics, s’expressen en el cervell, i set d’ells s’han associat prèviament amb trastorns psiquiàtrics específics. En resum, hem identificat vuit gens que contribuirien a trastorns psiquiàtrics, assenyalats per SNPs que tenen un efecte sobre la metilació del DNA i sobre l’expressió gènica al cervell.

Referència:

Pineda-Cirera L, Cabana-Domínguez J, Cormand B, Fernández-Castillo N. Identification of genetic variants influencing methylation in brain with pleiotropic effects on psychiatric disorders. (Manuscrit en preparació).

Identification of genetic variants influencing methylation in brain with pleiotropic effects on psychiatric disorders

Running title: Allele-specific methylation and psychiatric disorders

Laura Pineda-Cirera^{1,2,3,4}, Judit Cabana-Domínguez Ph. D.^{1,2,3,4}, Noèlia Fernández-Castillo Ph. D.^{1,2,3,4,&}, Bru Cormand Ph. D.^{1,2,3,4,&}

¹ Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain

² Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Spain

³ Institut de Biomedicina de la Universitat de Barcelona (IBUB), Catalonia, Spain

⁴ Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Barcelona, Catalonia Spain

&Senior authors that equally contributed

CORRESPONDING AUTHORS:

Bru Cormand: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain, Avinguda Diagonal 643, edifici Prevosti, 3^a planta, 08028, Barcelona, Catalonia, Spain. Tel.: (+34) 93 4021013. Fax: (+34) 93 4034420; Email: bcormand@ub.edu

Noèlia Fernández Castillo: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain, Avinguda Diagonal 643, edifici Prevosti, 1^a planta, sala 134. 08028, Barcelona, Catalonia, Spain. Tel.: (+34) 93 4037082. Fax: (+34) 93 4034420; Email: noefernandez@ub.edu

ABSTRACT

Psychiatric disorders affect 29.2% of the global population at least once in their life. Shared genetics for psychiatric disorders has been proved by twin, family and GWAS studies, but the underlying molecular mechanisms still remain largely unknown. Work over the past decade has demonstrated a crucial role of DNA methylation in complex disorders, and remarkably, an enrichment of common genetic variants influencing allele-specific methylation (ASM) has been reported among variants associated with specific psychiatric disorders. In the present study we have assessed the contribution of ASM to a set of eight psychiatric disorders by combining genetic (cross-disorder meta-analysis of eight psychiatric disorders), epigenetic (ASM in brain) and expression data (eQTLs in brain). We identified 80 SNPs with pleiotropic effects on psychiatric disorders that show an opposite effect on methylation and gene expression, pointing at eight candidate genes: *ZSCAN29*, *ZSCAN31*, *BTN3A2*, *DDAH2*, *HAPLN4*, *ARTN*, *FAM109B* and *NAGA*. *ZSCAN29* is the gene with the broadest pleiotropic effect, showing associations with five out of eight psychiatric disorders considered, followed by *ZSCAN31* and *BTN3A2*, associated with three disorders. Interestingly, all these genes overlap with CNVs related to cognitive phenotypes and

psychiatric traits or conditions, they are expressed in brain, and seven of them have previously been associated with specific psychiatric disorders, supporting the results obtained in the present study. To sum up, we have identified eight candidate genes contributing to psychiatric disorders highlighted by SNPs influencing DNA methylation and gene expression in brain.

INTRODUCTION

Psychiatric disorders are complex disorders that encompass multiple neurologic conditions characterized by changes in mood, thinking and/or behavior that affect 29.2% of the worldwide population at least once in lifetime (Steel et al., 2014). The classification and diagnostic criteria for all psychiatric disorders are collected in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013). Psychiatric disorders show in general a high heritability, and they share part of the underlying genetic basis, as proven by family, twin and more recently, genome-wide association studies (GWAS) (Brainstorm Consortium et al., 2018; Smoller et al., 2019). A recent GWAS meta-analysis performed across eight psychiatric disorders has detected 146 risk *loci*, 109 of them associated with at least two psychiatric disorders, confirming a pleiotropic effect of genetic risk variants. This study also reports high genetic correlations between most pairs of studied disorders, the highest between schizophrenia and bipolar disorder ($r_g = 0.70 \pm 0.02$, $p\text{-value} < 1E-06$) (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). However, most of the underlying genetic basis of these conditions still remain undetermined, and it is known as missing heritability. The omnigenic model proposes that sub-threshold variants described in GWAS can explain part of this missing heritability of complex disorders (Boyle, Li and Pritchard, 2017; Liu, Li and Pritchard, 2019). Interestingly, genetic variants associated with different complex disorders (either with genome-wide significance or subthreshold) are enriched in active chromatin regions, including promoters and enhancers (Maurano et al., 2012; Kundaje et al., 2015; Wang et al., 2016), stressing the importance of considering the functional effect of identified variants.

Epigenetics plays an important role in complex disorders like psychiatric conditions, as it mediates the interaction between genetics and the environment. DNA methylation is the most stable epigenetic mechanism and it is influenced by multiple environmental exposures (pollution, tobacco smoke, or nutritional factors, among others) leading to new DNA methylation patterns (Li and Zhang, 2014; Martin and Fry, 2018). Allele-specific methylation (ASM) occurs when the different alleles of one SNP are associated with differential levels of methylation at a CpG site, which may lead to changes in gene expression (Meaburn, Schalkwyk and Mill, 2010). ASM is

heritable, dynamic and remarkably, an enrichment of ASM has been reported in regulatory regions like the ones overlapping with histone marks related to promoter and enhancer elements (Hannon et al., 2015; Gagliano et al., 2016; Benton et al., 2019). Furthermore, multiple studies have assessed the contribution of ASM to the susceptibility to specific psychiatric disorders using different approaches (Chuang et al., 2013; Hannon et al., 2015; Gagliano et al., 2016; Andrews et al., 2017; Pineda-Cirera et al., 2019; Wu and Pan, 2019), and interestingly, some of them have reported an enrichment of ASM variants in GWAS data for attention-deficit/hyperactivity disorder, autism and schizophrenia (Hannon et al., 2015; Gagliano et al., 2016; Andrews et al., 2017; Pineda-Cirera et al., 2019).

Considering the importance of ASM in specific psychiatric disorders and its impact on gene expression, the aim of the present study is to assess the contribution of ASM to the underlying genetics shared across a set of eight psychiatric disorders by using a systematic approach that combines genetic, epigenetic and expression data.

MATERIAL AND METHODS

Allele-Specific Methylation SNPs selection

The SNP selection is fully described in a previous study of our group in which we assessed the contribution of ASM to ADHD susceptibility (Pineda-Cirera et al., 2019). Briefly, we combined the information from two previous studies in *postmortem* brain samples (Gibbs et al., 2010; Zhang et al., 2010) that describe allele-specific methylation in four areas (cerebellum, frontal cortex, caudal pons and temporal cortex) and obtained 43,132 SNP-CpG pairs (Figures S1 and S2). As we observed redundancy between studies and areas (Figure S2), as well as between SNPs (in linkage disequilibrium), we performed a SNP selection process ending up with 3,896 tagSNPs that influence methylation in brain (Figures S1 and S3). The SNP selection criteria and process is described in Figure S1, further details can be found in our previous paper (Pineda-Cirera et al., 2019).

Case-control GWAS datasets

Our selection of 3,896 ASM tagSNPs was explored in the summary statistics from a GWAS meta-analysis performed by the Cross-Disorder Group of the Psychiatric Genomics Consortium that includes data of eight psychiatric disorders: attention-deficit/hyperactivity disorder (ADHD) (Demontis et al., 2019), anorexia nervosa (AN) (Duncan et al., 2017), autism spectrum disorder (ASD) (Grove et al., 2019), bipolar disorder (BIP) (Stahl et al., 2019), major depression (MD) (Wray

et al., 2018), obsessive-compulsive disorder (OCD) (Arnold et al., 2018), schizophrenia (SCZ) (Ripke et al., 2014) and Tourette syndrome (TS) (Yu et al., 2019). This GWA meta-analysis consists of 6,786,994 autosomal SNPs, 162,151 cases (11.8% ADHD, 2.2% AN, 11.3% ASD, 12.6% BIP, 36.9% MD, 1.7% OCD, 20.7% SCZ and 2.9% OCD) and 276,846 controls with European ancestry. Diagnose criteria for each of the psychiatric disorders is detailed in Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019. The publicly available summary statistics that we used do not include 23andMe data.

Statistical analysis

First, we performed an enrichment analysis using the Fisher exact test in R (R Core Team, 2014) to assess whether ASM SNPs are overrepresented among the identified risk variants in the cross-disorder GWAS meta-analysis (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). We considered p-value thresholds ranging from 5E-02 to 5E-08 and the total number of ASM SNPs present in the cross-disorder meta-analysis (31,486 out of 33,944 SNPs) (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019).

We then explored the contribution of ASM SNPs to different psychiatric disorders by following different steps (see Figure 1). First, we inspected the selection of 3,896 tagSNPs in the summary statistics of the cross-disorder GWAS meta-analysis (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). 302 of these SNPs were not present in the GWAS data, so we applied the Bonferroni correction for multiple testing for 3,594 SNPs ($p \leq 1.39E-05$, $0.05/3,594$ tagSNPs). We then discarded possible artifacts in the results by removing those CpG sites that were detected by probes that lie in genomic regions with SNPs variants. We then retrieved tagged ASM SNPs in high LD ($r^2 \geq 0.85$) with the previous ones that also correlated in *cis* with the methylation levels of the same CpG sites ($R^2 \geq 0.2$). From all those ASM SNPs, we selected only the ones that show association with at least two psychiatric disorders by using the posterior probability of association with each disorder (m -value ≥ 0.9) (Figure 1) (Han and Eskin, 2012). Finally, we selected those ASM SNPs that are eQTLs in brain for genes containing at least one of our selected CpG sites within their possible promoter region (-5Kb to +1Kb from the transcription start site) and which risk allele shows opposite effects on methylation and gene expression (Figure 1). To do that, we used GTEx data (Release V8) to assess whether the associated ASM SNPs were described to be eQTLs in any of the available brain areas (The GTEx Consortium, 2013).

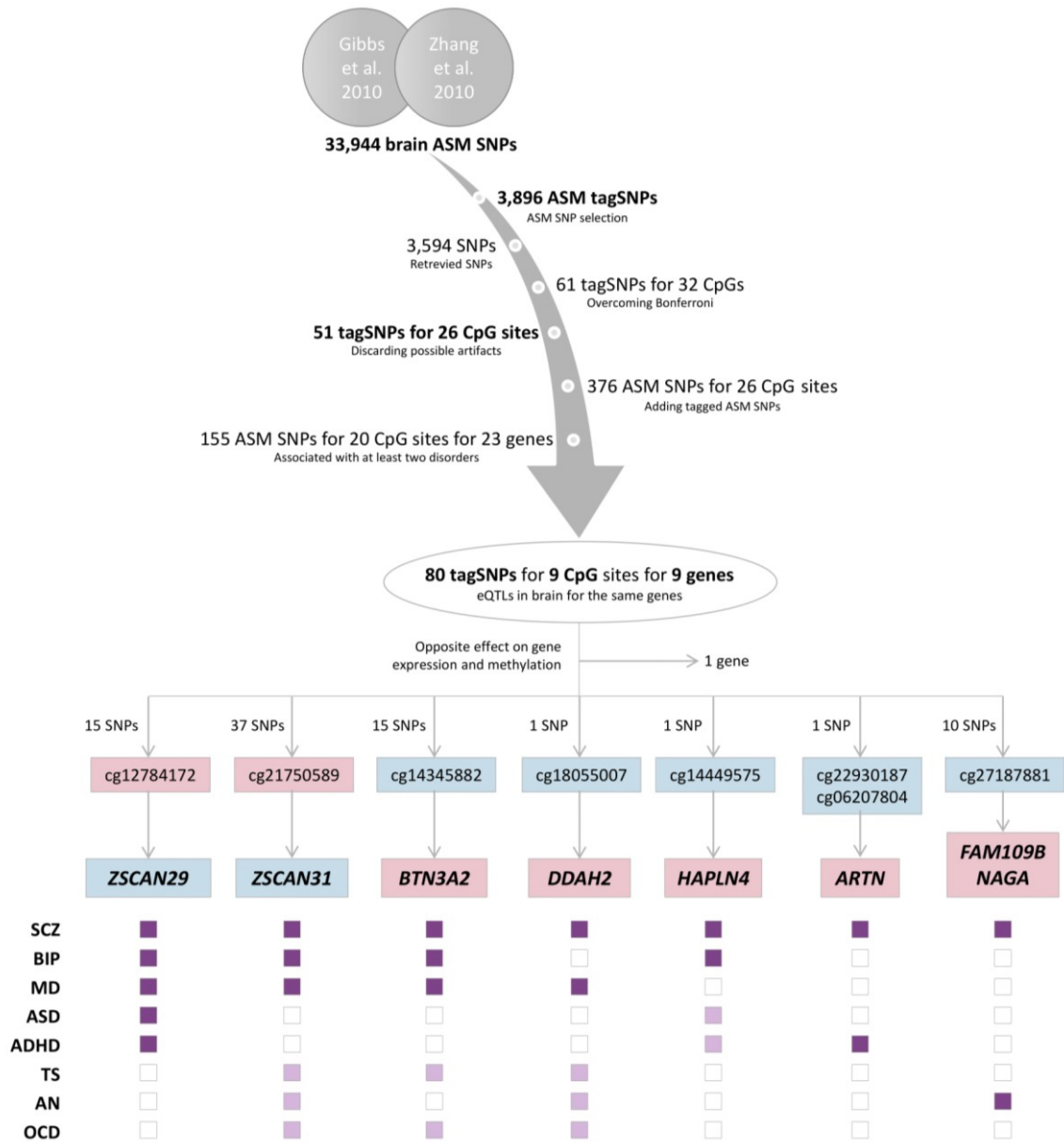


Figure 1. Outline of the association study carried out starting from a selection of 3,896 SNPs previously described to display ASM in brain. The brain areas used to assess ASM are cerebellum, caudal pons, frontal cortex and temporal cortex (Gibbs et al., 2010; Zhang et al., 2010). SNPs were tested in the summary statistics of the cross-disorder GWAS meta-analysis (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019) and Bonferroni multiple testing correction was applied ($p \leq 1.39E-05$, $0.05/3,594$ SNPs). The color of genes and CpG sites indicate the effect on methylation or gene expression of the ADHD risk allele (red: over-expression, blue: under-expression). The colored squares indicate the probability of association with each disorder considering dark purple: $m\text{-value} \geq 0.9$, light purple: $0.7 \leq m\text{-value} < 0.9$ and white: $m\text{-value} < 0.7$).

Functional annotation

We assessed functional annotations of the ASM SNPs identified using three methods. First, we inspected if the associated ASM SNPs lie in regions enriched in histone marks related to enhancer (H3K4me1 and H3K27ac) or promoter regions (H3K4me3 and H3K9ac) using the Haploreg v4.1.

tool in any of the available brain areas (Ward and Kellis, 2016). Second, we assessed if the estimated expression of the highlighted genes, modulated by the pinpointed SNPs, is altered in the phenotypes of study using MetaXcan (Barbeira et al., 2018). Prediction models were constructed considering the SNPs located within ± 1 Mb from the transcription start site (TSS) of the highlighted genes and were trained with the RNA-Seq data of 13 GTEx brain tissues (The GTEx Consortium, 2013). The SNP covariance matrices were generated using the 1000 Genomes Project Phase 3 (The 1000 Genomes Project Consortium, 2015) EUR genotypes of the SNPs included in the prediction model. Bonferroni correction for multiple testing was considered ($p \leq 9.25E-04$; $0.05/54$ tests (54 gene - brain area combinations)). Finally, we explored if the identified variants have an effect on subcortical brain volumes using the summary statistics of a GWAS meta-analysis of eight MRI volumetric measures (Hibar et al., 2015). This GWAS meta-analysis consists of seven million markers inspected in 13,171 subjects of European ancestry (Hibar et al., 2015). We applied the Bonferroni correction ($p \leq 8.93E-05$; $0.05/560$ (80 SNPs and 7 brain areas)).

RESULTS

In the present study we assessed the contribution of ASM in brain to psychiatric disorders. We first investigated whether the genetic risk variants identified in a GWAS meta-analysis of eight psychiatric disorders (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019) are enriched in variants influencing methylation by inspecting all the SNPs displaying ASM identified in two previous studies on post-mortem brains (Gibbs et al., 2010; Zhang et al., 2010). Remarkably, we observed an enrichment of ASM SNPs at all the association thresholds evaluated ($p < 2.2E-16$), together with an increase of the odds ratio (OR) when lowering the p-value threshold ($1.75 < OR < 8.90$) (Table S1).

We then inspected a selection of 3,896 tagSNPs displaying ASM in brain in the same cross-disorder GWAS meta-analysis retrieving the information of 3,594 tagSNPs (92.2%). We obtained 61 ASM tagSNPs showing association with the phenotype after applying the Bonferroni correction for multiple testing ($p \leq 1.39E-05$, $0.05/3,594$ tagSNPs) that correlate with methylation levels of 32 CpG sites in *cis*. After discarding six CpG sites which methylation was detected with probes that recognize genomic regions with SNP variants, we ended up with 51 ASM tagSNPs that correlate with methylation levels of 26 CpG sites in *cis* (Figure 1 and Table S2). Considering only tagSNP may overlook true causal SNPs, so we included in the analysis 328 additional ASM SNPs (LD, $r^2 \geq 0.85$), tagged by the initial 51 tagSNPs and correlating with the same 26 CpG sites, and retrieved the association p-values for 325 of them (99%) (Figure 1). We then selected those ASM SNPs that were

strongly associated with at least two psychiatric disorders ($m\text{-value} \geq 0.9$), obtaining 155 ASM SNPs that correlate with the methylation levels of 20 CpG sites (Figure 1).

Moreover, we explored whether those 155 ASM SNPs are eQTLs for those genes which possible promoter region contains at least one of the highlighted CpG sites. We identified 80 SNPs to be eQTLs in brain for nine genes. It is known that methylation in promoter regions inversely correlates with gene expression, and the risk alleles of all the 80 ASM SNPs were concordant with this statement for eight genes: *ZSCAN29*, *ZSCAN31*, *BTN3A2*, *DDAH2*, *HAPLN4*, *ARTN*, *FAM109B* and *NAGA*, pointing them as possible contributors to susceptibility to psychiatric disorders (Figures 1 to 3 and S4 to S8 and Table S3). Interestingly, 47 out of the 80 ASM SNPs lie in regions with histone marks related to promoter or enhancer regions (Table S4), and six out of the 80 SNPs show nominal associations with volume changes of pallidum (Table S5).

The risk alleles of the 80 ASM SNPs belong to 12 LD blocks, one of them associated with five disorders, two of them with three disorders and nine of them with two disorders (Table S3 and Figures S9 to S12). Remarkably, all the SNPs from different LD blocks that correlate with one CpG site have the same direction of the effect on methylation and gene expression. Four LD blocks correlate with increased methylation and diminished gene expression of *ZSCAN29* or *ZSCAN31* (Figures 2 and 3 and Table S3) and eight of them correlate with decreased methylation and increased gene expression of at least one of the following genes: *BTN3A2*, *DDAH2*, *HAPLN4*, *ARTN*, *FAM109B* and/or *NAGA* (Figures S4 to S8 and Table S3). All the variants identified in *ZSCAN29*, *ZSCAN31*, *BTN3A2* and *DDAH2* show association with SCZ and BIP, the two disorders showing the highest genetic correlation ($r_g = 0.70 \pm 0.02$, $p\text{-value} < 1E-06$) and several of the variants for *BTN3A2*, *ZSCAN29* and *ZSCAN31* are also associated with MD (Table S3 and Figures 1 to 3, S5 and S6). SCZ, BIP and MD are all included in the mood and psychotic disorders group as defined by factor analysis (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). Interestingly, most *ZSCAN29* variants also show associations with ASD and ADHD, conditions of the early-onset neurodevelopmental disorder group, making it the gene with broader pleiotropic effects in our study (Table S3 and Figures 1 and 2). The SNP for *HAPLN4* is associated with SCZ and MD (Table S3 and Figures 1 and S7). Finally, the SNP for *ARTN* is associated with SCZ and ADHD (Table S3 and Figures 1 and S4) and the SNPs for *FAM109B* and *NAGA* to SCZ and AN, disorders that belong to different groups (Table S3 and Figure 1 and S8). Considering that all the SNPs from each block are in high LD ($r^2 \geq 0.85$), we assessed the putative causal SNP at each LD block by selecting those variants showing more functional annotations (Table 1).

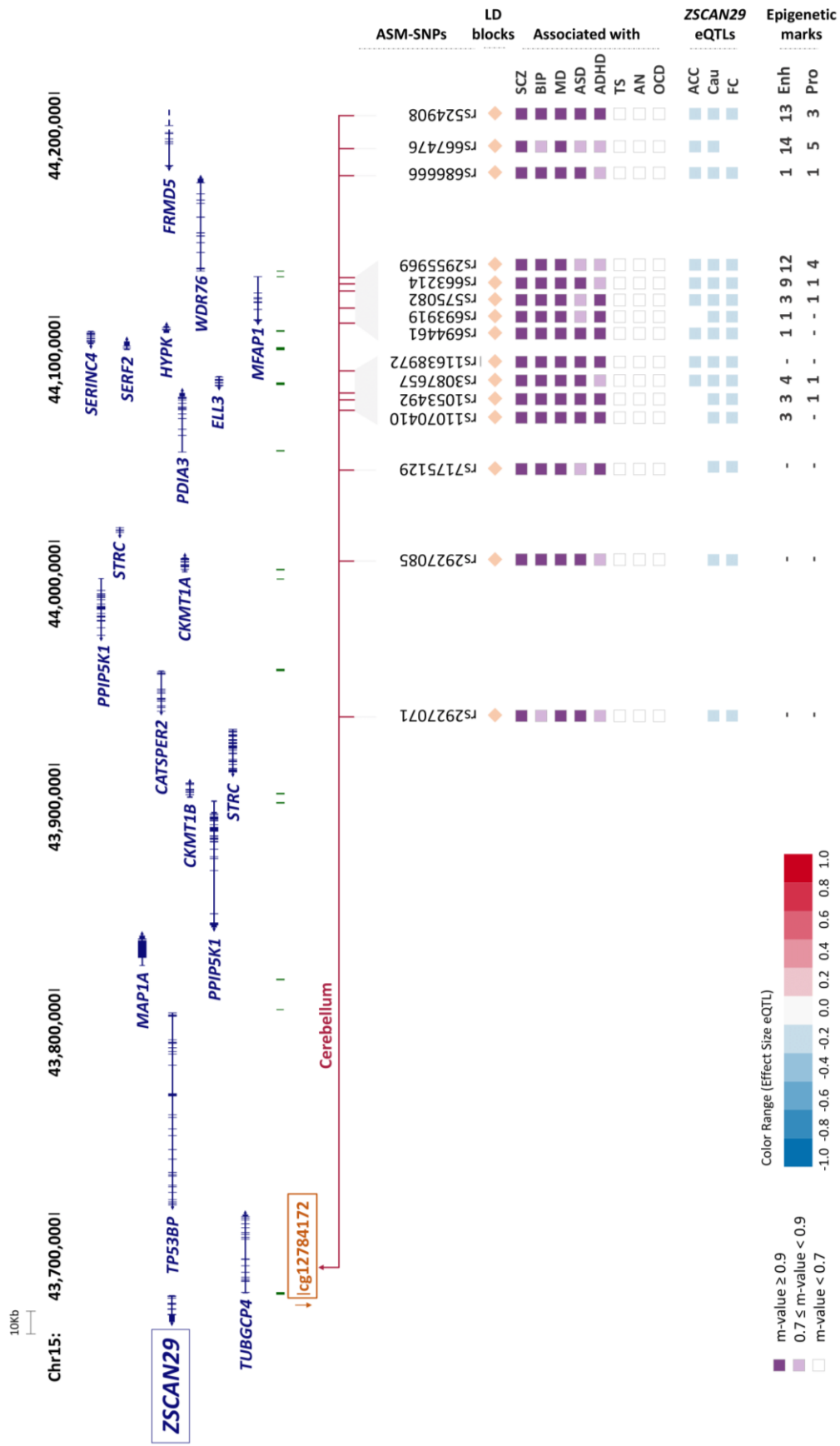


Figure 2. Genomic context of ASM variants, and methylation and eQTL information for the CpG site cg12784172. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The tagSNPs are underscored. The colored rhombuses show the LD blocks present in each region. The framed squares indicate probability of association with each disorder considering dark purple: $m\text{-value} \geq 0.9$, light purple: $0.7 \leq m\text{-value} < 0.9$ and white: $m\text{-value} < 0.7$. The colored squares for eQTLs indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). ACC: anterior cingulate cortex; Cau: caudate basal ganglia; FC: frontal cortex. The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. ‘ - ’ indicates no known enhancer or promoter histone marks. Enh: enhancer; Pro: promoter.

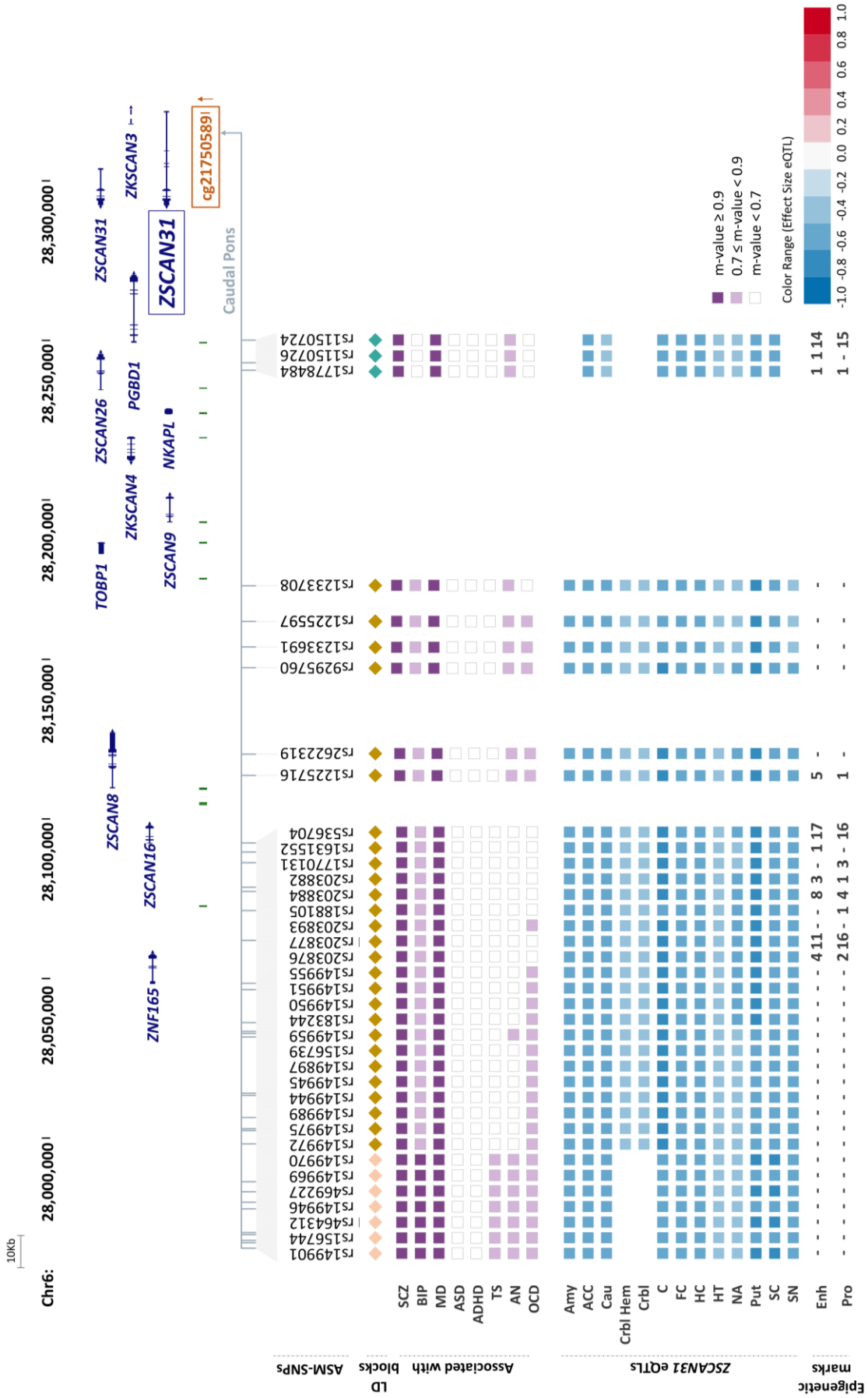


Figure 3. Genomic context of ASM variants, and methylation and eQTL information for the CpG site cg2.1750589. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The colored rhombuses show the LD blocks present in each region. The framed squares indicate probability of association with each disorder considering dark purple: $m\text{-value} \geq 0.9$, light purple: $0.7 \leq m\text{-value} < 0.9$ and white: $m\text{-value} < 0.7$. The colored squares for eQTLs indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). ACC: anterior cingulate cortex; Cau: caudate basal ganglia; FC: frontal cortex. The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. ' - ' indicates no known enhancer or promoter histone marks. Enh: enhancer; Pro: promoter.

Table 1. Selection of putative causal ASM SNP for each LD block according to functional annotations.

SNP	SNP information		¹ Association values from the cross-disorder GWAS meta-analysis										² Correlation with methylation				³ Epigenetic marks		⁴ Effect on gene expression	⁵ Effect on brain volumes			
	Chr	BP	A1	A2	OR	p-value	SCZ	BIP	MD	ASD	ADHD	TS	AN	OCD	Effect on methylation	Tissue	p-value	R ²			Enhancer	Promoter	
rs2906458*	1	44336389	A	G	0.95	8.04E-07	0.990	0.509	0.234	0.612	0.970	0.374	0.514	0.421	↓ cg22930187	crbl#	5.55E-10	0.23	6	-	-	↑ ARTN	-
rs9467704	6	26319486	T	C	0.91	3.11E-18	1	0.693	0.930	0.187	0.006	0.309	0.352	0.703	↓ cg06207804	crbl#	1.10E-12	0.29	6	-	-	-	-
rs149946	6	27970031	T	G	0.95	1.94E-11	1	0.959	1	0.593	0.626	0.716	0.806	0.789	↓ cg14345882	pons	7.44E-09	0.24	2	-	-	-	↑ Pallidum
rs536704	6	28092603	T	G	1.06	7.45E-12	1	0.778	1	0.354	0.589	0.518	0.662	0.653	↓ cg21750589	pons	1.41E-10	0.29	1	2	-	-	↑ Pallidum
rs1150724	6	28250236	T	C	0.95	1.31E-11	1	0.034	0.926	0.306	0.061	0.447	0.714	0.366	↓ cg18055007	pons	3.65E-08	0.22	8	6	-	-	↑ Pallidum
rs1043618	6	31783507	C	G	0.95	1.05E-14	1	0.33	0.986	0.057	0.016	0.759	0.702	0.827	↓ cg12784172	crbl	1.81E-08	0.26	7	16	-	-	?
rs524908	15	44187991	A	C	1.04	5.11E-06	0.997	0.939	0.983	0.91	0.908	0.530	0.523	0.673	↑ cg14449575	crbl#	2.19E-09	0.29	13	3	-	-	-
rs2905421*	19	19457908	A	C	1.05	9.21E-08	1	0.910	0	0.819	0.832	0.122	0.385	0.366	↓ cg27187881	pons	6.74E-10	0.22	4	3	-	-	-
rs2143139	22	42614401	C	G	1.06	3.72E-07	0.999	0.051	0.007	0.023	0.164	0.026	0.908	0.313	↓ cg18055007	pons	4.30E-10	0.27	16	6	-	-	?

ASM: Allele-specific methylation; ¹Data obtained from the GWA meta-analysis of eight psychiatric disorders (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019); ²Described in Zhang et al., 2010 and Gibbs et al., 2010; ³eQTL information for brain tissues for the genes with a highlighted CpG site in their possible promoter region; ⁴Histone marks found in brain areas; ⁵Data from the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) consortium (Hibar et al. 2015); SNP: Single Nucleotide Polymorphism; Chr: Chromosome; BP: Position (build hg19); A1: Allele 1; A2: Allele 2; All alleles are reported in the forward strand; OR: Odds Ratio (calculated on A1); SCZ: Schizophrenia; BIP: Bipolar Disorder; MD: Major Depression; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit Hyperactivity Disorder; TS: Tourette Syndrome; AN: Anorexia Nervosa; OCD: Obsessive-Compulsive Disorder; Effect: Direction of the risk allele effect on DNA methylation levels; *: Significant tagSNPs overcoming Bonferroni multiple testing correction, the other SNPs are ASM SNPs in LD with these significant tagSNPs; Underlined allele: Risk allele; In bold: Significant associations between ASM tagSNPs and each psychiatric disorder (m-value ≥ 0.9); ⁶Information from Zhang et al., 2010; Crbl: Cerebellum; Tctx: Temporal cortex; Fctx: Frontal cortex.

Finally, we explored whether the eight genes identified by us were predicted to be differentially expressed in cases versus controls when considering not only the ASM SNP but also all the SNPs located within ± 1 MB from the TSS of each gene, and we obtained the same effect direction for *ZSCAN29* in hypothalamus (Z-score = -3.79), for *DDAH2* in frontal cortex (Z-score = 5.87), for *HAPLN4* in cerebellum (Z-score = 4.41) and for *NAGA* in caudate basal ganglia and cerebellum (Z-score = 3.53 and 3.62 respectively) (Table S6).

DISCUSSION

In the present study we have evaluated the contribution of ASM variants to the genetic basis of a set of eight psychiatric disorders (ADHD, AN, ASD, BIP, MD, OCD, SCZ and TS). We identified 80 SNPs influencing methylation and gene expression in brain with pleiotropic effects on at least two psychiatric disorders, that highlighted eight candidate genes: *ZSCAN29*, *ZSCAN31*, *BTN3A2*, *DDAH2*, *HAPLN4*, *ARTN*, *FAM109B* and *NAGA*.

We observed an enrichment of ASM variants in the set of SNPs of the cross-disorder GWAS meta-analysis that are associated with the phenotype. These results are consistent with previous studies in which an enrichment of ASM variants was observed in a GWAS for ASD, ADHD and SCZ (Hannon et al., 2015; Gagliano et al., 2016; Andrews et al., 2017; Pineda-Cirera et al., 2019). Our results also support the omnigenic model for complex disorders that states that sub-threshold variants are needed to decipher the missing heritability of these conditions (Boyle, Li and Pritchard, 2017; Liu, Li and Pritchard, 2019). Furthermore, regions with histone marks related to promoter and enhancer regions have previously been found enriched in ASM variants (Benton et al., 2019), and almost 60% of the identified variants in our study lie in regions with histone marks, either enhancers or promoters. All these evidences suggest that ASM is a key functional mechanism that helps to explain the causality of common variation related to complex disorders, identified through GWAS analyses.

Psychiatric disorders share genetic risk factors, and a recent study identified three groups of highly genetically-related disorders based on genetic correlations and factor analysis: mood and psychotic disorders (SCZ, BIP and MD), early-onset neurodevelopmental disorders (ASD, ADHD and TS) and disorders with compulsive behaviors (AN, OCD and TS) (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). Most of the strong associations (m -value ≥ 0.9) detected in the present study belong to the mood and psychotic disorders group, and also correspond to the disorders with larger sample sizes. Interestingly, the genes *ZSCAN31*, *BTN3A2*, *DDAH2* and *HAPLN4* are also associated with psychiatric disorders classified in another group if we apply a less

strict threshold ($0.7 \leq m\text{-value} < 0.9$), mainly TS, AN and OCD. This is possibly due to the limited sample sizes of the GWAS performed in these disorders, which reduces the odds to detect significant associations. Further GWAS meta-analysis with larger sample sizes should be carried out to confirm these suggestive associations.

The 80 ASM SNPs identified as associated with the phenotype consistently show opposite effects on methylation and expression for eight out of nine genes, as expected for methylation occurring in promoter regions, which is the case. Consistently, the direction of the effect on gene expression observed for the ASM SNPs is the same as that observed when considering not only ASM SNPs for *ZSCAN29*, *DDAH2*, *HAPLN4* and *NAGA* in some of the evaluated brain areas. Moreover, all eight genes are expressed in brain (The GTEx Consortium, 2013) and overlap with CNVs related to cognitive phenotypes or psychiatric traits like autism, intellectual disability or aggressive behavior among others (Firth et al., 2009) (Table S7).

ZSCAN29 (associated with five disorders) and *ZSCAN31* (associated with three disorders) encode zinc finger transcription factor proteins that belong to the krueppel C2H2-type zinc-finger protein family. Zinc finger proteins are essential in multiple regulatory mechanisms and some have been associated with psychiatric disorders as reviewed by Squassina et al., 2019, such as *ZNF804A* with SCZ and BIP (O'Donovan et al., 2008; Steinberg et al., 2011; Sun et al., 2015) or *ZEB2*, *ZKSCAN4* and *ZNF750* associated with SCZ (Yue et al., 2011; Ripke et al., 2013; Guan et al., 2016; Khan et al., 2016). Furthermore, *ZSCAN31* is located in the region of chromosome 6 known as major histocompatibility complex (MHC, 6p21.32-p22.1), enriched in histone-related genes and immune system genes that has previously been associated with SCZ in multiple GWAS (Mokhtari and Lachman, 2016). Remarkably, zinc finger genes have also been associated with response reactions to psychotropic drugs, which points to an interesting role to be further explored in the understanding of the contribution of these proteins to psychiatric disorders (Squassina et al., 2019).

BTN3A2 is associated with three mood and psychotic disorders (SCZ, BIP and MD), it is also located in the MHC region and it is involved in the immune response (Messal et al., 2011). Interestingly, several studies have highlighted an involvement of the immune system in psychiatric disorders (Dantzer, 2018; Liberman et al., 2018; Bennett and Molofsky, 2019) because of the crosstalk that exists between immunological function and the brain (Bennett and Molofsky, 2019). Furthermore, three of the variants located near *BTN3A2* had previously been associated with cognitive phenotypes or psychiatric disorders. Specifically, the rs9467714 and rs3799378 SNPs were

associated with general cognitive function (p-value = 9.13E-11 and 2.24E-11, respectively) (Davies et al., 2018) and the rs3799380 SNP was associated with MD (p-value = 7.8E-09) (Nagel et al., 2018).

All the other highlighted genes are associated with two psychiatric disorders, schizophrenia being one of the two conditions. The results presented in this work are concordant with previous SCZ GWAS (studies that are part of the cross-disorder GWAS meta-analysis (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019)) that pointed out the *HAPLN4*, *FAM109B* and *NAGA* as candidate genes for the disorder (Ripke et al., 2013, 2014), and *DDAH2* is located in the MHC region, extensively associated with SCZ (Mokhtari and Lachman, 2016). The functions of *DDAH2* and *ARTN* are also of interest. *DDAH2* encodes an enzyme that regulates nitric oxide generation (Leiper et al., 1999). Nitric oxide has multiple well known functions in brain (Picón-Pagès, Garcia-Buendia and Muñoz, 2019) and it has been suggested to have a key role in psychotic disorders like SCZ (Nasyrova et al., 2015; Maia-de-Oliveira et al., 2016). *ARTN* encodes artemin, a protein that is necessary for sensory and sympathetic peripheral neuron survival *in vitro* and possibly also for the survival of dopaminergic neurons of the ventral mid-brain (Baloh et al., 1998). Furthermore, a knock-out mouse model for *ARTN* presents aberrations in the migration and axonal projections in the sympathetic nervous system (Honma et al., 2002). Lastly, the *ARTN* expression was altered in blood of MD patients (Otsuki et al., 2008) and was also associated with ADHD both in the largest GWAS performed on ADHD (Demontis et al., 2019) and in a work carried out by our group in which we studied the contribution of ASM to ADHD (Pineda-Cirera et al., 2019).

There are several strengths and limitations in the present study that should be discussed. Strengths: Although GWAS of psychiatric disorders have been quite successful in identifying novel associations, one of their major weaknesses has been the lack of knowledge on the functionality of the identified variants and so, the difficulty in deciphering the true causal genetic risk factors. In the present study we have overcome this major limitation by using SNP variants influencing ASM, therefore focusing our attention only on variants with a possible functional effect. For instance, 50 out of the 80 SNP variants identified, which point to five specific genes (*ZSCAN31*, *HAPLN4*, *ARTN*, *NAGA* and/or *FAM109B*) lie inside the 146 significant *loci* reported in the cross-disorder GWAS meta-analysis (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). However, only *ZSCAN31* but none of the other four was named in this GWAS (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). Other studies have also used ASM information in order to get new insights on the functional effect of SNP variants reported in GWAS

for specific psychiatric disorders (Chuang et al., 2013; Hannon et al., 2015; Gagliano et al., 2016; Andrews et al., 2017; Pineda-Cirera et al., 2019; Wu and Pan, 2019), and some of the results described are concordant with the ones described in the present study: 1) There is an enrichment of brain ASM variants among sub-threshold and genome-wide significant variants identified in GWAS for specific psychiatric disorders (Hannon et al., 2015; Gagliano et al., 2016; Andrews et al., 2017; Pineda-Cirera et al., 2019); 2) ASM SNPs are overrepresented in regulatory regions as promoter or enhancer regions (Hannon et al., 2015; Gagliano et al., 2016) and 60% of the reported variants lie in regions with histone marks related to enhancer and promoter regions; 3) The *ARTN* gene was pointed out to be associated specifically with ADHD in a study performed by us in which we used the same integrative approach (Pineda-Cirera et al., 2019); 4) Exploration of ASM SNPs has been useful in the assessment of the functional role of variants described in GWAS for psychiatric disorders (Gagliano et al., 2016; Andrews et al., 2017; Pineda-Cirera et al., 2019; Wu and Pan, 2019). Furthermore, three SNPs (rs9467704, rs9467714 and rs3799378) and multiple genes identified in the present study have previously been associated with specific cognitive phenotypes or psychiatric disorders. Finally, it is known that methylation in promoter regions inversely correlates with gene expression, and all reported variants are concordant with this statement. Limitations: We have not tested all possible ASM SNPs as the studies performed by Gibbs et al., 2010 and Zhang et al., 2010 were carried out by using genotyping platforms that do not include all the SNPs present in the genome. Furthermore, we only considered *cis* ASM variants and discarded from the study *trans* ASM variants (less common). In addition, the number of ASM SNPs described in cerebellum is higher compared to the other tissues studied and lastly, there is an imbalance of the sample sizes used in the cross-disorder GWAS meta-analysis, and this could limit the power to detect associations for the disorders with smaller samples.

To sum up, we used an integrative approach combining genetic, epigenetic and expression data that allowed us to highlight 80 SNPs influencing methylation and gene expression in brain and having pleiotropic effects on several psychiatric disorders. These SNPs point at eight genes, namely *ZSCAN29*, *ZSCAN31*, *BTN3A2*, *DDAH2*, *HAPLN4*, *ARTN*, *FAM109B* and *NAGA*. Among them, *ZSCAN29* is associated with several mood and psychotic disorders and with early-onset neurodevelopmental disorders, turning it into the gene with the broadest pleiotropic effect followed by *ZSCAN31* and *BTN3A2*, associated with mood and psychotic disorders. Further investigation should be carried out to understand the specific role of these genes in the susceptibility to psychiatric disorders.

FUNDING AND DISCLOSURE

The authors declare no conflict of interest.

This work was supported mainly by the Spanish 'Ministerio de Economía y Competitividad' [grant number SAF2015-68341-R and RTI2018-100968-B-100] and AGAUR, 'Generalitat de Catalunya' [grant number 2017-SGR-738]. BC also received funding from the European Union H2020 Program [H2020/2014-2020, grant agreement 667302 and 728018]. LP-C and JC-D were supported by 'Generalitat de Catalunya' [grant numbers 2016 FI_B 00728 and 2015 FI_B 00448, respectively]. LP-C was also supported by 'Ministerio de Educación, Cultura y Deporte' [grant number FPU15/03867]. NF-C is supported by a contract of the 'Centro de Investigación Biomédica en Red de Enfermedades Raras' (CIBERER).

ACKNOWLEDGMENTS

We are also grateful to the Cross-Disorder Group of the Psychiatric Genomics Consortium (PGC) for distributing the summary statistics of the cross-disorder GWAS meta-analysis. We also thank the ENIGMA consortium for sharing the summary statistics of genome-wide association meta-analyses of MRI phenotypes.

AUTHOR CONTRIBUTIONS

LP-C and JC-D performed all the analysis and included in this study; LP-C, JC-D, BC and NF-C designed the study; LP-C prepared the first draft of the manuscript and all figures and tables; LP-C, JC-D, BC and NF-C reviewed and edited the manuscript. BC and NF-C coordinated the study and supervised the manuscript preparation. All authors contributed to and approved the final version of the manuscript.

BIBLIOGRAPHY

- American Psychiatric Association (2013) *Diagnostic and statistical manual of mental disorders* (5th ed.), Arlington, VA: American Psychiatric Publishing.
- Andrews, S. V. et al. (2017) 'Cross-tissue integration of genetic and epigenetic data offers insight into autism spectrum disorder', *Nature Communications*, 8, p. 1011.
- Arnold, P. D. et al. (2018) 'Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis', *Molecular psychiatry*, 23(5), pp. 1181–1188. doi: 10.1038/mp.2017.154.
- Baloh, R. H. et al. (1998) 'Artemin, a novel member of the GDNF ligand family, supports peripheral and central neurons and signals through the GFR α 3-RET receptor complex', *Neuron*, 21(6), pp. 1291–1302. doi: 10.1016/S0896-6273(00)80649-2.
- Barbeira, A. N. et al. (2018) 'Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics', *Nature Communications*, 9(1), p. 1825. doi: 10.1038/s41467-018-03621-1.
- Bennett, F. C. and Molofsky, A. V. (2019) 'The immune system and psychiatric disease: a basic science perspective', *Clinical and Experimental Immunology*, 197(3), pp. 294–307. doi: 10.1111/cei.13334.
- Benton, M. C. et al. (2019) 'Genome-wide allele-specific methylation is enriched at gene regulatory regions in a multi-generation pedigree from the Norfolk Island isolate', *Epigenetics and Chromatin*, 12, p. 60. doi: 10.1186/s13072-019-0304-7.
- Boyle, E. A., Li, Y. I. and Pritchard, J. K. (2017) 'An Expanded View of Complex Traits: From Polygenic to Omnigenic', *Cell*, 169(7), pp. 1177–1186. doi: 10.1016/j.cell.2017.05.038.
- Brainstorm Consortium et al. (2018) 'Analysis of shared heritability in common disorders of the brain', *Science*, 360, p. eaap8757.
- Chuang, L.-C. et al. (2013) 'Pathway analysis using information from allele-specific gene methylation in genome-wide association studies for bipolar disorder', *PLoS ONE*, 8(1), p. e53092.
- Cross-Disorder Group of the Psychiatric Genomics Consortium (2019) 'Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders', *Cell*, 179(7), pp. 1469–1482. doi: 10.1016/j.cell.2019.11.020.
- Dantzer, R. (2018) 'Neuroimmune interactions: From the brain to the immune system and vice versa', *Physiological Reviews*, 98(1), pp. 477–504. doi: 10.1152/physrev.00039.2016.
- Davies, G. et al. (2018) 'Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function', *Nature Communications*, 9, p. 2098. doi: 10.1038/s41467-018-04362-x.
- Demontis, D. et al. (2019) 'Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder', *Nature Genetics*, 51(1), pp. 63–75.
- Duncan, L. et al. (2017) 'Significant Locus and Metabolic Genetic Correlations Revealed in Genome-Wide Association Study of Anorexia Nervosa', *The American journal of psychiatry*, 174(9), pp. 850–858. doi: 10.1176/appi.ajp.2017.16121402.
- Firth, H. V. et al. (2009) 'DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources', *American journal of human genetics*, 84(4), pp. 524–533.
- Gagliano, S. A. et al. (2016) 'Allele-Skewed DNA Modification in the Brain: Relevance to a Schizophrenia GWAS', *American journal of human genetics*, 98(5), pp. 956–962.
- Gibbs, J. R. et al. (2010) 'Abundant quantitative trait loci exist for DNA methylation and gene expression in Human Brain', *PLoS Genetics*, 6(5), p. e1000952.

- Grove, J. et al. (2019) 'Identification of common genetic risk variants for autism spectrum disorder', *Nature Genetics*, 51(3), pp. 431–444. doi: 10.1038/s41588-019-0344-8.
- Guan, L. et al. (2016) 'Common variants on 17q25 and gene-gene interactions conferring risk of schizophrenia in Han Chinese population and regulating gene expressions in human brain', *Molecular psychiatry*, 21(9), pp. 1244–1250. doi: 10.1038/mp.2015.204.
- Han, B. and Eskin, E. (2012) 'Interpreting meta-analyses of genome-wide association studies', *PLoS Genetics*, 8(3), p. e1002555. doi: 10.1371/journal.pgen.1002555.
- Hannon, E. et al. (2015) 'Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci', *Nature Neuroscience*, 19(1), pp. 48–54.
- Hibar, D. P. et al. (2015) 'Common genetic variants influence human subcortical brain structures', *Nature*, 520, pp. 224–229.
- Honma, Y. et al. (2002) 'Artemin is a vascular-derived neurotrophic factor for developing sympathetic neurons', *Neuron*, 35(2), pp. 267–282. doi: 10.1016/S0896-6273(02)00774-2.
- Khan, R. A. W. et al. (2016) 'A new risk locus in the ZEB2 gene for schizophrenia in the Han Chinese population', *Progress in neuro-psychopharmacology & biological psychiatry*, 66, pp. 97–103. doi: 10.1016/j.pnpbp.2015.12.001.
- Kundaje, A. et al. (2015) 'Integrative analysis of 111 reference human epigenomes', *Nature*, 518, pp. 317–329. doi: 10.1038/nature14248.
- Leiper, J. M. et al. (1999) 'Identification of two human dimethylarginine dimethylaminohydrolases with distinct tissue distributions and homology with microbial arginine deiminases', *The Biochemical journal*, 343, pp. 209–214.
- Li, E. and Zhang, Y. (2014) 'DNA methylation in mammals', *Cold Spring Harbor perspectives in biology*, 6, p. a019133.
- Liberman, A. C. et al. (2018) 'Neuroimmune and inflammatory signals in complex disorders of the central nervous system', *NeuroImmunoModulation*, 25(5–6), pp. 246–270.
- Liu, X., Li, Y. I. and Pritchard, J. K. (2019) 'Trans Effects on Gene Expression Can Drive Omnigenic Inheritance', *Cell*, 177, pp. 1022–1034.
- Maia-de-Oliveira, J. et al. (2016) 'Nitric Oxide's Involvement in the Spectrum of Psychotic Disorders', *Current Medicinal Chemistry*, 23, pp. 2680–2691. doi: 10.2174/0929867323666160721144549.
- Martin, E. M. and Fry, R. C. (2018) 'Environmental Influences on the Epigenome: Exposure-Associated DNA Methylation in Human Populations', *Annual Review of Public Health*, 39, pp. 309–333. doi: 10.1146/annurev-publhealth-040617-014629.
- Maurano, M. T. et al. (2012) 'Systematic localization of common disease-associated variation in regulatory DNA', *Science*, 337, pp. 1190–1195. doi: 10.1126/science.1222794.
- Meaburn, E. L., Schalkwyk, L. C. and Mill, J. (2010) 'Allele-specific methylation in the human genome: Implications for genetic studies of complex disease', *Epigenetics*, 5(7), pp. 578–582.
- Messal, N. et al. (2011) 'Differential role for CD277 as a co-regulator of the immune signal in T and NK cells', *European Journal of Immunology*, 41, pp. 3443–3454. doi: 10.1002/eji.201141404.
- Mokhtari, R. and Lachman, H. M. (2016) 'The Major Histocompatibility Complex (MHC) in Schizophrenia: A Review', *Journal of Clinical & Cellular Immunology*, 7, p. 479. doi: 10.4172/2155-9899.1000479.
- Nagel, M. et al. (2018) 'Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways', *Nature genetics*, 50(7), pp. 920–927. doi: 10.1038/s41588-018-0151-7.

- Nasyrova, R. F. et al. (2015) 'Role of nitric oxide and related molecules in schizophrenia pathogenesis: Biochemical, genetic and clinical aspects', *Frontiers in Physiology*, 6, p. 139. doi: 10.3389/fphys.2015.00139.
- O'Donovan, M. C. et al. (2008) 'Identification of loci associated with schizophrenia by genome-wide association and follow-up', *Nature Genetics*, 40(9), pp. 1053–1055. doi: 10.1038/ng.201.
- Otsuki, K. et al. (2008) 'Altered expression of neurotrophic factors in patients with major depression', *Journal of Psychiatric Research*, 42, pp. 1145–1153. doi: 10.1016/j.jpsychires.2008.01.010.
- Picón-Pagès, P., Garcia-Buendia, J. and Muñoz, F. J. (2019) 'Functions and dysfunctions of nitric oxide in brain', *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1865(8), pp. 1949–1967. doi: 10.1016/j.bbadis.2018.11.007.
- Pineda-Cirera, L. et al. (2019) 'Exploring genetic variation that influences brain methylation in attention-deficit/hyperactivity disorder', *Translational psychiatry*, 9(1), p. 242. doi: 10.1038/s41398-019-0574-7.
- R Core Team (2014) 'R: A Language and Environment for Statistical Computing'. R Foundation for Statistical Computing, Vienna, Austria. Available at: <https://www.r-project.org/>.
- Ripke, S. et al. (2013) 'Genome-wide association analysis identifies 13 new risk loci for schizophrenia', *Nature Genetics*, 45(10), pp. 1150–1159.
- Ripke, S. et al. (2014) 'Biological insights from 108 schizophrenia-associated genetic loci', *Nature*, 511, pp. 421–427.
- Smoller, J. W. et al. (2019) 'Psychiatric genetics and the structure of psychopathology', *Molecular Psychiatry*, 24(3), pp. 409–420. doi: 10.1038/s41380-017-0010-4.
- Squassina, A. et al. (2019) 'Zinc finger proteins in psychiatric disorders and response to psychotropic medications', *Psychiatric genetics*, 29(5), pp. 132–141. doi: 10.1097/YPG.0000000000000231.
- Stahl, E. A. et al. (2019) 'Genome-wide association study identifies 30 loci associated with bipolar disorder', *Nature Genetics*, 51(5), pp. 793–803. doi: 10.1038/s41588-019-0397-8.
- Steel, Z. et al. (2014) 'The global prevalence of common mental disorders: A systematic review and meta-analysis 1980-2013', *International Journal of Epidemiology*, 43(2), pp. 476–493. doi: 10.1093/ije/dyu038.
- Steinberg, S. et al. (2011) 'Expanding the range of ZNF804A variants conferring risk of psychosis', *Molecular psychiatry*, 16(1), pp. 59–66. doi: 10.1038/mp.2009.149.
- Sun, Y. et al. (2015) 'Association between variants of zinc finger genes and psychiatric disorders: systematic review and meta-analysis', *Schizophrenia research*, 162, pp. 124–37. doi: 10.1016/j.schres.2015.01.036.
- The 1000 Genomes Project Consortium (2015) 'A global reference for human genetic variation', *Nature*, 526, pp. 68–74. doi: 10.1038/nature15393.
- The GTEx Consortium (2013) 'The Genotype-Tissue Expression (GTEx) project', *Nature Genetics*, 45(6), pp. 580–585.
- Wang, X. et al. (2016) 'Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures', *eLife*, 5, p. e10557. doi: 10.7554/eLife.10557.
- Ward, L. D. and Kellis, M. (2016) 'HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease', *Nucleic Acids Research*, 44, pp. D877–D881.
- Wray, N. R. et al. (2018) 'Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression', *Nature Genetics*, 50(5), pp. 668–681. doi: 10.1038/s41588-018-0090-3.
- Wu, C. and Pan, W. (2019) 'Integration of methylation QTL and enhancer–target gene maps with schizophrenia GWAS summary results identifies novel genes', *Bioinformatics*, 35(19), pp. 3576–3583.

- Yu, D. et al. (2019) 'Interrogating the Genetic Determinants of Tourette's Syndrome and Other Tic Disorders Through Genome-Wide Association Studies', *The American journal of psychiatry*, 176(3), pp. 217–227. doi: 10.1176/appi.ajp.2018.18070857.
- Yue, W. H. et al. (2011) 'Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2', *Nature Genetics*, 43, pp. 1228–1231. doi: 10.1038/ng.979.
- Zhang, D. et al. (2010) 'Genetic Control of Individual Differences in Gene-Specific Methylation in Human Brain', *American Journal of Human Genetics*, 86(3), pp. 411–419.

Supplementary Material

Identification of genetic variants influencing methylation in brain with pleiotropic effects on psychiatric disorders

Laura Pineda-Cirera^{1,2,3,4}, Judit Cabana-Domínguez Ph. D.^{1,2,3,4}, Noèlia Fernández-Castillo Ph. D.^{1,2,3,4,&}, Bru Cormand Ph. D.^{1,2,3,4,&}

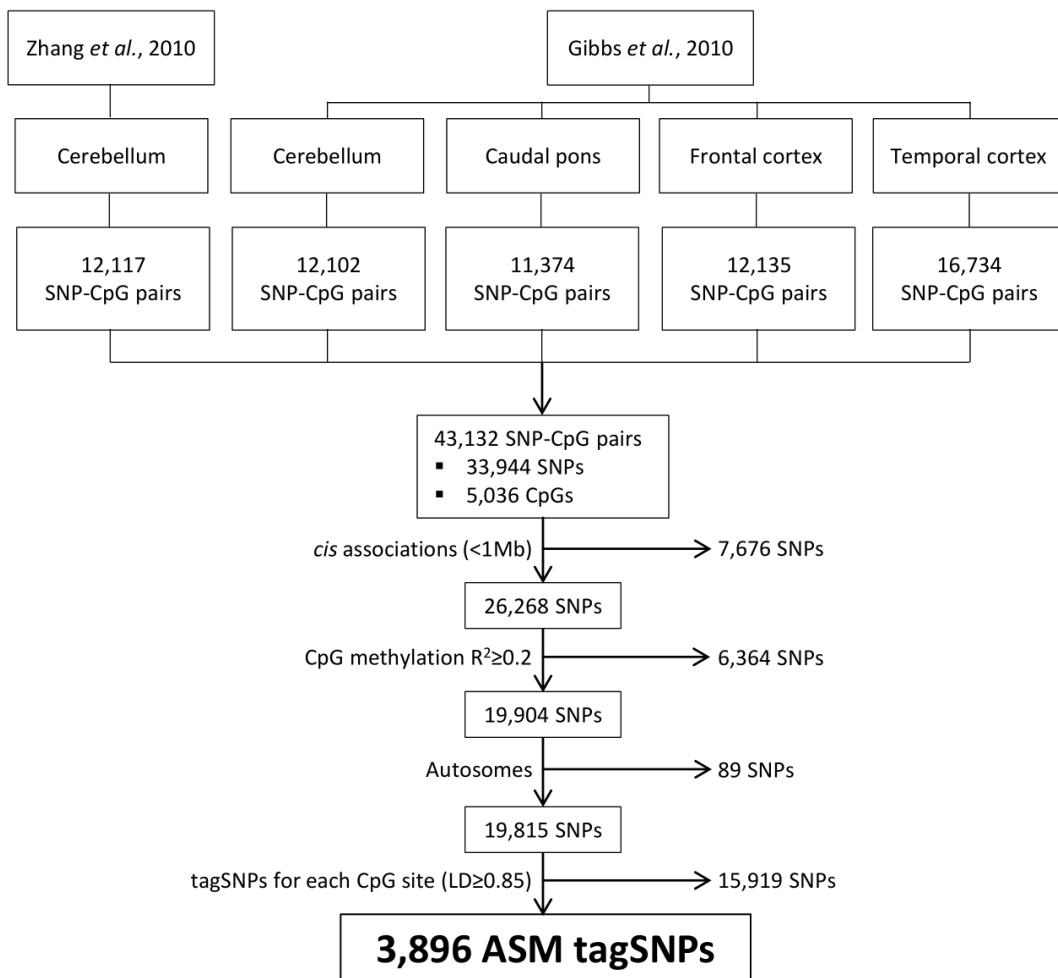


Figure S1. Selection of allele-specific methylation (ASM) SNPs. Adapted with permission from Translational Psychiatry (Pineda-Cirera et al., 2019).

Pineda-Cirera, L. et al. (2019) 'Exploring genetic variation that influences brain methylation in attention-deficit/hyperactivity disorder', *Translational psychiatry*. Springer US, 9(1), p. 242. doi: 10.1038/s41398-019-0574-7.

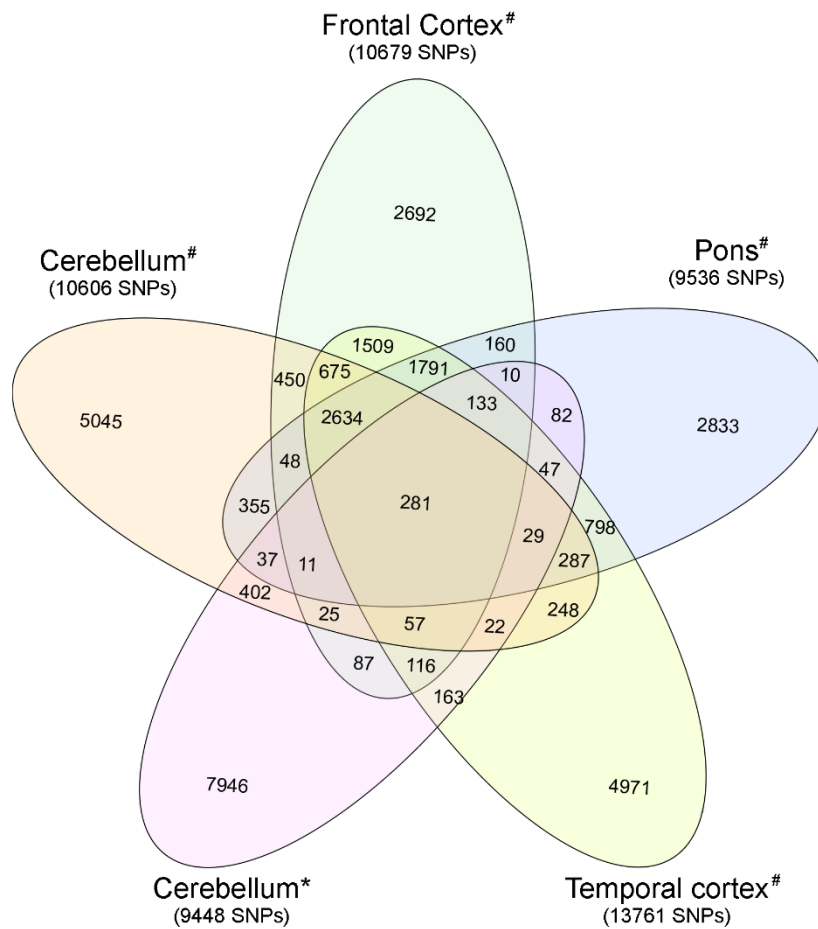


Figure S2. Venn diagram showing the overlaps among the initial 33,944 ASM SNPs selected in the different brain areas. *Cerebellum from the study by Zhang *et al.* 2010. #Brain areas from the study by Gibbs *et al.* 2010. The overlap accounts for 31% of the ASM SNPs; 12.5% ASM SNPs are shared between two tissues, 9% are shared between three tissues, 8.5% are shared between four tissues and only 1% are shared between all the tissues. Sixteen percent of the 9,448 ASM SNPs identified by Zhang *et al.*, 2010 overlap with the ASM SNPs identified in the Gibbs *et al.*, 2010 study. Reproduced with permission from Translational Psychiatry (Pineda-Cirera *et al.*, 2019).

Pineda-Cirera, L. *et al.* (2019) 'Exploring genetic variation that influences brain methylation in attention-deficit/hyperactivity disorder', *Translational psychiatry*. Springer US, 9(1), p. 242. doi: 10.1038/s41398-019-0574-7.

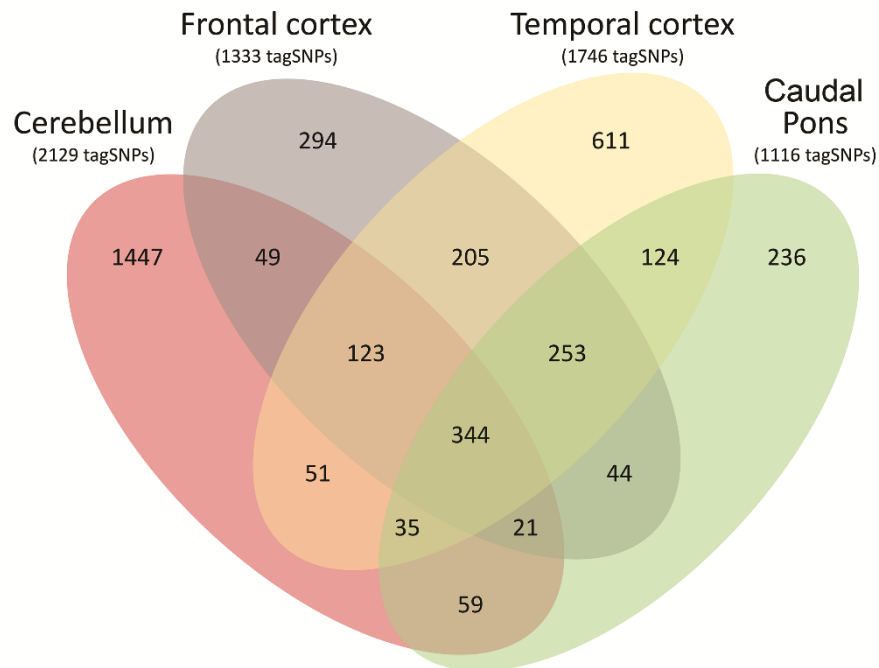


Figure S3. Venn diagram showing the overlaps among the 3,896 ASM tagSNPs selected in the different brain areas. The overlap accounts for 33.5% of tagSNPs; 13.5% tagSNPs are shared between two tissues, 11% are shared between three tissues and 9% are shared between all the tissues. Reproduced with permission from Translational Psychiatry (Pineda-Cirera et al., 2019).

Pineda-Cirera, L. et al. (2019) 'Exploring genetic variation that influences brain methylation in attention-deficit/hyperactivity disorder', *Translational psychiatry*. Springer US, 9(1), p. 242. doi: 10.1038/s41398-019-0574-7.

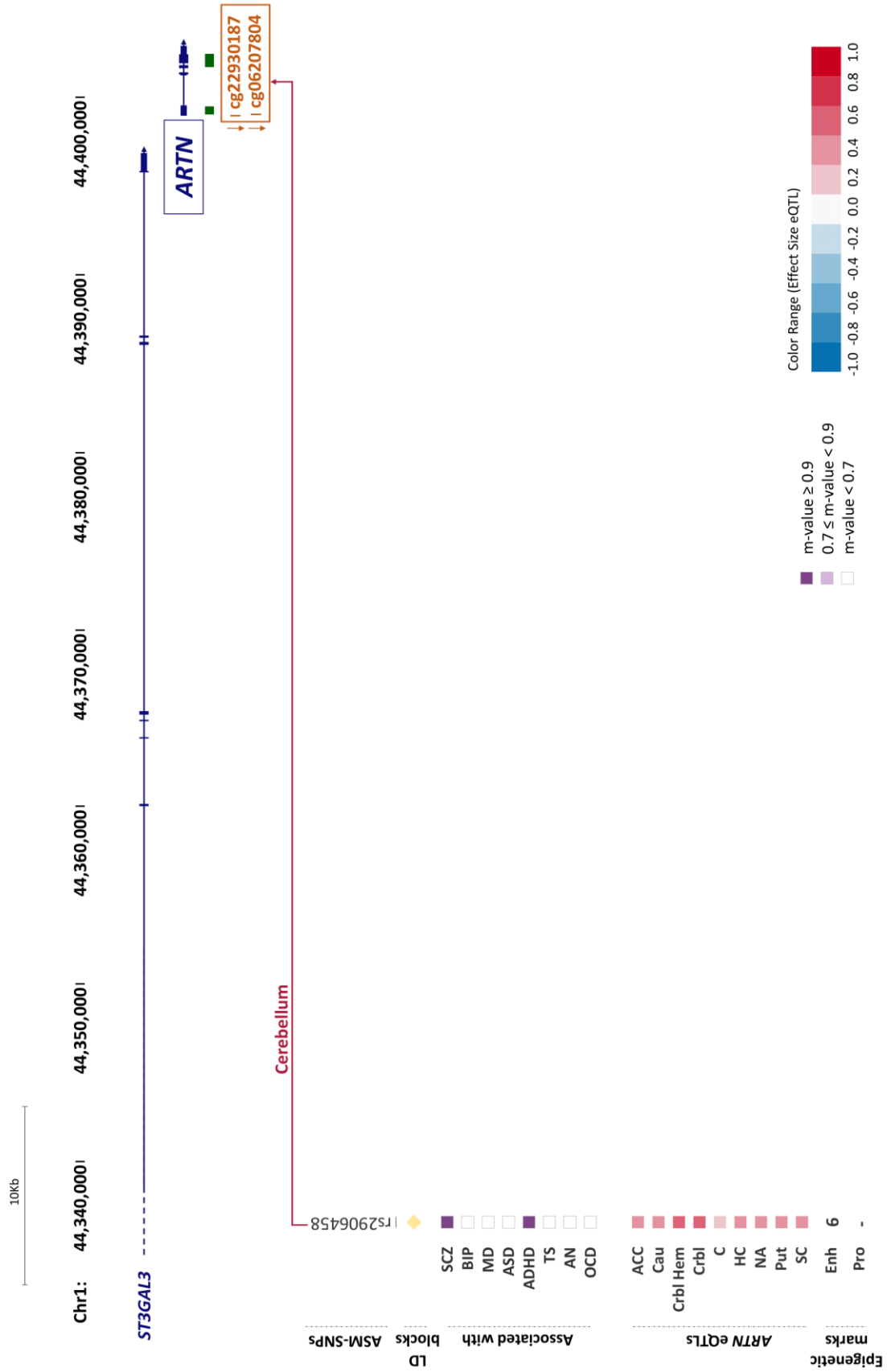


Figure S4. Genomic context of ASM variants, and methylation and eQTL information for the CpG sites cg22930187 and cg06207804. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The tagSNPs are underscored. The colored rhombuses show the LD blocks present in each region. The framed squares indicate probability of association with each disorder considering dark purple: $m\text{-value} \geq 0.9$, light purple: $0.7 \leq m\text{-value} < 0.9$ and white: $m\text{-value} < 0.7$. The colored squares for eQTLs indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). ACC: anterior cingulate cortex; Cau: caudate basal ganglia; FC: frontal cortex. The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. ' - ' indicates no known enhancer or promoter histone marks. Enh: enhancer; Pro: promoter.

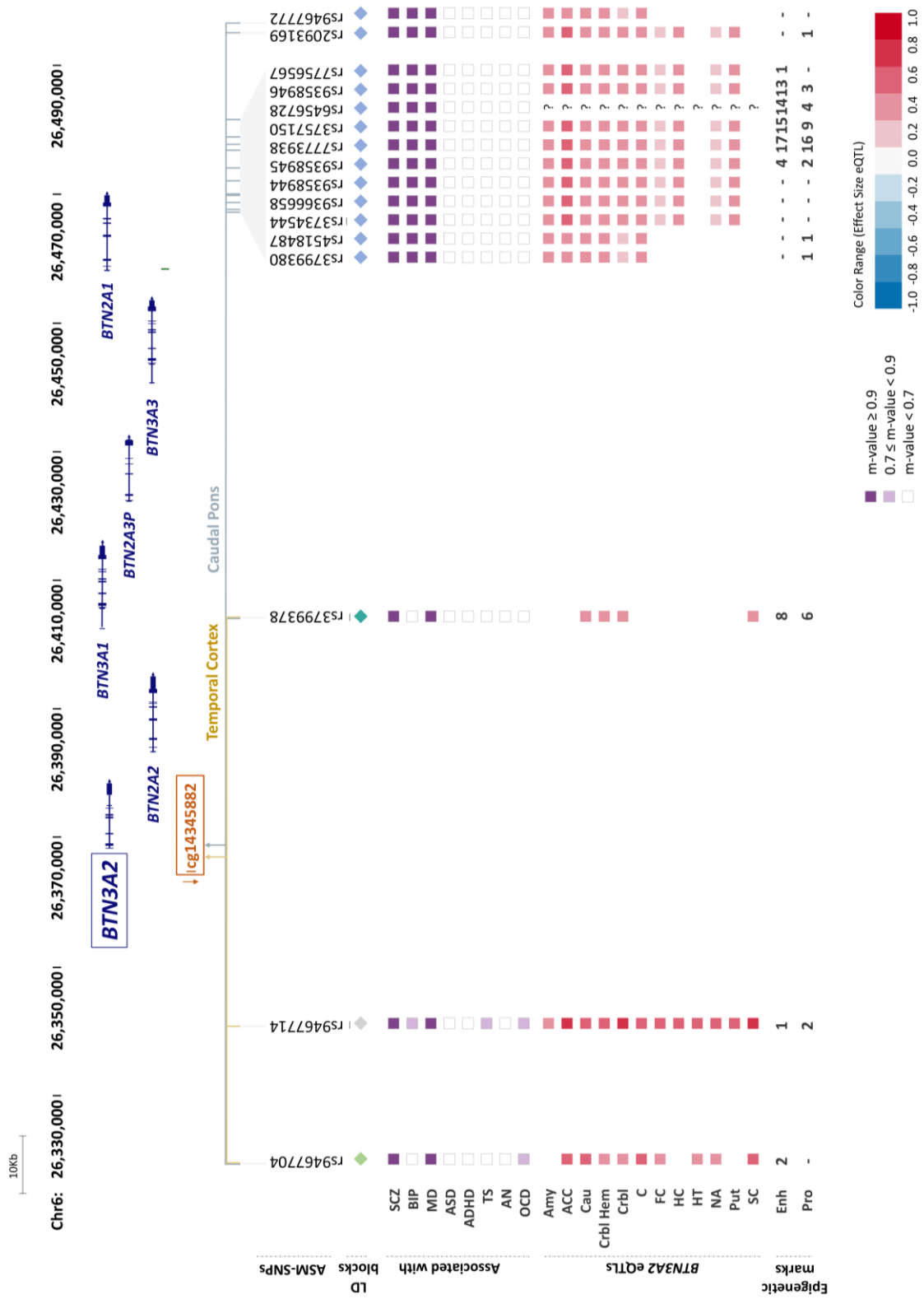


Figure S5. Genomic context of ASM variants, and methylation and eQTL information for the CpG site cg14345882. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The tagSNPs are underscored. The colored rhombuses show the LD blocks present in each region. The framed squares indicate probability of association with each disorder considering dark purple: $m\text{-value} \geq 0.9$, light purple: $0.7 \leq m\text{-value} < 0.9$ and white: $m\text{-value} < 0.7$. The colored squares for eQTLs indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). ACC: anterior cingulate cortex; Cau: caudate basal ganglia; FC: frontal cortex. The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. ' - ' indicates no known enhancer or promoter histone marks. Enh: enhancer; Pro: promoter.

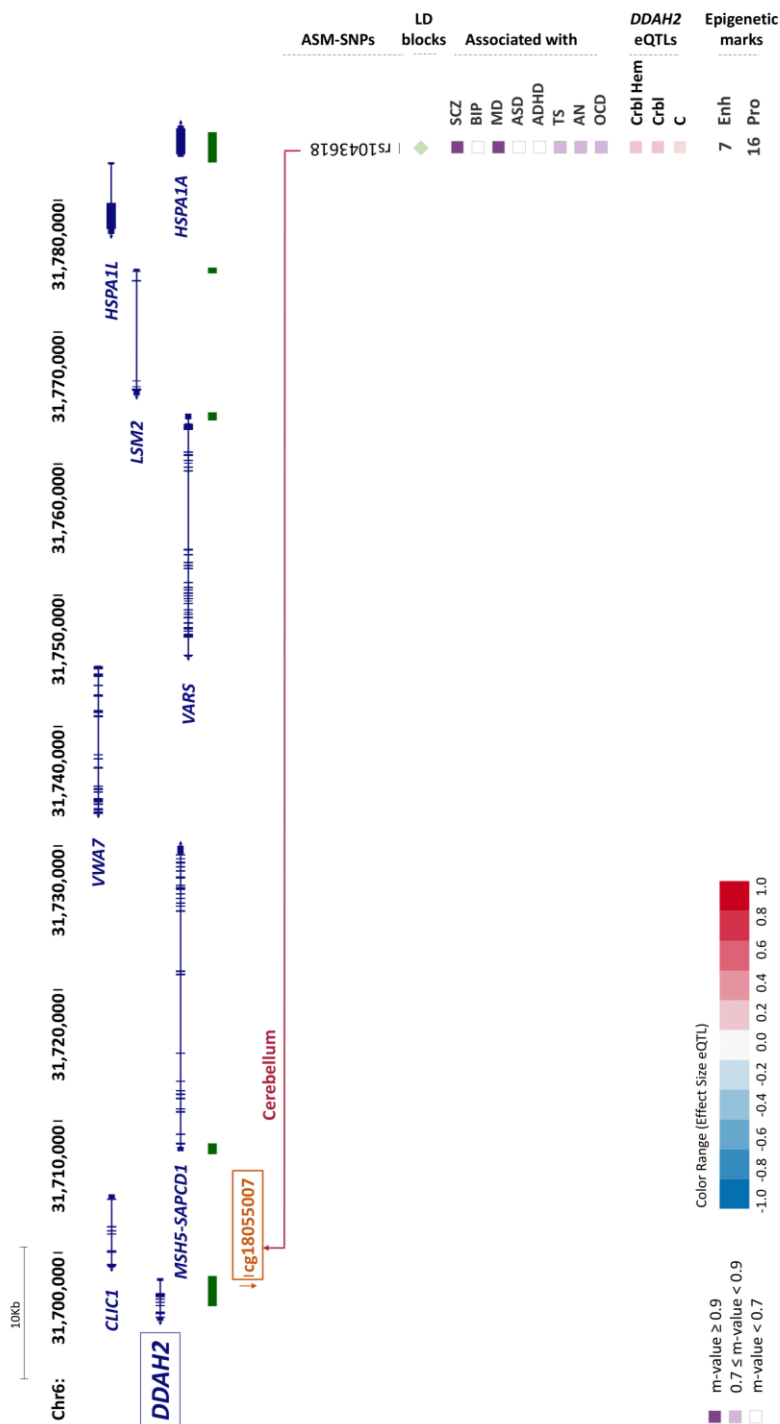


Figure S6. Genomic context of ASM variants, and methylation and eQTL information for the CpG site cg18055007. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The tagSNPs are underscored. The colored rhombuses show the LD blocks present in each region. The framed squares indicate probability of association with each disorder considering dark purple: $m\text{-value} \geq 0.9$, light purple: $0.7 \leq m\text{-value} < 0.9$ and white: $m\text{-value} < 0.7$. The colored squares for eQTLs indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). ACC: anterior cingulate cortex; Cau: caudate basal ganglia; FC: frontal cortex. The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. ' - ' indicates no known enhancer or promoter histone marks. Enh: enhancer; Pro: promoter.

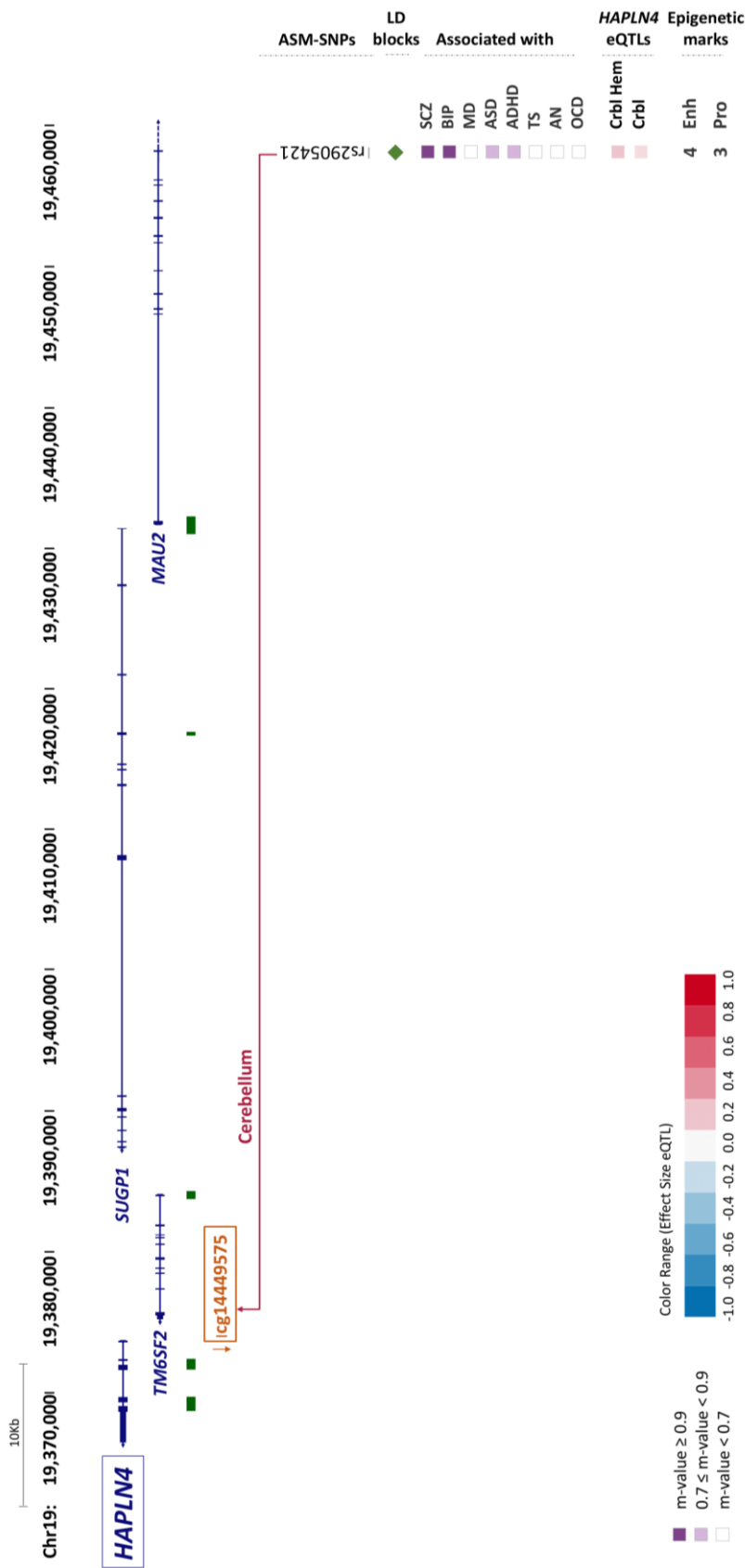


Figure S7. Genomic context of ASM variants, and methylation and eQTL information for the CpG site cg12784172. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The tagSNPs are underscored. The colored rhombuses show the LD blocks present in each region. The framed squares indicate probability of association with each disorder considering dark purple: $m\text{-value} \geq 0.9$, light purple: $0.7 \leq m\text{-value} < 0.9$ and white: $m\text{-value} < 0.7$. The colored squares indicate probability indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). ACC: anterior cingulate cortex; Cau: caudate basal ganglia; FC: frontal cortex. The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. ' - ' indicates no known enhancer or promoter histone marks. Enh: enhancer; Pro: promoter.

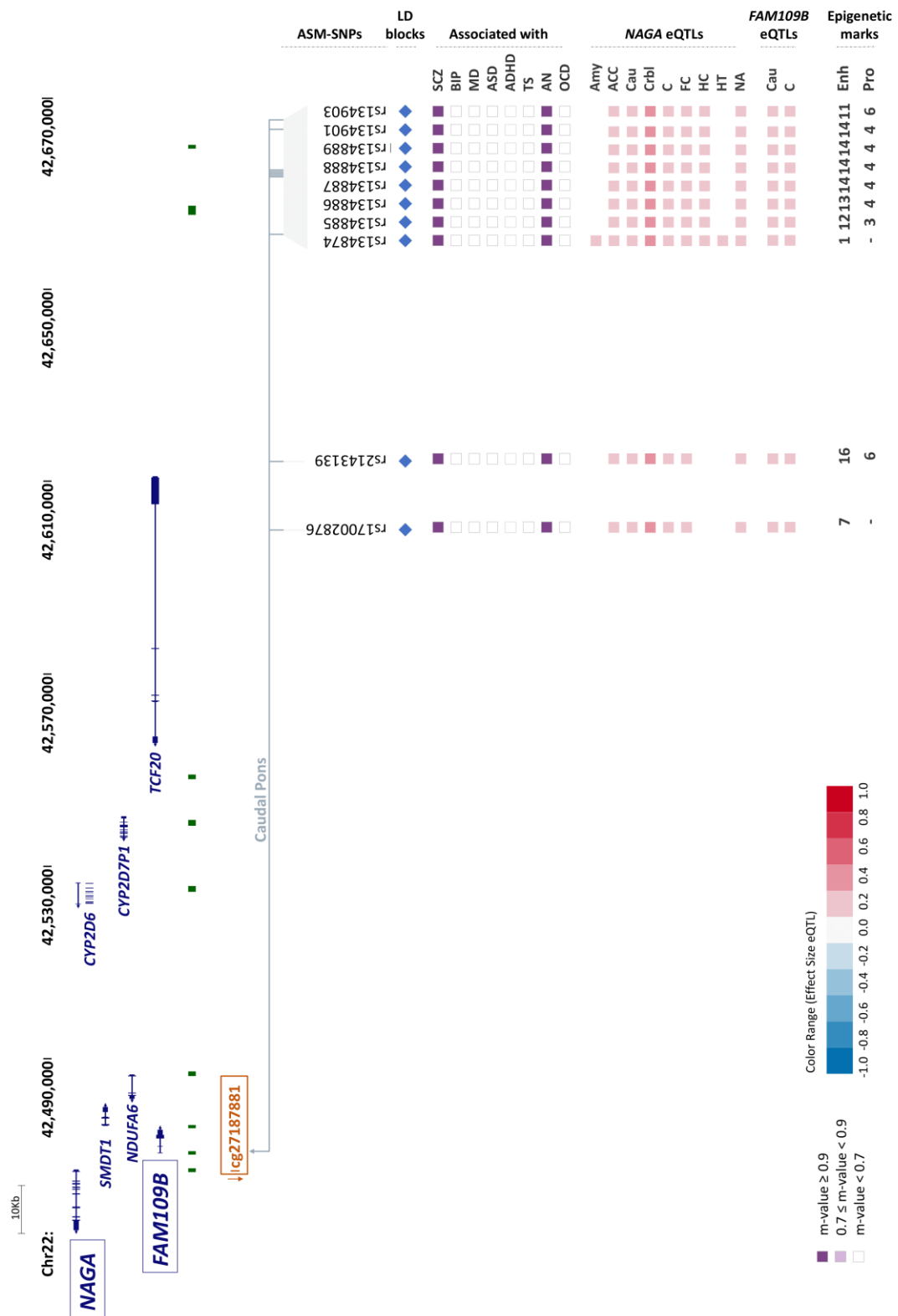


Figure S8. Genomic context of ASM variants, and methylation and eQTL information for the CpG site cg27187881. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The tagSNPs are underscored. The colored rhombuses show the LD blocks present in each region. The framed squares indicate probability of association with each disorder considering dark purple: $m\text{-value} \geq 0.9$, light purple: $0.7 \leq m\text{-value} < 0.9$ and white: $m\text{-value} < 0.7$. The colored squares for eQTLs indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). ACC: anterior cingulate cortex; Cau: caudate basal ganglia; FC: frontal cortex. The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. ‘ - ’ indicates no known enhancer or promoter histone marks. Enh: enhancer; Pro: promoter.

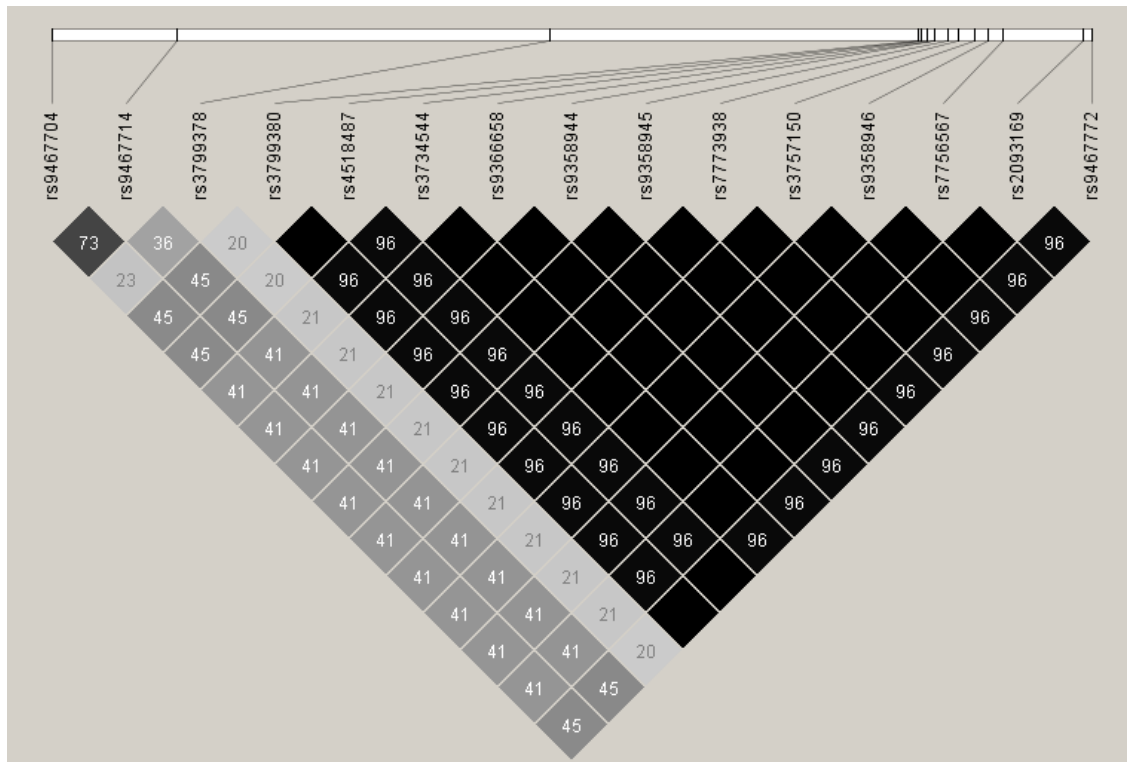


Figure S9. Linkage disequilibrium (LD, r^2) patterns between the 15 SNPs associated with schizophrenia, bipolar disorder and major depression that also correlate with differential methylation at the CpG site cg14345882, located in the possible promoter region of *BTN3A2*.

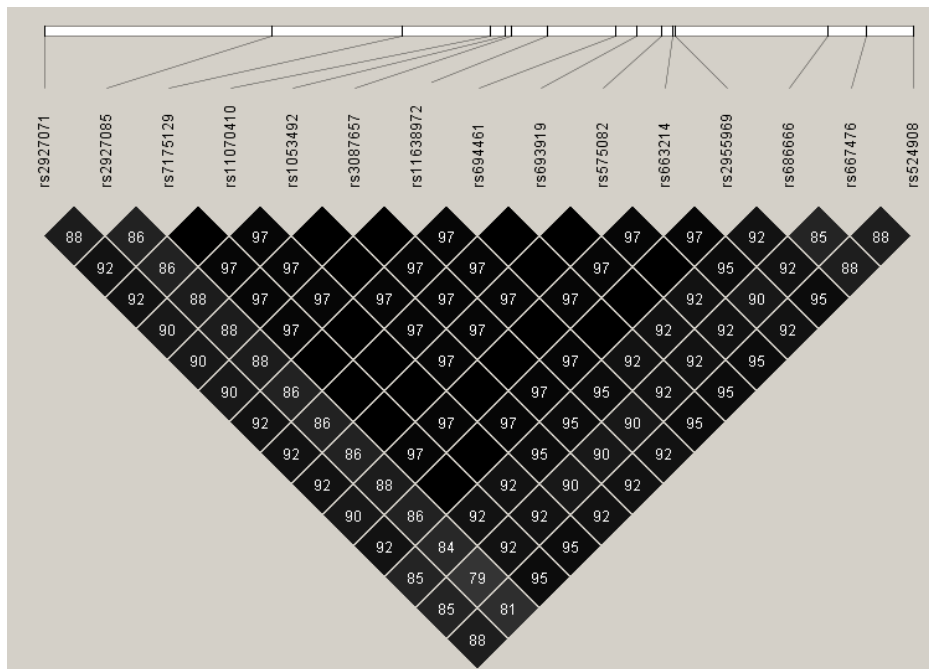


Figure S11. Linkage disequilibrium (LD, r^2) patterns between the 15 SNPs associated with schizophrenia, bipolar disorder, major depression, autism spectrum disorder and attention-deficit/hyperactivity disorder that also correlate with differential methylation at the CpG site cg12784172, located in the possible promoter region of *ZSCAN29*.

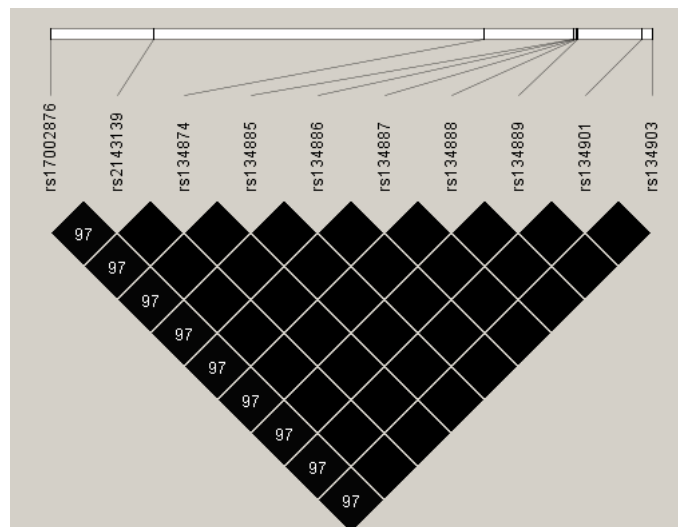


Figure S12. Linkage disequilibrium (LD, r^2) patterns between the 10 SNPs associated with schizophrenia and anorexia nervosa that also correlate with differential methylation at the CpG site cg27187881, located in the possible promoter region of *FAM109B* and *NAGA*.

Pineda-Cirera, L. et al. (2019) 'Exploring genetic variation that influences brain methylation in attention-deficit/hyperactivity disorder', *Translational psychiatry*. Springer US, 9(1), p. 242. doi: 10.1038/s41398-019-0574-7.

Table S1. Enrichment analysis of ASM SNPs at different significance thresholds in the cross-disorder GWAS meta-analysis.

Significance				
Threshold	N SNPs	N ASM SNPs	p-value	OR
5.00E-08	8751	347	<u>< 2.2E-16</u>	8.90
5.00E-07	13479	418	<u>< 2.2E-16</u>	6.91
5.00E-06	21509	501	<u>< 2.2E-16</u>	5.16
5.00E-05	39397	709	<u>< 2.2E-16</u>	3.98
5.00E-04	86357	1145	<u>< 2.2E-16</u>	2.94
5.00E-03	232676	2340	<u>< 2.2E-16</u>	2.26
5.00E-02	817392	6108	<u>< 2.2E-16</u>	1.75

ASM: Allele-specific methylation; N SNPs: Number of significant SNPs in the Cross-Disorder GWAS meta-analysis (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019) for the corresponding significance threshold; N ASM SNPs: Number of significant ASM SNPs in the GWAS meta-analysis; Underlined: Significant enrichment of ASM SNPs in the list of cross-disorder associated SNPs; OR: Odds ratio.

Table S2. Association values and methylation information for the 51 tagSNPs.

SNP information		¹ Association values from the cross-disorder GWAS meta-analysis										² Correlation with methylation						
SNP	Chr	BP	Alleles		OR	p-value	m-value				CpG site	Tissue	Effect	p-value				
			A1	A2			SCZ	BIP	MD	ASD					ADHD	TS	AN	OCD
rs2906458	1	44336389	A	G	0.95	8.04E-07	0.99	0.509	0.234	0.612	0.97	0.374	0.514	0.421	cg22930187	Crbl*	↓	5.55E-10
rs353547	3	52268866	T	C	0.95	5.94E-09	1	0.998	0.588	0.591	0.300	0.846	0.339	0.612	cg22484793	Tctx	↓	1.06E-11
rs3852064	3	52349412	T	G	0.95	2.46E-07	1	0.994	0.761	0.577	0.244	0.741	0.288	0.609	cg22484793	Tctx	↓	1.03E-08
rs10510760	3	52650348	A	G	0.96	1.94E-08	1	0.996	0.887	0.914	0.523	0.837	0.419	0.819	cg22484793	Tctx	↓	4.26E-08
rs13212936	6	25588815	T	C	0.87	4.72E-11	1	0.141	0.051	0.028	0.012	0.072	0.089	0.155	cg14345882	Pons	↓	1.32E-08
rs972087	6	25872579	A	G	1.15	9.76E-13	1	0.137	0.124	0.024	0.004	0.078	0.101	0.437	cg14345882	Pons	↓	1.54E-09
rs13191445	6	26015489	A	G	0.82	6.24E-18	1	0.035	0.008	0.001	0	0.012	0.028	0.128	cg14345882	Pons	↓	2.20E-09
rs10484439	6	26309908	A	G	0.82	6.91E-21	1	0.475	0.405	0.039	0.001	0.201	0.109	0.374	cg14345882	Pons	↓	2.77E-09
rs9467701	6	26312170	C	G	0.85	8.47E-17	1	0.022	0.029	0.006	0	0.023	0.031	0.319	cg14345882	Pons	↓	2.77E-09
rs9467714	6	26340785	A	G	0.86	4.67E-17	1	0.880	0.973	0.455	0.032	0.731	0.639	0.899	cg14345882	Tctx	↓	2.56E-08
rs6933583	6	26355283	A	C	1.06	1.74E-13	1	0.786	0.971	0.141	0.014	0.057	0.132	0.203	cg14345882	Tctx	↓	3.28E-09
rs3734536	6	26365346	C	G	0.93	2.44E-10	1	0.418	0.523	0.071	0.126	0.112	0.161	0.133	cg14345882	Pons	↓	1.15E-08
rs9688946	6	26367445	A	C	0.93	8.06E-10	1	0.413	0.509	0.072	0.123	0.110	0.160	0.132	cg14345882	Tctx	↓	2.62E-09
rs3799378	6	26404374	A	G	1.06	2.64E-10	1	0.288	0.984	0.310	0.042	0.239	0.214	0.340	cg14345882	Pons	↓	7.44E-09
rs3734544	6	26468660	A	G	0.93	1.34E-18	1	0.986	0.999	0.113	0.015	0.267	0.578	0.643	cg14345882	Pons	↓	2.35E-08
rs3734536	6	26365346	C	G	0.93	2.44E-10	1	0.418	0.523	0.071	0.126	0.112	0.161	0.133	cg01269795	Crbl*	↑	4.44E-09
rs464312	6	27967591	A	T	0.95	9.28E-13	1	0.975	1	0.619	0.694	0.753	0.862	0.817	cg21750589	Pons	↑	1.03E-08
rs203877	6	28048624	T	C	1.05	3.23E-11	1	0.794	1	0.433	0.637	0.571	0.677	0.691	cg21750589	Pons	↑	4.77E-11
rs1225715	6	28113373	A	G	1.07	3.87E-09	1	0.132	0.519	0.297	0.345	0.519	0.670	0.523	cg21750589	Pons	↑	4.56E-11
rs1150724	6	28250236	T	C	0.95	1.31E-11	1	0.034	0.926	0.306	0.061	0.447	0.714	0.366	cg21750589	Pons	↑	3.65E-08
rs213238	6	28321993	I	C	1.08	8.27E-11	1	0.083	0.425	0.197	0.221	0.317	0.641	0.445	cg21750589	Pons	↑	3.07E-09
rs7754960	6	28346945	C	G	0.91	1.40E-14	1	0.205	0.276	0.182	0.045	0.379	0.690	0.382	cg21750589	Pons	↑	4.87E-08
rs966660	6	29646165	T	C	0.94	8.26E-06	0.998	0.067	0.001	0.043	0.006	0.060	0.228	0.164	cg23244913	Pons	↑	4.98E-12

rs2747430	6	29648506	A	<u>G</u>	0.94	1.24E-05	0.998	0.067	0.001	0.041	0.006	0.059	0.225	0.163	cg23244913	Tctx	↑	1.54E-10
															cg23244913	Fctx	↑	7.85E-11
															cg04623837	Pons	↑	4.77E-11
															cg04623837	Fctx	↑	2.26E-09
															cg23244913	Pons	↑	4.98E-12
															cg23244913	Tctx	↑	1.54E-10
															cg23244913	Fctx	↑	7.85E-11
															cg04623837	Pons	↑	4.77E-11
															cg04623837	Fctx	↑	2.26E-09
rs1043618	6	31783507	C	<u>G</u>	0.95	1.05E-14	1	0.330	0.986	0.057	0.016	0.759	0.702	0.827	cg18055007	CrbI	↓	1.81E-08
rs644827	6	31838441	T	<u>C</u>	0.94	1.24E-05	0.999	0.708	0.496	0.369	0.258	0.423	0.424	0.531	cg07363637	Tctx	↑	8.39E-16
															cg07363637	Fctx	↑	5.78E-12
rs494620	6	31838713	A	G	1.06	8.03E-07	1	0.752	0.494	0.333	0.359	0.779	0.511	0.596	cg07363637	Tctx	↑	2.96E-08
rs2734335	6	31893944	A	G	1.04	1.21E-05	0.999	0.602	0.668	0.422	0.580	0.750	0.598	0.598	cg07363637	Tctx	↑	2.67E-10
rs3020644	6	31894626	A	<u>G</u>	0.94	6.32E-07	1	0.716	0.556	0.592	0.321	0.508	0.772	0.765	cg07363637	Tctx	↑	1.85E-08
rs2072633	6	31919578	A	<u>G</u>	0.95	3.30E-09	1	0.963	0.504	0.258	0.201	0.950	0.845	0.785	cg07363637	Fctx	↑	2.12E-08
rs2071295	6	32038700	I	C	1.06	1.66E-06	0.998	0.513	0.494	0.352	0.259	0.578	0.805	0.854	cg03543593	Pons	↑	7.65E-10
rs174383	6	111914700	A	<u>G</u>	0.96	2.55E-06	0.999	0.939	0.133	0.210	0.083	0.476	0.769	0.823	cg26904406	CrbI	↓	6.12E-09
rs2898674	8	38110215	I	C	1.05	5.93E-09	0.999	0.999	0.020	0.916	0.631	0.054	0.616	0.833	cg25510610	Pons	↓	3.90E-08
															cg25510610	CrbI*	↓	7.84E-12
rs11777067	8	38298647	T	<u>C</u>	0.95	8.78E-09	1	0.999	0.045	0.920	0.784	0.068	0.735	0.729	cg25510610	CrbI*	↓	5.88E-13
rs7096169	10	104618695	A	G	1.09	1.12E-16	1	0.978	0	0.001	0.010	0.020	0.032	0.007	cg00035347	Tctx	↑	3.36E-10
rs1046778	10	104661484	I	C	1.08	9.07E-10	1	0.417	0.001	0	0.017	0.028	0.052	0.026	cg00035347	Tctx	↑	6.34E-13
rs3977751	10	104920232	A	G	1.09	4.78E-13	1	0.249	0	0	0.011	0.007	0.020	0.027	cg00035347	Tctx	↑	2.30E-18
															cg00035347	Pons	↑	9.97E-10
															cg00035347	Fctx	↑	3.57E-09
rs4131364	11	812188	A	G	1.07	1.21E-05	0.026	0.040	0.039	0.138	0.889	0.586	0.211	0.101	cg20225915	Pons	↓	2.49E-10
rs11570190	11	57560452	A	<u>C</u>	0.97	2.70E-10	1	0.332	1	0.958	0.974	0.784	0.724	0.921	cg06200697	CrbI*	↓	2.13E-09
rs320109	11	63680045	A	C	1.05	2.29E-07	0.996	0.998	0.456	0.092	0.047	0.704	0.113	0.241	cg06204948	CrbI*	?	4.40E-09
rs11070410	15	44057045	A	<u>G</u>	0.97	2.43E-06	0.998	0.932	0.981	0.907	0.906	0.476	0.586	0.554	cg12784172	CrbI	↑	1.46E-10
															cg12784172	CrbI*	↑	2.02E-18
rs11638972	15	44074691	C	<u>G</u>	0.96	5.86E-07	0.998	0.923	0.993	0.922	0.901	0.454	0.570	0.516	cg12784172	CrbI	↑	1.46E-10
rs11070410	15	44057045	A	<u>G</u>	0.97	2.43E-06	0.998	0.932	0.981	0.907	0.906	0.476	0.586	0.554	cg03679734	CrbI*	↑	2.57E-09

rs4788209	16	30013266	T	C	0.95	9.45E-11	1	0.992	0.543	0.910	0.665	0.587	0.682	0.951	cg00410921	Crbl*	↑	6.63E-11
rs11152003	18	52476469	A	G	1.04	1.12E-05	0.812	0.382	0.997	0.035	0.146	0.159	0.518	0.269	cg24235633	Crbl	↓	2.99E-08
rs2905421	19	19457908	A	C	1.05	9.21E-08	1	0.910	0	0.819	0.832	0.122	0.385	0.366	cg22105582	Fctx	↓	1.87E-20
rs2905421	19	19457908	A	C	1.05	9.21E-08	1	0.910	0	0.819	0.832	0.122	0.385	0.366	cg22105582	Tctx	↓	3.81E-14
rs4808964	19	19603692	A	C	1.04	1.54E-07	1	0.767	0	0.575	0.761	0.069	0.227	0.479	cg14449575	Crbl*	↓	6.74E-10
rs4820378	22	39869209	I	C	1.06	7.14E-06	0.996	0.015	0.002	0.009	0.065	0.120	0.720	0.720	cg10669058	Crbl*	↑	4.67E-09
rs9306336	22	39873470	A	I	0.94	1.02E-06	0.999	0.005	0	0.001	0.003	0.071	0.067	0.775	cg10669058	Crbl*	↑	1.95E-11
rs2899319	22	39874314	A	C	0.94	9.50E-06	0.995	0.082	0.002	0.006	0.011	0.097	0.136	0.751	cg03972838	Pons	↑	3.73E-08
rs1033415	22	39888774	A	G	0.94	4.74E-07	0.999	0.247	0.009	0.008	0.017	0.032	0.071	0.902	cg03972838	Pons	↑	2.88E-09
rs133370	22	42465260	T	C	0.94	2.73E-06	0.998	0.218	0.013	0.075	0.139	0.032	0.817	0.277	cg03972838	Pons	↑	3.13E-15
rs134889	22	42674350	A	I	0.94	2.54E-07	0.999	0.047	0.007	0.024	0.185	0.022	0.927	0.279	cg27187881	Pons	↓	2.36E-15
															cg27187881	Tctx	↓	1.21E-10
															cg27187881	Pons	↓	2.82E-09
															cg27187881	Pons	↓	4.30E-10

ASM: Allele-specific methylation; ¹Data obtained from the cross disorder GWAS meta-analysis (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019); ²Described in Zhang et al., 2010 and Gibbs et al., 2010; SNP: Single Nucleotide Polymorphism; Chr: Chromosome; BP: Position (build hg19); A1: Allele 1; A2: Allele 2; All alleles are reported in the forward strand; OR: Odds Ratio (calculated on A1); SCZ: Schizophrenia; BIP: Bipolar Disorder; MD: Major Depression; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit Hyperactivity Disorder; Tourette Syndrome; AN: Anorexia Nervosa; OCD: Obsessive-Compulsive Disorder; Effect: Direction of the risk allele effect on DNA methylation levels; Underlined allele: Risk allele for Cross-Disorder; In bold: Significant associations for the association between ASM tagSNPs and each psychiatric disorder (m-value ≥ 0.9); Crbl*: Cerebellum from Zhang et al., 2010; Crbl: Cerebellum; Tctx: Temporal cortex; Fctx: Frontal cortex; "?": data not available.

Table S3. Association values, methylation information and functional annotations for the 80 highlighted ASM SNPs.

SNP information			m-value										Correlation with methylation			Effect on expression			Epigenetic marks					
SNP	Chr	BP	Alleles		OR	p-value	SCZ	BIP	MD	ASD	ADHD	TS	AN	OCD	Tissue	CpG site	Effect	p-value	R ²	R ²	3 Effect on expression	Enh	Pro	5 Effect on brain volumes
			A1	A2																				
rs2906458*	1	44336389	A	G	0.95	8.04E-07	0.99	0.509	0.234	0.612	0.970	0.374	0.514	0.421	crbl#	cg06207804	↓	5.55E-10	0.23	0.23	↑ ARTN	6		
rs9467704	6	26319486	T	C	0.91	3.11E-18	1	0.693	0.930	0.187	0.006	0.309	0.352	0.703	pons	cg14345882	↓	2.77E-09	0.25	0.25		2		↑ Pallidum
rs9467714*	6	26340785	A	G	0.86	4.67E-17	1	0.880	0.973	0.455	0.032	0.731	0.639	0.899	tctx	cg14345882	↓	7.62E-09	0.24	0.24		1		↑ Pallidum
rs3799378*	6	26404374	A	G	1.06	2.64E-10	1	0.288	0.984	0.310	0.042	0.239	0.214	0.340	pons	cg14345882	↓	2.56E-08	0.22	0.22		1		↑ Pallidum
rs3799380	6	26467182	T	C	1.07	4.28E-19	1	0.977	0.999	0.137	0.024	0.267	0.606	0.645	tctx	cg14345882	↓	7.44E-09	0.24	0.24		8	6	↑ Pallidum
rs4518487	6	26467531	A	G	1.07	5.00E-19	1	0.977	0.999	0.130	0.024	0.264	0.614	0.650	pons	cg14345882	↓	2.35E-08	0.22	0.22		1		
rs3734544*	6	26468660	A	G	0.93	1.34E-18	1	0.986	0.999	0.113	0.015	0.267	0.578	0.643	pons	cg14345882	↓	2.35E-08	0.22	0.22		1		
rs9366658	6	26469866	A	G	0.94	6.17E-18	1	0.986	0.999	0.119	0.016	0.265	0.579	0.643	pons	cg14345882	↓	2.35E-08	0.22	0.22		4	2	
rs9358944	6	26469875	A	C	1.07	5.29E-18	1	0.986	0.999	0.118	0.016	0.268	0.579	0.643	pons	cg14345882	↓	2.35E-08	0.22	0.22		17	16	
rs9358945	6	26472114	A	G	1.07	6.61E-18	1	0.983	0.999	0.124	0.017	0.266	0.581	0.643	pons	cg14345882	↓	2.35E-08	0.22	0.22		15	9	?
rs7773938	6	26474044	T	C	0.94	7.26E-18	1	0.985	0.999	0.119	0.016	0.264	0.58	0.642	pons	cg14345882	↓	2.35E-08	0.22	0.22		13	3	
rs3757150	6	26476782	A	I	0.93	9.06E-19	1	0.984	0.999	0.121	0.015	0.261	0.579	0.640	pons	cg14345882	↓	2.35E-08	0.22	0.22		1		
rs9358946	6	26478927	A	G	0.94	9.71E-18	1	0.985	0.999	0.119	0.016	0.262	0.569	0.645	pons	cg14345882	↓	2.35E-08	0.22	0.22		4	2	
rs7756567	6	26481642	I	G	1.08	9.03E-19	1	0.985	0.999	0.116	0.015	0.265	0.578	0.652	pons	cg14345882	↓	2.35E-08	0.22	0.22		1		
rs2093169	6	26495099	T	C	0.93	1.82E-18	1	0.987	0.999	0.122	0.017	0.279	0.594	0.633	pons	cg14345882	↓	2.35E-08	0.22	0.22		1		
rs9467772	6	26496578	A	T	1.07	3.50E-18	1	0.978	0.999	0.133	0.026	0.257	0.602	0.653	pons	cg14345882	↓	2.35E-08	0.22	0.22		1		
rs149901	6	27965503	T	C	0.95	2.83E-12	1	0.975	1	0.573	0.659	0.740	0.852	0.813	pons	cg21750589	↑	1.03E-08	0.23	0.23				↑ Pallidum
rs156744	6	27967274	A	G	0.95	2.10E-11	1	0.959	1	0.589	0.625	0.714	0.806	0.776	pons	cg21750589	↑	1.03E-08	0.23	0.23				?
rs464312*	6	27967591	A	I	0.95	9.28E-13	1	0.975	1	0.619	0.694	0.753	0.862	0.817	pons	cg21750589	↑	1.03E-08	0.23	0.23				↑ Pallidum
rs149946	6	27970031	T	G	0.95	1.94E-11	1	0.959	1	0.593	0.626	0.716	0.806	0.789	pons	cg21750589	↑	1.03E-08	0.23	0.23				
rs469227	6	27970705	A	C	1.06	9.04E-13	1	0.975	1	0.612	0.691	0.731	0.859	0.828	pons	cg21750589	↑	1.03E-08	0.23	0.23				
rs149969	6	27977737	I	C	1.05	1.25E-11	1	0.959	1	0.570	0.612	0.708	0.801	0.778	pons	cg21750589	↑	1.03E-08	0.23	0.23				↑ Pallidum
rs149970	6	27980220	A	G	0.95	1.71E-12	1	0.976	1	0.608	0.685	0.748	0.859	0.820	pons	cg21750589	↑	1.03E-08	0.23	0.23				
rs149972	6	27983227	T	C	0.95	8.56E-11	1	0.847	1	0.362	0.633	0.570	0.669	0.718	pons	cg21750589	↑	3.01E-10	0.28	0.28				
rs149975	6	27986340	A	G	0.95	1.01E-10	1	0.843	1	0.361	0.631	0.569	0.661	0.717	pons	cg21750589	↑	3.01E-10	0.28	0.28				
rs149989	6	27998184	A	I	0.95	8.99E-11	1	0.849	1	0.356	0.632	0.567	0.657	0.714	pons	cg21750589	↑	3.01E-10	0.28	0.28				?
rs149944	6	28002818	T	C	0.95	1.00E-10	1	0.847	1	0.359	0.633	0.568	0.660	0.717	pons	cg21750589	↑	9.41E-11	0.29	0.29				
rs149945	6	28002853	T	C	1.05	9.14E-11	1	0.842	1	0.357	0.631	0.566	0.659	0.716	pons	cg21750589	↑	9.41E-11	0.29	0.29				
rs149897	6	28006650	A	G	1.05	9.19E-11	1	0.852	1	0.353	0.635	0.555	0.655	0.713	pons	cg21750589	↑	8.06E-09	0.24	0.24				
rs156739	6	28013410	A	C	0.95	3.25E-11	1	0.835	1	0.355	0.628	0.558	0.653	0.718	pons	cg21750589	↑	9.41E-11	0.29	0.29				
rs149959	6	28013854	I	C	1.05	7.32E-11	1	0.881	1	0.394	0.663	0.581	0.716	0.739	pons	cg21750589	↑	9.41E-11	0.29	0.29				
rs183244	6	28031838	A	C	0.95	2.26E-11	1	0.86	1	0.439	0.666	0.570	0.687	0.737	pons	cg21750589	↑	1.20E-10	0.29	0.29				
rs149950	6	28033039	I	G	1.05	7.39E-11	1	0.861	1	0.434	0.666	0.566	0.686	0.736	pons	cg21750589	↑	9.00E-11	0.29	0.29				
rs149951	6	28033087	I	C	1.05	8.24E-11	1	0.853	1	0.434	0.665	0.568	0.687	0.737	pons	cg21750589	↑	9.00E-11	0.29	0.29				

rs149955	6	28036225	C	G	0.95	9.29E-11	1	0.858	1	0.455	0.659	0.577	0.691	0.731	poins	cg21750589	↑	9.00E-11	0.29	↓ ZSCAN31	?
rs203876	6	28046673	I	C	1.05	3.84E-11	1	0.788	1	0.426	0.640	0.573	0.671	0.697	poins	cg21750589	↑	1.41E-10	0.29	4	2
rs203877*	6	28048624	I	C	1.05	3.23E-11	1	0.794	1	0.433	0.637	0.571	0.677	0.691	poins	cg21750589	↑	4.77E-11	0.31	11	16
rs203893	6	28062066	A	C	0.95	5.59E-11	1	0.809	1	0.451	0.650	0.591	0.676	0.702	poins	cg21750589	↑	1.41E-10	0.29	1	1
rs188105	6	28071393	T	C	0.95	3.64E-11	1	0.806	1	0.435	0.652	0.556	0.683	0.685	poins	cg21750589	↑	1.41E-10	0.29	8	4
rs203884	6	28077374	A	C	0.95	5.49E-11	1	0.792	1	0.448	0.660	0.590	0.687	0.695	poins	cg21750589	↑	1.41E-10	0.29	3	1
rs203882	6	28078502	A	G	0.95	6.82E-11	1	0.811	1	0.443	0.659	0.564	0.689	0.676	poins	cg21750589	↑	1.41E-10	0.29	3	1
rs1770131	6	28086413	T	C	0.95	1.56E-10	1	0.815	1	0.449	0.664	0.602	0.692	0.697	poins	cg21750589	↑	1.41E-10	0.29	1	3
rs1631552	6	28089699	A	C	0.95	9.56E-11	1	0.797	1	0.430	0.652	0.584	0.684	0.692	poins	cg21750589	↑	1.41E-10	0.29	17	16
rs536704	6	28092603	I	G	1.06	7.45E-12	1	0.778	1	0.354	0.589	0.518	0.662	0.653	poins	cg21750589	↑	1.41E-10	0.29	5	1
rs1225716	6	28113746	A	C	1.05	2.38E-10	1	0.803	1	0.485	0.634	0.607	0.741	0.706	poins	cg21750589	↑	1.41E-10	0.29	?	
rs2622319	6	28120401	C	G	0.95	2.13E-10	1	0.799	1	0.479	0.633	0.608	0.744	0.705	poins	cg21750589	↑	1.41E-10	0.29		
rs9295760	6	28147385	A	G	0.95	2.47E-10	1	0.802	1	0.481	0.634	0.614	0.750	0.705	poins	cg21750589	↑	1.41E-10	0.29		
rs1233691	6	28153897	I	G	1.05	2.74E-11	1	0.783	1	0.515	0.669	0.636	0.811	0.715	poins	cg21750589	↑	3.17E-09	0.25		
rs1225597	6	28162087	T	C	0.95	2.15E-11	1	0.765	1	0.503	0.651	0.602	0.808	0.703	poins	cg21750589	↑	3.17E-09	0.25		
rs1233708	6	28173219	A	G	0.95	1.23E-11	1	0.721	1	0.473	0.615	0.607	0.793	0.680	poins	cg21750589	↑	3.17E-09	0.25		
rs1778484	6	28240798	A	G	1.05	6.53E-12	1	0.036	0.927	0.315	0.063	0.473	0.71	0.366	poins	cg21750589	↑	3.34E-08	0.22	1	1
rs1150726	6	28243042	A	G	0.95	2.07E-11	1	0.035	0.918	0.314	0.064	0.47	0.714	0.368	poins	cg21750589	↑	3.34E-08	0.22	14	15
rs1150724	6	28250236	T	C	0.95	1.31E-11	1	0.034	0.926	0.306	0.061	0.447	0.714	0.366	poins	cg21750589	↑	3.65E-08	0.22	7	16
rs1043618	6	31783507	C	G	0.95	1.05E-14	1	0.33	0.986	0.057	0.016	0.759	0.702	0.827	crbl	cg18055007	↓	1.81E-08	0.26	↑ DDAH2	?
rs2927071	15	43919081	I	C	1.04	2.78E-06	0.998	0.897	0.985	0.907	0.898	0.420	0.519	0.471	crbl	cg12784172	↑	5.46E-12	0.36		
rs2927085	15	43989530	I	C	1.04	9.57E-06	0.997	0.935	0.978	0.913	0.854	0.545	0.534	0.518	crbl	cg12784172	↑	9.79E-12	0.36		
rs7175129	15	44030018	A	G	1.03	1.85E-05	0.997	0.924	0.967	0.900	0.903	0.480	0.580	0.556	crbl	cg12784172	↑	1.46E-10	0.32		
rs11070410	15	44057045	A	G	0.97	2.43E-06	0.998	0.932	0.981	0.907	0.906	0.476	0.586	0.554	crbl	cg12784172	↑	1.46E-10	0.32	3	
rs1053492	15	44061802	I	C	1.04	8.34E-07	0.999	0.923	0.99	0.921	0.903	0.447	0.539	0.497	crbl ^{ff}	cg12784172	↑	2.02E-18	0.40	3	1
rs3087657	15	44063859	A	G	1.04	5.16E-07	0.999	0.925	0.992	0.925	0.899	0.461	0.561	0.514	crbl	cg12784172	↑	1.46E-10	0.32	3	1
rs11638972*	15	44074691	C	G	0.96	5.86E-07	0.998	0.923	0.993	0.922	0.901	0.454	0.570	0.516	crbl	cg12784172	↑	1.46E-10	0.32	4	1
rs694461	15	44096063	I	C	1.03	5.28E-06	0.996	0.909	0.987	0.907	0.905	0.486	0.597	0.564	crbl	cg12784172	↑	2.11E-10	0.32	↓ ZSCAN29	?
rs693919	15	44102328	A	G	1.03	3.36E-06	0.996	0.917	0.987	0.897	0.905	0.469	0.570	0.560	crbl	cg12784172	↑	2.11E-10	0.32	1	1
rs575082	15	44110142	A	G	0.97	1.25E-05	0.996	0.911	0.988	0.894	0.900	0.486	0.588	0.546	crbl	cg12784172	↑	2.11E-10	0.32	3	1
rs663214	15	44113782	T	G	0.96	6.16E-07	0.998	0.915	0.995	0.915	0.897	0.468	0.565	0.500	crbl	cg12784172	↑	2.11E-10	0.32	9	1
rs2955969	15	44114527	A	G	1.03	4.76E-06	0.996	0.911	0.988	0.896	0.899	0.486	0.587	0.547	crbl	cg12784172	↑	2.11E-10	0.32	12	4
rs686666	15	44161550	A	G	0.97	1.52E-06	0.997	0.922	0.994	0.901	0.897	0.476	0.565	0.539	crbl	cg12784172	↑	2.49E-10	0.32	1	1
rs667476	15	44173755	A	C	1.03	3.49E-05	0.993	0.874	0.973	0.836	0.831	0.477	0.492	0.556	crbl	cg12784172	↑	1.42E-08	0.26	14	5
rs524908	15	44187991	A	C	1.04	5.11E-06	0.997	0.939	0.983	0.910	0.908	0.530	0.523	0.673	crbl	cg12784172	↑	2.19E-09	0.29	13	3
rs2905421*	19	19457908	A	C	1.05	9.21E-08	1	0.910	0	0.819	0.832	0.122	0.385	0.366	crbl ^{ff}	cg14449575	↓	6.74E-10	0.22	↑ HAPLN4	?
rs17002876	22	42599647	C	G	1.06	4.08E-07	0.999	0.051	0.007	0.025	0.169	0.022	0.902	0.321	poins	cg27187881	↓	4.30E-10	0.27	7	?
rs2143139	22	42614401	C	G	1.06	3.72E-07	0.999	0.051	0.007	0.023	0.164	0.026	0.908	0.313	poins	cg27187881	↓	4.30E-10	0.27	16	6
rs134874	22	42661144	A	G	0.94	7.39E-08	1	0.033	0.005	0.016	0.157	0.022	0.926	0.292	poins	cg27187881	↓	4.30E-10	0.27	1	1
rs134885	22	42673811	C	G	0.94	2.13E-07	1	0.042	0.007	0.022	0.172	0.021	0.924	0.273	poins	cg27187881	↓	4.30E-10	0.27	12	3
rs134886	22	42673856	A	G	0.94	2.43E-07	0.999	0.048	0.018	0.028	0.179	0.026	0.925	0.285	poins	cg27187881	↓	4.30E-10	0.27	↑ FAM109B /	?
rs134887	22	42674239	A	G	0.94	1.91E-07	1	0.043	0.007	0.023	0.184	0.022	0.928	0.271	poins	cg27187881	↓	4.30E-10	0.27	↑ NAGA	?
rs134888	22	42674281	I	C	1.07	9.49E-08	1	0.032	0.005	0.019	0.181	0.019	0.940	0.318	poins	cg27187881	↓	4.30E-10	0.27	14	4

rs134889*	22	42674350	A	I	0.94	2.54E-07	0.999	0.047	0.007	0.024	0.185	0.022	0.927	0.279	pons	cg27187881	↓	4.30E-10	0.27	14	4	?
rs134901	22	42683520	I	C	1.07	2.40E-07	1	0.039	0.006	0.023	0.189	0.016	0.915	0.325	pons	cg27187881	↓	4.30E-10	0.27	14	4	
rs134903	22	42684913	A	C	1.06	2.71E-07	1	0.060	0.008	0.028	0.183	0.028	0.914	0.284	pons	cg27187881	↓	4.30E-10	0.27	11	6	

ASM: Allele-specific methylation; ¹Data obtained from the GWA meta-analysis of eight psychiatric disorders (Cross-Disorder Group of the Psychiatric Genomics Consortium., 2019); ²Described in Zhang et al., 2010 and Gibbs et al., 2010; ³eQTL information for the genes with a highlighted CpG site in their possible promoter region; ⁴Histone marks found in brain areas; ⁵Data from the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) consortium (Hibar et al. 2015); SNP; Single Nucleotide Polymorphism; Chr: Chromosome; BP: Position (build hg19); A1: Allele 1; A2: Allele 2; All alleles are reported in the forward strand; OR: Odds Ratio (calculated on A1); SCZ: Schizophrenia; BIP: Bipolar Disorder; MD: Major Depression; ASD: Autism Spectrum Disorder; ADHD: Attention Deficit Hyperactivity Disorder; TS: Tourette Syndrome; AN: Anorexia Nervosa; OCD: Obsessive-Compulsive Disorder; Effect: Direction of the risk allele effect on DNA methylation levels; Enh: Enhancer; Pro: promoter; *: Significant tagSNPs overcoming Bonferroni multiple testing correction, the other SNPs are ASM SNPs in LD with these significant tagSNPs; Underlined allele: Risk allele; In bold: Significant associations between ASM tagSNPs and each psychiatric disorder (m-value ≥ 0.9); ⁶Information from Zhang et al., 2010; Crbl: Cerebellum; Tctx: Temporal cortex; "?: No values retrieved for these SNPs.

Table S4. Histone marks for the 80 highlighted ASM SNPs.

SNP	Chr	BP		Hippocampus Middle	Substantia Nigra	Anterior Caudate	Cingulate Gyrus	Inferior Temporal Lobe	Angular Gyrus	Dorsolateral Prefrontal Cortex	Germinal Matrix	Fetal Brain Female	Fetal Brain Male	
rs2906458	1	44336389	Chromatin state	15-state model										
			25-state model											
			Enhancer histone marks	H3K4me1				H3K4me1					H3K4me1	
			Promoter histone marks	H3K27ac				H3K27ac						
rs9467704	6	26319486	Chromatin state	15-state model										
			25-state model											
			Enhancer histone marks	H3K4me1									H3K4me1	
			Promoter histone marks	H3K27ac										
rs9467714	6	26340785	Chromatin state	15-state model										
			25-state model											
			Enhancer histone marks	H3K4me1										
			Promoter histone marks	H3K27ac										
rs3799378	6	26404374	Chromatin state	15-state model										
			25-state model											
			Enhancer histone marks	H3K4me1										
			Promoter histone marks	H3K27ac										
rs3799380	6	26467182	Chromatin state	15-state model										
			25-state model											
			Enhancer histone marks	H3K4me1										
			Promoter histone marks	H3K27ac										
rs4518487	6	26467531	Chromatin state	15-state model										
			25-state model											
			Enhancer histone marks	H3K4me1										
			Promoter histone marks	H3K27ac										
rs3734544	6	26468660	Chromatin state	15-state model										
			25-state model											
			Enhancer histone marks	H3K4me1										
			Promoter histone marks	H3K27ac										
rs9366658	6	26469866	Chromatin state	15-state model										
			25-state model											
			Enhancer histone marks	H3K4me1										
			Promoter histone marks	H3K27ac										
rs9358944	6	26469875	Chromatin state	15-state model										
			25-state model											
			Enhancer histone marks	H3K4me1										
			Promoter histone marks	H3K27ac										

rs17002876	22	42599647	Chromatin state 15-state model 25-state model	Enh	TxEnhW	TxEnhW	TxEnhW	TxEnhW	TxEnhW	Enh	Enh	Enh
			Enhancer histone marks H3K4me1 H3K27ac	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1
			Promoter histone marks H3K4me3 H3K9ac									
rs2143139	22	42614401	Chromatin state 15-state model 25-state model	Enh	TxEnh5	TxEnh5	TxEnh5	TxEnh5	TxEnh5	Enh	Enh	Enh
			Enhancer histone marks H3K4me1 H3K27ac	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K27ac	H3K27ac	H3K4me1
			Promoter histone marks H3K4me3 H3K9ac	H3K4me3						H3K9ac	H3K9ac	
rs134874	22	42661144	Chromatin state 15-state model 25-state model								Enh	
			Enhancer histone marks H3K4me1 H3K27ac	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K4me1	H3K27ac	H3K4me1	H3K4me1
			Promoter histone marks H3K4me3 H3K9ac									
rs134885	22	42673811	Chromatin state 15-state model 25-state model									
			Enhancer histone marks H3K4me1 H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K27ac	H3K4me1	H3K4me1
			Promoter histone marks H3K4me3 H3K9ac							H3K9ac	H3K9ac	
rs134886	22	42673856	Chromatin state 15-state model 25-state model									
			Enhancer histone marks H3K4me1 H3K27ac	H3K4me1	H3K27ac	H3K27ac	H3K4me1	H3K27ac	H3K4me1	H3K27ac	H3K4me1	H3K4me1
			Promoter histone marks H3K4me3 H3K9ac									
rs134887	22	42674239	Chromatin state 15-state model 25-state model								Enh	Enh
			Enhancer histone marks H3K4me1 H3K27ac	H3K4me1	H3K27ac	H3K27ac	H3K4me1	H3K27ac	H3K4me1	H3K27ac	H3K4me1	H3K4me1
			Promoter histone marks H3K4me3 H3K9ac									
rs134888	22	42674281	Chromatin state 15-state model 25-state model								Enh	Enh
			Enhancer histone marks H3K4me1 H3K27ac	H3K4me1	H3K27ac	H3K27ac	H3K4me1	H3K27ac	H3K4me1	H3K27ac	H3K4me1	H3K4me1
			Promoter histone marks H3K4me3 H3K9ac									
rs134889	22	42674350	Chromatin state 15-state model								Enh	Enh
			Enhancer histone marks H3K4me1 H3K27ac	H3K4me1	H3K27ac	H3K27ac	H3K4me1	H3K27ac	H3K4me1	H3K27ac	H3K4me1	H3K4me1
			Promoter histone marks H3K4me3 H3K9ac									

Table S5. Correlations with sub-cortical brain volumes for the 80 highlighted ASM SNPs.

SNP	ASM for CpG site	Nucleus															
		Accumbens		Amygdala		Caudate nucleus		Hippocampus		Pallidum		Putamen		Thalamus			
		A1	A2	Effect	p-value	Effect	p-value	Effect	p-value	Effect	p-value	Effect	p-value	Effect	p-value		
rs2906458*	cg22930187/cg06207804	A	G	0.181	8.85E-01	0.743	7.84E-01	0.838	8.81E-01	-4.12	4.42E-01	3.948	7.27E-02	8.94	1.91E-01	3.477	6.28E-01
rs9467704		T	C	2.806	1.28E-01	5.56	1.56E-01	5.433	5.04E-01	10.6	1.75E-01	8.261	9.34E-03	4.168	6.76E-01	12.44	2.33E-01
rs9467714*		A	G	1.622	3.46E-01	6.97	5.84E-02	5.456	4.74E-01	8.048	2.72E-01	6.371	3.19E-02	7.213	4.40E-01	10.64	2.77E-01
rs3799378*		A	G	1.787	1.70E-01	4.173	1.33E-01	-3.75	5.17E-01	0.987	8.59E-01	5.526	1.53E-02	3.448	6.27E-01	7.518	3.11E-01
rs3799380		T	C	1.44	2.93E-01	1.205	6.83E-01	2.719	6.58E-01	1.194	8.39E-01	2.852	2.34E-01	1.002	8.94E-01	8.099	3.04E-01
rs4518487		A	G	1.466	2.84E-01	1.166	6.93E-01	2.586	6.73E-01	1.236	8.34E-01	2.841	2.35E-01	1.296	8.63E-01	8.197	2.97E-01
rs3734544*		A	G	1.257	3.70E-01	1.787	5.54E-01	-1.33	8.32E-01	2.416	6.88E-01	2.77	2.58E-01	0.998	8.97E-01	10.48	1.94E-01
rs9366658		A	G	1.213	3.87E-01	1.829	5.45E-01	-1.54	8.07E-01	2.367	6.94E-01	2.819	2.50E-01	0.917	9.05E-01	10.47	1.95E-01
rs9358944	cg14345882	A	C	1.213	3.87E-01	1.829	5.45E-01	-1.54	8.07E-01	2.365	6.95E-01	2.82	2.50E-01	0.917	9.05E-01	10.49	1.94E-01
rs9358945		A	G	1.304	3.52E-01	1.986	5.11E-01	-1.35	8.30E-01	2.535	6.74E-01	2.791	2.54E-01	1.662	8.29E-01	11.21	1.65E-01
rs7773938		T	C	1.247	3.74E-01	1.883	5.33E-01	-1.37	8.27E-01	2.566	6.70E-01	2.841	2.46E-01	0.94	9.03E-01	10.64	1.88E-01
rs3757150		A	I	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs9358946		A	G	1.276	3.62E-01	1.718	5.69E-01	-1.83	7.71E-01	2.584	6.68E-01	2.809	2.51E-01	0.874	9.10E-01	10.41	1.97E-01
rs7756567		T	G	1.287	3.58E-01	1.736	5.65E-01	-1.76	7.79E-01	2.604	6.65E-01	2.811	2.51E-01	0.941	9.03E-01	10.38	1.98E-01
rs2093169		T	C	1.211	3.87E-01	1.696	5.74E-01	-2.56	6.85E-01	2.607	6.65E-01	2.686	2.72E-01	0.559	9.42E-01	10.22	2.05E-01
rs9467772		A	T	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs149901		T	C	1.178	0.3552	-1.96	0.4698	4.914	0.3875	-1.84	0.7353	4.361	0.05162	-7.17	0.3041	2.793	0.7019
rs156744		A	G	1.248	0.3329	-1	0.7192	5.872	0.3074	-1.88	0.7334	4.979	0.02799	-8.28	0.2408	4.413	0.5501
rs464312		A	I	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs149946		T	G	1.19	0.3554	-0.98	0.7247	6.003	0.2965	-1.71	0.7557	4.984	0.02772	-8.44	0.2314	4.288	0.5612
rs469227		A	C	1.148	0.3699	-2.07	0.4522	4.732	0.4078	-2.37	0.6651	4.289	0.05738	-7.63	0.2773	2.406	0.7431
rs149969		I	C	1.215	0.3459	-1.03	0.7106	5.873	0.3075	-1.54	0.7797	4.93	0.02956	-8.38	0.2355	4.21	0.5688
rs149970		A	G	1.081	0.3956	-2.04	0.4525	4.941	0.3845	-2.02	0.711	4.254	0.05734	-8.04	0.2491	2.537	0.7279
rs149972		T	C	0.252	0.8342	-0.72	0.7771	3.78	0.4784	-1.42	0.7806	2.891	0.1648	-9.34	0.1545	0.498	0.9419
rs149975		A	G	0.284	0.8134	-0.65	0.8008	3.748	0.4821	-1.42	0.7819	2.928	0.1595	-9.28	0.1572	0.737	0.9142
rs149989		A	I	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs149944		T	C	0.282	0.8149	-0.76	0.7661	3.705	0.4877	-1.49	0.7718	2.858	0.1703	-9.5	0.1482	0.93	0.892
rs149945		I	C	0.199	0.8686	-0.99	0.6993	3.44	0.5192	-1.73	0.736	2.836	0.1735	-9.82	0.1349	0.818	0.9049
rs149897		A	G	0.22	0.8548	-1.02	0.6902	3.565	0.5045	-1.82	0.7224	2.882	0.167	-9.8	0.1359	0.739	0.9144
rs156739		A	C	0.309	0.7974	-0.59	0.8175	4.024	0.4512	-1.26	0.8057	2.9	0.164	-9.16	0.1633	1.635	0.8113
rs149959		I	C	0.293	0.8073	-0.63	0.8058	4.214	0.43	-1.1	0.8297	2.877	0.1674	-9.14	0.1643	1.759	0.7972
rs183244		A	C	0.315	0.7937	-0.59	0.8167	3.836	0.4724	-1.39	0.7865	2.965	0.1548	-8.78	0.1811	1.847	0.7873
rs149950		I	G	0.314	0.7943	-0.59	0.8171	3.813	0.4749	-1.41	0.7835	2.96	0.1553	-8.75	0.1825	1.865	0.7853
rs149951		I	C	0.313	0.7948	-0.59	0.8175	3.812	0.4751	-1.41	0.7828	2.962	0.1551	-8.75	0.1827	1.86	0.7858
rs149955	cg21750589	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?

rs203876	I	C	0.134	0.9118	-0.8	0.7557	3.293	0.5384	-2.55	0.6199	2.679	0.1997	-9.56	0.1466	-0.71	0.9178
rs203877	I	C	0.123	0.9186	-0.87	0.7332	3.211	0.5486	-2.65	0.6054	2.619	0.2098	-9.36	0.155	-0.86	0.9002
rs203893	A	C	0.055	0.9638	-0.89	0.7301	3.279	0.5401	-2.77	0.5893	2.56	0.2203	-9.58	0.1458	-1.26	0.8539
rs188105	T	C	0.062	0.9592	-1	0.6966	3.18	0.5522	-2.76	0.5903	2.538	0.2242	-9.69	0.1408	-1.3	0.8496
rs203884	A	C	0.052	0.9655	-0.99	0.6987	2.961	0.58	-2.73	0.5949	2.51	0.2293	-9.63	0.1435	-1.09	0.8738
rs203882	A	C	0.094	0.9377	-0.91	0.7215	3.372	0.5283	-2.71	0.5977	2.601	0.2128	-9.56	0.1462	-0.8	0.9069
rs1770131	T	C	0.08	0.9471	-0.97	0.7056	3.464	0.517	-2.61	0.6104	2.577	0.217	-9.39	0.1535	-1.04	0.88
rs1631552	A	C	0.079	0.9477	-1.05	0.6822	3.151	0.5558	-2.61	0.6104	2.513	0.2287	-9.61	0.1441	-1.15	0.8669
rs536704	I	G	0.216	0.8607	-1.39	0.5946	2.075	0.7051	-2.52	0.6294	2.294	0.2808	-9.98	0.1369	-0.94	0.893
rs1225716	A	C	0.114	0.9248	-1.15	0.6536	3.52	0.5109	-1.95	0.7036	2.698	0.1964	-8.09	0.2194	-0.65	0.9243
rs2622319	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs9295760	A	G	0.12	0.921	-1	0.6976	3.603	0.5015	-1.55	0.7633	2.814	0.1785	-7.63	0.2471	-0.3	0.9658
rs1233691	I	G	-0.01	0.9951	-0.46	0.8565	3.8	0.4751	-0.11	0.9824	2.765	0.1834	-5.92	0.3655	0.098	0.9885
rs1225597	T	C	-0.02	0.9835	-0.4	0.8755	4.047	0.4469	0.033	0.9948	2.745	0.1867	-5.71	0.3828	0.198	0.9769
rs1233708	A	G	0.012	0.9919	-0.1	0.9693	4.055	0.4462	0.284	0.9557	2.941	0.1571	-5.48	0.4025	0.29	0.9661
rs1778484	A	G	-0.74	0.5213	1.59	0.5167	4.373	0.3898	0.033	0.9946	-2.19	0.2714	0.948	0.8789	-8.64	0.1848
rs1150726	A	G	-0.83	0.4714	1.523	0.5344	4.229	0.4058	0.081	0.9867	-2.16	0.2777	0.633	0.919	-8.82	0.1761
rs1150724	T	C	-0.83	0.4698	1.508	0.5386	4.048	0.4262	0.047	0.9924	-2.13	0.2845	0.467	0.9402	-8.69	0.1822
rs1043618*	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
cg18055007																
rs2927071	I	C	-1.28	2.90E-01	0.659	7.99E-01	-5.35	3.20E-01	-1.44	7.81E-01	-0.72	7.32E-01	-4.79	4.67E-01	1.831	7.91E-01
rs2927085	I	C	-1.54	2.16E-01	0.453	8.65E-01	-5.76	2.97E-01	-1.77	7.40E-01	-1.27	5.55E-01	-3.93	5.59E-01	0.433	9.51E-01
rs7175129	A	G	-1.31	2.85E-01	0.135	9.59E-01	-2.83	6.02E-01	-2.88	5.81E-01	-0.99	6.40E-01	-5.39	4.17E-01	0.996	8.87E-01
rs11070410	A	G	-1.27	3.06E-01	0.518	8.44E-01	-1.5	7.82E-01	-1.14	8.28E-01	-0.44	8.37E-01	-4.48	5.01E-01	1.644	8.14E-01
rs1053492	I	C	-1.12	3.52E-01	0.386	8.81E-01	-3.96	4.59E-01	-3.03	5.56E-01	-1.22	5.58E-01	-6.34	3.34E-01	-1.04	8.79E-01
rs3087657	A	G	-1.11	3.57E-01	0.482	8.52E-01	-3.97	4.58E-01	-3	5.60E-01	-1.17	5.75E-01	-6.3	3.37E-01	-1.32	8.48E-01
rs11638972*	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs694461	I	C	-1.26	3.00E-01	0.192	9.41E-01	-2.92	5.87E-01	-2.74	5.97E-01	-0.93	6.59E-01	-5.43	4.11E-01	0.374	9.57E-01
rs693919	A	G	-1.02	4.07E-01	0.122	9.63E-01	-2.48	6.47E-01	-1.93	7.11E-01	-0.61	7.74E-01	-4.19	5.27E-01	2.754	6.93E-01
rs575082	A	G	-1.16	3.41E-01	0.134	9.59E-01	-2.62	6.27E-01	-2.56	6.21E-01	-0.86	6.83E-01	-5.21	4.30E-01	0.866	9.01E-01
rs663214	T	G	-1.09	3.68E-01	0.424	8.70E-01	-3.89	4.67E-01	-3.1	5.47E-01	-1.28	5.39E-01	-6.23	3.42E-01	-1.02	8.82E-01
rs2955969	A	G	-1.16	3.41E-01	0.169	9.48E-01	-2.87	5.94E-01	-2.74	5.96E-01	-0.91	6.62E-01	-5.29	4.23E-01	0.848	9.03E-01
rs686666	A	G	-0.98	4.22E-01	0.646	8.04E-01	-3.35	5.38E-01	-3.12	5.47E-01	-1.39	5.09E-01	-5.94	3.70E-01	-0.51	9.42E-01
rs667476	A	C	-1.28	2.99E-01	0.495	8.50E-01	-2.65	6.27E-01	-3.27	5.34E-01	-1.33	5.37E-01	-4.47	5.04E-01	0.965	8.90E-01
rs524908	A	C	-0.91	4.51E-01	1.38	5.91E-01	0.656	9.02E-01	-1.33	7.95E-01	-0.36	8.64E-01	-1.73	7.92E-01	5.881	3.91E-01
rs2905421*	A	C	1.567	1.67E-01	0.068	9.78E-01	1.811	7.25E-01	-2.58	6.02E-01	2.723	1.78E-01	5.283	4.01E-01	9.846	1.37E-01
rs17002876	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs2143139	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?
rs134874	A	G	0.096	9.34E-01	1.718	4.84E-01	1.698	7.38E-01	2.356	6.28E-01	-0.19	9.23E-01	-2.8	6.53E-01	-8.46	1.94E-01
rs134885	C	G	?	?	?	?	?	?	?	?	?	?	?	?	?	?

Table S6. MetaXcan prediction of gene expression effects on the set of psychiatric disorders considered in the cross-disorder GWAS meta-analysis for multiple brain tissues.

Gene	Brain tissue*	Z-score	p-value	N SNPs in model	N SNPs used	N ASM SNPs	Predicted R ²
<i>ARTN</i>	Caudate basal ganglia	-1.65	9.85E-02	19	16	1	0.1052
	Cerebellar hemisphere	-1.60	1.09E-01	15	15	5	0.3702
	Cerebellum	-3.66	<u>2.54E-04</u>	31	31	5	0.3670
	Cortex	-4.01	<u>6.18E-05</u>	5	4	2	0.1434
	Frontal cortex	-3.46	<u>5.37E-04</u>	29	24	2	0.0904
	Hippocampus	-2.53	1.14E-02	140	131	2	0.1574
	Spinal cord	1.89	5.83E-02	19	19	0	0.1510
<i>BTN3A2</i>	Amygdala	-0.91	3.61E-01	96	90	7	0.2039
	Anterior cingulate	-0.66	5.12E-01	72	67	5	0.4768
	Caudate basal ganglia	-2.37	1.79E-02	88	82	8	0.3768
	Cerebellar hemisphere	-1.64	1.01E-01	43	38	5	0.2924
	Cerebellum	-0.62	5.34E-01	71	63	6	0.4083
	Cortex	-2.17	3.01E-02	99	93	6	0.3580
	Frontal cortex	2.54	1.12E-02	105	98	6	0.2675
	Hippocampus	-1.80	7.12E-02	51	43	6	0.3658
	Hypothalamus	-1.02	3.08E-01	119	113	11	0.1200
	Nucleus accumbens	1.71	8.64E-02	126	117	6	0.5888
	Putamen	-1.83	6.79E-02	103	95	15	0.2854
	Spinal Cord	-1.70	8.97E-02	51	46	15	0.5080
	Substantia nigra	-1.18	2.36E-01	104	97	6	0.2892
<i>DDAH2</i>	Cerebellar hemisphere	-3.23	1.26E-03	39	3	0	0.0724
	Cerebellum	-7.15	<u>8.93E-13</u>	21	3	0	0.2509
	Cortex	1.97	4.85E-02	16	4	1	0.1292
	Frontal cortex	5.87	<u>4.42E-09</u>	53	5	2	0.1196
	Hippocampus	-3.99	<u>6.67E-05</u>	15	1	0	0.0437
	Hypothalamus	-0.87	3.86E-01	17	5	0	0.0422
<i>ZSCAN29</i>	Amygdala	1.33	1.83E-01	14	8	3	0.1288
	Caudate basal ganglia	-0.84	4.04E-01	43	37	7	0.1069
	Frontal cortex	-0.85	3.97E-01	42	40	8	0.1473
	Hypothalamus	-3.79	<u>1.53E-04</u>	4	2	1	0.0794
<i>HAPLN4</i>	Cerebellum	4.41	<u>1.05E-05</u>	22	20	1	0.1685
	Frontal cortex	2.87	4.12E-03	23	22	1	0.0611
<i>FAM109B</i>	Nucleus accumbens	1.30	1.95E-01	8	7	0	0.0778
<i>NAGA</i>	Anterior cingulate	1.62	1.06E-01	37	35	13	0.2230
	Caudate basal ganglia	3.53	<u>4.23E-04</u>	12	10	3	0.2134
	Cerebellar hemisphere	1.36	1.73E-01	30	30	1	0.4652
	Cerebellum	3.62	<u>3.00E-04</u>	22	22	3	0.4838
	Cortex	2.68	7.34E-03	26	25	5	0.0950
	Frontal cortex	1.54	1.23E-01	19	18	0	0.2633
	Nucleus accumbens	2.80	5.18E-03	37	33	3	0.1484
	Putamen	2.45	1.44E-02	12	11	4	0.1244
	Amygdala	0.28	7.78E-01	61	30	13	0.2095
	Anterior cingulate	2.69	7.11E-03	45	31	11	0.2855
	Caudate basal ganglia	-1.74	8.24E-02	72	43	17	0.3584
	Cerebellar hemisphere	0.74	4.61E-01	58	26	9	0.2171
	Cerebellum	-2.40	1.64E-02	20	13	1	0.1536

	Cortex	0.12	9.06E-01	83	54	19	0.3814
ZSCAN31	Frontal cortex	-1.98	4.73E-02	26	26	8	0.3504
	Hippocampus	-2.63	8.44E-03	40	37	9	0.2494
	Hipothamalus	0.32	7.51E-01	62	51	8	0.2124
	Nucleus accumbens	1.81	6.99E-02	69	51	12	0.1200
	Putamen	2.09	3.63E-02	77	33	13	0.3896
	Spinal Cord	1.03	3.03E-01	48	40	13	0.2234
	Substantia nigra	-1.45	1.46E-01	47	38	6	0.0877

*Prediction models were only available for some tissues and genes; Z-score: Number of standard deviations change in gene expression; p-value: Significance of the association between predicted expression levels and a set of psychiatric disorders; N SNPs in model: Number of SNPs used in the training of prediction models for each gene; N SNPs used: Number of SNPs used from the cross-disorder GWAS meta-analysis summary statistics; ASM: Allele-specific methylation; N ASM SNPs: Number of ASM SNPs included in the model; Predicted R²: Correlation between the predicted and observed gene expression during prediction model training; Underlined p-values: Significant p-values overcoming Bonferroni correction for multiple testing.

Table S7. Copy number variants overlapping with the eight highlighted genes.

Gene	Variant	Type	Sex	Inheritance	Phenotype(s)
ARTN	1:36831723-61417768	Deletion	46XX	Unknown	Hypoplasia of the corpus callosum, Intellectual disability
	1:41126365-45133712	Deletion	46XY	De novo constitutive	Congenital microcephaly
	1:43827703-45107379	Deletion	46XX	De novo constitutive	Autism, Intellectual disability
	1:43870211-45107379	Deletion	46XX	De novo constitutive	Autism, Delayed speech and language development, Intellectual disability
BTN3A2	6:26116330-26472702	Duplication	46XX	Paternally inherited, constitutive in father	Delayed speech and language development, Impaired social interactions
	6:26031691-28849914	Deletion	46XX	Unknown	Delayed speech and language development, Poor coordination
ZSCAN31	6:26031691-28849914	Duplication	46XX	Unknown	Delayed speech and language development, Poor coordination
	6:7055001-30092021	Deletion	46XY	De novo constitutive	Delayed speech and language development, Intellectual disability
	6:7055001-30092021	Deletion	46XY	De novo constitutive	Abnormality of brain morphology, Delayed speech and language development, Intellectual disability
DDAH2	6:31348684-32768762	Deletion	46XY	De novo constitutive	Delayed speech and language development, Growth delay, Motor delay
	6:31185755-32497381	Duplication	46XY	Unknown	Intellectual disability
ZSCAN29	15:41285354-45826511	Deletion	46XX	De novo constitutive	Delayed speech and language development, Intellectual disability, Truncal obesity
	15:41140284-60397505	Duplication	46XY	Unknown	Delayed speech and language development, Intellectual disability
	19:19254086-19437079	Deletion	46XX	Inherited from normal parent	Delayed speech and language development, Intellectual disability, Microcephaly
HAPLN4	19:18052555-19433673	Deletion	46XX	Unknown	Aggressive behavior, Dyscalculia, Dysgraphia, Emotional lability, Hyperactivity, Intellectual disability, Short attention span
	19:13993244-24301987	Duplication	46XX	Unknown	Global developmental delay, Hyperactivity, Macrocephaly, Motor delay
	19:10636529-31625537	Duplication	46XY	De novo constitutive	Global developmental delay
	22:32539375-51169045	Duplication	46XX	Unknown	Growth delay, Intellectual disability, Intrauterine growth retardation, Microcephaly, Postnatal growth retardation
FAM109B and NAGA	22:32674605-47095548	Duplication	unknown	Unknown	Intellectual disability
	22:37369198-51218950	Amplification	46XY	De novo constitutive	Intellectual disability, Microcephaly, Self-mutilation
	22:38573008-51103900	Duplication	46XX	Unknown	Microcephaly, Postnatal growth retardation
	22:40542666-42883431	Duplication	46XX	De novo constitutive	Intellectual disability
	22:41673826-42889510	Deletion	46XX	Unknown	Delayed speech and language development, Intellectual disability, Cerebral cortical atrophy, Global developmental delay, Mild microcephaly
	22:41701951-42924628	Deletion	46XX	De novo constitutive	Intellectual disability
	22:41834214-45630709	Deletion	46XX	De novo constitutive	Delayed speech and language development, Intellectual disability
	22:41895397-43567966	Deletion	46XX	De novo constitutive	Delayed speech and language development
22:41992973-43298580	Deletion	46XX	De novo constitutive	Autistic behavior, Delayed speech and language development, Dolichocephaly, Global developmental delay	
	22:41999974-51203704	Deletion	46XX	Unknown	Global developmental delay, Macrocephaly
	22:42029866-48224587	Deletion	46XX	Unknown	Global developmental delay, Macrocephaly

22:42139078-42567090	Deletion	46XY	De novo constitutive	Delayed speech and language development, Macrocephaly, Plagiocephaly
22:42160270-45160369	Deletion	46XX	De novo constitutive	Intellectual disability, Macrocephaly
22:42160470-42593988	Deletion	46XY	De novo constitutive	Global developmental delay, Obesity, Polyphagia
22:42176139-47543953	Deletion	46XY	De novo constitutive	Delayed speech and language development, Intellectual disability
22:42238027-43233333	Duplication	46XX	Unknown	Delayed speech and language development, Intellectual disability
22:42252765-51106944	Deletion	unknown	Unknown	Intellectual disability
22:42284638-44063259	Deletion	46XY	De novo constitutive	Intellectual disability
22:42355570-50988545	Duplication	46XX	Unknown	Cognitive impairment
22:42464812-42630809	Duplication	46XY	Maternally inherited, constitutive in mother	Intellectual disability, Language impairment, Microcephaly

Variant: CNV position (build hg19); Source: DECIPHER (v9.27) database; The reported phenotypes have been restricted to those related with the eight psychiatric disorders considered.

DISCUSSIÓ

L'addicció és un trastorn neuropsiquiàtric complex associat a comportaments destructius i maladaptatius a conseqüència de l'abús persistent, compulsiu i incontrolat d'un estímul addictiu. En el cas de les drogues d'abús, per exemple, només una part dels individus que consumeixen aquestes substàncies desenvoluparan l'addicció, i no tots ho faran en el mateix moment. En funció de com es produeixin els canvis necessaris al cervell per l'establiment de l'addicció, alguns la desenvoluparan poc temps després d'haver iniciat el consum i altres al cap de molt temps. Contribueixen a la seva etiologia factors genètics i ambientals, amb l'epigenètica actuant probablement com a medidora entre tots dos. A més, l'addicció presenta una elevada comorbiditat amb la majoria de trastorns psiquiàtrics i, de fet, diversos estudis epidemiològics descriuen que aproximadament la meitat de les persones diagnosticades amb un trastorn psiquiàtric desenvoluparan addicció a les drogues en algun moment de la seva vida i viceversa (Ross i Peselow, 2012).

Durant els primers anys, els estudis genètics de les malalties complexes s'han centrat principalment en l'exploració de gens candidats mitjançant metodologies de genotipació a petita escala seguides d'estudis d'associació cas-control. També s'han estat fent estudis d'expressió gènica dirigits a gens concrets o a grups de gens. Això ha permès identificar variants de risc en alguns gens i detectar alteracions de l'expressió en d'altres. En els últims 10 o 15 anys, el desenvolupament de tecnologies de genotipació d'alt rendiment ha obert les portes als estudis d'associació a escala genòmica (GWAS), i tant els xips d'expressió com la seqüenciació massiva de RNA (RNAseq), han permès abordar anàlisis transcriptòmiques per interrogar l'expressió de tots els gens del genoma. Tot això ha fet canviar radicalment el panorama científic i hem passat de la investigació basada en hipòtesis als estudis sense hipòtesis prèvies, la qual cosa ha permès identificar gens que difícilment haurien estat assenyalats a partir dels coneixements previs que teníem d'aquests trastorns. Però, fins a quin punt són robustos els resultats obtinguts amb aquestes metodologies? Es necessiten estudis de rèplica de les variants obtingudes? Quins són els mecanismes moleculars, cel·lulars i funcionals del cervell que connecten les variants i els gens identificats amb el fenotip que observem en els pacients? Quins estudis es poden dur a terme per estudiar l'impacte funcional de les variants i gens de risc assenyalats pels estudis genètics a gran escala? Els mecanismes epigenètics (metilació del DNA, miRNAs o modificacions d'histones, entre d'altres) poden explicar part de l'heretabilitat perduda d'aquests trastorns? Aquests mecanismes contribueixen també a la base genètica comuna entre els trastorns psiquiàtrics? En aquesta Tesi hem intentat respondre a algunes d'aquestes preguntes a través de l'ús de diferents metodologies, començant per la identificació de factors epigenètics de risc a les addiccions tant en estudis a petita escala (Articles 3 i 4, Capítol 2) o utilitzant tecnologies de seqüenciació d'alt

rendiment (Article 5, Capítol 2). A continuació hem identificat factors epigenètics de susceptibilitat a trastorns comòrbids a les addiccions, com el TDAH o altres patologies psiquiàtriques (Articles 6 i 7, Capítol 3). També hem intentat replicar a la nostra mostra de dependència de drogues senyals d'associació procedents d'estudis GWAS previs (Article 1, Capítol 1). I, finalment, hem realitzat estudis funcionals a nivell cel·lular (Article 2, Capítol 1) i en un model murí (Annex de l'Article 6, Capítol 3) per analitzar gens concrets identificats per nosaltres com a candidats a l'addicció o al TDAH.

1. Consideracions sobre les metodologies emprades

En aquesta Tesi hem utilitzat diferents metodologies per estudiar i identificar factors de risc genètics i epigenètics a les addiccions i a trastorns comòrbids. Totes aquestes metodologies, malgrat tenir avantatges, també presenten certes limitacions que es discuteixen a continuació.

1.1. Estudis d'associació cas-control poblacionals

Els estudis d'associació cas-control poblacionals tenen com a objectiu identificar factors genètics de risc per malalties complexes mitjançant la comparació de les freqüències al·lèliques o genotípiques de variants d'interès entre un grup de casos i un de controls no relacionats entre ells. En les addiccions, així com en els trastorns comòrbids que són objecte d'aquesta Tesi, s'han considerat variants comunes del genoma, de les quals s'espera una contribució al fenotip individualment petita, però globalment important.

En un inici, els estudis d'associació se centraven en gens candidats que s'havien relacionat prèviament amb el fenotip en estudis clínics, farmacològics, genètics o bioquímics, entre d'altres. Posteriorment, gràcies a l'aparició de les tecnologies de genotipació d'alt rendiment, aquesta tendència ha anat canviant, primer estudiant grups de gens candidats i posteriorment analitzant sistemàticament centenars de milers de SNPs situats al llarg del genoma, mitjançant estudis GWAS. En tractar-se d'un enfoc en què no calen hipòtesis prèvies perquè s'explora tot, s'han pogut identificar gens o vies funcionals que no havien estat considerades mai abans.

Els estudis d'associació cas-control tenen diverses limitacions que són inherents a la metodologia. Per una banda hi ha els falsos positius (errors de tipus I), que apareixen quan detectem associacions que no són reals, i per l'altra hi hauria els falsos negatius (errors de tipus II), que apareixen quan no detectem associacions reals. Per tal de minimitzar aquest tipus d'errors i augmentar la probabilitat d'identificar variants de risc és important tenir en compte les consideracions exposades a continuació.

Poder estadístic i mida mostral

El poder estadístic és la probabilitat de detectar una associació en el cas que aquesta existeixi, i és un factor clau en el disseny dels estudis d'associació. Hi ha diversos factors que influeixen en l'estimació del poder estadístic, com la prevalença de la malaltia a la població general, la mida de la mostra, la freqüència al·lèlica del marcador associat a la malaltia, el risc que aquest confereix (mesurat amb l'*odds-ratio*, OR) i el seu model d'herència (Nsengimana i Bishop, 2017). D'entre tots ells, la mida mostral és el factor principal a tenir en compte per obtenir un poder estadístic suficient. A més, la mida mostral presenta una correlació inversa amb l'impacte fenotípic de la variant, de manera que, per detectar variants amb una contribució molt subtil a un fenotip necessitarem mides mostrals més grans.

Tres dels estudis d'associació inclosos en aquesta Tesi (Article 1, Capítol 1; Articles 3 i 4, Capítol 2) s'han realitzat amb una mostra d'individus dependents de drogues i controls espanyols no relacionats, d'ascendència caucàsica reclutats a Barcelona. La mida mostral d'aquests tres estudis ha augmentat progressivament al llarg de la Tesi perquè hem anat reclutant nous pacients i controls, i amb això també ha augmentat el poder estadístic. Tot i així, aquest segueix sent limitat per detectar senyals sota models on es requereixen mides mostrals més grans, com el model recessiu, o per variants amb un efecte petit en la malaltia o amb una MAF (freqüència de l'al·lel menor) petita. Per altra banda, també hem realitzat dos estudis d'associació utilitzant dades GWAS d'estudis previs amb mides mostrals molt més grans i per tant, amb més poder estadístic, que ens han permès identificar variants epigenètiques de risc per diferents trastorns psiquiàtrics comòrbids a l'addicció (Articles 6 i 7, Capítol 3).

Homogeneïtat del fenotip estudiat

En dissenyar un estudi d'associació cas-control d'una malaltia complexa és molt important que el fenotip sigui el més homogeni possible per evitar una reducció del poder estadístic. Per tal de reduir l'heterogeneïtat de la nostra mostra de casos de dependència de drogues, tots ells han estat avaluats al mateix centre (Hospital Vall d'Hebron de Barcelona) per un únic equip de professionals clínics, i diagnosticats en base als mateixos criteris, definits en el DSM-IV-TR (American Psychiatric Association, 2013) i utilitzant les dades recollides a través de l'entrevista estandarditzada *Structured Clinical Interview* (SCID).

En els estudis d'associació realitzats per a la dependència de drogues, el factor d'heterogeneïtat més difícil de controlar és la concurrència amb l'abús o dependència d'altres drogues, així com l'alta comorbiditat amb altres trastorns psiquiàtrics (Ross i Peselow, 2012), que pot actuar com a

factor de confusió donant lloc a falsos negatius. La dificultat d'obtenir individus dependents a una sola substància fa que molts estudis d'associació considerin totes les drogues com un conjunt homogeni (dependència de substàncies), ja que totes elles actuen directament o indirecta sobre el sistema de recompensa del cervell, i per tant compartien una base biològica comuna (Wise, 1996; MacNicol, 2017). Tot i això, cada droga també activa vies específiques (Graziane *et al*, 2016; Russo *et al*, 2009; Zhou *et al*, 2011), i és per això que també hi ha estudis que intenten identificar variants per a la dependència d'una droga específica, enriquint la mostra amb individus dependents de la droga d'interès. En els estudis d'associació inclosos en aquesta Tesi hem considerat dos fenotips d'estudi, la dependència de drogues en general i la dependència de cocaïna en particular utilitzant les aproximacions abans explicades (Article 1, Capítol 1; Articles 3 i 4, Capítol 2).

Selecció de la mostra control

En els estudis d'associació és fonamental seleccionar correctament la mostra control, sempre intentant que el nombre de controls sigui similar o més elevat que el nombre de casos (Nsengimana i Bishop, 2017) per incrementar el poder estadístic i evitar l'aparició de falsos positius, tal i com hem fet en els estudis d'associació inclosos en aquesta Tesi. També és important que la mostra control tingui característiques similars a la mostra de casos pel que fa a l'ètnia, la proporció de sexes i l'edat, per evitar falsos positius. La nostra mostra de casos i controls per la dependència de drogues, presenta diferències significatives pel que fa a l'edat, ja que la mostra de casos és molt més jove. És per aquest motiu que en tots els estudis considerem l'edat com una covariable (Article 1, Capítol 1; Articles 3 i 4, Capítol 2). Per altra banda, sempre hem mantingut la mateixa proporció de sexes en casos i controls ja que la dependència de drogues és més prevalent en homes.

També és important controlar l'estratificació poblacional, que es dona quan hi ha diferències geogràfiques, socials o culturals entre les mostres de casos i controls seleccionades (Hellwege *et al.*, 2017). Aquest factor de confusió sol donar lloc a falsos positius, a causa de les diferències en les freqüències al·lèliques no relacionades amb el fenotip d'estudi. L'estratificació poblacional però, sol tenir un impacte molt limitat si els casos i els controls són homogenis respecte al seu origen ètnic, com és el cas de la nostra mostra, on casos i controls són d'ètnia caucàsica, nacionalitat espanyola i fills de progenitors també espanyols. A més, en un estudi previ del nostre grup es va descartar la presència d'estratificació poblacional utilitzant dades genètiques en un subgrup de les mostres incloses en els estudis d'associació d'aquesta Tesi (Fernández-Castillo *et al.*, 2013).

Finalment, en el cas de la dependència de drogues hi ha certa controvèrsia en relació a la selecció dels controls, ja que hi ha autors que defensen que els millors controls són aquells que han estat exposats a la droga com a mínim una vegada a la seva vida però que no han desenvolupat la dependència (Bierut *et al.*, 2010; Gelernter *et al.*, 2014). Aquesta postura però, no permet l'estudi dels factors de risc que empenyen als individus a tenir el seu primer contacte amb les drogues (Camchong, Endres i Fein, 2014; Mitchell i Potenza, 2014), com per exemple la impulsivitat o la tendència a les pràctiques de risc, ambdues amb una elevada component genètica (Bevilacqua i Goldman, 2013). És per aquesta raó que nosaltres i altres autors fem servir controls poblacionals que poden haver estat exposats o no a la droga (Ikeda *et al.*, 2013; Johnson *et al.*, 2015). És també important que el fenotip d'estudi s'hagi exclòs a la mostra control per evitar falsos negatius, sobretot si el trastorn és molt prevalent en la població. En la nostra mostra control hem pogut descartar el fenotip de consum de drogues injectables, ja que és un criteri d'exclusió de donació en el Banc de Sang i Teixits, on hem reclutat els controls. Malauradament però, no hem pogut descartar el fenotip de dependència de drogues. Tot i així, tenint en compte que només un 0,7% de la població mundial adulta presenta algun trastorn per consum de drogues il·lícites (UNODC, 2019) i un 5,1% presenta algun trastorn per consum d'alcohol (World Health Organization, 2018), l'impacte que pugui tenir una contaminació de la mostra de controls amb individus afectats és molt petit.

Errors de genotipació

En els estudis d'associació també hem de tenir en compte els errors de genotipació, que normalment solen ser deguts a la qualitat del DNA o a la manipulació de dades. Per aquest motiu és important afegir duplicats i controls negatius entre les mostres que s'han de genotipar (Abou-Sleiman, Hanna i Wood, 2006), tal i com hem fet en els estudis d'associació en què hem utilitzat la nostra mostra de dependents de drogues (Article 1, Capítol 1; Articles 3 i 4, Capítol 2). A més, en tots els estudis hem calculat la taxa de genotipació per cada individu i variant i hem eliminat aquells individus o variants que no superessin un determinat llindar, sempre per sobre del 90%. Finalment, també hem considerat l'equilibri de Hardy-Weinberg, que es calcula tot comparant les distribucions dels genotips i les freqüències al·lèliques, una eina útil per detectar errors de forma sistemàtica, ja que desviacions d'aquest equilibri ens poden informar d'errors en la genotipació. En els estudis mencionats anteriorment es va calcular l'equilibri de Hardy-Weinberg (HWE) en la mostra control i es van eliminar aquells polimorfismes que no complien aquest requisit.

Rèpliques

Els estudis de rèplica són una eina molt utilitzada per confirmar l'associació i la direcció de l'efecte de variants assenyalades per estudis previs (Chanock *et al.*, 2007). Quan una associació es replica en una o més mostres independents, es reforça la seva associació amb el fenotip avaluat. En canvi, en el cas de no replicar la senyal avaluada, és difícil discernir si l'associació estudiada és un fals positiu o si el que impedeix replicar la variant són les diferències existents entre els estudis de descoberta i rèplica (Chanock *et al.*, 2007; Greene *et al.*, 2009; Bühler *et al.*, 2015). Tot i així, els estudis de rèplica, com l'estudi inclòs en l'Article 1 del Capítol 1, segueixen sent importants per demostrar les associacions descrites prèviament, sobretot per aquelles variants assenyalades en estudis amb mides mostrals limitades, com és el cas de la majoria d'estudis d'associació per la dependència de drogues.

Alguns estudis d'associació utilitzen el disseny en dues etapes: una part de la mostra s'utilitza com a descoberta per genotipar una gran quantitat de variants i només aquelles que presenten associació amb el trastorn es genotipen en una mostra de rèplica independent, disseny utilitzat en dos dels estudis d'associació inclosos en aquesta Tesi (Articles 3 i 4, Capítol 2). Aquesta estratègia permet una reducció dels costos de genotipació, ja que en la descoberta s'analitzen totes les variants candidates, i només aquelles variants que mostren una associació significativa s'analitzen en la mostra de rèplica independent (Satagopan i Elston, 2003; Wason i Dudbridge, 2012). La baixada de costos de genotipació però, fa que cada vegada sigui més freqüent l'anàlisi conjunt de les mostres o les metanàlisis per tal d'obtenir una mida mostral més gran i així augmentar el poder estadístic per detectar associacions significatives (Satagopan i Elston, 2003; McCarthy *et al.*, 2008; Evangelou i Ioannidis, 2013).

Correccions per tests múltiples

En els estudis d'associació se solen analitzar múltiples variants genètiques a la vegada. Això fa que augmenti el nombre de comparacions que es fan i que, per tant, incrementi la probabilitat de trobar falsos positius. És per això que és molt important realitzar correccions per tests múltiples per tal d'augmentar l'astringència del llindar de significació. La correcció més utilitzada és la de Bonferroni, que consisteix a establir un nou llindar de significació que té en compte el nombre total de tests/variants estudiades/models analitzats (n), de manera que el llindar de significació passa a ser $p\text{-valor} = 0,05/n$. Tot i això, la correcció de Bonferroni assumeix independència entre els tests, i en el cas dels estudis d'associació sovint els SNPs no són del tot independents, ja que alguns d'ells estan en desequilibri de lligament. Això fa que aquesta correcció acabi sent

excessivament rigorosa i que la probabilitat d'obtenir falsos negatius augmenti considerablement, sobretot quan es realitzen molts tests. Un segon mètode àmpliament utilitzat i menys astringent és la taxa de falsos descobriments (*False Discovery Rate*, FDR), que calcula la proporció de resultats que són falsos positius entre tots els resultats positius obtinguts, considerant la distribució dels nivells de significació assolits com un tot i discriminant entre la hipòtesi nul·la i aquells valors que se'n desvien, ajustant així el nivell de significació. Aquest mètode és més adequat quan es realitza una gran quantitat de tests. Finalment, també hem utilitzat la *Genetic Type I Error Calculator* (GEC) (M. X. Li *et al.*, 2012), una aplicació desenvolupada per adreçar les correccions per tests múltiples que té en compte la dependència entre els SNPs presents en l'estudi d'associació.

En els estudis d'associació inclosos en aquesta Tesi s'han aplicat mètodes de correcció quan el tipus d'estudi ho requeria. En els estudis on avaluàvem un número elevat de variants vam aplicar sempre correccions per tests múltiples, en el cas de l'Article 6 del Capítol 3 vam aplicar FDR i GEC i, en canvi, en l'Article 7 del Capítol 3 vam utilitzar la correcció de Bonferroni. En els estudis on avaluàvem un nombre petit de variants independents vam utilitzar la correcció de Bonferroni per reforçar els resultats obtinguts (Article 1, Capítol 1; Article 4, Capítol 2), però en els estudis amb descoberta i rèplica la necessitat d'aplicar aquest tipus de correccions no és tant evident, ja que el propi disseny d'aquests estudis redueix la probabilitat d'obtenir falsos positius en requerir la rèplica dels resultats en una mostra independent.

Selecció de polimorfismes candidats

En els estudis d'associació basats en hipòtesis se seleccionen i es genotipen aquelles variants que hom creu que poden influir en la susceptibilitat a la malaltia estudiada. Pot tractar-se de variants que cauen en gens o sistemes funcionals candidats, o bé variants que poden tenir un impacte funcional segons estudis previs o determinades eines bioinformàtiques. Al llarg dels anys s'han utilitzat diferents tipus de marcadors genètics en els estudis d'associació. En l'actualitat, però, les variants de tipus SNP són les que més s'utilitzen perquè són abundants en el genoma i a més són fàcils de genotipar. En els estudis d'associació inclosos en aquesta Tesi ens hem centrat en els SNPs amb un possible impacte funcional, concretament en variants que tenen influència sobre la metilació de llocs CpG (Article 3, Capítol 2; Articles 6 i 7, Capítol 3) o que poden alterar la unió de molècules de miRNA a dianes de mRNA (Article 4, Capítol 2).

Els estudis d'associació basats en hipòtesis prèvies però, s'estan quedant enrere a causa de l'aparició de les metodologies de genotipació d'alt rendiment, que permeten analitzar sistemàticament i sense hipòtesis prèvies milions d'SNPs distribuïts per tot el genoma.

1.2. Estudis d'associació a escala genòmica (GWAS)

L'addicció i els trastorns psiquiàtrics comòrbids tenen una elevada càrrega genètica, tal i com ho demostren els estudis d'heretabilitat clàssics que es realitzen amb bessons. Tot i així, els resultats obtinguts fins ara a través de diferents aproximacions experimentals queda lluny d'explicar tota l'heretabilitat d'aquests trastorns. En l'inici de l'era de la genotipació d'alt rendiment es va pensar que els GWAS ens ajudarien a identificar l'heretabilitat perduda de les malalties complexes gràcies a l'alta cobertura d'aquest tipus d'estudis i a la gran sensibilitat per identificar variants d'efectes petits sobre el fenotip (Hirschhorn i Daly, 2005). La realitat, però, no ha satisfet totes les expectatives, ja que malgrat que s'han pogut descriure moltes variants i gens que prèviament no s'havien relacionat amb aquestes malalties, i a la vegada replicar resultats prèviament identificats en estudis de gens candidats, com el cas de l'alcohol deshidrogenasa 1B en relació al consum d'alcohol (Clarke *et al.*, 2017; Walters *et al.*, 2018), encara queda part de la base genètica d'aquestes malalties sense explicar (Manolio *et al.*, 2009; Young, 2019). Un exemple concret: El TDAH té una heretabilitat del 76% segons una exhaustiva metanàlisi d'estudis de bessons (Faraone *et al.*, 2005), tot indicant que la genètica explica aproximadament $\frac{3}{4}$ de la variació fenotípica en aquest trastorn. Els estudis d'heretabilitat basada en SNPs, que parteixen de dades GWAS i que permeten identificar la fracció de la variància fenotípica deguda als SNPs, donen una xifra del 22% (Demontis *et al.*, 2019). És a dir, un 76% del TDAH té un origen genètic i un 22% del TDAH es pot explicar per variacions en SNPs; o, en altres paraules, aproximadament una tercera part de la genètica del TDAH està en els SNPs. On cau la resta? No ho sabem bé encara.

Una de les grans limitacions dels estudis GWAS és la necessitat d'una mida mostral molt gran a causa del número de variants polimòrfiques testades, que requereix un llindar de significació molt astringent per assegurar-nos que no obtenim falsos positius. Els GWAS utilitzen un llindar de significació preestablert (p -valor $< 5E-08$; $0,05/1E06$), basat en la correcció de Bonferroni per un milió de tests independents, tot i que no sempre es genotipa un milió d'SNPs, i que els SNPs genotipats no sempre són independents en termes de desequilibri de lligament. A més, també s'ha establert un segon llindar per a variants suggestives (p -valor $< 1E-05$) (Johnson *et al.*, 2010). En el cas dels GWAS de dependència de drogues publicats fins ara s'han obtingut poques senyals d'associació que superin el llindar de significació en comparació amb altres trastorns psiquiàtrics, probablement per un poder estadístic insuficient degut a una mida mostral petita i a

l'heterogeneïtat existent entre els casos (M. J. Li *et al.*, 2012, 2016). Per tal de fer front a aquesta limitació, s'estan creant múltiples consorcis, com el *Psychiatric Genomics Consortium* (PGC), que està invertint molts esforços en unir investigadors d'arreu del món per dur a terme estudis genètics en mostres de grans dimensions per diversos trastorns psiquiàtrics (Sullivan *et al.*, 2018).

Una altra limitació dels GWAS és la dificultat en determinar quines són les variants causals entre les variants identificades, així com el gen diana afectat per la variant causal. Contràriament a les expectatives inicials, la gran majoria de les variants identificades en GWAS cauen en regions no codificants i lluny de gens coneguts (Maurano *et al.*, 2012; Schaub *et al.*, 2012; Tak i Farnham, 2015), de manera que la hipòtesi més probable és que les variants causals tinguin un efecte regulador en l'expressió de gens situats en *cis* o en *trans*. Aquesta hipòtesi es recolza en diversos estudis que demostren que aquestes variants associades que cauen en regions no codificants solen estar situades preferentment en regions reguladores (Ernst *et al.*, 2011; Maurano *et al.*, 2012; Schaub *et al.*, 2012; Hannon *et al.*, 2015; Gagliano *et al.*, 2016). El seu impacte podria, doncs tenir a veure amb la interferència sobre la unió d'elements reguladors, o amb possibles alteracions de la metilació, entre d'altres mecanismes. És en aquest punt on l'ús de dades epigenòmiques podria resultar clau per esbrinar quin és l'impacte funcional d'aquestes variants (Wang *et al.*, 2016). A més, en els últims anys s'han desenvolupat múltiples eines, models i anàlisis integratives que permeten incorporar diferents fonts d'informació externes a les anàlisis GWAS per identificar el possible potencial regulador de les variants obtingudes (Boyle, Li i Pritchard, 2017; Watanabe *et al.*, 2017; Wu *et al.*, 2018; Wu i Pan, 2019). En aquesta Tesi hem decidit 'buscar' en dades GWAS de diferents trastorns psiquiàtrics senyals de la contribució de la metilació específica d'al·lel (ASM) a la patologia (Articles 6 i 7, Capítol 3).

1.3. Estudis transcriptòmics i metilòmics

En l'inici dels estudis transcriptòmics i metilòmics, la tecnologia més utilitzada per a obtenir les dades eren els *microarrays* (o xips), amb els quals s'anàlitzava l'expressió d'un conjunt de (o tots els) gens o de mecanismes epigenètics predeterminats. Actualment però, malgrat que els *microarrays* han evolucionat molt i se segueixen utilitzant, estan quedant en un segon pla per deixar pas a les tecnologies basades en la seqüenciació d'alt rendiment. Aquestes tecnologies permeten analitzar l'expressió de tots els gens (transcriptoma) i dels diferents mecanismes de regulació epigenètics (epigenoma) amb una gran sensibilitat. A més, permeten obtenir tant dades qualitatives (p.ex. la presència de transcrits alternatius), com quantitatives (nivells d'expressió o de metilació) (Mutz *et al.*, 2013). Tot i això, la implementació d'aquestes metodologies també comporta certes dificultats, ja que suposa un increment de la quantitat de dades a emmagatzemar

així com més complexitat d'anàlisi, que en part s'han resolt a través del desenvolupament d'eines bioinformàtiques (Costa-Silva, Domingues i Lopes, 2017). En aquesta Tesi hem realitzat un estudi transcriptòmic i metilòmic (Article 5, Capítol 2) en el qual hem utilitzat algunes d'aquestes eines bioinformàtiques per: i) analitzar la qualitat de les seqüències obtingudes i alinear-les contra el genoma de referència, ii) analitzar els patrons de metilació del DNA i iii) per avaluar el patró d'expressió de miRNAs.

En aquest tipus d'estudis també és molt important definir la mostra d'estudi, que en general solen ser teixits de models animals, teixits *post-mortem* d'humans, o bé cultius cel·lulars. La majoria d'estudis, però, recorren als teixits de models animals i als cultius cel·lulars, que tot i que no mimetitzen a la perfecció els fenotips d'estudi, permeten controlar millor les variables experimentals, un punt clau per a l'obtenció de resultats significatius i robustos. En l'actualitat, un dels models animals més utilitzat és el ratolí, ja que és econòmicament assequible, és relativament fàcil de manipular genèticament, és un animal petit i per tant fàcil de mantenir en un estabulari, i el temps entre generacions és curt, entre d'altres avantatges. A més, el perfil d'expressió gènica també està força conservat respecte als humans, sobretot en aquells teixits on és freqüent l'expressió de gens específica de teixit, com el cervell, el cor, el fetge, els ronyons o els testicles (Breschi, Gingeras i Guigó, 2017). En un dels estudis d'aquesta Tesi (Article 5, Capítol 2) hem aprofitat aquesta similitud del perfil d'expressió que presenta el cervell entre ambdues espècies per identificar factors genètics i epigenètics de susceptibilitat a l'addicció al menjar. Concretament hem realitzat un estudi transcriptòmic (mRNA i miRNA) i metilòmic en dues àrees cerebrals implicades en l'addicció, NAc i mPFC, en un model de ratolí per a l'addicció al menjar. Per tal de reduir al màxim la variabilitat entre mostres vam utilitzar una soca consanguínia i tots els ratolins es van sotmetre a les mateixes condicions experimentals i es van sacrificar en el mateix moment. A més, el disseny d'aquest estudi comptava amb una mostra biològica de rèplica que reforça la validesa dels resultats obtinguts.

1.4. Models animals

El desenvolupament de models animals d'addicció i de trastorns psiquiàtrics comòrbids, així com l'ús de models animals genoanul·lats, és imprescindible per millorar el coneixement de la neurobiologia d'aquests trastorns i el descobriment de noves dianes terapèutiques. Malauradament, la generació de models animals que mimetitzin a la perfecció els trastorns psiquiàtrics ha estat per ara impossible a causa de l'alta complexitat i la gran heterogeneïtat genètica i fenotípica d'aquestes malalties, i per les diferències fonamentals que hi ha entre els

humans i els animals que s'usen com a model en molts aspectes relacionats amb el comportament o les capacitats cognitives.

En el cas de les addiccions, l'ús de models animals ens permet controlar moltes variables experimentals que no podem controlar en els estudis en humans, com l'edat i el sexe dels animals, la dosi de droga, el temps d'exposició, etc. A més, hi ha múltiples paradigmes que s'utilitzen per modelar diferents aspectes de la conducta addictiva, com la persistència, la motivació, el reforç, la pèrdua de control o la recaiguda (García *et al.*, 2017). Un dels models més utilitzats és el model d'autoadministració operant, descrit a la introducció, en el qual l'animal aprèn el comportament de cerca de la droga en unes caixes equipades amb dues palanques, una activa i una inactiva. Cada vegada que l'animal activi la palanca o forat actiu aquest rebrà una dosi de la droga i també s'activarà l'estímul condicionat (llum o so) que l'animal associarà a la droga. Aquest model està molt relacionat amb l'abús en humans, tant per la forma d'administració com per la resposta conductual que es genera en consumir la substància (O'Connor *et al.*, 2011) i s'ha utilitzat àmpliament per estudiar l'efecte de drogues d'abús. Recentment també s'ha validat un model d'autoadministració de característiques similars per estudiar l'addicció al menjar (Mancino *et al.*, 2015; Domingo-Rodríguez *et al.*, 2020), model que hem utilitzat en un dels estudis inclosos en aquesta Tesi (Article 5, Capítol 2).

Els diferents paradigmes disponibles, com el de l'autoadministració operant, es poden testar en models animals genoanul·lats per estudiar la implicació d'un gen concret en un trastorn. Tot i que els ratolins genoanul·lats són una eina molt valuosa en la investigació, hi ha una sèrie de limitacions que cal tenir en compte (Nelson, 1997): 1) Un 15% dels models genoanul·lats complets són letals durant el desenvolupament. Aquesta limitació es pot resoldre mitjançant l'ús de mutacions condicionals en el temps o segons el teixit. 2) L'eliminació completa del gen pot donar lloc a fenotips molt incapacitants que no permeten una correcta avaluació de la implicació real del gen en la malaltia. Una possible solució és l'ús d'animals genoanul·lats heterozigots que, en general, expressen un 50% de la proteïna. És important recordar que en les malalties complexes, cadascun dels gens implicats solen presentar alteracions funcionals o d'expressió moderades o lleus. 3) En alguns casos, l'eliminació total d'un gen pot ser que no produeixi cap canvi observable en el fenotip presentat pels ratolins, o que doni lloc a un fenotip diferent del de la patologia que estem modelant. Cal tenir en compte que, en general, darrere de cada malaltia complexa hi ha un conjunt de gens alterats, i és la combinació d'aquests defectes que dona lloc a un fenotip determinat. En el model animal, en general només estem actuant sobre un d'aquests gens. 4) Per alguns gens és molt complicat l'obtenció d'un model genoanul·lat. En aquesta Tesi hem utilitzat

un model de ratolí genoanul·lat per estudiar la contribució del gen *C2orf82* en el TDAH (Annex de l'Article 6, Capítol 3). En el nostre cas, l'eliminació d'aquest gen no donava lloc a un fenotip letal ni a un fenotip altament incapacitant, de manera que hem pogut realitzar experiments tant en els ratolins homozigots com en els heterozigots per la mutació.

Un cop discutides les principals limitacions i avantatges de les metodologies emprades en aquesta Tesi, a continuació es discuteixen els principals resultats obtinguts.

2. Identificació de factors epigenètics de susceptibilitat

Inicialment, els estudis de cerca de factors de susceptibilitat a les malalties complexes se centraven en l'anàlisi de gens o de sistemes de gens candidats, sovint prioritzant variants intragèniques amb un efecte directe sobre la funció de la proteïna codificada per aquell gen. Aquesta tendència però, va anar canviant amb l'aparició de les tecnologies de genotipació a gran escala, que van permetre la realització d'estudis massius en què es va observar que moltes de les variants associades no tenien un efecte funcional obvi o no estaven relacionades amb gens o sistemes prèviament assenyalats com a candidats. Així doncs, a partir d'aquests estudis va sorgir la necessitat d'entendre l'impacte funcional d'aquestes variants i la seva relació amb gens concrets, situats a vegades a grans distàncies en el genoma. És aquí on entra en joc l'epigenètica, proposada des de fa temps com a peça clau per entendre les bases genètiques de les malalties complexes. L'epigenètica té un paper fonamental en el desenvolupament de malalties complexes, ja que pot actuar com a moderadora de la interacció entre la genètica i l'ambient. A nivell molecular, els mecanismes epigenètics regulen l'expressió gènica sense modificar la seqüència de nucleòtids, essent les modificacions de les histones, la metilació del DNA i els miRNAs els tres mecanismes principals. En aquesta Tesi ens hem centrat en la metilació del DNA i en els miRNAs.

2.1. Variants genètiques que tenen influència sobre la metilació del DNA

La metilació específica d'al·lel (*Allele-Specific Methylation, ASM*) és un mecanisme epigenètic heretable, dinàmic i específic de teixit, bastant freqüent en el genoma humà. Es caracteritza per diferències en els nivells de metilació d'un dinucleòtid CpG en funció de l'al·lel que està present en una variant genètica generalment situada en *cis*, tot i que també pot estar en *trans* (Meaburn, Schalkwyk i Mill, 2010; Gaur *et al.*, 2013). En tractar-se d'un mecanisme específic de teixit, en els últims anys s'han realitzat múltiples estudis per descriure SNPs que mostren ASM en diferents teixits (Do *et al.*, 2017), entre ells cervells *post-mortem* humans (Gibbs *et al.*, 2010; Zhang *et al.*, 2010). També és interessant destacar que hi ha un enriquiment de variants que mostren ASM en regions amb marques d'histones relacionades amb promotors i *enhancers* (Hannon *et al.*, 2015; Gagliano *et al.*, 2016; Benton *et al.*, 2019), tot destacant el seu possible paper en la regulació

gènica. A més, també s'ha observat que hi ha un enriquiment de variants que mostren ASM entre les variants obtingudes en GWAS de diferents trastorns psiquiàtrics (Hannon *et al.*, 2015; Gagliano *et al.*, 2016). En base a aquestes evidències vam pensar que seria interessant explorar quina era la contribució de l'ASM en l'addicció a drogues i en trastorns comòrbids (Article 3, Capítol 2; Articles 6 i 7, Capítol 3).

Per aquests estudis vam partir dels resultats obtinguts en dos treballs previs realitzats per altres autors en què es van identificar variants que mostren ASM en múltiples regions de cervell humà *post-mortem* (Gibbs *et al.*, 2010; Zhang *et al.*, 2010). A partir d'aquests dos estudis vam elaborar dos llistats d'SNPs amb un efecte en la ASM, una més reduïda (182 SNPs) que vam inspeccionar a la nostra mostra de dependència de drogues, i una més extensa (3.896 SNPs) que vam explorar en els resultats de dos estudis GWAS, un de TDAH (Demontis *et al.*, 2019) i un que inclou vuit trastorns psiquiàtrics (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). El nombre d'SNPs utilitzat per la dependència de drogues va ser més petit ja que aquestes variants s'havien de genotipar en la nostra mostra, i això suposava un cost econòmic elevat. Així doncs, en un futur, seria interessant ampliar l'estudi quan hi hagi dades GWAS disponibles per aquest trastorn amb mides mostrals grans, ja que encara no hi ha cap GWAS en dependència de drogues il·lícites que superi els 5.000 pacients.

En el cas de la dependència de drogues vam realitzar un estudi d'associació utilitzant la nostra mostra de dependència de drogues, obtenint dues variants associades amb el trastorn, rs6020251 i rs11585570. Els genotips per aquestes dues variants presenten correlació amb els nivells de metilació de tres llocs CpG situats en regions reguladores de *CTNBL1* (rs6020251) o de *SCP2* i *ECHDC2* (rs11585570). A més, la variant rs6020251 cau en una regió rica en marques d'histones relacionades amb promotors i *enhancers* en cervell, la variant rs11585570 és un eQTL en cervell i en sang per *SCP2* i *ECHDC2* i ambdues variants poden afectar l'afinitat d'unió de diferents factors de transcripció, reforçant la idea del d'un paper regulador d'aquestes. Tots tres gens s'expressen en cervell i tenen funcions molt diverses: *CTNBL1* codifica un component del complex de tall i unió (*spliceosome*) i és necessari per activar l'*splicing* de pre-mRNAs, *SCP2* està involucrat en la transferència de lípids entre membranes i *ECHDC2* està relacionat amb el metabolisme dels àcids grassos. Finalment, és interessant destacar que els gens *CTNBL1* i *SCP2* s'havien vinculat prèviament amb trets i patologies molt relacionades amb la dependència de drogues (Myers-Payne *et al.*, 1996; Liu *et al.*, 2008; Papassotiropoulos *et al.*, 2013). Així doncs, aquesta aproximació ens ha permès assenyalar tres gens amb una possible contribució a la dependència de drogues: *CTNBL1*, *SCP2* i *ECHDC2*.

A continuació ens vam centrar en el TDAH, un trastorn en què el grup de recerca té experiència i que presenta una alta comorbiditat amb la dependència de drogues (De Alwis *et al.*, 2014; Harstad *et al.*, 2014; van de Glind *et al.*, 2014; Chen *et al.*, 2018), de manera que vam pensar que variants amb un efecte en la metilació també podien contribuir a aquest trastorn. En aquest cas, vam decidir explorar els resultats d'un GWAS de TDAH (Demontis *et al.*, 2019) utilitzant la nostra estratègia de priorització de variants amb efectes sobre la metilació en cervell. D'aquesta manera potser podríem trobar possibles explicacions funcionals a algunes de les associacions identificades que no semblaven tenir una connexió clara amb el fenotip estudiat. Vam utilitzar els resultats obtinguts en el GWAS de TDAH abans assenyalat (Demontis *et al.*, 2019), que compta amb una mida mostral força gran (20.183 casos i 35.191 controls) i és el primer GWAS de TDAH amb associacions que superen el llindar de significació a escala genòmica. Primerament vam poder observar que les variants genètiques de risc per al TDAH (en la majoria de llindars de significació considerats) mostraven un enriquiment en variants amb un efecte en la metilació de llocs CpG, resultat que també s'havia observat en GWAS d'altres trastorns psiquiàtrics (Gagliano *et al.*, 2016; Andrews *et al.*, 2017). A continuació vam explorar la nostra llista de 3.896 SNPs en el GWAS de TDAH (Demontis *et al.*, 2019) identificant un total de 60 SNPs associats amb el TDAH que presentaven correlació amb els nivells de metilació de cinc llocs CpG situats a la regió promotora de cinc gens. També vam avaluar si els SNPs eren eQTLs en cervell pels mateixos gens (és a dir, si tenien un impacte sobre la seva expressió). Així, vam observar que tots els SNPs que presentaven correlació amb quatre dels llocs CpG destacats complien aquesta premissa i que, a més, la direcció de l'efecte en l'expressió i la metilació era oposada, tal i com s'espera quan la metilació cau en la regió promotora de gens. Aquest estudi ens va permetre assenyalar tres gens candidats pel TDAH: *ARTN*, *C2orf82* i *PIDD1*, amb funcions diverses. *ARTN* estaria implicat en la supervivència de diferents tipus de neurones, la funció de *C2orf82* encara no està ben caracteritzada, i *PIDD1* s'ha relacionat amb l'apoptosi. També és interessant destacar que tant *ARTN* com *C2orf82* havien estat prèviament relacionats amb altres trastorns psiquiàtrics comòrbids al TDAH (Otsuki *et al.*, 2008; Ripke *et al.*, 2013, 2014; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019), convertint-los en bons candidats a ser estudiats a fons.

Finalment, a partir de totes les evidències obtingudes en els darrers anys on es corrobora la importància de les variants amb un efecte en l'ASM en els diferents trastorns psiquiàtrics, tant en estudis previs (Gamazon *et al.*, 2013; Hannon *et al.*, 2015; Van Eijk *et al.*, 2015; Gagliano *et al.*, 2016; Andrews *et al.*, 2017; Wu i Pan, 2019) com en dos dels estudis inclosos en aquesta Tesi (Article 3, Capítol 2; Article 6, Capítol 3), vam decidir explorar quina era la seva contribució a la base genètica compartida entre diferents trastorns psiquiàtrics (Article 7, Capítol 3) utilitzant una

metanàlisi GWAS que incloïa vuit trastorns psiquiàtrics: TDAH, anorèxia nerviosa, trastorns de l'espectre autista, trastorn bipolar, depressió, trastorn obsessiu-compulsiu, esquizofrènia i síndrome de Tourette (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). És interessant destacar que en aquest estudi GWAS previ s'havien calculat les correlacions genètiques entre totes les parelles possibles d'aquests vuit trastorns, constatant que en molts casos hi ha una base genètica compartida (a nivell de SNPs) molt significativa. Per tant no era estrany pensar que les variants que mostren ASM, que havíem confirmat que tenien una contribució tant en la dependència de drogues com en el TDAH, també contribuirien a la base genètica compartida dels trastorns psiquiàtrics. En aquest estudi (Article 7, Capítol 3), vam tornar a observar un enriquiment de variants que mostren ASM entre les variants genètiques de risc identificades (per tots els llindars de significació avaluats), confirmant la implicació d'aquest mecanisme epigenètic en els trastorns psiquiàtrics. Així, doncs, vam explorar la selecció de 3,896 SNPs realitzada en l'estudi anterior (Article 6, Capítol 3) en els resultats obtinguts en aquesta metanàlisi (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). Mitjançant una estratègia similar a la de l'anterior estudi, vam assenyalar 80 variants associades a com a mínim dos trastorns que mostraven un efecte oposat en la metilació de llocs CpG i l'expressió de gens en cervell (considerant aquells gens que tenen almenys una de les CpG assenyalades a la seva regió promotora). Aquesta estratègia ens va permetre identificar un total de vuit gens candidats a contribuir a la base genètica compartida entre diferents trastorns psiquiàtrics: *ZSCAN29*, *ZSCAN31*, *BTN3A2*, *DDAH2*, *HAPLN4*, *ARTN*, *FAM109B* i *NAGA*, la majoria d'ells prèviament relacionats amb altres trastorns psiquiàtrics (Otsuki *et al.*, 2008; Ripke *et al.*, 2013, 2014; Mokhtari i Lachman, 2016).

Aquests tres estudis (Article 3, Capítol 2; Articles 6 i 7, Capítol 3), ens han permès corroborar la importància de les variants amb un efecte en la metilació de llocs CpG en diferents trastorns psiquiàtrics, tot assenyalant gens candidats per la dependència de drogues, pel TDAH i pels trastorns psiquiàtrics en general. Tot i així, encara és necessari realitzar estudis funcionals *in vivo* i *in vitro* per entendre quina és contribució d'aquests gens i variants assenyalades en els trastorns, tal i com hem fet pel gen *C2orf82*, candidat pel TDAH (Annex Article 6, Capítol 3).

2.2. Variants genètiques en llocs d'unió de miRNAs

Els miRNAs són molècules de RNA no codificants que actuen com a reguladors postranscripcionals de l'expressió gènica a través de la unió per complementarietat de bases a mRNAs (principalment a la regió 3' UTR) per inhibir la seva traducció o induir la seva degradació (Bartel, 2004). Aproximadament un 11% dels SNPs coneguts estan situats a les regions 3' UTR de gens, de manera

que podrien alterar la unió de miRNAs tot alterant la regulació de l'expressió gènica (Moszyńska *et al.*, 2017). Tenint en compte la implicació dels miRNAs en nombrosos mecanismes del sistema nerviós central (Schratt, 2009; Ceman i Saugstad, 2011; Meza-Sosa *et al.*, 2012; Wang, Kwon i Tsai, 2012; Sun i Shi, 2015; Luoni i Riva, 2016), així com el seu paper en el desenvolupament dels canvis neurobiològics induïts per les drogues (Smith i Kenny, 2018), vam decidir explorar la contribució d'aquests SNPs que poden alterar el lloc d'unió de miRNAs en la susceptibilitat a la dependència de drogues (Article 4, Capítol 2).

En aquest estudi vam utilitzar els resultats d'un treball previ que descriu un total de 7000 regions d'unió de miRNAs en cervell que contenen prop de 1000 SNPs (Boudreau *et al.*, 2014). D'entre aquests SNPs vam seleccionar-ne 59 que vam avaluar en un estudi d'associació en dues etapes utilitzant la nostra mostra d'individus dependents de drogues. En aquest estudi vam identificar i replicar una associació de la variant rs1047383, localitzada a la regió 3' UTR del gen *PLCB1*, tant amb la dependència de drogues en general, com amb la dependència de cocaïna en particular. A més, aquesta variant presenta un elevat desequilibri de lligament amb una variant que també alteraria llocs d'unió de miRNAs segons diferents prediccions, rs708910, situada també a la regió 3' UTR del gen *PLCB1*. A continuació vam fer experiments per comprovar si aquests SNPs afectaven la unió de miRNAs i no vam poder observar diferències segons els al·lels en cap dels miRNAs testats. Malgrat tot, no vam poder descartar que aquests SNPs afectessin la unió d'altres miRNAs no testats en aquest estudi, ja que molt sovint, diversos miRNAs reconeixen una mateixa regió en un mRNA concret. Finalment, vam inspeccionar l'expressió del gen *PLCB1* utilitzant dades procedents d'estudis previs i vam observar un increment de la seva expressió en neurones dopaminèrgiques en cultiu tractades amb cocaïna i també en mostres *post-mortem* de nucli accumbens d'individus amb abús o dependència de cocaïna (Albertson *et al.*, 2004), tot donant suport a una contribució del gen *PLCB1* a la dependència de cocaïna. A més, aquest gen ja havia estat prèviament relacionat amb altres trastorns psiquiàtrics i neurològics (Drgon *et al.*, 2011; Udawela *et al.*, 2011; Rita Lo Vasco, Cardinale i Polonia, 2012; Girirajan *et al.*, 2013; St Pourcain *et al.*, 2014; Schoonjans *et al.*, 2016). Així doncs, malgrat que no vam poder demostrar que l'SNP associat alteri la unió de miRNAs, els resultats obtinguts en aquest article i en treballs anteriors ens van permetre assenyalar el gen *PLCB1* com a possible candidat a contribuir a la susceptibilitat a l'addicció a drogues.

2.3. Estudi transcriptòmic en un model d'addicció al menjar

Els models animals han estat àmpliament utilitzats per estudiar trastorns i trets psiquiàtrics humans, entre ells l'addicció a substàncies. Aquests models, juntament amb l'ús de tecnologies

d'anàlisi transcriptòmica d'alt rendiment (com RNAseq), han permès identificar vies moleculars implicades en diferents etapes de l'addicció, com ara la resposta aguda i crònica a les drogues, l'efecte de l'administració de diferents dosis de droga, canvis produïts durant l'abstinència i la recaiguda en el consum.

En un dels estudis inclosos en aquesta Tesi (Article 5, Capítol 2) ens vam proposar identificar factors de susceptibilitat que podien contribuir al desenvolupament de l'addicció al menjar. Per realitzar aquest estudi vam treballar amb una soca consanguínia (*inbred*), és a dir, ratolins genèticament homogenis que van ser exposats a un paradigma d'autoadministració de menjar saborós (pèl·lets amb sabor de xocolata) per poder mimetitzar l'addicció al menjar (Mancino *et al.*, 2015; Domingo-Rodriguez *et al.*, 2020). Tot i la identitat genètica, el comportament dels ratolins envers el menjar no és igual, la qual cosa suggereix que hi ha factors de susceptibilitat diferencials en aquests animals que no estan determinats en la seqüència de nucleòtids, com ara diferències en mecanismes epigenètics que porten a una expressió diferencial de gens.

Per realitzar el nostre estudi vam seleccionar dues poblacions extremes d'animals amb diferent comportament envers al menjar en què vam avaluar si hi havia canvis en l'expressió d'mRNAs (per RNAseq) o miRNAs (smallRNAseq) en dues regions cerebrals que sabem que estan implicades en l'addicció: l'escorça prefrontal i el nucli accumbens. Amb aquesta estratègia vam identificar múltiples gens diferencialment expressats, molts d'ells prèviament associats a trastorns psiquiàtrics o a l'índex de massa corporal. A més, també vam explorar si aquestes diferències d'expressió es podien explicar per canvis dels nivells de metilació de les regions promotores dels gens destacats, ja que és sabut que la metilació en la regió promotora té una relació inversa amb els nivells d'expressió. Així vam observar metilació diferencial a la regió promotora d'alguns d'aquests gens que podrien explicar els canvis d'expressió observats. D'entre tots els gens, destaquem *Adamts2*, *Arsj* i *Klhl40*, ja que mostraven un canvi d'expressió destacable i tenien regions diferencialment metilades en la seva regió promotora que queien en zones del genoma amb marques de cromatina relacionades amb elements reguladors, recolzant la seva possible funció reguladora. *Adamts2* s'encarrega del tall de propèptids de col·lagen de tipus I i II, *Arsj* és un membre de la família de les arilsulfatases, enzims lisosomals, i *Klhl40* forma part d'un complex d'una ubiquïtina lligassa E3 que actua com a regulador clau del desenvolupament del múscul esquelètic. A més, *Adamts2* i *Arsj* havien estat prèviament relacionats amb trets i trastorns psiquiàtrics que tenen relació amb l'addicció al menjar, entre ells el consum de sucre, la impulsivitat motora o la dependència de l'alcohol (Lasky-Su, Anney, *et al.*, 2008; Kalsi *et al.*, 2010; Crespo-Facorro, Prieto i Sainz, 2015; Hwang *et al.*, 2019; Ruso-Julve *et al.*, 2019; Sanchez-Roige *et*

al., 2019). També vam descriure un enriquiment de gens relacionats amb la via de senyalització d'oxitocina. És interessant destacar aquesta via ja que l'oxitocina és un neuropèptid implicat en la regulació de moltes conductes associades a la recompensa (Carter, 2014) que s'ha relacionat àmpliament tant amb l'addicció a les drogues com amb trastorns alimentaris (Sarnyai i Kovács, 1994, 2014; Kovács, Sarnyai i Szabó, 1998; Giel, Zipfel i Hallschmid, 2018; Leong *et al.*, 2018; Leslie *et al.*, 2019). Un mal desenvolupament del sistema d'oxitocina en nens es pot relacionar amb una major susceptibilitat a l'addicció a les drogues en adults (Buisman-Pijlman *et al.*, 2014). En estudis amb animals s'havia observat que l'oxitocina inhibeix el desenvolupament de la tolerància als opioides (Sarnyai i Kovács, 1994; Kovács, Sarnyai i Szabó, 1998), tot suggerint un efecte directe en l'eficàcia de la recompensa als opioides (Panksepp, Knutson i Burgdorf, 2002). A més, el sucre, i probablement també altres menjars saborosos, incrementen els nivells endògens d'opioïdes (Volkow, Wang i Baler, 2011), i agonistes i antagonistes d'opioïdes s'han relacionat amb un increment i una reducció de la ingesta d'aliments saborosos (aliments implicats en l'addicció al menjar), respectivament (Stice *et al.*, 2013). Així doncs, aquesta relació de l'oxitocina amb els opioïdes ens obre les portes a una possible via a ser estudiada per l'addicció al menjar.

2.4. Gens i factors epigenètics de risc identificats en els nostres estudis

Aquests últims anys s'ha vist que la plasticitat neuronal a nivell cel·lular i sinàptic juga un paper essencial en la memòria, l'aprenentatge i en el procés adaptatiu subjacent a les respostes a diferents estímuls (Duman i Newton, 2007). Dins del procés adaptatiu, la formació de records en relació a recompenses és un factor molt important i, a la vegada, també és un component central en diferents trastorns mentals com l'addicció. La formació d'aquests records va lligada a canvis de plasticitat de les neurones dopaminèrgiques situades a l'àrea tegmental ventral (VTA) (Stuber *et al.*, 2008), de manera que les senyals que indiquen futures recompenses (estímuls condicionats) adquireixen la capacitat d'activar les neurones dopaminèrgiques i d'augmentar l'alliberació de dopamina en el sistema mesolímbic (Schultz, Dayan i Montague, 1997; Day *et al.*, 2007; Stuber *et al.*, 2008; Flagel *et al.*, 2011). Hi ha molts mecanismes moleculars que contribueixen a aquests canvis de plasticitat neuronal en resposta a una varietat d'estímuls (comportaments, factors ambientals, fàrmacs, drogues, etc.), que inclouen la regulació de l'expressió gènica, i és aquí on s'ha vist que els diferents mecanismes epigenètics tenen un rol essencial (Duman i Newton, 2007). El punt més important rau en el fet que els processos epigenètics poden canviar de forma dinàmica en resposta a factors externs, sent així el mecanisme clau pel qual l'ambient pot exercir la seva influència sobre l'expressió gènica i per tant també sobre el fenotip (Wong, Mill i Fernandes, 2011). L'epigenètica, doncs, seria el mediador entre gens i ambient, tots dos agents clau en l'etiologia de les malalties complexes.

El sistema de recompensa és un element clau en els trastorns psiquiàtrics, ja que està involucrat en la recompensa, la motivació, la presa de decisions i el control inhibitori (Volkow *et al.*, 2013a), factors centrals en la patologia mental. De fet, molts trastorns psiquiàtrics representen estats en què la satisfacció o l'eufòria són difícils d'obtenir a causa d'un desequilibri dels nivells dels diferents neurotransmissors (dopamina, norepinefrina, GABA, oxitocina, etc.) i neuropèptids (oxitocina, ghrelina, etc.) que fan augmentar el requeriment químic en el sistema mesolímbic per aconseguir el plaer i la recompensa (Neumann i Landgraf, 2012; Michels *et al.*, 2014). És per això que en la majoria de trastorns psiquiàtrics es pot parlar d'un dèficit en el sistema de recompensa que ens porta a un model hipodopaminèrgic (Gondré-Lewis, Bassey i Blum, 2020). En aquesta Tesi ens hem centrat principalment en l'addicció (a drogues i al menjar) i en el TDAH, dos trastorns psiquiàtrics altament comòrbids, en què el sistema de recompensa és clau. En el cas de les addiccions, totes les drogues d'abús i els menjars saborosos actuen directament o indirecta sobre el sistema de recompensa tot produint un increment de l'alliberació de dopamina a l'espai sinàptic, i per això no és estrany que diversos autors hagin proposat que una deficiència del sistema de recompensa pot ser un factor de risc clau per l'addicció a drogues i també al menjar (Comings i Blum, 2000; Volkow, Fowler i Wang, 2003; Blum, Thanos i Gold, 2014; Leyton, 2014). En el cas del TDAH, el model hipodopaminèrgic és una de les teories proposades per explicar la fisiopatologia del trastorn, proposant un dèficit de la transferència d'alliberació de dopamina a l'estímul condicionat (Tripp i Wickens, 2009). Totes aquestes evidències fan que no sigui descabellat pensar que hi ha unes bases epigenètiques específiques de cada trastorn psiquiàtric però també hi hauria unes bases epigenètiques comunes que podrien jugar un paper important en la seva etiologia, i que actuarien com a intermediaris entre els factors ambientals de risc i el genoma (Tsankova *et al.*, 2007; Kofink *et al.*, 2013; Peña *et al.*, 2014).

En aquesta Tesi hem pogut corroborar la implicació de mecanismes epigenètics (metilació de l'ADN i miRNAs) en diferents trastorns psiquiàtrics humans, i també hem pogut observar que un mateix mecanisme epigenètic, com la metilació específica d'al·lel, està implicat en diferents trastorns. L'anàlisi genètica d'aquest mecanisme ens va permetre identificar gens candidats a la dependència de drogues (*CTNBL1*, *SCP2* i *ECHDC2*), al TDAH (*ARTN*, *C2orf82* i *PIDD1*) i als trastorns psiquiàtrics en general (*ZSCAN29*, *ZSCAN31*, *BTN3A2*, *DDAH2*, *HAPLN4*, *ARTN*, *FAM109B* i *NAGA*) (Article 3, Capítol 2; Articles 6 i 7, Capítol 3). En el cas de l'addicció a drogues, també vam poder demostrar la implicació de variants que poden alterar la unió de llocs de miRNAs tot identificant el gen *PLCB1* com a candidat per a aquesta patologia (Article 4, Capítol 2).

Finalment, vam realitzar un estudi sense hipòtesi prèvia a escala genòmica per identificar factors de susceptibilitat genètics i epigenètics per l'addicció al menjar específicament, en aquest cas en un model murí (Article 5, Capítol 2). En aquest estudi vam trobar un nombre força elevat de gens diferencialment expressats a nucli accumbens i a l'escorça prefrontal, que en alguns casos mostraven canvis en els nivells de metilació a la regió promotora, proporcionant així un possible mecanisme per explicar els canvis d'expressió observats. A més, també vam observar un enriquiment de gens diferencialment expressats relacionats amb la via de senyalització d'oxitocina. L'oxitocina és un neuropèptid àmpliament relacionat amb l'addicció a les drogues (Sarnyai i Kovács, 1994, 2014; Kovács, Sarnyai i Szabó, 1998; Jurek i Neumann, 2018) que interacciona directament amb el sistema de recompensa (Love, 2014; Jurek i Neumann, 2018). Així doncs, una alteració de l'expressió de gens de la via de senyalització d'oxitocina podria portar-nos cap a un dèficit del sistema de recompensa (Gondré-Lewis, Bassey i Blum, 2020), que tal i com hem comentat anteriorment, és un element central en la majoria de trastorns psiquiàtrics.

3. Factors genètics de risc a l'addicció a drogues i trastorns comòrbids i estudis funcionals

Gràcies a totes les metodologies explicades anteriorment, en els darrers anys s'han generat molts gens i variants candidates tant per a l'addicció com per a diversos trastorns comòrbids. Però això no és un final de trajecte sinó tot el contrari. S'obren ara moltes vies, començant pels estudis de rèplica que han de confirmar o no els resultats trobats, així com estudis funcionals *in vitro* i *in vivo* per entendre el rol dels gens candidats assenyalats en el trastorn d'estudi. En aquesta Tesi hem realitzat diferents estudis funcionals per variants i gens identificats en estudis previs realitzats per altres grups (Article 1, Capítol 1) o per nosaltres (Article 2, Capítol 1; Annex Article 6, Capítol 3).

Estudi de rèplica de senyals GWAS de dependència de drogues

En els últims anys s'han realitzat múltiples estudis d'associació que han permès identificar noves variants genètiques per a la dependència de drogues, tot i que només algunes d'elles s'han pogut replicar en mostres independents (Bühler *et al*, 2015). Un dels objectius d'aquesta Tesi ha consistit a intentar replicar les senyals obtingudes en diversos GWAS previs de dependència de drogues (p -valor $< 5E-08$) duent a terme un estudi d'associació amb la nostra mostra de dependència de drogues, recollida a l'Hospital Vall d'Hebron. A causa de la metodologia utilitzada en aquests estudis previs, on es realitzen un elevat nombre de tests, podria ser que algunes d'aquestes variants genètiques corresponguin en realitat a falsos positius (Chenock *et al*, 2007; Kraft *et al*, 2009).

En el moment en què vam realitzar l'estudi de rèplica (Article 1, Capítol 1) només s'havien publicat tres GWAS de dependència de drogues en general amb resultats que superaven el llindar de significació a escala genòmica, i que assenyalaven tres variants de risc en individus amb ascendència europea i una variant en individus amb ascendència africana (Johnson *et al.*, 2015). Pel que fa a la dependència de cocaïna, només s'havia descrit una associació en una metanàlisi d'individus d'ascendència europea i africana (Gelernter *et al.*, 2014). Degut a les característiques d'etnicitat de la nostra mostra, vam decidir intentar replicar únicament aquelles variants associades al fenotip en mostres d'ascendència europea. Així doncs, les variants que vam seleccionar van ser rs1868152 i rs2952621, associades amb la dependència de drogues en individus europeus (McGue *et al.*, 2013; Wetherill *et al.*, 2014) però vam descartar la variant rs2567261 ja que els autors de l'estudi no van poder replicar-la en una mostra independent (Wetherill *et al.*, 2014). Per la dependència de cocaïna vam seleccionar l'única variant descrita, rs2629540. De les tres variants avaluades, només vam poder replicar l'associació de l'SNP rs2952621 amb la dependència de drogues sota el model dominant, amb el mateix al·lel de risc, T, en totes les tres mostres estudiades (Wetherill *et al.*, 2014), tal i com es va corroborar amb la metanàlisi realitzada. La relació funcional d'aquesta variant amb la predisposició a la dependència de drogues no està clara, tot i que a través de diverses eines bioinformàtiques sabem que està prop d'un gen encara no caracteritzat, *LINC01854*, en una regió amb marques d'histones relacionades amb *enhancers* i promotors en cervell. Aquests resultats ens tornen a assenyalar la importància d'estudiar la possible contribució de l'epigenètica en l'addicció a les drogues: en aquest cas, podria ser que la variant associada tingui un impacte sobre determinats elements reguladors, tot afectant la regulació de l'expressió gènica.

Tot i que no hem pogut replicar les associacions amb les altres dues variants avaluades, això no descarta necessàriament la seva validesa. Poden haver-hi diferents raons, degudes a limitacions intrínseques dels estudis d'associació que ja s'han discutit, com ara les característiques de la mostra que afecten directament al poder estadístic (heterogeneïtat clínica dels casos, ètnia, efecte modest de la variant avaluada, mida mostral, etc) o el disseny de l'estudi (covariables considerades, mètodes de correcció aplicats, model d'herència considerat, etc), tant en els estudis originals (McGue *et al.*, 2013; Gelernter *et al.*, 2014) com en l'estudi realitzat per nosaltres. Per exemple, la variant rs2629540 fou prèviament associada amb la dependència de cocaïna en un estudi que combinava individus americans d'ascendència europea i africana, mentre que la nostra mostra conté exclusivament individus espanyols d'ascendència europea. D'altra banda, les mostres utilitzades en l'estudi original no mostren un equilibri en el número de casos i controls utilitzats (o un predomini dels controls), sent quatre vegades més gran el número de casos que de

controls, al contrari del patró que segueixen la majoria de dissenys, com el nostre, i que es recomana per tenir un bon poder estadístic (Nsengimana i Bishop, 2017). Finalment, els controls utilitzats en l'estudi previ són individus que han estat prèviament exposats a cocaïna almenys un cop a la seva vida però que no han desenvolupat la dependència, i en canvi els nostres controls són individus de la població general. Totes aquestes diferències en el disseny experimental podrien explicar la no replicació de les senyals, i és per això que no podem concloure clarament que les senyals no replicades siguin falsos positius.

Estudi funcional *in vitro* del gen *NFAT5*, candidat a contribuir a la dependència de drogues

En un estudi previ realitzat pel grup es va assenyalar *NFAT5* com a candidat a la susceptibilitat a la dependència de cocaïna (Fernández-Castillo *et al.*, 2015). *NFAT5* mostrava una expressió incrementada en cèl·lules SH-SY5Y (model neuronal dopaminèrgic) després d'una exposició aguda a cocaïna i es va demostrar que variants dins d'aquest gen estaven associades a la dependència de cocaïna (Fernández-Castillo *et al.*, 2015). Un estudi previ realitzat en cèl·lules del tub renal havia descrit que *NFAT5* induïa la transcripció de *DDC*, que codifica l'enzim encarregat de transformar la L-DOPA en dopamina (Hsin *et al.*, 2011). Tal i com s'ha comentat en la Introducció, l'alliberació de dopamina al sistema mesolímbic implica l'activació de la sensació de plaer i recompensa en resposta a estímuls naturals i addictius, sent doncs un neurotransmissor clau en la dependència de cocaïna (Volkow, Wise i Baler, 2017). En base a tota aquesta informació, ens vam plantejar si *NFAT5* també podia modular l'alliberació de dopamina al cervell a través de la regulació de l'expressió de la isoforma neuronal de *DDC*, la qual cosa reforçaria la seva candidatura a un trastorn psiquiàtric. Per tal de respondre aquesta pregunta vam dur a terme un estudi *in vitro* utilitzant dues línies cel·lulars neurals dopaminèrgiques, PC12 i SH-SY5Y (Article 2, Capítol 1).

En els estudis *in vitro* és molt important controlar la variabilitat entre les mostres i l'ús de rèpliques biològiques i controls, ja que la fiabilitat i replicabilitat dels resultats obtinguts dependrà d'això. Per tal de controlar la variabilitat, tots els experiments es van realitzar sota les mateixes condicions metodològiques, en cultius cel·lulars amb el mateix número de passades cel·lulars i es van utilitzar rèpliques biològiques per a cada condició experimental. Per contrastar la nostra hipòtesi vam induir l'expressió i l'activació d'*NFAT5* (mitjançant un xoc hipertònic) i vam avaluar el possible efecte sobre l'expressió del nostre 'gen problema', *DDC*, i de dues dianes conegudes d'*NFAT5* (*TAUT* i *SMIT*) que ens servien de control positiu. Després del xoc hipertònic vam observar un augment de l'expressió d'*NFAT5* i de les seves dianes *TAUT* i *SMIT*, tal i com s'havia descrit en estudis previs (Loyher *et al.*, 2004; Aramburu *et al.*, 2006; Halterman, Kwon i Wamhoff, 2012) i, curiosament, una disminució de l'expressió de *DDC*. A continuació vam silenciar *NFAT5* utilitzant

un shRNA específic i vam seguir observant la mateixa disminució de l'expressió de *DDC*, tot indicant que el canvi d'expressió era degut a un mecanisme independent d'*NFAT5*. Així doncs, a través d'aquest estudi vam poder concloure que *NFAT5* no regula l'expressió de la isoforma neuronal de *DDC* en cap de les dues línies cel·lulars utilitzades i que per tant no modula l'alliberació de dopamina a través d'aquest mecanisme en aquests models cel·lulars. En base aquests resultats doncs, serien necessaris altres estudis per entendre quin és el rol d'*NFAT5* en l'addicció a cocaïna. De fet, malgrat que la funció principal descrita per *NFAT5* és la de regulació homeostàtica sota estrès osmòtic (Haltermann, Kwon i Wamhoff, 2012), també s'han descrit diferents funcions per aquesta proteïna en cervell com la regulació de l'expressió de *AQP4* en astròcits, la protecció de neurones i astròcits contra danys isquèemics o la participació en la neuroinflamació (Mak *et al.*, 2012; Yi *et al.*, 2013; Jeong *et al.*, 2016; Xia *et al.*, 2017; Yang, Wang i Peng, 2018), funcions que podrien abordar-se per tal d'entendre la contribució d'aquest gen a la dependència de cocaïna.

Model animal genoanul·lat del gen *C2orf82*, candidat pel TDAH

En un dels estudis inclosos en aquesta Tesi (Article 6, Capítol 3) vam assenyalar el gen *C2orf82*, juntament amb dos altres gens, com a candidats a estar implicats en el TDAH. El gen *C2orf82*, també conegut com a *SNORC* (Small NOvel Rich in Cartilage), és un gen molt poc estudiat pel qual la majoria d'estudis s'han centrat en la seva funció en cartílag, on s'ha suggerit que pot estar involucrat en la maduració dels condrocits i l'ossificació endocondral postnatal (Heinonen *et al.*, 2011, 2017), i on s'ha vist que la seva expressió està regulada per *SOX9* (Jaiswal *et al.*, 2020). En aquests estudis també s'ha suggerit una expressió específica d'aquest gen en cartílag mitjançant la tècnica de transferència *Northern* (Heinonen *et al.*, 2011; Jaiswal *et al.*, 2020), tot i que en la present Tesi hem pogut constatar que *C2orf82* també s'expressa al cervell mitjançant la tècnica de qRT-PCR, mostrant les expressions més altes en el nucli accumbens i el nucli estriat dorsal. Aquesta diferència en els resultats probablement pot ser deguda a les diferències entre ambdues tecnologies (Dallas *et al.*, 2005). Concretament, la RT-qPCR és entre 1.000 i 10.000 vegades més sensible que altres tècniques basades en la transferència d'RNA (O'Driscoll *et al.*, 1993; Dean, Goodwin i Hsiang, 2002).

L'any 2011 es va realitzar el primer estudi centrat en la proteïna *C2orf82* (Heinonen *et al.*, 2011). Aquest estudi inclou una predicció de l'estructura secundària i terciària d'aquesta proteïna, que indicava que es tractava d'un proteoglicà de transmembrana curt de tipus I que contenia un pèptid senyal, un domini transmembrana prop de la regió C-terminal i un domini intracel·lular. A més, també es va predir que en el domini extracel·lular hi havia nou treonines/serines propenses a la

modificació O-N-acetylgalactosamina (O-GalNAc) i que hi havia un lloc d'unió de glicosaminoglicans (GAG) (Heinonen *et al.*, 2011). En aquesta Tesi també vam realitzar un estudi de predicció de l'estructura secundària i terciària de la proteïna ja que des del 2011 fins l'actualitat s'han desenvolupat eines bioinformàtiques i bases de dades amb més poder predictiu. Mitjançant aquesta anàlisi vam arribar a conclusions similars, però amb la diferència que ara no quedaria clar que la proteïna contingui un únic domini transmembrana: hem predit una segona hèlix transmembrana a una regió més proximal a l'extrem N-terminal de la proteïna, que tindria associada una probabilitat elevada (Annex de l'Article 6, Capítol 3). Finalment, en un estudi publicat recentment (Jaiswal *et al.*, 2020), s'ha constatat que l'estructura de la proteïna C2orf82 conté un únic domini transmembrana, on l'extrem N-terminal es troba a la regió extracel·lular i l'extrem C-terminal és intracel·lular, confirmant doncs que la segona hèlix predita en les nostres anàlisis realment no existeix.

Tot i que la funció en cervell de C2orf82 és desconeguda, se sap que els proteoglicans són crucials en diferents processos bàsics del neurodesenvolupament i manteniment del SNC, com la proliferació, migració i especificació cel·lular, l'orientació axonal, la sinaptogènesi, la plasticitat i la regeneració (Schwartz i Domowicz, 2018). A més, diferents estudis també indiquen que els proteoglicans poden tenir un paper en diferents trastorns psiquiàtrics com l'esquizofrènia, el trastorn bipolar o la depressió (Goossens *et al.*, 2003; Cichon *et al.*, 2011; Shi *et al.*, 2011; Schwartz i Domowicz, 2018). Pel que fa al gen *C2orf82* específicament, s'ha associat prèviament amb esquizofrènia (Ripke *et al.*, 2013; Li *et al.*, 2017), i a més, també apareix com una de les troballes destacades en un GWAS que inclou vuit trastorns psiquiàtrics. Concretament, en aquest estudi *C2orf82* va ser assenyalat per la variant rs778353 (p -valor = $2,06E-11$, associada amb esquizofrènia i trastorn bipolar (m -value > 0,9)), situada 47 kB *downstream* del gen (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). Totes aquestes evidències, juntament amb els resultats de l'Article 6 del Capítol 3 d'aquesta Tesi, converteixen el gen *C2orf82* en un bon candidat per ser estudiat en un model animal, com hem fet.

En la caracterització fenotípica dels ratolins genoanul·lats per *C2orf82* vam realitzar tant tests de caracterització fenotípica general com tests específics per trets i trastorns psiquiàtrics. En la fenotipació general vam poder confirmar que els ratolins no mostraven problemes de salut general, ni anhedònia (incapacitat per sentir plaer), ni problemes de coordinació motora. A més, també vam poder observar hipolocomoció en els ratolins KO en comparar-los amb els WT en diversos tests realitzats, així com un possible fenotip d'ansietat i una menor curiositat exploratòria. En la *Continuous Performance Task*, un test específic per avaluar diferents tipus

d'atenció així com la impulsivitat, tots els ratolins van mostrar una capacitat d'aprenentatge correcta, i no es van observar diferències en l'atenció però sí una major impulsivitat per part dels ratolins KO en comparació amb els WT. Finalment, en els resultats d'autoadministració de cocaïna vam observar que els ratolins KO mostraven una menor motivació per la droga. També vam observar diferències en la fase de restabliment de la resposta a la droga, que simula la situació de recaiguda dels humans. En aquesta etapa vam observar que tots els ratolins, sense importar el genotip, recordaven l'estímul associat als efectes reforçants de la cocaïna, però que tant els HET com els homozigots KO cessaven abans el comportament de cerca de la droga, fet que es pot interpretar com una disminució de la probabilitat de recaiguda. Aquests resultats obtinguts en el test d'autoadministració de cocaïna anirien en consonància amb la teoria hipodopaminèrgica discutida anteriorment, de manera que els ratolins KO pel gen *C2orf82*, igual que els individus diagnosticats amb TDAH, poden presentar aquesta disfunció del sistema de recompensa, tal i com s'observa amb la menor motivació per part dels ratolins KO per aconseguir la recompensa (cocaïna) quan això implica un esforç extra. A més, el dèficit de transferència d'alliberació de dopamina a l'estímul condicionat també explicaria el cessament primerenc del comportament de cerca de la droga, en què s'estimula l'animal únicament amb l'estímul condicionat. Tot i així, el número d'animals utilitzats encara és molt limitat, i per tant hauríem d'incrementar la mida mostral per poder confirmar els fenotips descrits.

4. Perspectives de futur

Al llarg dels anys s'han utilitzat múltiples aproximacions per estudiar base genètica de les addiccions i trastorns comòrbids. Al principi aquestes eren estratègies basades en hipòtesis, que poc a poc han evolucionat cap a metodologies sense hipòtesi prèvia, com els GWAS o els estudis transcriptòmics i epigenòmics, que han permès identificar centenars de variants i gens prèviament no relacionats amb els trastorns.

En l'inici de l'era dels GWAS, fa prop de 15 anys, els trastorns que s'estudiaven comptaven amb mostres d'uns milers d'individus, que en molts casos no van permetre detectar senyals d'associació que superessin el llindar de significació genòmica, probablement perquè l'etiologia és poligènica amb efectes menors o moderats de cadascuna de les variants implicades. Així, va sorgir la necessitat d'incrementar les mides mostrals d'aquests estudis, i amb aquest objectiu es van començar a crear grans consorcis internacionals com el *Psychiatric Genomics Consortium* (PGC, <https://www.med.unc.edu/pgc/>) que han permès la realització de GWAS i metanàlisis en mostres molt més grans (desenes o centenars de milers d'individus afectats) que difícilment s'haurien pogut reclutar localment. El PGC treballa en aquests moments amb onze trastorns

psiquiàtrics diferents, entre ells l'addicció a drogues i el TDAH (Sullivan *et al.*, 2018). L'última metanàlisi publicada en TDAH comptava amb uns 20.000 casos i 35.000 controls i es van identificar 12 associacions que superaven el llindar de significació (Demontis *et al.*, 2019). En el passat *World Congress of Psychiatric Genetics* (2019) es van presentar els resultats d'un nou estudi GWAS en TDAH que supera els cent mil individus entre casos i controls, i que incrementa el número d'associacions significatives a prop de 100. Pel que fa a l'addicció a drogues, tot i que s'han fet estudis destacables per a l'addicció a l'alcohol (Walters *et al.*, 2018; Kranzler *et al.*, 2019), encara falta molt camí per recórrer en relació a l'addicció a les drogues il·legals, un fenotip poc estudiat a nivell genètic. L'últim GWAS publicat per aquest fenotip va utilitzar una mostra de més de 4.000 americans d'origen europeu o africà amb dependència de l'alcohol o a drogues il·lícites, i uns 6.200 controls, identificant només tres variants genètiques associades al fenotip considerant només els individus americans d'origen africà (Wetherill *et al.*, 2019). Tot i així, el poder estadístic segueix sent limitat degut a la mida mostral, i és per això que un dels objectius del PGC en un futur proper és recollir al voltant de 10.000 casos amb dependència de cocaïna i els controls corresponents per poder fer un GWAS amb suficient poder estadístic. A més, aquest consorci no només ha treballat amb la genètica específica de cada trastorn, sinó que també ha realitzat estudis per explorar les bases genètiques compartides entre els trastorns psiquiàtrics (Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.*, 2013; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2019). Els resultats mostren de forma molt clara que el grau de solapament genètic entre trastorns és important, i això podria explicar, almenys en part, la comorbiditat que s'observa entre diferents malalties psiquiàtriques. Fins ara, la majoria d'estudis s'han centrat en identificar les bases genètiques de trastorns psiquiàtrics específics, però aquests resultats mostren que també és interessant buscar quines són les bases genètiques compartides entre ells.

L'era GWAS doncs, ha permès la identificació de centenars de *loci* que s'han associat al risc de patir diferents trastorns psiquiàtrics. Tot i això, hi ha diversos factors que poden complicar trobar la relació entre les variants genètiques identificades i el trastorn associat, així com entendre l'impacte funcional d'aquestes variants. En primer lloc, l'associació d'un *locus* amb una malaltia no específica quina variant (o variants) és en realitat la causal, ni quin gen (o gens) està afectat per la variant causal. En conseqüència, és necessària una validació empírica per determinar quines de les variants associades són realment funcionals (Schaub *et al.*, 2012; Edwards *et al.*, 2013). I és aquí on l'epigenètica ens pot ajudar a entendre l'impacte funcional d'almenys part d'aquestes variants. Però l'ús de l'epigenètica no acaba aquí, ja que també pot assenyalar biomarcadors molt robustos per poder classificar els pacients més enllà de l'expressió fenotípica que presenten, i així

poder adequar millor el tractament que han de rebre (Kular i Kular, 2018). Finalment, també s'han assenyalat com a possibles dianes terapèutiques, recolzant-se en resultats obtinguts en models animals (Kular i Kular, 2018; Nestler i Lüscher, 2019). Un exemple són els inhibidors de desacetilases d'histones (HDAC), que s'han proposat com a possible tractament tant per la dependència de drogues com per la depressió (Fuchikami *et al.*, 2016; Misztak, Pańczyszyn-Trzewik i Sowa-Kućma, 2018; Nestler i Lüscher, 2019). Tot i així, el problema principal que presenten aquestes molècules és que tenen com a diana proteïnes d'expressió ubíqua, cosa que podria conduir a efectes secundaris fora de l'objectiu desitjat (Nestler i Lüscher, 2019). Així doncs, encara fa falta molta recerca en aquest camp per assenyalar molècules amb dianes més específiques.

Finalment, un cop s'han identificat les variants causals i quin pot ser el seu impacte funcional o quin és el gen afectat per aquestes variants, cal anar un pas més enllà i estudiar a fons quina és la seva connexió amb el fenotip mitjançant estudis funcionals, tant a nivell cel·lular com en models animals. El ratolí és normalment el model seleccionat ja que el seu genoma presenta una gran identitat amb l'humà (superior al 95%), a banda que hi ha nombroses tècniques de manipulació genètica que permeten mimetitzar el trastorn estudiat (Ermann i Glimcher, 2012) i un gran nombre de tests per estudiar el comportament dels ratolins. El peix zebra i la *Drosophila* són altres organismes model que cada vegada són més populars perquè presenten una sèrie d'avantatges respecte els ratolins: la facilitat de manipulació genètica, la capacitat de produir ràpidament gran quantitat d'organismes amb un genotip específic i la capacitat d'estudi de l'expressió gènica específica de teixit en animals vius (Edwards *et al.*, 2013). Malgrat que ambdues espècies s'han utilitzat per estudiar trastorns psiquiàtrics o conductes complexes lligades a aquests trastorns (Kalueff *et al.*, 2013; Van Alphen i Van Swinderen, 2013), el nombre de tests existents per estudiar el comportament d'aquests models és encara limitat. Així doncs, en els pròxims anys la recerca també s'ha de centrar en idear nous tests per poder millorar la caracterització fenotípica d'aquests models.

En resum, el panorama científic està canviant, primerament amb l'aparició de les noves tecnologies de genotipació i seqüenciació a escala genòmica i a continuació amb la integració de dades epigenòmiques per entendre la funcionalitat de les variants destacades en aquests estudis. Tot i això, encara queden moltes incògnites que cal abordar en els propers anys per tenir un coneixement complet de la genètica de les addiccions i dels trastorns comòrbids, així com de la genètica compartida entre ells.

CONCLUSIONS

Study of genetic risk factors for drug addiction identified previously

1. The variant rs2952621, previously identified in a GWAS study of drug dependence, has been replicated in our Spanish sample.
2. NFAT5, previously identified by us as involved in cocaine dependence, does not modulate dopamine synthesis through *DDC* expression regulation in dopaminergic neural cell lines.

Identification of genes and epigenetic risk factors for drug and food addiction

3. We have identified and replicated variants with an epigenetic effect (methylation or miRNA regulation) that contribute to drug dependence susceptibility.
4. The *CTNBL1*, *SCP2*, *ECHDC2* and *PLCB1* genes have been pointed by these epigenetic variants as possible contributors to drug dependence.
5. There are specific mRNA or miRNA expression profiles in the NAc and mPFC of mice that are related to vulnerability/resilience to food addiction, some of them explained by changes in methylation.
6. The oxytocin signalling pathway, previously related to addiction and eating disorders, is highlighted as a possible contributor to food addiction.

Identification of genes and epigenetic risk factors for psychiatric disorders that are comorbid to drug addiction

7. Genetic risk factors for ADHD are enriched in variants influencing methylation (ASM). Some of them alter the expression of the *ARTN*, *C2orf82* and *PIDD1* genes.
8. The *C2orf82* gene is expressed in the mouse brain with the highest expression in the NAc, and its depletion in mice causes hypolocomotion, impulsivity and decreased motivation for cocaine.
9. We have identified SNPs influencing methylation and expression of eight genes in brain that show pleiotropic effects on several psychiatric disorders.

BIBLIOGRAFIA

- Abou-Sleiman, P. M., Hanna, M. G. i Wood, N. W. (2006) «Genetic association studies of complex neurological diseases», *Journal of Neurology, Neurosurgery and Psychiatry*, 77(12), p. 1302-1304. doi: 10.1136/jnnp.2005.082024.
- Afonso-Grunz, F. i Müller, S. (2015) «Principles of miRNA-mRNA interactions: Beyond sequence complementarity», *Cellular and Molecular Life Sciences*, 72(16), p. 3127-3141. doi: 10.1007/s00018-015-1922-2.
- Agrawal, A. *et al.* (2006) «Association of GABRA2 with drug dependence in the collaborative study of the genetics of alcoholism sample», *Behavior Genetics*, 7, 36(5), p. 640-650. doi: 10.1007/s10519-006-9069-4.
- Agrawal, A. *et al.* (2012) «The genetics of addiction-a translational perspective», *Translational Psychiatry*, 2, p. e140. doi: 10.1038/tp.2012.54.
- Al-Eitan, L. N. *et al.* (2014) «Characterization of serotonin transporter gene (SLC6A4) polymorphisms and its association with drug dependence in a Jordanian Arab population», *Toxicology and Industrial Health*, 30(7), p. 598-610. doi: 10.1177/0748233712462446.
- Albertson, D. N. *et al.* (2004) «Gene expression profile of the nucleus accumbens of human cocaine abusers: evidence for dysregulation of myelin», *Journal of neurochemistry*, 88(5), p. 1211-1219. doi: 10.1046/J.1471-4159.2003.02247.X.
- Allis, C. D. i Jenuwein, T. (2016) «The molecular hallmarks of epigenetic control», *Nature Reviews Genetics*, 17(8), p. 487-500. doi: 10.1038/nrg.2016.59.
- Almagro Armenteros, J. J. *et al.* (2019) «SignalP 5.0 improves signal peptide predictions using deep neural networks», *Nature Biotechnology*, 37(4), p. 420-423. doi: 10.1038/s41587-019-0036-z.
- Van Alphen, B. i Van Swinderen, B. (2013) «Drosophila strategies to study psychiatric disorders», *Brain Research Bulletin*, 92, p. 1-11. doi: 10.1016/j.brainresbull.2011.09.007.
- De Alwis, D. *et al.* (2014) «Attention-deficit/hyperactivity disorder subtypes and substance use and use disorders in NESARC», *Addictive Behaviors*, 39(8), p. 1278-1285. doi: 10.1016/j.addbeh.2014.04.003.
- American Psychiatric Association (2013) *Diagnostic and statistical manual of mental disorders (5th ed.)*, Arlington, VA: American Psychiatric Publishing.
- Andrews, S. V. *et al.* (2017) «Cross-tissue integration of genetic and epigenetic data offers insight into autism spectrum disorder», *Nature Communications*, 8, p. 1011.
- Anier, K. *et al.* (2010) «DNA methylation regulates cocaine-induced behavioral sensitization in mice», *Neuropsychopharmacology*, 35(12), p. 2450-2461. doi: 10.1038/npp.2010.128.
- Aoki, Y., Cortese, S. i Castellanos, F. X. (2018) «Diffusion tensor imaging studies of attention-deficit/hyperactivity disorder: meta-analyses and reflections on head motion», *Journal of child psychology and psychiatry, and allied disciplines*, 59(3), p. 193-202. doi: 10.1111/jcpp.12778.
- Aramburu, J. *et al.* (2006) «Regulation of the hypertonic stress response and other cellular functions by the Rel-like transcription factor NFAT5», *Biochemical Pharmacology*, 72(11), p. 1597-1604.
- Bahi, A., Chandrasekar, V. i Dreyer, J. L. (2014) «Selective lentiviral-mediated suppression of microRNA124a in the hippocampus evokes antidepressants-like effects in rats», *Psychoneuroendocrinology*, 46, p. 78-87. doi: 10.1016/j.psyneuen.2014.04.009.
- Bai, M. *et al.* (2012) «Abnormal Hippocampal BDNF and miR-16 Expression Is Associated with Depression-Like Behaviors Induced by Stress during Early Life», *PLoS ONE*, 7(10), p. e46921. doi: 10.1371/journal.pone.0046921.
- Bai, M. *et al.* (2014) «Anhedonia was associated with the dysregulation of hippocampal HTR4 and microRNA Let-7a in rats», *Physiology and Behavior*, 129, p. 135-141. doi: 10.1016/j.physbeh.2014.02.035.
- Baik, J. H. (2013) «Dopamine signaling in reward-related behaviors», *Frontiers in Neural Circuits*, 7, p. 152. doi: 10.3389/fncir.2013.00152.

Bibliografia

- Bali, P. i Kenny, P. J. (2013) «MicroRNAs and drug addiction», *Frontiers in Genetics*, 4, p. 43. doi: 10.3389/fgene.2013.00043.
- Barkus, E. i Murray, R. M. (2010) «Substance Use in Adolescence and Psychosis: Clarifying the Relationship», *Annual Review of Clinical Psychology*, 6(1), p. 365-389. doi: 10.1146/annurev.clinpsy.121208.131220.
- Bartel, D. P. (2004) «MicroRNAs: Genomics, Biogenesis, Mechanism, and Function», *Cell*, 116(2), p. 281-297. doi: 10.1016/S0092-8674(04)00045-5.
- Benton, M. C. *et al.* (2019) «Genome-wide allele-specific methylation is enriched at gene regulatory regions in a multi-generation pedigree from the Norfolk Island isolate», *Epigenetics and Chromatin*, 12, p. 60. doi: 10.1186/s13072-019-0304-7.
- Berridge, K. C. (2009) «“Liking” and “wanting” food rewards: Brain substrates and roles in eating disorders», *Physiology and Behavior*, 97(5), p. 537-550. doi: 10.1016/j.physbeh.2009.02.044.
- Bevilacqua, L. i Goldman, D. (2013) «Genetics of impulsive behaviour», *Philosophical Transactions of the Royal Society B*, 368(1615), p. 20120380. doi: 10.1098/rstb.2012.0380.
- Biederman, J. *et al.* (1999) «Pharmacotherapy of attention-deficit/hyperactivity disorder reduces risk for substance use disorder», *Pediatrics*, 104(2), p. e20. doi: 10.1542/peds.104.2.e20.
- Bierut, L. J. *et al.* (2010) «A genome-wide association study of alcohol dependence», *PNAS*, 107(11), p. 5082-5087. doi: 10.1073/pnas.0911109107.
- Bird, A. P. i Wolffe, A. P. (1999) «Methylation-induced repression-belts, braces, and chromatin», *Cell*, 99(5), p. 451-454. doi: 10.1016/S0092-8674(00)81532-9.
- Björklund, A. i Dunnett, S. B. (2007) «Fifty years of dopamine research», *Trends in Neurosciences*, 30(5), p. 185-187. doi: 10.1016/j.tins.2007.03.004.
- Blum, K. *et al.* (2008) «Attention-deficit-hyperactivity disorder and reward deficiency syndrome», *Neuropsychiatric Disease and Treatment*, 4(5), p. 893-917. doi: 10.2147/ndt.s2627.
- Blum, K., Thanos, P. K. i Gold, M. S. (2014) «Dopamine and glucose, obesity and reward deficiency syndrome», *Frontiers in Psychology*, 5, p. 919. doi: 10.3389/fpsyg.2014.00919.
- Bodetto, S. P. *et al.* (2013) «Cocaine represses protein phosphatase-1C β through DNA methylation and Methyl-CpG Binding Protein-2 recruitment in adult rat brain», *Neuropharmacology*, 73, p. 31-40. doi: 10.1016/j.neuropharm.2013.05.005.
- Bönsch, D. *et al.* (2012) «Methylation matters? Decreased methylation status of genomic DNA in the blood of schizophrenic twins», *Psychiatry Research*, 198(3), p. 533-537. doi: 10.1016/j.psychres.2011.09.004.
- Botia, B. *et al.* (2012) «Expression of ethanol-induced behavioral sensitization is associated with alteration of chromatin remodeling in mice», *PLoS ONE*, 7(10), p. e47527. doi: 10.1371/journal.pone.0047527.
- Boudreau, R. L. *et al.* (2014) «Transcriptome-wide discovery of microRNA binding sites in Human Brain», *Neuron*, 81(2), p. 294-305. doi: 10.1016/j.neuron.2013.10.062.
- Boyle, E. A., Li, Y. I. i Pritchard, J. K. (2017) «An Expanded View of Complex Traits: From Polygenic to Omnigenic», *Cell*, 169(7), p. 1177-1186. doi: 10.1016/j.cell.2017.05.038.
- Brady, K. T. *et al.* (2013) «Substance use disorders and anxiety: A treatment challenge for social workers», *Social Work in Public Health*, 28, p. 407-423. doi: 10.1080/19371918.2013.774675.
- Breschi, A., Gingeras, T. R. i Guigó, R. (2017) «Comparative transcriptomics in human and mouse», *Nature Reviews Genetics*, 18(7), p. 425-440. doi: 10.1038/nrg.2017.19.
- Bühler, K.-M. *et al.* (2015) «Common single nucleotide variants underlying drug addiction: more than a decade of research», *Addiction Biology*, 20, p. 845-871.
- Buisman-Pijlman, F. T. A. *et al.* (2014) «Individual differences underlying susceptibility to addiction: Role for the endogenous oxytocin system», *Pharmacology Biochemistry and Behavior*, 119, p. 22-38. doi: 10.1016/j.pbb.2013.09.005.

- Caballero-Puntiverio, M. *et al.* (2019) «Effect of ADHD medication in male C57BL/6J mice performing the rodent Continuous Performance Test», *Psychopharmacology*, 236(6), p. 1839-1851. doi: 10.1007/s00213-019-5167-x.
- Cabana-Domínguez, J. *et al.* (2019) «Genome-wide association meta-analysis of cocaine dependence: Shared genetics with comorbid conditions», *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 94, p. 109667. doi: 10.1016/j.pnpbp.2019.109667.
- Camchong, J., Endres, M. i Fein, G. (2014) «Decision making, risky behavior, and alcoholism», en Sullivan, E. V. i Pfefferbaum, A. (ed.) *Handbook of Clinical Neurology: Alcohol and the Nervous System*, p. 227-236. doi: 10.1016/B978-0-444-62619-6.00014-8.
- Capusan, A. J. *et al.* (2017) «Genetic and environmental aspects in the association between attention-deficit hyperactivity disorder symptoms and binge-eating behavior in adults: A twin study», *Psychological Medicine*, 47(16), p. 2866-2878. doi: 10.1017/S0033291717001416.
- Carey, M. P. *et al.* (1998) «Differential distribution, affinity and plasticity of dopamine D-1 and D-2 receptors in the target sites of the mesolimbic system in an animal model of ADHD», *Behavioural Brain Research*, 94(1), p. 173-185. doi: 10.1016/S0166-4328(97)00178-2.
- Carter, A. *et al.* (2016) «The Neurobiology of “Food Addiction” and Its Implications for Obesity Treatment and Policy», *Annual Review of Nutrition*, 36(1), p. 105-128. doi: 10.1146/annurev-nutr-071715-050909.
- Carter, C. S. (2014) «Oxytocin Pathways and the Evolution of Human Behavior», *Annual Review of Psychology*, 65(1), p. 17-39. doi: 10.1146/annurev-psych-010213-115110.
- Catalanotto, C., Cogoni, C. i Zardo, G. (2016) «MicroRNA in control of gene expression: An overview of nuclear functions», *International Journal of Molecular Sciences*, 17(10), p. 1712. doi: 10.3390/ijms17101712.
- Ceman, S. i Saugstad, J. (2011) «MicroRNAs: Meta-controllers of gene expression in synaptic activity emerge as genetic and diagnostic markers of human disease», *Pharmacology and Therapeutics*, 130(1), p. 26-37. doi: 10.1016/j.pharmthera.2011.01.004.
- Cerdá, M. *et al.* (2010) «Genetic and environmental influences on psychiatric comorbidity: A systematic review», *Journal of Affective Disorders*, 126, p. 14-38. doi: 10.1016/j.jad.2009.11.006.
- Chandrasekar, V. i Dreyer, J. L. (2009) «microRNAs miR-124, let-7d and miR-181a regulate Cocaine-induced Plasticity», *Molecular and Cellular Neuroscience*, 42(4), p. 350-362. doi: 10.1016/j.mcn.2009.08.009.
- Chandrasekar, V. i Dreyer, J. L. (2011) «Regulation of MiR-124, Let-7d, and MiR-181a in the accumbens affects the expression, extinction, and reinstatement of cocaine-induced conditioned place preference», *Neuropsychopharmacology*, 36(6), p. 1149-1164. doi: 10.1038/npp.2010.250.
- Chanock, S. J. *et al.* (2007) «Replicating genotype-phenotype associations», *Nature*, 447, p. 655-660. doi: 10.1038/447655a.
- Chase, K. A. *et al.* (2013) «Histone methylation at H3K9: Evidence for a restrictive epigenome in schizophrenia», *Schizophrenia Research*, 149, p. 15-20. doi: 10.1016/j.schres.2013.06.021.
- Chen, A. C. H. *et al.* (2002) «Potentially functional polymorphism in the promoter region of prodynorphin gene may be associated with protection against cocaine dependence or abuse», *American Journal of Medical Genetics*, 114(4), p. 429-435. doi: 10.1002/ajmg.10362.
- Chen, L. *et al.* (2016) «A systematic review and meta-analysis of tract-based spatial statistics studies regarding attention-deficit/hyperactivity disorder», *Neuroscience and Biobehavioral Reviews*, p. 838-847. doi: 10.1016/j.neubiorev.2016.07.022.
- Chen, Q. *et al.* (2018) «Common psychiatric and metabolic comorbidity of adult attention-deficit/hyperactivity disorder: A population-based cross-sectional study», *PLoS ONE*, 13(9), p. e0204516. doi: 10.1371/journal.pone.0204516.
- Chen, W. *et al.* (2008) «DSM-IV combined type ADHD shows familial association with sibling trait scores: A sampling strategy for QTL linkage», *American Journal of Medical Genetics Part B*, 147B, p. 1450-1460. doi: 10.1002/ajmg.b.30672.

Bibliografia

- Di Chiara, G. *et al.* (2004) «Dopamine and drug addiction: The nucleus accumbens shell connection», *Neuropharmacology*, 47, p. 227-241.
- Chuang, L.-C. *et al.* (2013) «Pathway analysis using information from allele-specific gene methylation in genome-wide association studies for bipolar disorder», *PLoS ONE*, 8(1), p. e53092.
- Cichon, S. *et al.* (2011) «Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder», *American Journal of Human Genetics*, 88(3), p. 372-381. doi: 10.1016/j.ajhg.2011.01.017.
- Clarke, T.-K. *et al.* (2013) «Further evidence for association of polymorphisms in the CNR1 gene with cocaine addiction: confirmation in an independent sample and meta-analysis», *Addiction biology*, 18(4), p. 702-708.
- Clarke, T.-K. *et al.* (2017) «Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N=112 117)», *Molecular psychiatry*, 22, p. 1376-1384. doi: 10.1038/mp.2017.153.
- Collins, A. L. *et al.* (2004) «Mild overexpression of MeCP2 causes a progressive neurological disorder in mice», *Human Molecular Genetics*, 13(21), p. 2679-2689. doi: 10.1093/hmg/ddh282.
- Comings, D. E. *et al.* (1994) «The dopamine D2 receptor gene: a genetic risk factor in substance abuse», *Drug and alcohol dependence*, 34(3), p. 175-180. doi: 10.1016/0376-8716(94)90154-6.
- Comings, D. E. i Blum, K. (2000) «Reward deficiency syndrome: genetic aspects of behavioral disorders», *Progress in brain research*, 126, p. 325-341.
- Cooper, S., Robison, A. J. i Mazei-Robison, M. S. (2017) «Reward Circuitry in Addiction», *Neurotherapeutics*, 14(3), p. 687-697. doi: 10.1007/s13311-017-0525-z.
- Cornelis, M. C. *et al.* (2016) «A genome-wide investigation of food addiction», *Obesity*, 24(6), p. 1336-1341. doi: 10.1002/oby.21476.
- Cortese, S. *et al.* (2016) «Functional Decoding and Meta-analytic Connectivity Modeling in Adult Attention-Deficit/Hyperactivity Disorder», *Biological psychiatry*, 80(12), p. 896-904. doi: 10.1016/j.biopsych.2016.06.014.
- Costa-Silva, J., Domingues, D. i Lopes, F. M. (2017) «RNA-Seq differential expression analysis: An extended review and a software tool», *PLoS ONE*, 12, p. e0190152. doi: 10.1371/journal.pone.0190152.
- Costa, E. *et al.* (2007) «Reviewing the Role of DNA (Cytosine-5) Methyltransferase Overexpression in the Cortical GABAergic Dysfunction Associated with Psychosis Vulnerability», *Epigenetics*, 2(1), p. 29-36. doi: 10.4161/epi.2.1.4063.
- Crabbe, J. C. *et al.* (1996) «Elevated alcohol consumption in null mutant mice lacking 5HT(1B) serotonin receptors», *Nature Genetics*, 14(1), p. 98-101. doi: 10.1038/ng0996-98.
- Crespo-Facorro, B., Prieto, C. i Sainz, J. (2015) «Schizophrenia Gene Expression Profile Reverted to Normal Levels by Antipsychotics», *International Journal of Neuropsychopharmacology*, 18, p. pyu066.
- Cross-Disorder Group of the Psychiatric Genomics Consortium *et al.* (2013) «Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs», *Nature genetics*, 45(9), p. 984-994.
- Cross-Disorder Group of the Psychiatric Genomics Consortium (2019) «Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders», *Cell*, 179(7), p. 1469-1482. doi: 10.1016/j.cell.2019.11.020.
- Cubells, J. F. *et al.* (2000) «A haplotype at the DBH locus, associated with low plasma dopamine β -hydroxylase activity, also associates with cocaine-induced paranoia», *Molecular Psychiatry*, 5(1), p. 56-63. doi: 10.1038/sj.mp.4000657.
- Cunningham, C. L., Gremel, C. M. i Groblewski, P. A. (2006) «Drug-induced conditioned place preference and aversion in mice», *Nature Protocols*, 1(4), p. 1662-1670. doi: 10.1038/nprot.2006.279.
- D'Addario, C. *et al.* (2014) «Endocannabinoid signaling and food addiction», *Neuroscience and Biobehavioral*

- Reviews*, 47(1), p. 203-224. doi: 10.1016/j.neubiorev.2014.08.008.
- Dahl, J. P. *et al.* (2006) «Analysis of variations in the tryptophan hydroxylase-2 (TPH2) gene in cocaine dependence», *Addiction Biology*, 11(1), p. 76-83. doi: 10.1111/j.1369-1600.2006.00005.x.
- Dallas, P. B. *et al.* (2005) «Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR - How well do they correlate?», *BMC Genomics*, 6, p. 59. doi: 10.1186/1471-2164-6-59.
- Davis, C. i Bonder, R. (2019) «Genetics and epigenetics of food addiction», en Cottone, P. *et al.* (ed.) *Compulsive Eating Behavior and Food Addiction: Emerging Pathological Constructs*. York University, Toronto, ON, Canada: Academic Press, p. 293-327. doi: 10.1016/B978-0-12-816207-1.00010-X.
- Day, J. J. *et al.* (2007) «Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens», *Nature Neuroscience*, 10(8), p. 1020-1028. doi: 10.1038/nn1923.
- Deacon, R. M. J. (2013) «Measuring motor coordination in mice», *Journal of visualized experiments : JoVE*, (75). doi: 10.3791/2609.
- Deak, J. D., Miller, A. P. i Gizer, I. R. (2019) «Genetics of alcohol use disorder: a review», *Current Opinion in Psychology*, 27, p. 56-61. doi: 10.1016/j.copsy.2018.07.012.
- Dean, J. D., Goodwin, P. H. i Hsiang, T. (2002) «Comparison of relative RT-PCR and northern blot analyses to measure expression of β -1,3-glucanase in *Nicotiana benthamiana* infected with *Colltotrichum destructivum*», *Plant Molecular Biology Reporter*, 20(4), p. 347-356. doi: 10.1007/BF02772122.
- Demontis, D. *et al.* (2019) «Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder», *Nature Genetics*, 51(1), p. 63-75.
- Deng, J. V. *et al.* (2014) «MeCP2 phosphorylation limits psychostimulant-induced behavioral and neuronal plasticity», *Journal of Neuroscience*, 34(13), p. 4519-4527. doi: 10.1523/JNEUROSCI.2821-13.2014.
- Deutch, A. Y. i Roth, R. H. (1990) «The determinants of stress-induced activation of the prefrontal cortical dopamine system», *Progress in Brain Research*, 85, p. 367-403. doi: 10.1016/S0079-6123(08)62691-6.
- Dilger, S. *et al.* (2003) «Brain activation to phobia-related pictures in spider phobic humans: an event-related functional magnetic resonance imaging study», *Neuroscience letters*, 348(1), p. 29-32. doi: 10.1016/s0304-3940(03)00647-5.
- Do, C. *et al.* (2017) «Genetic-epigenetic interactions in cis: a major focus in the post-GWAS era», *Genome biology*, 18(1), p. 120.
- Domcke, S. *et al.* (2015) «Competition between DNA methylation and transcription factors determines binding of NRF1», *Nature*, 528, p. 575-579. doi: 10.1038/nature16462.
- Domingo-Rodríguez, L. *et al.* (2020) «A specific prelimbic-nucleus accumbens pathway controls resilience versus vulnerability to food addiction», *Nature Communications*, 11(1), p. 782. doi: 10.1038/s41467-020-14458-y.
- Dong, E. *et al.* (2007) «Histone hyperacetylation induces demethylation of reelin and 67-kDa glutamic acid decarboxylase promoters», *PNAS*, 104(11), p. 4676-4681. doi: 10.1073/pnas.0700529104.
- Drgon, T. *et al.* (2011) «“Replicated” genome wide association for dependence on illegal substances: genomic regions identified by overlapping clusters of nominally positive SNPs», *American Journal of Medical Genetics Part B*, 156(2), p. 125-138. doi: 10.1002/ajmg.b.31143.
- Drozdetskiy, A. *et al.* (2015) «JPred4: A Protein Secondary Structure Prediction Server», *Nucleic Acids Research*, 43, p. W389-W394.
- Ducci, F. i Goldman, D. (2012) «The genetic basis of addictive disorders», *The Psychiatric clinics of North America*, 35(2), p. 495-519.
- Duman, R. S. i Newton, S. S. (2007) «Epigenetic Marking and Neuronal Plasticity», *Biological Psychiatry*, 62, p. 1-3. doi: 10.1016/j.biopsych.2007.04.037.

Bibliografia

- Ebejer, J. L. *et al.* (2013) «Genome-wide association study of inattention and hyperactivity-impulsivity measured as quantitative traits», *Twin Research and Human Genetics*, 16(2), p. 560-574. doi: 10.1017/thg.2013.12.
- Edwards, S. L. *et al.* (2013) «Beyond GWASs: Illuminating the dark road from association to function», *American Journal of Human Genetics*, 93(5), p. 779-797. doi: 10.1016/j.ajhg.2013.10.012.
- Egervari, G. *et al.* (2018) «Shaping vulnerability to addiction – the contribution of behavior, neural circuits and molecular mechanisms», *Neuroscience and Biobehavioral Reviews*, 85, p. 117-125. doi: 10.1016/j.neubiorev.2017.05.019.
- Van Eijk, K. R. *et al.* (2015) «Identification of schizophrenia-associated loci by combining DNA methylation and gene expression data from whole blood», *European Journal of Human Genetics*, 23(8), p. 1106-1110. doi: 10.1038/ejhg.2014.245.
- Ellison-Wright, I., Ellison-Wright, Z. i Bullmore, E. (2008) «Structural brain change in Attention Deficit Hyperactivity Disorder identified by meta-analysis», *BMC Psychiatry*, 8, p. 1-8. doi: 10.1186/1471-244X-8-51.
- Enoch, M. A. (2012) «The influence of gene-environment interactions on the development of alcoholism and drug dependence», *Current Psychiatry Reports*, 14(2), p. 150-158. doi: 10.1007/s11920-011-0252-9.
- Ermann, J. i Glimcher, L. H. (2012) «After GWAS: Mice to the rescue?», *Current Opinion in Immunology*, p. 564-570. doi: 10.1016/j.coi.2012.09.005.
- Ernst, J. *et al.* (2011) «Mapping and analysis of chromatin state dynamics in nine human cell types», *Nature*, 473(7345), p. 43-49. doi: 10.1038/nature09906.
- Evangelou, E. i Ioannidis, J. P. A. (2013) «Meta-analysis methods for genome-wide association studies and beyond», *Nature Reviews Genetics*, 14(6), p. 379-389. doi: 10.1038/nrg3472.
- Faraone, S. V. *et al.* (2005) «Molecular genetics of attention-deficit/hyperactivity disorder», *Biological Psychiatry*, 57(11), p. 1313-1323. doi: 10.1016/j.biopsych.2004.11.024.
- Faraone, S. V. *et al.* (2015) «Attention-deficit/hyperactivity disorder», *Nature Reviews Disease Primers*, 1, p. 15020. doi: 10.1038/nrdp.2015.20.
- Fernández-Castillo, N. *et al.* (2010) «Association study between the DAT1, DBH and DRD2 genes and cocaine dependence in a Spanish sample», *Psychiatric Genetics*, 20(6), p. 317-320. doi: 10.1097/YPG.0b013e32833b6320.
- Fernández-Castillo, N. *et al.* (2012) «Candidate pathway association study in cocaine dependence: The control of neurotransmitter release», *World Journal of Biological Psychiatry*, 13(2), p. 126-134. doi: 10.3109/15622975.2010.551406.
- Fernández-Castillo, N. *et al.* (2013) «Association study of 37 genes related to serotonin and dopamine neurotransmission and neurotrophic factors in cocaine dependence», *Genes, brain, and behavior*, 12(1), p. 39-46.
- Fernández-Castillo, N. *et al.* (2015) «Transcriptomic and genetic studies identify NFAT5 as a candidate gene for cocaine dependence», *Translational psychiatry*, 5, p. e667.
- Flagel, S. B. *et al.* (2011) «A selective role for dopamine in stimulus-reward learning», *Nature*, 469(7328), p. 53-59. doi: 10.1038/nature09588.
- Fliers, E. A. *et al.* (2012) «Genome-wide association study of motor coordination problems in ADHD identifies genes for brain and muscle function», *World Journal of Biological Psychiatry*, 13(3), p. 211-222. doi: 10.3109/15622975.2011.560279.
- Forman, J. J. i Collier, H. A. (2010) «The code within the code: MicroRNAs target coding regions», *Cell Cycle*, 9(8), p. 1533-1541. doi: 10.4161/cc.9.8.11202.
- Foroud, T. *et al.* (2008) «The tachykinin receptor 3 is associated with alcohol and cocaine dependence», *Alcoholism: Clinical and Experimental Research*, 32(6), p. 1023-1030. doi: 10.1111/j.1530-

- 0277.2008.00663.x.
- Franke, B. *et al.* (2010) «Multicenter analysis of the SLC6A3/DAT1 VNTR haplotype in persistent ADHD suggests differential involvement of the gene in childhood and persistent ADHD», *Neuropsychopharmacology*, 35(3), p. 656-664. doi: 10.1038/npp.2009.170.
- Friedman, R. C. *et al.* (2009) «Most mammalian mRNAs are conserved targets of microRNAs», *Genome Research*, 19(1), p. 92-105. doi: 10.1101/gr.082701.108.
- Frodl, T. i Skokauskas, N. (2012) «Meta-analysis of structural MRI studies in children and adults with attention deficit hyperactivity disorder indicates treatment effects», *Acta Psychiatrica Scandinavica*, 125(2), p. 114-126. doi: 10.1111/j.1600-0447.2011.01786.x.
- Fuchikami, M. *et al.* (2016) «The potential use of histone deacetylase inhibitors in the treatment of depression», *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 64, p. 320-324. doi: 10.1016/j.pnpbp.2015.03.010.
- Gagliano, S. A. *et al.* (2016) «Allele-Skewed DNA Modification in the Brain: Relevance to a Schizophrenia GWAS», *American journal of human genetics*, 98(5), p. 956-962.
- Gallo, E. F. i Posner, J. (2016) «Moving towards causality in attention-deficit hyperactivity disorder: Overview of neural and genetic mechanisms», *The Lancet Psychiatry*, 3(6), p. 555-567. doi: 10.1016/S2215-0366(16)00096-1.
- Gamazon, E. R. *et al.* (2013) «Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants», *Molecular Psychiatry*, 18(3), p. 340-346. doi: 10.1038/mp.2011.174.
- Garcia-Martínez, I. *et al.* (2016) «Preliminary evidence for association of genetic variants in pri-miR-34b/c and abnormal miR-34c expression with attention deficit and hyperactivity disorder», *Translational Psychiatry*, 6, p. e879. doi: 10.1038/tp.2016.151.
- García, M. P. *et al.* (2017) «Animal Models of Drug Addiction», *Adicciones*, 29(4), p. 278-292.
- Gardner, E. . (1997) «Brain reward mechanisms», en Lowinson, J. *et al.* (ed.) *Substance abuse: a comprehensive textbook*, p. 48-61.
- Gaur, U. *et al.* (2013) «Research progress in allele-specific expression and its regulatory mechanisms», *Journal of Applied Genetics*, 54(3), p. 271-283.
- Geaghan, M. i Cairns, M. J. (2015) «MicroRNA and Posttranscriptional Dysregulation in Psychiatry», *Biological Psychiatry*, 78(4), p. 231-239. doi: 10.1016/j.biopsych.2014.12.009.
- Gearhardt, A. N., Corbin, W. R. i Brownell, K. D. (2009) «Preliminary validation of the Yale Food Addiction Scale», *Appetite*, 52(2), p. 430-436. doi: 10.1016/j.appet.2008.12.003.
- Gearhardt, A. N., Corbin, W. R. i Brownell, K. D. (2016) «Development of the Yale Food Addiction Scale Version 2.0», *Psychology of Addictive Behaviors*, 30(1), p. 113-121. doi: 10.1037/adb0000136.
- Gelernter, J. *et al.* (2014) «Genome-wide association study of cocaine dependence and related traits: FAM53B identified as a risk gene», *Molecular psychiatry*, 19(6), p. 717-723.
- Geyer, M. A. i Swerdlow, N. R. (1998) «Measurement of Startle Response, Prepulse Inhibition, and Habituation», *Current Protocols in Neuroscience*, 3(8.7), p. 1-15. doi: 10.1002/0471142301.ns0807s03.
- Gibbs, J. R. *et al.* (2010) «Abundant quantitative trait loci exist for DNA methylation and gene expression in Human Brain», *PLoS Genetics*, 6(5), p. e1000952.
- Giel, K., Zipfel, S. i Hallschmid, M. (2018) «Oxytocin and Eating Disorders: A Narrative Review on Emerging Findings and Perspectives», *Current Neuropharmacology*, 16(8), p. 1111-1121. doi: 10.2174/1570159x15666171128143158.
- Girault, J.-A. i Greengard, P. (2004) «The Neurobiology of Dopamine», *Archives of Neurology & Psychiatry*, 61, p. 641-644. doi: 10.1176/ajp.137.12.1633-b.

Bibliografia

- Girirajan, S. *et al.* (2013) «Refinement and discovery of new hotspots of copy-number variation associated with autism spectrum disorder», *American Journal of Human Genetics*, 92(2), p. 221-237. doi: 10.1016/j.ajhg.2012.12.016.
- Giros, B. *et al.* (1996) «Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter», *Nature*, 379(6566), p. 606-612. doi: 10.1038/379606a0.
- Gizer, I. R., Ficks, C. i Waldman, I. D. (2009) «Candidate gene studies of ADHD: A meta-analytic review», *Human Genetics*, 126(1), p. 51-90. doi: 10.1007/s00439-009-0694-x.
- van de Glind, G. *et al.* (2014) «Variability in the prevalence of adult ADHD in treatment seeking substance use disorder patients: Results from an international multi-center study exploring DSM-IV and DSM-5 criteria», *Drug and Alcohol Dependence*, 134(1), p. 158-166. doi: 10.1016/j.drugalcdep.2013.09.026.
- Goldman, D., Oroszi, G. i Ducci, F. (2005) «The genetics of addictions: uncovering the genes», *Nature reviews. Genetics*, 6(7), p. 521-532.
- Gómez-Grau, M. *et al.* (2017) «New murine Niemann-Pick type C models bearing a pseudoexon-generating mutation recapitulate the main neurobehavioural and molecular features of the disease», *Scientific Reports*, 7, p. 41931. doi: 10.1038/srep41931.
- Gondré-Lewis, M. C., Basse, R. i Blum, K. (2020) «Pre-clinical models of reward deficiency syndrome: A behavioral octopus», *Neuroscience and Biobehavioral Reviews*, In Press. doi: 10.1016/j.neubiorev.2020.04.021.
- Goossens, D. *et al.* (2003) «A novel CpG-associated brain-expressed candidate gene for chromosome 18q-linked bipolar disorder», *Molecular Psychiatry*, 8(1), p. 83-89. doi: 10.1038/sj.mp.4001190.
- Gordon, E. L. *et al.* (2018) «What is the evidence for “food addiction?” A systematic review», *Nutrients*, 10, p. 477. doi: 10.3390/nu10040477.
- Grant, B. F. *et al.* (2016) «Epidemiology of DSM-5 drug use disorder results from the national epidemiologic survey on alcohol and related conditions-III», *JAMA Psychiatry*, 73(1), p. 39-47. doi: 10.1001/jamapsychiatry.2015.2132.
- Graziane, N. M. *et al.* (2016) «Opposing mechanisms mediate morphine- and cocaine-induced generation of silent synapses», *Nature Neuroscience*, 19(7), p. 915-925. doi: 10.1038/nn.4313.
- Greene, C. S. *et al.* (2009) «Failure to replicate a genetic association may provide important clues about genetic architecture», *PLoS ONE*, 4(6), p. e5639. doi: 10.1371/journal.pone.0005639.
- Grucza, R. A. *et al.* (2008) «A risk allele for nicotine dependence in CHRNA5 is a protective allele for cocaine dependence», *Biological psychiatry*, 64(11), p. 922-929.
- Gulyaeva, L. F. i Kushlinskiy, N. E. (2016) «Regulatory mechanisms of microRNA expression», *Journal of Translational Medicine*, 14(1), p. 143. doi: 10.1186/s12967-016-0893-x.
- Guy, J. *et al.* (2001) «A mouse Mecp2-null mutation causes neurological symptoms that mimic rett syndrome», *Nature Genetics*, 27(3), p. 322-326. doi: 10.1038/85899.
- Hall, F. S. *et al.* (2013) «Implications of genome wide association studies for addiction: Are our a priori assumptions all wrong?», *Pharmacology and Therapeutics*, 140(3), p. 267-279. doi: 10.1016/j.pharmthera.2013.07.006.
- Halterman, J. A., Kwon, H. M. i Wamhoff, B. R. (2012) «Tonicity-independent regulation of the osmosensitive transcription factor TonEBP (NFAT5)», *American Journal of Physiology. Cell Physiology*, 302, p. C1-C8.
- Hamilton, P. J. i Nestler, E. J. (2019) «Epigenetics and addiction», *Current Opinion in Neurobiology*, 59, p. 128-136. doi: 10.1016/j.conb.2019.05.005.
- Hannon, E. *et al.* (2015) «Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci», *Nature Neuroscience*, 19(1), p. 48-54.
- Harstad, E. *et al.* (2014) «Attention-deficit/hyperactivity disorder and substance abuse», *Pediatrics*, 134(1),

- p. e293-e301. doi: 10.1542/peds.2014-0992.
- Hart, H. *et al.* (2012) «Meta-analysis of Functional Magnetic Resonance Imaging Studies of Inhibition and Attention in Attention-deficit/Hyperactivity Disorder», *JAMA Psychiatry*, 70(2), p. 185. doi: 10.1001/jamapsychiatry.2013.277.
- Hasin, D. S. *et al.* (2013) «DSM-5 criteria for substance use disorders: Recommendations and rationale», *American Journal of Psychiatry*, 170(8), p. 834-851. doi: 10.1176/appi.ajp.2013.12060782.
- Hauck, C. *et al.* (2017) «Prevalence of “Food Addiction” as Measured with the Yale Food Addiction Scale 2.0 in a Representative German Sample and Its Association with Sex, Age and Weight Categories», *Obesity Facts*, 10(1), p. 12-24. doi: 10.1159/000456013.
- Heinonen, J. *et al.* (2011) «Snorc is a novel cartilage specific small membrane proteoglycan expressed in differentiating and articular chondrocytes», *Osteoarthritis and Cartilage*, 19(8), p. 1026-1035. doi: 10.1016/j.joca.2011.04.016.
- Heinonen, J. *et al.* (2017) «Defects in chondrocyte maturation and secondary ossification in mouse knee joint epiphyses due to Snorc deficiency», *Osteoarthritis and Cartilage*, 25(7), p. 1132-1142. doi: 10.1016/j.joca.2017.03.010.
- Hellwege, J. N. *et al.* (2017) «Population Stratification in Genetic Association Studies», *Current Protocols in Human Genetics*, 95(1), p. 1.22.1-1.22.23. doi: 10.1002/cphg.48.
- Higgins, G. A., Sellers, E. M. i Fletcher, P. J. (2013) «From obesity to substance abuse: Therapeutic opportunities for 5-HT 2C receptor agonists», *Trends in Pharmacological Sciences*, 34(10), p. 560-570. doi: 10.1016/j.tips.2013.08.001.
- Hinney, A. *et al.* (2011) «Genome-wide association study in German patients with attention deficit/hyperactivity disorder», *American Journal of Medical Genetics, Part B*, 156(8), p. 888-897. doi: 10.1002/ajmg.b.31246.
- Hirschhorn, J. N. i Daly, M. J. (2005) «Genome-wide association studies for common diseases and complex traits», *Nature Reviews Genetics*, 6(2), p. 95-108. doi: 10.1038/nrg1521.
- Hoehe, M. R. *et al.* (2000) «Sequence variability and candidate gene analysis in complex disease: association of mu opioid receptor gene variation with substance dependence», *Human Molecular Genetics*, 9(19), p. 2895-2908. doi: 10.1093/hmg/9.19.2895.
- Hoogman, M. *et al.* (2017) «Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis», *Lancet Psychiatry*, 4(4), p. 310-319. doi: 10.1016/S2215-0366(17)30049-4.
- Hope, B. T. (1998) «Cocaine and the AP-1 transcription factor complex», *Annals of the New York Academy of Sciences*, 844, p. 1-6. doi: 10.1111/j.1749-6632.1998.tb08216.x.
- Hser, Y. I. *et al.* (2001) «An evaluation of drug treatments for adolescents in 4 US cities», *Archives of General Psychiatry*, 58(7), p. 689-695. doi: 10.1001/archpsyc.58.7.689.
- Hsin, Y. H. *et al.* (2011) «The role of TonEBP in regulation of AAD expression and dopamine production in renal proximal tubule cells upon hypertonic challenge», *Biochemical and Biophysical Research Communications*, 414, p. 598-603.
- Hunsberger, J. G. *et al.* (2009) «MicroRNAs in mental health: from biological underpinnings to potential therapies», *Neuromolecular medicine*, 11(3), p. 173-182. doi: 10.1007/s12017-009-8070-5.
- Hunt, G. E. *et al.* (2018) «Prevalence of comorbid substance use in schizophrenia spectrum disorders in community and clinical settings, 1990–2017: Systematic review and meta-analysis», *Drug and Alcohol Dependence*, 191, p. 234-258. doi: 10.1016/j.drugalcdep.2018.07.011.
- Hwang, L. D. *et al.* (2019) «New insight into human sweet taste: A genome-wide association study of the perception and intake of sweet substances», *American Journal of Clinical Nutrition*, 109(6), p. 1724-1737. doi: 10.1093/ajcn/nqz043.

Bibliografia

- Ikeda, M. *et al.* (2013) «Evidence for Shared Genetic Risk Between Methamphetamine-Induced Psychosis and Schizophrenia», *Neuropsychopharmacology*, 38(10), p. 1864-1870. doi: 10.1038/npp.2013.94.
- Im, H. I. *et al.* (2010) «MeCP2 controls BDNF expression and cocaine intake through homeostatic interactions with microRNA-212», *Nature Neuroscience*, 13(9), p. 1120-1127. doi: 10.1038/nn.2615.
- Ishida, T. i Kinoshita, K. (2007) «PrDOS: Prediction of Disordered Protein Regions From Amino Acid Sequence», *Nucleic Acid Research*, 35, p. W460-W464.
- Issler, O. *et al.* (2014) «MicroRNA 135 is essential for chronic stress resiliency, antidepressant efficacy, and intact serotonergic activity», *Neuron*, 83(2), p. 344-360. doi: 10.1016/j.neuron.2014.05.042.
- Issler, O. i Chen, A. (2015) «Determining the role of microRNAs in psychiatric disorders», *Nature Reviews Neuroscience*, 16(4), p. 201-212. doi: 10.1038/nrn3879.
- Jacobs, E. H. *et al.* (2003) «Neuroadaptive effects of active versus passive drug administration in addiction research», *Trends in Pharmacological Sciences*, 24(11), p. 566-573. doi: 10.1016/j.tips.2003.09.006.
- Jaiswal, P. K. *et al.* (2020) «Biochemical characteristics of the chondrocyte-enriched SNORC protein and its transcriptional regulation by SOX9», *Scientific Reports*, 10(1). doi: 10.1038/s41598-020-64640-x.
- Javaras, K. N. (2007) «Epidemiology of Eating Disorders : Creating Opportunities», *International Journal of Eating Disorders*, 41(2), p. 174-179. doi: 10.1002/eat.
- Jayanthi, S. *et al.* (2014) «Methamphetamine downregulates striatal glutamate receptors via diverse epigenetic mechanisms», *Biological Psychiatry*, 76(1), p. 47-56. doi: 10.1016/j.biopsych.2013.09.034.
- Jensen, K. P. (2016) «A Review of Genome-Wide Association Studies of Stimulant and Opioid Use Disorders», *Molecular neuropsychiatry*, 2(1), p. 37-45.
- Jeong, G. R. *et al.* (2016) «Inflammatory signals induce the expression of tonicity-responsive enhancer binding protein (TonEBP) in microglia», *Journal of Neuroimmunology*, 295, p. 21-29.
- Johnson, E. O. *et al.* (2015) «KAT2B polymorphism identified for drug abuse in African Americans with regulatory links to drug abuse pathways in human prefrontal cortex», *Addiction Biology*, 21(6), p. 1217-1232.
- Johnson, R. C. *et al.* (2010) «Accounting for multiple comparisons in a genome-wide association study (GWAS)», *BMC Genomics*, 11(1), p. 724. doi: 10.1186/1471-2164-11-724.
- Jurek, B. i Neumann, I. D. (2018) «The oxytocin receptor: From intracellular signaling to behavior», *Physiological Reviews*, 98(3), p. 1805-1908. doi: 10.1152/physrev.00031.2017.
- Kalsi, G. *et al.* (2010) «A systematic gene-based screen of chr4q22-q32 identifies association of a novel susceptibility gene, DKK2, with the quantitative trait of alcohol dependence symptom counts», *Human Molecular Genetics*, 19(12), p. 2497-2506. doi: 10.1093/hmg/ddq112.
- Kalueff, A. V. *et al.* (2013) «Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond», *Zebrafish*, 10(1), p. 70-86. doi: 10.1089/zeb.2012.0861.
- Kamegai, J. *et al.* (2000) «Central Effect of Ghrelin, an Endogenous Growth Hormone Secretagogue, on Hypothalamic Peptide Gene Expression», *Endocrinology*, 141(12), p. 4797-4800. doi: 10.1210/endo.141.12.7920.
- Kandemir, H. *et al.* (2014) «Evaluation of several micro RNA (miRNA) levels in children and adolescents with attention deficit hyperactivity disorder», *Neuroscience Letters*, 580, p. 158-162. doi: 10.1016/j.neulet.2014.07.060.
- Kelly, T. M. i Daley, D. C. (2013) «Integrated treatment of substance use and psychiatric disorders», *Social Work in Public Health*, 28(3-4), p. 388-406. doi: 10.1080/19371918.2013.774673.
- Kim, C. H. *et al.* (2015) «The continuous performance test (rCPT) for mice: A novel operant touchscreen test of attentional function», *Psychopharmacology*, 232, p. 3947-3966. doi: 10.1007/s00213-015-4081-0.
- Kim, V. N., Han, J. i Siomi, M. C. (2009) «Biogenesis of small RNAs in animals», *Nature Reviews Molecular Cell*

- Biology*, 10(2), p. 126-139. doi: 10.1038/nrm2632.
- Klose, R. J. i Bird, A. P. (2006) «Genomic DNA methylation: The mark and its mediators», *Trends in Biochemical Sciences*, 31(2), p. 89-97. doi: 10.1016/j.tibs.2005.12.008.
- Kofink, D. *et al.* (2013) «Epigenetic dynamics in psychiatric disorders: Environmental programming of neurodevelopmental processes», *Neuroscience and Biobehavioral Reviews*, 37(5), p. 831-845. doi: 10.1016/j.neubiorev.2013.03.020.
- Komada, M., Takao, K. i Miyakawa, T. (2008) «Elevated plus maze for mice», *Journal of Visualized Experiments*, 22(e1088). doi: 10.3791/1088.
- Koob, G. F. i Le Moal, M. (2001) «Drug Addiction, Dysregulation of Reward, and Allostasis», *Neuropsychopharmacology*, 24(2), p. 97-129. doi: 10.1016/S0893-133X(00)00195-0.
- Koob, G. F. i Volkow, N. D. (2010) «Neurocircuitry of addiction», *Neuropsychopharmacology*, 35(1), p. 217-238. doi: 10.1038/npp.2009.110.
- Koob, G. F. i Volkow, N. D. (2016) «Neurobiology of addiction: a neurocircuitry analysis», *The Lancet Psychiatry*, 3(8), p. 760-773. doi: 10.1016/S2215-0366(16)00104-8.
- Kornberg, R. (1974) «Chromatin Structure : A Repeating Unit of Histones and DNA Chromatin structure is based on a repeating unit of eight», *Science*, 184, p. 868-871.
- Kovács, G. L., Sarnyai, Z. i Szabó, G. (1998) «Oxytocin and addiction: A review», *Psychoneuroendocrinology*, 23(8), p. 945-962. doi: 10.1016/S0306-4530(98)00064-X.
- Kraft, P., Zeggini, E. i Ioannidis, J. P. A. (2009) «Replication in genome-wide association studies», *Statistical science*, 24(4), p. 561-573. doi: 10.1214/09-STS290.
- Kranzler, H. R. *et al.* (1998) «Association of alcohol or other drug dependence with alleles of the mu opioid receptor gene (OPRM1)», *Alcoholism: Clinical and Experimental Research*, 22(6), p. 1359-1362. doi: 10.1111/j.1530-0277.1998.tb03919.x.
- Kranzler, H. R. *et al.* (2019) «Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations», *Nature Communications*, 10, p. 1499. doi: 10.1038/s41467-019-09480-8.
- Krebs, M. O. *et al.* (1998) «Dopamine D3 receptor gene variants and substance abuse in schizophrenia», *Molecular Psychiatry*, 3(4), p. 337-341. doi: 10.1038/sj.mp.4000411.
- Krogh, A. *et al.* (2001) «Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes», *Journal of Molecular Biology*, 305(3), p. 567-580. doi: 10.1006/jmbi.2000.4315.
- Kuehner, J. N. *et al.* (2019) «Epigenetic regulations in neuropsychiatric disorders», *Frontiers in Genetics*, 10, p. 268. doi: 10.3389/fgene.2019.00268.
- Kular, L. i Kular, S. (2018) «Epigenetics applied to psychiatry: Clinical opportunities and future challenges», *Psychiatry and Clinical Neurosciences*, 72(4), p. 195-211. doi: 10.1111/pcn.12634.
- Kumar, A. *et al.* (2005) «Chromatin remodeling is a key mechanism underlying cocaine-induced plasticity in striatum», *Neuron*, 48(2), p. 303-314. doi: 10.1016/j.neuron.2005.09.023.
- Kundaje, A. *et al.* (2015) «Integrative analysis of 111 reference human epigenomes», *Nature*, 518, p. 317-329. doi: 10.1038/nature14248.
- Kundu, T. K. i Rao, M. R. S. (1999) «CpG islands in chromatin organization and gene expression», *Journal of Biochemistry*, 125(2), p. 217-222. doi: 10.1093/oxfordjournals.jbchem.a022276.
- Kurita, M. *et al.* (2012) «HDAC2 regulates atypical antipsychotic responses through the modulation of mGlu2 promoter activity», *Nature Neuroscience*, 15(9), p. 1245-1254. doi: 10.1038/nn.3181.
- de la Peña, J. B. *et al.* (2018) «Exploring the Validity of Proposed Transgenic Animal Models of Attention-Deficit Hyperactivity Disorder (ADHD)», *Molecular Neurobiology*, 55(5), p. 3739-3754. doi: 10.1007/s12035-017-0608-1.

Bibliografia

- Labouèbe, G. *et al.* (2013) «Insulin induces long-term depression of ventral tegmental area dopamine neurons via endocannabinoids», *Nature neuroscience*, 16(3), p. 300-308. doi: 10.1038/nn.3321.
- Landaas, E. T. *et al.* (2010) «An international multicenter association study of the serotonin transporter gene in persistent ADHD», *Genes, Brain and Behavior*, 9(5), p. 449-458. doi: 10.1111/j.1601-183X.2010.00567.x.
- Laplant, Q. *et al.* (2010) «Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens», *Nature Neuroscience*, 13(9), p. 1137-1143. doi: 10.1038/nn.2619.
- Lasky-Su, J., Neale, B. M., *et al.* (2008) «Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations», *American Journal of Medical Genetics Part B*, 147B(8), p. 1345-1354. doi: 10.1002/ajmg.b.30867.
- Lasky-Su, J., Anney, R. J. L., *et al.* (2008) «Genome-wide association scan of the time to onset of attention deficit hyperactivity disorder», *American Journal of Medical Genetics Part B*, 147B(8), p. 1355-1358. doi: 10.1002/ajmg.b.30869.
- Ledent, C. *et al.* (1999) «Unresponsiveness to cannabinoids and reduced addictive effects of opiates in CB1 receptor knockout mice», *Science*, 283, p. 401-404. doi: 10.1126/science.283.5400.401.
- Lee, I. *et al.* (2009) «New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites», *Genome Research*, 19(7), p. 1175-1183. doi: 10.1101/gr.089367.108.
- Lee, S. S. *et al.* (2011) «Prospective association of childhood attention-deficit/hyperactivity disorder (ADHD) and substance use and abuse/dependence: A meta-analytic review», *Clinical Psychology Review*, 31(3), p. 328-341. doi: 10.1016/j.cpr.2011.01.006.
- Leo, D. i Gainetdinov, R. R. (2013) «Transgenic mouse models for ADHD», *Cell and Tissue Research*, p. 259-271. doi: 10.1007/s00441-013-1639-1.
- Leong, K. C. *et al.* (2018) «Oxytocin and Rodent Models of Addiction», en Olive, M. F. i Tomek, S. E. (ed.) *International Review of Neurobiology. Animal Models for Examining Social Influences on Drug Addiction*, p. 201-247. doi: 10.1016/bs.irn.2018.07.007.
- Lesch, K. P. *et al.* (2008) «Molecular genetics of adult ADHD: Converging evidence from genome-wide association and extended pedigree linkage studies», *Journal of Neural Transmission*, 115(11), p. 1573-1585. doi: 10.1007/s00702-008-0119-3.
- Leslie, M. *et al.* (2019) «The influence of oxytocin on eating behaviours and stress in women with bulimia nervosa and binge eating disorder», *Molecular and Cellular Endocrinology*, 497, p. 110354. doi: 10.1016/j.mce.2018.12.014.
- Lever, C., Burton, S. i O'Keefe, J. (2006) «Rearing on Hind Legs, Environmental Novelty, and the Hippocampal Formation», *Reviews in the Neurosciences*, 17(1-2). doi: 10.1515/REVNEURO.2006.17.1-2.111.
- Leyton, M. (2014) «What's deficient in reward deficiency?», *Journal of Psychiatry and Neuroscience*, 39(5), p. 291-293. doi: 10.1503/jpn.140204.
- Li, M. J. *et al.* (2012) «GWASdb: a database for human genetic variants identified by genome-wide association studies», *Nucleic acids research*, 40, p. D1047-D1054. doi: 10.1093/nar/gkr1182.
- Li, M. J. *et al.* (2016) «GWASdb v2: an update database for human genetic variants identified by genome-wide association studies», *Nucleic acids research*, 44, p. D869-D876. doi: 10.1093/nar/gkv1317.
- Li, M. X. *et al.* (2012) «Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets», *Human Genetics*, 131, p. 747-756.
- Li, Q. *et al.* (2007) «The usefulness of the spontaneously hypertensive rat to model attention-deficit/hyperactivity disorder (ADHD) may be explained by the differential expression of dopamine-related genes in the brain», *Neurochemistry International*, 50(6), p. 848-857. doi: 10.1016/j.neuint.2007.02.005.
- Li, Y. *et al.* (2016) «MiR-182 (microRNA-182) suppression in the hippocampus evokes antidepressant-like effects in rats», *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 65, p. 96-103. doi:

- 10.1016/j.pnpbp.2015.09.004.
- Li, Z. *et al.* (2017) «Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia», *Nature Genetics*, 49(11), p. 1576-1583.
- Liu, C. *et al.* (2018) «DNA Methylation and Psychiatric Disorders», *Progress in Molecular Biology and Translational Science*, 157, p. 175-232. doi: 10.1016/bs.pmbts.2018.01.006.
- Liu, X., Li, Y. I. i Pritchard, J. K. (2019) «Trans Effects on Gene Expression Can Drive Omnigenic Inheritance», *Cell*, 177, p. 1022-1034.
- Liu, Y. J. *et al.* (2008) «Genome-wide association scans identified CTNBL1 as a novel gene for obesity», *Human Molecular Genetics*, 17(12), p. 1803-1813. doi: 10.1093/hmg/ddn072.
- Llewellyn, C. H. *et al.* (2013) «Finding the missing heritability in pediatric obesity: The contribution of genome-wide complex trait analysis», *International Journal of Obesity*, 37(11), p. 1506-1509. doi: 10.1038/ijo.2013.30.
- Lohoff, F. W., Weller, A. E., *et al.* (2008) «Association between the catechol-O-methyltransferase Val158Met polymorphism and cocaine dependence», *Neuropsychopharmacology*, 33(13), p. 3078-3084. doi: 10.1038/npp.2008.126.
- Lohoff, F. W., Bloch, P. J., *et al.* (2008) «Genetic variants in the cocaine- and amphetamine-regulated transcript gene (CARTPT) and cocaine dependence», *Neuroscience Letters*, 440(3), p. 280-283. doi: 10.1016/j.neulet.2008.05.073.
- Love, T. M. (2014) «Oxytocin, motivation and the role of dopamine», *Pharmacology, Biochemistry and Behavior*, 119, p. 49-60.
- Loyher, M. L. *et al.* (2004) «Transcription factor tonicity-responsive enhancer-binding protein (TonEBP) which transactivates osmoprotective genes is expressed and upregulated following acute systemic hypertonicity in neurons in brain», *Neuroscience*, 124, p. 89-104.
- Lueptow, L. M. (2017) «Novel Object Recognition Test for the Investigation of Learning and Memory in Mice», *Journal of Visualized Experiments*, 126, p. e55718. doi: 10.3791/55718.
- Luo, X. *et al.* (2003) «Haplotypes at the OPRM1 locus are associated with susceptibility to substance dependence in European-Americans», *American Journal of Medical Genetics Part B*, 120B, p. 97-108. doi: 10.1002/ajmg.b.20034.
- Luoni, A. i Riva, M. A. (2016) «MicroRNAs and psychiatric disorders: From aetiology to treatment», *Pharmacology and Therapeutics*, 167, p. 13-27. doi: 10.1016/j.pharmthera.2016.07.006.
- Lv, J. *et al.* (2013) «The epigenetic switches for neural development and psychiatric disorders», *Journal of Genetics and Genomics*, 40(7), p. 339-346. doi: 10.1016/j.jgg.2013.04.007.
- Lytle, J. R., Yario, T. A. i Steitz, J. A. (2007) «Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR», *PNAS*, 104(23), p. 9667-9672. doi: 10.1073/pnas.0703820104.
- MacNicol, B. (2017) «The biology of addiction», *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*, 64, p. 141-148. doi: 10.1007/s12630-016-0771-2.
- Magidson, J. F. *et al.* (2012) «Comparison of the course of substance use disorders among individuals with and without generalized anxiety disorder in a nationally representative sample», *Journal of Psychiatric Research*, 46(5), p. 659-666. doi: 10.1016/j.jpsychires.2012.02.011.
- Mak, K. M. C. *et al.* (2012) «Nuclear Factor of Activated T Cells 5 Deficiency Increases the Severity of Neuronal Cell Death in Ischemic Injury», *Neurosignals*, 20, p. 237-251.
- Maldonado, R. *et al.* (1997) «Absence of opiate rewarding effects in mice lacking dopamine D2 receptors», *Nature*, 388, p. 586-589. doi: 10.1038/41567.
- Malvaez, M. *et al.* (2011) «CBP in the nucleus accumbens regulates cocaine-induced histone acetylation and is critical for cocaine-associated behaviors», *Journal of Neuroscience*, 31(47), p. 16941-16948. doi:

Bibliografia

- 10.1523/JNEUROSCI.2747-11.2011.
- Mancino, S. *et al.* (2015) «Epigenetic and proteomic expression changes promoted by eating addictive-like behavior», *Neuropsychopharmacology*, 40(12), p. 2788-2800. doi: 10.1038/npp.2015.129.
- Manolio, T. A. *et al.* (2009) «Finding the missing heritability of complex diseases», *Nature*, 461, p. 747-753. doi: 10.1038/nature08494.
- Marouli, E. *et al.* (2017) «Rare and low-frequency coding variants alter human adult height», *Nature*, 542, p. 186-190. doi: 10.1038/nature21039.
- Martín-García, E. *et al.* (2011) «New operant model of reinstatement of food-seeking behavior in mice», *Psychopharmacology*, 215(1), p. 49-70. doi: 10.1007/s00213-010-2110-6.
- Matte, B. *et al.* (2015) «ADHD in DSM-5 : a field trial in a large , representative sample of 18- to 19-year-old adults», *Psychological Medicine*, 45, p. 361-373. doi: 10.1017/S0033291714001470.
- Maurano, M. T. *et al.* (2012) «Systematic localization of common disease-associated variation in regulatory DNA», *Science*, 337, p. 1190-1195. doi: 10.1126/science.1222794.
- McCarthy, M. I. *et al.* (2008) «Genome-wide association studies for complex traits: Consensus, uncertainty and challenges», *Nature Reviews Genetics*, 9(5), p. 356-369. doi: 10.1038/nrg2344.
- McGue, M. *et al.* (2013) «A genome-wide association study of behavioral disinhibition», *Behavior Genetics*, 43(5), p. 363-373.
- Meaburn, E. L., Schalkwyk, L. C. i Mill, J. (2010) «Allele-specific methylation in the human genome: Implications for genetic studies of complex disease», *Epigenetics*, 5(7), p. 578-582.
- Melas, P. A. *et al.* (2012) «Epigenetic aberrations in leukocytes of patients with schizophrenia: Association of global DNA methylation with antipsychotic drug treatment and disease onset», *FASEB Journal*, 26(6), p. 2712-2718. doi: 10.1096/fj.11-202069.
- Mellios, N. *et al.* (2008) «A set of differentially expressed miRNAs, including miR-30a-5p, act as post-transcriptional inhibitors of BDNF in prefrontal cortex», *Human Molecular Genetics*, 17(19), p. 3030-3042. doi: 10.1093/hmg/ddn201.
- Merikangas, K. R. *et al.* (1998) «Familial transmission of substance use disorders», *Archives of General Psychiatry*, 55(11), p. 973-979. doi: 10.1001/archpsyc.55.11.973.
- Messas, G. *et al.* (2005) «Association study of dopamine D2 and D3 receptor gene polymorphisms with cocaine dependence», *Psychiatric Genetics*, 15(3), p. 171-174. doi: 10.1097/00041444-200509000-00006.
- Meule, A. (2011) «How prevalent is “food addiction”?», *Frontiers in Psychiatry*, 2, p. 61. doi: 10.3389/fpsy.2011.00061.
- Meule, A. (2019) «A history of “food addiction”», en Cottone, P. *et al.* (ed.) *Compulsive Eating Behavior and Food Addiction: Emerging Pathological Constructs*. Academic Press. doi: 10.1016/b978-0-12-816207-1.00001-9.
- Meule, A. i Gearhardt, A. N. (2014) «Five years of the Yale Food Addiction Scale: Taking stock and moving forward», *Current Addiction Reports*, 1(3), p. 193-205. doi: 10.1007/s40429-014-0021-z.
- Meza-Sosa, K. F. *et al.* (2012) «Role of microRNAs in central nervous system development and pathology», *Journal of Neuroscience Research*, 90(1), p. 1-12. doi: 10.1002/jnr.22701.
- Michels, L. *et al.* (2014) «Prefrontal GABA and glutathione imbalance in posttraumatic stress disorder: Preliminary findings», *Psychiatry Research: Neuroimaging*, 224, p. 288-295. doi: 10.1016/j.pscychresns.2014.09.007.
- Mick, E. *et al.* (2010) «Family-based genome-wide association scan of attention-deficit/ hyperactivity disorder», *Journal of the American Academy of Child and Adolescent Psychiatry*, 49, p. 898-905.e3. doi: 10.1016/j.jaac.2010.02.014.
- Misztak, P., Pańcyszyn-Trzewik, P. i Sowa-Kućma, M. (2018) «Histone deacetylases (HDACs) as therapeutic

- target for depressive disorders», *Pharmacological Reports*, 70(2), p. 398-408. doi: 10.1016/j.pharep.2017.08.001.
- Mitchell, M. R. i Potenza, M. N. (2014) «Addictions and Personality Traits: Impulsivity and Related Constructs», *Current behavioral neuroscience reports*, 1(1), p. 1-12. doi: 10.1007/s40473-013-0001-y.
- Mokhtari, R. i Lachman, H. M. (2016) «The Major Histocompatibility Complex (MHC) in Schizophrenia: A Review», *Journal of Clinical & Cellular Immunology*, 7, p. 479. doi: 10.4172/2155-9899.1000479.
- Moser, M. *et al.* (1988) «Component analyses differentiate between exploratory behaviour of spontaneously hypertensive rats and Wistar Kyoto rats in a two-compartment free-exploration open field», *Scandinavian Journal of Psychology*, 29(3-4), p. 200-206. doi: 10.1111/j.1467-9450.1988.tb00792.x.
- Moszyńska, A. *et al.* (2017) «SNPs in microRNA target sites and their potential role in human disease», *Open Biology*, 7(4), p. 170019. doi: 10.1098/rsob.170019.
- Moy, S. S. *et al.* (2004) «Sociability and preference for social novelty in five inbred strains: An approach to assess autistic-like behavior in mice», *Genes, Brain and Behavior*, 3(5), p. 287-302. doi: 10.1111/j.1601-1848.2004.00076.x.
- Müller, C. P. i Homberg, J. R. (2015) «The role of serotonin in drug use and addiction», *Behavioural Brain Research*, 277, p. 146-192. doi: 10.1016/j.bbr.2014.04.007.
- Murphy, T. M. *et al.* (2015) «Anxiety is associated with higher levels of global DNA methylation and altered expression of epigenetic and interleukin-6 genes», *Psychiatric Genetics*, 25(2), p. 71-78. doi: 10.1097/YPG.0000000000000055.
- Mutz, K. O. *et al.* (2013) «Transcriptome analysis using next-generation sequencing», *Current Opinion in Biotechnology*, 24(1), p. 22-30. doi: 10.1016/j.copbio.2012.09.004.
- Myers-Payne, S. C. *et al.* (1996) «Effects of chronic ethanol consumption on sterol transfer proteins in mouse brain», *Journal of neurochemistry*, 66(1), p. 313-320.
- Myers, M. M., Whittmore, S. R. i Hendley, E. D. (1981) «Changes in catecholamine neuronal uptake and receptor binding in the brains of spontaneously hypertensive rats (SHR)», *Brain Research*, 220(2), p. 325-338. doi: 10.1016/0006-8993(81)91221-X.
- Nagarajan, R. P. *et al.* (2006) «Reduced MeCP2 expression is frequent in autism frontal cortex and correlates with aberrant MECP2 promoter methylation», *Epigenetics*, 1, p. e1-e11. doi: 10.4161/epi.1.4.3514.
- Nakao, T. *et al.* (2011) «Gray matter volume abnormalities in ADHD: voxel-based meta-analysis exploring the effects of age and stimulant medication», *The American journal of psychiatry*, 168(11), p. 1154-63. doi: 10.1176/appi.ajp.2011.11020281.
- Neale, B. M., Medland, S., *et al.* (2010) «Case-control genome-wide association study of attention-deficit/hyperactivity disorder», *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(9), p. 906-920. doi: 10.1016/j.jaac.2010.06.007.
- Neale, B. M., Medland, S. E., *et al.* (2010) «Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder», *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(9), p. 884-897. doi: 10.1016/j.jaac.2010.06.008.
- Nelson, R. J. (1997) «The use of genetic “knockout” mice in behavioral endocrinology research», *Hormones and Behavior*, 31(3), p. 188-196. doi: 10.1006/hbeh.1997.1381.
- Németh, N. *et al.* (2013) «Association of impulsivity and polymorphic MicroRNA-641 target sites in the SNAP-25 gene», *PLoS ONE*, 8(12). doi: 10.1371/journal.pone.0084207.
- Nestler, E. J. (2001) «Molecular neurobiology of addiction», *The American Journal on Addictions*, p. 201-217. doi: 10.1080/105504901750532094.
- Nestler, E. J. (2005) «Is there a common molecular pathway for addiction?», *Nature Neuroscience*, 8(11), p. 1445-1449. doi: 10.1038/nn1578.

Bibliografia

- Nestler, E. J. (2014) «Epigenetic mechanisms of drug addiction», *Neuropharmacology*, 76, p. 259-268. doi: 10.1016/j.neuropharm.2013.04.004.
- Nestler, E. J. i Hyman, S. E. (2010) «Animal models of neuropsychiatric disorders», *Nature Neuroscience*, 13(10), p. 1161-1169. doi: 10.1038/nn.2647.
- Nestler, E. J. i Lüscher, C. (2019) «The Molecular Basis of Drug Addiction: Linking Epigenetic to Synaptic and Circuit Mechanisms», *Neuron*, 102(1), p. 48-59. doi: 10.1016/j.neuron.2019.01.016.
- Neumann, I. D. i Landgraf, R. (2012) «Balance of brain oxytocin and vasopressin: Implications for anxiety, depression, and social behaviors», *Trends in Neurosciences*, 35(11), p. 649-659. doi: 10.1016/j.tins.2012.08.004.
- Nishioka, M. *et al.* (2013) «Comprehensive DNA methylation analysis of peripheral blood cells derived from patients with first-episode schizophrenia», *Journal of Human Genetics*, 58(2), p. 91-97. doi: 10.1038/jhg.2012.140.
- Noble, E. P. *et al.* (1993) «Allelic association of the D2 dopamine receptor gene with cocaine dependence», *Drug and alcohol dependence*, 33(3), p. 271-285.
- Norman, L. J. *et al.* (2016) «Structural and functional brain abnormalities in attention-deficit/hyperactivity disorder and obsessive-compulsive disorder: A comparative meta-analysis», *JAMA Psychiatry*, 73(8), p. 815-825. doi: 10.1001/jamapsychiatry.2016.0700.
- Nsengimana, J. i Bishop, D. T. (2017) «Design considerations for genetic linkage and association studies», en Elston, R. (ed.) *Statistical Human Genetics. Methods in Molecular Biology*. Humana Press, New York, NY, p. 257-281. doi: 10.1007/978-1-4939-7274-6_13.
- O'Connor, E. C. *et al.* (2011) «The predictive validity of the rat self-administration model for abuse liability», *Neuroscience & Biobehavioral Reviews*, 35(3), p. 912-938. doi: 10.1016/J.NEUBIOREV.2010.10.012.
- O'Driscoll, L. *et al.* (1993) «The use of reverse transcriptase-polymerase chain reaction (RT-PCR) to investigate specific gene expression in multidrug-resistant cells», *Cytotechnology*, 12(1-3), p. 289-314. doi: 10.1007/BF00744669.
- Okamoto, K. i Aoki, K. (1963) «Development of a strain of spontaneously hypertensive rats», *Japanese Circulation Journal*, 27(3), p. 282-293. doi: 10.1253/jcj.27.282.
- Olive, M. F. i Kalivas, P. W. (2010) «Conditioning of Addiction», en Johnson B. (ed.) *Addiction Medicine*, p. 159-178. doi: 10.1007/978-1-4419-0338-9_8.
- Onishchenko, N. *et al.* (2008) «Long-lasting depression-like behavior and epigenetic changes of BDNF gene expression induced by perinatal exposure to methylmercury», *Journal of Neurochemistry*, 106(3), p. 1378-1387. doi: 10.1111/j.1471-4159.2008.05484.x.
- Otsuki, K. *et al.* (2008) «Altered expression of neurotrophic factors in patients with major depression», *Journal of Psychiatric Research*, 42, p. 1145-1153. doi: 10.1016/j.jpsychires.2008.01.010.
- Pagerols, M. *et al.* (2017) «Pharmacogenetics of methylphenidate response and tolerability in attention-deficit/hyperactivity disorder», *Pharmacogenomics Journal*, 17(1), p. 98-104. doi: 10.1038/tpj.2015.89.
- Panksepp, J., Knutson, B. i Burgdorf, J. (2002) «The role of brain emotional systems in addictions: A neuro-evolutionary perspective and new "self-report" animal model», *Addiction*, 97(4), p. 459-469. doi: 10.1046/j.1360-0443.2002.00025.x.
- Papassotiropoulos, A. *et al.* (2013) «A genome-wide survey and functional brain imaging study identify CTNBL1 as a memory-related gene», *Molecular Psychiatry*, 18(2), p. 255-263.
- Parakh, P. i Basu, D. (2013) «Cannabis and psychosis: Have we found the missing links?», *Asian Journal of Psychiatry*, 6(4), p. 281-287. doi: 10.1016/j.ajp.2013.03.012.
- Peña, C. J. *et al.* (2014) «Epigenetic signaling in psychiatric disorders», *Journal of Molecular Biology*, 426(20), p. 3389-3412. doi: 10.1016/j.jmb.2014.03.016.

- Persico, A. M. *et al.* (1996) «D2 dopamine receptor gene TaqI A1 and B1 restriction fragment length polymorphisms: Enhanced frequencies in psychostimulant-preferring polysubstance abusers», *Biological Psychiatry*, 40(8), p. 776-784. doi: 10.1016/0006-3223(95)00483-1.
- Pettinati, H. M., O'Brien, C. P. i Dundon, W. D. (2013) «Current status of co-occurring mood and substance use disorders: A new therapeutic target», *American Journal of Psychiatry*, 170(1), p. 23-30. doi: 10.1176/appi.ajp.2012.12010112.
- Plichta, M. M. i Scheres, A. (2014) «Ventral-striatal responsiveness during reward anticipation in ADHD and its relation to trait impulsivity in the healthy population: a meta-analytic review of the fMRI literature», *Neuroscience and biobehavioral reviews*, 38, p. 125-134. doi: 10.1016/j.neubiorev.2013.07.012.
- Polanczyk, G. *et al.* (2007) «The worldwide prevalence of ADHD: A systematic review and meta-regression analysis», *American Journal of Psychiatry*, 164(6), p. 942-948. doi: 10.1176/appi.ajp.164.6.942.
- Raman, S. *et al.* (2009) «Structure prediction for CASP8 with all-atom refinement using Rosetta», *Proteins: Structure, Function and Bioinformatics*, 77, p. 89-99. doi: 10.1002/prot.22540.
- Rauch, S. L. *et al.* (2000) «Exaggerated amygdala response to masked facial stimuli in posttraumatic stress disorder: A functional MRI study», *Biological Psychiatry*, 47(9), p. 769-776. doi: 10.1016/S0006-3223(00)00828-3.
- Razin, A. i Kantor, B. (2005) «DNA methylation in epigenetic control of gene expression», *Progress in molecular and subcellular biology*, 38, p. 151-167. doi: 10.1007/3-540-27310-7_6.
- Rege, S. (2019) *Neurobiology of Attention Deficit Hyperactivity Disorder (ADHD)- A Primer*. Disponible a: <https://psychscenehub.com/psychinsights/neurobiology-of-adhd/> (Accedit: 7 juny 2020).
- Reichborn-Kjennerud, T. *et al.* (2004) «Genetic and environmental influences on binge eating in the absence of compensatory behaviors: A population-based twin study», *International Journal of Eating Disorders*, 36(3), p. 307-314. doi: 10.1002/eat.20047.
- Renthal, W. *et al.* (2007) «Histone Deacetylase 5 Epigenetically Controls Behavioral Adaptations to Chronic Emotional Stimuli», *Neuron*, 56(3), p. 517-529. doi: 10.1016/j.neuron.2007.09.032.
- Renthal, W. *et al.* (2009) «Genome-wide Analysis of Chromatin Regulation by Cocaine Reveals a Role for Sirtuins», *Neuron*, 62(3), p. 335-348. doi: 10.1016/j.neuron.2009.03.026.
- Ribasés, M., Sánchez-Mora, C., *et al.* (2012) «An association study of sequence variants in the forkhead box P2 (FOXP2) gene and adulthood attention-deficit/hyperactivity disorder in two European samples», *Psychiatric Genetics*, 22(4), p. 155-160. doi: 10.1097/YPG.0b013e328353957e.
- Ribasés, M., Ramos-Quiroga, J. A., *et al.* (2012) «Candidate system analysis in ADHD: Evaluation of nine genes involved in dopaminergic neurotransmission identifies association with DRD1», *World Journal of Biological Psychiatry*, 13(4), p. 281-292. doi: 10.3109/15622975.2011.584905.
- Ripke, S. *et al.* (2013) «Genome-wide association analysis identifies 13 new risk loci for schizophrenia», *Nature Genetics*, 45(10), p. 1150-1159.
- Ripke, S. *et al.* (2014) «Biological insights from 108 schizophrenia-associated genetic loci», *Nature*, 511, p. 421-427.
- Rita Lo Vasco, V., Cardinale, G. i Polonia, P. (2012) «Deletion of PLCB1 gene in schizophrenia-affected patients», *Journal of Cellular and Molecular Medicine*, 16(4), p. 844-851. doi: 10.1111/j.1582-4934.2011.01363.x.
- Robison, A. J. i Nestler, E. J. (2011) «Transcriptional and epigenetic mechanisms of addiction», *Nature Reviews Neuroscience*, 12(11), p. 623-637. doi: 10.1038/nrn3111.
- Rocha, B. A. *et al.* (1998) «Increased vulnerability to cocaine in mice lacking the serotonin-1B receptor», *Nature*, 393, p. 175-178.
- Rogge, G. A. i Wood, M. A. (2013) «The role of histone acetylation in cocaine-induced neural plasticity and behavior», *Neuropsychopharmacology*, 38(1), p. 94-110. doi: 10.1038/npp.2012.154.

Bibliografia

- Romo, L. *et al.* (2018) «Attention-deficit hyperactivity disorder and addictions (substance and behavioral): Prevalence and characteristics in a multicenter study in France», *Journal of Behavioral Addictions*, 7(3), p. 743-751. doi: 10.1556/2006.7.2018.58.
- Ross, S. i Peselow, E. (2012) «Co-occurring psychotic and addictive disorders: Neurobiology and diagnosis», *Clinical Neuropharmacology*, 35(5), p. 235-243. doi: 10.1097/WNF.0b013e318261e193.
- Ruso-Julve, F. *et al.* (2019) «Dopaminergic control of ADAMTS2 expression through cAMP/CREB and ERK: molecular effects of antipsychotics», *Translational Psychiatry*, 9(1), p. 306. doi: 10.1038/s41398-019-0647-7.
- Russell, V. A. (2000) «The nucleus accumbens motor-limbic interface of the spontaneously hypertensive rat as studied in vitro by the superfusion slice technique», *Neuroscience and Biobehavioral Reviews*, 24(1), p. 133-136. doi: 10.1016/S0149-7634(99)00056-1.
- Russo, S. J. *et al.* (2009) «Neurotrophic factors and structural plasticity in addiction», *Neuropharmacology*, 56(Suppl 1), p. 73-82. doi: 10.1016/j.neuropharm.2008.06.059.
- Sagvolden, T. (2000) «Behavioral validation of the spontaneously hypertensive rat (SHR) as an animal model of attention-deficit/hyperactivity disorder (AD/HD)», *Neuroscience and Biobehavioral Reviews*, 24(1), p. 31-39. doi: 10.1016/S0149-7634(99)00058-5.
- Samochowiec, J. *et al.* (2014) «Genetics of Alcohol Dependence: A Review of Clinical Studies», *Neuropsychobiology*, 70(2), p. 77-94.
- Sánchez-Mora, C. *et al.* (2010) «Meta-analysis of brain-derived neurotrophic factor p.Val66Met in adult ADHD in four European populations», *American Journal of Medical Genetics, Part B*, 153(2), p. 512-523. doi: 10.1002/ajmg.b.31008.
- Sánchez-Mora, C. *et al.* (2011) «Exploring DRD4 and its interaction with SLC6A3 as possible risk factors for adult ADHD: A meta-analysis in four European populations», *American Journal of Medical Genetics, Part B. Am J Med Genet B Neuropsychiatr Genet*, 156B(5), p. 600-612. doi: 10.1002/ajmg.b.31202.
- Sánchez-Mora, C. *et al.* (2013) «Evaluation of single nucleotide polymorphisms in the miR-183-96-182 cluster in adulthood attention-deficit and hyperactivity disorder (ADHD) and substance use disorders (SUDs)», *European Neuropsychopharmacology*, 23(11), p. 1463-1473. doi: 10.1016/j.euroneuro.2013.07.002.
- Sánchez-Mora, C. *et al.* (2015) «Case-control genome-wide association study of persistent attention-deficit hyperactivity disorder identifies FBXO33 as a novel susceptibility gene for the disorder», *Neuropsychopharmacology*, 40(4), p. 915-926. doi: 10.1038/npp.2014.267.
- Sanchez-Roige, S. *et al.* (2019) «Genome-wide association studies of impulsive personality traits (BIS-11 and UPPS-P) and drug experimentation in up to 22,861 adult research participants identify loci in the CACNA11 and CADM2 genes», *Journal of Neuroscience*, 39(13), p. 2562-2572. doi: 10.1523/JNEUROSCI.2662-18.2019.
- Sanchis-Segura, C. i Spanagel, R. (2006) «Behavioural assessment of drug reinforcement and addictive features in rodents: An overview», *Addiction Biology*, 11(1), p. 2-38. doi: 10.1111/j.1369-1600.2006.00012.x.
- Santucci, K. (2012) «Psychiatric disease and drug abuse», *Current Opinion in Pediatrics*, 24(2), p. 233-237. doi: 10.1097/MOP.0b013e3283504fbf.
- Sarnyai, Z. i Kovács, G. L. (1994) «Role of oxytocin in the neuroadaptation to drugs of abuse», *Psychoneuroendocrinology*, 19(1), p. 85-117. doi: 10.1016/0306-4530(94)90062-0.
- Sarnyai, Z. i Kovács, G. L. (2014) «Oxytocin in learning and addiction: From early discoveries to the present», *Pharmacology Biochemistry and Behavior*, 119, p. 3-9. doi: 10.1016/j.pbb.2013.11.019.
- Satagopan, J. M. i Elston, R. C. (2003) «Optimal two-stage genotyping in population-based association studies», *Genetic Epidemiology*, 25(2), p. 149-157. doi: 10.1002/gepi.10260.
- Schaub, M. A. *et al.* (2012) «Linking disease associations with regulatory information in the human genome», *Genome Research*, 22(9), p. 1748-1759. doi: 10.1101/gr.136127.111.

- Schmitt, A. *et al.* (2014) «The impact of environmental factors in severe psychiatric disorders», *Frontiers in Neuroscience*, 8, p. 19. doi: 10.3389/fnins.2014.00019.
- Schoonjans, A. S. *et al.* (2016) «PLCB1 epileptic encephalopathies; Review and expansion of the phenotypic spectrum», *European Journal of Paediatric Neurology*, 20(3), p. 474-479. doi: 10.1016/j.ejpn.2016.01.002.
- Schratt, G. (2009) «microRNAs at the synapse», *Nature reviews. Neuroscience*, 10(12), p. 842-849. doi: 10.1038/nrn2763.
- Schroeder, F. A. *et al.* (2008) «Drug-induced activation of dopamine D1 receptor signaling and inhibition of class I/II histone deacetylase induce chromatin remodeling in reward circuitry and modulate cocaine-related behaviors», *Neuropsychopharmacology*, 33(12), p. 2981-2992. doi: 10.1038/npp.2008.15.
- Schultz, W. (2002) «Getting formal with dopamine and reward», *Neuron*, 36(2), p. 241-263. doi: 10.1016/S0896-6273(02)00967-4.
- Schultz, W., Dayan, P. i Montague, P. R. (1997) «A neural substrate of prediction and reward», *Science*, 275(5306), p. 1593-1599. doi: 10.1126/science.275.5306.1593.
- Schwartz, N. B. i Domowicz, M. S. (2018) «Proteoglycans in brain development and pathogenesis», *FEBS Letters*, 592(23), p. 3791-3805. doi: 10.1002/1873-3468.13026.
- Seibenhener, M. L. i Wooten, M. C. (2015) «Use of the open field maze to measure locomotor and anxiety-like behavior in mice», *Journal of Visualized Experiments*, 96, p. e52434. doi: 10.3791/52434.
- Shen, H. Y. *et al.* (2008) «Additive effects of histone deacetylase inhibitors and amphetamine on histone H4 acetylation, cAMP responsive element binding protein phosphorylation and Δ FosB expression in the striatum and locomotor sensitization in mice», *Neuroscience*, 157(3), p. 644-655. doi: 10.1016/j.neuroscience.2008.09.019.
- Sherva, R. *et al.* (2010) «Variation in nicotinic acetylcholine receptor genes is associated with multiple substance dependence phenotypes», *Neuropsychopharmacology*, 35(9), p. 1921-1931.
- Shi, H., Kichaev, G. i Pasaniuc, B. (2016) «Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data», *American Journal of Human Genetics*, 99(1), p. 139-153. doi: 10.1016/j.ajhg.2016.05.013.
- Shi, J. *et al.* (2011) «Genome-wide association study of recurrent early-onset major depressive disorder», *Molecular Psychiatry*, 16(2), p. 193-201. doi: 10.1038/mp.2009.124.
- Shimabukuro, M. *et al.* (2007) «Global hypomethylation of peripheral leukocyte DNA in male patients with schizophrenia: A potential link between epigenetics and schizophrenia», *Journal of Psychiatric Research*, 41(12), p. 1042-1046. doi: 10.1016/j.jpsychires.2006.08.006.
- Simon, V. *et al.* (2009) «Prevalence and correlates of adult attention-deficit hyperactivity disorder: Meta-analysis», *British Journal of Psychiatry*, 194(3), p. 204-211. doi: 10.1192/bjp.bp.107.048827.
- Smalheiser, N. R. *et al.* (2012) «MicroRNA Expression Is Down-Regulated and Reorganized in Prefrontal Cortex of Depressed Suicide Subjects», *PLoS ONE*, 7(3), p. e33201. doi: 10.1371/journal.pone.0033201.
- Smith, A. C. W. i Kenny, P. J. (2018) «MicroRNAs regulate synaptic plasticity underlying drug addiction», *Genes, Brain and Behavior*, 17(3), p. 1-11. doi: 10.1111/gbb.12424.
- Song, Y. *et al.* (2013) «High-resolution comparative modeling with RosettaCM», *Structure*, 21(10), p. 1735-1742. doi: 10.1016/j.str.2013.08.005.
- Sora, I. *et al.* (2010) «Transgenic mice in the study of drug addiction and the effects of psychostimulant drugs», *Annals of the New York Academy of Sciences*, 1187, p. 218-246. doi: 10.1111/j.1749-6632.2009.05276.x.
- Soria, G. *et al.* (2005) «Lack of CBI cannabinoid receptor impairs cocaine self-administration», *Neuropsychopharmacology*, 30(9), p. 1670-1680. doi: 10.1038/sj.npp.1300707.
- Srivastav, S., Walitza, S. i Grünblatt, E. (2018) «Emerging role of miRNA in attention deficit hyperactivity

Bibliografia

- disorder: a systematic review», *ADHD Attention Deficit and Hyperactivity Disorders*, 10(1), p. 49-63. doi: 10.1007/s12402-017-0232-y.
- St Pourcain, B. *et al.* (2014) «Variability in the common genetic architecture of social-communication spectrum phenotypes during childhood and adolescence», *Molecular Autism*, 5, p. 18. doi: 10.1186/2040-2392-5-18.
- Steenfot, C. *et al.* (2013) «Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology», *EMBO Journal*, 32(10), p. 1478-1488. doi: 10.1038/emboj.2013.79.
- Stein, M. B. *et al.* (2002) «Increased amygdala activation to angry and contemptuous faces in generalized social phobia», *Archives of General Psychiatry*, 59(11), p. 1027-1034. doi: 10.1001/archpsyc.59.11.1027.
- Stice, E. *et al.* (2013) «The contribution of brain reward circuits to the obesity epidemic», *Neuroscience and biobehavioral reviews*, 37, p. 2047-2058. doi: 10.1016/j.neubiorev.2012.12.001.
- Stuber, G. D. *et al.* (2008) «Reward-predictive cues enhance excitatory synaptic strength onto midbrain dopamine neurons», *Science*, 321(5896), p. 1690-1692. doi: 10.1126/science.1160873.
- Sullivan, P. F. *et al.* (2018) «Psychiatric genomics: An update and an Agenda», *American Journal of Psychiatry*, 175(1), p. 15-27. doi: 10.1176/appi.ajp.2017.17030283.
- Sun, E. i Shi, Y. (2015) «MicroRNAs: Small molecules with big roles in neurodevelopment and diseases», *Experimental Neurology*, 268, p. 46-53. doi: 10.1016/j.expneurol.2014.08.005.
- Swerdlow, N. R. *et al.* (2008) «Realistic expectations of prepulse inhibition in translational models for schizophrenia research», *Psychopharmacology*, 199(3), p. 331-388. doi: 10.1007/s00213-008-1072-4.
- Tahiliani, M. *et al.* (2009) «Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1», *Science*, 324, p. 930-935. doi: 10.1126/science.1170116.
- Tak, Y. G. i Farnham, P. J. (2015) «Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome», *Epigenetics & chromatin*, 8(1), p. 57. doi: 10.1186/s13072-015-0050-4.
- Takao, K. i Miyakawa, T. (2006) «Light/dark transition test for mice», *Journal of Visualized Experiments*, 1, p. 104. doi: 10.3791/104.
- Tang, T. *et al.* (2010) «A mouse knockout library for secreted and transmembrane proteins», *Nature Biotechnology*, 28(7), p. 749-755. doi: 10.1038/nbt.1644.
- Taniguchi, M. *et al.* (2012) «Histone Deacetylase 5 Limits Cocaine Reward through cAMP-Induced Nuclear Import», *Neuron*, 73(1), p. 108-120. doi: 10.1016/j.neuron.2011.10.032.
- Tawa, E. A., Hall, S. D. i Lohoff, F. W. (2016) «Overview of the Genetics of Alcohol Use Disorder», *Alcohol and Alcoholism National Institutes of Health*, 51(5), p. 507-514.
- Torrens, M., Gilchrist, G. i Domingo-Salvany, A. (2011) «Psychiatric comorbidity in illicit drug users: Substance-induced versus independent disorders», *Drug and Alcohol Dependence*, 113, p. 147-156. doi: 10.1016/j.drugalcdep.2010.07.013.
- Tremolizzo, L. *et al.* (2002) «An epigenetic mouse model for molecular and behavioral neuropathologies related to schizophrenia vulnerability», *PNAS*, 99(26), p. 17095-17100. doi: 10.1073/pnas.262658999.
- Tremolizzo, L. *et al.* (2014) «Decreased whole-blood global DNA methylation is related to serum hormones in anorexia nervosa adolescents», *World Journal of Biological Psychiatry*, 15(4), p. 327-333. doi: 10.3109/15622975.2013.860467.
- Tripp, G. i Wickens, J. R. (2009) «Neurobiology of ADHD», *Neuropharmacology*, 57, p. 579-589. doi: 10.1016/j.neuropharm.2009.07.026.
- Tristán-Noguero, A. *et al.* (2013) «Lack of association between the LPR and VNTR polymorphisms of the serotonin transporter gene and cocaine dependence in a Spanish sample», *Psychiatry Research*, 210(3), p. 1287-1289. doi: 10.1016/j.psychres.2013.09.004.

- Tsankova, N. *et al.* (2007) «Epigenetic regulation in psychiatric disorders», *Nature Reviews Neuroscience*, p. 355-367. doi: 10.1038/nrn2132.
- Tsapakis, E. M. i Travis, M. J. (2002) «Glutamate and psychiatric disorders», *Advances in Psychiatric Treatment*, 8(3), p. 189-197. doi: 10.1192/apt.8.3.189.
- Tseng, P. T. *et al.* (2014) «Age-associated decrease in global DNA methylation in patients with major depression», *Neuropsychiatric Disease and Treatment*, 10, p. 2105-2114. doi: 10.2147/NDT.S71997.
- Tycko, B. (2010a) «Allele-specific DNA methylation: Beyond imprinting», *Human Molecular Genetics*, 19(2), p. R210-R220. doi: 10.1093/hmg/ddq376.
- Tycko, B. (2010b) «Mapping Allele-Specific DNA Methylation: A New Tool for Maximizing Information from GWAS», *American Journal of Human Genetics*, 86(2), p. 109-112. doi: 10.1016/j.ajhg.2010.01.021.
- Udawela, M. *et al.* (2011) «Phospholipase C beta 1 expression in the dorsolateral prefrontal cortex from patients with schizophrenia at different stages of illness», *Australian and New Zealand Journal of Psychiatry*, 45(2), p. 140-147. doi: 10.3109/00048674.2010.533364.
- Uhl, G. R., Koob, G. F. i Cable, J. (2019) «The neurobiology of addiction», *Annals of the New York Academy of Sciences*, 1451(1), p. 5-28. doi: 10.1111/nyas.13989.
- UNODC (2019) «World Drug Report», p. Sales No. E.19.XI.9.
- Valera, E. M. *et al.* (2007) «Meta-Analysis of Structural Imaging Findings in Attention-Deficit/Hyperactivity Disorder», *Biological Psychiatry*, 61(12), p. 1361-1369. doi: 10.1016/j.biopsych.2006.06.011.
- Valle, F. (1970) «Effects of Strain, Sex, and Illumination on Open-Field Behavior of Rats», *The American Journal of Psychology*, 83, p. 103-111.
- Valzania, A. *et al.* (2017) «Histone deacetylase 5 modulates the effects of social adversity in early life on cocaine-induced behavior», *Physiology and Behavior*, 171, p. 7-12. doi: 10.1016/j.physbeh.2016.12.027.
- Vandenbergh, D. *et al.* (1997) «High-activity catechol-O-methyltransferase allele is more prevalent in polysubstance abusers», *American Journal of Medical Genetics*, 74(4), p. 439-442.
- Vandenbergh, D. J. *et al.* (2000) «Long forms of the dopamine receptor (DRD4) gene VNTR are more prevalent in substance abusers: No interaction with functional alleles of the catechol-o-methyltransferase (COMT) gene», *American Journal of Medical Genetics*, 96(5), p. 678-683.
- de Villiers, A. S. *et al.* (1995) « β 2 Mediated inhibition of [3H]dopamine release from nucleus accumbens slices and monoamine levels in a rat model for attention-deficit hyperactivity disorder», *Neurochemical Research*, 20(4), p. 427-433. doi: 10.1007/BF00973098.
- Volkow, N. D. *et al.* (2002) «Role of dopamine, the frontal cortex and memory circuits in drug addiction: insight from imaging studies», *Neurobiology of learning and memory*, 78(3), p. 610-24.
- Volkow, N. D. *et al.* (2013a) «Obesity and addiction: Neurobiological overlaps», *Obesity Reviews*, 14(1), p. 2-18. doi: 10.1111/j.1467-789X.2012.01031.x.
- Volkow, N. D. *et al.* (2013b) «The addictive dimensionality of obesity», *Biological Psychiatry*, 73(9), p. 811-818. doi: 10.1016/j.biopsych.2012.12.020.
- Volkow, N. D., Fowler, J. S. i Wang, G. J. (2003) «The addicted human brain: Insights from imaging studies», *Journal of Clinical Investigation*, 111(10), p. 1444-1451. doi: 10.1172/JCI18533.
- Volkow, N. D., Koob, G. F. i McLellan, A. T. (2016) «Neurobiologic advances from the brain disease model of addiction», *New England Journal of Medicine*, 374(4), p. 363-371. doi: 10.1056/NEJMr1511480.
- Volkow, N. D. i Morales, M. (2015) «The Brain on Drugs: From Reward to Addiction», *Cell*, 162, p. 712-725. doi: 10.1016/j.cell.2015.07.046.
- Volkow, N. D., Wang, G. J. i Baler, R. D. (2011) «Reward, dopamine and the control of food intake: Implications for obesity», *Trends in Cognitive Sciences*, 15(1), p. 37-46. doi: 10.1016/j.tics.2010.11.001.

Bibliografia

- Volkow, N. D., Wise, R. A. i Baler, R. (2017) «The dopamine motive system: Implications for drug and food addiction», *Nature Reviews Neuroscience*, 18(12), p. 741-752. doi: 10.1038/nrn.2017.130gvf.
- Vorspan, F. *et al.* (2015) «Anxiety and Substance Use Disorders: Co-occurrence and Clinical Issues», *Current Psychiatry Reports*, 17(2). doi: 10.1007/s11920-014-0544-y.
- Walker, D. M. i Nestler, E. J. (2018) «Neuroepigenetics and addiction», en Geschwind, D. H., Paulson, H. L., i Klein, C. (ed.) *Handbook of Clinical Neurology. Neurogenetics, Part II*, p. 747-765. doi: 10.1016/B978-0-444-64076-5.00048-X.
- Walters, R. K. *et al.* (2018) «Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders», *Nature Neuroscience*, 21(12), p. 1656-1669. doi: 10.1038/s41593-018-0275-1.
- Wang, W., Kwon, E. J. i Tsai, L. H. (2012) «MicroRNAs in learning, memory, and neurological diseases», *Learning and Memory*, 19(9), p. 359-368. doi: 10.1101/lm.026492.112.
- Wang, X. *et al.* (2016) «Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures», *eLife*, 5, p. e10557. doi: 10.7554/eLife.10557.
- Wason, J. M. S. i Dudbridge, F. (2012) «A General framework for two-stage analysis of genome-wide association studies and its application to case-control studies», *American Journal of Human Genetics*, 90(5), p. 760-773. doi: 10.1016/j.ajhg.2012.03.007.
- Watanabe, K. *et al.* (2017) «Functional mapping and annotation of genetic associations with FUMA», *Nature Communications*, 8, p. 1826. doi: 10.1038/s41467-017-01261-5.
- Weissenberger, S. *et al.* (2017) «ADHD, lifestyles and comorbidities: A call for an holistic perspective - From medical to societal intervening factors», *Frontiers in Psychology*, 8, p. 454. doi: 10.3389/fpsyg.2017.00454.
- Wetherill, L. *et al.* (2008) «Neuropeptide Y receptor genes are associated with alcohol dependence, alcohol withdrawal phenotypes, and cocaine dependence», *Alcoholism: Clinical and Experimental Research*, 32(12), p. 2031-2040. doi: 10.1111/j.1530-0277.2008.00790.x.
- Wetherill, L. *et al.* (2014) «Association of substance dependence phenotypes in the COGA sample», *Addiction Biology*, 20(3), p. 617-627.
- Wetherill, L. *et al.* (2019) «Genome-wide association study identifies loci associated with liability to alcohol and drug dependence that is associated with variability in reward-related ventral striatum activity in African- and European-Americans», *Genes, Brain and Behavior*, 18(6), p. e12580. doi: 10.1111/gbb.12580.
- Wickens, J. R., Hyland, B. I. i Tripp, G. (2011) «Animal models to guide clinical drug development in ADHD: Lost in translation?», *British Journal of Pharmacology*, 164(4), p. 1107-1128. doi: 10.1111/j.1476-5381.2011.01412.x.
- Wise, R. A. (1996) «Addictive Drugs and Brain Stimulation Reward», *Annual Review of Neuroscience*, 19(1), p. 319-340. doi: 10.1146/annurev.ne.19.030196.001535.
- Wise, R. A. (2008) «Dopamine and reward: The anhedonia hypothesis 30 years on», *Neurotoxicity Research*, 14(i), p. 169-183. doi: 10.1007/BF03033808.
- Wong, C. C. Y., Mill, J. i Fernandes, C. (2011) «Drugs and addiction: an introduction to epigenetics», *Addiction*, 106(3), p. 480-489. doi: 10.1111/j.1360-0443.2010.03321.x.
- Wood, A. i Shilatifard, A. (2004) «Posttranslational modifications of histones by methylation», *Advances in Protein Chemistry*, 67, p. 201-222. doi: 10.1016/S0065-3233(04)67008-2.
- World Health Organization (2018) *Global status report on alcohol and health*. Geneva. Licence: BY-NC-SA 3.0 IGO.
- Wu, C. i Pan, W. (2019) «Integration of methylation QTL and enhancer–target gene maps with schizophrenia GWAS summary results identifies novel genes», *Bioinformatics*, 35(19), p. 3576–3583.
- Wu, L. *et al.* (2010) «A novel function of MicroRNA Let-7d in regulation of galectin-3 expression in attention

- deficit hyperactivity disorder rat brain», *Brain Pathology*, 20(6), p. 1042-1054. doi: 10.1111/j.1750-3639.2010.00410.x.
- Wu, L. H. *et al.* (2015) «Circulating MicroRNA Let-7d in Attention-Deficit/Hyperactivity Disorder», *NeuroMolecular Medicine*, 17(2), p. 137-146. doi: 10.1007/s12017-015-8345-y.
- Wu, L. H. *et al.* (2017) «Nr3C1-Bhlhb2 Axis Dysregulation Is Involved in the Development of Attention Deficit Hyperactivity», *Molecular Neurobiology*, 54(2), p. 1196-1212. doi: 10.1007/s12035-015-9679-z.
- Wu, Y. *et al.* (2018) «Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits», *Nature Communications*, 9, p. 918.
- Xia, X. *et al.* (2017) «NFAT5 protects astrocytes against oxygen–glucose–serum deprivation/restoration damage via the SIRT1/Nrf2 pathway», *Journal of Molecular Neuroscience*, 61(1), p. 96-104. doi: 10.1007/s12031-016-0849-x.
- Xing, B., Li, Y. C. i Gao, W. J. (2016) «Norepinephrine versus dopamine and their interaction in modulating synaptic function in the prefrontal cortex», *Brain Research*, 1641, p. 217-233. doi: 10.1016/j.brainres.2016.01.005.
- Yang, J. *et al.* (2010) «Common SNPs explain a large proportion of the heritability for human height», *Nature Genetics*, 42(7), p. 565-569. doi: 10.1038/ng.608.
- Yang, J. i Li, M. D. (2014) «Association and interaction analyses of 5-HT3 receptor and serotonin transporter genes with alcohol, cocaine, and nicotine dependence using the SAGE data», *Human Genetics*, 133(7), p. 905-918. doi: 10.1007/s00439-014-1431-7.
- Yang, L. *et al.* (2013) «Polygenic transmission and complex neuro developmental network for attention deficit hyperactivity disorder: Genome-wide association study of both common and rare variants», *American Journal of Medical Genetics Part B*, 162B, p. 419-430. doi: 10.1002/ajmg.b.32169.
- Yang, X.-L., Wang, X. i Peng, B.-W. (2018) «NFAT5 Has a Job in the Brain», *Review Dev Neurosci*, 40, p. 289-300.
- Yi, M.-H. *et al.* (2013) «NFAT5-Dependent Expression of AQP4 in Astrocytes», *Cellular and Molecular Neurobiology*, 33(2), p. 223-232.
- Yong, W. S., Hsu, F. M. i Chen, P. Y. (2016) «Profiling genome-wide DNA methylation», *Epigenetics and Chromatin*, 9, p. 26. doi: 10.1186/s13072-016-0075-3.
- Young, A. I. (2019) «Solving the missing heritability problem», *PLoS Genetics*, 15(6), p. e1008222. doi: 10.1371/journal.pgen.1008222.
- Zallar, L. J. *et al.* (2017) «The Role of the Ghrelin System in Drug Addiction», en *International Review of Neurobiology*, p. 89-119. doi: 10.1016/bs.irn.2017.08.002.
- Zayats, T. *et al.* (2015) «Genome-wide analysis of attention deficit hyperactivity disorder in Norway», *PLoS ONE*, 10(4), p. e0122501. doi: 10.1371/journal.pone.0122501.
- Zhang, D. *et al.* (2010) «Genetic Control of Individual Differences in Gene-Specific Methylation in Human Brain», *American Journal of Human Genetics*, 86(3), p. 411-419.
- Zhang, H. *et al.* (2008) «The OPRD1 and OPRK1 loci in alcohol or drug dependence: OPRD1 variation modulates substance dependence risk», *Molecular Psychiatry*, 13(5), p. 531-543. doi: 10.1038/sj.mp.4002035.
- Zhang, L. *et al.* (2012) «ADHDgene: A genetic database for attention deficit hyperactivity disorder», *Nucleic Acids Research*, 40(D1), p. 1003-1009. doi: 10.1093/nar/gkr992.
- Zhou, Z. *et al.* (2011) «Substance-specific and shared transcription and epigenetic changes in the human hippocampus chronically exposed to cocaine and alcohol», *PNAS*, 108(16), p. 6626-6631. doi: 10.1073/pnas.1018514108.
- Zuo, L. *et al.* (2009) «Interaction between Two Independent CNR1 Variants Increases Risk for Cocaine

Bibliografia

Dependence in European Americans: A Replication Study in Family-Based Sample and Population-Based Sample», *Neuropsychopharmacology*, 34(6), p. 1504-1513.

ANNEX

Article

Involvement of the 14-3-3 Gene Family in Autism Spectrum Disorder and Schizophrenia: Genetics, Transcriptomics and Functional Analyses

Bàrbara Torrico ^{1,2,3,4,†}, Ester Antón-Galindo ^{1,2,3,4,†}, Noèlia Fernández-Castillo ^{1,2,3,4,†}, Eva Rojo-Francàs ^{1,2,3,4}, Sadaf Ghorbani ⁵, Laura Pineda-Cirera ^{1,2,3,4}, Amaia Hervás ^{6,7}, Isabel Rueda ⁶, Estefanía Moreno ^{3,8}, Janice M. Fullerton ^{9,10}, Vicent Casadó ^{3,8}, Jan K. Buitelaar ^{11,12}, Nanda Rommelse ^{12,13}, Barbara Franke ^{13,14}, Andreas Reif ¹⁵, Andreas G. Chiocchetti ¹⁶, Christine Freitag ¹⁶, Rune Kleppe ^{5,17}, Jan Haavik ⁵, Claudio Toma ^{1,9,10,18,*} and Bru Cormand ^{1,2,3,4,*}

- ¹ Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Preosti Building, floor 2, Av. Diagonal 643, 08028 Barcelona, Spain; barticoa@gmail.com (B.T.); eantongalindo@ub.edu (E.A.-G.); noefernandez@ub.edu (N.F.-C.); erojofra20@alumnes.ub.edu (E.R.-F.); l.pineda@ub.edu (L.P.-C.)
- ² Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, 28029 Madrid, Spain
- ³ Institut de Biomedicina de la Universitat de Barcelona (IBUB), 08028 Barcelona, Spain; estefaniamoreno@ub.edu (E.M.); vcasado@ub.edu (V.C.)
- ⁴ Institut de Recerca Sant Joan de Déu (IR-SJD), 08950 Esplugues de Llobregat, Spain
- ⁵ Centre for Neuropsychiatric Disorders, Department of Biomedicine, University of Bergen, N5009 Bergen, Norway; Sadaf.Ghorbani@uib.no (S.G.); Rune.Kleppe@uib.no (R.K.); Jan.Haavik@uib.no (J.H.)
- ⁶ Child and Adolescent Mental Health Unit, Hospital Universitari Mútua de Terrassa, 08221 Terrassa, Spain; ahervas@mutuaterrassa.cat (A.H.); irueda@sjdhospitalbarcelona.org (I.R.)
- ⁷ IGAÍN, Global Institute of Integral Attention to Neurodevelopment, 08007 Barcelona, Spain
- ⁸ Department of Biochemistry and Molecular Biomedicine, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain
- ⁹ Neuroscience Research Australia, Sydney, NSW 2031, Australia; j.fullerton@neura.edu.au
- ¹⁰ School of Medical Sciences, University of New South Wales, Sydney, NSW 2052, Australia
- ¹¹ Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, 6525 HR Nijmegen, The Netherlands; Jan.Buitelaar@radboudumc.nl
- ¹² Karakter Child and Adolescent Psychiatry University Centre, 6525 GC Nijmegen, The Netherlands; Nanda.Lambrechts-Rommelse@radboudumc.nl
- ¹³ Department of Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, 6525 HR Nijmegen, The Netherlands; Barbara.Franke@radboudumc.nl
- ¹⁴ Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, 6525 HR Nijmegen, The Netherlands
- ¹⁵ Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt, 60590 Frankfurt am Main, Germany; andreas.reif@kgu.de
- ¹⁶ Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Autism Research and Intervention Center of Excellence Frankfurt, JW Goethe University, 60323 Frankfurt am Main, Germany; andreas.chiocchetti@kgu.de (A.G.C.); christinemargarete.freitag@kgu.de (C.F.)
- ¹⁷ Division of Psychiatry, Haukeland University Hospital, 5021 Bergen, Norway
- ¹⁸ Centro de Biología Molecular “Severo Ochoa”, Universidad Autónoma de Madrid/CSIC, C/Nicolás Cabrera, 1, Campus UAM, 28049 Madrid, Spain
- * Correspondence: claudio.toma@cbm.csic.es (C.T.); bcormand@ub.edu (B.C.); Tel.: +34-911-964-731 (C.T.); +34-934-021-013 (B.C.)
- † These authors equally contributed to this work.
- ‡ These authors equally supervised this work.

Received: 30 April 2020; Accepted: 10 June 2020; Published: 13 June 2020



Abstract: The 14-3-3 protein family are molecular chaperones involved in several biological functions and neurological diseases. We previously pinpointed *YWHAZ* (encoding 14-3-3 ζ) as a candidate gene for autism spectrum disorder (ASD) through a whole-exome sequencing study, which identified a frameshift variant within the gene (c.659-660insT, p.L220Ffs*18). Here, we explored the contribution of the seven human 14-3-3 family members in ASD and other psychiatric disorders by investigating the: (i) functional impact of the 14-3-3 ζ mutation p.L220Ffs*18 by assessing solubility, target binding and dimerization; (ii) contribution of common risk variants in 14-3-3 genes to ASD and additional psychiatric disorders; (iii) burden of rare variants in ASD and schizophrenia; and iv) 14-3-3 gene expression using ASD and schizophrenia transcriptomic data. We found that the mutant 14-3-3 ζ protein had decreased solubility and lost its ability to form heterodimers and bind to its target tyrosine hydroxylase. Gene-based analyses using publicly available datasets revealed that common variants in *YWHAZ* contribute to schizophrenia ($p = 6.6 \times 10^{-7}$), whereas ultra-rare variants were found enriched in ASD across the 14-3-3 genes ($p = 0.017$) and in schizophrenia for *YWHAZ* (meta- $p = 0.017$). Furthermore, expression of 14-3-3 genes was altered in post-mortem brains of ASD and schizophrenia patients. Our study supports a role for the 14-3-3 family in ASD and schizophrenia.

Keywords: autism; 14-3-3 gene family; rare variants; common variants; transcriptomics; schizophrenia; *YWHAZ*; *YWHAZ*

1. Introduction

Autism spectrum disorder (ASD) is characterized by impairments in communication and social interactions, and the presence of repetitive and restrictive behaviours [1]. However, the clinical picture is often accompanied by additional features, such as intellectual disability (ID), epilepsy, language impairment, anxiety, sleep disorders, and attention-deficit hyperactivity disorder (ADHD) [2]. ASD has a prevalence of approximately 1.5% in the general population [3,4] and large studies suggest an unequivocal genetic contribution to its aetiology. Indeed, family and twin studies indicate a heritability of around 80%, which represents one of the highest amongst neuropsychiatric disorders [5–7]. Despite the substantial role of genetic factors in the disorder, the genetic architecture is not fully dissected and many of the underlying genes are yet to be identified. Also, the genetic relationships amongst comorbid phenotypes remain largely unknown.

Genetic studies suggest a multi-hit model of inheritance in which a combination of risk alleles, including both common variants of small effect size and rare variants with higher penetrance, contribute to the phenotype [8,9]. Despite the large number of association studies in the last decade, the robust identification of common risk alleles for ASD has been elusive, and the first genome-wide association studies (GWAS) performed with reasonably large samples [10–12] pinpointed associations with single nucleotide polymorphisms (SNPs) that were subsequently not replicated in a large European sample [13]. A GWAS recently performed in 18,381 ASD patients and 27,969 controls identified five genome-wide significant loci [14], suggesting the need of large samples to identify common risk variants. Regarding rare variants in ASD, the first whole-exome sequencing (WES) studies focused on de novo variants and suggested a substantial role for this class of damaging mutations in the aetiology of the disorder [15–17]. These approaches were crucial to pinpoint novel genes involved in the disorder [18], and recently the largest WES study reported 102 ASD risk genes in a sample of 11,896 cases [19]. WES studies were also performed to address the impact of rare inherited variants [18,20–22]. We performed the first WES study in multiplex families with autism that suggested a role for genome-wide truncating mutations in the aetiology of ASD [22]. A current genetic model would implicate a higher number of gene-disrupting variants in severe ASD phenotypes, increasing symptom severity and comorbidity with ID [2,19].

In our first WES study [22] we identified a truncating mutation in the *YWHAZ* gene (c.659-660insT, p.L220Ffs*18) that was transmitted from a mother with depression to an ASD sib-pair. In that previous study, the *YWHAZ* gene was the main interconnected node in a network including mutated genes identified by previous WES studies and other genes implicated in ASD. This gene encodes 14-3-3 ζ , a protein involved in a range of biological processes including cell cycle, transcription, neuronal development, migration and neurite outgrowth [23–25]. 14-3-3 ζ is one of the seven members (β , γ , ϵ , η , ζ , σ , θ) of the highly conserved 14-3-3 protein family. These proteins exert their function as homo- and heterodimers through protein-protein interaction with a wide range of target proteins, typically binding to phosphorylated serine and threonine residues [26]. Several studies have demonstrated the role of *YWHAZ* and other family members in neurogenesis and neurodifferentiation, and its possible implication in neurodevelopmental disorders [24,27]. Recent studies suggested that decreased 14-3-3 ζ protein levels in ASD may be responsible for deficits in melatonin synthesis observed in ASD via the downregulation of the aralkylamine N-acetyltransferase (AANAT) and acetylserotonin o-methyltransferase (ASMT) enzymatic activity [28–30]. Moreover, *YWHAZ* knock-out mice show impaired cortical development and larger lateral ventricles, reduced dendritic and synaptic density, aberrant neuronal migration in hippocampus, abnormal mossy fibers connectivity, and cognitive deficits [31–33].

Interestingly, several members of the 14-3-3 gene family have been implicated in psychiatric disorders, including reported associations of *YWHAB*, *YWHAE*, *YWHAZ* and *YWHAH* with schizophrenia [34–39]; *YWHAE* and *YWHAQ* with ADHD [39]; *YWHAH*, *YWHAG* and *YWHAE* with bipolar disorder [39,40]; *YWHAE* and *YWHAQ* with major depressive disorder [39,41]; and *YWHAE* with suicide attempts [42]. However, most of these studies were performed in small samples and lacked replication in larger populations. Furthermore, altered levels of 14-3-3 proteins were found in the blood or brains of patients with ASD [29,43,44], schizophrenia [45–48] and bipolar disorder [48]. Interestingly, microduplications of *YWHAE*, which encodes 14-3-3 ϵ that form stable heterodimers with 14-3-3 ζ , were reported in ASD patients; whereas microdeletions involving both *YWHAE* and *PFAH1B1* genes cause Miller–Dieker syndrome, a form of lissencephaly with ID and seizures [49–53], and deletions including *YWHAG* and *HIP1* were related to epilepsy, learning problems and ID [54].

Animal models deficient for 14-3-3 proteins show a variety of behavioural manifestations related to psychiatric disorders: *Ywhaz* (14-3-3 ζ) deficient mice present hyperactivity, impaired memory, lower anxiety and impaired sensorimotor gating [32,33]; *Ywhae* (14-3-3 ϵ) deficient mice present enhanced anxiety-like behaviour, defects in working memory, increased locomotor activity and sociability [34,55]; and heterozygous knock-out mice deficient for *Ywhag* (14-3-3 γ) show hyperactivity and depressive-like behaviour [56]. Furthermore, a recent study reported that the specific inhibition of 14-3-3 proteins in the hippocampus of mice is sufficient to cause hyperactivity, reduce sensorimotor gating and impair associative learning and memory [57].

Based on these aforementioned reports, we hypothesize that the 14-3-3 gene family may play an important role in the susceptibility to psychiatric disorders. Thus, here we aim to: (i) gain molecular insight for the role of the *YWHAZ* truncating mutation p.L220Ffs*18 previously reported in a family with distinct psychiatric disorders; (ii) assess the contribution of common and rare variants of the 14-3-3 gene family (*SFN*, *YWHAB*, *YWHAE*, *YWHAG*, *YWHAH*, *YWHAQ* and *YWHAZ*) to ASD and other psychiatric disorders; and (iii) explore possible altered expression levels of this gene family in psychiatric disorders.

2. Experimental Section

2.1. Expression, Purification and Solubility Testing of Recombinant 14-3-3 ζ Wild-Type and Mutated Proteins

The recombinant human *YWHAZ* wild-type (WT) and mutant were expressed in *E. coli* using pGEX-2TK expression system, and purification of the soluble fractions as fusion proteins with glutathione-S-transferase (GST) was performed according to previous protocols [58], see also

supplementary information for details. The 14-3-3ζ mutated protein replicating the C-terminal amino acid alteration from the p.L220Ffs*18 insertion mutation, was generated using site-directed mutagenesis as detailed in the supplementary information.

Due to low solubility of the GST-14-3-3ζ_{mut} protein we tested expression and solubility at different temperatures (induction time, h) of 30 (4 h), 25 (5 h) and 20 (6 h) °C and compared total and soluble lysate to that of WT on sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE, see supplementary information for details).

2.2. Functional Assessment of 14-3-3ζ WT and Mutant Proteins by Surface Plasmon Resonance

One of the best-characterized binding targets of 14-3-3 proteins is tyrosine hydroxylase (TH) phosphorylated at serine 19 (THpSer19) [58,59]. Human TH was purified and phosphorylated on serine 19 using p38 regulated/activated protein kinase (PRAK, also referred to as MK2) as previously described [60]. The WT and mutated 14-3-3ζ proteins were assessed for their binding to THpS19 or non-phosphorylated TH using surface plasmon resonance (Biacore 3000, Cytiva, Marlborough, MA, USA) as previously described [58,60]. GST-14-3-3 proteins were immobilized on CM5 sensor chips according to manufacturer's instructions using the GST-capture kit (Cytiva) giving similar amounts of immobilized GST-14-3-3ζ_{WT} and GST-14-3-3ζ_{mut}. Target protein binding was assessed at 25 °C, using the HBS-P Buffer provided by the manufacturer (10 mM HEPES pH 7.4, 150 mM NaCl and 0.005% polysorbate 20). Different concentrations of Ser19 phosphorylated TH were injected multiple times for multiple immobilizations. We used the unphosphorylated TH as a negative control of the binding. The resulting sensograms were analyzed with the BIAevaluation v3.2 software (Biacore AB, Uppsala, Sweden).

2.3. Assessing the Dimerization of Mutant and Wild-Type Proteins (14-3-3ζ and 14-3-3σ) in Cells by Bioluminescence Resonance Energy Transfer (BRET) Assay

Plasmids were obtained for expressing fusion proteins of different 14-3-3 members (14-3-3ζ, 14-3-3ε, 14-3-3σ) with Rluc (Renilla luciferase, donor) and EYFP (enhanced yellow variant of GFP, acceptor), as described in the supplementary information. The ability of mutant 14-3-3ζ to form heterodimers with 14-3-3ε, and of mutant 14-3-3σ to form homodimers, was assessed using a bioluminescence resonance energy transfer (BRET) assay.

Human embryonic kidney cells (HEK-293T) were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, 100 U/mL penicillin and 100 µg/mL streptomycin (GIBCO, Carlsbad, CA, USA), in a 5% CO₂ humidified atmosphere at 37 °C. The cell line was grown in 6-well plates (35-mm-diameter wells) at a density of 5.0×10^5 cells/well for transfection using CaCl₂ as described in the supplementary information. Cells were co-transfected with the cDNA construct coding for Rluc-target-protein-1, acting as BRET donor, and increasing amounts of the cDNA construct coding for YFP-target-protein-2 as BRET acceptor (see Table S1 for the amounts of cDNA used). D(1A) dopamine receptor (DRD1) fusion protein with YFP or Rluc [61] was used as a negative control of the dimerization, and co-transfected with the corresponding tested constructs (Table S1). After 48 h upon transfection, cells were washed twice with Hanks' balanced salt solution pH 7.4 (HBSS, 137 mM NaCl, 5 mM KCl, 1.26 mM CaCl₂, 0.4 mM MgSO₄, 0.5 mM MgCl₂, 0.34 mM Na₂HPO₄, 0.44 mM KH₂PO₄, 10 mM HEPES) supplemented with 1% glucose (*w/v*), detached and resuspended in the same buffer. Protein concentration was determined using the Bradford assay kit (Bio-Rad, Munich, Germany) and all cell suspensions were diluted with HBSS to obtain a final concentration of 0.2 mg/mL of protein.

In order to quantify fluorescence, cell suspensions (20 µg of protein) were distributed in duplicates in a 96-well black microplate with a transparent bottom (Porvair, King's Lynn, UK). Fluorescence was then measured using a FLUOstar Optima fluorimeter (BMG Labtechnologies, Offenburg, Germany) equipped with a high-energy xenon flash lamp, and a 10 nm bandwidth excitation filter at 400 nm reading. A PHERAstar Flagship FSX fluorimeter (BMG Labtechnologies, Offenburg, Germany) was used for BRET and luminescence measurements. Cell suspensions (20 µg of protein) were distributed

in duplicates in a 96-well white opaque microplate (Porvair, King's Lynn, UK) and coelenterazine H (Molecular Probes Europe, Leiden, The Netherlands) was added at a final concentration of 5 mM. One minute after adding coelenterazine H, luminescence readings were collected using sequential integration of signals detected at 440–500 nm and 510–590 nm. Luminescence measurements of the same samples were performed after 10 min of incubation with coelenterazine H. Cells expressing BRET donors alone were used to determine background. The BRET ratio is defined as: $((\text{emission at } 510\text{--}590 \text{ nm})/(\text{emission at } 440\text{--}500 \text{ nm}))\text{--}C_f$; where C_f corresponds to $(\text{emission at } 510\text{--}590 \text{ nm})/(\text{emission at } 440\text{--}500 \text{ nm})$ for the donor construct expressed alone in the same experiment. Data were fitted to a non-linear regression equation, assuming a single phase saturation curve with GraphPad Prism software (San Diego, CA, USA).

2.4. Common Variants in the 14-3-3 Family in Autism Spectrum Disorder (ASD): Association Study in Our Sample

2.4.1. Subjects of Our ASD Cohorts

The cohort used in the case-control association study consisted of 727 ASD patients and 714 gender-matched controls from three European populations: Spanish, Dutch and German (Table S2). The cohort used for mutational screening employed a subset of 288 ASD patients from the same three populations (Table S2). All individuals had European ancestry. ASD patients were assessed using the ADI-R (Autism-Diagnostic Interview-Revised) [62] and, where possible, also the ADOS (Autism Diagnostic Observation Schedule) [63]. Cytogenetic abnormalities or a positive Fragile X test were considered exclusion criteria. The study was approved by the relevant ethics committee from each center and written informed consent was obtained from all parents/guardians or, where possible, by affected individuals, according to the Helsinki Declaration. Genomic DNA was extracted from peripheral blood samples using the standard salting-out method [64].

2.4.2. Common Variant Association Study of the 14-3-3 Family in Our ASD Sample

SNP selection was performed to encompass common genetic variants across the seven genes of the 14-3-3 family (*SFN*, *YWHAQ*, *YWHAG*, *YWHAZ*, *YWHAB*, *YWHAH* and *YWHAE*). Each gene included a 5kb flanking region (both 5' and 3'), and patterns of linkage disequilibrium (LD) were considered using the Central European (CEU) panel of HapMap project data (www.hapmap.org; phases 1, 2, 3; release 28). A total of 42 tagSNPs were selected using the Tagger implementation in HaploView v4.2 [65], according to the following criteria: $r^2 \leq 0.8$ and minor allele frequency (MAF) ≥ 0.05 . The sample of 1441 subjects was genotyped using iPlex-Sequenom technology (Sequenom, San Diego, CA, USA) at the Spanish National Genotyping Center (CeGen). Duplicates were included as controls for genotyping quality. After quality control procedures, 36 individuals with genotyping rate lower than 90% were removed from the study, setting the final sample to 1405 individuals (713 cases and 692 controls). From the 42 SNPs initially genotyped, four assays failed, and one SNP was monomorphic. The genotyping rate for the 37 remaining SNPs was 99.2%. One of the SNPs was excluded for quality reasons, but all the remaining 36 SNPs were in Hardy–Weinberg equilibrium (threshold set at $p < 0.01$ in controls). Thirty-four SNPs showed a MAF > 0.05 in our sample and were examined for association with ASD (Table S3). LD patterns and D' values were determined in our sample data with Haploview v4.2 [65] (Figure S1). A case-control association study under the additive model was run with the PLINK package [66].

2.5. Common Variant Associations with the 14-3-3 Gene Family across Psychiatric Disorders Using Public Genome-Wide Association Studies (GWAS) Data

We assessed the contribution of common variation in the seven 14-3-3 genes to psychiatric disorders using GWAS summary statistics from the Psychiatric Genomics Consortium (PGC), Broad Antisocial Behaviour Consortium (BroadABC) and Integrative Psychiatric Research Consortium (iPSYCH). We considered the following phenotypes: attention-deficit/hyperactivity disorder (ADHD) [67], anti-social behaviour (ASB) [68], anxiety [69], autism spectrum disorder (ASD) [14], bipolar disorder (BD) [70], major depressive disorder (MDD) [71], obsessive-compulsive disorder (OCD) [72], schizophrenia [73] and cross-disorder meta-analysis [74] (see details in Table S4). All summary statistics used for subsequent analysis had a MAF \geq 0.01 and info-score for imputation quality \geq 0.6. A gene-based association study was performed with MAGMA (v1.06) [75] using the 1000 Genomes Project Phase 3 (Build 37/European data only) as a reference panel. We also performed a self-contained gene-set analysis considering the whole 14-3-3 gene family for each of the eight psychiatric phenotypes and for the cross-disorder meta-analysis.

2.6. Rare Variant Analysis of the 14-3-3 Gene Family: Mutational Screening in Our ASD Sample

The seven genes of the 14-3-3 family were analyzed in 288 Caucasian ASD patients by high-throughput sequencing using the Ion Torrent platform (ThermoFisher Scientific, Waltham, MA, USA) at the Centre for Research in Agricultural Genomics (CRAG). A total of 57 tagged-primer pairs were designed with the Ion Ampliseq Designer (ThermoFisher Scientific) and covered 96.3% of all coding exons across the 14-3-3 gene family (Table S5), including the splice sites and part of the 5' and 3' untranslated regions (UTR) (Figure S2). The corresponding amplicons were sequenced in 288 ASD patients (182 Spanish, 94 Dutch, 12 Germans), including the MT_160.3 patient, heterozygous carrier for the previously described c.659_660insT mutation leading to p.L220Ffs*18 in the *YWHAZ* gene [22], as a positive control. Genomic DNA was pooled in groups of three subjects to minimize costs. The results were processed using the Ion Reporter Software (ThermoFisher Scientific) under somatic variant calling parameters to identify low-frequency variant calls (average read depth 1400X), and the function Variant Analysis of the Ingenuity Pathway Analysis software (<http://www.ingenuity.com/products/ipa>) was employed to assess the identification and the molecular nature of the variants identified in our sample. To handle analysis of pooled sequencing, every single change identified in a pool was subjected to Sanger validation in the three subjects present in the reaction, in order to confirm the variant and determine the carrier status of each individual.

2.7. Rare Variant Analysis of the 14-3-3 Gene Family: ASD and Schizophrenia Public Datasets

The impact of rare variants across the seven genes of the 14-3-3 family was assessed and further extended using publicly available sequencing data of schizophrenia, ASD and control cohorts from the following sources: (i) whole-exome sequencing (WES) from the Sweden-Schizophrenia population-based case-control (database of Genotypes and Phenotypes (dbGAP) accession: phs000473.v2.p2) (6135 cases and 6245 controls); (ii) ARRA Autism Sequencing Collaboration (dbGAP accession: phs000298.v3.p2) (1288 unrelated ASD probands); (iii) European ASD samples, which included our previously sequenced sample of 288 ASD patients plus 348 additional ASD patients from Germany; (iv) Medical Genome Reference Bank (2845 healthy Caucasian Australians aged > 75). The selection of variants was based on: (1) their predicted pathogenicity using the Variant Effect Predictor annotation tool software (<https://www.ensembl.org/Tools/VEP>): missense mutations predicted to be damaging in both SIFT (Sorting Intolerant From Tolerant) and Polyphen2 (Polymorphism Phenotyping v2)), and with CADD (Combined Annotation Dependent Depletion) > 20 for canonical splice site variants, stop codon mutations and indels leading to frameshift; and (2) minor allele frequency (MAF < 0.0001) in non-Finnish European populations from the Genome Aggregation Database (<http://gnomad.broadinstitute.org/>) as previously described [18]. A chi-square statistic was used to compare the schizophrenia patient sample (6135 cases) and combined ASD datasets (1924 cases)

with the combined control datasets (9090 individuals). Additionally, we explored the impact of rare variants for each of the 14-3-3 genes as reported by the Autism Sequencing Consortium (ASC) (<https://asc.broadinstitute.org/>) and the Schizophrenia Exome-sequencing Meta-Analysis (SCHEMA) (<https://schema.broadinstitute.org/>). ASC represents the largest source of data of rare variants for genetic studies of ASD, which includes de novo variant calls for 6430 probands and 2179 unaffected siblings (family-based dataset), and rare variants identified in 5556 ASD patients and 8809 controls (case-control dataset). Similarly, SCHEMA combines data for rare variants across several world-wide populations with a joint sample of 24,248 schizophrenia patients and 97,322 controls (case-control dataset), and 3444 schizophrenia trios (family-based dataset).

2.8. Expression of the 14-3-3 Genes in Autism Spectrum Disorder and Schizophrenia

Differential expression of the seven 14-3-3 genes was assessed using transcriptomic data from post-mortem brain regions of ASD and schizophrenia patients in publicly available human datasets, either in the Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo>) or in published articles. We found 39 studies with available information on gene expression in brain: 12 in ASD patients and 27 in schizophrenia patients. In particular, in the case of ASD we analysed data on altered gene expression in brain from 11 papers and one GEO dataset (PubMed ID: 27919067, 29859039, 21614001, 22457638, 22984548, 18006270, 18378158, 25494366, 27219343, 27685936, 30545856 and GSE38322). For schizophrenia we analysed data on brain gene expression from 14 papers (30545856, 25113377, 24287731, 21091092, 24167345, 23904455, 24686180, 24886351, 22031440, 18778695, 26818902, 22212594, 22954356, 29931221) and 13 GEO datasets (GSE46509, GSE37981, GSE21935, GSE21138, GSE17612, GSE12654, GSE87610, GSE53987, GSE62191, GSE35977, GSE12649, GSE12679, GSE35978). Differential expression was assessed in multiple brain areas, including hippocampus, cerebellum or cortex, depending on the dataset.

3. Results

3.1. Functional Effect of the YWHAZ (14-3-3 ζ) Mutation p.L220Ffs*18

We firstly aimed to assess the functional consequences of a truncating variant (c.659_660insT, p.L220Ffs*18) previously identified in the YWHAZ gene and transmitted from a mother with depression, phobia and fibromyalgia to two siblings both with ASD and attention-deficit hyperactivity disorder (ADHD) [22]. The maternal grandmother was diagnosed with schizophrenia (Figure 1A), although DNA from this case was unfortunately not available for mutation analysis.

This truncating variant does not meet criteria to trigger degradation of the corresponding mRNA by nonsense-mediated RNA decay (NMD), since it is located 22 nucleotides before the last exon-exon junction (Figure 1B) [76,77]. Therefore, we performed experiments to investigate the possible functional effect of p.L220Ffs*18 on the protein by assessing: (i) its solubility; (ii) its ability to bind tyrosine hydroxylase (TH), one of its natural targets when phosphorylated at its 14-3-3 binding site (THpSer19); and (iii) its capacity to form heterodimers with 14-3-3 ϵ .

The mutated 14-3-3 ζ protein (p.L220Ffs*18) showed decreased solubility compared to the WT 14-3-3 ζ when it was expressed in prokaryotes (*E. coli*) (Figure 1C,D). The solubility did not show any improvement at different temperatures for the mutated protein (Figure 1C). The fraction of soluble mutated protein was lower than 5% (compared to 90% for the WT protein), with the vast majority remaining insoluble in the lysate fraction (Figure 1D).

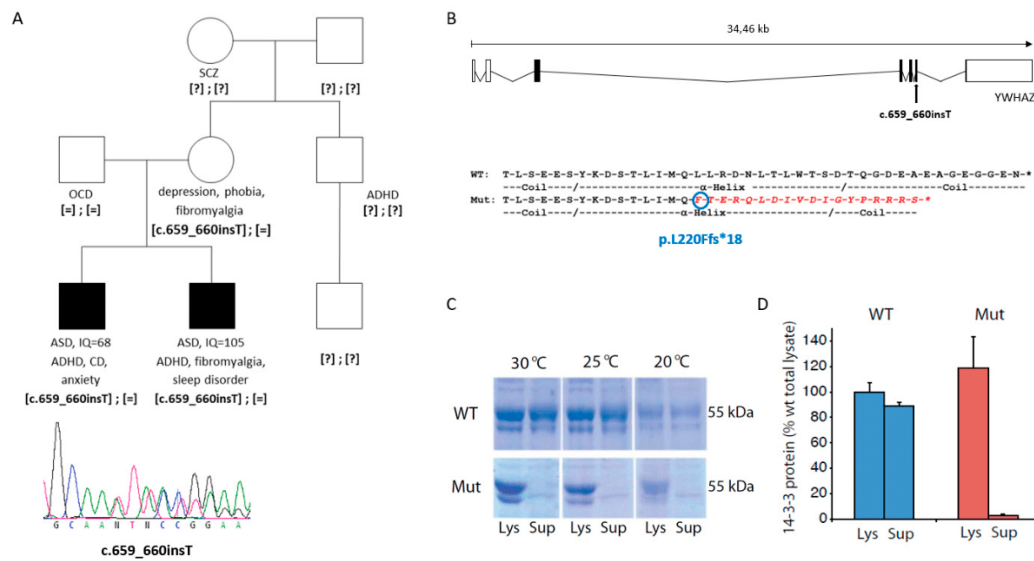


Figure 1. Mutation identified in *YWHAZ* in two siblings with ASD. (A) Pedigree of the family carrying the c.659-660insT mutation in *YWHAZ* and below the Sanger sequence of the truncating variant. Abbreviations: ADHD, attention-deficit hyperactivity disorder; ASD, autism spectrum disorder; CD, conduct disorder; IQ, intelligence quotient; OCD, obsessive-compulsive disorder; SCZ, schizophrenia. [=], wild-type allele; [?], unknown genotype. (B) Location of the mutation in the *YWHAZ* gene, and comparison of the last amino acids of the wild-type (WT) and mutant (Mut) 14-3-3 ζ protein, showing in red the amino acids that diverge in the mutant protein. (C) Coomassie stained sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) of glutathione-S-transferase (GST)-14-3-3 ζ (WT and mutant) expressed in *E. coli* strain at different temperatures. Aliquots of the bacteria were lysed, and the amount of soluble protein was assessed by comparing total lysate (Lys) to supernatant after centrifugation of the lysate (Sup) for wild-type 14-3-3 ζ (WT) and mutant 14-3-3 ζ (Mut). (D) Quantification of solubility by measuring the major GST-14-3-3 band (55 kDa) for WT and Mut expressed at 30 °C. Data presented as means and error bars denote the standard deviation ($n = 3$; $p = 1.7 \times 10^{-3}$ for Mut Lys vs. Mut Sup; $p = 1.4 \times 10^{-4}$ for WT Sup vs. Mut Sup; t-test, two sided).

The 14-3-3 proteins exert their function as homo- or heterodimers through binding to their target proteins, usually in a Ser/Thr phosphorylation-dependent manner [78]. We assessed the ability of 14-3-3 ζ p.L220Ffs*18 to bind tyrosine hydroxylase (TH), the rate limiting enzyme in the synthesis of dopamine and one of its canonical target proteins. WT 14-3-3 ζ has previously been reported to bind with nM affinity to Ser19-phosphorylated TH (THpSer19) (THpSer19) [58,79]. By using surface plasmon resonance, we found that the mutant protein expressed in prokaryotes (*E. coli*) lost its ability to bind the THpSer19 compared to the WT protein (Figure 2A,B).

We further assessed the capacity of both 14-3-3 ζ _WT and 14-3-3 ζ _mut proteins to interact with 14-3-3 ϵ and form heterodimers, all expressed in a human cell line (HEK 293T cells). Through BRET assays we were able to detect the interaction between WT 14-3-3 ζ and 14-3-3 ϵ , as shown by a positive and saturable BRET signal (Figure 2C). Our results showed that the mutant 14-3-3 ζ lost its capacity to form heterodimers with 14-3-3 ϵ , since the linear relationship with the acceptor/donor ratio (YFP/Rluc) suggested lack of interaction (Figure 2C). We used DRD1, not known to be a 14-3-3 target, fused either to luciferase or to yellow fluorescent protein (DRD1-Rluc and DRD1-YFP) as negative controls. No interactions were identified, obtaining linear non-specific BRET signals, confirming the specificity of the WT 14-3-3 ζ and 14-3-3 ϵ interaction (Figure S3). Remarkably, when expressed in HEK293T cells for BRET assays, the 14-3-3 ζ mutant protein only showed a detectable signal when transfecting at least 10 times the amount of plasmid required for the WT (Table S1), indicating likely protein degradation, in line with the decreased solubility observed in *E. coli*.

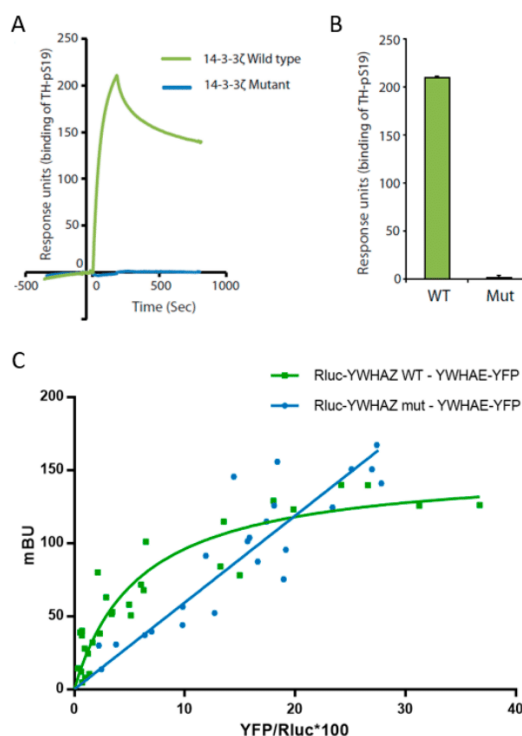


Figure 2. Characterization of mutant 14-3-3 ζ binding capacity. (A) Binding of Ser19 phosphorylated human tyrosine hydroxylase (25 nM) to immobilized wild-type GST-14-3-3 ζ (green) or mutant GST-14-3-3 ζ (blue) using surface plasmon resonance (Biacore 3000, see Materials and Methods for details). The proteins were expressed in the BL21 Codon Plus *E. coli* strain and purified prior to the analysis. (B) The binding response for Ser19 phosphorylated tyrosine hydroxylase (25 nM) at the end of the injection was compared between 14-3-3 ζ WT or Mut and Ser19 phosphorylated tyrosine hydroxylase (TH-pS19) (25 nM) repeated by several immobilizations and injections. Data presented as means and error bars denote the standard deviation ($n = 3$, $p = 1.61 \times 10^{-8}$ WT vs. Mut, two-sided t-test). (C) Characterization of wild-type 14-3-3 ζ (YWHAZ WT) and mutant 14-3-3 ζ (YWHAZ mut) interaction with 14-3-3 ϵ (YWHAE) using a bioluminescence resonance energy transfer (BRET) assay. Rluc-YWHAZ WT co-transfection with an increasing amount of YWHAE-YFP gives a saturable positive signal (green), whereas the signal obtained co-transfecting Rluc-YWHAZ mut with an increasing amount of YWHAE-YFP fits with a linear regression (blue). mBU, BRET ratio expressed in milli-BRET units. The relative amount of BRET is given as a function of YFP/Rluc*100, where YFP corresponds to the fluorescence signal due to the increasing amount of donor and Rluc corresponds to the stable luminescence signal measured at 10 min. Values shown correspond to independent experiments ($n = 4$).

Thus, a damaging effect of the truncating mutation p.L220Ffs*18 on 14-3-3 ζ was confirmed, with loss of function of the altered protein.

3.2. Common Variants Across the 14-3-3 Gene Family in ASD and Other Psychiatric Disorders

We sought for common genetic risk variants in the gene family encoding the 14-3-3 proteins in ASD and in other psychiatric disorders.

A case-control study was first performed in our sample of 713 ASD patients and 692 controls, both with European ancestry, investigating the common genetic variability of the 14-3-3 gene family (*SFN*, *YWHAB*, *YWHAE*, *YWHAG*, *YWHAH*, *YWHAQ* and *YWHAZ*) tagged by 34 SNPs. Only the variant rs1883660, located at the 3'UTR of the *SFN* gene, showed a nominal association with ASD ($p = 0.01$), but it did not remain significant after Bonferroni correction for multiple testing (Table S3).

We further extended the analysis for contribution of common variants in the seven genes of the 14-3-3 family to large GWAS datasets of eight psychiatric disorders, using summary statistics from the PGC, BroadABC and iPSYCH GWAS datasets (Table S4). The gene-based association study showed nominal associations for four out of seven genes with several psychiatric disorders: *YWHAB* with ADHD and schizophrenia, *YWHAE* with bipolar disorder and schizophrenia, *YWHAZ* with MDD and schizophrenia, and *SFN* with anxiety (Table 1).

Table 1. Gene-based association analysis of each of the seven 14-3-3 genes across several psychiatric disorders.

Gene Symbol	Entrez ID	ADHD	ASB	Anxiety	ASD	BD	MDD	OCD	SCZ	Cross-Disorder
<i>YWHAB</i>	7529	0.024	0.923	0.098	0.711	0.629	0.777	0.170	0.001	0.050
<i>YWHAE</i>	7531	0.086	0.933	0.190	0.468	0.006	0.154	0.262	<u>1.35×10^{-6}</u>	<u>1.01×10^{-5}</u>
<i>YWHAG</i>	7532	0.081	0.054	0.111	0.929	0.065	0.854	0.454	0.119	0.169
<i>YWHAH</i>	7533	0.153	0.520	0.218	0.741	0.104	0.658	0.243	0.291	0.410
<i>YWHAQ</i>	10971	0.801	0.438	0.927	0.835	0.637	0.441	0.054	0.300	0.900
<i>YWHAZ</i>	7534	0.262	0.415	0.311	0.125	0.168	0.029	0.771	0.001	0.095
<i>SFN</i>	2810	0.834	0.767	0.045	0.272	0.648	0.849	0.375	0.979	0.966

p-values were calculated using MAGMA (v1.06) software. Nominal associations are highlighted in bold. Underlined values survived Bonferroni correction for multiple testing, $p = 7.9 \times 10^{-4}$ (7 genes and 9 phenotypes). ADHD: attention-deficit hyperactivity disorder; ASB: antisocial behaviour; ASD: autism spectrum disorder; BD: bipolar disorder; MDD: major depression disorder; OCD: obsessive-compulsive disorder; SCZ: schizophrenia.

However, the association found between *YWHAE* and schizophrenia ($p = 1.35 \times 10^{-6}$; 33,640 cases and 43,456 controls) was the only surviving Bonferroni correction for 7 genes and 9 phenotypes. The gene-based results from the cross-disorder meta-analysis combining data across eight psychiatric disorders also showed a significant association with *YWHAE* ($p = 1.01 \times 10^{-5}$). We also performed a combined gene-set association analysis that indicates a nominal association of the whole 14-3-3 gene family with schizophrenia ($p = 0.018$) (Table S6).

3.3. Rare Variants in the 14-3-3 Gene Family in ASD and Schizophrenia

We investigated the role of rare variants in the 14-3-3 gene family in ASD and schizophrenia. First, we used next-generation sequencing to search for rare variants in the seven 14-3-3 genes in 288 ASD patients from our European collection, identifying nine rare variants (Table S7). All variants were validated by Sanger sequencing and the parental origin was assessed when possible (Table S7). Two detected variants were predicted to be deleterious and were found in the *SFN* gene in the same patient (p.E75del and p.T165S) together with a third variant predicted as benign (p.S149L). All three variants were present on the same chromosome and transmitted from the mother to an affected ASD proband (Figure S4A,B). We subsequently assessed the functional effect of 14-3-3 σ carrying these three changes using BRET assays, observing that the mutant protein was still able to interact with both the WT and the mutated 14-3-3 σ protein, as shown by positive saturable signals with similar BRET₅₀ and BRET_{max} (Figure S4C). The specificity of this interaction was confirmed by negative controls, obtaining linear non-specific BRET signals (Figure S4D). These results suggest that these rare variants do not impact dimerization.

We also explored the impact of ultra-rare variants (URV) across the 14-3-3 gene family in an extended sample of ASD patients, as well as in a sample of schizophrenia patients. For this objective we used publicly available sequencing data, which comprised 1924 ASD probands, 6135 schizophrenia patients, and 9090 control individuals. The high degree of evolutionary conservation of the seven 14-3-3 genes [80,81] and the limited sample size, resulted in relatively limited numbers of genetic variants. This prompted us to combine the data for the whole gene family. Interestingly, a significant burden of URVs was observed for ASD in the 14-3-3 family when compared to controls (7 in 1924 ASD cases vs. 11 in 9090 controls, $p = 0.017$), driven by a splice site variant in *YWHAE* (rs756213490), which was found in four unrelated ASD probands and not in controls (Table 2).

Table 2. List of ultra-rare variants (URVs) across the 14-3-3 gene family found in large public datasets.

Chromosome	Ref/Alt	Gene	Data_Set (Controls/Cases)	gnomAD NFE AF	Impact	Amino Acids	SIFT	PolyPhen-2	CADD	Existing Variation
17:1264594	T/A	YWHAE	ARRA (0/4)	2.19×10^{-5}	splice_acceptor_variant	canonical splice site	-	-	34	rs756213490
8:101961051	C/A	YWHAZ	German_ASD (0/1)		missense_variant	A/S	deleterious low_confidence (0.03)	possibly_damaging (0.799)	26.9	rs774415799
1:27189925	GGA/-	SFN	Spanish_ASD (0/1)	2.65×10^{-5}	inframe_deletion	SE/S	-	-	21.5	rs773116730
8:101936203	A/AT	YWHAZ	Spanish_ASD (0/1)		frameshift_variant	frameshift	-	-	-	-
2:9731521	G/A	YWHAQ	Swedish_SCZ (0/1)	8.81×10^{-6}	stop_gained	Q/*	-	-	42	rs769768341
8:101936511	G/C	YWHAZ	Swedish_SCZ (0/1)	8.84×10^{-6}	stop_gained	S/*	-	-	40	rs754522887
22:32352395	CAAGGTGTTT TACCTGA/C	YWHAH	Swedish_SCZ (0/1)	8.80×10^{-6}	frameshift_variant	KVFFYLK/X	-	-	35	rs759467778
1:27189840	T/C	SFN	MRGB (1/0)	2.64×10^{-5}	missense_variant	V/A	deleterious low_confidence (0.01)	probably_damaging (0.954)	28.6	rs77608477
1:27190047	A/C	SFN	MRGB (1/0)		missense_variant	E/A	deleterious low_confidence (0)	probably_damaging (0.929)	26.5	-
1:27190145	C/T	SFN	MRGB (1/0)		missense_variant	R/W	deleterious low_confidence (0.03)	possibly_damaging (0.776)	33	-
2:9725474	G/GTGTAGG TTGT	YWHAQ	MRGB (1/0)		frameshift_variant	frameshift	-	-	-	-
8:101961101	A/C	YWHAZ	MRGB (1/0)		missense_variant	L/R	deleterious low_confidence (0.01)	probably_damaging (0.909)	27.4	-
22:32352337	T/C	YWHAH	MRGB (1/0)	8.79×10^{-6}	missense_variant	V/A	deleterious low_confidence (0)	probably_damaging (0.999)	27.9	rs1196036662
22:32352724	A/G	YWHAH	MRGB (2/0)		missense_variant	N/S	deleterious low_confidence (0)	probably_damaging (0.991)	26.5	-
1:27189780	T/G	SFN	MRGB (1/0)	8.80×10^{-6}	missense_variant	M/R	deleterious low_confidence (0.01)	possibly_damaging (0.631)	27.7	rs747687239
1:27190388	CTG/C	SFN	Swedish_control (1/0)	8.81×10^{-6}	frameshift_variant	frameshift	-	-	35	rs774524068
17:1264594	T/A	YWHAE	Swedish_control (1/0)	2.19×10^{-5}	splice_acceptor_variant	canonical splice site	-	-	34	rs756213490

ARRA_c1, ASD cases from ARRA c1 data set (dbGAP accession: phs000298.v3.p2); German_ASD, ASD samples from Germany; MGRB, Medical Genome Reference Bank; Spanish_ASD, ASD samples from Spain; Swedish_SCZ, Sweden-Schizophrenia population-based Case-Control (dbGAP accession: phs000473.v2.p2); All the URVs variants are selected to be rare (MAF < 0.0001 in Non-Finnish European population in gnomAD), <https://gnomad.broadinstitute.org/> and predicted to be pathogenic both in SIFT and, and CADD > 20.

No significant burden was identified for schizophrenia (3 in 6135 cases vs. 11 in 9090 controls, $p = 0.71$). It is noteworthy that truncating variants in *YWHAZ* are reported here only in patients (i.e., the frameshift c.659_660insT found in the ASD family functionally investigated here, and a stop mutation rs754522887 found in a Swedish schizophrenia patient), and they were not observed in the control group. Furthermore, we also explored large available datasets of ASD and schizophrenia from the Broad Institute for enrichment of rare variants, and observed that when the 14-3-3 genes were considered individually, only *YWHAZ* reached significance for a higher number of SNVs in schizophrenia (meta-analysis $p = 0.017$).

4.4. Altered Expression of the 14-3-3 Genes in ASD and Schizophrenia

Finally, we explored potential alterations of expression of the 14-3-3 genes in post-mortem brains of ASD and schizophrenia patients using transcriptomic datasets or literature reports. From the 19 studies with available information on gene expression in the brain (12 in ASD patients and 27 in schizophrenia patients), we found 11 studies reporting significant altered expression in 14-3-3 genes ($p < 0.05$) in ASD or schizophrenia. Seven of them reported altered 14-3-3 genes expression with a false discovery rate (FDR) < 0.1 .

We found altered expression of six of the seven 14-3-3 genes in at least one of the two phenotypes compared to control subjects ($p < 0.05$, FDR < 0.1) (Table 3). All six genes, except for *SFN*, showed decreased expression in different brain areas. Interestingly, five of these genes showed alterations in expression in both ASD and schizophrenia: *YWHAB*, *YWHAE*, *YWHAH* and *YWHAZ* showed decreased expression whether *SFN* showed increased expression in both disorders. In the case of *YWHAQ*, we found decreased expression only in ASD patients compared to controls (Table 3).

Table 3. Altered expression of the 14-3-3 genes in individuals with schizophrenia or autism spectrum disorder.

Gene Symbol	Disorder	FC &	p-Value	FDR	Probe	Tissue	Study (PMID) or GEO ID	Sample, Cases vs. Controls
<i>YWHAB</i>	SCZ	-1.06	2.51×10^{-3}	0.10	8062880	cerebellum	GSE35978	44 SCZ vs. 50 control
<i>YWHAB</i>	SCZ	-1.41	0.001	0.02	217717_s_at *	hippocampus	GSE53987	15 SCZ vs. 18 control
<i>YWHAB</i>	ASD	N/A	N/A	1.43×10^{-3}	N/A	cortex (BA19, BA10, BA44)	Gupta et al., 2014 (25494366)	32 ASD vs. 40 control
<i>YWHAE</i>	ASD	-1.34	0.003	0.06	ILMN_1807535	cerebellum	GSE38322	15 ASD vs. 12 control
<i>YWHAE</i>	SCZ	-1.52	1.93×10^{-4}	0.01	210317_s_at	hippocampus	GSE53987	15 SCZ vs. 18 control
<i>YWHAE</i>	SCZ	-1.10	3.31×10^{-4}	0.04	11753092_s_at *	DLPFC	GSE87610	65 SCZ vs. 72 control
<i>YWHAH</i>	SCZ	-1.68	6.69×10^{-5}	0.01	201020_at	hippocampus	GSE53987	15 SCZ vs. 18 control
<i>YWHAH</i>	ASD	N/A	< 0.01	< 0.078	N/A	DLPFC	Liu et al., 2016 (27685936)	34 ASD vs. 40 control
<i>YWHAQ</i>	SCZ	-1.02	0.004	0.03	N/A	frontal and temporal cortex	Gandal et al., 2018 (30545856)	560 SCZ vs. 936 control
<i>YWHAQ</i>	SCZ	-1.34	2.34×10^{-5}	0.01	200693_at *	hippocampus	GSE53987	15 SCZ vs. 18 control
<i>YWHAZ</i>	ASD	-1.61	3.01×10^{-5}	0.01	ILMN_1669286	cerebellum	GSE38322	14 ASD vs. 12 control
<i>YWHAZ</i>	SCZ	-1.87	3.07×10^{-5}	0.01	200641_s_at *	hippocampus	GSE53987	15 SCZ vs. 18 control
<i>SFN</i>	ASD	2.18	0.001	0.03	N/A	frontal and temporal cortex	Gandal et al., 2018 (30545856)	51 ASD vs. 936 control
<i>SFN</i>	SCZ	1.34	7.69×10^{-5}	1.33×10^{-3}	N/A	frontal and temporal cortex	Gandal et al., 2018 (30545856)	559 SCZ vs. 936 control
<i>SFN</i>	SCZ	1.53	0.010	0.08	33323_r_at *	hippocampus	GSE53987	15 SCZ vs. 18 control

ASD, autism spectrum disorder; SCZ, schizophrenia; FDR, false discovery rate; & FC, fold change was calculated when log2FC was provided in the study; N/A, data not available; * Genes showing significant differential expression in independent probe sets targeting different transcripts/exons of the same gene, data shown corresponding to the probe with the highest fold-change.

4. Discussion

The 14-3-3 gene family encodes seven proteins that act as effectors of signaling-regulated proteins. They are highly expressed in the brain during development [82], and are involved in several neuronal processes, such as differentiation, migration, synaptogenesis, and axon guidance [24,27,31], as well as metabolic regulation [83]. Their important role in neuronal functions makes them plausible candidate genes for ASD and other psychiatric disorders. In a previous WES study we identified a truncating mutation in *YWHAZ* (encoding 14-3-3ζ), one of the members of the 14-3-3 family, present in two brothers with ASD [22]. Both individuals were diagnosed also with ADHD. One proband had mild intellectual disability (IQ of 68) and presented severe conduct disorder, aggressive behaviour and anxiety. The other proband had a normal IQ (105) and presented fibromyalgia and sleep disorder. Furthermore, in this family the mother of the ASD sib-pair, carrier of the *YWHAZ* mutation, presented with depression and other conditions, including phobia, fibromyalgia, hypothyroidism, asthma and obesity. The maternal grandmother of the ASD sib-pair was diagnosed with schizophrenia, and a maternal uncle of the two sibs had ADHD. Unfortunately, the potential segregation of this truncating variant in *YWHAZ* in the broader family could not be tested, as DNA collection was not possible. Interestingly, genes involved in dopamine neurotransmission have been reported to be associated with ADHD and ASD [84–86], and fibromyalgia and chronic pain have been related to decreased dopaminergic activity [87,88]. The 14-3-3 proteins act as regulators of dopamine synthesis by binding and stabilizing tyrosine hydroxylase [89]. Our study showed that the mutated protein 14-3-3ζ (p.L220Ffs*18) was not able to bind the phosphorylated tyrosine hydroxylase, a well-established molecular partner, leading likely to altered dopamine synthesis. Moreover, the mutated protein presented a decreased solubility when expressed in *E. coli* and it was not able to form heterodimers with 14-3-3ε when expressed in a human cell line. Our results are interesting also from the perspective of understanding the structural basis of 14-3-3 protein functions, as they suggest that the far C-terminal region may be involved in dimerization and phospho-target interaction, although further experiments would be needed to confirm it.

Previous research indicates that 14-3-3ζ deficient mice (*Ywhaz* knock-out) display cognitive and behavioural deficits possibly related to the dopamine system, altered hippocampal development, defective migration of pyramidal and granular neurons [33,90]. Also, the *Ywhaz* and *Ywhae* double knock-out mice show impaired neurogenesis, neuronal proliferation, differentiation and migration, and present severe seizures, indicating that both 14-3-3ζ and 14-3-3ε play a critical role during cortical development [31]. Indeed, 14-3-3 inhibition in certain brain regions in mice leads to impaired learning, working memory and long-term synaptic plasticity, symptoms that are associated with schizophrenia-like behaviours [91]. Altogether, these findings support an essential role for the *YWHAZ* gene in brain function and development, and together with our current report, highlight its contribution to neurodevelopmental disorders.

Given the functional evidence of this truncating mutation in *YWHAZ*, and several studies that consistently relate the 14-3-3 gene family to behavioural deficits in animal models, we further investigated the possible contribution of common and rare variants in this gene family to psychiatric disorders. The association study in our ASD sample failed to identify SNPs associated with the disorder, although the sample size was limited considering the small effect sizes typical of common variants in psychiatric disorders [13]. However, a previous study with similar sample size of adult ADHD patients identified a significant epistatic effect between *YWHAE* and two other members of this gene family, *YWHAZ* and *YWHAQ* [39]. When extending our analyses into larger samples of the PGC, BroadABC and iPSYCH, we found that common variants were gene-based associated with several psychiatric phenotypes in four of the seven 14-3-3 genes (*SFN*, *YWHAB*, *YWHAE* and *YWHAZ*), although only the association between *YWHAE* and schizophrenia survived correction for multiple testing. In line with this result, several previous studies found an association of genetic variants in *YWHAE* with schizophrenia [34,55].

The possible implication of 14-3-3 members in schizophrenia was also suggested in animal studies, which showed that 14-3-3ε deficient mice present alterations in hippocampal and cortical structures

due to defects in neurogenesis and neuronal migration [92,93]. In addition, 14-3-3 ϵ deficient mice exhibit behavioural phenotypes, such as increased motor activity and decreased working memory and are used as schizophrenia-related models [34,55]. Furthermore, 14-3-3-mediated signaling seems to be strongly affected in schizophrenia patients [47,48], supporting the association of this protein family with the disorder. Interestingly, a polymorphism (rs28365859) in *YWHAE* associated with schizophrenia [34] correlates with differences in volumes of different brain regions in patients [94,95]. Taken together, these data support the contribution of *YWHAE* to schizophrenia and general brain development.

We also investigated the impact of rare variants in a small European sample of ASD sequencing all seven 14-3-3 genes. We identified 9 rare variants in 3 genes (*YWHAE*, *YWHAB* and *SFN*), including two rare variants in *SFN* that were predicted to be pathogenic (p.E75del and p.T165S) and were present on the same chromosome of the same patient. However, the functional characterization of the effect of these latter variants showed no effect on dimerization, although they may have other functional consequences.

We then expanded the analysis of rare variants to larger samples using available sequencing datasets of ASD and schizophrenia, which pinpointed a plausible impact for rare variants in ASD when we combined data of all 14-3-3 genes. Interestingly, the effect of this association was driven by URVs from *YWHAE* and *YWHAZ*. Indeed, several exome sequencing studies in ASD patients found rare variants in *YWHAZ* and *YWHAG* [19,96,97]. We also explored sequencing data from the ASC and SCHEMA consortia for each of the 14-3-3 genes individually and an enrichment was observed only for SNVs in *YWHAZ* in the schizophrenia dataset.

Interestingly, both the expression and the splicing of several 14-3-3 genes are regulated by RNA-binding proteins that are relevant to ASD and other psychiatric disorders. In particular, RBFOX1 (RNA binding fox-1 homolog 1) regulates *YWHAE*, *YWHAG*, *YWHAQ* and *YWHAZ*, whereas FMRP (Fragile X mental retardation protein) binds *YWHAG* [98,99]. Interestingly, *RBFOX1* has been reported in several ASD genetic studies [97,100–103], as well as in studies of aggressive behaviour and several psychiatric disorders [74,104–106]. Also, several targets of FMRP have been suggested to play a role in ASD [98,107,108].

Thus, our results indicate that combined rare variants in 14-3-3 genes may contribute to ASD, and that common variants from 14-3-3 family members significantly associate with schizophrenia. Also, when 14-3-3 genes are considered individually, rare variants in the *YWHAZ* gene are shown to contribute to schizophrenia.

Finally, we have systematically gathered and analysed previous transcriptomic data that suggest significant alteration of the expression of 14-3-3 genes in ASD and schizophrenia in several brain regions. Altered levels of the 14-3-3 family have previously been reported in ASD and in schizophrenia patients: 14-3-3 protein levels are diminished in platelets and pineal glands in ASD patients [29,44] and 14-3-3 ζ levels are reduced in post-mortem brains of schizophrenia patients [109]. Moreover, an increased expression of *SFN* and decreased expression of *YWHAB*, *YWHAE*, *YWHAG* and *YWHAQ* mRNA have been reported in leukocytes of schizophrenia patients [45]. Our analysis of available transcriptomic data is in line with these results, as we found an altered brain expression of six of the seven 14-3-3 genes in ASD and schizophrenia patients: a robust increased expression of *SFN* in both ASD and schizophrenia (reflecting the top 9th and 36th differentially expressed gene in the transcriptome, respectively [110]); and a decreased expression of the other five 14-3-3 genes).

5. Conclusions

Our work combines functional studies, association studies, sequencing of a European sample, and interrogation of available genetic datasets that implicate the 14-3-3 gene family in ASD and schizophrenia, suggesting shared genetics between these two disorders.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2077-0383/9/6/1851/s1>: Supplementary file containing: Additional methodological procedures; Table S1. Acceptor and donor combinations of plasmids for BRET experiments, Table S2. European ASD samples genotyped in the case-control association study and used in the mutation screening. After quality control procedures the final genotyped sample consisted of 713 ASD cases and 692 controls, Table S3. Results from the ASD case-control association study (713 ASD cases and 692 controls) with tagSNPs across the 14-3-3 gene family in the overall sample under the additive model, Table S4. Description of the summary statistics of publicly available GWAS data of eight psychiatric disorders and the corresponding cross-disorder dataset used for gene-based and gene-set analyses, Table S5. Experimental design of targeted next-generation sequencing: The coding regions of 14-3-3 genes was covered by 57 amplicons, Table S6. Gene-set association results of the 14-3-3 family set of genes on eight different psychiatric phenotypes and in the cross-disorder meta-analysis dataset, Table S7. Rare variants identified in the seven 14-3-3 family genes in 287 288 European ASD patients, Figure S1. Linkage disequilibrium values among the 37 tagSNPs analyzed in this study, calculated from the whole sample (1441 individuals with European ancestry) with the Haploview software. D' values between all the possible SNP pairs are shown, Figure S2. The 57 amplicons used in the mutational screening are depicted in green and cover the coding regions of the 14-3-3 genes. For each of the seven genes we show the amplicon ID name, the number of overlapping amplicons per exon and the genomic region, Figure S3. Negative controls in the BRET experiments of the interaction of YWHAZ WT or YWHAZ mutant (mut) with YWHAZ using D(1A) dopamine receptor as a donor (DRD1-Rluc) or an acceptor (DRD1-YFP) and adjusted to a linear regression, Figure S4. Characterization of three rare inherited *SFN* variants identified in an ASD patient.

Author Contributions: Conceptualization: B.T., E.A.-G., N.F.-C., C.T., B.C.; Investigation, formal analysis and validation: B.T., E.A.-G., N.F.-C., E.R.-F., S.G., R.K., E.M., L.P.-C., C.T.; Resources and patients: J.M.F., J.H., V.C., A.H., I.R., J.K.B., N.R., B.F., A.G.C., C.F., A.R.; B.C.; Writing—Original Draft Preparation: E.A.-G., N.F.-C.; Writing—Review and Editing: J.M.F., R.K., C.T., B.C.; Supervision: N.F.-C., R.K., C.T., B.C.; Project Administration, B.C.; Funding Acquisition: C.T. and B.C. All authors have read and agreed to the published version of the manuscript.

Funding: Financial support was received from Fundació La Marató de TV3 (092010), Fundación Alicia Koplowitz, AGAUR (2017SGR738, 2017SGR1497), the Spanish Ministerio de Economía y Competitividad with FEDER funds (SAF2015-68341-R, SAF2017-87629-R, RTI2018-100968-B-I00) and the Australian National Medical and Health Research Council (NHMRC) Project Grant 1063960 and 1066177, and Program Grant 1037196. The research leading to these results also received funding from the European Union H2020 Program (H2020/2014-2020) under grant agreements 667302 (CoCA) and 643051 (MiND). B. Torricó was supported by AGAUR (Generalitat de Catalunya), E. Antón-Galindo by the Ministerio de Economía y Competitividad (Spanish Government), N. Fernández-Castillo by the Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER) and L. Pineda-Cirera by the Spanish Ministerio de Educación, Cultura y Deporte (FPU15/03867). C. Toma was supported by the European Union (Marie Curie, PIEF-GA-2009-254930) and currently is a recipient of a 'Ramón y Cajal' fellowship (RyC2018-024106-I) from the Spanish MINECO. J.M. Fullerton was supported by the Janette Mary O'Neil Research Fellowship.

Acknowledgments: The authors are grateful to all patients and their families for their participation and thank all clinical collaborators who contributed to the diagnosis of probands. The authors acknowledge the contribution of data from dbGAP: phs000473.v2.p2 (Sweden-Schizophrenia population-based case-control) and phs000298.v3.p2 (ARRA Autism Sequencing Collaboration). Sequencing services were provided by the Centre for Research in Agricultural Genomics (CRAG), Barcelona, Spain.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*; American Psychiatric Association: Washington, DC, USA, 2013.
2. Toma, C. Genetic variation across phenotypic severity of autism. *Trends Genet.* **2020**, *36*, 228–231. [[CrossRef](#)]
3. Hodges, H.; Fealko, C.; Soares, N. Autism spectrum disorder: Definition, epidemiology, causes, and clinical evaluation. *Transl. Pediatr.* **2020**, *9*, S55–S65. [[CrossRef](#)] [[PubMed](#)]
4. Christensen, D.L.; Maenner, M.J.; Bilder, D.; Constantino, J.N.; Daniels, J.; Durkin, M.S.; Fitzgerald, R.T.; Kurzius-Spencer, M.; Pettygrove, S.D.; Robinson, C.; et al. Prevalence and characteristics of Autism spectrum disorder among children aged 4 years—early autism and developmental disabilities monitoring network, seven sites, United States, 2010, 2012, and 2014. *MMWR Surveill. Summ.* **2019**, *68*, 1–19. [[CrossRef](#)]
5. Ronald, A.; Hoekstra, R.A. Autism spectrum disorders and autistic traits: A decade of new twin studies. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2011**, *156*, 255–274. [[CrossRef](#)] [[PubMed](#)]
6. Tick, B.; Bolton, P.; Happé, F.; Rutter, M.; Rijdsdijk, F. Heritability of autism spectrum disorders: A meta-analysis of twin studies. *J. Child Psychol. Psychiatry Allied Discip.* **2016**, *57*, 585–595. [[CrossRef](#)] [[PubMed](#)]

7. Sandin, S.; Lichtenstein, P.; Kuja-Halkola, R.; Hultman, C.; Larsson, H.; Reichenberg, A. The heritability of Autism spectrum disorder analysis method B. *JAMA* **2017**, *318*, 1182–1184. [[CrossRef](#)]
8. Bourgeron, T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat. Rev. Neurosci.* **2015**, *16*, 551–563. [[CrossRef](#)]
9. De Rubeis, S.; Buxbaum, J.D. Recent advances in the genetics of Autism spectrum disorder. *Curr. Neurol. Neurosci. Rep.* **2015**, *15*, 36. [[CrossRef](#)]
10. Wang, K.; Zhang, H.; Ma, D.; Bucan, M.; Glessner, J.T.; Abrahams, B.S.; Salyakina, D.; Imielinski, M.; Bradfield, J.P.; Sleiman, P.M.A.; et al. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **2009**, *459*, 528–533. [[CrossRef](#)]
11. Anney, R.; Klei, L.; Pinto, D.; Regan, R.; Conroy, J.; Magalhaes, T.R.; Correia, C.; Abrahams, B.S.; Sykes, N.; Pagnamenta, A.T.; et al. A genome-wide scan for common alleles affecting risk for autism. *Hum. Mol. Genet.* **2010**, *19*, 4072–4082. [[CrossRef](#)]
12. Weiss, L.A.; Arking, D.E.; Daly, M.J.; Chakravarti, A.; Brune, C.W.; West, K.; O'Connor, A.; Hilton, G.; Tomlinson, R.L.; West, A.B.; et al. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* **2009**, *461*, 802–808. [[CrossRef](#)] [[PubMed](#)]
13. Torricco, B.; Chiochetti, A.G.; Bacchelli, E.; Trabetti, E.; Hervás, A.; Franke, B.; Buitelaar, J.K.; Rommelse, N.; Yousaf, A.; Duketis, E.; et al. Lack of replication of previous autism spectrum disorder GWAS hits in European populations. *Autism Res.* **2017**, *10*, 202–211. [[CrossRef](#)] [[PubMed](#)]
14. Grove, J.; Ripke, S.; Als, T.D.; Mattheisen, M.; Walters, R.K.; Won, H.; Pallesen, J.; Agerbo, E.; Andreassen, O.A.; Anney, R.; et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **2019**, *51*, 431–444. [[CrossRef](#)]
15. Geisheker, M.R.; Heymann, G.; Wang, T.; Coe, B.P.; Turner, T.N.; Stessman, H.A.F.; Hoekzema, K.; Kvarnung, M.; Shaw, M.; Friend, K.; et al. Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nat. Neurosci.* **2017**, *20*, 1043–1051. [[CrossRef](#)] [[PubMed](#)]
16. Iossifov, I.; O'Roak, B.J.; Sanders, S.J.; Ronemus, M.; Krumm, N.; Levy, D.; Stessman, H.A.; Witherspoon, K.T.; Vives, L.; Patterson, K.E.; et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **2014**, *515*, 216–221. [[CrossRef](#)]
17. Roak, O.B.J.; Vives, L.; Girirajan, S.; Karakoc, E.; Krumm, N.; Coe, B.P.; Levy, R.; Ko, A.; Lee, C.; Smith, J.D.; et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **2012**, *485*, 246–250. [[CrossRef](#)]
18. Torricco, B.; Shaw, A.D.; Mosca, R.; Vivó-Luque, N.; Hervás, A.; Fernández-Castillo, N.; Aloy, P.; Bayés, M.; Fullerton, J.M.; Cormand, B.; et al. Truncating variant burden in high-functioning autism and pleiotropic effects of LRP1 across psychiatric phenotypes. *J. Psychiatry Neurosci.* **2019**, *44*, 350–359. [[CrossRef](#)]
19. Satterstrom, F.K.; Kosmicki, J.A.; Wang, J.; Breen, M.S.; De Rubeis, S.; An, J.-Y.; Peng, M.; Collins, R.; Grove, J.; Klei, L.; et al. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell* **2020**, *180*, 568–584. [[CrossRef](#)]
20. Bacchelli, E.; Loi, E.; Cameli, C.; Moi, L.; Vega Benedetti, A.; Blois, S.; Fadda, A.; Bonora, E.; Mattu, S.; Fadda, R.; et al. Analysis of a sardinian multiplex family with autism spectrum disorder points to post-synaptic density gene variants and identifies CAPG as a functionally relevant candidate gene. *J. Clin. Med.* **2019**, *8*, 212. [[CrossRef](#)]
21. Chapman, N.H.; Nato, A.Q.; Bernier, R.; Ankenman, K.; Sohi, H.; Munson, J.; Patowary, A.; Archer, M.; Blue, E.M.; Webb, S.J.; et al. Whole exome sequencing in extended families with autism spectrum disorder implicates four candidate genes. *Hum. Genet.* **2015**, *134*, 1055–1068. [[CrossRef](#)]
22. Toma, C.; Torricco, B.; Hervás, A.; Valdés-Mas, R.; Tristán-Noguero, A.; Padillo, V.; Maristany, M.; Salgado, M.; Arenas, C.; Puente, X.S.; et al. Exome sequencing in multiplex autism families suggests a major role for heterozygous truncating mutations. *Mol. Psychiatry* **2014**, *19*, 784–790. [[CrossRef](#)] [[PubMed](#)]
23. Aghazadeh, Y.; Papadopoulos, V. The role of the 14-3-3 protein family in health, disease, and drug development. *Drug Discov. Today* **2016**, *21*, 278–287. [[CrossRef](#)] [[PubMed](#)]
24. Cornell, B.; Toyooka, K. 14-3-3 proteins in brain development: Neurogenesis, neuronal migration and neuromorphogenesis. *Front. Mol. Neurosci.* **2017**, *10*, 1–17. [[CrossRef](#)] [[PubMed](#)]
25. Hermeking, H.; Benzinger, A. 14-3-3 Proteins in cell cycle regulation. *Semin. Cancer Biol.* **2006**, *16*, 183–192. [[CrossRef](#)] [[PubMed](#)]
26. Aitken, A. 14-3-3 proteins: A historic overview. *Semin. Cancer Biol.* **2006**, *16*, 162–172. [[CrossRef](#)]

27. Kaplan, A.; Kent, C.B.; Charron, F.; Fournier, A.E. Switching responses: Spatial and temporal regulators of axon guidance. *Mol. Neurobiol.* **2014**, *49*, 1077–1086. [[CrossRef](#)]
28. Toma, C.; Rossi, M.; Sousa, I.; Blasi, F.; Bacchelli, E.; Alen, R.; Vanhala, R.; Monaco, A.P.; Järvelä, I.; Maestrini, E. Is ASMT a susceptibility gene for autism spectrum disorders? A replication study in European populations. *Mol. Psychiatry* **2007**, *12*, 977–979. [[CrossRef](#)]
29. Pagan, C.; Goubran-Botros, H.; Delorme, R.; Benabou, M.; Lemièrre, N.; Murray, K.; Amsellem, F.; Callebort, J.; Chaste, P.; Jamain, S.; et al. Disruption of melatonin synthesis is associated with impaired 14-3-3 and miR-451 levels in patients with autism spectrum disorders. *Sci. Rep.* **2017**, *7*, 2096. [[CrossRef](#)]
30. Melke, J.; Goubran Botros, H.; Chaste, P.; Betancur, C.; Nygren, G.; Anckarsäter, H.; Rastam, M.; Ståhlberg, O.; Gillberg, I.C.; Delorme, R.; et al. Abnormal melatonin synthesis in autism spectrum disorders. *Mol. Psychiatry* **2008**, *13*, 90–98. [[CrossRef](#)]
31. Toyo-Oka, K.; Wachi, T.; Hunt, R.F.; Baraban, S.C.; Taya, S.; Ramshaw, H.; Kaibuchi, K.; Schwarz, Q.P.; Lopez, A.F.; Wynshaw-Boris, A. 14-3-3E and Z regulate neurogenesis and differentiation of neuronal progenitor cells in the developing brain. *J. Neurosci.* **2014**, *34*, 12168–12181. [[CrossRef](#)]
32. Xu, X.; Jaehne, E.J.; Greenberg, Z.; McCarthy, P.; Saleh, E.; Parish, C.L.; Camera, D.; Heng, J.; Haas, M.; Baune, B.T.; et al. 14-3-3ζ deficient mice in the BALB/c background display behavioural and anatomical defects associated with neurodevelopmental disorders. *Sci. Rep.* **2015**, *5*. [[CrossRef](#)] [[PubMed](#)]
33. Cheah, P.S.; Ramshaw, H.S.; Thomas, P.Q.; Toyo-Oka, K.; Xu, X.; Martin, S.; Coyle, P.; Guthridge, M.A.; Stomski, F.; Van Den Buuse, M.; et al. Neurodevelopmental and neuropsychiatric behaviour defects arise from 14-3-3ζ deficiency. *Mol. Psychiatry* **2012**, *17*, 451–466. [[CrossRef](#)] [[PubMed](#)]
34. Ikeda, M.; Hikita, T.; Taya, S.; Uraguchi-asaki, J.; Toyo-Oka, K.; Wynshaw-boris, A.; Ujike, H.; Inada, T.; Takao, K.; Miyakawa, T.T.; et al. Identification of YWHAE, a gene encoding 14-3-3epsilon, as a possible susceptibility gene for schizophrenia. *Hum. Mol. Genet.* **2008**, *17*, 3212–3222. [[CrossRef](#)] [[PubMed](#)]
35. Jia, Y.; Yu, X.; Zhang, B.; Yuan, Y.; Xu, Q.; Shen, Y.; Shen, Y. An association study between polymorphisms in three genes of 14-3-3 (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein) family and paranoid schizophrenia in northern Chinese population. *Eur. Psychiatry* **2004**, *19*, 377–379. [[CrossRef](#)]
36. Li, Z.; Chen, J.; Yu, H.; He, L.; Xu, Y.; Zhang, D.; Yi, Q.; Li, C.; Li, X.; Shen, J.; et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat. Genet.* **2017**, *49*, 1576–1583. [[CrossRef](#)] [[PubMed](#)]
37. Oldmeadow, C.; Mossman, D.; Evans, T.-J.; Holliday, E.G.; Tooney, P.A.; Cairns, M.J.; Wu, J.; Carr, V.; Attia, J.R.; Scott, R.J. Combined analysis of exon splicing and genome wide polymorphism data predict schizophrenia risk loci. *J. Psychiatr. Res.* **2014**, *52*, 44–49. [[CrossRef](#)]
38. Wong, A.H.C.; Macciardi, F.; Klempan, T.; Kawczynski, W.; Barr, C.L.; Lakatoo, S.; Wong, M.; Buckle, C.; Trakalo, J.; Boffa, E.; et al. Identification of candidate genes for psychosis in rat models, and possible association between schizophrenia and the 14-3-3η gene. *Mol. Psychiatry* **2003**, *8*, 156–166. [[CrossRef](#)] [[PubMed](#)]
39. Jacobsen, K.K.; Kleppe, R.; Johansson, S.; Zayats, T.; Haavik, J. Epistatic and gene wide effects in YWHA and aromatic amino hydroxylase genes across ADHD and other common neuropsychiatric disorders: Association with YWHAE. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2015**, *168*, 423–432. [[CrossRef](#)]
40. Grover, D.; Verma, R.; Goes, F.S.; Belmonte-Mahon, P.L.; Gershon, E.S.; McMahon, F.J.; Potash, J.B.; Steele, J.; Kassem, L.; Lopez, V.; et al. Family-based association of YWHAH in psychotic bipolar disorder. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2009**, *150*, 977–983. [[CrossRef](#)]
41. Liu, J.; Zhang, H.X.; Li, Z.Q.; Li, T.; Li, J.Y.; Wang, T.; Li, Y.; Feng, G.Y.; Shi, Y.Y.; He, L. The YWHAE gene confers risk to major depressive disorder in the male group of Chinese Han population. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* **2017**, *77*, 172–177. [[CrossRef](#)]
42. Yanagi, M.; Shirakawa, O.; Kitamura, N.; Okamura, K.; Sakurai, K.; Nishiguchi, N.; Hashimoto, T.; Nishida, H.; Ueno, Y.; Kanbe, D.; et al. Association of 14-3-3 ε gene haplotype with completed suicide in Japanese. *J. Hum. Genet.* **2005**, *50*, 210–216. [[CrossRef](#)] [[PubMed](#)]
43. Garbett, K.; Ebert, P.J.; Mitchell, A.; Lintas, C.; Manzi, B.; Mirnics, K.; Persico, A.M. Immune transcriptome alterations in the temporal cortex of subjects with autism. *Neurobiol. Dis.* **2008**, *30*, 303–311. [[CrossRef](#)] [[PubMed](#)]

44. Pagan, C.; Delorme, R.; Callebort, J.; Goubran-Botros, H.; Amsellem, F.; Drouot, X.; Boudebessé, C.; Dudal, K.L.; Ngo-Nguyen, N.; Laouamri, H.; et al. The serotonin-N-acetylserotonin-melatonin pathway as a biomarker for autism spectrum disorders. *Transl. Psychiatry* **2014**, *4*, e479. [[CrossRef](#)] [[PubMed](#)]
45. Qing, Y.; Sun, L.; Yang, C.; Jiang, J.; Yang, X.; Hu, X.; Cui, D.; Xu, Y.; He, L.; Han, D.; et al. Dysregulated 14-3-3 family in peripheral blood leukocytes of patients with Schizophrenia. *Sci. Rep.* **2016**, *6*. [[CrossRef](#)]
46. Saia-Cereda, V.M.; Cassoli, J.S.; Schmitt, A.; Falkai, P.; Nascimento, J.M.; Martins-de-Souza, D. Proteomics of the corpus callosum unravel pivotal players in the dysfunction of cell signaling, structure, and myelination in schizophrenia brains. *Eur. Arch. Psychiatry Clin. Neurosci.* **2015**, *265*, 601–612. [[CrossRef](#)]
47. Saia-Cereda, V.M.; Cassoli, J.S.; Martins-de-Souza, D.; Nascimento, J.M. Psychiatric disorders biochemical pathways unraveled by human brain proteomics. *Eur. Arch. Psychiatry Clin. Neurosci.* **2017**, *267*, 3–17. [[CrossRef](#)]
48. Schubert, K.O.; Föcking, M.; Cotter, D.R. Proteomic pathway analysis of the hippocampus in schizophrenia and bipolar affective disorder implicates 14-3-3 signaling, aryl hydrocarbon receptor signaling, and glucose metabolism: Potential roles in GABAergic interneuron pathology. *Schizophr. Res.* **2015**, *167*, 64–72. [[CrossRef](#)]
49. Bruno, D.L.; Anderlid, B.M.; Lindstrand, A.; Van Ravenswaaij-Arts, C.; Ganesamoorthy, D.; Lundin, J.; Martin, C.L.; Douglas, J.; Nowak, C.; Adam, M.P.; et al. Further molecular and clinical delineation of co-locating 17p13.3 microdeletions and microduplications that show distinctive phenotypes. *J. Med. Genet.* **2010**, *47*, 299–311. [[CrossRef](#)]
50. Capra, V.; Mirabelli-Badenier, M.; Stagnaro, M.; Rossi, A.; Tassano, E.; Gimelli, S.; Gimelli, G. Identification of a rare 17p13.3 duplication including the BHLHA9 and YWHAE genes in a family with developmental delay and behavioural problems. *BMC Med. Genet.* **2012**, *13*, 93. [[CrossRef](#)]
51. Nagamani, S.C.S.; Zhang, F.; Shchelochkov, O.A.; Bi, W.; Ou, Z.; Scaglia, F.; Probst, F.J.; Shinawi, M.; Eng, C.; Hunter, J.V.; et al. Microdeletions including YWHAE in the Miller-Dieker syndrome region on chromosome 17p13.3 result in facial dysmorphisms, growth restriction, and cognitive impairment. *J. Med. Genet.* **2009**, *46*, 825–833. [[CrossRef](#)]
52. Blazejewski, S.M.; Bennison, S.A.; Smith, T.H.; Toyo-oka, K. Neurodevelopmental genetic diseases associated with microdeletions and microduplications of chromosome 17p13.3. *Front. Genet.* **2018**, *9*, 9. [[CrossRef](#)] [[PubMed](#)]
53. Curry, C.J.; Rosenfeld, J.A.; Grant, E.; Gripp, K.W.; Anderson, C.; Aylsworth, A.S.; Saad, T.B.; Chizhikov, V.V.; Dybose, G.; Fagerberg, C.; et al. The duplication 17p13.3 phenotype: Analysis of 21 families delineates developmental, behavioral and brain abnormalities, and rare variant phenotypes. *Am. J. Med. Genet. Part A* **2013**, *161*, 1833–1852. [[CrossRef](#)] [[PubMed](#)]
54. Ramocki, M.B.; Bartnik, M.; Szafranski, P.; Kołodziejaska, K.E.; Xia, Z.; Bravo, J.; Miller, G.S.; Rodriguez, D.L.; Williams, C.A.; Bader, P.I.; et al. Recurrent distal 7q11.23 deletion including HIP1 and YWHAG identified in patients with intellectual disabilities, epilepsy, and neurobehavioral problems. *Am. J. Hum. Genet.* **2010**, *87*, 857–865. [[CrossRef](#)] [[PubMed](#)]
55. Wachi, T.; Cornell, B.; Toyo-oka, K. Complete ablation of the 14-3-3epsilon protein results in multiple defects in neuropsychiatric behaviors. *Behav. Brain Res.* **2017**, *319*, 31–36. [[CrossRef](#)]
56. Kim, D.E.; Cho, C.H.; Sim, K.M.; Kwon, O.; Hwang, E.M.; Kim, H.W.; Park, J.Y. 14-3-3T haploinsufficient mice display hyperactive and stress-sensitive behaviors. *Exp. Neurol.* **2019**, *28*, 43–53. [[CrossRef](#)]
57. Graham, K.; Zhang, J.; Qiao, H.; Wu, Y.; Zhou, Y. Region-specific inhibition of 14-3-3 proteins induces psychomotor behaviors in mice. *NPJ Schizophr.* **2019**, *5*, 1. [[CrossRef](#)]
58. Ghorbani, S.; Fossbakk, A.; Jorge-Finnigan, A.; Flydal, M.I.; Haavik, J.; Kleppe, R. Regulation of tyrosine hydroxylase is preserved across different homo- and heterodimeric 14-3-3 proteins. *Amino Acids* **2016**, *48*, 1221–1229. [[CrossRef](#)]
59. Itagaki, C.; Isobe, T.; Taoka, M.; Natsume, T.; Nomura, N.; Horigome, T.; Omata, S.; Ichinose, H.; Nagatsu, T.; Greene, L.A.; et al. Stimulus-coupled interaction of tyrosine hydroxylase with 14-3-3 proteins. *Biochemistry* **1999**, *38*, 15673–15680. [[CrossRef](#)]
60. Kleppe, R.; Rosati, S.; Jorge-Finnigan, A.; Alvira, S.; Ghorbani, S.; Haavik, J.; Valpuesta, J.M.; Heck, A.J.R.; Martinez, A. Phosphorylation dependence and stoichiometry of the complex formed by tyrosine hydroxylase and 14-3-3γ. *Mol. Cell. Proteom.* **2014**, *13*, 2017–2030. [[CrossRef](#)]

61. Moreno, E.; Vaz, S.H.; Cai, N.S.; Ferrada, C.; Quiroz, C.; Barodia, S.K.; Kabbani, N.; Canela, E.I.; McCormick, P.J.; Lluis, C.; et al. Dopamine-galanin receptor heteromers modulate cholinergic neurotransmission in the rat ventral hippocampus. *J. Neurosci.* **2011**, *31*, 7412–7423. [[CrossRef](#)]
62. Rutter, M.; LeCouteur, A.; Lord, C. *Autism Diagnostic Interview-Revised (ADI-R)*; Western Psychological Services: Torrance, CA, USA, 2003.
63. Lord, C.; Rutter, M.; DiLavore, P.; Risi, S.; Gotham, K.; Bishop, L.; RJ, L.; Guthrie, W. *Autism Diagnostic Observation Schedule*, 2nd ed.; Western Psychological Services: Torrance, CA, USA, 2012.
64. Miller, S.A.; Dykes, D.D.; Polesky, H.F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **1988**, *6*, 1215. [[CrossRef](#)] [[PubMed](#)]
65. Barrett, J.C.; Fry, B.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2005**, *21*, 263–265. [[CrossRef](#)] [[PubMed](#)]
66. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)] [[PubMed](#)]
67. Demontis, D.; Walters, R.K.; Martin, J.; Mattheisen, M.; Als, T.D.; Agerbo, E.; Baldursson, G.; Belliveau, R.; Bybjerg-Grauholm, J.; Bækvad-Hansen, M.; et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **2019**, *51*, 63–75. [[CrossRef](#)] [[PubMed](#)]
68. Tielbeek, J.J.; Johansson, A.; Polderman, T.J.C.; Rautiainen, M.R.; Jansen, P.; Taylor, M.; Tong, X.; Lu, Q.; Burt, A.S.; Tiemeier, H.; et al. Genome-wide association studies of a broad spectrum of antisocial behavior. *JAMA Psychiatry* **2017**, *74*, 1242–1250. [[CrossRef](#)] [[PubMed](#)]
69. Meier, S.M.; Tronetti, K.; Purves, K.L.; Als, T.D.; Grove, J.; Laine, M.; Pedersen, M.G.; Bybjerg-Grauholm, J.; Bækved-Hansen, M.; Sokolowska, E.; et al. Genetic variants associated with anxiety and stress-related disorders: A genome-wide association study and mouse-model study. *JAMA Psychiatry* **2019**, *76*, 924–932. [[CrossRef](#)]
70. Stahl, E.A.; Breen, G.; Forstner, A.J.; McQuillin, A.; Ripke, S.; Trubetskoy, V.; Mattheisen, M.; Wang, Y.; Coleman, J.R.I.; Gaspar, H.A.; et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **2019**, *51*, 793–803. [[CrossRef](#)]
71. Wray, N.R.; Ripke, S.; Mattheisen, M.; Trzaskowski, M.; Byrne, E.M.; Abdellaoui, A.; Adams, M.J.; Agerbo, E.; Air, T.M.; Andlauer, T.M.F.; et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **2018**, *50*, 668–681. [[CrossRef](#)]
72. Arnold, P.D.; Askland, K.D.; Barlassina, C.; Bellodi, L.; Bienvenu, O.J.; Black, D.; Bloch, M.; Brentani, H.; Burton, C.L.; Camarena, B.; et al. Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol. Psychiatry* **2018**, *23*, 1181–1188. [[CrossRef](#)]
73. Ripke, S.; Neale, B.M.; Corvin, A.; Walters, J.T.R.; Farh, K.H.; Holmans, P.A.; Lee, P.; Bulik-Sullivan, B.; Collier, D.A.; Huang, H.; et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **2014**, *511*, 421–427. [[CrossRef](#)]
74. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* **2019**, *179*, 1469–1482. [[CrossRef](#)] [[PubMed](#)]
75. De Leeuw, C.A.; Mooij, J.M.; Heskes, T.; Posthuma, D. MAGMA: Generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **2015**, *11*, 4. [[CrossRef](#)] [[PubMed](#)]
76. Kurosaki, T.; Maquat, L.E. Nonsense-mediated mRNA decay in humans at a glance. *J. Cell Sci.* **2016**, *129*, 461–467. [[CrossRef](#)] [[PubMed](#)]
77. Khajavi, M.; Inoue, K.; Lupski, J.R. Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease. *Eur. J. Hum. Genet.* **2006**, *14*, 1074–1081. [[CrossRef](#)] [[PubMed](#)]
78. Obsil, T.; Obsilova, V. Structural basis of 14-3-3 protein functions. *Semin. Cell Dev. Biol.* **2011**, *22*, 663–672. [[CrossRef](#)]
79. Halskau, Ø.J.; Ying, M.; Baumann, A.; Kleppe, R.; Rodriguez-Larrea, D.; Almås, B.; Haavik, J.; Martinez, A. Three-way interaction between 14-3-3 proteins, the N-terminal region of tyrosine hydroxylase, and negatively charged membranes. *J. Biol. Chem.* **2009**, *284*, 32758–32769. [[CrossRef](#)]
80. Uhart, M.; Bustos, D.M. Human 14-3-3 paralogs differences uncovered by cross-talk of phosphorylation and lysine acetylation. *PLoS ONE* **2013**, *8*, 1–16. [[CrossRef](#)]

81. Rosenquist, M.; Sehnke, P.; Ferl, R.J.; Sommarin, M.; Larsson, C. Evolution of the 14-3-3 protein family: Does the large number of isoforms in multicellular organisms reflect functional specificity? *J. Mol. Evol.* **2000**, *51*, 446–458. [[CrossRef](#)]
82. Berg, D.; Holzmann, C.; Riess, O. 14-3-3 Proteins in the nervous system. *Nat. Rev. Neurosci.* **2003**, *4*, 752–762. [[CrossRef](#)]
83. Kleppe, R.; Martinez, A.; Døskeland, S.O.; Haavik, J. The 14-3-3 proteins in regulation of cellular metabolism. *Semin. Cell Dev. Biol.* **2011**, *22*, 713–719. [[CrossRef](#)]
84. Cabana-Domínguez, J.; Torrico, B.; Reif, A.; Fernández-Castillo, N.; Cormand, B. Exploring the genetic contribution of the dopaminergic and serotonergic pathways in psychiatric disorders. *Transl. Psychiatry* **2019**, *9*, 242. [[CrossRef](#)]
85. Ribasés, M.; Ramos-Quiroga, J.A.; Hervás, A.; Sánchez-Mora, C.; Bosch, R.; Bielsa, A.; Gastaminza, X.; Lesch, K.P.; Reif, A.; Renner, T.J.; et al. Candidate system analysis in ADHD: Evaluation of nine genes involved in dopaminergic neurotransmission identifies association with DRD1. *World J. Biol. Psychiatry* **2012**, *13*, 281–292. [[CrossRef](#)] [[PubMed](#)]
86. Toma, C.; Hervás, A.; Balmaña, N.; Salgado, M.; Maristany, M.; Vilella, E.; Aguilera, F.; Orejuela, C.; Cuscó, I.; Gallastegui, E.; et al. Neurotransmitter systems and neurotrophic factors in autism: Association study of 37 genes suggests involvement of DDC. *World J. Biol. Psychiatry* **2013**, *14*, 516–527. [[CrossRef](#)] [[PubMed](#)]
87. Treister, R.; Pud, D.; Ebstein, R.P.; Laiba, E.; Gershon, E.; Haddad, M.; Eisenberg, E. Associations between polymorphisms in dopamine neurotransmitter pathway genes and pain response in healthy humans. *Pain* **2009**, *147*, 187–193. [[CrossRef](#)]
88. Wood, P.B.; Holman, A.J. An elephant among us: The role of dopamine in the pathophysiology of fibromyalgia. *J. Rheumatol.* **2009**, *36*, 221–224. [[CrossRef](#)]
89. Ghorbani, S.; Szigetvari, P.D.; Haavik, J.; Kleppe, R. Serine 19 phosphorylation and 14-3-3 binding regulate phosphorylation and dephosphorylation of tyrosine hydroxylase on serine 31 and serine 40. *J. Neurochem.* **2020**, *152*, 29–47. [[CrossRef](#)]
90. Ramshaw, H.; Xu, X.; Jaehne, E.J.; McCarthy, P.; Greenberg, Z.; Saleh, E.; McClure, B.; Woodcock, J.; Kabbara, S.; Wiszniak, S.; et al. Locomotor hyperactivity in 14-3-3ζ KO mice is associated with dopamine transporter dysfunction. *Transl. Psychiatry* **2013**, *3*, e327. [[CrossRef](#)]
91. Qiao, H.; Foote, M.; Graham, K.; Wu, Y.; Zhou, Y. 14-3-3 proteins are required for hippocampal long-term potentiation and associative learning and memory. *J. Neurosci.* **2014**, *34*, 4801–4808. [[CrossRef](#)]
92. Toyo-Oka, K.; Shionoya, A.; Gambello, M.J.; Cardoso, C.; Leventer, R.; Ward, H.L.; Ayala, R.; Tsai, L.H.; Dobyns, W.; Ledbetter, D.; et al. 14-3-3ε is important for neuronal migration by binding to NUDEL: A molecular explanation for Miller-Dieker syndrome. *Nat. Genet.* **2003**, *34*, 274–285. [[CrossRef](#)]
93. Cornell, B.; Wachi, T.; Zhukarev, V.; Toyo-oka, K. Regulation of neuronal morphogenesis by 14-3-3epsilon (Ywhae) via the microtubule binding protein, doublecortin. *Hum. Mol. Genet.* **2016**, *25*, 4405–4418. [[CrossRef](#)]
94. Takahashi, T.; Nakamura, Y.; Nakamura, Y.; Aleksic, B.; Takayanagi, Y.; Furuichi, A.; Kido, M.; Nakamura, M.; Sasabayashi, D.; Ikeda, M.; et al. The polymorphism of YWHAE, a gene encoding 14-3-3epsilon, and orbitofrontal sulcogyral pattern in patients with schizophrenia and healthy subjects. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* **2014**, *51*, 166–171. [[CrossRef](#)] [[PubMed](#)]
95. Kido, M.; Nakamura, Y.; Nemoto, K.; Takahashi, T.; Aleksic, B.; Furuichi, A.; Nakamura, Y.; Ikeda, M.; Noguchi, K.; Kaibuchi, K.; et al. The polymorphism of YWHAE, a gene encoding 14-3-3epsilon, and brain morphology in Schizophrenia: A voxel-based morphometric study. *PLoS ONE* **2014**, *9*, E103571. [[CrossRef](#)] [[PubMed](#)]
96. Lim, E.T.; Uddin, M.; De Rubeis, S.; Chan, Y.; Kamumbu, A.S.; Zhang, X.; D’Gama, A.M.; Kim, S.N.; Hill, R.S.; Goldberg, A.P.; et al. Rates, distribution and implications of postzygotic mosaic mutations in Autism spectrum disorder. *Nat. Neurosci.* **2017**, *20*, 1217–1224. [[CrossRef](#)]
97. De Rubeis, S.; He, X.; Goldberg, A.P.; Poultney, C.S.; Samocha, K.; Cicek, A.E.; Kou, Y.; Liu, L.; Fromer, M.; Walker, S.; et al. Synaptic, transcriptional and chromatin genes disrupted in Autism. *Nature* **2014**, *515*, 209–215. [[CrossRef](#)] [[PubMed](#)]
98. Darnell, J.C.; Van Driesche, S.J.; Zhang, C.; Hung, K.Y.S.; Mele, A.; Fraser, C.E.; Stone, E.F.; Chen, C.; Fak, J.J.; Chi, S.W.; et al. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and Autism. *Cell* **2011**, *146*, 247–261. [[CrossRef](#)]

99. Lee, J.A.; Damianov, A.; Lin, C.H.; Fontes, M.; Parikshak, N.N.; Anderson, E.S.; Geschwind, D.H.; Black, D.L.; Martin, K.C. Cytoplasmic Rbfox1 regulates the expression of synaptic and autism-related genes. *Neuron* **2016**, *89*, 113–128. [[CrossRef](#)]
100. Hamada, N.; Ito, H.; Nishijo, T.; Iwamoto, I.; Morishita, R.; Tabata, H.; Momiyama, T.; Nagata, K.I. Essential role of the nuclear isoform of RBFOX1, a candidate gene for Autism spectrum disorders, in the brain development. *Sci. Rep.* **2016**, *6*, 30805. [[CrossRef](#)]
101. Turner, T.N.; Hormozdiari, F.; Duyzend, M.H.; McClymont, S.A.; Hook, P.W.; Iossifov, I.; Raja, A.; Baker, C.; Hoekzema, K.; Stessman, H.A.; et al. Genome sequencing of Autism-affected families reveals disruption of putative noncoding regulatory DNA. *Am. J. Hum. Genet.* **2016**, *98*, 58–74. [[CrossRef](#)]
102. Voineagu, I.; Wang, X.; Johnston, P.; Lowe, J.K.; Tian, Y.; Horvath, S.; Mill, J.; Cantor, R.M.; Blencowe, B.J.; Geschwind, D.H. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **2011**, *474*, 380–386. [[CrossRef](#)]
103. Kanduri, C.; Kantojärvi, K.; Salo, P.M.; Vanhala, R.; Buck, G.; Blancher, C.; Lähdesmäki, H.; Järvelä, I. The landscape of copy number variations in Finnish families with autism spectrum disorders. *Autism Res.* **2016**, *9*, 9–16. [[CrossRef](#)]
104. Fernández-Castillo, N.; Gan, G.; van Donkelaar, M.M.J.; Vaht, M.; Weber, H.; Retz, W.; Meyer-Lindenberg, A.; Franke, B.; Harro, J.; Reif, A.; et al. RBFOX1, encoding a splicing regulator, is a candidate gene for aggressive behavior. *Eur. Neuropsychopharmacol.* **2020**, *30*, 44–55. [[CrossRef](#)] [[PubMed](#)]
105. Lal, D.; Pernhorst, K.; Klein, K.M.; Reif, P.; Tozzi, R.; Toliat, M.R.; Winterer, G.; Neubauer, B.; Nürnberg, P.; Rosenow, F.; et al. Extending the phenotypic spectrum of RBFOX1 deletions: Sporadic focal epilepsy. *Epilepsia* **2015**, *56*, e129–e133. [[CrossRef](#)] [[PubMed](#)]
106. Wen, M.; Yan, Y.; Yan, N.; Chen, X.S.; Liu, S.Y.; Feng, Z.H. Upregulation of RBFOX1 in the malformed cortex of patients with intractable epilepsy and in cultured rat neurons. *Int. J. Mol. Med.* **2015**, *35*, 597–606. [[CrossRef](#)] [[PubMed](#)]
107. Budimirovic, D.B.; Kaufmann, W.E. What can we learn about autism from studying fragile X syndrome? *Dev. Neurosci.* **2011**, *33*, 379–394. [[CrossRef](#)]
108. Lai, A.; Valdez-Sinon, A.N.; Bassell, G.J. Regulation of RNA granules by FMRP and implications for neurological diseases. *Traffic* **2020**. [[CrossRef](#)]
109. English, J.A.; Pennington, K.; Dunn, M.J.; Cotter, D.R. The neuroproteomics of Schizophrenia. *Biol. Psychiatry* **2011**, *69*, 163–172. [[CrossRef](#)]
110. Gandal, M.J.; Zhang, P.; Hadjimichael, E.; Walker, R.L.; Chen, C.; Liu, S.; Won, H.; Van Bakel, H.; Varghese, M.; Wang, Y.; et al. Transcriptome-wide isoform-level dysregulation in ASD, Schizophrenia, and bipolar disorder. *Science* **2018**, *362*. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

