The Mouse Cortex Regulome

Effects of environmental enrichment on postnatal brain development

Sergio Espeso Gil

TESI DOCTORAL UPF / 2016

DIRECTOR DE LA TESI

Dr. Stephan Ossowski

GENOMIC AND EPIGENOMIC VARIATION IN DISEASE GROUP
BIOINFORMATICS AND GENOMICS DEPARTMENT

CENTRE FOR GENOMIC REGULATION

I received a FPI doctoral fellowship from the Spanish Ministry of Economy and Competitiveness (MINECO), BES-2012-056499 during the period 2012-2016.
Also, granted with a research fellowship from the Spanish Ministry of Economy and Competitiveness (MINECO), EEBB-I-1509744 during the period August to December 2015 to perform an internship at the Dept. Psychiatry-Icahn School of Medicine at Mount Sinai, New York, USA

Acknowledgements

Hoy, hace algo más de cuatro años me embarqué en esta aventura. Jamás hubiera podido pensar que el tiempo pasase tan rápido. Pero lo cierto es que sí, llegó el momento. La prueba, la foto de mi tarjeta de entrada al PRBB. Cualquiera diría que esa era mi cara. Quizá haya sido yo mismo la cobaya de mi propia investigación. Así que aquí estoy en frente del ordenador, sin mucho tiempo y pensando en cómo voy a estructurar estos agradecimientos. Siempre he tenido la sensación que es la parte más aburrida, pero trataré de hacerla amena al curioso/a .

No estaría probablemente aquí, delante de este ordenador varias veces reparado, sin la confianza y el apoyo de los/as muchos/as que estáis leyendo estas líneas. Este trabajo ha sido duro, lo sabéis, he puesto muchísima ilusión, tesón y esfuerzo, pero sin vosotros/as hubiese sido probablemente mucho peor. Gracias

Quiero empezar agradeciendo el apoyo de mi familia, en especial el de mi madre. Creo que sin lugar a dudas, has sido la persona que más ha sufrido hacer una tesis, a parte de mi mismo claro. Tal es la sinergia que podrían también poner tu nombre en la portada (rio). Por eso gran parte de este trabajo te lo debo a ti. Gracias mamá por todo el apoyo y por todos esos tappers ricos que nunca volvieron a Madrid (rio de nuevo). No quiero olvidar a mi padre y a mis hermanos. Gracias hermanos por encenderme la llama de la ciencia. Dos físicos nacidos el mismo día de diferente año era suficiente razón para que vuestro hermano estudiase Neurociencia y tener la voluntad de estudiar vuestro cerebro (jaja). Gracias

Hay amigos de los que a pesar de la distancia, siempre estarán ahí. Gracias María, por haber siempre estado a mi lado, por tu apoyo incondicional, por ser casi mi segunda madre. También gracias a ti Andrés, mil gracias, de verdad. Gracias por haberme hecho reír a carcajadas tantas veces, y gracias por estar siempre ahí, aunque estés lejos de mí. Gracias a los dos por haberme escuchado tanto, de verdad, os llevo en el corazón. También he de extender esto a ti Javi, el señor de Lago! Mil gracias por todo el cariño que me has dado galleguiño. No me iba a olvidar de ti Sergio, gracias por estar siempre ahí. Tampoco me iba a olvidar de ti Pata, hubiese estado genial haber sido compañeros de piso en Barcelona. Desde luego hiciste honor a tu nombre rompiéndote las dos piernas cruzando un paso de cebra (jaja). Siempre voy a recordar los momentos que he podido disfrutar de tu presencia, ¡estás loco! Bueno...más o menos como yo. Contigo va el tren del Z: Mon, Elo, Pablo, Rocio, Elena, Charlie, Ester, Chiki... Gracias Mon por tus consejos de amiga (rio). Gracias Elo por todas las conversaciones de neuro que hemos tenido, la verdad es que las echo de menos. Escribiendo esto me estoy poniendo terriblemente melancólico, os echo mucho de menos.

Hace casi 10 años Nuria y yo creamos La Biothèque, una organización sin animo de lucro constituida por jóvenes científicos para impulsar la ciencia entre nuestra sociedad. Gracias por acompañarme en este camino de divulgación, gracias por brindarme tu confianza plena. Sé que ha sido y es difícil , pero en fin.. ahí estamos aportando nuestro granito de arena.

No me podía olvidar de ti Pablo, en fin...sabes que te quiero un montón y me alegro mucho de que mi segunda madre nos haya presentado aquel día en Lavapiés. Gracias por acompañarme durante todos estos años. Gracias.

Luiño mil gracias por tus consejos, por tus correcciones, por nuestras conversaciones científicas filosofando sobre la vida. Ha sido un placer conocerte, pasear contigo desvariando sobre el Universo y CRISPR (dos temas muy parecidos jaja) por las calles de Nueva York. Sin olvidar el Pro-thai!. Tenemos que volver a ese sitio algún día, por favor.

En este viaje también me han acompañado unos cuantos compañeros del Máster de Neurociencia de la UAM. En especial Marta, Ana, Helen y Alba. Tengo tan buenos recuerdos, ays, os echo de menos! A vosotras os sigo viendo a pesar de la distancia, pero me gustaría que nos pudiéramos juntar algún día todos! Con Jesús, Débora, Adolfo… No me quiero olvidar tampoco la gente del Instituto Cajal, en especial de Alejandra y Paty. Qué grandes sois!! Gracias Ale, de verdad. Gracias por tu apoyo, por tus consejos, por tus correcciones de esta tesis. También a Pamper, Simona.. molaría volver a ser compañeros! Y también de la Facultad! En especial a ti Luchy!!

Aquí en Barcelona he podido conocer gente que me ha hecho pasar momentos muy divertidos y me han cuidado mucho. Gracias Adrián , Ángel y Alessandro S. Hay sido genial conoceros. También me gustaría agradecer a mi compañero de piso Rafa! Gracias! Ha sido un placer compartir casa juntos, me has mimado mucho la verdad, gracias.

Si alguien me ha acompañado más que nadie en estos últimos meses has sido tú Roberto. Gracias por llevarme a nadar a Picornell y por sentarte a mi lado durante los largos fines de semana escribiendo esta tesis. Pero también a ti Joan por compartir nuestros momentos más sedentarios (jaja). Gracias a los dos!

Quiero agradecer a todos los compañeros de deporte de estos últimos años. Gracias Luisa, Oliver, Michael, y Carlos por las cursas y triatlones que nos hemos compartido. Gracias además por ser grandes compañeros de trabajo.

Je suis arrivé ici après un long chemin et de nombreux efforts auxquels d'importants accents francophones ont participé. Il y a très longtemps, je me suis plongé dans le monde des Neurosciences, principalement dans le domaine de la physiologie animale. L'envie d'apprendre était telle qu'elle m'a donné l'énergie de continuer à étudier ce monde riche de mystères et de surprises. Mais la seule manière d'y parvenir était de partir en Erasmus, puisque ma faculté n'offrait pas ce genre d'études. Après une éducation scolaire liée à la langue et à la culture française, l'Université Pierre et Marie Curie me semblait un excellent choix pour continuer mes études en Neurosciences. J'ai rassemblé toutes mes forces, à l'époque éparpillées, pour me présenter aux entretiens qui me permettraient de partir en Erasmus. Ce jour là, mon interlocuteur fut Vicente Mazimpaka; je le remercie de m'avoir offert cette opportunité, en changeant la seule place déjà prise pour y aller. Je garderai toujours un souvenir amusé de tes expressions comme "Tu as un français macaronique!"; probablement que je n'avais rien appris à l'école. Nicole! Qu'est-ce que tu m'as appris? Hehehe, merci à toi aussi. J'y ai connu des personnalités qui se sont révélées importantes et qui m´on accompagné tout au long de mon séjour. Merci à Diana pour tous les moments que l'on a vécus ensemble; je sais que je devrai aller à Paris, elle le comprendra à raison comme un futur très proche. Merci aussi à Jérémy, ma petite naine (haha), que je remercie pour tous les moments offerts en dansant, en riant, en partageant enfin la nuit parisienne et me faire rencontrer

des gens aussi cools que Naëlle et Jean-Baptiste. Merci à tous. Je voudrais aussi faire une mention tout à fait spéciale à ma petite famille belge: merci beaucoup à Hubert et Jérôme pour leur amour, pour prendre soin de moi, me faire découvrir la campagne belge, me traîter comme un roi; ils savent qu'une partie de mon cœur reste près d'eux. Et merci à Karl, malgré nos rencontres manquées à New York ou à Angre, alors que j´étais déjà parti, il sait que j'ai hâte de le revoir.

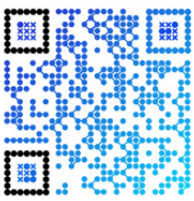
Well...almost there…I would really like to thank a lot of people at the CRG and PRBB. Good work colleagues but also friends. Especially to Mimma, how nice is to make you smile! Mimmi!!! Hehehe, It has been such a nice and great experience to share the space with you and to doctorate in fact together! Also Davide, how sweaty you get when you do GRIT! Hahaha, just joking, thanks to invite me to this kind of exhausting cross-fit experience, and also to make me smile! Such a nice guy! Thanks also to Alessandro D. it is great to have you around. Having such nice moments such as all those moments speaking secretly in French when everyone could understand us, hahaha. Thanks also to you Elias, Júlia A, Mara D, Ivica, Teresa, Estel, Vicky, Silvia P, Amaya, Álvaro, Silvina, Maria, Irene, Nuria, Maik, Ana M, Johen, Oscar F, Erika, Eva ... I probably forget people, sorry! Thanks all to your support and help!

I would like also to thank Elvira, Maruxa y Mónica from the PRBB Communications department.

Special thanks to Schahram Akbarian, for giving me the opportunity to be part of your lab during my internship. I have learnt a lot. Thanks Amanda for your patience. Thanks to the rest of the lab, especially to Prashanth, Sandhya, Leanne, Royce, Ben, Jenny… to all, thanks! You rock!

Finally, I would really like to thank all the people from Ossowski lab. Thanks Stephan for giving the space to develop my creativity. It has been great to share with you guys this last four years. Thanks Oli and Charlotte for your corrections and comments about my thesis dissertation, and to be always there available. I would like to also thank specially to Shalu to be my crazy Indian lab mate, to Fran to be my military gym trainer together with German, to Aliaksei for all the computational help, to Katja to be the best master student I have ever had (I didn't have anyone else haha), to Michael to help me with the wetlab (and corrections) but also to have such a great and funny moments at the lab! (I am a …), to Mattia for the computational help and support, and to Dani for some wetlab useful tips and the super powerful energy cookies!

Gracias a todos, de verdad, me he sentido muy querido.



Abstract

The regulome constitutes a complex system of factors that control the molecular phenotype of the cell, which is influenced by the environment. Any disturbance can trigger a set of changes involving dysfunctional regulation. The brain constantly integrates a multitude of motor, sensory and cognitive information. This engagement is particularly important in postnatal development when the brain must establish the molecular commitments needed to adapt to a changing environment. The aim of this study is to investigate how environmental factors influence the cerebral cortex regulome during postnatal development. In order to study the interaction between the regulome and the environment, we used the paradigm of environmental enrichment (EE) in which mice received constant and novel stimulation during a month. Next Generation Sequencing (NGS) –based techniques were employed to analyze the epigenome, gene accessibility, chromosomal interactions, the transcriptome and the proteome. Notably, dynamic changes in neuronal H3K79me2 coverage were observed, together with a general gain of promoter and enhancer accessibility of learning-associated genes. These changes were also supported by transcriptomic and proteomic data. We followed a flow cytometry strategy that allowed us to highlight differences in EE-induced changes in the cerebral cortex and cortical neurons. Our research reveals for the first time that EE induces changes in the regulatory mechanisms related with synaptic fine-tuning in cortical neurons during postnatal development.

Resum

El reguloma està constituït per un sistema complex de factors que controlen el fenotip molecular de la cèl·lula, que al seu torn està influenciada pel medi ambient. Qualsevol pertorbació pot desencadenar canvis que poden implicar una regulació disfuncional. El cervell integra constantment una quantitat considerable d'informació motora, sensorial i cognitiva. Aquesta integració és particularment important en el desenvolupament postnatal, en què el cervell ha d'establir els compromisos moleculars necessaris per adaptar-se a un entorn canviant. L'objectiu d'aquest estudi és investigar com els factors ambientals poden influenciar en el reguloma de l'escorça cerebral durant el desenvolupament postnatal. Per tal d'estudiar la interacció entre el reguloma i el medi ambient, s'ha utilitzat el paradigma d'enriquiment ambiental en què s'exposa els ratolins a estímuls freqüentment canviants de manera constant durant un mes. En aquest estudi s'ha emprat seqüenciació d'última generació per poder analitzar l'epigenoma, regions obertes de la cromatina, interaccions cromosòmiques, el transcriptoma i el proteoma. En particular, s'observen canvis dinàmics en la cobertura de la modificació H3K79me2 neuronal, juntament amb un augment general d'accessibilitat en regions promotores i *enhancers* associats a gens importants per a l'aprenentatge. Complementàriament, les dades de transcriptòmica i proteòmica recolzen aquests resultats. Així mateix, s'ha implementat una estratègia particular en citometria de flux que ha permès esbrinar quines són les majors diferències en els canvis induïts per l'enriquiment ambiental en l'escorça cerebral i neurones corticals. En conjunt, i per primer cop, aquests estudis apunten que l'enriquiment ambiental indueix una sèrie de canvis en els mecanismes de regulació de les neurones corticals, relacionats amb un minuciós ajust sinàptic durant el desenvolupament postnatal.

Preface

This current project represents a small part of the total datasets that I have produced during my PhD. The full project consisted to study the epigenome, chromatin accessibility and conformation, the transcriptome and the proteome of four different experimental conditions: standard maze animals, environmental enriched, epigallocatechin gallate (EGCG) treated and combined treatment (EGCG +EE) (Figure i). Each of the experimental condition was applied to wild type and TgDyrk1A, a mouse model of Down Syndrome(Altafaj et al., 2001). I have performed all the ChIP-seq libraries for the 8 different groups, consisting in 7 histone marks: H3K27ac, H3K4me3, H3K4me1, H3K79me2, H3K36me3, H3K27me3, H3K9me3 and CTCF (2 biological replicates per condition, sequenced at Centre for Genomic Regulation-CRG, Barcelona). Charlotte Hor performed the ATAC-seq libraries and the DNA extraction for bisulfite sequencing. Together with Justo González, they both performed the RNA extraction for RNA-seq libraries (sequenced at Centro Nacional de Análisis Genómico-CNAG, Barcelona). The Proteomic Facility (CRG, Barcelona) performed the protein extraction and performed the iTRAQ.

On top of that, I have performed the ChIP-seq computational analysis. I have worked with Jekaterina Kokatjuhha and Aliaksei Holik in developping the peak-independent methods (CSAW and in-house tool). I have performed the differential analysis of the ATAC-seq data (mapped by a benchmark pipeline of Stephan Ossowski), the removal of the bacth effect of the RNA-seq, the differential analysis of the RNA-seq, the microRNA analysis (mapped by Marc Friedländer) and the proteome analysis.

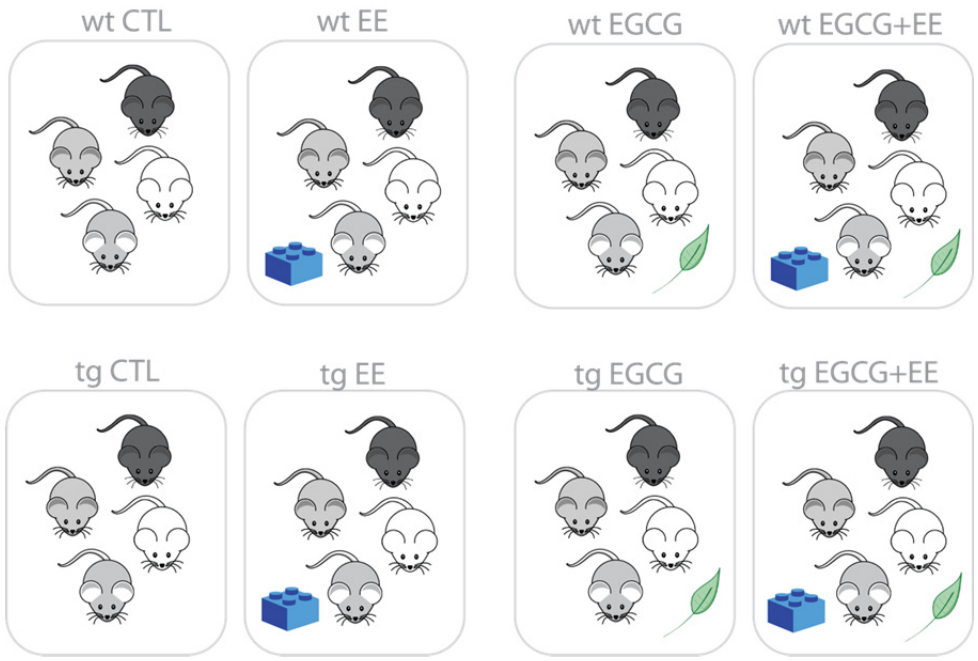


Figure i. Overview of the full project

As the amount of data was initially big enough, we decided then to focus our attention into the contrast of wild type non-enriched versus EE. Thus, complementary experiments were performed and new groups of animals were bred (3 biological

replicates for the 2nd group and 2 biological replicates for the 3rd group Tg(Thy-1YFP) mice). I took care of them with the assistance of Júlia Albaigès, changing every 3 days the objects to maintain the novelty. I performed the dissections and some of the perfusions. The second group of animals was bred to validate part of the data: the RNA-seq, microRNAs and the proteome. But also to perform FAC sorting of total neurons (NeuN+, Rbfox3) followed by ChIP-seq (H3K79me2), ATAC-seq, 3C and in situ HiC. I performed all validations, I optimized the FACS-ChIPseq, FACS-ATACseq, FACS-3C and FACS-*in situ*-HiC protocols, I performed the experiments and I have analyzed the data. The computational analysis of the *in situ*-HiC libraries has been done together with Aliaksei Holik (data here not shown).

The aim of the third group, corresponding to Tg(Thy1-YFP) mice was to perform ChIP-seq (H3K79me2) and ATAC-seq in pyramidal neurons. Michael Maher assisted me to perform the FAC-sorting. I also performed the computational analysis of these libraries.

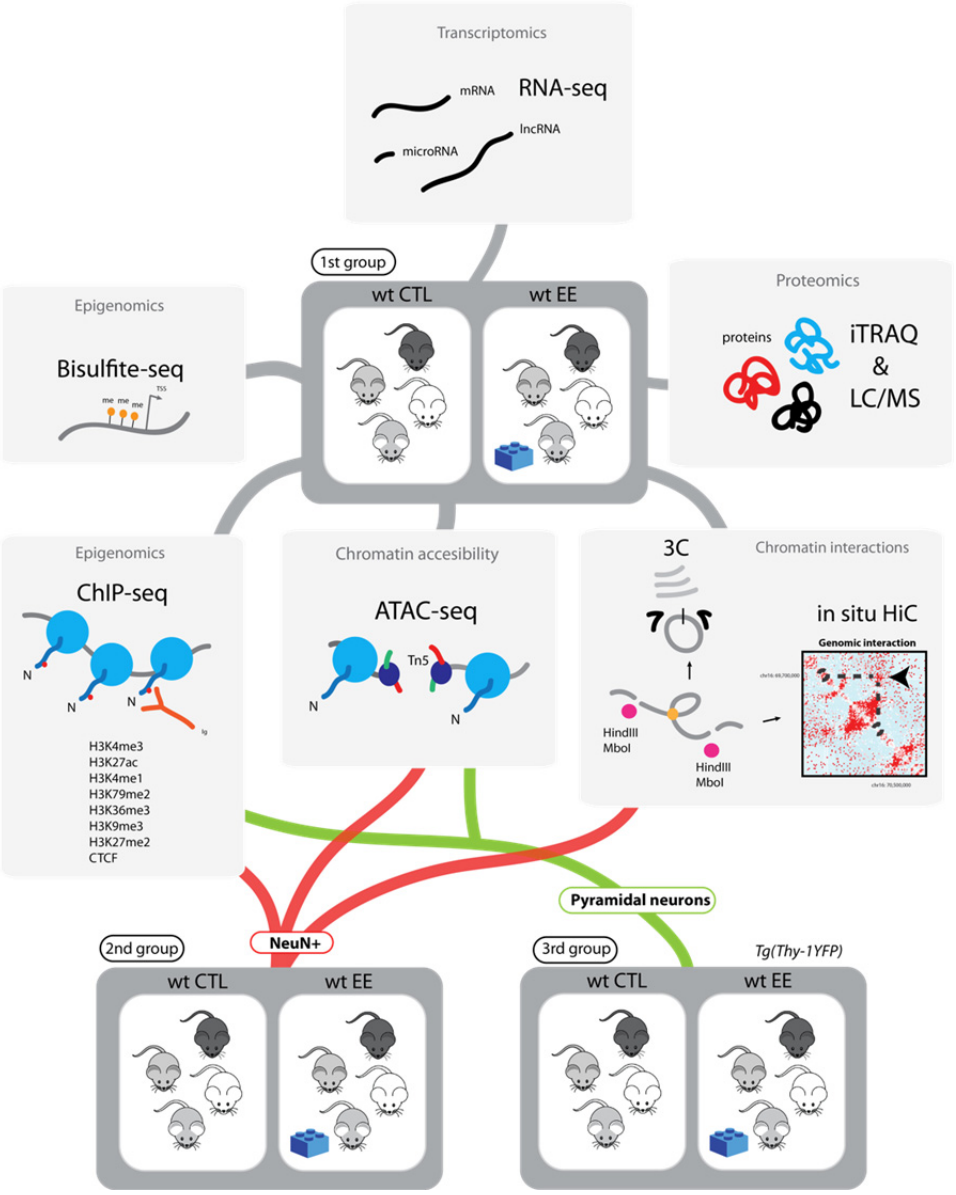


Figure ii. Overview of the Environmental Enrichment study

This work is supported by complementary data found at the QR codes in figures and at the beginning of some chapters. QRs can be read with any application available at your favorite application manager, but the data can be accessed via:

www.mousecortexregome.org

Index

Chapter 1
Introduction

We are constantly exposed to environmental factors in our daily lives. Although attempts are underway to understand the health impact of those factors, there is still much to learn concerning the complex gene regulatory modifications during cognitive development in the crucial early periods of life. Early experience-dependent learning is a key step in neuronal postnatal development and will determine an animal's adaptation to their environment.

The first evidence of the environmental importance in cognition arose well before Donald Hebb claimed his famous "Hebbian learning" hypothesis, in which he investigated whether a repeated stimulus can induce enhanced synaptic strength. He realized that the rats he took as pets performed the experimental tasks much faster and more accurately compared to standard maze rats (Hebb 1947). This observation was the origin of the concept of environmental enrichment (EE). EE consists of a combination of external factors that stimulate the brain by providing cognitive, motor and somato-sensorial stimulation and by maintaining constant novelty and complexity in the environment, thus boosting learning processes (Figure 1.1). Since that discovery, EE has been a factor widely used to study the molecular and electrophysiological mechanisms that allow neurons to achieve the goal of new memory constitution. The first evidenced effects of EE were linked to volumetric changes in the brain (Bennett et al., 1969) demonstrating increased gliogenesis (Szeligo and Leblond, 1977) and neurogenesis (Fabel et al., 2009; Kempermann et al., 1997; Kronenberg et al., 2003; Steiner et al., 2008; van Praag et al., 1999; Young et al., 1999) as well as increased dendritic branching (Bose et al., 2010; Greenough and Volkmar, 1973; Holloway, 1966), fiber axonal sprouting (Caleo et al., 2009) and synaptogenesis (Bednarek and Caroni, 2011). Despite all those observable effects it remains to be elucidated how an environmental factor paradigm such as EE can trigger cerebral changes in gene and chromatin regulation, i.e. how the 'regulome' is modified to allow adaptation to a constantly changing environment.

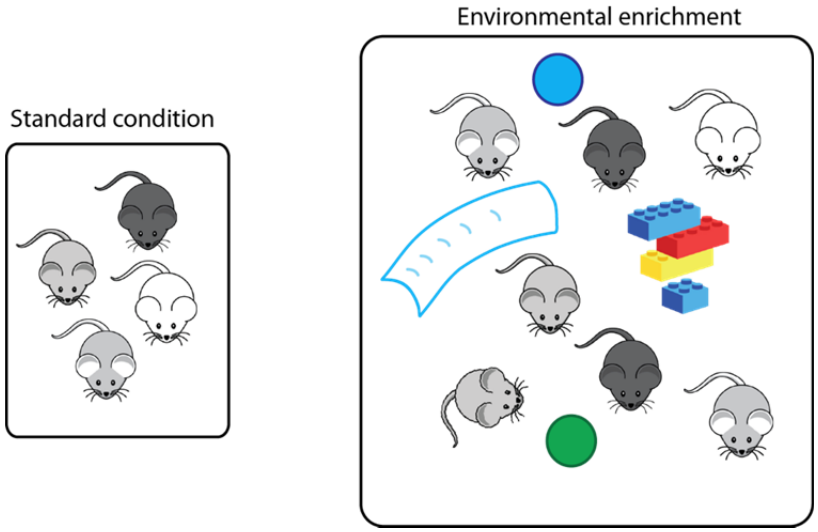


Figure 1. Environmental enrichment paradigm
Schematic view of mice under standard maze condition versus environmental enrichment (EE). Mice under standard conditions have less social interaction and smaller cages. On the contrary, EE mice have greater social interaction and they are exposed to constant novelty inducing motor, sensorial and cognitive stimulation.

1.1 Molecular consequences of environmental enrichment training in the brain

The so-called Hebbian learning process is brought about by the synchrony of signals from the presynaptic to the postsynaptic neurons that are induced by EE-dependent increased neuronal activity. This theory claims that due to increased synapse efficiency by repeated stimulation of the presynaptic to the postsynaptic neuron, the connection gains strength (Hebb, 1949). This is accomplished by molecular cascades occurring at the synaptic button, which then trigger morphological changes that induce cytoskeletal rearrangements and increased receptor and ionic channel insertions at the synaptic membrane (Priel et al., 2010; Voglis and Tavernarakis, 2006). This molecular process is known as long-term potentiation (LTP): a molecular mechanism of memory formation (Luscher and Malenka, 2012).

Although a large amount of publications have studied the molecular events leading to experience-dependent LTP, the mechanism is not fully understood. The best characterization involves the NMDA receptor activation and subsequent insertion of AMPA receptors into the postsynaptic membrane (Luscher and Malenka, 2012). This requires the activation of calmodulin and Camk kinases upon the calcium influx into the synaptic terminal (Herring and Nicoll, 2016) (Figure 1.2). On top of that, the process requires signaling transduction via the ERK/MAPK pathway (English and Sweatt, 1996, 1997) triggering the expression of early-expressed genes to support the potentiation mechanism (Flavell and Greenberg, 2008; Sweatt, 2001). As more receptors and ionic channels are recruited into the synaptic membrane, the terminal needs to grow. This is achieved by the Rho signaling pathway, which interacts with the cytoskeleton inducing the rearrangements needed to gain synaptic volume (Auer et al., 2011; Sweatt, 2001; Tolias et al., 2011). Finally, all of these dynamic changes would not be possible without energy, which in this case is supplied by the synaptic mitochondria (Vos et al., 2010). These organelles provide energy in the form of ATP and GTP to be used by the kinases that are required to phosphorylate substrates that boost cytoskeletal and receptor contribution changes. This dynamic process needs molecular blocks which are provided by synaptic ribosomes (Schuman et al., 2006).

The ultimate neuronal effect of EE is to induce changes in neuronal strength; hence it is expected that changes in neurotransmission will be observed. Glutamate and GABA release have been shown to be the main increased neurotransmitters in the hippocampus of aged rats due to EE (Segovia et al., 2006). This is supported by a higher contribution of the ionotropic glutamate receptors, NDMAR1 (Hu et al., 2013), NMDAR2A ((Andin et al., 2007), the metabotropic glutamate receptors, mGluR1 and mGluR5 in the prefrontal cortex (Melendez et al., 2004) and the AMPA receptors, GluR2 and GluR4 in the hippocampus (Naka et al., 2005). However, other studies could not find any significant increases of AMPA receptors subunits upon EE exposure (Andin et al., 2007; Gagne et al., 1998). Interestingly, it has been reported a GABA decrease in the early adult mice hippocampus compared to previous results in aged rats (Begenisic et al., 2011). This could point to a distinct effect of EE during different periods of life. Among other neurotransmitters, studies have reported altered levels of serotonin (Rasmuson et al., 1998), dopamine and acetylcholine (Del Arco et al., 2007a; Faherty et al., 2005; Wagner et al., 2005)

Another important breakthrough for the understanding of synaptic plasticity was the identification of growth factors that are induced upon EE exposure. These include nerve growth factor (NGF)(Mohammed et al., 1993; Pham et al., 1999), brain derived neurotrophic factor (BDNF)(Falkenberg et al., 1992; Hu et al., 2013; Mosaferi et al., 2015), glial-cell-derived neurotrophic factor (GDNF)(Young et al., 1999), and vascular endothelial growth factor (VEGF)(Cao et al., 2004). All of these factors are thought to mediate synaptic plasticity as well as cell survival, neurogenesis, and angiogenesis.

Developmental pathways such as Wingless signaling (Wnt) also appear to be targets of EE molecular effects. Though it is not deeply studied in this context, it has been shown that the ligand *Wnt7* is up-regulated upon EE exposure in the CA3 hippocampal region (Gogolla et al., 2009). Both isoforms of Wnt7 have been associated with neuro-architectural changes. The isoform *Wnt7a* triggers dendrite growth and synaptic strength activity by activating CamKII (Ciani et al., 2011). Moreover, *Wnt7b* has been previously reported to increase dendritic branching through the non-canonical pathway activating the subfamily Rho GTPase (Rac) and c-Jun N-terminal kinase (JNK) (Rosso et al., 2005). Other less explored developmental signals such as Bone Morphogenic Protein (BMP) may be also involved. For instance, exercise induces a down-regulation of *Bmp4* levels and seems to be important for neurogenesis and memory formation in mice hippocampus (Gobeske et al., 2009). Similarly, Meyers et al. reported an important up-regulation of the ligand *Bmp4* during aging, decreasing neurogenesis and promoting cognitive decline (Meyers et al., 2016).

All of the previous studies have focused on the effect of EE on synaptic plasticity inducing LTP processes. Despite their importance, the reality in postnatal development may be more complex. After birth neurons are enriched with a higher number of synapses than what is needed, in preparation to quickly learn about the surrounding environment (Stiles and Jernigan, 2010). As such, synapses are neuronal plastic elements that are capable to change upon experience and integrate outside information. In order to form new memories, not only is the reinforcement of existing synapses needed but also a fine-tuning or even removal of those no longer needed. This process is called synaptic scaling and it is essential for the brain physiology to adapt to the constant changing environment (Turrigiano, 2008). The molecular process triggering the retraction of synapses is known as long-term depression (LTD, (Collingridge et al., 2010).The research of Bednarek & Caroni (2011) showed that synaptic turnover together with the gaining of synaptic strength are two events that occur upon EE exposure.

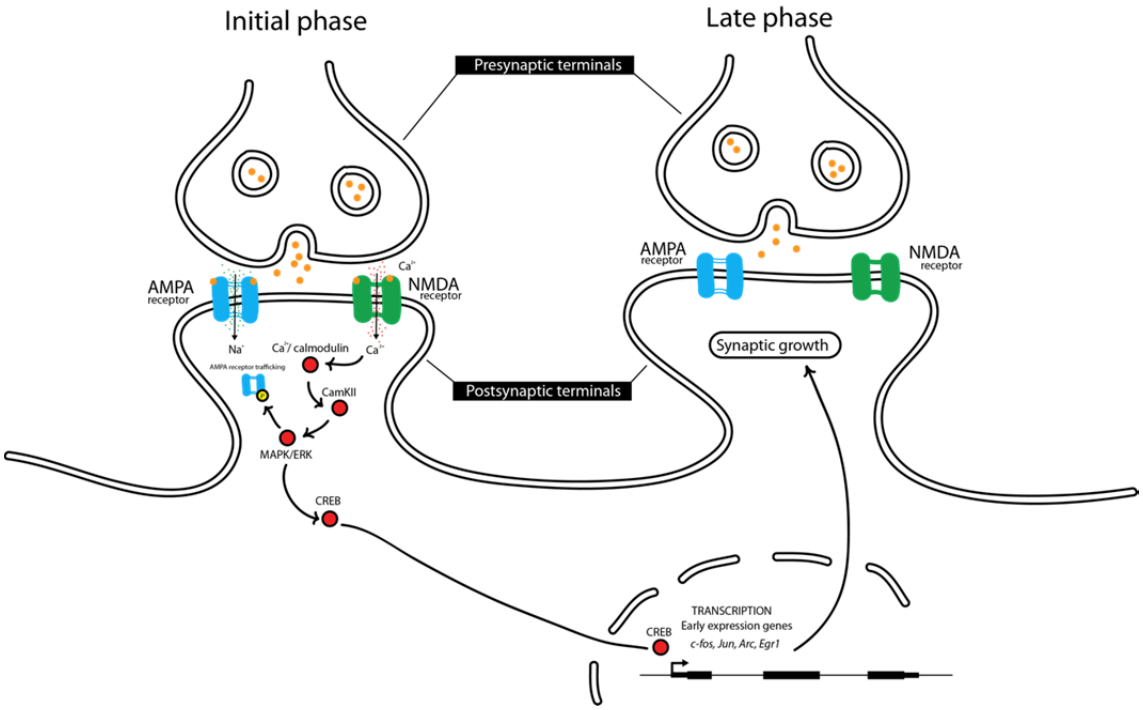


Figure 1.2. NMDA dependent LTP mechanism
Simplified model of long-term potentiation. At the initial phase the presynaptic terminal releases glutamate causing the depolarization of the postsynaptic neuron via the activation of AMPA receptors. This causes the displacement of magnesium from the NMDA receptor, that maintains this receptor blocked. Then, calcium enters to the postsynaptic terminal activating calmodulin and CamKII phosphorylating MAPK kinases. Those can phosphorylate AMPA subunits to integrate more receptors into the postsynaptic membrane. The MAPK/ERK pathway also transduces the signal to the nucleus activating transcription factors such as CREB that will activate the transcription of early-expression genes. The late phase might require a battery of new translated proteins to trigger the synaptic growth and maintain the potentiation.

1.2 EE as a means to maintain health and to partially recover from disease

The biological relevance of EE is evidenced in its ability of boosting learning and memory formation, as previously described. It has also studied how this non-invasive and non-harmful treatment can elicit its beneficial effects in a disease context.

The most studies on the effect of EE in disease have been focused on Alzheimer's Disease (AD) models. Jankowsky et al. showed that EE could ameliorate cognitive performance of mice overexpressing β-amyloid or presenilin-1 in a mouse model of AD (Jankowsky et al., 2005). However, this research also found an unexpected increase of β-amyloid levels that could not be replicated by others (Arendash et al., 2004; Wolf et al., 2006), while others studies found reduced levels due to EE (Lazarov et al., 2005). EE might mediate its effects by a neuroprotective mechanism elucidated by different studies. For instance, Xu et al. demonstrated that EE treatment was able to prevent microglia neuroinflamation after intraventricularly injecting soluble human β-amyloid (Xu et al., 2016). The molecular mechanism behind this neuroprotection could be explained in part by the increased expression of glial- and brain-derived neurotrophic factors, which are well known for their neuroprotective roles in avoiding apoptotic processes (Lindvall et al., 1994; Young et al., 1999). Costa et al. reported that these improvements are triggered by amyloid-dependent and -independent mechanisms, in which some of the overexpressed genes are related with β-amyloid sequestration (Costa et al., 2007).

Neuroprotective roles of EE have also been reported in the context of Parkinson's disease (Faherty et al., 2005; Jadavji et al., 2006) and cognitive amelioration in Huntington's disease (Hockly et al., 2002) as well as benefits in depression (Grippo et al., 2014), stress (Francis et al., 2002; Widman et al., 1992), autism (Woo et al., 2015) and other neurological disorders. However, not all studies conducted up to now have reported improved cognitive performance. For example, transgenic mice with mutations in the major human risk allele of AD do not show cognitive amelioration upon EE exposure (Levi et al., 2003). Similar results were also found in mice overexpressing APP, PS1 and Tau resulting in just mild effects (Blazquez et al., 2014).

As benefits of EE in neuronal disease, findings reported in other contexts are of interest as well. For instance, the research of Cao et al. described how EE could reduce tumorigenesis by activating the hypothalamic-pituitary-adrenal (HPA) axis pathway leading to decreased leptin in the bloodstream, thus reducing tumor growth (Cao et al., 2010). Similar results have been found in reducing breast tumor size (Nachat-Kappes et al., 2012). Taken together, it has been evidenced that EE has the ability to confer health benefits. As such it has potential as a non-invasive treatment to be used therapeutically for a wide number of diseases and as a means to maintain health.

1.3 Introducing the Regulome

Our understanding of the term *genome* has shifted notably since the Human Genome Project was completed fifteen years ago (Moraes and Goes, 2016; Venter et al., 2001). Slowly we are gaining new understanding of the huge non-coding part of our genetic material – previously knowns as *junk DNA*. The emerging picture integrates the more familiar protein-encoding *genes* with non-coding regions constituting the *regulome, i.e. any region facilitating regulation of transcriptional activity and chromatin conformation*. The genome is partnered with an array of DNA-binding proteins and post-translational modifications (PTM) constituting the chromatin. Heritable modifications of the chromatin that do not affect the primary DNA sequence constitute the epigenome (arising from the Greek '*epi*-', meaning 'above'), which plays a key role in the control of gene expression (Turner, 2007). Traditionally the epigenome comprises DNA methylation and histone modifications (Kouzarides, 2007; Smith and Meissner, 2013). However, it is important to note that other molecules such as non-coding RNAs (e.g. miRNA and lncRNAs) or DNA interacting proteins play an important role in regulating the genome and are often considered epigenetic factors as well (Allis and Jenuwein, 2016).

DNA methylation consists of the covalent union of methyl groups to cytosine DNA nucleotides and it tends to be enriched at CpG islands, areas of repeated units of cytosines and guanines (Ramsahoye, 2002; Ziller et al., 2013). In mammals, methylated promoters usually correspond to genes that are down-regulated (Deaton and Bird, 2011). A wide number of studies focus on brain methylation and especially DNA-binding proteins such as MeCP2, whose disruption leads to a severe neurological disorder called Rett Syndrome (Gabel et al., 2015). After DNA-methylation was described a novel subclass has been discovered, that is 5-hydroximethylation, highly enriched in brain tissues (Jin et al., 2011; Kriaucionis and Heintz, 2009; Munzel et al., 2010). Recent research has focused on understanding why the central nervous system is

more enriched with this kind of methylation, targeting roles in synaptic plasticity and neuronal differentiation (Santiago et al., 2014; Wen and Tang, 2014).

Histone modifications are post-translational modifications covalently linked to histone tails, mainly at the N-terminal. They can tag different aspects of the genome, modulating chromatin regulation and structure (Rao et al., 2014). Among the different modifications, acetylation is usually a measure of gene activity, as it occurs at active promoters and enhancers. Histone methylation is involved in both the activation and repression of genomic regions (Kouzarides, 2007; Shahbazian and Grunstein, 2007) (Figure 3). Among many others, one modification that is less explored is citrunillation, potentially having a role in chromatin decondensation (Mohanan et al., 2012; Wang et al., 2009) (Figure 1.3).

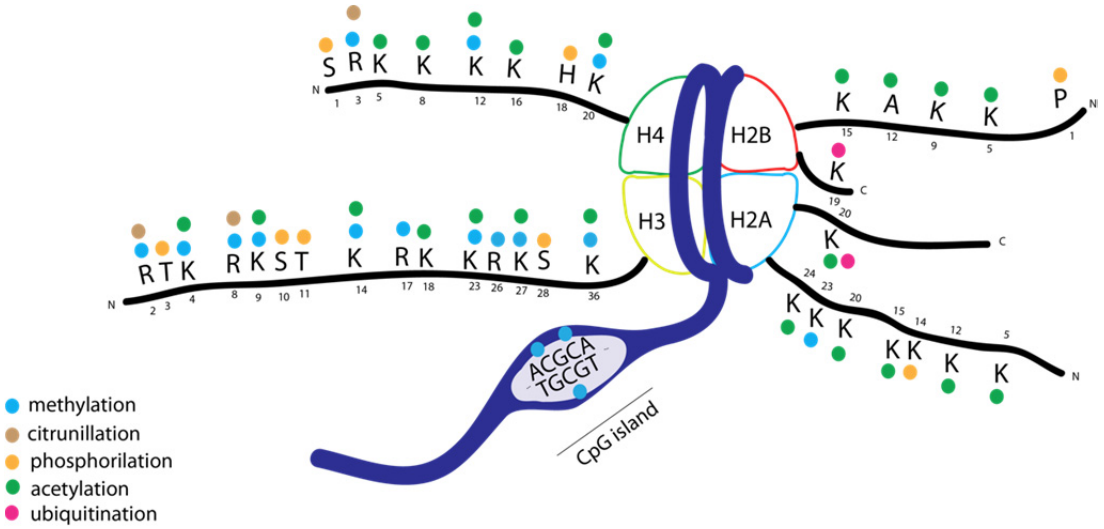


Figure 1.3. DNA-methylation and histone modifications overview
Schematic view of a nucleosome representing the histone modification patterns and the DNA methylation on a CpG island. Modified from (Rodriguez-Paredes and Esteller, 2011)

Epigenomics is a growing field of interest, as it constitutes an important part of the regulome. As a concept, we can understand the regulome as the total number of regulatory events that control the molecular phenotype of a cell. Among the different players, there are other important regulatory features. For example the destabilization of histone marks and reduced histone density around promoter regions can allow greater accessibility for transcription factors to bind to a genetic region and trigger transcription (Tsompana and Buck, 2014). Other important regulome features include the action of non-coding elements such as microRNAs that buffer the total yield of transcripts inside the cell (Jonas and Izaurralde, 2015). These can have an important impact on the molecular readout of the DNA. Non-epigenomic effects that influence the molecular phenotype of the cell and thus are part of the regulome include translation efficiency, RNA splicing and degradation rate as well as protein degradation. Indeed the challenge of current research is to understand how the environment can affect each of the contributors of the cell's regulome in order to better understand cell molecular physiology and its implications in disease. Evidence has been accumulated pointing to the importance of using a spatio-temporal and multi-dimensional perspective (i.e. measuring multiple omics features over time and space) whilst studying regulatory

components affected by environmental factors. The best example is found in twin studies. It has been shown that differing dietary habits during the lifetime of two genetically identical twin can induce changes in DNA-methylation and acetylation, thus impacting on gene expression, or the 'static genome' (Fraga et al., 2005).

1.4. Regulome changes induced by EE

Environmental enrichment affecting regulatory elements has not been deeply studied in epigenetic research. However, some reports reveal important discoveries. For instance the research of Irier et al. shows that EE can modulate 5-hydroximethilation by reducing its gene body levels in the hippocampus of both young and old mice (Irier et al., 2014). Interestingly this effect was significantly accentuated in hippocampus and not in cerebral cortex. Moreover, it was found that those differentially methylated genes were highly enriched for neuro-architectural changes. Regarding histone modifications Fischer et al. showed an increase in histone acetylation and methylation upon different EE exposure times in both cortex and hippocampus (Fischer et al., 2007) . Notably the changes seen in hippocampus were also greater than in cortex, similar to results found for the 5-hydroximethilation (Irier et al., 2014).

Other studies focused on a locus-specific analysis. For instance the research of Morse et al. reported a H3K4me3 coverage change in exon 4 of *Bdnf* regulating the already reported higher expression of that isoform upon EE training (Morse et al., 2015; Tao et al., 1998). Similar results were observed in the work of Kuzumaki et al. together with a down-regulation of H3K9me3 and H3K27me3 at the same exon (Kuzumaki et al., 2011). Alternative evidence of epigenetic changes comes from altered post-translational modifiers. For instance, the research of Rampon et al. reveals a dramatic increase of DNA-methyltransferases upon EE training (Rampon et al., 2000). Also, mice deficient for an epigenetic enzyme such as the acetyltransferase CBP presented marked deficiencies in EE-induced neurogenesis and were less responsive to EE-mediated enhancement of spatial navigation capabilities (Lopez-Atalaya and Barco, 2014).

1.5 Epigenetic mechanisms involved in memory formation

One of the main induced changes during EE training is an increase in learning and memory processes. The majority of studies concerning memory have been performed using fear-conditioning mechanisms. These studies found conserved mechanisms of memory consolidation when compared to other environmental factors, such as EE. In this context, it has been shown by Miller & Sweatt that DNA methyltransferases (DNMTs) are highly expressed upon exposure to a threatening factor, similar as in the study of Rampon et al. after EE training (Miller and Sweatt, 2007; Rampon et al., 2000). This study showed that DNMT is needed for memory formation at the hippocampus. The experience induced a differential DNA methylation in important genes in memory and cognition, such as increased levels at the promoter region of the memory suppressor gene *PP1* promoter as well as a methylation decrease of the synaptic plasticity gene *Reln.* Later, the same group discovered DNA methylation changes at the cortical level at CpG islands of the early expression gene *Egr1*, the synaptic plasticity gene *Reln* and the memory suppressor *CaN* (Miller et al., 2010). Recently the research of Halder et al. reported genome-wide DNA-methylation changes and histone post-translational modifications in learning-associated genes in both neuronal and non-neuronal populations of hippocampal CA1 and anterior cingulate cortex (Halder et al., 2016). In

fact, previous reports have shown also up-regulation of H3K9me2, H3K4me3, H3S10p and H3S10/H3K14p-ac in the CA1 region of the hippocampus (Chwang et al., 2006; Gupta et al., 2010; Peleg et al., 2010). Moreover, the research of Levenson et al. reported increased acetylation and phosphorylation of H3 in mice hippocampal CA1, highlighting the importance of histone deacetylase inhibitors as interesting pharmacological treatments for enhancing memory processes (Levenson et al., 2004; Vecsey et al., 2007). These changes in histone methylation can be due to cross-talking between DNA methylation binding proteins such as MeCP2 and histone modifiers such as histone deacetylases (HDAC) and acetylates (HAC) (Cross et al., 1997; D'Alessio et al., 2007; Jones et al., 1998; Nan et al., 1998)

Other memory consolidation paradigms have been used to study epigenetic changes in the brain. For example object recognition tasks have demonstrated an increase in H3 acetylation and increased activity of HDACS (Alarcon et al., 2004; Bousiges et al., 2010; Fischer et al., 2007; Korzus et al., 2004; Reul et al., 2009). Interestingly, Graff et al showed increased H3S10 phosphorylation, H3K14 and H3K5 acetylation, and H3K36 methylation after the task in hippocampus, finding a delay in increased changes in the prefrontal cortex (Graff et al., 2012). Furthermore, when inhibiting acetylation, phosphorylation or methylation processes by chemical blockers this processes of memory consolidation was disrupted, impairing the object recognition. Similar results were also observed when DNMTs were inhibited in the prefrontal cortex, preventing the recall of the existing memories (Miller et al., 2010). These results show that dynamic changes in histone or DNA modifications are needed for memory formation, thus highlighting the importance of the current study.

Evidenced here are the key roles the regulome components, post-translational modifications and DNA methylation, play in acquiring new memories and in brain function. This overlay of information could reveal how memories are formed, constituting an important milestone of research in neuroepigenetics.

1.6 Aims of the present study

Notwithstanding the attempts to understand the impact of environmental factors in our health, there is still much to understand concerning gene regulation complexity during cognitive changes in early periods of life. Experience-dependent learning is a key step in neuronal postnatal development to integrate all the surrounding new information useful for the correct adaptation and survival. Thus, EE is a powerful paradigm to understand all the orchestrated gene regulation changes that might occur in the early brain to achieve that goal.

We have designed an experimental workflow that could allow us to understand all the possible layers of regulation by studying the epigenome, namely chromatin modification, accessibility and conformation, as well as the transcriptome, and the proteome in the mouse cerebral cortex, neuron populations and pyramidal neuron populations. Our challenge is to analyze how these layers of information can be integrated to better understand the mechanisms affected by the changes of the regulome triggered by EE.

Cell heterogeneity, i.e. the complex cell composition of tissues, prevents the accurate identification of regulome changes in a specific cell type. However, it has the advantage of probing many cell types at once, thus making it more likely to discover potential target genes and regions. Nonetheless, findings need to be replicated in a cell-type specific population. Therefore we followed a top-down experimental design, starting with a discovery phase in whole-cortex, followed by genome-wide neuron-specific and pyramidal-neuron-specific experiments and finally locus-specific evaluation. We have generated a comprehensive landscape regulome changes induced by environmental enrichment.

Chapter 2 - Preface of ARTICLE 1

Contributions

My contribution consisted in the optimization and experimental work concerning of whole cortex ChIP-seq. Together with the sorting of neuronal nuclei (NeuN+, NeuN-) and pyramidal cells (GFP+ of Tg(Thy(YFP)) followed by ChIP-seq library preparation. I took care of the animals and I have performed the dissection for FACS-ChIP datasets myself. Moreover I have performed the mapping, the computational benchmarking of the differential binding analysis, by using both peak-dependent and independent methodologies with the assistance of Jekaterina Kokatjuhha and Aliaksei Holik.

Jekaterina Kokatjuhha worked and produced the peak-independent pipeline and Aliaksei Holik worked in the peak-independent window approach using CSAW.

Júlia Albaigues help me with the new breedings and some of the dissections.

Michael Maher helped me during the long days of sorting.

Meritxel Pons bred the first group of animals and the whole cortex dataset dissections.

Neuronal sorting took place at the FAC-sorting Facility leaded by Oscar Fornás at the UPF-CRG. Finally, sequencing was performed at the Genomics Unit leaded by Jochen Hecht at the CRG

Objectives

The main objectives of the present study are:

- To study epigenetic changes due environmental enrichment.
- To determine if FAC-sorted cells could allow understanding better epigenetic changes due a given treatment.

Highlights

- Peak-independent strategy allowed to better associated differential epigenetic due environmental enrichment (EE) in broad histone marks.
- H3K79me2 showed the greatest dynamic changes of all histone marks here analyzed and its differential activity could be linked to altered transcription rate production due EE.
- FAC-sorting revealed different epigenetic contribution for the different cerebral cortex populations.
- Neuronal H3K79me2 differential activity profile shows enhanced activity in genes related with synaptic function. Also supported by whole cortex transcriptomic changes.
- The NeuN negative population enriched by glial cells shows similar behavior when compared to the total cortex. This points to a great contribution of this population in the results of the total cortex differential analysis output.

Acronyms

Environmental Enrichment (EE), Fluorescent Activated Cell Sorting (FACS), Transcription Start Site (TSS), Polymerase Chain Reaction (PCR), Differential Binding Sites (DBS), Genomic Regions Enrichment of Annotations Tool (GREAT), Nitric Oxide (NO), Mitogen-activated Protein Kinases (MAPK), α-amino-3-hydroxy-*5-methyl-4-isoxazolepropionic acid* receptor (AMPA), *N-methyl-D-aspartate* receptor (NMDA), *Bone morphogenetic proteins (BMP), Differential Expressed Genes (DEGs), Gene Ontology (GO), False Discovery Rate (FDR), Postranslational Modifications (PTM).*

Chapter 2
De-convoluting epigenetic changes in mouse cortex induced by environmental enrichment

2.1 Abstract

Epigenetic modifications modulate the regulatory and transcriptional activity of different regions of the genome. Thus, they have a key role in regulating the genome by tagging active regions to be transcribed or to be repressed by compaction in non-expressed heterochromatin segments. It has been always questioned how this layer of information interacts with the environment. In this study we have used Environmental Enrichment (EE) to determine how several histone marks and CTCF change their binding and how chromatin conformation is modified upon the motor, sensorial and cognitive stimulation of the treatment in the mouse cerebral cortex. To distinguish cell type specific effects from changes introduced by cell type composition of the complex cortex tissue we additionally analyzed the histone mark H3K79me2 in both neuronal and non-neuronal populations derived by FAC-sorting. We found significant changes of the epigenetic profile after exposure to EE for one week or one month that boost learning and memory processes.

2.2 Introduction

Histone modifications are covalent changes that modulate the structure of chromatin by their electrochemical properties. This role of demarking, known as "the histone code" (Jenuwein and Allis, 2001), can prime genomic regions to be: transcribed, repressed, or to allow proteins to bind for regulatory purposes. The method most commonly used in order to study the "code" is called Chromatin Immuno-precipitation followed by sequencing or simply: ChIP-seq, introduced by Barski and colleagues in 2007 (Barski et al., 2007) (see Methods).

The Environmental Enrichment (EE) paradigm stands for a constant novelty of factors to induce the stimulation of different cerebral regions and trigger learning and memory. Despite the big number of studies reporting molecular and phenotypic changes due the exposure to this treatment (Nithianantharajah and Hannan, 2006; van Praag et al., 2000), EE has not been deeply studied in an epigenetic context. However some studies reported global histone acetylation and methylation changes in mice hippocampus (Fischer et al., 2007) or locus-specific changes, such as an increase of H3K4me3 at *Bdnf* exon IV (Morse et al., 2015). Our study provides the opportunity to advance the field and to shed light on the changes of the epigenetic landscape induced by the treatment.

The main goal of our study is to obtain a full perspective of the mouse cortex epigenome under the influence of EE, thus, providing evidence of all the regulatory mechanisms that might change under the treatment condition. To that end we designed a workflow that includes ChIP-seq based analysis of seven histone marks and CTCF allowing a comprehensive profile of all possible regulatory regions that can be either active or repressive (Figure 2.1). In order to better distinguish cell-type specific signals from the noise introduced by cell type heterogeneity in complex tissues we sorted into neuronal and non-neuronal populations by using the neuronal marker NeuN.

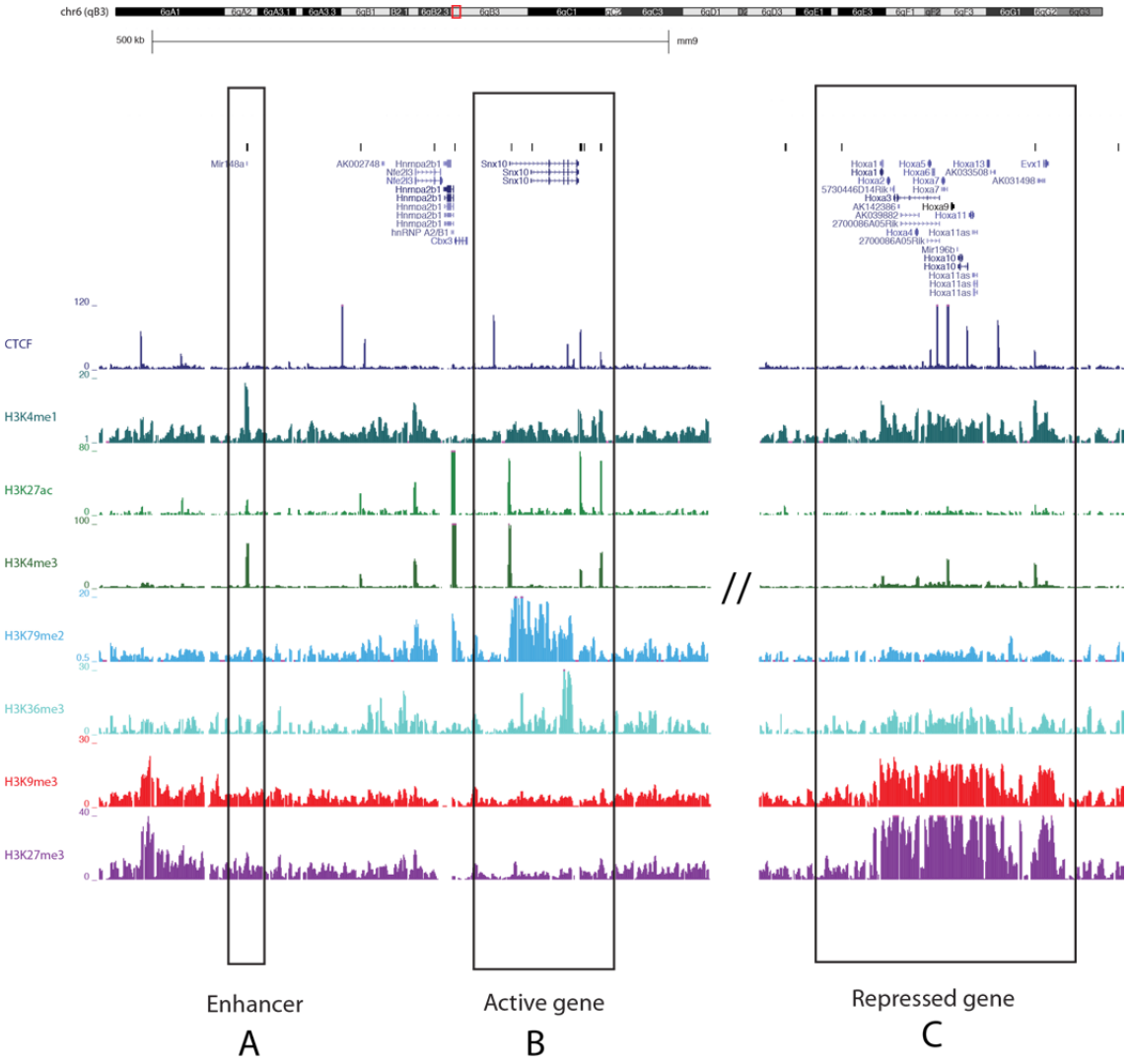


Figure 2.1. ChIP-seq landscape
Genome Browser tracks of all the histone modifications and CTCF used in the present study. It clearly shows differential activity around the *Hox* gene locus (C). **A)** Illustrates H3K4me1 signal upstream the track, denoting an enhancer regulatory region, showing in some extent H3K27ac and H3K4me3 occupancy. Downstream appear the broad gene body marks: H3K79me2 and H3K36me3 tagging gene activity highly correlated with transcription (Dong and Weng, 2013). In this track it is evident how their coverage intersect along the gene, finding the highest H3K79me2 density upstream, close to the TSS; complementary H3K36me3 signal is enriched downstream, close to the end of the gene(Huff et al., 2010) **(B).** As *Hox* genes are mainly repressed in the adulthood repressive marks such as H3K27me3 and H3K9me3 show higher occupancy in this region **(C)**.

2.3 Workflow and bioinformatics pipeline

In this study we have profiled 7 different histone modifications (H3K4me3, H3K27ac, H3K4me1, H3K79me2, H3K36me3, H3K27me3 and H3K9me3) and CTCF to obtain a comprehensive map of epigenomic changes induced by EE. To achieve this we have first bred 2 biological replicates per condition. Each of them corresponds to 5 different cortices pooled and homogenized. We have performed ChIP-seq for each of them using ~ 50k cells, while using 700k cells for CTCF. Cell heterogeneity in complex tissues is a problem because different signals from a vast number of distinct neuronal and glial populations are mixed, possibly leading to misinterpretations. For that reason, we decided to sort both neuronal and non-neuronal populations by the neuronal marker NeuN+ (also called Rbfox3) (Matevossian and Akbarian, 2008) in samples taken after one week or one month of EE and to perform ChIP-seq for the histone mark H3K79me2. (Figure 2.2 A)

ChIP-seq is a powerful method to study gene regulation, but it can result in a noisy output due technical biases such as un-specific binding, efficiency of the procedure, composition biases of the final library, etc.. Because these potential technical issues can arise, different pipelines have been designed in order to get the best results and better assess the epigenetic changes due to the given treatment (Bailey et al., 2013). We designed two main methodologies: a peak-dependent method, using the most common peak-callers to map the occupancy of each mark; and a peak-independent method. In the latter we have performed a window-based approach (data not shown here) and an annotation based approach in which we compare the coverage in predefined regions such as promoters, gene body or enhancers. In this chapter we will describe the results of the epigenome analysis. We will benchmark all methods to highlight their strengths and weakness and to evaluate, which method gives the most consistent and significant results (Figure 2.2 B).

For more details regarding the implementation of the analysis pipelines please see the Methods Chapter.

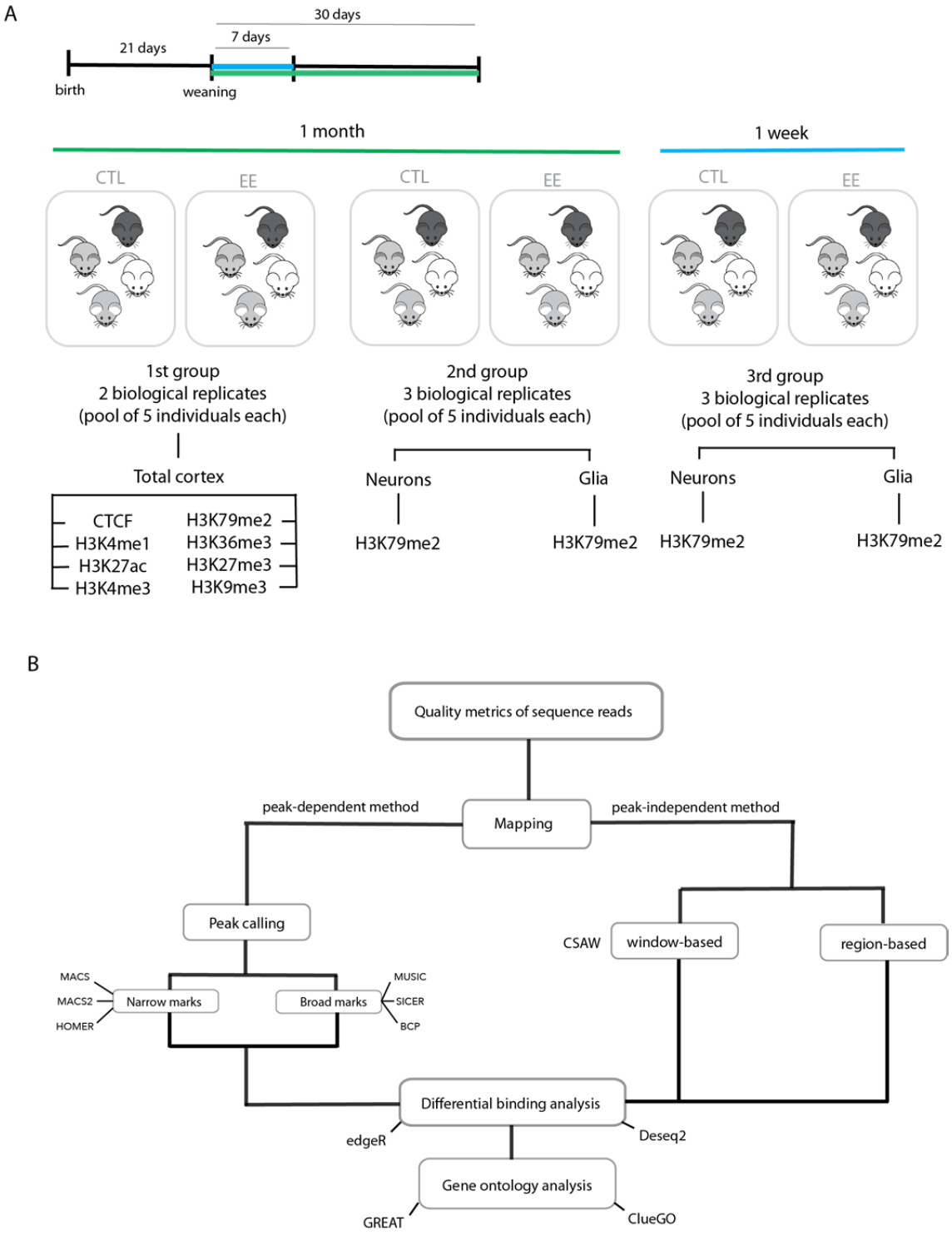


Figure 2.2. Experimental and bioinformatics epigenetic study workflow
A) Schematic timeline of the treatment. Immediately after weaning animals started the environmental enrichment period, either during a month or a week. Animals were divided into 3 groups. The first group of animals has two biological replicates per condition. Each of the replicates corresponds to homogenized and pooled cortices of 5 different animals. This first group belongs to an initial study where nuclei preparation was performed followed by ChIP-seq against 7 histone marks: H3K4me1, H3K27ac, H3K4me3, H3K36me3, H3K79me2, H3K9me3, K3K27me3 and a non-histonic mark: CTCF. A second group, corresponding to three biological replicates this time, was also enriched during a month and cortex samples were FAC-sorted using NeuN marker to get neuronal and non-neuronal populations. Therefore, H3K79me2 was performed for both populations. Finally a third group was bred, again three biological replicates, followed by the same procedure as the second group. However non-neuronal population data is not shown in the present study. **B)** Computational benchmarking pipeline performed for ChIP-seq data. A peak-dependent and independent method was designed. The first uses different algorithms for peak detection and annotation followed by differential analysis using DiffBind and gene ontology analysis using GREAT. The second analysis approach was peak-calling independent. Two strategies were used, the first one binning the genome into windows (CSAW) and the second looking exclusively to regions of

interest based on pre-defined genome annotation and the occupancy pattern of each mark. The differential analysis was performed using edgeR and/or Deseq2 followed by the gene ontology analysis using the Cytoscape application: ClueGO+Clupedia.

2.4 Results

2.4.1 Peak-dependent strategy to study epigenetic changes induced by Environmental Enrichment

Taking into account the different strategies that peak-callers use to report and annotate the protein occupancy of interest, we first asked how the differential binding analysis would be influenced by the peak annotation provided by different peak-callers. Therefore we compared the results of differential binding analysis performed by the tool DiffBind (a wrapper for the two tools edgeR and Deseq)(Ross-Innes et al., 2012), using as input the peak annotation from the peak callers: MACS, MACS2 (Zhang et al., 2008), SICER(Xu et al., 2014b), HOMER (Heinz et al., 2010), MUSIC (Harmanci et al., 2014) and BCP (Xing et al., 2012).

The results show consistent differential binding regions for H3K4me3 and H3K79me2 for all peak-callers. However, this is not the case for the remaining five histone marks and for CTCF. For these marks significant differential binding was only found with some peak-callers, and not consistently (Figure 2.3 A). This output makes us aware that peak-calling and subsequent differential binding analysis in annotated peaks may not be the best strategy to analyze the data, a reason for designing our own peak-independent methods (discussed bellow, see Methods for more details). The reason behind the discordant outputs can be linked to the different methodology used by each peak caller and also to technical problems such as library complexity, cell heterogeneity and other potential batch biases. Furthermore peak calling in different cohorts and replicates can lead to partially overlapping peaks, complicating the differential occupancy analysis. Finally the peak length varies substantially between the different histone marks, ranging from very short peaks only found in transcription start sites (TSS) or enhancers for H3K4me1/3 to long peaks typically observed for gene body and repressive marks. Some peak callers have thus been specialized to short or long peaks, making their results less comparable.

Low library complexity leads to PCR duplicates and hence redundant reads originating from the same DNA fragment, which map to the exact same position in the genome. Some peak-callers simply remove duplicate reads, while others follow more complex statistical models to reduce their influence on the results. It is common for narrow peaks to contain high redundancy values, especially when pooling low cellular yields. The issue is intensified for histones such as H3K4me1 that are concentrated in a small part of the genome compared to other marks, forming strong peaks e.g. only around TSS. When these histone modifications are pulled low yields of DNA material are obtained. Thus, higher PCR amplification cycles are needed in order to obtain a proper library for sequencing, inducing a bias in duplicate numbers and affecting dramatically the complexity of the library. The direct consequence of this phenomenon is that often almost half of the reads are lost after filtering (Figure 2.3 B). Despite the technical improvements to solve the problem it could be partially ameliorated by performing a deeper sequencing, providing more unique reads to the differential analysis and potentially getting more results for those peak-datasets.

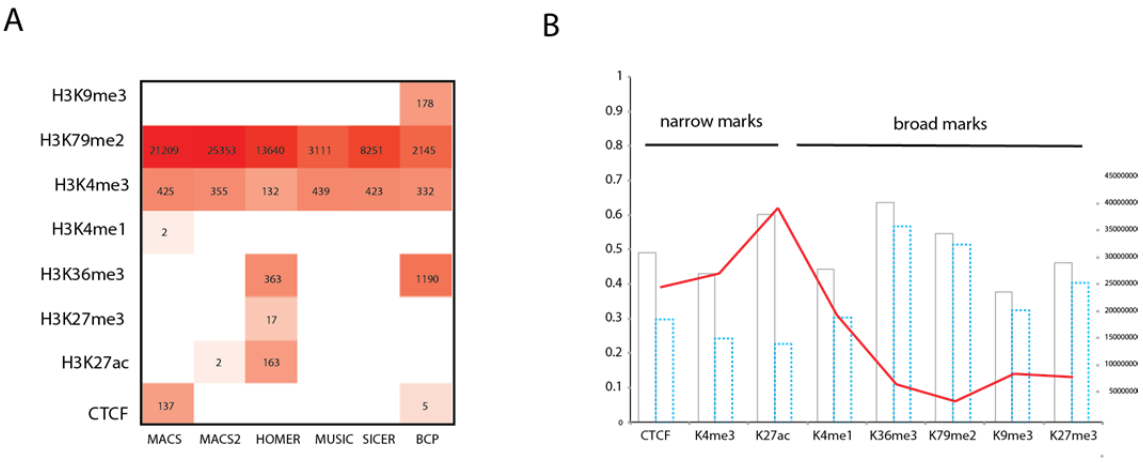


Figure 2.3. Peak-dependent benchmarking and library complexity analysis.
A) Differential binding sites (DBS) reported by edgeR-DiffBind using as input six commonly used peak-callers. H3K4me3 and H3K79me2 report DBS of each of the peak-callers used. **B)** Redundancy analysis performed by MACS, red line reports the redundancy score percentage and bars show the amount of reads available before (gray) and after (blue) the filtering. The Figure clearly illustrates how narrow marks are highly affected by library complexity.

2.4.1.1 Promoter activity changes induced by EE: increasing norepinephrine, and glucocorticoid signaling whereas decreasing BMP and ionic channels

The major histone mark showing consistent results linked to promoter activity is H3K4me3 (Figure 2.3A). Although differential binding sites (DBS) were reported by using different peak callers, there was almost no intersection between the reported DBS. For that reason, the strategy followed focused on the highest DBS output computed with DiffBind, that corresponded to the usage of MACS, MACS2 and MUSIC. SICER was excluded as it optimized for broad marks Figure2.3 A. Next, we used GREAT to annotated the DBS relative to the mouse gene annotation mm9 using the association rule: single nearest gene +100kb distal (McLean et al., 2010). All regions associated to genes were loaded into ClueGO (Bindea et al., 2009) to perform a gene ontology analysis.

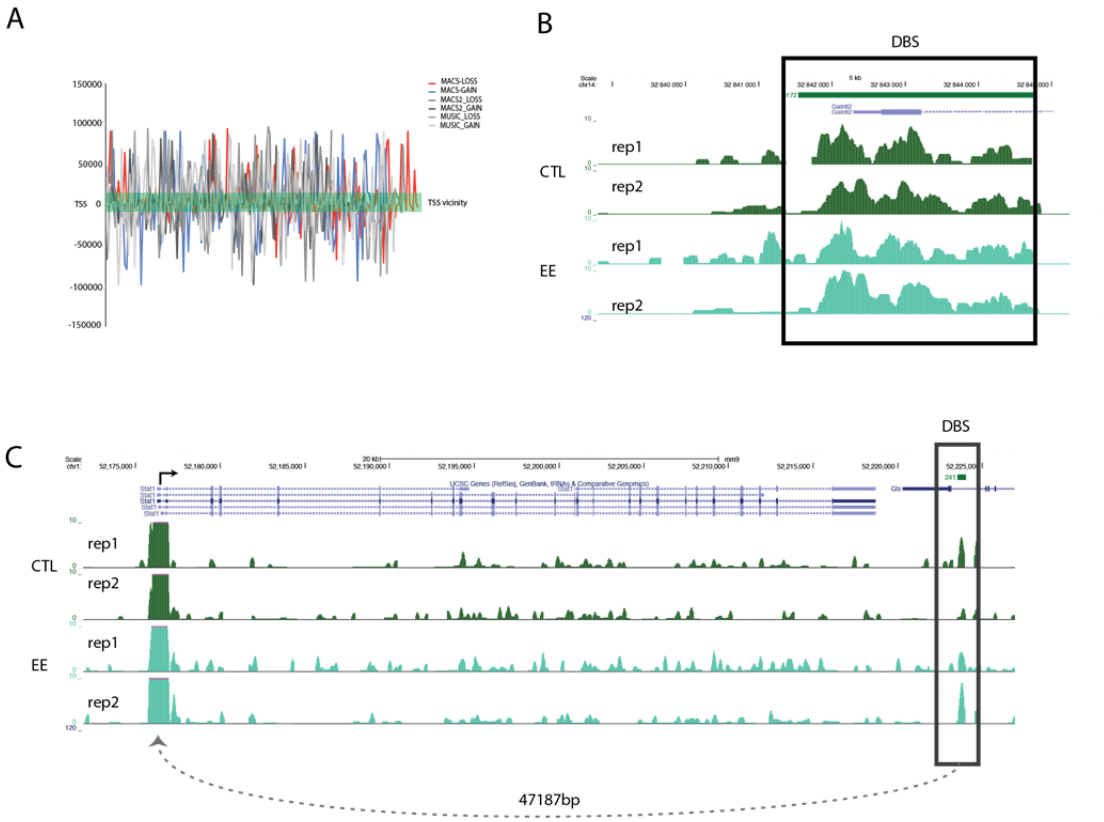


Figure 2.4. H3K4me3 DBS distribution along the genome
A) Distance in base pairs to TSS (y-axis) of the DBS (x-axis) reported by the differential analysis using the peak-callers MACS, MACS2 and MUSIC. Green color denotes the close vicinity of DBS to the TSS. **B)** Example of a DBS falling into the promoter region of *Galntl2* **C)** Distal H3K4me3 DBS found 47kb downstream to its associated gene *Stat1*.

A main problem encountered with the H3K4me3 DBS output is that a large fraction of DBS regions is located far away from the closest TSS (>50kb) (Figure 2.4 A). However, this mark is thought to be enriched in the vicinity of TSS, making this result unexpected. On top of that, these regions, although being significant, show low coverage, raising doubts about the validity of these DBS regions (see examples of DBS Figure 2.4 B, C).

Despite the possibly low quality of the dataset, the union of DBS reported by MACS, MACS2 and MUSIC revealed some interesting terms in the gene ontology analysis related with known effects of EE (Figure 2.5 A). For example DBS terms are enriched for response to norepinephrine stimulus (p=4.1E-3) that has been previously reported to be part of monoaminergic neuronal contribution on plasticity maintenance (Naka et al., 2002). Moreover, glucocorticoid metabolic processes (p=6.1E-3) play important roles reducing anxiety (Durairaj and Koilmani, 2014) having protective roles under stress (Zanca et al., 2015). Interestingly, a contradiction is here highlighted as increased terms show astrocyte fate commitment (p=28E-3), but glucocorticoids at the same time inhibit astrocyte proliferation (Unemura et al., 2012). Growth factors also play important roles in proliferation and differentiation and the link with previous terms arises by the increased promoter activity of negative regulation of neurotrophin TRK receptors (p=20E-3). Indeed, this can point to a possible negative feedback loop of the pathway followed by a reported increased level of neurotrophins after enrichment (Ickes et al., 2000).

One important element enriched is the regulation of the mitochondrial outer membrane by the term: transport permeabilization in mitochondria (p=23E-3). As a matter of fact, it has been shown that mitochondria not only play a key role in providing the source of energy at the synapse, but also in maintaining calcium homeostasis during neurotransmission (Vos et al., 2010). In addition, neurotransmission is also enriched by an enhanced L-glutamate transport (p=39.0E-3). Denoting the importance of glutamatergic involvement in the response to EE, that is a substantial molecular part of the main form of memory and learning processes (Segovia et al., 2006).

At first sight, it is striking to find a decreased neuronal fate commitment (p=83.0E-15), but it is important to highlight, as we will see later on, that different cellular contributions to the final output mix in this analysis. Second, the BMP signaling pathway is decreased due to the EE exposure (p=1.9E-3). In spite of other major signaling pathways induced by the treatment (such as Wnt signaling), recently Meyers et al. discovered that this pathway is gained during aging (Meyers et al., 2016), making for the first time a link of BMP between EE and aging. In addition lost signals also show synaptic transmission effectors, by the decreased activity on genes involved in the activation of G protein gated potassium channels (p=170E-3) that contribute to hyperpolarization resting period after action potentials (Luscher and Slesinger, 2010). Moreover there is a down-regulation of inhibition of voltage gated Ca^{2+} channels via Gbeta/gamma subunits (p=170E-3) also linked to synaptic function (Voglis and Tavernarakis, 2006). Furthermore related to neuronal transmission, we found a decrease on H3K4me3 activity of GABA B receptor activation (p= 270.0E-3), in consonance with previous results showing a decrease GABA release induced by EE (Begenisic et al., 2011).

Notably it is interesting to find the enriched term eicosanoids (p=200E-3) that has ben shown in some reports to reduce tumor growth (Garofalo et al., 2015; Nachat-Kappes et al., 2012). Nevertheless there are other studies that claim the contrary (Westwood et al., 2013). In any case the gene enriched for this term corresponds to the cytochromes that mediate the hydroxylation of eicosanoids (Cui et al., 2001), not directly tagging effectors for the previous effect.

Other minor enrichments found in the gene ontology analysis are the decrease of negative regulation of oligodendrocytes differentiation (p= 200E-3) and Notch signaling pathway (p=260E-3). Gliogenesis, and especially oligodendrocyte contribution, is one of the reported phenomena that EE might induce (Szeligo and Leblond, 1977).

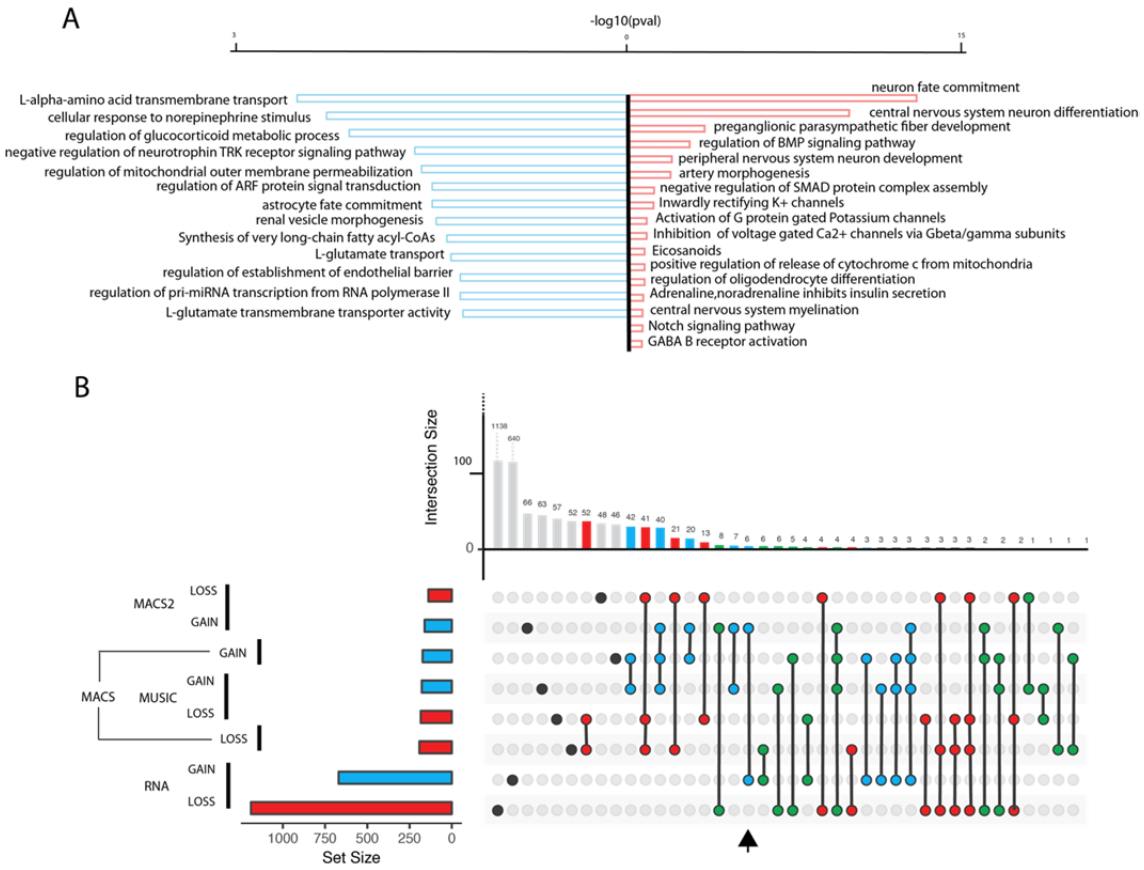


Figure 2.5. Biological processes induced by EE and gene intersection between methods and RNA-seq
A) Gene ontology analysis using GREAT followed by ClueGO. It illustrates biological enriched terms gained (blue) or lost (red) by the effect of EE and ranked by the –log10(p) values. **B)** Intersections of H3K4me3 DBS found using different peak callers with RNA-seq. Left graph shows the amount of elements that is on each dataset. Blue bars correspond to gains, red to losses. The upper graph shows the number of genes intersected that shared or exclusive (in this case dots do not connect with any other set). Shared genes between datasets are spotted in the color of the direction (if both increased – blue, if decreased –red, discrepancies are shown in green).

2.4.1.2 H3K4me3 do not intersect with gene expression

In order to explore the relationship between H3K4me3 activity change and transcription, an intersection analysis was performed (Figure 2.5 B). The figure clearly illustrates that differential binding of H3K4me3 (based on the peak callers MACS, MACS2, MUSIC) and transcriptional activity are not well correlated. This means that changes in promoter activity indicated by differential binding of H3K4me3 are not necessarily translated into changes in transcriptional activity; only 6 up-regulated genes were found in common at the RNA-seq level and by the unique usage of MACS2 (black arrow, Figure 2.5 B). Notably influenced by the limited power of discovery of this assay, redundancy could play a key part. This result can be however also explained, because promoter histone modifications do not correlate well with transcription. In fact it has been shown that promoter marks are more related to transcription factor binding rather than a direct effect on transcription (Dong and Weng, 2013). Furthermore the quantity of a transcript found in a cell is not only influenced by transcriptional activity, but is to various degrees influenced by RNA transcription, processing and degradation efficiency in a cell-type and developmental stage specific manor (Rabani et al., 2014). However, the different peak-caller methods intersect properly in the same contrast direction in a considerable number of shared hits (that constitutes around 25% of each dataset). Especially, the peak results of MACS and MUSIC show a large intersection.

2.4.1.3 Gene body mark induced DBS

Another mark that displays significant output of DBS using different peak-callers was the broad mark H3K79me2 (Figure 2.3.A). This modification is a gene body mark linked to gene activity and has been reported as the histone mark that shows the highest correlation with transcriptional activity (Dong and Weng, 2013). This present study is the first time that H3K79me2 is analyzed in the context of a treatment, manifesting some peculiar behaviors of the mark that will be discussed in the next sections.

The main disadvantage of the peak-dependent pipeline is that it can fragment the DBS analysis of a single genomic region (e.g. gene) depending on the number of distinct peaks that were called. In the case of the H3K79me2 dataset this leads to a large number of DBS, which typically occur in clusters. As expected, the majority of DBS were found in the gene body, but not in a single peak but in multiple adjacent peaks spanning the full gene body (Figure 2.6). This represents a problem when trying to associate the genes, as some of them might end as false positives being associated to other genes upstream or downstream. Even if DBS clustering could be a solution, it might also produce false positives or negatives, needing to optimize each time the correct threshold to be established for different histone marks. Moreover, the fragmentation of DBS leads to multiple results per gene, which reduces the significance level that can be achieved. For that reason we decided to develop or use other methods to study both histone and non-histonic proteins, using a peak-independent strategy: either sliding window-based or using predefined annotations of genomic features like genes (Figure 2.2.B).

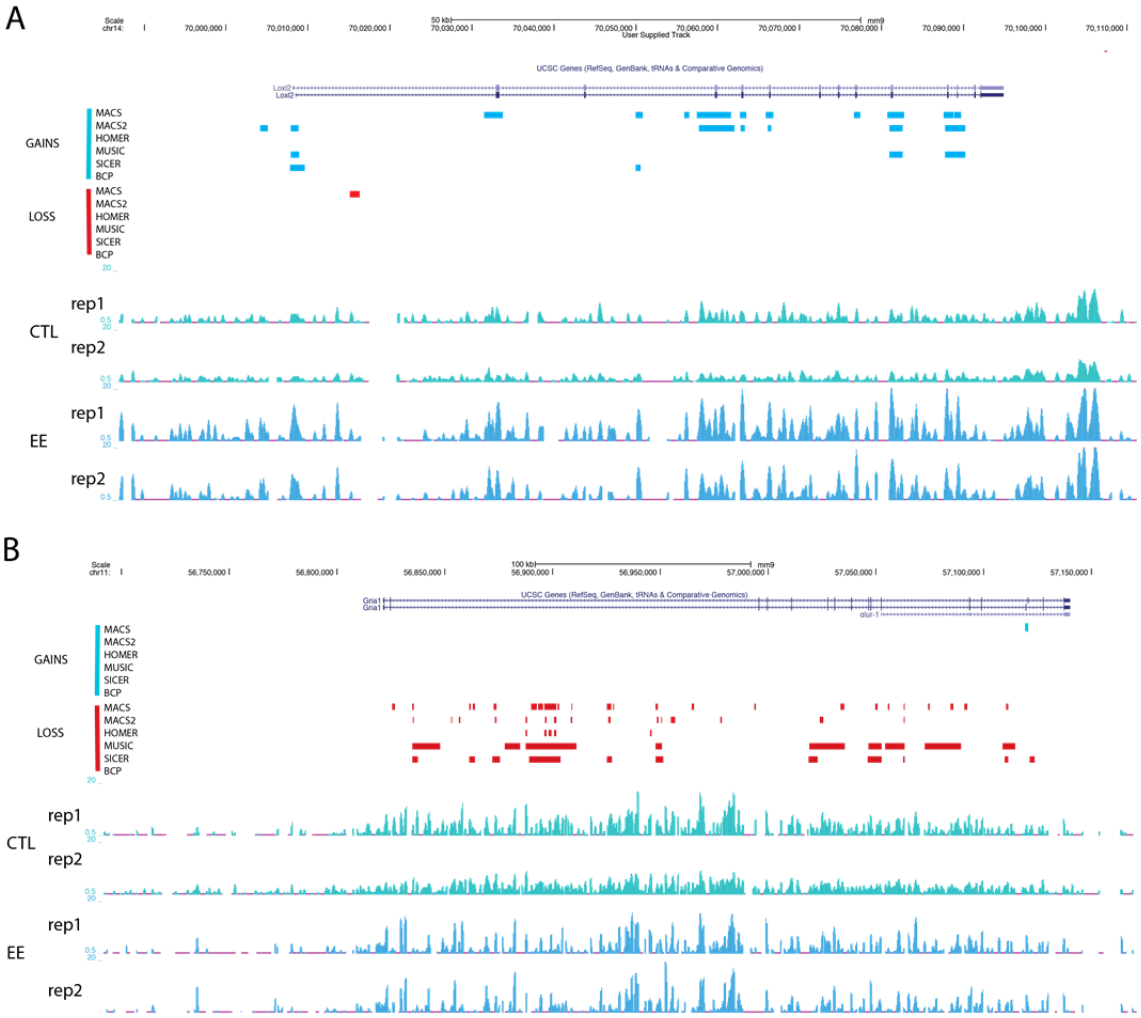


Figure 2.6. H3K79me2 coverage
USCS tracks of the broad histone mark H3K79me2. **A)** Gain of binding due to EE in *Loxl2* gene, Light blue tracks represent each of the control (CTL) biological replicates and in dark blue the EE samples. Blue peaks represent the gain DBS calls for different peak-callers and in red are represented the losses. **B)** Loss of H3K79me2 activity on *Gria1* gene.

2.5 Peak-independent strategy: window or annotation based method

As previously mentioned, one of the main issues of the peak-dependent approach is the large amount of fragmented DBS reported for broad marks such as H3K79me2, making it difficult to obtain highly significant results and to perform the gene association analysis with GREAT (McLean et al., 2010). For that reason we decided to look into different ways to analyze ChIP-seq data. The first approach consists of a window-based analysis where the genome is binned into windows in which a read count is obtained. Read counts per window are normalized for all possible ChIP-seq biases, and are subsequently used for differential binding analysis. This method is implemented by the tool CSAW (Lun and Smyth, 2016). Initial tests show good sensitivity of the method, however the analysis has not yet been finished and the preliminary results will not be discussed here. We developed a second approach consisting of a histone occupancy analysis in predefined regions of interest. These regions could be chosen from gene and regulatory region annotations, e.g. mm9 or mm10 for mouse. For instance, we selected the complete gene body for the broad marks H3K36me3 and H3K79me2; 500bp upstream and 1500bp downstream the TSS for H3K4me3, H3K27ac and CTCF; and 700bp upstream the TSS to the end of the corresponding gene for the repressive marks.

H3K4me1 was excluded, as its occupancy pattern mostly overlaps with enhancers. Next, reads were counted per region and DBS analysis per region was performed using edgeR (Robinson et al., 2010) or Deseq2 (Love et al., 2014). The majority of DBS were found for the gene body marks H3K79me2 and H3K36me3, which will therefore be the main object of discussion in the next sections.

2.6 Effects of EE on total cortex gene activity and its relationship with transcription

Both broad gene body histone marks are known for their link to gene activity and their high correlation with transcription (Dong and Weng, 2013). The coverage of the two marks across the gene is anti-correlated, with high levels of H3K79me2 at the 5' end, decreasing towards the 3' end vice-versa for H3K36me3. Here it will be discussed how this coverage pattern changes due to the treatment, as well as a preliminary overview of the main functional GO term enrichments among differentially active genes and finally how this change in epigenomic activity is be related to transcription.

2.6.1 Dynamic changes of histone coverage due to environmental enrichment

Plotting the average coverage through the differentially bound regions found by edgeR (Robinson et al., 2010), we observed differences among gains and losses histone profiles (Figure 2.7.A). Notably, genes gaining H3K79me2 occupancy show their highest read coverage centered at the gene body, instead of close to the TSS as observed across the majority of genes. Interestingly, H3K79me2 losses show the common coverage distribution (Huff et al., 2010), the number of reads decreasing from the TSS towards the 3' end of the gene. Nevertheless, looking carefully at the losses, we detected an increased number of reads compared to the control samples at the very beginning of the gene (see black arrow, Figure 2.7.A right). This dynamic change could elicit the functional role of H3K79me2, as it has been postulated to be enriched in long upstream antisense transcripts associated genes (Lepoivre et al., 2013), demarking a role in non-coding regulatory elements. It has also been suggested to be important for early elongation of transcripts and for mediating RNApol II transcription speed (Bintu et al., 2012; Jonkers and Lis, 2015). However, it seems that the gene length could affect this behavior. Surprisingly the average gene length of the losses is almost 6 times longer than the regions of increase (134kb and 23kb respectively). Finally, the dynamics of H3K36me3 follows the expected distribution of the reads, both for gains and losses. Although for H3K36me3 it shows a similar tendency for the first part of the gene body when looking the losses (Figure 2.8 A).

2.6.2 Gene ontology analysis: neuronal architecture, signaling, proliferation and angiogenesis epigenetic signal changes induced by EE

Gene ontology analysis gives a full perspective of gene functions by clustering the genes discovered within the differential analysis according to their function and by testing for significantly enriched functions. However when the number of genes is large, as is the case for this dataset, it tends to show general terms containing large number of genes, which are difficult to interpret. For this reason we have removed large GO terms

from the analysis to simplify the hypothesis generation and only important concepts will be discussed.

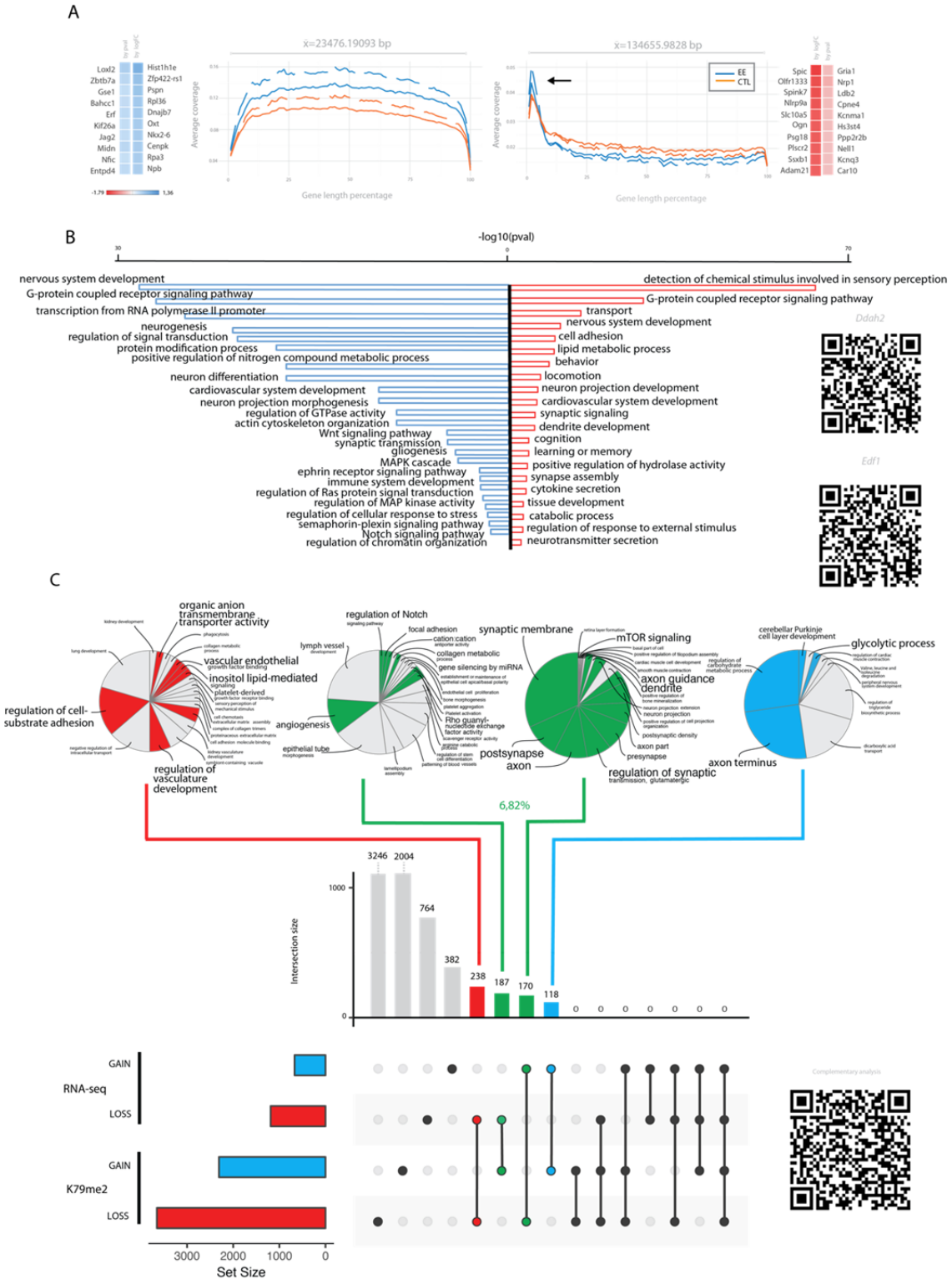


Figure 2.7. Total cortex H3K79me2 overview

A) Gene coverage analysis for both gain and losses together with the top 10 results by fold change or FDR ranking. Top gains on the heatmap are represented in blue and losses in red. Centered plots shows in dark blue EE average coverage for all significant gains (left) or losses (right) and orange for the controls. **B)** Gene ontology analysis for biological process (with term parental fusion and with p<0.01). The graph shows a ranked term by the score of their p-value (–log10(p)). Blue bars (left) represent the gains and in red the losses. Right botton QR code gives access complementary analysis. Right up QR codes gives access to the Human Protein Atlas server information for Ddah2 and Edf1 **C)** Intersection of H3K79me2 data with RNA-seq. Left plot show the set size of each dataset, gains are represented in blue and losses in red. Dots represents the intersections; when a dot is unwired means exclusive number of genes in that data set that are not present in the rest; when wired it means that those datasets shared a particular number

of genes, if both are gained then blue dots, if both are lost then red dots and discrepancies are represented on green. Upper graph represents the number of elements in each intersection with the corresponding color of the contrast direction. Upper pie plot represents the biological process terms belonging to the most important intersections: both increased (blue), both decreased (red), discrepancies (6,82% , green). Highlighted color sections are the elements covered by the text.

The differential analysis reported greater number of H3K79me2 losses (3667) than gains (2315), whereas H3K36me3 shows the opposite trend: more gains (1716) than losses (1286). It seems contradictory that decreased binding of an activating mark could lead to more activity. However in the gene ontology analysis we found that terms enriched in both directions are conceptually related. This contradiction could be explained by the fact the EE produces changes on targets that belong to the same pathway, down-regulating the activity of some of genes meanwhile others need to be up-regulated. Nevertheless we can find some exclusive terms, such as positive regulation of nitrogen compound metabolic process, both in H3K79me2 and H3K36me3 (p=10E-18 and 56E-15, respectively) related with the production of nitric oxide (NO), relevant for synaptic plasticity (Hardingham et al., 2013)(Figure 1.7A and 1.8A). However, it seems (for the current case) that this term is more related to endothelial production as some of the genes found in this cluster are *Ddah2* (that produces inhibitors of the nitric oxide synthase in endothelial cells) (Sakurada et al., 2008) and *Edf1* (that regulates the endothelial nitric oxide synthase activity, (Leidi et al., 2010), that could also licit the term appearance of one of the NO effects that is angiogenesis or cardiovascular system development (H3K79me2 p=120E-12). Despite the fact that NO seems more related to endothelium rather than neuronal production, even if it is not endorsed by the literature, Edf1 and especially Ddah2 are highly present in neuronal cells (human protein atlas, see QR codes, Figure 2.7 B). Also unique for gains, but also exclusive for H3K79me2, are the terms enriched for proliferation, such as neurogenesis (p=7,5E-22) and gliogenesis (p=140E-6), well-known effects of EE (Kempermann et al., 1997; van Praag et al., 1999, 2000)(Figure 2.7 A).

Moreover, regulation of chromatin organization is enriched in both datasets, illustrated by the top H3K79me2 finding of the *Hist1h1e* gene (logFC= 1.36, Figure 1.7A) and the *Hist1h2aa* gene from the H3K36me3 top results (logFC= 1.11, Figure 1.8A). Histone 1 is a protein that functions as a linker between histone nucleosomes, and it has been reported that there are way more dynamic structures than core histones, making them important for chromatin architecture (Misteli et al., 2000).

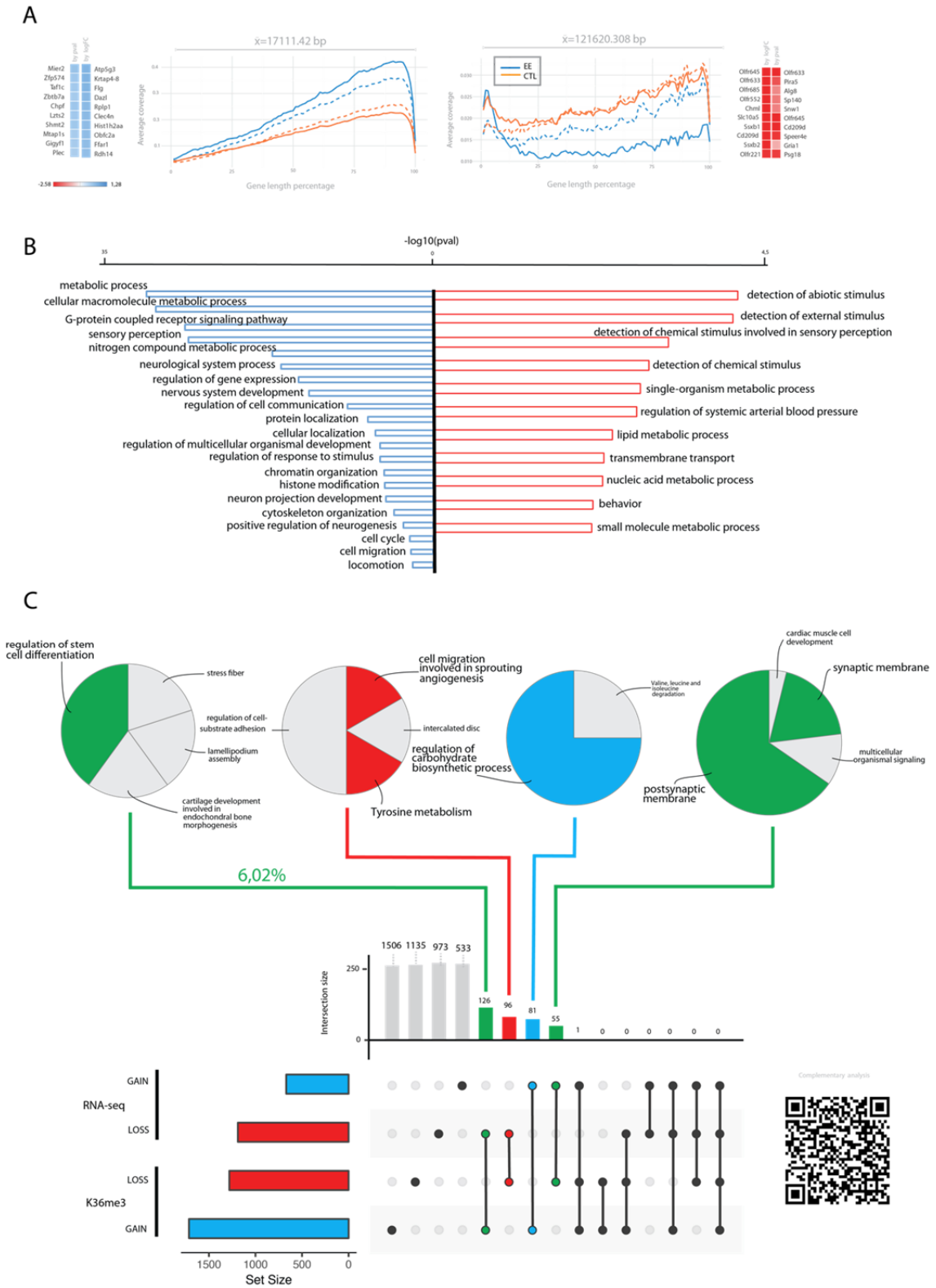


Figure 2.8. Total cortex H3K36me3 overview

A) Gene coverage analysis for both gain and losses together with the top 10 results by fold change or FDR ranking. Top gains on the heatmap are represented in blue and losses in red. Centered graphs shows in dark blue EE average coverage for all significant gains (left) or losses (right) and orange for the controls. **B)** Gene ontology analysis for biological process (with term parental fusion and only showing terms with p<0.01). The graph shows a ranked term by the score of their p-value (–log10(p)). Blue bars (left) represent the gains and red bars (right) the losses. **C)** Intersection of H3K36me3 data with RNA-seq. Left plot show the set size of each dataset, gains are represented in blue and losses in red. Dots represents the intersections; when a dot is unwired means exclusive number of genes in that data set that are not present in the rest; when wired it means that those datasets shared a particular number of genes, if both are gained then blue dots, if both are lost then red dots and discrepancies are represented on green. Upper graph represents the number of elements in each intersection with the corresponding color of the contrast direction. Upper pie plot represents the biological process terms belonging to the most important intersections: both gained (blue), both lost (red),

discrepancies (6,02% green). Highlighted color sections are the elements covered by the text. QR code gives access to complementary analysis

Signaling pathways are also present in both datasets by the gain of activity of G-protein coupled receptors activation, important for synapse function (Gainetdinov et al., 2004). However, H3K79me3 shows more specific terms gained by EE, that are: GTPases activation (Auer et al., 2011), that leads to an enhanced phosphorylation, the Wnt signaling pathway, important for dendritogenesis (Ciani et al., 2011; Gogolla et al., 2009), as well as activation of Ras proteins, the MAPK kinases pathway, the ephrin receptor pathway and the Notch signaling.

Moving to the exclusive terms that are found among genes associated with regions of decreased activity, it is important to highlight the terms of learning (H3K79me2 p= 4.7E-3), cognition (H3K79me2 p= 130E-6) and neurotransmitter release (H3K79me2 p=5.6E-3), specially glutamatergic ion channels such as: *Gria1* (H3K79me2 logFC= -0.554), *Grin2a* (H3K79me2 logFC= -0.329), *Grin2b* (H3K79me2 logFC=-0.426) (see Figure 1.7A). Surprisingly, EE induces a loss of activity of these genes, which is contradictory to previous reports on the field. This might be explained by cell type heterogeneity and tissue composition, as those genes are not exclusive for neurons (Dzamba et al., 2015).

Finally, describing briefly the shared terms for both gains and losses in both datasets, we find neurodevelopmental genes. EE is well known to induce an enhanced neuronal activity and for gaining new synaptic forms and strength (van Praag et al., 2000). We find some neurodevelopmental genes that show increased activity (H3K79me2 p = 54E-30, H3K36me3 p=47E-15) but also genes with decreased activity (H3K79me2 p= 13.0E-9, Figure 1.7A and 1.8A). Interestingly we find a gain of the Wnt signaling pathway (genes such as: Wnt11, Wnt4, Wnt7b, Wnt9a), but in losses a decreased gene activity of BMP, similar to our finding in H3K4me3 promoter activity (genes: Bmp5, Bmpr1b, Bmpr2). One of the genes found clustered in this neurodevelopmental term also appears in the top genes gained in H3K79me2, that is oxytocin (Oxt, logFC=0,83, Figure 2.7 A). Its role has been implicated in sensorial cortical plasticity (Zheng et al., 2014). Moreover it was found that structural neuronal architecture changes enriched by terms such as neuron projection morphogenesis, actin cytoskeleton organization (important for the synaptic volume consolidation), and semaphorin-plexin signaling pathway (important for axogenesis).

2.6.3 Gene body marks and transcription

As shown by the research of Dong et al. H3K79me2 and H3K36me3 gene body marks are the histone modifications with the strongest correlation to transcription (Dong and Weng, 2013). However this study did not explore how changes in those activity marks lead to increased transcription. Our study opens the opportunity to assess the dynamic changes of these epigenetic marks and to analyze how they can explain the differential transcription process due EE.

Starting with the H3K79me2 dataset, the intersection with differentially expressed genes (DEGs, see chapter 3) corresponds to a loss of activity/expression in 238 mutually

shared genes that are enriched for terms linked to angiogenesis but also cell adhesion, inositol lipid mediated signaling and platelet-derived growth factor binding (Figure 2.7 C). Regarding common gains found in both datasets, it identified 118 genes related with carbohydrate metabolic process, important for the generation of energy, as well as cellular components such as axon terminus, and some unexpected terms such as Purkinje cell layer development. Purkinje cell layers are not present in the neocortex but might share some pathways with the neuronal tissue studied here.
The H3K36me3 dataset shows similar common gain and loss terms, pointing at an increased activity of metabolic processes and a decreased activity of genes linked to angiogenesis and tyrosine metabolism (Figure 2.8 C).

Notably, the discrepancies show important terms that drove the complementary experiments we performed in order to better understand the complexity of the data. Especially striking is to find terms linked to learning and memory processes in both H3K79me2 and H3K36me2 datasets, where genes such as the AMPA subunit receptor *Gria1* or the NMDA subunit receptors *Grin2a* and *Grin2b* show a loss of H3K79me2 binding but increased transcription (Figure 2.7 C & Figure 2.8 C).

2.7 Subtracting cellular complexity and de-convoluting total cortex signal contribution

As previous results clearly showed, we observe a discrepancy between gene body histone marks and the corresponding RNA-seq upon treatment, which could be due to the large cell heterogeneity in the complex brain tissue. In order to understand better this phenomenon, we sorted neuronal and non-neuronal cells by using the neuronal marker NeuN+ (Rbfox3)(Matevossian and Akbarian, 2008). H3K79me2 was prioritized, as it was by far the histone that reported the best signal. However it is important to note that the marker NeuN (Rbfox3) is highly expressed in neurons, but to a small extent (~10%) is not present in all neurons. These neurons then fall into the negative population, together with e.g. glial cells (NeuN -) (Mullen et al., 1992; Sarnat et al., 1998). For simplification the NeuN- population will be called non-neuronal, despite the small fractions of contained neurons. To avoid this problem, we also enriched a 4th group of animals that express GFP at exclusively pyramidal neurons (Tg(Thy-YFP)). However even after extensive optimization of the FACS-sorting process, especially the sonication, FACS sorting using Tg caused problems and the data is not available yet. Nevertheless, for chromatin accessibility using ATAC-seq (an experiment discussed in the next chapter), the technique was successful.

2.7.1 Neuronal and non neuronal dynamic coverage changes induced by EE

The coverage changes previously described in the H3K79me2 total cortex dataset are found to be different in the neuronal population. First, looking at the average coverage of the gain-DBS in EE compared to control samples (Figure 2.9.A, left coverage plot), we observe a peak close to the TSS that is common to both treated and non-treated samples. Notably, a second peak that covers the full gene body is found for EE samples. In contrast, control samples simply show a constant decrease until the end of the gene body. Contrary to the gains, losses span for both conditions the full gene length, not showing the first characteristic peak. This change of coverage distribution compared to

the whole-tissue analysis was not found to be highly affected by the gene length, as for the gains the average is 22kb and for the losses 37kb. Interestingly, the non-neuronal population behaves similar to the whole cortex data, showing similar coverage profiles.

2.7.2 Neuronal changes due to EE and their relationship with transcription

Biological process analysis for neuronal populations shows a clear recapitulation of neurodevelopmental genes, illustrated by a gain of neuronal system development (p= 19.0E-27) followed by neuron projection development (p= 1.9E-18). In addition, posttranslational modifications appear to be enriched by the term of protein modification process (p=51E-9). As well as a negative regulation of nitrogen (p=2.4E-3) and an increase in cell adhesion molecules (p=3.0E-3). Concerning neuronal H3K79me2 losses, there is also an enrichment of developmental genes, followed by angiogenesis and term of cardiovascular system development (p= 5.9E-9). Moreover there is a decreased activity of protein translational modification by the citrullination (p= 13.0E-6) and peptydil-tyrosine phosphorylation (p= 260.0E-6). Furthermore it is interesting the appearance of immune related terms such as: cytokine production (p= 2.1E-3) and regulation of T cell differentiation (p= 7.6E-3).

The highest intersection with RNA-seq arises for those genes that are mutually lost. Exactly, 81 genes were lost and terms highlighted are linked to vasculature development (angiogenesis). Additionally, EE seems to induce a decrease of BMP signaling by loss of SMAD phosphorylation. Finally, the respiratory chain mitochondrial machinery seems affected by the loss of transcription and activity of the cytochrome P450.

Making possible to sort neurons and to study their epigenome allows revealing the cellular origin of the neuronal epigenetic and transcriptomic signal. This is obvious by the fact that mutually gained terms for neuronal H3K79me2 and RNA show important synaptic elements that were lost in the total cortex H3K79me2 and H3K36me3 dataset and that were related to memory and learning processes. This observation makes it difficult to come to any trustworthy conclusion about findings of epigenetic changes in whole cortex, as our RNA-seq data, the neuron-specific epigenomic data and the majority of literature point to the contrary (Rampon et al., 2000). This also points to the possibility that this processes have decreased activity in glial cells leading a loss in total cortex dataset. By sorting we were able to find 42 up-regulated genes in both neuronal H3K79me2 and RNA-seq that cluster into synapse function, such as ion sodium and potassium voltage-gated channels, and synaptic glutamatergic transmission being both important for plasticity processes (Voglis and Tavernarakis, 2006).

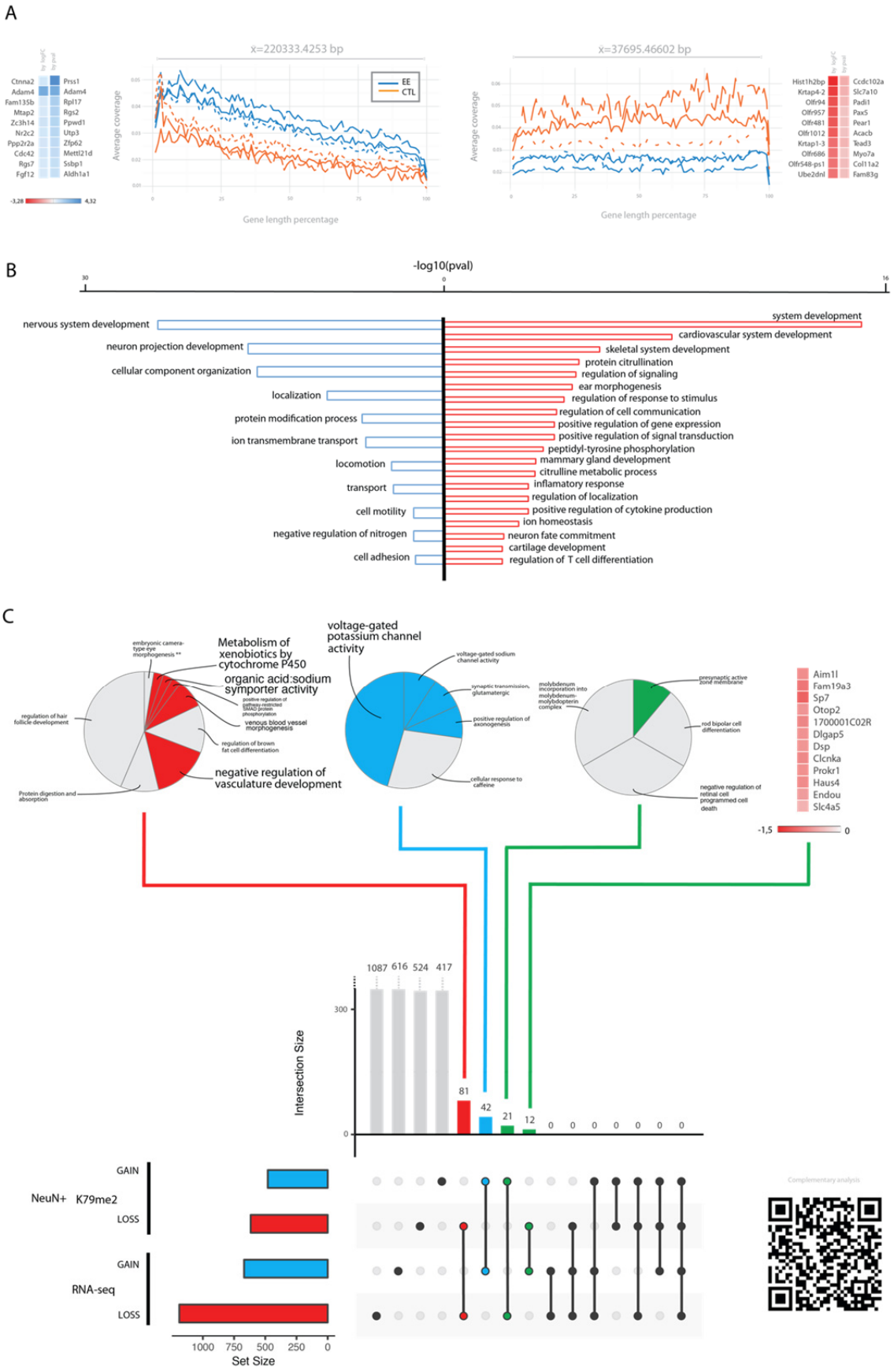


Figure 2.9. Neuronal H3K79me2 overview
A) Gene coverage analysis for both gain and losses together with the top 10 results by fold change or FDR ranking. Top gains on the heatmap are represented in blue and losses in red. Centered graphs shows in dark blue EE average coverage for all significant gains

(left) or losses (right) and orange for the controls. **B)** Gene ontology analysis for biological process (with term parental fusion and only showing terms with $p<0.01$). The graph shows a ranked term by the score of their p-value ($-\log10(p)$). Blue bars (left) represent the gains and in red the losses. QR code gives access to gene ontology analysis performed on cellular components, molecular function and KEGG pathways. **C)** Intersection of neuronal H3K79me2 data with RNA-seq. Left plot show the set size of each dataset, gains are represented in blue and losses in red. Dots represents the intersections; when a dot is unwired means exclusive number of genes in that data set that are not present in the rest; when wired it means that those datasets shared a particular number of genes, if both are gained then blue dots, if both are lost then red dots and discrepancies are represented on green. Upper graph represents the number of elements in each intersection with the corresponding color of the contrast direction. Upper pie plot represents the biological process terms belonging to the most important intersections: both gained (blue), both lost (red), discrepancies (green). Highlighted color sections are the elements covered by the text.

2.7.3 Non neuronal changes due EE and its relationship with RNA

The aim to study this population was to see if the negative population could drive the preliminary results of total cortex, by showing reduced activity for genes that in neurons are expected to be activated by the effect of environmental enrichment. This is indeed what we found, firstly because coverage changes described in 1.4.1 section are similar between whole cortex and NeuN negative population. Also because a big part of the genes of both datasets behave similarly concerning the direction of the fold change, as we will see in next section 1.6.

The Non-neuronal population has the drawback of being heterogeneous mix of glial cells, endothelium and a few neuronal cell types (Mullen et al., 1992; Sarnat et al., 1998). Overall, both gains and losses for this population are semantically highly related making it difficult to establish a direction for a particular biological process.

Briefly, regarding the intersection of DBRs with the DEGs from RNA-seq there are 61 genes mutually showing losses in both experiments that are related with the JAK/STAT signaling pathway illustrated by the term: tyrosine phosphorylation of the STAT protein (Figure 2.10.C). Together with other terms, there is also a loss of response to nicotine and platelet-derived growth factor activation pathway. Finally, 21 genes were mutually gained in both dataset related with the fatty acid biosynthesis by the enriched term: [Acyl-carrier-protein] S-malonyltransferase activity.

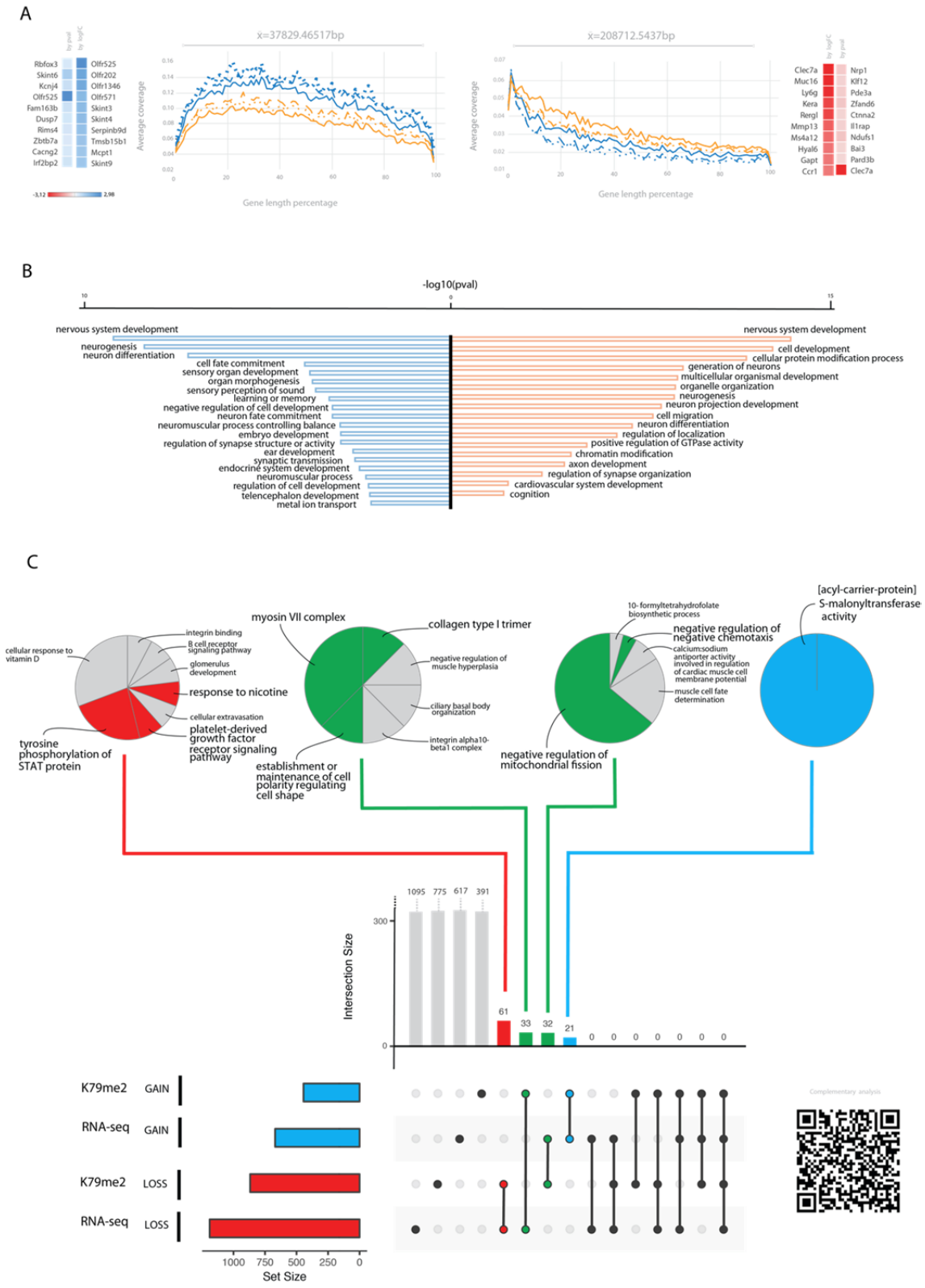


Figure 2.10. Non-neuronal H3K79me2 overview

A) Gene coverage analysis for both gain and losses together with the top 10 results by fold change or FDR ranking. Top gains on the heatmap are represented in blue and losses in red. Centered graphs shows in dark blue EE average coverage for all significant gains (left) or losses (right) and orange for the controls. **B)** Gene ontology analysis for biological process (with term parental fusion and only showing terms with p<0.01). The graph shows a ranked term by the score of their p-value (–log10(p)). Blue bars (left) represent the gains and in red the losses. **C)** Intersection of non-neuronal H3K79me2 data with RNA-seq. Left plot show the set size of each dataset, gains are represented in blue and losses in red. Dots represents the intersections; when a dot is unwired means exclusive number of genes in that data set that are not present in the rest; when wired it means that those datasets shared a particular number of genes, if both are gained then blue dots, if both are lost then red dots and discrepancies are represented on green. Upper graph represents the number of elements in each intersection with the corresponding color of the contrast direction. Upper pie plot

represents the biological process terms belonging to the most important intersections: both gained (blue), both lost (red), discrepancies (green). Highlighted color sections are the elements covered by the text. QR code gives access to gene ontology analysis performed on cellular components, molecular function and KEGG pathways

2.7.4 Early effect of EE

After a week of EE treatment a smaller number of DBS was found compared to the treatment for one month. However, some of the hits show bigger fold changes, especially the gains where the top gene reaches a logFC of 9,38 for to the transcription factor *Znrf4* (Figure 2.11 A).

The gene ontology analysis resulted in a lot of gained terms related with angiogenesis together with other minor concepts (Figure 2.11B). Losses however, show an enrichment of synapse related terms, especially those related to glutamatergic transmission. Notably by intersecting it with the transcriptome data only three genes are shared: *Dnahc6*, *Abca17*, and *Lmx1a* (Figure 2.11C). The rest corresponds to discrepancies: 44 shared genes that are gained in H3K79me2 but lost at the transcriptome level. Terms of interest for those genes are response to magnesium ion, production of energy via glycogen synthesis, cytosketelon by the enrichment of myosin filaments, and finally citrulline metabolic process.

Interestingly the comparison of the early with the late effects could reveal what are the main H3K79me2 DBS genes that change their activity over time (Figure 2.11 C). Notably 61 genes of up-regulated activity at the very beginning of the treatment being later down-regulated show GO terms related to an enhanced calcium ion transport, together with changes in the cytoskeleton components such as the myosin complex. However, immediate losses upon EE exposure trigger a down-regulation of NMDA receptors clustering that later are up-regulated.

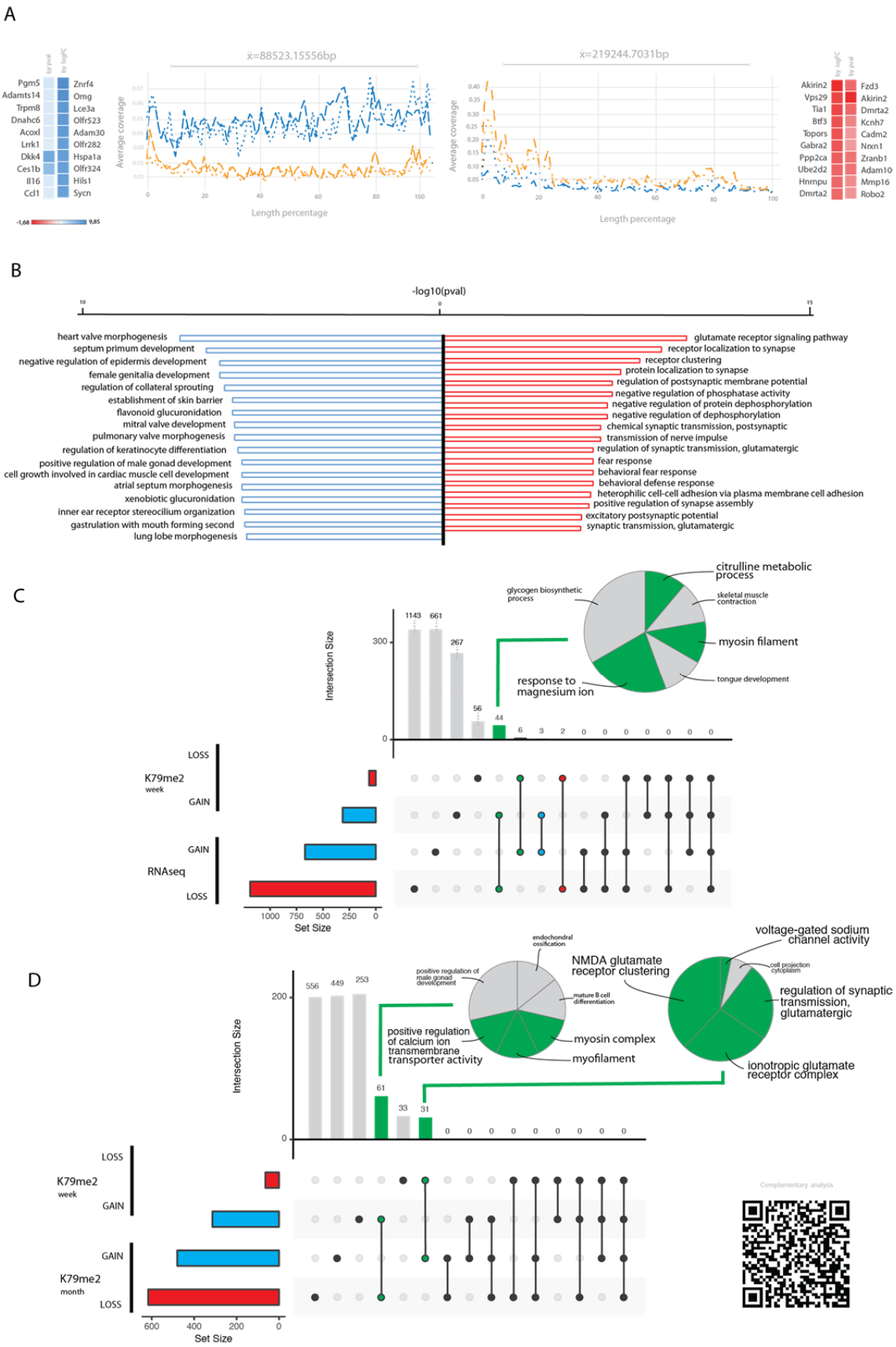


Figure 2.11. Early effects of EE on H3K79me2

A) Gene coverage analysis for both gain and losses together with the top 10 results by fold change or FDR ranking. Top gains on the heatmap are represented in blue and losses in red. Centered graphs shows in dark blue EE average coverage for all significant gains (left) or losses (right) and orange for the controls. **B)** Gene ontology analysis for biological process (with term parental fusion and with $p<0.01$). The graph shows a ranked term by the score of their p-value (–log10(p)). Blue bars (left) represent the gains and in red the losses. QR code gives access to gene ontology analysis performed on cellular components, molecular function and KEGG

pathways. **C)** Intersection of early effect H3K79me2 output data with RNA-seq. Left plot shows the set size of each dataset, gains are represented in blue and losses in red. Dots represent the intersections; when a dot is unwired means exclusive number of genes in that data set that are not present in the rest; when wired it means that those datasets shared a particular number of genes, if both are gained then blue dots, if both are lost then red dots and discrepancies are represented on green. Upper graph represents the number of elements in each intersection with the corresponding color of the contrast direction. Upper pie plot represents the biological process terms belonging to the most important intersections: both gained (blue), both lost (red), discrepancies (green). Highlighted color sections are the elements covered by the text. **D)** Intersection between early and late effects of EE on H3K79me2 output.

2.8 Intersecting epigenetic datasets

To better understand the regulatory dynamics of the broad histone marks and the RNA-seq in a greater perspective, a full intersection was performed with all datasets. As a result, the whole cortex H3K79me2 and H3K36me3 datasets are the most closely related (black arrows, Figure 1.12) sharing 927 up-regulated and 714 down-regulated genes. This constitutes almost half of the total H3K79me2 gains and a third of the losses, making this intersection a set of consistent ChIPseq results for these two histone modifications.

Notably another substantial remark is that the NeuN negative population shares 250 losses with the H3K79me2 total cortex (94 if H3K36me3 is included) and 114 gains for both H3K79me2 and H3K36me3 total cortex dataset (dashed arrows, Figure 1.12). This observation demonstrates that the negative population drives the major changes found at the total cortex.

Moreover, the total cortex ChIP-seq dataset could explain better the changes in the transcriptome, sharing 124 losses (for H3K79me2) and 58 gains (for both H3K79me2 and H3K36me3). But it also exhibits the most discrepancies: 82 genes with increased expression are found to have reduced H3K79me2 occupancy and 81 DEG losses are gains in both histone marks.

Finally, as expected, FAC-sorted neurons are the most discrepant datasets. In addition the amount of total induced differences by EE training in those datasets (1 week and 1 month of EE exposure) is low compared to the rest of the datasets. This makes an interesting observation that many of the changes triggered by the treatment might impact strongly in other cell types more than neurons.

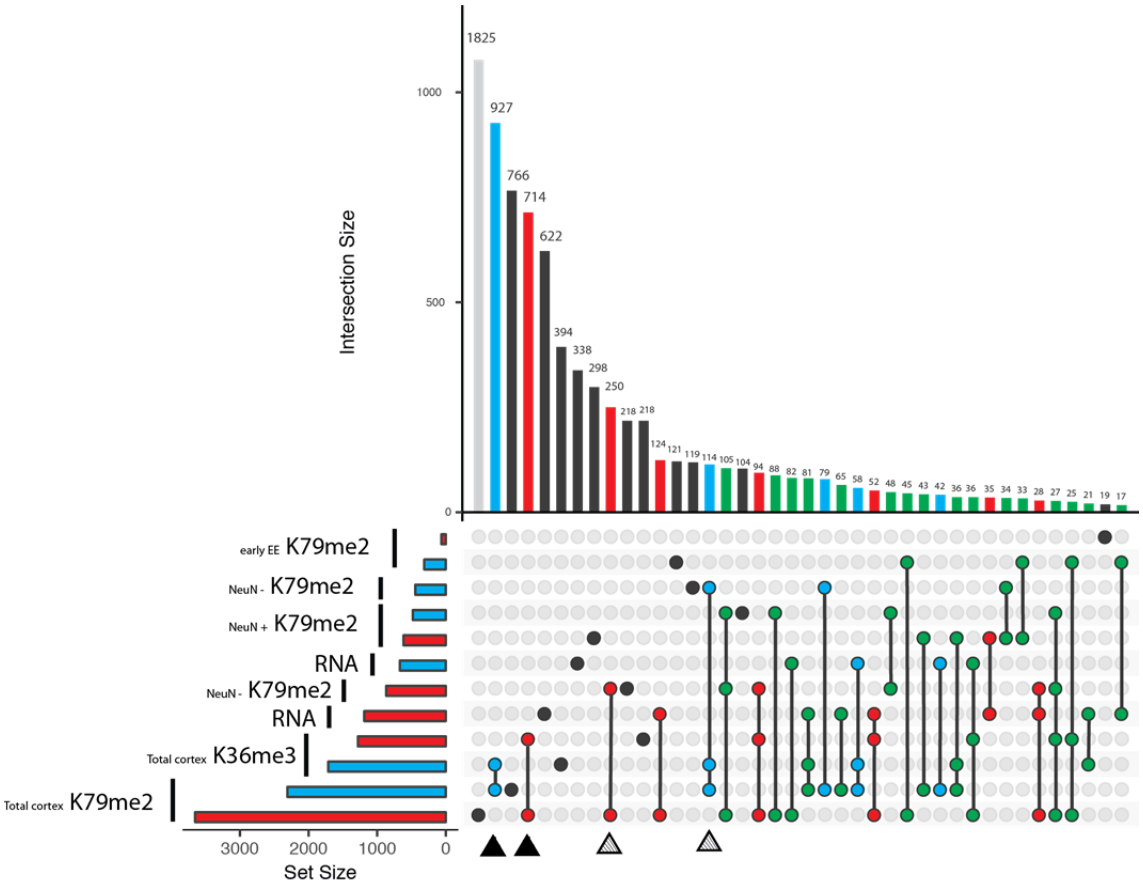


Figure 2.12. Broad marks intersection with RNA
The intersection analysis ranks the different datasets by their size starting to intersect the biggest with the rest of sets. Left bar plot the size of each DBS set showing in light blue, the gains, and in red the losses for. Upper bar plot shows in black bars the unique number of genes that are in each set not shared with the rest (unique elements), together with blue if both data sets linked by a wire show the same direction on gains, red on losses and finally green for shared genes discrepancies.

2.9 Discussion and concluding remarks

Previous analysis provided a broad overview of the main GO terms that change activity upon EE treatment, both with respect to the epigenome and the transcriptome. The study does not provide conclusive information about the genes that contribute in each term. But, being the discovery phase of the study, it allowed us identify and select the most interesting and promising gene groups to be analyzed in future studies. In this last section I will provide specific information of groups of genes that previously appeared to give a better explanation of the biological relevant changes that might occur.

As previously introduced EE exposure triggers an enhanced neuronal activity that will trigger learning and memory processes to be consolidated. In order to better understand the epigenetic changes elicited by EE of genes involved in synaptic development we selected for deeper analysis groups of genes such as synaptic ionic channels, receptors and pathways. For simplification, fold changes have been sorted by NeuN+ FDR.

The first observation looking at the heatmaps is that total cortex data results from H3K79me2, H3K36me3 and non-neuronal population (NeuN-) are correlated (Figure 2.13).

As illustrated in Figure 2.13, the top ionic channel down-regulated in neuronal cells is the ionotropic glutamate receptor *Grin2c* that is mainly enriched in the cerebellum. It is similarly found down-regulated in the expression analysis. We previously saw that learning and memory GO terms had NMDA receptors that were both lost in the whole cortex gene-body histones, representing a contradiction with transcriptome data. Interestingly in the neuronal population (NeuN+) those genes are up-regulated, showing the same direction as the DEGs and thus agree with previous publications (Andin et al., 2007; Hu et al., 2013; Melendez et al., 2004; Naka et al., 2005; Rampon et al., 2000). Indeed, concerning LTP process NeuN+ H3K79me2 shows an increase of coverage in AMPA receptors *Gria2* and *Gria4*, endorsed by published transcriptomic changes due EE (Naka et al., 2005). These results could point to potential increase of AMPA receptor transcripts to be translated at the synapse and to be available for insertion at the synaptic membrane (Luscher and Malenka, 2012). Moreover, other interesting genes up-regulated in neuronal population are: the kainite ionic receptor *Grik2*, the voltage calcium receptor and the potassium channels *Kcnma1*, *Kcnmb3* and *Kcnn2*. Regarding the inhibitory contribution changes induced by EE, there is an increase of activity of GABA receptors: *Gabra2*, *Gabrb1*, *Gabrb2*, *Gabrb3* and *Gabrg3*. Moreover, concerning the early changes of EE it seems important to highlight the novelty of an increased activity of the nicotinic acetylcholine receptor *Chrna2*, previously never reported as an effect of this factor. For instance nicotin acetilcholine receptors might play an important role in facilitating plasticity both at presynaptic and postsynaptic level (Ji et al., 2001) but decrease of acetylcholine release has been also reported after EE training(Del Arco et al., 2007a; Del Arco et al., 2007b; Ji et al., 2001). Furthermore early effects of EE seem to also enhance the activity of the potassium ion channel *Kcnn4*. The rest of the significant changes fall into a decrease of activity of the kainate receptor *Grik2* (that show up-regulation after a month of EE), the AMPA receptor *Gria2* and the GABA ion receptor *Gabra2*.

Concerning the receptors there is considerable down-regulation in the neuronal population of the neurotrophic tyrosine kinase-1 receptor, *Ntkr1*, together with the beta-adrenergic receptor 2b, *Adra2b*. However there is an increased activity of the metabotropic glutamate receptor *Grm1* and *Grm7*. The latter is significantly down-regulated during the early effects of EE.

Ranking the pathway contribution changes due to the treatment, there is an obvious enrichment of Wnt signaling. Firstly by the up-regulation of the alpha-2-catenin: *Ctnna2*, together with a loss of Wnt ligand gene activity such as: *Wnt9b*, *Wnt3a*, *Wnt3*, *Wnt11* and *Wnt4*. However, there are no significant changes observed on Frizzled receptors (Wnt ligand receptors). Regarding the relationship with other developmental pathways we also observe the up-regulation of BMP signaling genes such as the receptors *Bmpr2* and *Bmpr1a*. Finally, there is also some Hedgehog signaling linked pathway, e.g. the decreased activity of *Smo*, a transmembrane receptor for signaling Shh transduction (Varjosalo and Taipale, 2008).

Another interesting contributor is the kinase activity appearing by the MAPK activity up-regulation of *Mapk1* and *Map2k4* together with a significant decreased activity of *Map3k6*. The MAPK pathway has been reported to be necessary for NMDA receptor-dependent LTP (Atkins et al., 1998; English and Sweatt, 1996, 1997). MAPK/ERK pathway is especially interesting because it can act as a biochemical evaluation for the neuron to trigger lasting changes of synaptic strength (Sweatt, 2001). It is also important

to highlight the up-regulation of *Cdc42* as well as the *CamKIV*. The latter has been shown to be important for plasticity (Jain and Bhalla, 2014), inducing the exposure of AMPA receptors at the synaptic cleft. In addition the JAK/STAT pathway seems to be down-regulated indicated by decreased H3K79me2 binding in the *Jak3* and *Stat5* genes. Concerning the Ras superfamily: the Rab family shows a decrease of *Rab37* and an increase of *Rab11a*, *Rab18* and *Rab10*. Moreover the only significant contribution from the Ras family belongs to *Rem1* that is down-regulated. Finally the Rho family shows an increase of *Rhobtb3* activity.

In the case of early EE effects, there is a clear enrichment again of Wnt signaling, this time showing the top significant result as a decrease activity of the Frizzled receptor *Flz3*, but and increase of the ligand *Wnt3* gene. Concerning the losses there is a decrease of BMP signaling by a decrease activity of the *Bmpr2* gene. In addition a loss of activity of the Rab family *Rab14* and a loss MAPK pathway by the decreased activity of *Mapk1*.

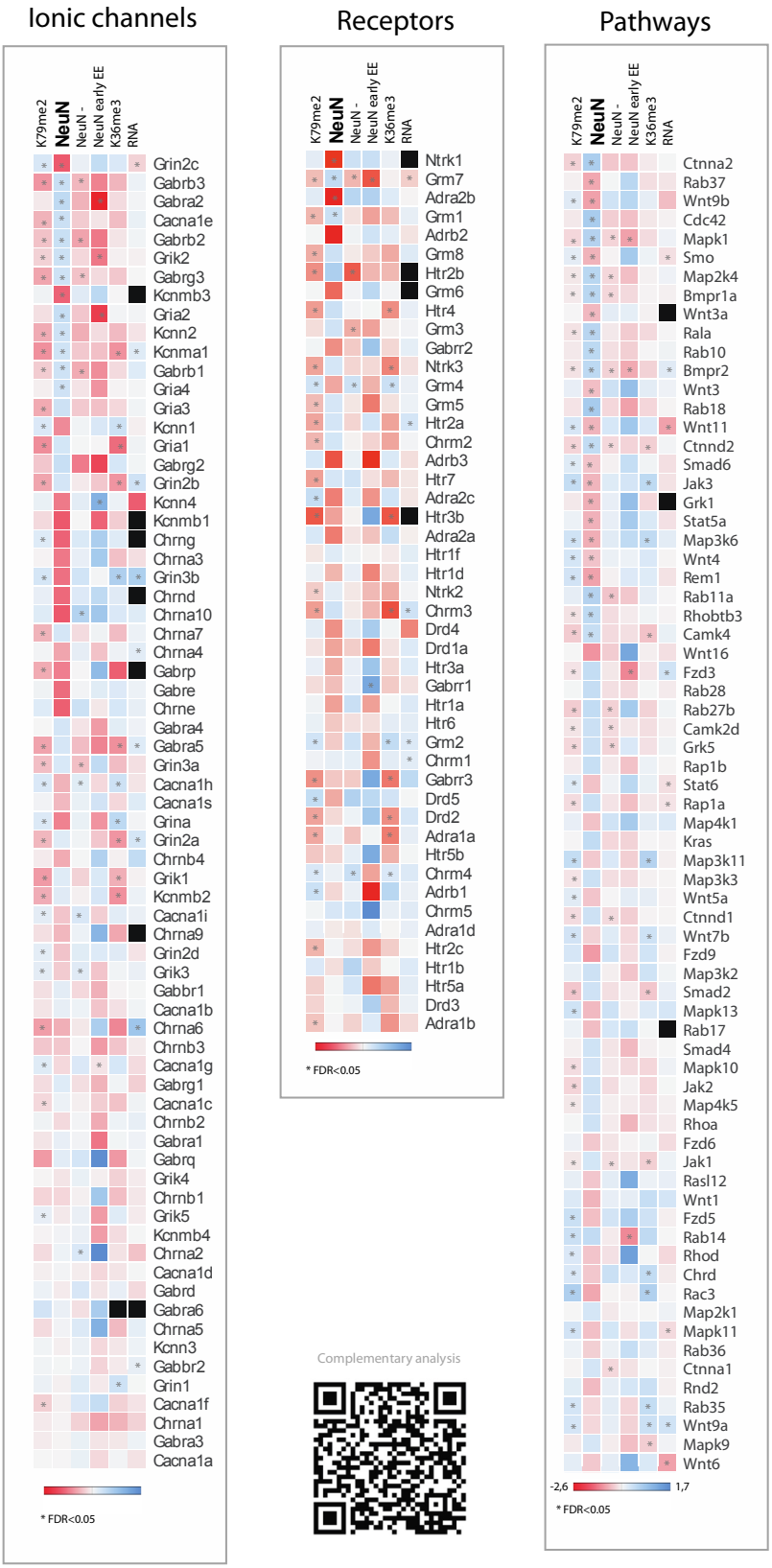


Figure 2.13. Synaptic function heatmap: ionic channels, receptors and signaling pathways
Heatmaps showing the contribution per gene of each broad mark dataset and RNA-seq, the heatmap is sorted by the best FDR value of the neuronal population, showing the significant changes with an asterisk (*). QR sends to complementary analysis to this section.

2.9.1 Concluding remarks

Here we reported the epigenomic changes induced by EE, which we have studied both in whole-cortex tissue as well as in FAC-sorted neuron and non-neuron populations.
Two distinct bioinformatics approaches were used to identify regions of differential binding from ChIP-seq data. One approach was optimized for finding DBS in broad marks. It uses a peak-independent sample comparison strategy avoiding heavy fragmentation of peak and DBS regions in large genes. On other hand it is obvious that this strategy removes the possibility to find DBS outside the selected regions. Therefore, we are currently working on another peak-independent methods using window-based approach as implemented in CSAW (Lun and Smyth, 2016). Nonetheless window-based approaches come with their own issues, e.g. they might misinterpret shifts in coverage distribution across a gene as DBS if the window size is small. Furthermore our study reveal that narrow peak datasets show high read redundancy resulting in a reduced number of informative reads and hence reduced significance of DBS.

The analysis of the broad histone marks revealed some interesting dynamics especially for H3K9me2 that could highlight the function of this mark and its relationship with RNA pol II speed of transcription. This might explain in part the discrepancies found between the epigenetic marks and the transcriptome, as different dynamics will be expected. A simple way to study this is to use gene body marks as a factory and the transcriptome as the resulting product. The speed of production does not follow always a direct proportion of change. Different signals could trigger a higher production that is not directly translated into more products at the very beginning of the enhanced activity of the factory. On top of that, an important factor to the final product is the RNA post-processing and degradation rate. Highlighting that different dynamics mix in the system. Simplifying, the dynamics determining the different datasets cannot be explained by a frozen picture in a particular time point. Understanding these dynamics is extremely important for comprehensive view of the mouse cortex regulome. They could explain the full perspective of the mechanistic molecular changes of the experience-dependent neuronal activity.

FAC-sorting allowed to de-convolute epigenetic signals of different cell types revealing their contribution to changes in the epigenomic landscape. We have demonstrated that the NeuN positive population shows a different epigenetic response to EE exposure, than the NeuN negative and the whole cortex populations, which are more similar to each other. Nonetheless, NeuN markers do not resolves the different contributions of each neuronal type, as even neurons are a heterogeneous population. The remaining challenge is to sort the different types of neurons in the brain, while obtaining enough material for sequencing methods like ChIP-seq, RNA-seq and ATAC-seq.

Chapter 3 - Preface of ARTICLE 2

Contributions

My contribution consisted in the experimental optimization of sorting neuronal nuclei (NeuN+) and pyramidal cells (GFP+ of Tg(Thy(EGFP)) followed by ATAC-seq library preparation. I took care of the animals and I have performed the dissection for those datasets myself. Moreover I have performed the mapping, the computational benchmarking of the differential gene accessibility analysis, the motif discovery analysis, the transcription factor prediction and the gene ontology analysis.

Charlotte Hor optimized and performed the libraries for the whole cortex. Currently she is working in complementary studies using the same datasets that will ultimately be fused with the current study for publication.

Shalu Jhanwar developed the enhancer tool prediction.

Júlia Albaigues help me with the new breedings and some of the dissections.

Michael Maher helped me during the long days of sorting and performed specifically some steps of the library prep for pyramidal neurons.

Meritxel Pons breed the first group of animals and the whole cortex dataset dissections.

Stephan Ossowski performed the mapping bechmarking of the first ATAC-seq libraries and revised the present work.

Neuronal sorting took place at the FAC-sorting Facility leaded by Oscar Fornás at the UPF-CRG. Finally, sequencing was performed at the Genomics Unit leaded by Jochen Hecht at the CRG

Objectives

The main objectives of the present study are:

- To study open-chromatin changes due to environmental enrichment in regulatory regions and characterize their gene-associated function in the brain.
- To analyze transcription factor motif enrichment in differential accessible regions due to the treatment.
- Finally determine if changes in gene accessibility could induce changes at the RNA level expression.

Highlights

- Environmental enrichment induces a general gain of accessibility in active promoters and enhancers of learning-responsive genes in mouse cerebral cortex.
- Motif discovery reported important early expressed transcription factors that are upregulated due to environmental enrichment playing important roles in the molecular mechanism involved in the up-regulation of learning-associated genes.

Acronyms

Differential accessibility sites (DAS), Transcription starting site (TSS), Environmental enrichment (EE), Assay for Transposase-Accessible Chromatin (ATAC-seq), Posttranslational modifications (PTMs), Green Fluorescent Protein (GFP), Gene Ontology (GO), Transcription Factor (TF), iRegulon motif enrichment score (NES).

Chapter 3
Environmental enrichment induces a general gain in promoter and enhancer accessibility in mouse cerebral cortex

3.1 Abstract

Chromatin accessibility is dynamically adjusted in a cell-type and tissue specific manor to allow for transcription factor binding and gene expression regulation. Changes in chromatin accessibility, also called open chromatin, can be induced by signals determined by environmental factors. Chromatin modifiers can influence the access of transcription factors (TFs) to promoters and enhancers by the re-positioning, disassembly or reassembly of nucleosomes. There is evidence that several chromatin remodelers can be recruited upon exposure to environmental conditions and regulate learning-responsive genes. Here we investigate if exposure to environmental enrichment (EE) can trigger chromatin accessibility changes in the mouse cerebral cortex. To better understand this key process in neuronal gene regulation we followed a top-down strategy by using ATAC-seq to study the accessibility of whole mouse cortex, populations of all neurons and purely pyramidal cells. For the first time the current study provides insight into neuronal activity-dependent global chromatin accessibility changes induced by EE in active promoters and enhancers.

3.2 Introduction

Mammalian transcription is a highly complex process that controls fundamental aspects of cell diversity and organismal adaptation. Neurons, as a particular model of study, exhibit remarkable specialization and plasticity, which is mediated, in part, by activity-dependent changes in gene expression (Greer and Greenberg, 2008). We have known for several decades that learning and memory formation requires gene expression. To launch these gene expression programs following learning requires access to DNA for transcription factors to bind (Alberini, 2009; Barrett and Wood, 2008). We can define gene accessibility as opening regions of the chromatin that allows DNA-interacting proteins to bind in regulatory regions and to establish their regulatory mechanisms. These mechanisms include promoting transcription (transcription factors) and induction of posttranslational modification (PTM), such as DNA methylation (Tsompana and Buck, 2014).

Learning, memory and plasticity are boosted by sustained exposure of rodents to environmental enrichment (EE). EE promotes various plasticity mechanisms in the hippocampus, including long-term potentiation, *Bdnf* gene up-regulation, enhanced dendritic branching, and stimulation of adult neurogenesis (Nithianantharajah and

Hannan, 2006; Novkovic et al., 2015; Simpson and Kelly, 2011; van Praag et al., 2000). An important step towards understanding the beneficial effects of EE is to fully understand the entire molecular events that take place. These include gene regulation changes, of which chromatin accessibility in promoters and enhancers is a key part.

We hypothesized that by knowing the expression changes induced by EE (Rampon et al., 2000), related regulatory regions need to be more open in order for DNA-binding proteins (transcription factors) to bind and prompt their action. The majority of chromatin accessibility studies have focused on its use for genome annotation, TF footprints and TF binding site (TFBS) prediction and its potential relationship with transcription. Only few studies have attempted to identify changes in chromatin accessibility induced by environmental changes, however, some examples do exist. For instance, the study describing single-cell ATAC-seq technique investigated variability in chromatin accessibility induced by different drug treatments in eukaryotic cell lines (Buenrostro et al., 2015). Other examples look into accessibility changes at different time-points of organogenesis. For instance, the research of Pajoro et al. demonstrated changes in accessibility in plant development during different time-points (Pajoro et al., 2014). (Stiles and Jernigan, 2010)

Here, we provide for the first time a full parse of how EE induces differential accessibility changes using ATAC-seq, that consists in the fragmentation and tagging with sequencing adaptors of open regions by the transposase Tn5 (Buenrostro et al., 2013) in a top-down strategy of stratified data belonging to three different cell populations: whole cortex, total number of sorted neurons (NeuN+) and a cell-type specific population of sorted pyramidal neurons (YFP+ neurons of Tg((Thy-YFP)). Our study shows that EE induces major gains in chromatin accessibility in regulatory regions for the three datasets despite a negligible number of accessibility losses. Importantly, regions that gain accessibility are associated with learning-response genes.

3.3 Workflow and bioinformatics pipeline

For the present study we used the Assay for Transposase-Accessible Chromatin or ATAC-seq to generate genome wide chromatin accessibility maps (Buenrostro et al., 2013). To fully understand how chromatin accessibility might change upon EE exposure we generated ATAC-seq data of mixed cell populations from the entire cerebral cortex (corresponding to 50k nuclei), the complete cortical neuronal population (by sorting 200k NeuN+ nuclei per library preparation) and all cortical pyramidal neurons (by sorting 100k GFP positive cells per library preparation in Tg(Thy-YFP)) (Figure 3.1 A)

To study how the accessibility might change during cognitive stimulation provided by EE we used a peak-calling based methodology. The idea behind ATAC-seq is that open chromatin allows the transposase to integrate together with the sequencing adapters. Hence, highly accessible regions are expected to show a higher number of reads. However, the distribution across fragmented sites can be uneven and is influenced by off-target reads (Buenrostro et al., 2013). For that reason we employed a density estimator tool called fseq to identify and quantify peaks (Boyle et al., 2008) as well as a second tool developed in-house for detecting active enhancers (Jhanwar et al. *in preparation*). To perform the differential analysis between treated and untreated samples we used DiffBind (Stark & Brown, 2011). We benchmarked the analysis using different background inputs: ChIPseq input (sonicated chromatin fragments without

protein enrichment followed by sequencing), ATAC-seq noise background ('tranposased' naked DNA followed by sequencing) and no background control (Figure 3.1 B). The best results were obtained by using the ChIPseq background, the other noise datasets did not report any result. Thus, ATAC-seq method is fast and accurate, but it is sensitive to technical biases introduced by the chromatin extraction protocol. Here we demonstrated that by providing ChIP-seq input background we could get results despite potential batch effects being found. The reason behind this effect may be related to 1) a sonication bias towards fragmentation points between nucleosomes, thus delimiting the positions of the nucleosomes, or 2) the removal of potential amplification bias or sequencing artifacts when providing this background rather than others. Further study of this effect needs to be performed in order to fully understand the phenomenon.

To demonstrate neuronal cell specificity, our data was compared to other tissues from ENCODE Consortium datasets (www.encodeproject.org). For instance pyramidal neurons do not show chromatin accessibility in genes that are typically not expressed in this cell population (Figure 3.1C) such as the GABAergic receptors (Figure 3.1C). These receptors are highly expressed in interneurons included in the NeuN+ and whole cortex population, but not in the excitatory neurons such as pyramidal cells.

For a comprehensive description of the experimental workflow and computational analysis, please see the Methods Chapter.

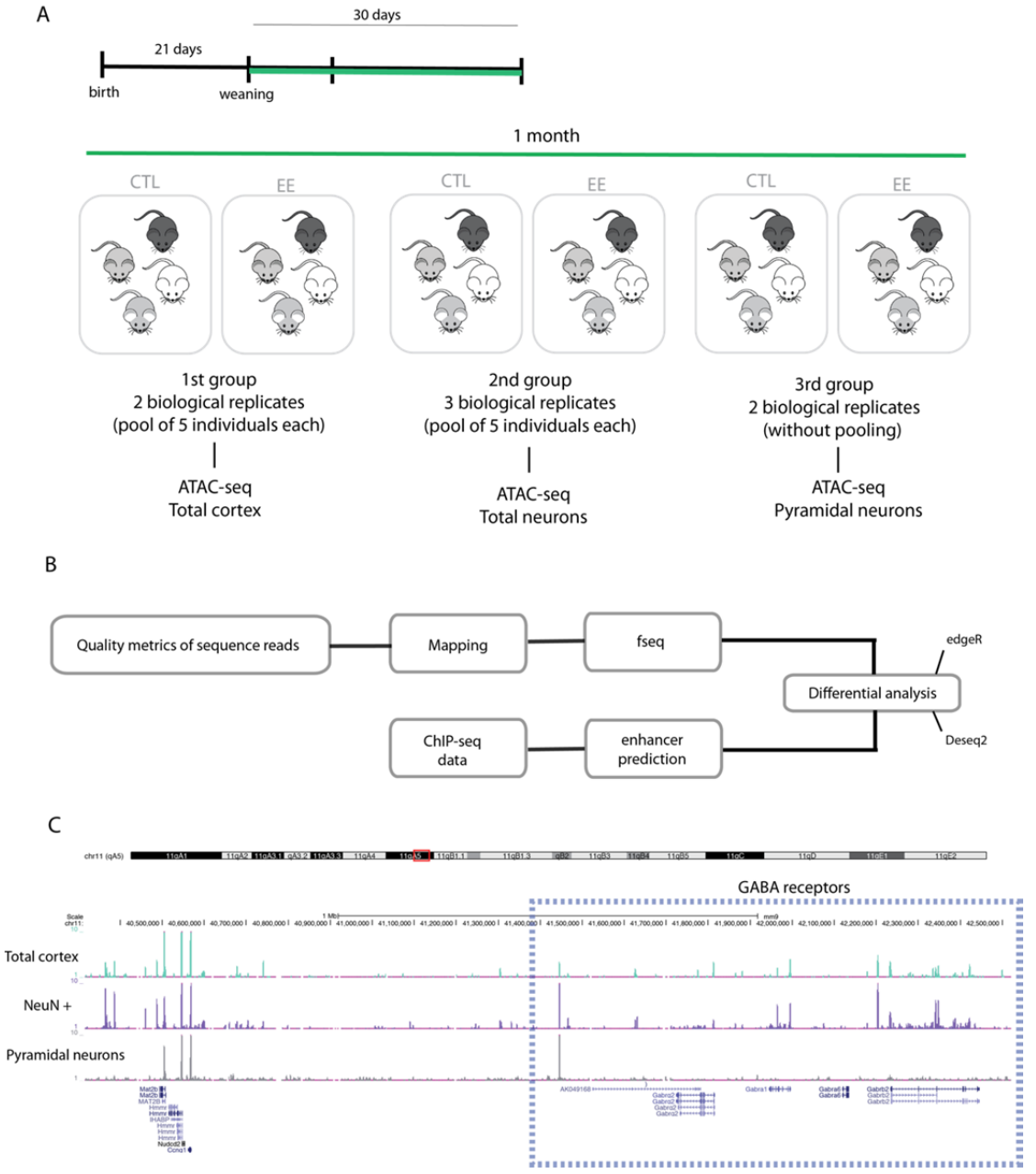


Figure 3.1. Accessibility study overview
A) Experimental workflow. Three different groups of animals were bred. First, two biological cortical replicates of five different individuals were pooled per condition and homogenized followed by nuclei preparation and ATAC-seq. Second, three biological cortical replicates of five different individuals were pooled per experimental condition and homogenized followed by FAC-sorting procedure (see methods for more details). Nuclei tagged with the neuronal marker NeuN (Rbfox3) were sorted to a final yield of 200,000 nuclei per ATAC-seq library. Finally a third group corresponding to two biological replicates per condition of transgenic mice expressing EGFP under control of the pyramidal promoter Thy-1 (Tg(Thy-EGFP)) were sorted to a final yield of 100,000 per ATAC-seq library. **B)** Computational pipeline. After sequencing, quality metrics were performed followed by mapping and peak calling by fseq. For both peak annotations: fseq and enhancer prediction set; a differential analysis was performed by DiffBind using both edgeR and Deseq2. **C)** USCS genome browser examples of ATAC-seq libraries for each group. This tracks clearly illustrates how pyramidal neurons do not show accessibility for GABAergic receptors *Gabrq2*, *Gabra6*, *Gabrab2* as expected (blue box).

3.4 Results and discussion

3.4.1 Environmental enrichment induces a general gain of chromatin accessibility in promoter regions showing highest accessibility changes in learning-associated genes, Wnt and BMP signals

EE is an environmental factor capable of inducing changes in gene expression, both at early and late exposures to the treatment (Rampon et al., 2000). This enhanced transcription activity needs the action of molecular machinery effectors that need to get access to open-chromatin regions to exert its action. As chromatin accessibility is a dynamic process where regions can act as gates, i.e. whereby they can be partially or fully open or closed for TF binding and transcription initiation, our first hypothesis was to check how these gates are affected by EE. It was found that EE induced a global gain of chromatin accessibility at the vicinity of the Transcription Start Site (TSS) for all three datasets: whole cortex, total cortical onset of neurons and cortical pyramidal cells (Figure 3.2). These up-regulated differential accessible sites (DAS) are enriched around 5 kb downstream of the TSS, corresponding to the close vicinity of promoter regions. While this observation could be influenced in part by the peak caller algorithm used in the analysis, it is important to note that the majority of open-chromatin regions in the genome are found at promoter or enhancer sites which are depleted of nucleosomes (Schones et al., 2008; Tsompana and Buck, 2014). Despite the accessibility bias towards DAS gains, we can also find some losses, which are low in number and show a random distance distribution with regard to the closest TSS compared to the gains, which enrich close to TSS (Figure 3.2).

Interestingly the DAS that we find by the analysis are already in active regions and also show coverage in control samples (Figure 2.4). This highlights the fact that EE is not acting to promote the opening of formerly inactive genomic regions, but rather to enhance the accessibility of the already active ones. This could point to a bigger amount of cells gaining accessibility for a particular region, the opening of the second allele or a higher likelihood of being open over time.

We found a total number of 961 DAS in whole cortex where activity in 885 DAS is increased and in 76 DAS is decreased by the effect of EE. Meanwhile the total set of neurons shows 223 DAS where 205 DAS show increased and 18 show decreased activity. Finally 397 DAS are found in pyramidal neurons of which 376 DAS are increased and 21 are decreased (Figure 3.3). These results show a tendency of finding fewer when narrowing the cell-specific population datasets, indicating that the analysis in whole cortex identifies changes in different cell-types.

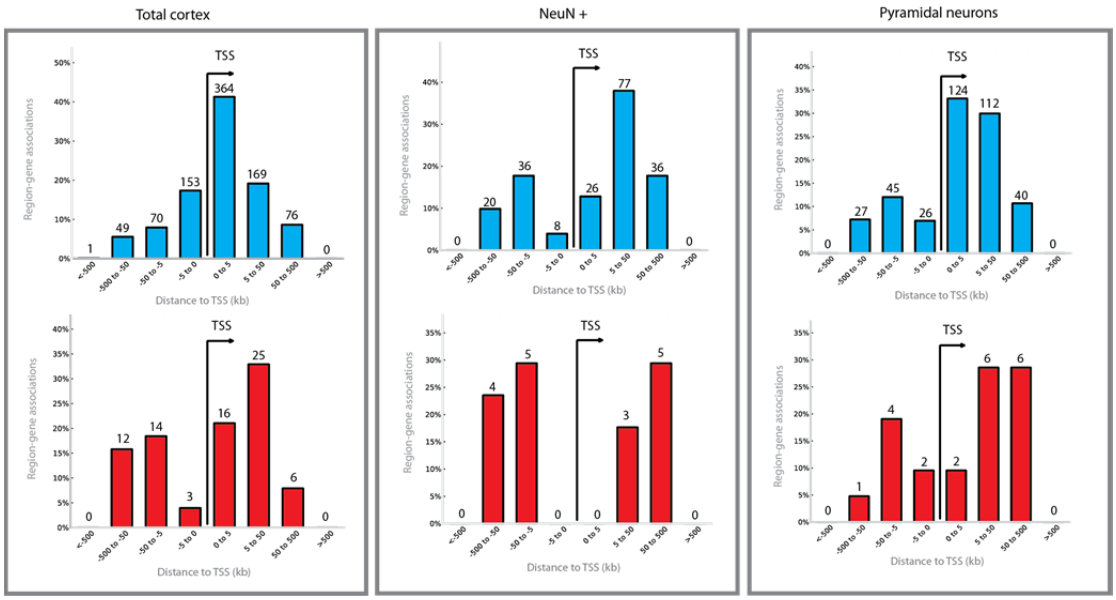


Figure 3.2. Promoter increased activity induced by EE
EE induces a general gain of activity on regions close to the TSS (blue bars). The effect is more prominent in total and pyramidal neurons. Losses (in red) are less in number and present a more spread distribution around the TSS.

Different strategies can be used to understand how a treatment triggers a particular mode of action in different targets. Next Generation Sequencing techniques allow to obtain an enormous amount of data that can be difficult to work with. Finding ways to compare and interpret different datasets from which to extract useful information is of vital importance. Having hundreds of DAS-associated genes, a first attempt to understand the main changes is to look into the biggest and the most significant changes induced by EE. Bearing in mind the assumption that DAS are *cis*-acting regions on their closest gene, we employed a strategy using the GREAT tool, which plotted the top 10 genes associated with each dataset. This tool associates any annotated regulatory region to its target gene (in contrast to the association rule followed for fseq dataset where the association is made with the single nearest gene within 500 kb of the DAS, http://bejerano.stanford.edu/great/public/html/,(McLean et al., 2010). Using the GREAT methodology, it was found that some of the strongest and most significant changes are linked to genes related with synaptic strength, function and plasticity. For example, whole cortex shows increased accessibility in the promoter regions of *Camk4* and *Grin2a*, which are cornerstones of synaptic plasticity (Figure 3.3 and 3.4 A)(Ho et al., 2000; Takao et al., 2010; Voglis and Tavernarakis, 2006). At the same time, despite the low number of losses, there is a decrease in accessibility in an intronic region of the GABAergic receptor *Gabbr2*. This loss of gene accessibility in this region could be associated with the discovery that EE induces decreased GABA ligand expression (Begenisic et al., 2011). Another EE-induced chromatin modification was the loss of accessibility in an intronic DAS of the demethylase *Dnmt3a*, that also plays an important role in maintaining synaptic plasticity (Feng et al., 2010; LaPlant et al., 2010).

Cortical neurons (NeuN+) also express relevant DAS-associated genes such as *Wnt3a*, involved in axogenesis (Purro et al., 2008), the potassium ion channel *Kcnn4*, implied in synaptic function (Figure 3.4B) (Voglis and Tavernarakis, 2006) and *Bmp7*, important for dendritogenesis (Withers et al., 2000). *Wnt3a* together with *Bmp7* and *Tmem106b*, *Foxd1*, and *Gdpd5* participate in neuron projection development as discussed bellow ($p_{\text{NeuN+}}$ = 4,5E-3, Figure 3.6). Of note among the losses is the kinesin *Kif26b* that

transports proteins such as Ab-1, which are important for cytoskeletal dynamics at the synaptic buttons (Heinrich et al., 2012).

Continuing with the main changes in pyramidal neurons, there is a substantial increase in accessibility in Wnt signaling contributors such as *Wnt5b* (Figure 3.3 and 3.4C). Moreover, *Tax1bp1* shows an increase in promoter accessibility and interestingly it has been shown to be exclusively present in synapses and to interact with the proteasome, playing potential roles in protein degradation (Tai et al., 2010) Also appealing is the finding of a DAS-associated with the *Fras1* gene. This is a transmembrane protein that is important in morphogenesis. *Fras1* interacts with the glutamate receptor interacting protein Grip1, potentially playing an important role in AMPA receptor clustering at the membrane (Takamiya et al., 2004). Concerning the DAS losses, one of the top hits found corresponds to an intronic region of the *Smo* gene (Figure 3.4 C), highlighting the possible role of Shh signaling pathway in neuronal architecture changes. In this case *Smo* is significantly down-regulated at the mRNA level (logFC= -0.21, $p = 0.036$). Also of interest was the appearance of a DAS associated with the down-regulated *Nfatc2*. This gene induces the expression of a transcription factor regulated by calcineurin which after depolarization and activation of the voltage-gated calcium receptors can trigger the transcription of early response genes such as *Fos* (Freeman et al., 2011; West and Greenberg, 2011).

Despite previous results it is important to be cautious with the DAS that show lower accessibility due to EE as the association rule that is followed could elicit false positive associations, both for gains and losses. But as losses are more enriched in regions that are farther away from the TSS, there may be enhancer regions that are difficult to be associated if no chromatin conformation data is available. In addition the read coverage for these regions is more spurious compared to gains, supporting the idea that losses are enrich for false positive peak predictions or false associations. For instance the DAS loss discovered for the *Gabbr2* gene seems to play an important role regulating the promoter of the close *Tbc1d2* gene, as observed by preliminary 3C experiments (data not shown).

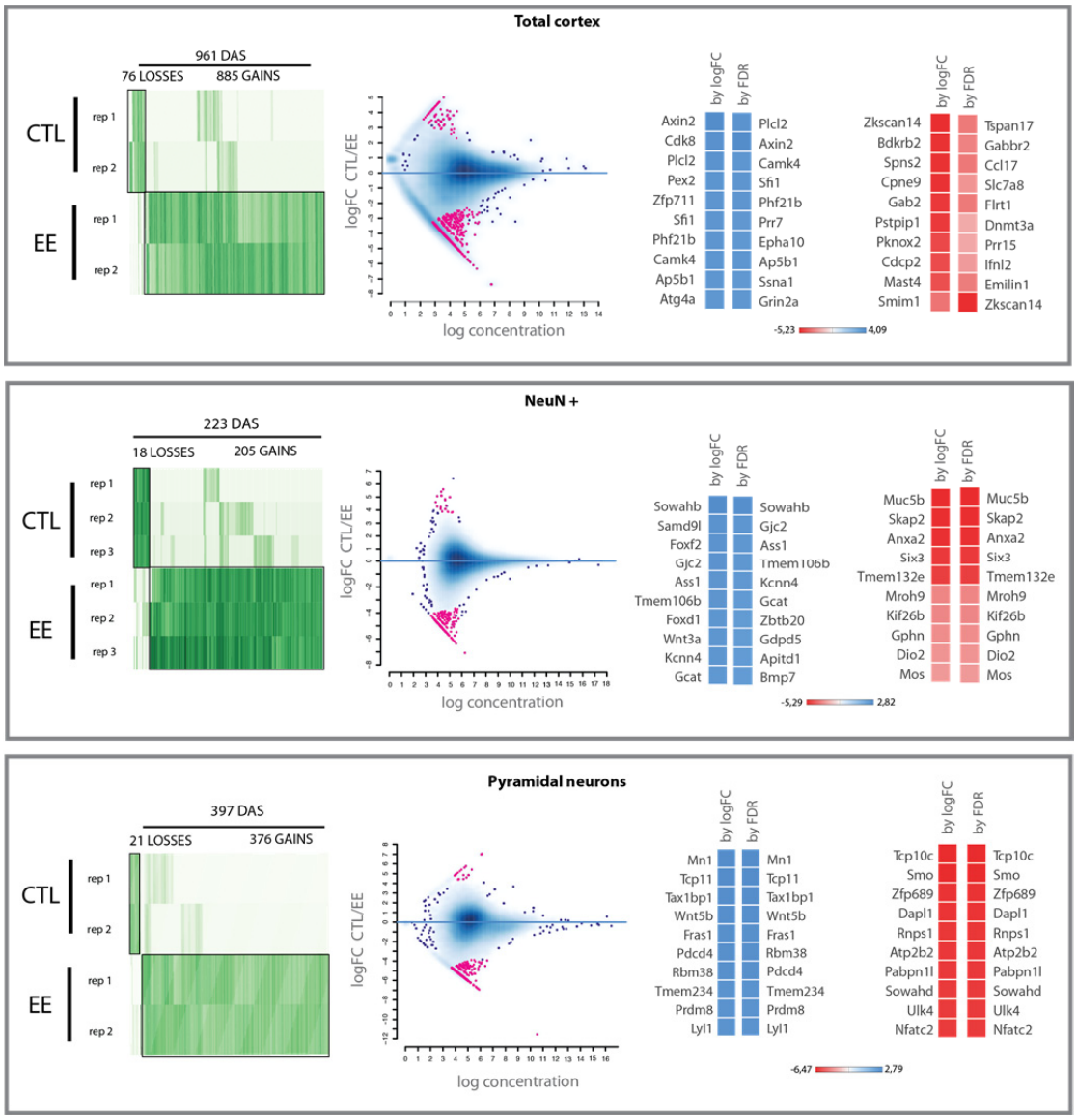


Figure 3.3. Enhanced activity and top differential accessibility sites.
Coverage heatmaps of DAS together with a log concentration vs logFC plots and the top ten results for each population. Both heatmaps and the log concentration plot clearly shows a bias toward gains rather than losses. It is interesting to realize that narrowing the cell population the number of DAS decreases. Top results are sorted by fold change (left) or FDR (right), however losses do not change the rank by sorting neither the fold change nor the FDR value.

3.4.2 Sorted neuronal cell types allow to dissect pathways revealed in whole tissue

Another strategy to understand the effects of EE in gene accessibility is to study the DAS gene-associated function by a gene ontology term enrichment analysis. As reported above, the main changes induced by EE are a general gain in open chromatin regions for all datasets studied here. Likewise, the three datasets show a general gain of similar biological processes with some exceptions (Figure 3.5A, 3.6A and 3.7A), especially those genes that recapitulate signaling pathways in developmental stages, such as the previously mentioned Wnt, BMP or Shh. These molecular signals are well known to be involved in neuronal architecture changes, such as synaptic structure (p_{cortex} = 3.5E-6) or neuroprojection development ($p_{\text{NeuN+}}$ = 4,5E-3, $p_{\text{pyramidal}}$ = 32.0E-3) (Figure 3.5 A, 3.6A, 3.7A). For the first term we find well-characterized genes such as *Bdnf*, *Camk2a* or *Grin2b* important in leading LTP processes and enhancing the synaptic strength and plasticity (Figure 3.5A)(Bramham and Messaoudi, 2005; Karpova, 2014;

Leal et al., 2014; Lu et al., 2008; Paoletti et al., 2013; Shonesy et al., 2014; Voglis and Tavernarakis, 2006). Second, it is interesting that NeuN+ and pyramidal neurons shared gains in *Bmp7* or *Wnt3a* as previously introduced. Both of them contribute to neuronal architecture changes (Figure 3.6A and 3.7A), the former to gaining synaptic strength (Withers et al., 2000) and the letter by its capability of increasing the neuron cone growth(Purro et al., 2008).

In fact the Wnt signaling term enrichment is prevalent for both whole cortex and pyramidal neurons (p_{cortex} = 4.7E-3, $p_{\text{pyramidal}}$ = 1.7E-3), with different ligand contributions (Figure 3.5A and 3.6A). In cerebral cortex we find canonical signaling pathways by enriched genes, such as the ligand *Wnt2*, which seems to control the proliferation of new-born neurons in the developing brain (Sousa et al., 2010), but also an enrichment of *Wnt7b* gene that controls synapse organization *(p* = 4.7E-3) (Rosso et al., 2005). However, in the case of pyramidal neurons the Wnt ligands that are enriched are different: *Wnt3a* linked to synaptic architecture (Purro et al., 2008) and *Wnt5b* with no known role reported in the brain (Oliva et al., 2013a, b; Rosso and Inestrosa, 2013). Even if NeuN+ do not show any specific term concerning Wnt signaling, some DAS-associated genes are found clustered in another term (axon development, *p* =1.4E-3) with *Wnt3a*. In addition, BMP signaling is also present in pyramidal neurons (*p* = 2.6E-3) by the gain of accessibility in *Bmp7* gene.

Taken together it appears that Wnt, BMP and to some extent Shh play a key role in the response to EE, with only minor differences in ligand enrichments found across the three datasets. This highlights the importance of these signals in modulating synaptic strength due to increased neuronal activity.

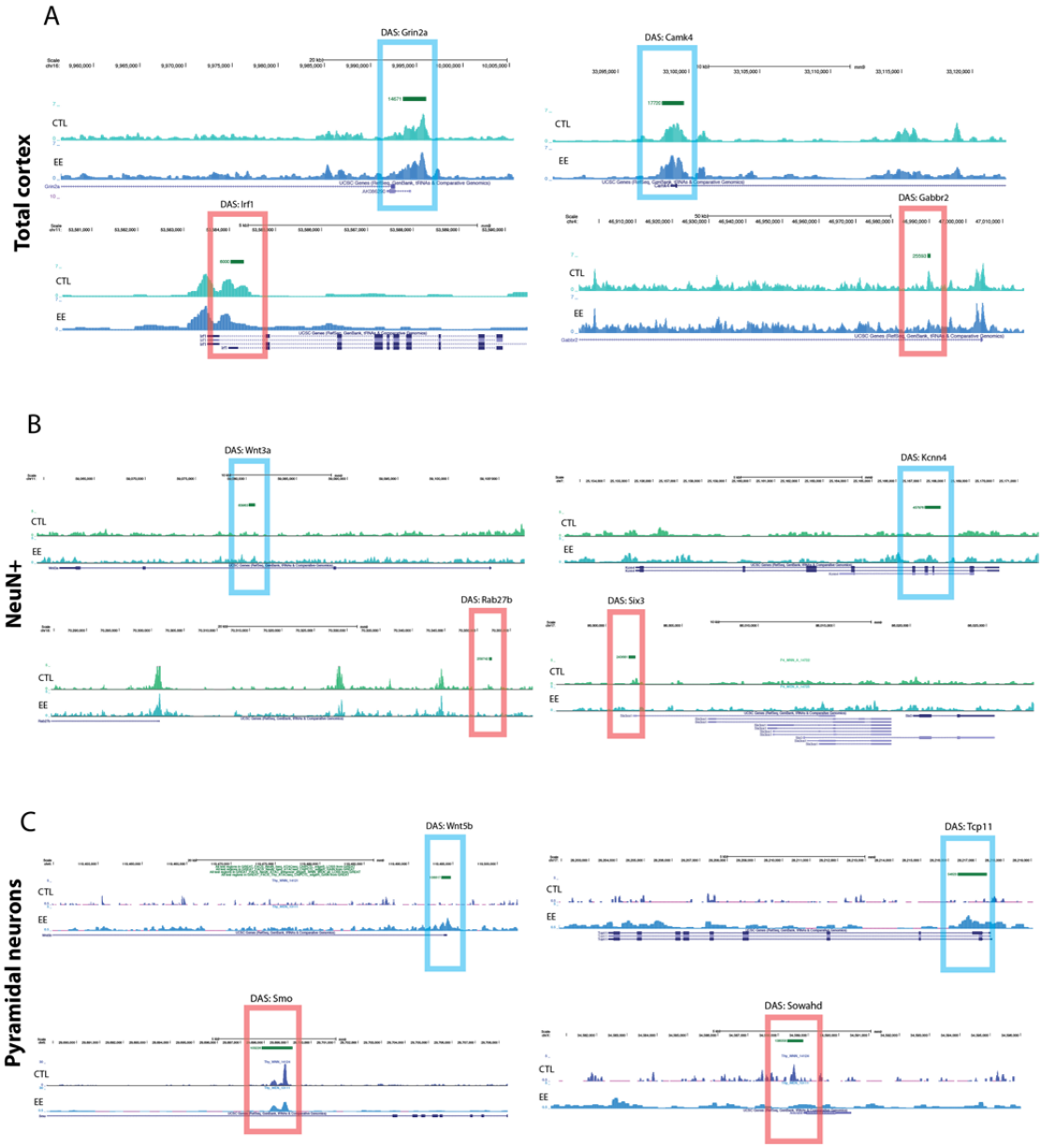


Figure 3.4. UCSC -Differential accessibility examples
Genome browser showing some DAS found by the pipeline fseq-DiffBind(edgeR). **A)** Tracks corresponding to the whole cortex, showing first two examples of increased accessibility in *Grin2a* (left) and *Camk4* (right) with a blue frame highlighting the DAS region. Bellow two more examples corresponding to *Irf1* and *Gabbr2* genes; a red frame is pointing to the loss of accessibility in EE samples. **B)** UCSC tracks examples of NeuN+ DAS, increased in EE: upper blue frames on *Wnt3a* and *Kcnn4* genes, decreased in EE: *Rab27b* and *Six3* gene. **C)** Finally UCSC genome browser tracks for increased DAS in EE on *Wnt5b* and *Tcp11* genes and decreased in *Smo* and *Sowahd* genes.

We found a substantial number of pathways in the GO term enrichment analysis, which are important for other key steps in gaining synaptic strength. For example, we find the Rho signaling pathway controlling the actin cytoskeleton (p = 9,3E-3) with genes such as the Rho/rac guanine exchange factors *Arhgef19* and *Arhgef2*, or the Rho GTPase-activating protein *Arhgaf32* (Figure 3.5 A) ((Auer et al., 2011; Tolias et al., 2011). There is also an increase in accessibility of MAPK signaling pathway genes in pyramidal neurons, possible effectors of the leading signal from NMDA receptors to induce the transcription of early-transcribed genes that will ultimately induce LTP responses (Figure 3.7 A) (Day and Sweatt, 2011).

Cellular component analysis determines on which parts of the cells the discovered DAS-associated genes may exert their action once they are transcribed and translated. This GO analysis indicates that there is a general enrichment of neuron projection parts (Figures 3.5 B, 3.6 B and 3.7 B). Furthermore, pyramidal neurons show enrichment of voltage-gated calcium channel activity, pointing to an enhanced calcium transport (Voglis and Tavernarakis, 2006). This term is again highlighted by GO molecular function analysis. Moreover there is an enrichment of an enhanced accessibility in the vicinity of genes linked to kinase activity, GTPase activity, enhanced transcription activity, Rho signaling and Ras activity (Figures 3.5 C, 3.6 C and 3.7 C).

Globally, GO cellular component and molecular function enrichment analysis revealed that EE induces enhanced activity of genes that are important for synaptic function, denoted by the activity of calcium and potassium ion transport. Also evidenced is a gain of accessibility of genes important for gaining synaptic strength, especially the Rho signaling with its role in regulating synaptic actin cytoskeleton (Auer et al., 2011; Tolias et al., 2011). Synaptic activity needs to be translated into signaling cascades for new transcripts to be induced. In the case of pyramidal neurons this process could be elicited by the MAPK signaling cascade after NMDA activation (Figure 3.8 A)(Day and Sweatt, 2011).

Additionally, KEGG pathway enrichment analysis in whole cortex reveals other interesting terms such as Oxytocin (Figure 3.5 D). Similarly, and also in the total cortex dataset, the ChIP-seq portion of the current study reveals increased H3K79me2 activity on the oxytocin-encoding gene. This molecule has been shown to play an important role in synaptic transmission in sensory cortices during postnatal development (Zheng et al., 2014). Moreover, cAMP and calcium signaling are also enriched, pointing again to DAS associated genes linked to synaptic activity. Finally, we found enrichment of angiogenesis by an increased DAS in genes related with VEGF signaling pathway.

3.4.3 DAS-associated genes and transcription

Concerning the relationship of ATAC-seq with the differential expressed gene (DEGs) results, the largest intersection corresponds to increased gene accessibility and loss of expression (green dots, Figures 3.5 E, 3.6 D and 3.7 D), an unexpected result. This discrepancy could potentially mean that accessibility changes are not directly correlated with transcriptional activity, although it has been shown that open-chromatin regions can correlate to some extent with mRNA-level (Dong and Weng, 2013) . However, no study investigated this correlation in treated and untreated groups, but all compared different tissue types. Thus gains in gene accessibility induced by environmental effects could simply provide increased probability of TF binding without a direct implication on transcription. Other factors might play a role, e.g. DNA methylation, or this may reflect regulatory dynamics over time: regions may be kept open although the related genes are not required anymore. For all three datasets, the genes in this intersection involve the presence of the previously described Wnt and BMP signaling genes. For the whole cortex we find an enrichment of angiogenesis by the enriched term, VEGF signaling pathway, where we can find the gene *Nos3*. NeuN+-mRNA intersected genes could not cluster into any GO, but ligands such as *Wnt11* and *Bmp7* appear (green dots, Figure 3.6D). Finally, in pyramidal neurons we find again again the enrichment for signaling pathways (green dots, Figure 3.7D).

The second largest intersection between DAS and DEG genes consists of up-regulated activity DASs and increased expression DEGs (blue dots, Figure 3.5E, 3.6D, and 3.7D), meaning some of the accessibility changes can be translated into a larger number of transcripts. These genes are related to signaling pathways such as the G-coupled receptor pathway in the whole cortex (Figure 3.5E), Rho members such as *Arhgef10* in total neurons (Figure 3.6D), or early expression genes such as *Arc* in pyramidal neurons (Figure 3.7D). In fact, it is not rare to find early expressed genes such as *Arc*, elucidating that the genes among these shared gains could have faster dynamics in response to neural activity elicited by EE.

Overall, the intersection of DASs and DEGs shows a low number of shared genes compared to the total number of genes of each set. This could be explained by the different powers of detection of each technique. For instance RNA-seq shows twice the number of differentially expressed genes than accessibility changes in ATAC-seq. On top of that it seems that accessibility could be better related to TF binding, rather than gene expression alone. This points to other regulatory mechanisms that could play an important role as transcription factor binding forms only part of the cascade that needs to take in place in order for transcription to be induced. Also we have seen that Wnt and BMP signals show accessibility gains in associated genes that are significantly down-regulated at the mRNA-level. Meanwhile genes that show both significant gains are more related with synaptic signal transduction pointing to a possible need of faster dynamics of transcription upon neuronal stimulation induced by EE. Moreover, the level of a transcript in a cell is defined by transcription, post-processing and degradation efficiencies, of which only the first can be directly influenced by chromatin accessibility. We therefore hypothesize that chromatin accessibility might be better correlated with gene body and promoter specific histone marks.

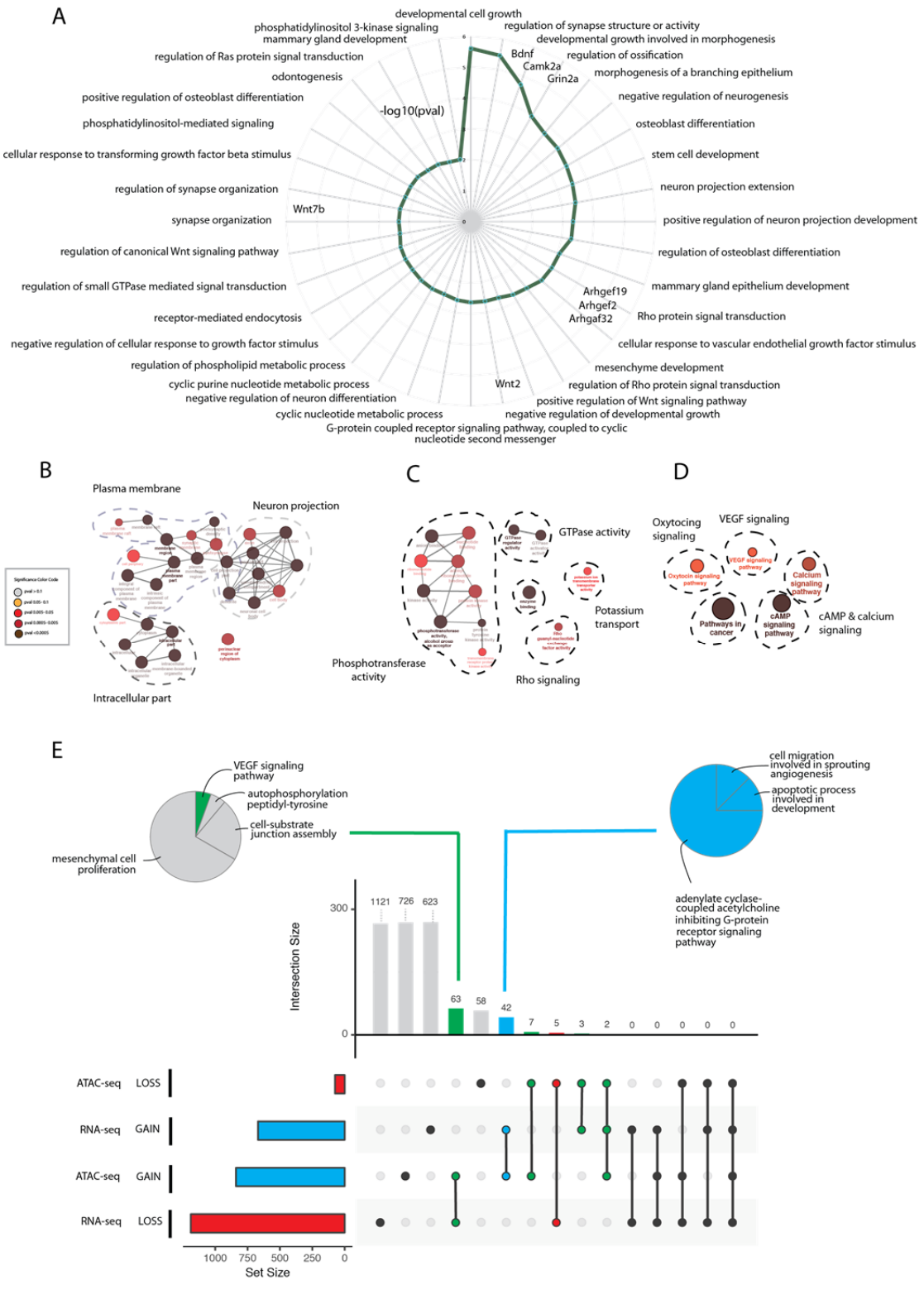


Figure 3.5. Whole cortex accessibility gene ontology analysis and its relationship with RNA-seq
Full overview GO results for whole cortex dataset with p-value < 0.05. **A)** Spider plot showing biological process that are enriched in increased DAS in EE samples. The plot is sorted by –log10(p-value). **B)** Cellular component GO analysis network. The color of the nodes represents the significance **C)** Molecular GO analysis network. Again the color represents the significance **D)** KEGG pathway GO analysis. **E)** Total cortex ATAC-seq and its relationship with mRNA. Bar plot at the left shows the data size for each set. Intersections are represented with wired dots. Green dots represent a discrepancy of the contrast direction. Blue dots correspond to an increase of those mutually shared genes in both datasets. Red dots highlight a common loss in both datasets for those shared genes. Upper bar plot represents the number of genes mutually (green, blue or red) or exclusively shared (in grey). A GO was also conducted for shared genes and represented in a pie-plot.

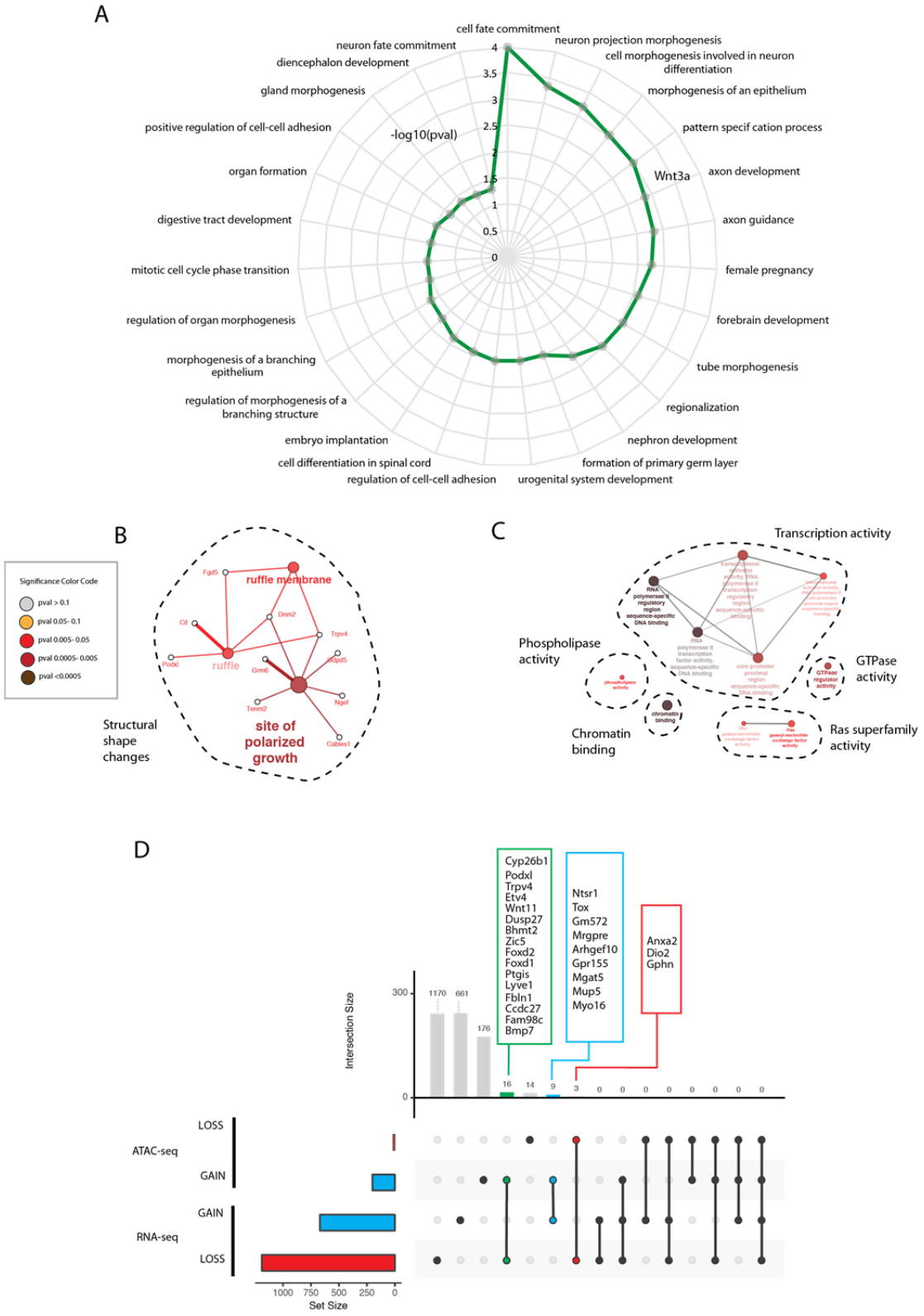


Figure 3.6. NeuN+ accessibility gene ontology analysis and its relationship with RNA-seq
Full overview GO results for NeuN+ dataset with p-value <0.05. **A)** Spider plot showing the biological process that are enriched on increased DAS in EE samples. The plot is sorted by –log10(p-value). **B)** Cellular component GO analysis network. **C)** Molecular GO analysis network. **D)** NeuN+ ATAC-seq and its relationship with RNA. Bar plot at the left shows the data size for each set. Intersections are represented with wired dots. Green dots represent a discrepancy of the contrast direction. Blue dots correspond to an increase of those mutually shared genes in both datasets. Red dots highlight a mutual loss in both dataset for those shared genes. Upper bar plot represents the number of genes mutually (green, blue or red) or exclusively shared (in grey). A GO was also conducted for shared genes and represented in a pie-plot.

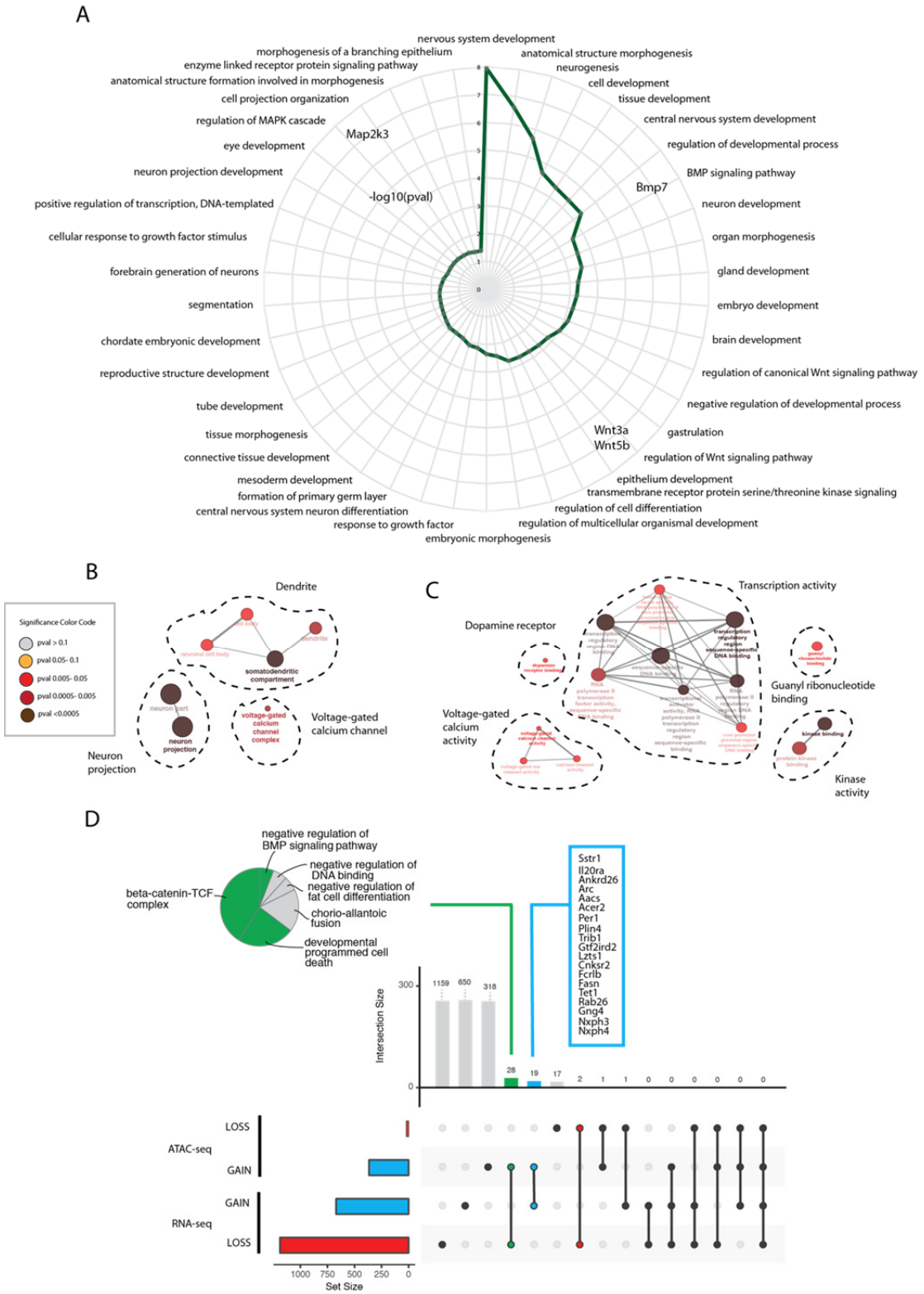


Figure 3.7. Pyramidal neurons accessibility gene ontology analysis and its relationship with RNA-seq
Full overview of GO results for pyramidal neuron dataset with p-value <0.05. **A)** Spider plot showing Biological process enriched on increased DAS in EE samples. The plot is sorted by –log10(p-value). **B)** Cellular component GO analysis network. The color of the nodes represents the significance **C)** Molecular GO analysis network **D)** Pyramidal neuron ATAC-seq and its relationship with RNA. Bar plot at the left shows the data size for each set. Intersections are represented with wired dots. Green dots represent a discrepancy of the contrast direction. Blue dots correspond to an increase of those mutually shared genes in both datasets. Red dots highlight a mutual loss in both dataset for those shared genes. A GO was also conducted for shared genes and represented in a pie-plot.

3.4.4 Motif enrichment found in differential activity sites discovered by fseq-DiffBind

Despite the previous finding of low intersection yield with changes in transcriptional activity, we wondered if the DAS that were discovered showed particular motifs which could enrich for particular transcription factors that could act as effectors of transcription.

Analysis with iRegulon (http://iregulon.aertslab.org) (Janky et al., 2014) allowed for the discovery of motif enrichment for the annotated increased DAS and also to obtain a prediction of transcription factors that might bind that particular motif. The biggest discovery motif output comes from NeuN+. Despite the big number of motifs found, only four of them show predicted binding of transcription factors that are differentially expressed in EE samples. Those transcription factors are *Zic1*, *Zic2*, *Zpf384* and *Tox* (motifs highlighted in green Figure 3.8 C). The first three are significantly down-regulated at the RNA level, meanwhile the last shows significant increase. Interestingly *Zic1* and *Zic2* can activate the transcription of the highly studied major risk variant, *Apoe*, involved in Alzheimer's disease development (Salero et al., 2001). Furthermore *Zic1* has been suggested to play an important role in neurogenesis(Aruga et al., 1994). Also *Zic2* seems to control the migration of neurons during development (Murillo et al., 2015) and, interestingly, the expression of key molecules in the axon guidance in some particular contexts (Escalante et al., 2013). Like *Zic1*, *Tox* seems to play important roles in neurogenesis (Artegiani et al., 2015; Karow and Berninger, 2015). Hence NeuN-positive cells show enriched transcription factors whose function is clearly linked to developmental processes such as neurogenesis, neuronal migration and neuronal projection.

In contrast, while motif discovery in whole cortex and pyramidal neuron reported a smaller number of DNA binding sequences, they were more interrelated (in blue, motifs mutually shared, Figure 3.8 A & B). Both datasets shared predicted transcription factors that are up-regulated in RNA-seq data, namely *Fos*, *Jun*, and *Lmx1a*. The first two play important roles in synaptic plasticity and they belong to the early-transcribed genes, well known as markers of neuronal activity(Flavell and Greenberg, 2008; Minatohara et al., 2015). *Lmx1a* determines the specification of dopaminergic and GABAergic neurons during development (Nakatani et al., 2010; Nefzger et al., 2012) but interestingly, it is known to induce the expression of targets such as *Wnt3a*, ligand gene of up-regulated accessibility in NeuN+ and pyramidal neurons. Furthermore some predicted transcription factors are commonly down-regulated in both datasets, such as: *Shox2*, *Isl1* and *Vax2*. The first one is a multifunctional transcript exerting a diverse amount of functions in different tissues (Blaschke et al., 1998), playing especially important roles in sensory neuron specification (Abdo et al., 2011; Scott et al., 2011). In addition, it can determine motor neuron specification (Liang et al., 2011) but also cholinergic neurons(Cho et al., 2014). *Vax2* in contrast seems to control the expression of genes important for axon guidance and plays important roles in the eye dorsal-ventral axis, as well as in the thalamocortical projections (Alfano et al., 2011; de la Torre-Ubieta and Bonni, 2011; Dufour et al., 2003). To sum up, shared predicted factors between the whole cortex and pyramidal neurons show up-regulated effectors whose function is related with learning-responsive genes, whereas effectors that show significant down-regulation at the mRNA level are more related to the developmental processes.

Concerning the non-shared predicted transcription factors, the whole cortex shows up-regulation in RNA-seq of *Nxph3* and *Mef2c*. The first one is a ligand of the adhesion molecules α–neurexins important in neurotransmitter release (Missler et al., 2003). Interestingly, *Nxph3* is expressed in excitatory neurons and it can be found in particular synapses in the cerebral cortical layer VIb playing a role together with α–neurexins in synaptic function (Beglopoulos et al., 2005). However, it is not endorsed by the literature as a transcription factor, but it could elicit this action by promoting dopaminergic specification (Nishimura et al., 2015). Likewise, *Mef2c* can also induce the differentiation of dopaminergic neurons (Cho et al., 2011). On top of that it has been shown by gain and loss of function experiments in the hippocampus to exert a substantial role in learning and memory processes. This is achieved by its action in limiting excessive synapse formation during activity-dependent refinement (Barbosa et al., 2008). Concerning the exclusive down-regulated factors in mRNA of the whole cortex, we find *Sox18*, *Prrx1*, *2310045NO1Rik*, *Als4*, *Dbx2* and *Emx2*. They all have key functions in neuronal development. *Sox18* induces gene expression related to angiogenesis (Cermenati et al., 2008). *Prrx1,* in contrast, maintains the adult neuronal stem cell niches(Shimozaki et al., 2013). *Dbx2* is important for ventral neural patterning (Ma et al., 2011; Pierani et al., 2001) and finally *Emx2* regulates neocortex development (Hamasaki et al., 2004) maintaining its expression during adulthood and playing a possible role in neuronal stem cell specification or maintenance (Gangemi et al., 2001). Finally, pyramidal neurons show a significantly decreased expression of the predicted transcription factors by iRegulon: *Six2*, *Rbm7* and *Zhx2* (Figure 3.8 B). Firstly, *Six2* seems to be mediate protective effects of GDNF in injured dopaminergic neurons and its overexpression can lead to a decrease in cell viability (Gao et al., 2016). *Rbm7* is a protein involved in both coding and non-coding RNAs transported to the exosome and its disruption can lead to motor neuron and cerebellum defects (Giunta et al., 2016). Finally *Zhx2* seems to be important for neuronal progenitor maintenance in the developing brain cortex (Wu et al., 2009).

To summarize, we found evidence that transcription factors with increased expression are related to responses in learning molecular processes or synaptic function, meanwhile losses are related with developmental processes.

Only six motifs are found to be mutually present in total cortex and pyramidal neurons sets (Figure 3.8 D). However, the predicted transcription factors could intersect better between the different samples, finding seven common effectors between the three datasets that are not differentially expressed at the mRNA level: *Pax7*, *Pax4*, *Pura*, *Hoxb4*, *Pdx1*, *Tfap2c*, and *Irf6* (Figure 3.8 E). In addition, using iRegulon we could sort the first 20 target genes for each of those transcription factors. However we could not report many changes at the mRNA level. In fact out of 14 differentially expressed target genes detected, only four were shown to be up-regulated: *Dusp6*, *Grb7*, *Cntn2*, and *Jup* (Figure 3.9). Again, similar to the motif intersection, the biggest number of shared predicted transcription factors belongs to the intersection between the whole cortex and pyramidal neuron dataset (see QR code, Figure 3.8 D).

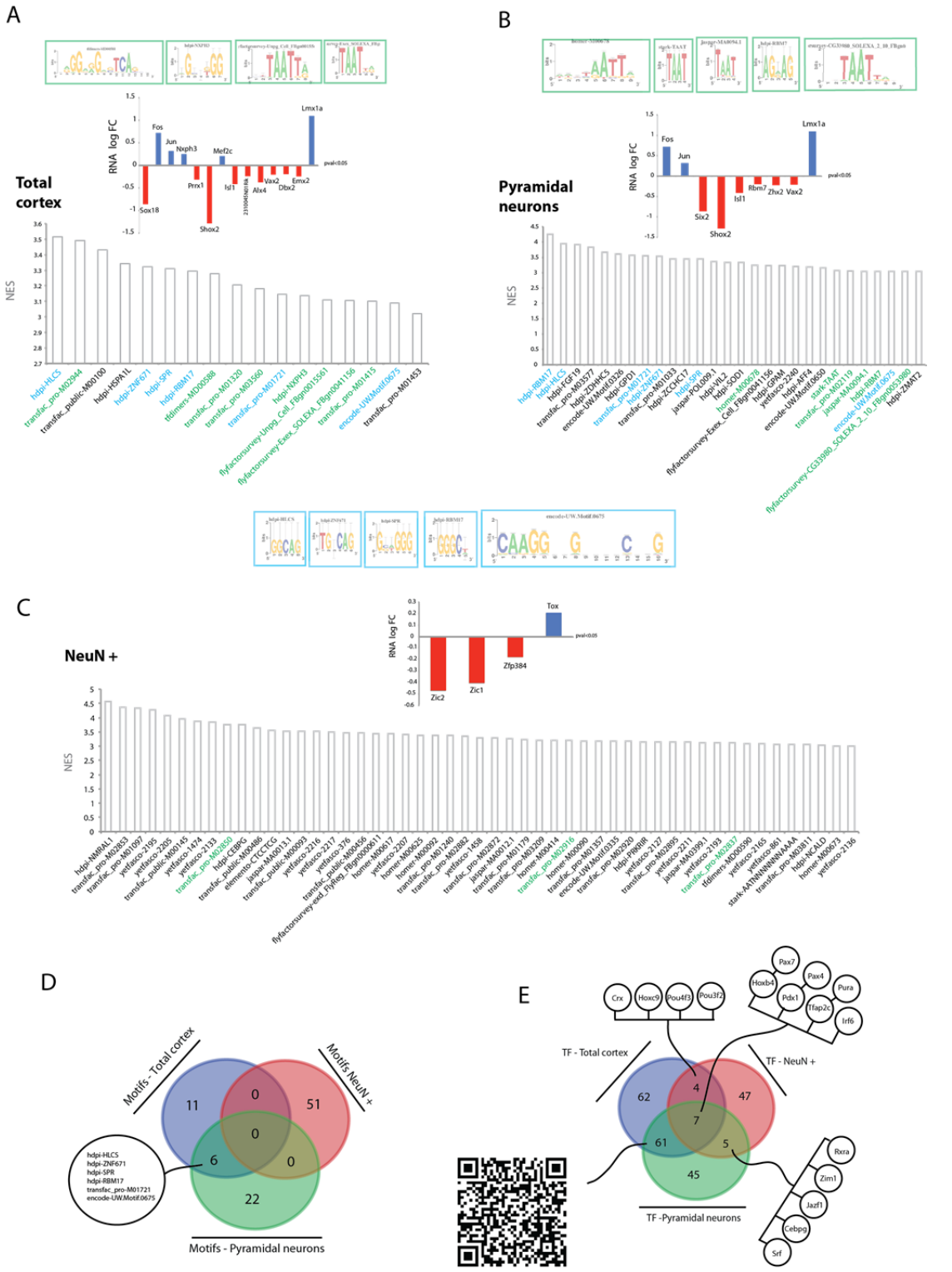


Figure 3.8. Motif analysis enrichment due to enhanced accessibility induced by EE
Motif enrichment discovery by iRegulon, NES=iRegulon enrichment score **A)** Motif discovery in DAS regions for whole cortex dataset. Green motifs show significant enhanced RNA expression of predicted transcription factors shown in a bar plot (blue bars increased and red decreased effectors in RNAseq data p<005). Blue motifs are the common shared motifs with pyramidal neuron **B)** Motif discovery in DAS regions for pyramidal neurons dataset **C)** Motif discovery in DAS regions for total onset of neurons dataset **D)** Intersection of the motifs discovered for all datasets **E)** Intersection of the predicted transcription factors discovered . QR gives access to an extended version of the figure .

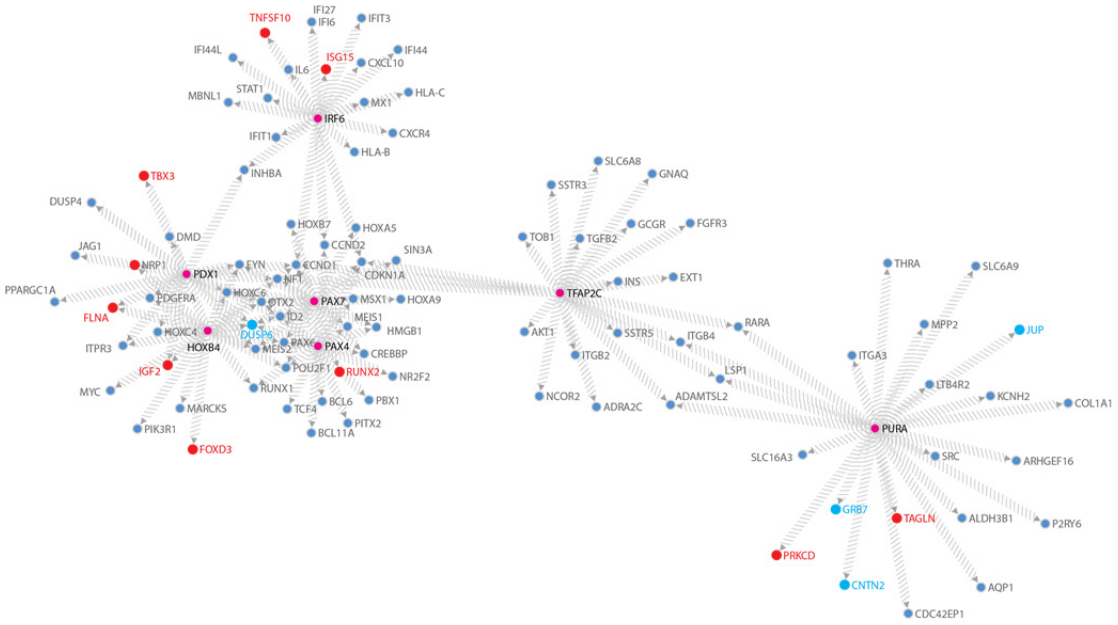


Figure 3.9. Shared transcription factor metatargetome prediction
Network showing the first 20 predicted genes regulated by the seven transcription factors shared by the whole cortex, neuronal population and pyramidal neurons. Transcription factors are shown in pink. Those genes that are significantly up-regulated are shown in blue and those down-regulated in red.

3.4.5 Action does not always induce a reaction: finding causality

Up to now, we have reported that changes in mRNA expression do not correlate well with changes in chromatin accessibility in the related cis-regulatory elements. In the previous section we discovered some enriched motifs that had greater accessibility. These could potentially provide binding access to some transcription factors that were significantly up-regulated or down-regulated due to EE exposure. Here we try to uncover whether these DAS-associated genes, where motifs and significantly expressed transcription factors were discovered, also show gene expression changes. This will allow the establishment of a mechanism of action-reaction: "*because the door is more open it might let significant effectors to enter [action] and establish their regulatory role [reaction]*".

For simplification we plot only the total cortex enriched motifs associated gene expression that also shows significant expression of their predicted transcription factors (to see the results for NeuN+ and pyramidal neuron visit the QR at Figure 3.10). We also show the top enriched motif (hdpi-HLCS) that actually does not show differential expression of its main effector *Hlcs*. The associated genes of this motif show a diverse response at the mRNA level. However, the number of significantly expressed genes clustered for this motif is substantial, but there is not a clear response towards a particular direction of expression. Then, when looking at those DAS-associated genes where the motifs discovered were enriched for differentially expressed factors due to EE, we could not find causality either. Nevertheless, there is some tendency to *Fos*-and *Junb*-enriched binding motifs (Figure 3.10).

To sum up these observations, we cannot find a general mechanism of action-reaction. One of the explanations of this behavior can be explained by the fact that some of the

transcription factors predicted can bind to different motifs. In addition, we can find transcription factors both up- and down-regulated for the same binding motif. For this reason it seems ambitious to establish a general mechanism of causality when many molecular mechanisms might be taking place (i.e. the recruitment of other factors that not necessarily bind to the DNA). However, this analysis could show evidence of particular mechanisms of causality that might require further study. For example it seems that the driving force of down-regulation of the genes *Skor1* and *Foxd3* could be the loss of transcription factors such as *Prrx1* and *Emx1*, as previously described (Figure 3.10).

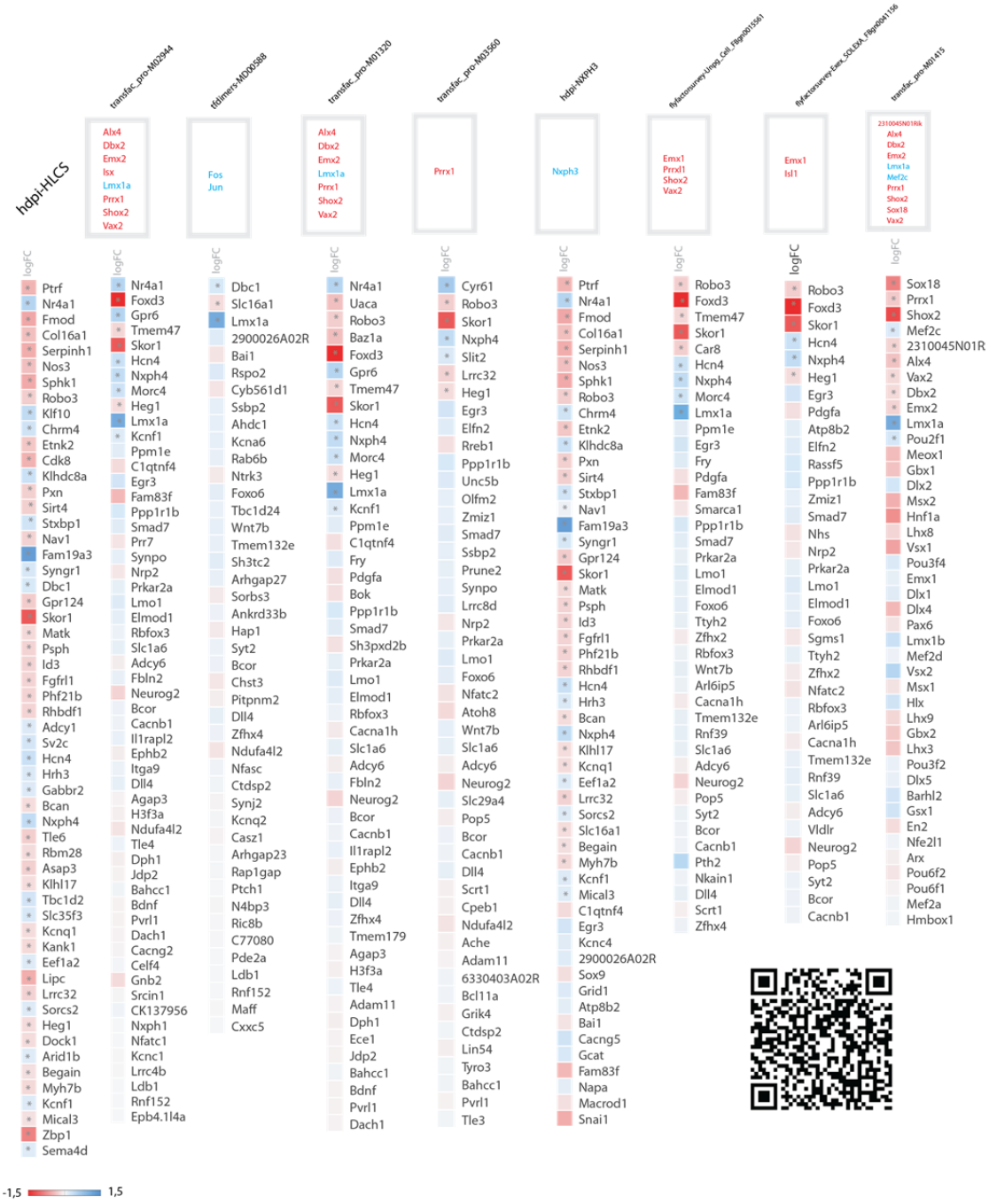


Figure 3.10. Attempts to find causality between gene accessibility, motifs, transcription factors and mRNA expression
Heatmaps of the RNA logFC of the DAS associated genes clustered for each motif discovered. Asterisks point to the significant changes (p-value<0.05). Grey boxes represent the predicted transcription factors that are differentially expressed due to EE

exposure. Blue TF denotes up-regulated and red down-regulated. Top of the plot shows the motif name discovered for the DAS associated genes.

3.4.6 EE induces higher activity in enhancer regulatory regions

An in-house enhancer prediction tool was developed to calculate accessibility changes within the enhancer regulatory regions due to EE. Illustrated by Figure 2.11 we found a global gain of chromatin accessibility in those regions in EE samples for all datasets. These results strongly suggest that EE induces an overall gain in chromatin accessibility in regulatory regions. It is also interesting to note that by narrowing the cell population specificity, the number of DAS in enhancer regions increases, contrary to what we observed using the fseq tool without prior annotation of regulatory regions (122 DAS for wholes cortex, 167 for NeuN+ and finally 343 for pyramidal neurons; Figure 3.11).

Generally enhancers tend to be located in between two genes, making it difficult to associate them with a particular locus without using chromatin conformation analysis methods such as chromosome conformation capture. Even though the GREAT tool allows computational association of regulatory elements and genes the majority of DAS enhancers retrieved in this study could not be uniquely associated to one gene (the association rule that was employed: 5kb upstream 1kb downstream – 200kb distal Figure 3.11 and 3.12).

Similar to a previous strategy using fseq, we first sort and rank the first 10 DAS enhancer changes by False Discovery Rate (FDR) value. Some interesting genes appear, such as *Camk2a* in whole cortex, or *Fos* in pyramidal neurons, both previously mentioned as important modulators of synaptic plasticity (Flavell and Greenberg, 2008; Lisman et al., 2002; Minatohara et al., 2015) (Figure 3.11 and 3.12 A and C).

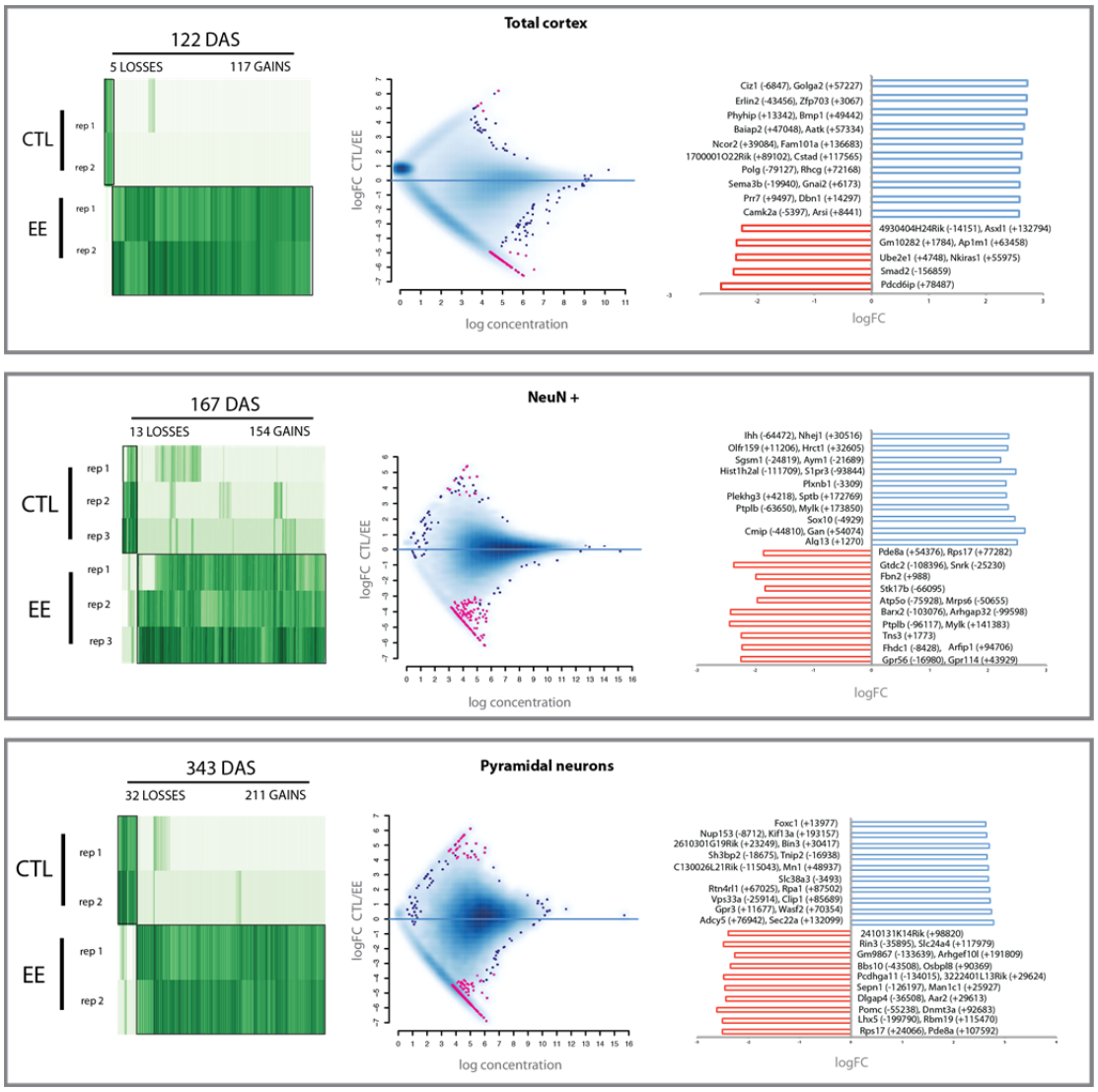


Figure 3.11. Changes in enhancer chromatin accessibility due to EE
Heatmaps of the genes cluster per each motif discovered that show a gain of accessibility but also a significant change in the motif promoters where their predicted transcription factors binds.

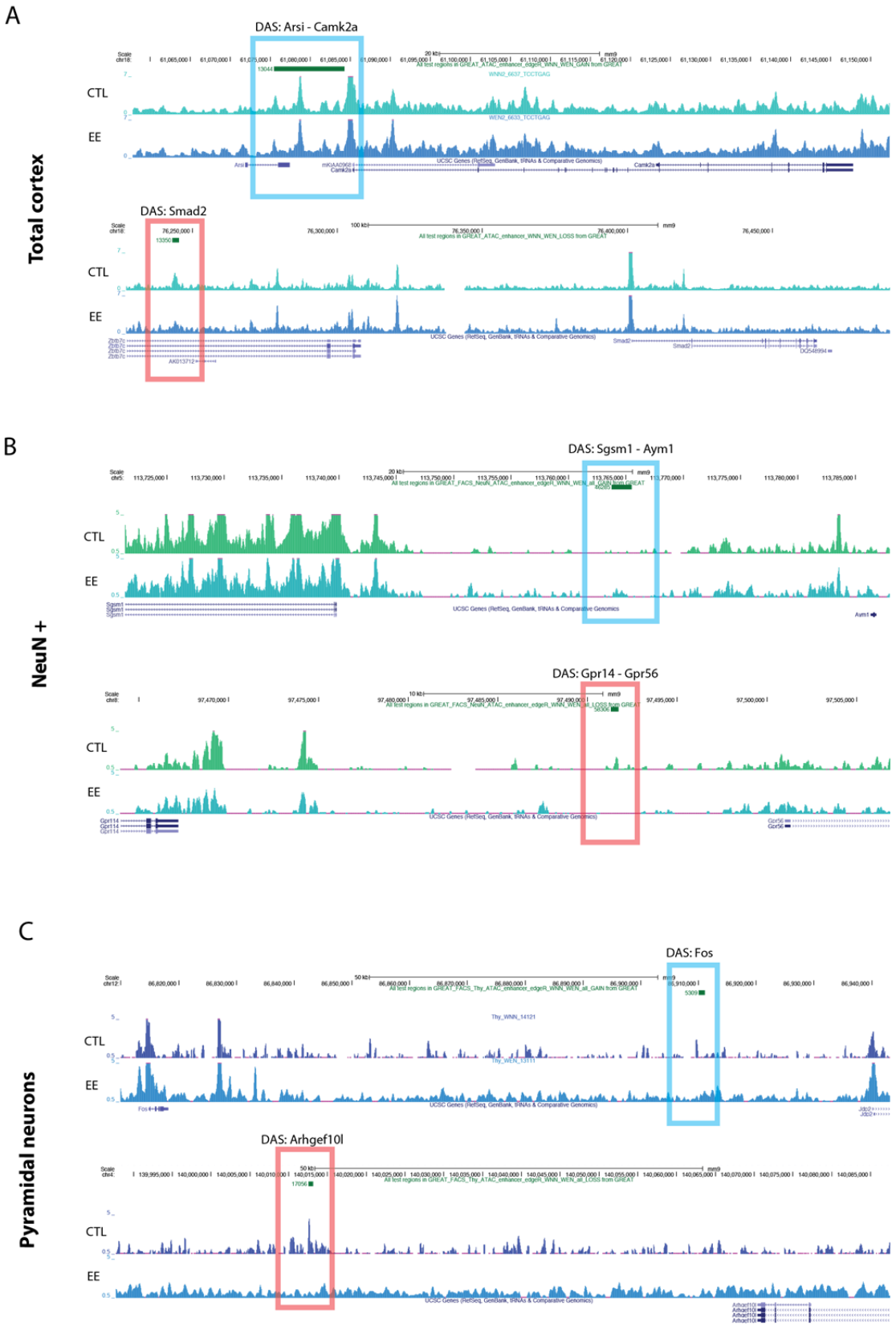


Figure 3.12. Enhancer DAS examples
UCSC genome browser tracks showing some examples of differential accessibility between control (CTL) and EE samples in whole cortex (A) , NeuN+ (B) and pyramidal neurons (C).

3.4.7 Gene ontology analysis in enhancer DAS induced by EE

Despite the issue with double associations reported by GREAT for most DAS enhancers, all the genes reported to be associated to a DAS region have been loaded into ClueGO(Bindea et al., 2009) to perform a GO term enrichment analysis (using terms from biological processes, cellular components, molecular function and KEGG pathways, Figure 3.13, 3.14, and 3.15). Here we also include some important terms in relation with the mRNA expression of their clustered genes.

Starting with the whole cortex dataset, the most significant GO term is BMP signaling (p = 1,1E-4), containing differentially expressed genes such as the ligand *Bmp1* (Figure 3.13). Moreover there is a gain in open enhancers associated to genes linked to synaptic plasticity (p = 2,7E-4), including *Shank2* or *Camk2a*. In this term we can only find the potassium voltage-gated channel *Kcnb1* that is also up-regulated at the RNA level. We found increased GABAergic gene enhancer signals (p = 20.0E-3), especially showing up-regulated expression of the GABAergic receptor *Gabbr2,* that was previously associated to a down-regulated intronic DAS by fseq. Furthermore more terms of increased accessibility fall into associated regions that are important for the control of negative regulation of axogenesis (p = 30.0E-3), containing associated genes such as *Sema3b*. We also find signals of up-regulated DAS-associated genes regulating long-term depression (LTD). Here we can find phosphatase *Ppp2r1b* clustered, but not significantly expressed. Finally, DAS chromatin modifiers associated genes appear by the enriched term protein deacetylase activity (p = 31.0E-3) which is an active field of research due to its ability to control memory formation (Graff and Tsai, 2013). Here we find important deacetylases such as *Sirt4* or *Hdac7* grouped together that are down-regulated at the mRNA level. Summing up, whole cortex DAS enhancers show associated genes related with BMP signal, regulation of synaptic plasticity, GABAergic synapses, axogenesis, LTD and protein deacetylases. The observed LTD signal might seem contradictory to previous findings, but it is important to mention that during post-natal development fine-tuning of the synapses is a key step. After birth, the brain shows an excess of synapses that might be rescaled due to the animal experience, gaining strength for some, whereas some others degenerate (Stiles and Jernigan, 2010).

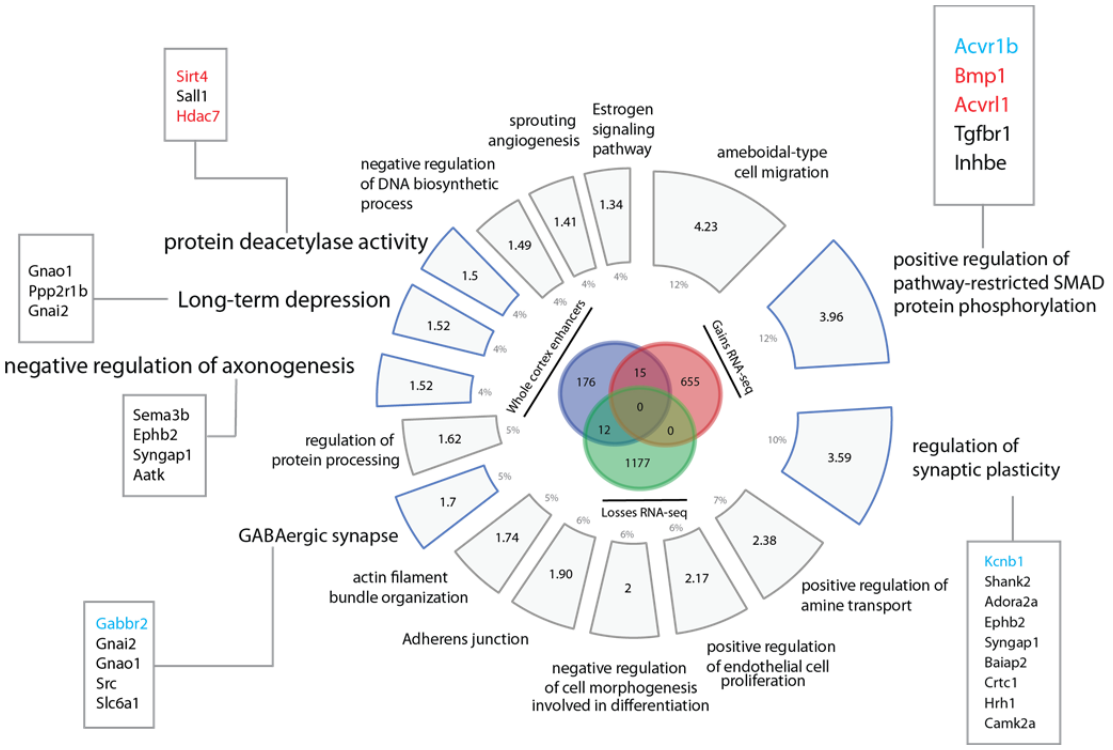


Figure 3.13. Gene ontology analysis of enhancer DAS in whole cortex
Pie plot showing the percentage and the –log10(p-value) of each GO term enriched. Terms were clustered with p-value <0.05 corresponding to both biological processes, cellular components, molecular function and KEGG pathway. Highlighted in dark blue are the terms discussed in the text. The genes there clustered that show gain in enhancer accessibility where intersected with the mRNA set showed in a grey boxed. Only those that are significant at the transcriptome are shown in blue or red if increased or decreased respectively. At the center there is an intersection of the total dataset with the RNA.

Concerning the NeuN+ population, identified enhancer DAS show the enriched GO terms cell adhesion molecules (p = 1.7E-3) and axon ensheathment (p = 45.0E-3) (Figure 3.14). A shared member of both terms is the adhesion molecule *Cntn2*. Also, the term 'intracellular ligand-gated calcium channel activity', represented by genes such as *Ryr1,* has been identified. Moreover there is overrepresentation of the Ras family pathway (p = 10.0E-3), including genes such as *Dennd3,* which regulates Rab family of GTPases (Marat et al., 2011). Thus, NeuN positive cells show a clear enrichment for terms of neuronal architecture changes and synaptic activity.

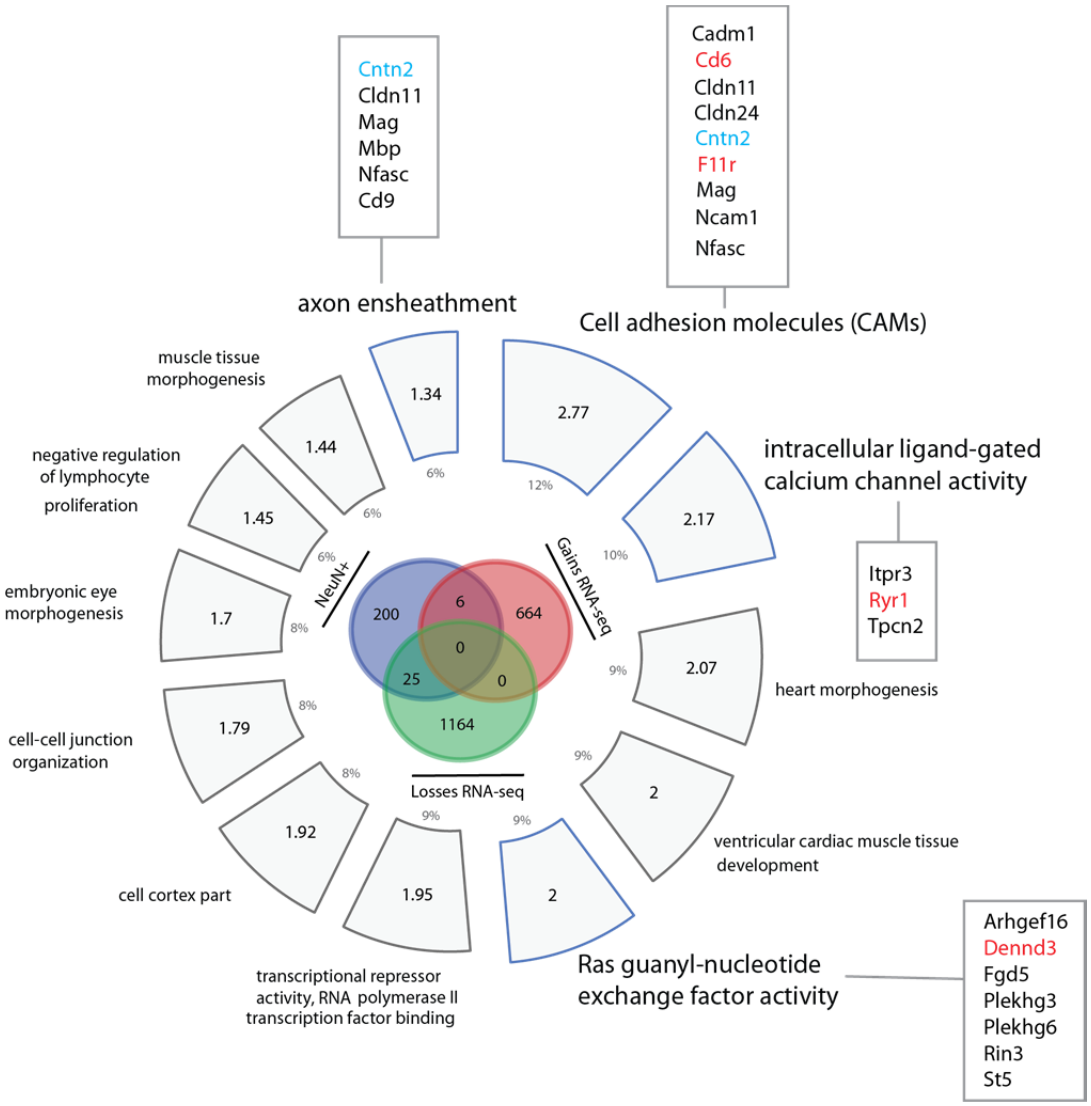


Figure 3.14. Gene ontology analysis of enhancer DAS in NeuN +
Pie plot showing the percentage and the –log10(p-value) of each enriched GO term. Terms were clustered with p-value <0.05 corresponding to both biological processes, cellular components, molecular function and KEGG pathway. Highlighted in dark blue are the terms discussed in the text. The clustered genes that show gain in enhancer accessibility and where intersected with the mRNA set are shown in the grey boxed. Only those that are significant at the transcriptome are shown in blue or red if increased or decreased respectively. At the center there is an intersection of the total dataset with the RNA.

Finally, we found the largest number of significantly enriched GO terms using the pyramidal neuron DASs, likely explained by having also the largest number of DASs. There is an enrichment of neurodevelopmental genes (p = 6.9E-12), neurogenesis (p = 4.7E-9) and neuron projection morphogenesis (p = 11.0E-6), sharing genes such as *Cntn2*, *Gas7* or *Myh10*. Similarly there is a gain in terms such as synapse organization (p = 220.0E-6), axon development (p = 690.0E-6) and actin cytoskeleton (p = 1.5E-3). On top of that up-regulated signals are enriched for terms such as voltage-gated cation channel activity (p = 9.6E-3), together with oxytocin (p = 700.0E-6) and MAPK signaling pathway (p = 1.1E-3) and GTPase activity (p = 4.0E-3) (Figure 3.15).

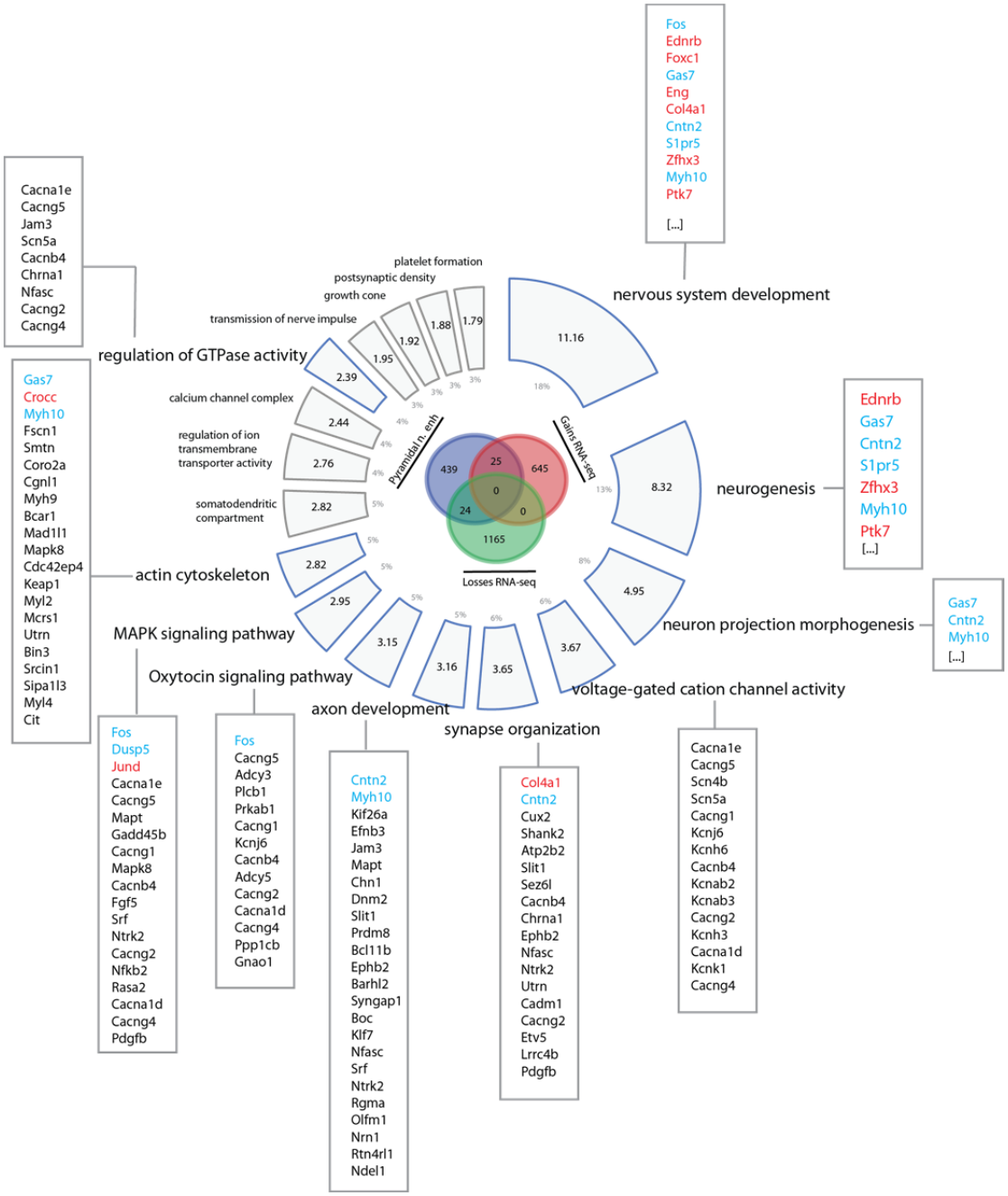


Figure 3.15. Gene ontology analysis of enhancer DAS in pyramidal neurons
Pie plot showing the percentage and the –log10(p-value) of each enriched GO term. Terms were clustered with p-value <0.05 corresponding to both biological processes, cellular components, molecular function and KEGG pathway. Highlighted in dark blue are the terms discussed in the text. The genes there clustered that show gain in enhancer accessibility where intersected with the mRNA set showed in a grey boxed. Only those that are significant at the transcriptome are shown in blue or red if increased or decreased respectively. At the center there is an intersection of the total dataset with the RNA.

Looking into the intersection of enhancer DASs and DEGs, only a small fraction of the related genes are shared, as was the case in the fseq based analysis. For both whole cortex and pyramidal neurons they share about the same number of genes for both gains and losses of expression. In contrast, NeuN+ can only intersect a small fraction of genes with expression gains, despite an increased bias to genes intersected with expression losses (Figure 3.14)

3.4.8 General Discussion: intersection between datasets and concluding remarks

Sorting different populations led to a better understanding of the specific effect of EE in a particular population, and thus avoiding potential constraints caused by mixing signals. Thus, our method is a top-down strategy with the initial hypothesis narrowing down each step of complexity. This will allow to better understand neuronal chromatin accessibility changes avoiding the intermingled signals that might be found in the whole cortex. Contrary to the epigenome, ATAC-seq analysis revealed similar GO terms enriched for all datasets despite the complexity of the library. In addition, it is surprising that pyramidal neurons better intersect with total cortex dataset than with NeuN+ population (Figure 3.16). This of course can reveal some important technical issues that ATAC-seq-NeuN+ procedure may have. For instance, nuclei need to be sorted without the crosslinking suffering damage during the sorting process (highlighting the reason of why we created these libraries with 200k instead of 50k as usual).

Fseq can call any significant peaks that show enough ATAC-seq coverage along the genome I at least one sample, hence, hence it is able to identify differentially active promoters, enhancers or any other type of regulatory regions. In light of the method intersection illustrated in Figure 2.16, it is obvious that fseq can call enhancer regulatory regions as both method share genes, especially for total cortex (third wired blue dots, Figure 3.16). However the amount of genes mutually shared is quite low compared to the total size of each dataset separately: the biggest intersection reported has 42 genes mutually shared between whole cortex and pyramidal neurons fseq that is around 4,2% of the whole cortex set (887 DAS gains).

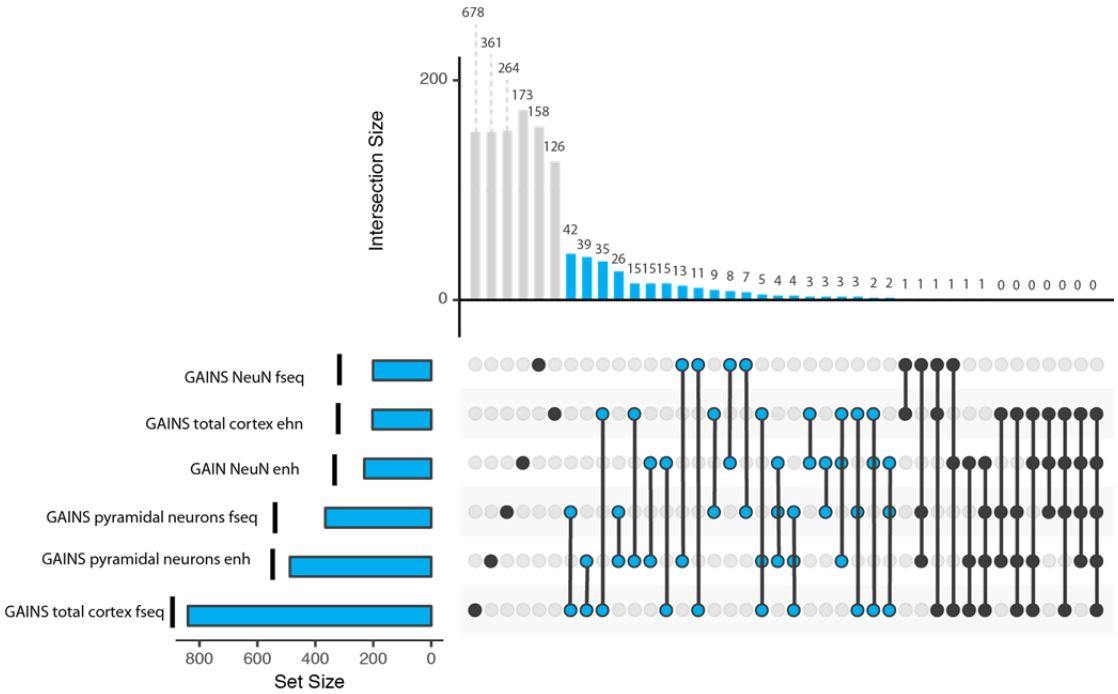


Figure 3.16. Intersection of fseq and enhancer prediction results
Gains of both fseq and enhancer prediction for each population are shown here intersected. Left bar plot shows the total number of genes in each dataset and upper plot shows the number of genes either exclusive or mutually shared. If sets have genes in common blue dotblots are linked by a wired.

In the present study (section 3.4.1 and 3.4.6) we have seen that EE induces a general gain of active open-chromatin regions linked to promoters and enhancers. The small yield of induced losses showed a random distribution around the TSS, which indicates an enrichment of false positive peaks in this group.

This global increase in accessibility can be interpreted as a priming effect for transcription allowing transcription factors to bind and exert their action. We find that gains in accessibility related with developmental signals such as Wnt or BMP are generally associated to down-regulation of gene expression, meanwhile mutual gains are mostly related with synaptic function and learning-responsive genes.

The low yield of intersection between ATAC and RNA-seq can be explained by their respective abilities to detect changes due to a given treatment, but also because of potential technical biases. However, this can also be explained by the fact that accessibility is more related with transcription factor binding rather than transcription itself (Dong and Weng, 2013). Another important factor is the dynamic change over time that both processes can undergo under EE. For example those common genes that show the same direction in both RNA-seq and ATAC-seq are related with early response expression genes, such as *Arc* (Davis et al., 2003; Flavell and Greenberg, 2008; Minatohara et al., 2015; Plath et al., 2006). This could point to a set of genes that requires faster chromatin accessibility and transcriptomic dynamics.

Concerning the biological relevance of this study, it is shown that gains in accessibility are related with an increased neuronal activity. Again a recapitulation of neurodevelopmental functions appeared, such as Wnt, BMP and to some extent Shh signaling. The changes in synaptic strength are explained by those factors, but also because the activation of signaling pathways such as Rho and Ras signaling is important for the modulation of synaptic actin cytoskeleton to gain synaptic strength. Additionally, gains in accessibility in early expressed genes take place as expected, such as *Fos* and *Arc*. Likewise, MAPK signaling accessibility genes are enhanced, potentially promoting the expression of early expressed genes. It is interesting to realize that the motif discovery herein described reveals predicted transcription factor genes significantly up-regulated which are important in mediating neuroplasticity, such as *Fos* and *Jun*. The loss of expression of predicted transcription factors is linked to the control of neuronal stem cell niches or developmental processes. Here, it is important to highlight that these mice are under postnatal development and fine-tuning of both circuits and populations might be taking place. This is illustrated by the appearance at the same time of neuroplasticity terms related with long term potentiation, but also depression; meaning some of the synapses will licit a gain of strength meanwhile other will disappear. The same might happen with the onset of neurogenesis, it is probable that gaining experience induced by EE reduces the neuron production in the subventricular regions, or turns on the apoptosis pathways to regulate the total number of neurons.

Chapter 4- Preface of ARTICLE 3

Contributions

My contribution to this paper consisted of, computational analysis of polyA-mRNA-seq data from all replicates of the mouse cortex epigenome study including batch effect removal, read mapping, differential expression and gene ontology enrichment analysis. Concerning the computational analysis of microRNAs in the same cohort I have performed the differential expression analysis, the targetome and the gene ontology analysis. I took care of the animals of the second wave of breeding and performed the dissections. I wrote this paper and did the figures.

Charlotte Hor and Justo Gonzalez performed the RNA extractions and prepared microRNA libraries.

Júlia Albaigues helped me with the new mouse breeding and some of the dissections.

Meritxel Pons bred the first group of animals and performed the whole cortex dataset dissections.

Long RNA library preparation and sequencing was performed at the Genomics Unit of the Centro Nacional de Análisis Genómico (CNAG) in Barcelona.

Objectives

The main objectives of the present study are:

- To analyse cerebral cortex mRNA changes upon exposure to Environmental Enrichment (EE).
- To analysis microRNA changes upon EE treatment and to find their targets.
- To find mRNA regulated by microRNA linked to EE exposure.

Highlights

- Transcriptomic analysis reveals a global change in gene expression linked to enhanced learning and memory-associated genes.
- Up-regulated microRNAs due to EE exposure target neurogenesis and cytoskeletal organization transcripts especially enriched for Wnt, BMP and FGF signals and Rho signalling
- We found that the non-coding transcript miR-125b-5p might specifically govern the expression of cytoskeletal organization genes upon EE exposure.
- MicroRNAs down-regulated upon EE exposure are enriched for let-7 family member and their targetome is related to angiogenesis, glycolisation and synaptogenesis.

Acronyms

Differentially Expressed Genes (DEG), Nitric Oxyde (NO), Environmental enrichment (EE), Gene Ontology (GO), Transcription Factor (TF), Wingless signaling (Wnt), Bone Morphogenic Protein (BMP), Fibroblast Growth Factor (FGF), Next Generation Sequencing (NGS), log-ratio of read count to median read count across sample (RLE), False Discovery Rate (FDR), Vascular Endothelial Growth Factor (VEGF)

Chapter 4
Coding and non-coding changes induced by environmental enrichment: linking molecular phenotypes and regulation

4.1 Abstract

Coding and non-coding transcripts are a dynamic set of elements highly exposed to environmental factors in both health and disease. Environmental enrichment impacts directly neuronal activity inducing changes that will ultimately trigger learning and memory processes. The multi-layered field of neuronal transcriptomics holds particular interest because spatio-temporal signal constrains might take place. Thus, synaptic activity needs to be translated into signals that are transduced into the nucleus in order to induce transcription. Then, transcribed molecules are sent back to the synaptic button where they are finally translated by the ribosome machinery. A large amount of regulatory elements might also be engaged in this process, especially non-coding transcripts that might act as molecular buffers of the incoming signals. To better understand how these DNA readouts might change under cognitive stimulation elicited by environmental enrichment (EE) we have used next generation sequencing (NGS) techniques to study both coding and non-coding elements under environmental enrichment.

4.2 Introduction

The transcriptome consists of a diverse set of elements of both coding and non-coding sequences that can be considered as phenotype readouts of the static DNA. Indeed, only a small section of the genome represents coding sequence regions (~1%, (Venter et al., 2001) that are finally transcribed. This process of transcription employs a large number of molecular regulatory mechanisms that interact with the DNA allowing production of nascent transcripts that will potentially go through a splicing checkpoint to produce different isoforms (Kapranov et al., 2007)

Further, the multi-layered transcriptome has enough intrinsic molecular tools to control and regulate gene expression by non-coding sequences, generally divided into long and small non-coding transcripts (Cech and Steitz, 2014). These can act as molecular buffers of expression controlling the dynamics of mRNA production or even having chromatin-remodeling features to repress some coding regions (Saxena and Carninci, 2011). Among the non-coding elements that are studied here, micro-RNAs are small RNA fragments of approximately 20 nucleotides long and are important for controlling gene expression. After a molecular process of maturation they usually bind to the 3'UTR of the target RNA to be silenced (Ha and Kim, 2014). Despite their diversity in gene expression regulation some studies have looked into particular effects of micro-RNA in

brain function, especially non-coding elements having neuronal-dependent expression (Eacker et al., 2013; Sim et al., 2014). Moreover, in the context of EE some studies have been published exploring the roles of microRNA and external factors. For example the microRNA-222 seems to mediate locomotor sensitivity to nicotine under EE (Gomez et al., 2016). Similarly, a small number of studies provided a general perspective of expression changes, rather than focusing on particular genes (Heinla et al., 2015; Reichmann et al., 2016; Rogers et al., 2016). However the most comprehensive study came from the research of Rampon et al. looking into a wide number of genes (~11,000) using microarrays, showing dynamic changes from early and late responses to EE in mouse cortex, enriching for important gene contributors of neuronal architecture (Rampon et al., 2000). To date, however, there are no studies, which provide an NGS-based overview of all possible mRNA changes that EE can trigger together with the non-coding regulatory elements that might change under EE treatment.

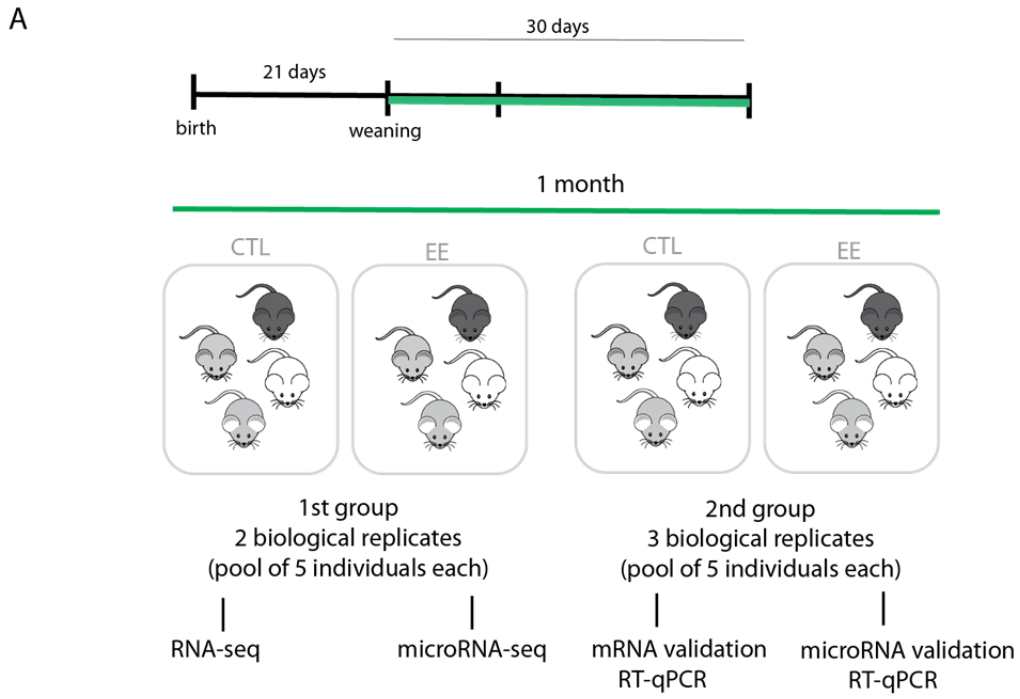
In our aim to fully understand transcriptomic regulatory changes induced by EE, we used NGS techniques for the first time to discover global changes in the mice cortex transcriptome and its regulation mediated by microRNAs.

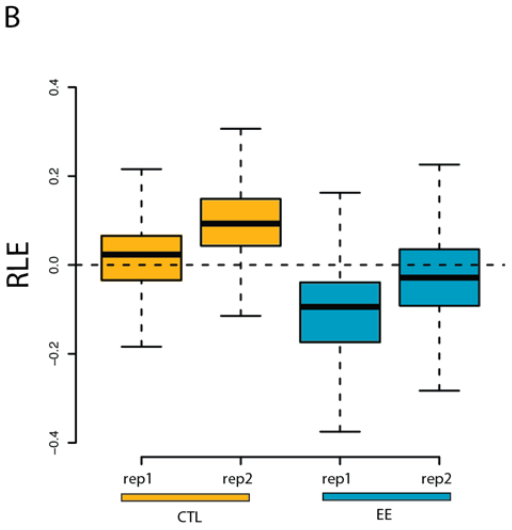
4.3 Transcriptome workflow and pipeline overview

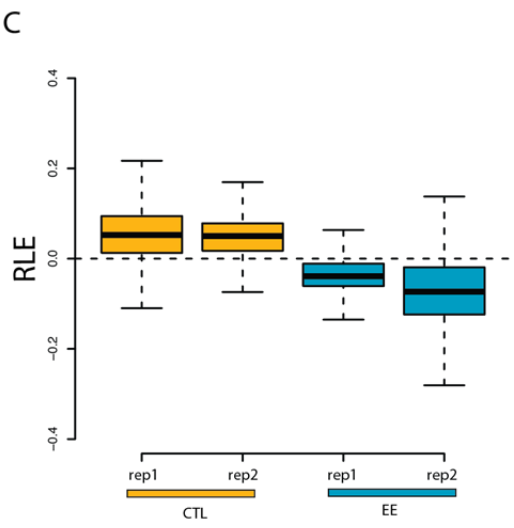
To fully determine the transcriptomic changes that might occur upon a month-long EE treatment, we performed RNA extraction of two different biological replicates per experimental condition (each consisting of a homogenized pool of 5 different cerebral cortices). Next, we sequenced and analyzed polyA mRNA, ribominus RNA (data here not included) and microRNA data (Figure 4.1 A). Furthermore, in order to validate previous results another group of 3 biological replicates per condition were bred (each biological replicate consisted of a pool of 5 cortices from individual animals, Figure 4.1). From a selected number of differentially expressing mRNA transcripts (21) and microRNAs (9), we performed the validation of previous results by carrying out RT-qPCRs (for primers used see Appendix – Table 4). Outlined here are the main expression changes found at the mRNA level and how small non-coding RNAs (microRNAs) can potentially buffer these induced changes brought about by EE. Still, more computational analysis needs to be carried out on the remaining data that corresponds to long non-coding RNAs, splicing events and isoforms.

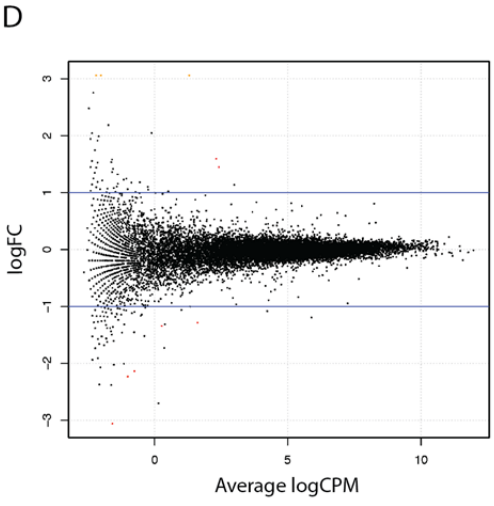
Due to a technical experimental issue whereby replicates were sequenced at different time points, the transcriptome data contained a significant batch effect. This is illustrated by the distribution of the samples showing a different scale of RLE values (log-ratio of read count to median read count across the sample) for biological replicates 1 and 2 (Figure 4.1 B). This issue made it difficult to acquire any differentially expressed genes after the first attempt of analysis (Figure 4.1 D). From the various computational tools that are available for removing unwanted variation between or within datasets (Conesa et al., 2016), the most useful of these was found to be RUV-seq (Risso et al., 2014). This approach inspects the less variable transcripts and considers them as artificial ‘spike-ins’. This rendering allows the program to normalize the data. Subsequently the resulting matrix is loaded into the differential analysis tools edgeR (Robinson et al., 2010) or Deseq2 (Love et al., 2014) tools. However even when the batch effect was partially removed, results clearly show a lower distribution of gene expression values for EE samples, which may influence the chances of detecting more losses than gains (Figure 4.1 C).

For more details about the experimental and computational analysis see ‘Methods’ chapter. All the tables can be found in the appendix or following the QR code at the header of the chapter.


A
30 days
21 days
birth
weaning
1 month
CTL
EE
CTL
EE
1st group
2 biological replicates
(pool of 5 individuals each)
2nd group
3 biological replicates
(pool of 5 individuals each)
RNA-seq
microRNA-seq
mRNA validation
RT-qPCR
microRNA validation
RT-qPCR


B
RLE
rep1
rep2
rep1
rep2
CTL
EE


C
RLE
rep1
rep2
rep1
rep2
CTL
EE


D
logFC
Average logCPM

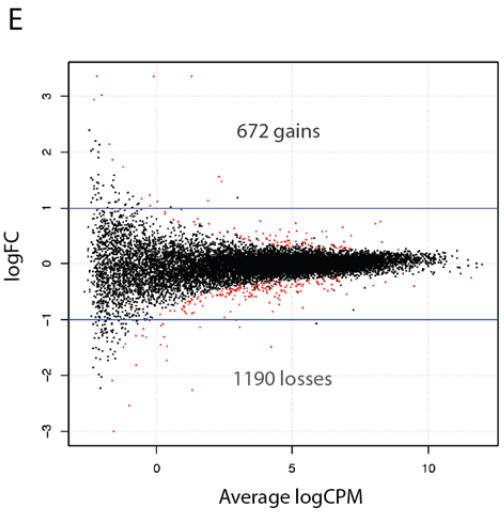

E
672 gains
1190 losses
logFC
Average logCPM

Figure 4.1. Transcriptome workflow and removal of unwanted variation
A) Schematic workflow illustrating the experimental steps for studying both coding and non-coding transcripts. **B and C)** Boxplots showing the RLE = log-ratio of read count to median read count across sample, before B) and after C) removal of the batch effect by RUVseq tool. **D and E)** Enrichment versus relative abundance. Significant logFC values due to EE are plotted as red dots. Values are ranked by the log value of their counts per million (CPM).

4.4 Results

4.4.1 Enhanced expression of neural transmission, learning and memory associated transcripts upon EE exposure.

The computational strategy chosen to perform the differential analysis (RUVg followed by edgeR) allowed to get a final set of 670 gains (increased gene expression) and 1190 losses (decreased gene expression) induced by the treatment (Figure 4.1 E)

Firstly we performed performed a gene ontology analysis in order to get a full perspective of the functional transcriptomic changes induced by EE. An increase of neuronal activity triggered by EE was clearly shown (Figure 4.2). The majority of the GO terms that appear are linked to synaptic transmission (pval= 1.0E-15), gaining synaptic strength and inducing learning (pval= 840.0E-9) and memory processes (pval= 970.0E-6). Those processes might be upregulated by involving the glutamatergic signaling pathway (pval= 3.3E-3) and also due to the molecular effects of nitric oxide (NO, pval= 110.0E-6) well known for its importance in synaptic plasticity (Prast and Philippu, 2001; Vincent, 2010). However losses are more diverse and less specific. Shown clearly are genes down-regulated in the angiogenesis pathway (pval= 30.0E-12) such as *Nos3* together with morphogene ligands such as *Wnt7a* , *Fgf2* and *Bmp4*. Indeed, BMP signalling terms appear (pval= 5.5E-3) clustering ligands such as *Bmp1*, *Bmp4*, *Bmp6*, and *Bmp7*. Other GO terms show non-neuronal down-regulation of gene expression in developmental genes that are shared between terms such as the renal system, lung or kidney. The reason why these terms appear is likely the wide diversity of actions that molecules such as Wnt or BMP can induce in different tissues. To sum up, the majority of the gains in gene transcription show enhancement of genes related with learning and memory responses involving the glutamatergic signaling pathway and NO. Concerning the losses an enrichment of developmental signals such as Wnt and BMP was found, which can also play important roles in synaptic signaling plasticity (Berke et al., 2013; Oliva et al., 2013a, b; Withers et al., 2000).

4.4.2 Top transcriptomic changes induced by EE

In order to identify the strongest and most significant differentially expressed genes (DEG), the top 20 genes were sorted and ranked by fold change and False Discovery Rate (FDR). Notably, we found an upregulation of neuronally active, early response genes such as *Fos*, *Junb*, *Arc*, *Egr1* or *Egr2* (Figure 4.2 B). This suggests that long term potentiation might be occuring (Davis et al., 2003; Flavell and Greenberg, 2008; Minatohara et al., 2015; Plath et al., 2006). Moreover there is an up-regulation of the neural growth factor *N4ra1* that regulates apoptosis and its gain of expression might be dependent on ERK/MAPK signalling activation (Mori Sequeiros Garcia et al., 2015). Furthermore it has been shown that this receptor is important for regulating spine density and distribution in pyramidal neurons (Chen et al., 2012). Interestingly, *Dusp1* and *Dusp5*, also found in the top upregulated transcripts, modulates MAPK signaling

pathway(Caunt and Keyse, 2013). In addition it has been shown that the up-regulation of these two phosphatases is dependent on vascular endothelial growth factor (VEGF) signaling pointing to a possible process of angiogenesis (Bellou et al., 2009)(Figure 4.2A). On the whole, EE induces the expression of early expressed genes important for synaptic plasticity, probably using the ERK/MAPK signaling pathway by the activation of glutamatergic signaling pathway (Figure 4.2A (Day and Sweatt, 2011) or alternatively by other signaling mechanisms such as VEGF (Bellou et al., 2009).

Interestingly we also find an up-regulation of the histone gene *Hist2h3b*, which seems to be a gene that is dependent on replication proceses, opening the possibility of an increase in proliferation(Banday et al., 2014). Similarly the up-regulation of *Btg2* can control neuron proliferation and differentiation (Tirone et al., 2013).

Continuing with the DEG that are down-regulated due the exposure of EE, the most significant findings involve the enrichment of fibrillar collagen tri-mer by the down-regulation of *Col1a2*, *Col3a1* and *Mmp13* expression, which may be related with changes in cell adhesion organization (Figure 4.2. B). Collagen extracelullar matrix is important for synapse morphology formation and may participate in synaptic activity (Bruses, 2010; Kurshan et al., 2014). In addition, it has been reported that down-regulation of *Foxd3*, found also down-regulated in our expression data, is important for avoiding inter-neuron differentiation (Dottori et al., 2001). Similar to gains, there are transcripts related with chromatin compaction by the down-regulation of *Hist2h2bb* and *Hist1h2bg* gene expression.

All previous results have been validated by RT-qPCR in three new biological replicates (Figure 4.2 C). From 21 genes tested, 19 were validated using specific primers (see Appendix-Table), showing that RNAseq results are consistent in spite of the batch effect.

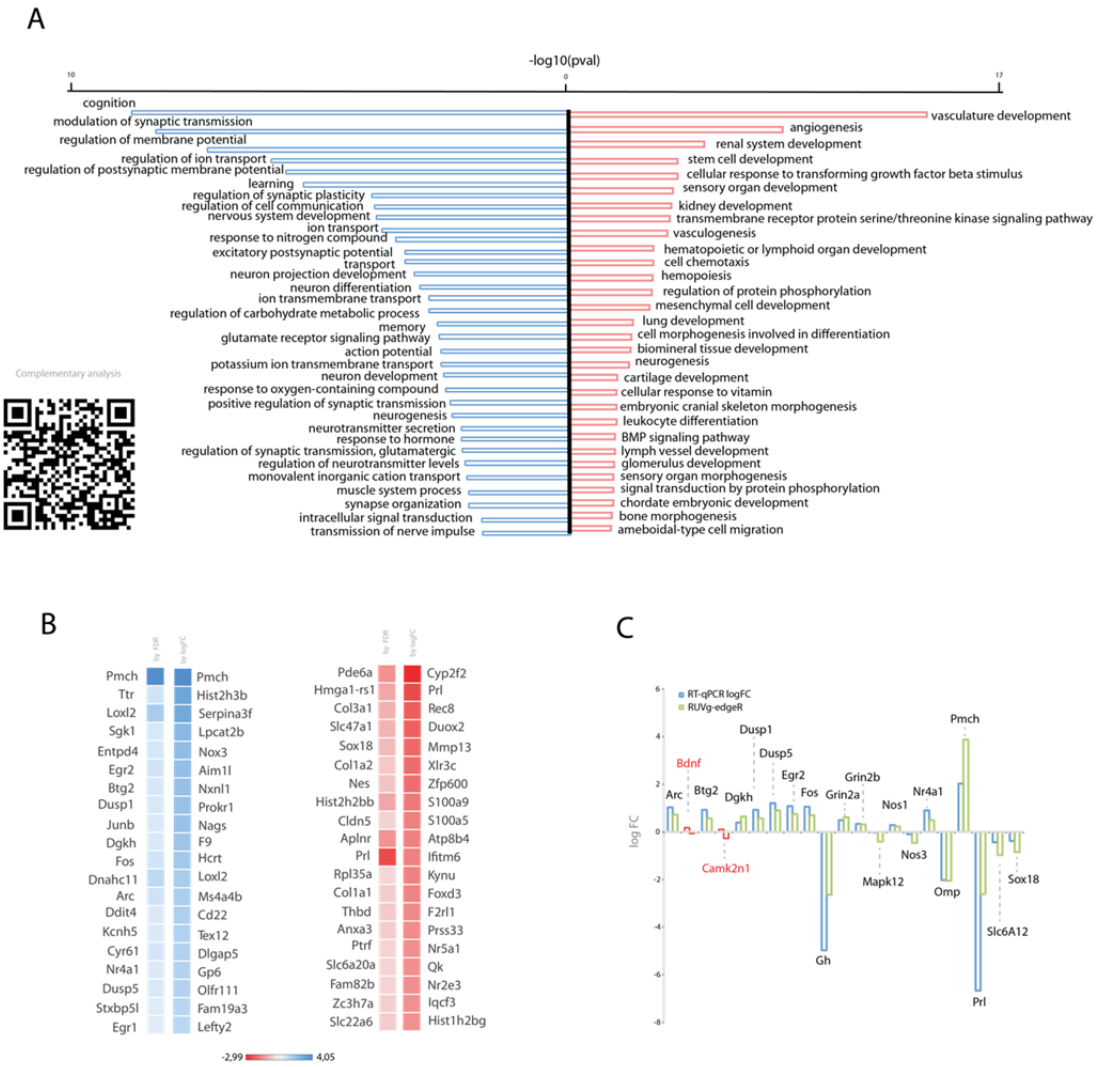


Figure 4.2. mRNA gene ontology analysis, top changes and validation
A) Biological process ontology analysis showing only gene ontology terms with pval< 0.01. Gains are plotted as blue bars (right) and losses in red (right). Results are ranked by their –log10(pval). **B)** Top 20 changes ranked by their FDR value or by their logFC. **C)** Validation by RT-qPCR of the main 21 changes found by the differential analysis. Blue bars represent the logFC of the RUVseq-edgeR results and green bars the logFC by using qPCR methodology. Only two transcripts that were differentially expressed could bot be validated: *Bdnf* and *Camk2n1*.

4.4.3 microRNA targetome: enrichment for neurodevelopmental genes

Small non-coding transcripts, or microRNA, are an important part of the regulatory elements that control RNA expression such that in recent years it has been a prompted focus of brain research both in health and in disease (Eacker et al., 2013; Sim et al., 2014). As previously highlighted, our dataset presented a strong batch bias and no tool tested was able to remove the unwanted variation completely. Therefore another strategy was carried out called fold change contrast conservation (see Methods), which found 104 up-regulated and 96 down-regulated microRNAs.

In order to know which microRNAs are regulating gene expression we used the miRoR-suite tool (Friedman et al., 2014)(http://www.proto.cs.huji.ac.il/mirror/) that provides a mouse cerebral cortex targetome dataset especially applicable for our research purposes in cortical studies. By using this tool, we identified the predicted targets followed by a gene ontology analysis that revealed a clear enrichment for gain of neurogenesis (pval= 410.0E-6), neuronal differentiation (pval= 390.0E-6) and neuron projection (pval=

18.0E-3) sharing genes such as *Wnt9a*, *Fgfr2*, or *Rab3a* (Figure 4.3 A). Noting the expression of the genes that clustered into those GO terms there are some significant targets that are lost in our own mRNA data, such as *Fgfr2*, *Cngb1*, *Nfatc4*, *Ngfr* and *Rom1*, pointing to a mechanism of active regulation of those transcripts due to EE. Moreover we find GO terms showing increased expression of microRNA targeting small GTPase-mediated signal transduction associated genes (pval= 15.0E-3), such as *Arhgef1*, *Arap3*, and again *Ngfrm,* which show significant down-regulation. Moreover, there is observable gene regulation related to calcium homeostasis mediated by the endoplasmic reticulum (pval= 34.0E-3).

While targets of down-regulated miRNAs show enrichment for genes related to angiogenesis (pval= 1.9E-3), the genes clustered in this GO term, *Nedd4*, *Pdcd6* and *Lect1*, do not show significant changes in gene expression. In addition there is down-regulation of genes linked to protein glycosylation (pval= 5.2E-3) showing *Vcp* as being significantly up-regulated in the RNA-seq and finally, microtubule cytoskeleton (pval= 3.3E-3) showing only *Inpp5j* that is significantly up-regulated.

To sum up previous results, we saw that target genes related with neurogenesis, neuronal projection, neuronal differentiation and small GTPase activity are possible being regulated by the gain of several microRNAs. Contrary, some of the transcripts being up-regulated are related with angiogenesis and protein glycosylation, possibly because of a lack of action of microRNAs that are down-regulated in our data. This output shows that despite the quality of our microRNA there is some clear biological relevance interesting to study.

4.4.4 Top microRNA changes induced by EE are involved in synaptic strength and morphology

To illustrate the biggest changes in microRNA expression, the top 10 microRNAs to have shown gains in expression were plotted and ranked. These non-coding transcripts show important roles in regulating brain function and can be essential in altered gene expression in disease (Figure 4.3 B). For example can target a wide number of transcripts. In the context of brain disorders this microRNA was reported to be up-regulated in the prefrontal auditory cortex in Autism (Ander et al., 2015). Meanwhile, in Schizophrenia it can be found to be down-regulated in blood cells (Yu et al., 2015). Moreover, we find up-regulation of the miR-138-2-3p that shows altered expression in Huntington Disease (Muller, 2014) and in epileptic rats (Bot et al., 2013). Concerning signaling pathway regulation, it has been shown that miR-582-3p regulates Wnt/β-catenin signaling in lung cancer cells (Fang et al., 2015), pointing to be a possible mediator of the down-regulation of neurodevelopment genes which we previously found enriched in the targetome of up-regulated miRNAs (Figure 4.2). Furthermore miR-99b-5p can regulate mTOR and mediate apoptosis under β-amyloid conditions in neuron cultures (Ye et al., 2015).

As previously mentioned EE triggers important changes in neuronal architecture. Interestingly we have found that upon EE exposure there is an up-regulation of miR-125b that can promote neuronal differentiation (Le et al., 2009) and also regulate synaptic shape in hippocampal neurons (Edbauer et al., 2010). In addition, this last publication shows other microRNAs that induce dendrite morphogenesis and that are also present in our dataset in the same fold change direction such as miR-132 and let-7c.

Also of interest is the up-regulation of miR-30d-5p that can control the expression of *Bdnf* in prefrontal cortex (Mellios et al., 2008). To summarize, up-regulated microRNA have important roles in brain function. Especially miR-125b and miR-132 up-regulation mediates synaptic strength, a role shared by miR-30d-5p as it controls the expression of *Bdnf*, a well-known inducer of synaptic plasticity (Bramham and Messaoudi, 2005; Karpova, 2014; Leal et al., 2014; Lu et al., 2008; Santos et al., 2010).

Concerning the top 10 down-regulated microRNAs, there is a general enrichment of several members of the let-7 (let-7a-5p, let-7g, let-7j, let-7f-5p, let-7i-5p) family, which are generally involved in developmental processes (Wang et al., 2012). They also have some other roles such as in opioid tolerance by regulating the μ opioid receptor (He et al., 2010). Moreover, this family together with miR-125 governs the formation of the neuromuscular junction in *Drosophila* (Caygill and Johnston, 2008) pointing to a potential role in synaptogenesis. Indeed, concerning the mechanisms that lead to synaptic strength we find miR-501-3p as an important regulator of AMPA receptor subunit *Gria1*. Its mechanism of action seems to be dependent on NMDA activation and is localized at the synaptic buttons (Hu et al., 2015). We also find a decreased expression of miR-34a-5p. Its gain of function has been shown to induce synaptic changes thus reducing synaptic function (Agostini et al., 2011). In the context of disease, it was found that a loss of miR-28a-5p upon exercise in a senescence mouse model (Cosin-Tomas et al., 2014). Moreover, miR-98-5p loss of expression has been shown to decrease apoptosis in SH-SY5Y neuronal-like cells under β-amyloid peptide conditions in the context of Alzheimer Disease (Li et al., 2016). Similarly to gains, losses show some important regulators of synaptogenesis such as let-7 microRNA members. On top of that, other microRNAs such as miR-501-3p impact synaptic strength by regulating the AMPA receptors, a key molecular step of the long-term potentiation process (Luscher and Malenka, 2012).

As evidenced by above results, both the top up- and down-regulated microRNAs have similar targets. The top up-regulated microRNAs target neuro-developmental transcripts pointing to synaptogenesis by the enrichment of miR-125-5p and miR-132. Top losses also show targets in developmental processes, especially enriching for let-7 family microRNAs.

Despite previous results, microRNA data could not be well validated. Only three microRNAs, miR-8114, miR702-3p and miR-7046-3p, show the same fold change direction values in the first group of breedings (i.e. those that were used to generate the present data) as the second group of animals bred. However separately, the first group RT-qPCR shows six validated microRNAs out of nine, and the second group another 6 of 9 (Figure 4.3C).

4.4.5 microRNAs targetome and RNA expression: finding causality

The intersection of microRNA targets discovered using the miRoR suite tool with the DEGs revealed potential mechanisms of transcript level regulation (Figure 4.3 D). First, we intersected targets of up-regulated microRNA with down-regulated mRNA transcripts in order to identify up-regulated microRNA whose target genes have been depleted. The targeted mRNA transcripts here are related with the BMP signaling pathway such as *Bmp1*, the Fibroblastic Growth Factor receptor *Fgfr2*, the neurotrophin TRKA receptor *Ngfr,* the Rho signaling by the appearance of *Arhgef1* transcript and adhesion molecules such as *Itga7*. Using the miROR suite tool to assess which microRNAs are regulating which genes, we see that there is an enrichment of miR-125b-5p. This is also one of the top microRNA found to be enriched in the top gains (Figure 4.3) that is regulating the following transcripts found to be down-regulated at the mRNA level: *Ap1g2*, *Arap3*, *Cdh22*, *Fgfr2*, *Mrc1*, *Tjap1*, *Tmem44*, *Tubd1* and *Speg*. By looking at these targets their main function is related with cytoskeleton organization by the enrichment of Rho signaling genes such as *Arap3* and elements such as the tubulin-related *Tubd1* or the cadherin-related *Cdh22*. We also find microRNAs enriched for this intersection such as miR-423-3p with similar targets. In conclusion, up-regulated microRNAs due the exposure of EE might target transcripts mainly related with cytoskeletal rearrangements and signals such as BMP and FGF. Our results are in support of previous research studies pointing to miR-125b-5p as a regulator of dendrite morphogenesis (Edbauer et al., 2010).

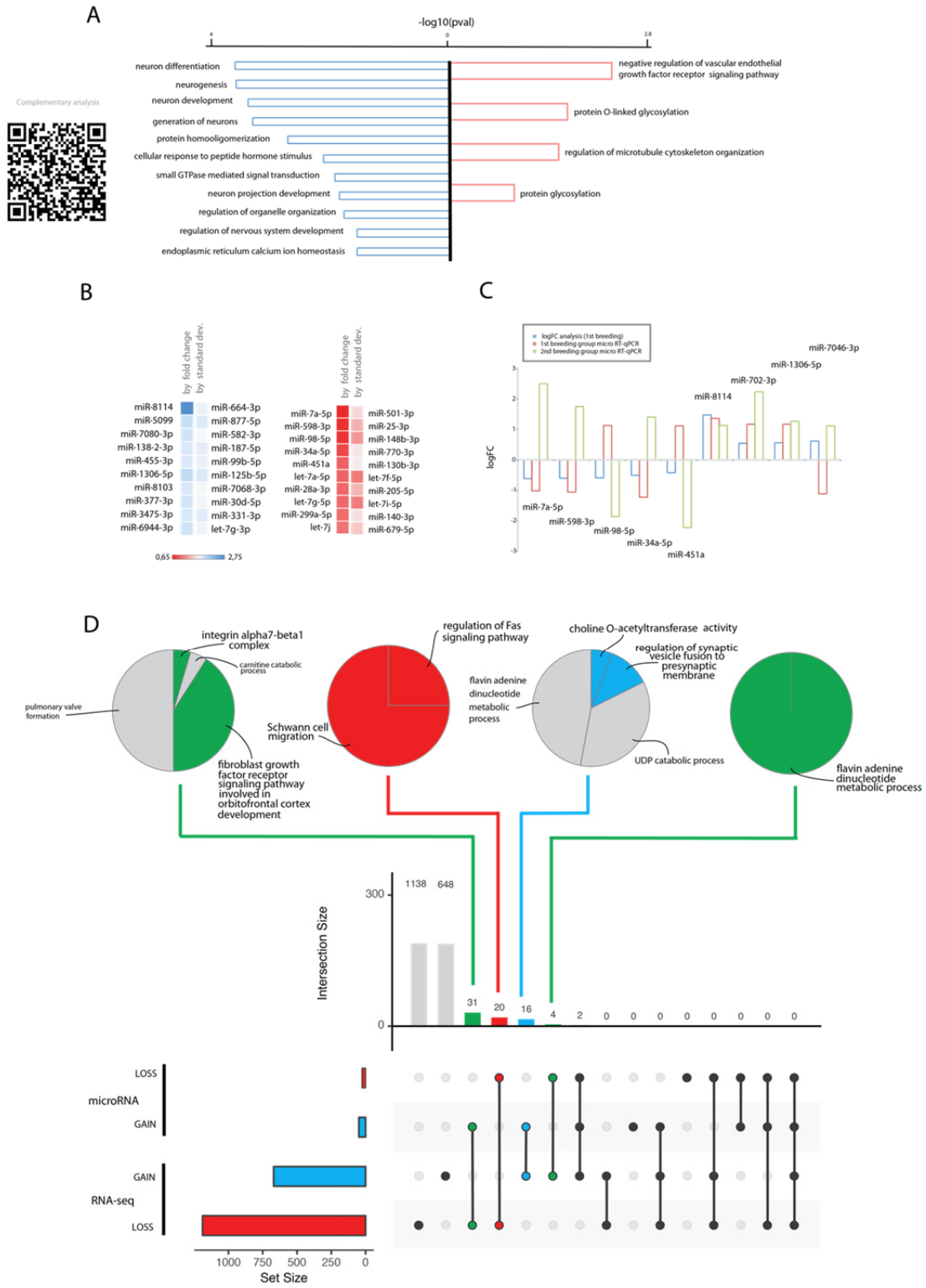


Figure 4.3. microRNA analysis: targetome gene ontology analysis, top fold changes, RT-qPCR validation and intersection with RNA-seq data.

A) Gene ontology analysis showing terms with a p-val<0.05 (using –log10(p-value)) of the predicted microRNA targets by miROR suite. **B)** Top fold changes in microRNA expression ranked and sorted by their fold change and by the standard deviation. **C)** RT-qPCR validation of the results using both the first group of animals to generate the data (blue and red bars) and a third group of animals corresponding to three new biological replicates (green bars) **D)** microRNA targetome and RNA-seq intersection. Left bar plot shows the size of each dataset (in blue: gains, in red: losses). Shared genes are connected by black bars. A green dot connected by a black bar represents a discrepancy (mechanism of microRNA-transcript regulation); blue refers to common shared gains; and red refers to losses. Upper plot show the size of each intersection.

4.4.6 General discussion and concluding remarks

The main objective of the current study is to determine the global transcriptomic changes upon EE exposure using NGS technology and thereby to find possible regulatory mechanisms that could trigger changes in expression.

EE induces enhanced neuronal activity by providing different signals that stimulate motor, sensorial and cognitive parts of cerebral cortex, which induce learning and memory processes. This observation is evidenced by the higher expression of early response genes that can induce enhanced synaptic strength (Davis et al. 2003; Plath et al.2006, Flavel & Greenberg, 2008; Minatohara et al.2015). These data are supported by gene ontology analysis (Figure 4.2A). To up-regulate those processes it is probable that glutamatergic signaling is involved. The signaling cascade that mediates transcription of these early genes may be achieved via the ERK/MAPK transduction pathway as some important modulators are also increased, such as *Dusp1* and *Dusp5* (Caunt and Keyse, 2013).

Interestingly we have reported for the first time up- and down-regulation of histone genes. Whereas some are up-regulated such as *Hist2h3b*, others as *Hist2h2bb* and *Hist1h2bg* are down-regulated after a month of EE exposure, pointing to a potential role in chromatin compaction or remodelling. Supporting this idea, Hake et al. showed that different histone variants tend to have different post-translational modifications (Hake et al., 2006). However these histones are also dependent on replication and might indicate a potential process of neurogenesis (Marzluff et al., 2002) . Little is known about EE-induced changes in histone expression, thus opening an important field of research to determine the reason behind these results.

Targetome analysis revealed that the main gains in microRNA expression affected neurodevelopmental genes such as the FGF or Wnt signaling. Effects were also observed in transduction signaling upon neuron activation, such as the small GTPases linked to Rho signaling, as previously described as a main effector to cytoskeletal rearrangements (Auer et al., 2011; Nakayama et al., 2000; Tolias et al., 2011). It is in fact not rare to find that the main regulators are found in the intersection (miR-125b-5p and miR-423-3p) with the differential expressed genes dataset (Figure 4.3 D). These microRNA specifically affect genes related with cytoskeletal organization, such as tubulin (*Tubd1*) or Rho members (*Arap3*).

Taking the RNA-seq and microRNA data together, they suggest a close relationship in the main functional changes in cerebral cortex induced by EE exposure. Gene expression reflects an enhanced neuronal activation of associated transcripts, which in turn may trigger morphological changes at(Tsai, 2014) the synapses. This may require appropriate homeostasis of cytoskeletal contributors. The buffering power brought about by the enhanced expression of microRNAs targeting those transcripts may mediate that control mechanism.

The microRNA analysis provided an interesting preamble for follow-up experiments, especially those targeting cytoskeletal organization changes. This could be tested by experiments in cell cultures to better validate the current results.

Preface of ARTICLE 5

Contributions

My contribution in this paper consisted of computational differential analysis of iTRAQ and LC/MS results together with gene ontology (GO) analysis. Concerning the experimental procedures, I took care of the animals, I carried out the pertinent dissections and I performed the protein extraction of the second wave of breedings. I also wrote this paper and produced the figures.

Aliaksei Holik performed the t-Test (two-sided) differential analysis of the first group of animals and the limma-loess analysis for the second group.

Charlotte Hor prepared the samples for the first group of animals.

Júlia Albaigues helped me with the new group of breedings and perform some of the dissections.

Meritxel Pons bred the first group of animals and performed the whole cortex dissections.

Proteomics Facility at the UPF-CRG performed the iTRAQ and LC/MS runs.

Objectives

The main objectives of the present study are:

- To determine the proteomic changes in whole cortex due to environmental enrichment (EE) exposure.
- To compare different proteomic techniques.
- To study how protein levels are related with the transcriptome.

Highlights

- Proteomic changes induced by EE exposure point to a shrinkage of synaptic buttons to rescale the connectome during postnatal development.
- iTRAQ and LC/MS have different sensibility of detection, but can be complementary procedures in proteomic research.
- The proteome and the transcriptome behave differently according to the treatment, pointing to spatiotemporal constrains that affect the biogenesis of new transcripts and proteins.

Acronyms

Environmental Enrichment (EE), Isobaric Tags for Relative and Absolute Quantitation (iTRAQ), Liquid Chromatography-Mass Spectrophotometry (LC-MS), Alzheimer Disease (AD)

Chapter 5
Mouse cortex proteomic changes induced by environmental enrichment: enhancing mitochondrial activity

5.1 Abstract

Environmental enrichment (EE) prompts increased neuronal activity that triggers signaling cascades. Transcriptional and translational programs of learning-associated genes accompanied by an increase in synaptic strength are then induced. The proteome can therefore be considered both as a molecular phenotype resulting from the above changes, and as the effector inducing all the structural, enzymatic, and signaling changes that drive these morphological learning-based changes. Thanks to the increasing power of proteome detection methods in recent years, the molecular phenotypic variations that may occur upon a month of EE exposure in post-natal cerebral mouse cortex are explored. We used two different techniques: iTRAQ and LC-MS in order to assess the sensibility of both procedures. This research reveals important organelles mediating in synaptic function upon the exposure to EE, such as enhanced mitochondrial activity and ribosomal processing. Furthermore, both proteins being up-regulated and down-regulated show elements related with neurotransmission, pointing at a role in experience-dependent synapse scaling.

5.2 Introduction

The power of protein detection has considerably increased during the last decade (Larance and Lamond, 2015). Similar to genomics, proteomics plays a key role in understanding the physical and chemical phenotypic changes that can occur in a system upon a given treatment.

Molecular phenotypes such as the proteome are in direct contact with the surrounding environment. For that reason scientific interest is growing to study how environmental factors interact with the proteome and to elucidate the mechanistic pathways involved in triggering changes in gene regulation. Thus, EE provides an appropriate study setting by triggering a series of motor, sensorial and cognitive components of the brain. These in turn induce neuronal activity prompting learning and memory associated gene expression (Rampon et al., 2000). One of the aims of this study is to specifically identify the main effectors that elicit expression changes. To this end young mice were exposed to one month of EE during postnatal development, followed by comprehensive analysis of epigenome, transcriptome and proteome of treatment and control groups.

Possibly due to its low power of detection, proteomics has not been widely used to study the effect of EE exposure on protein abundance, and less so in the mouse cerebral

cortex. Some studies have focused on other brain regions. For instance, the research of McNair et al. could only identify around 1-2% proteins of the total proteome with altered abundance in rat hippocampus upon 8 weeks of EE exposure. They discovered that a majority of changes affected biogenesis processes, energy metabolism and signal transduction pathways (McNair et al., 2007). Similar pathways have been reported regarding nucleus accumbens upon EE treatment in a cocaine drug context study(Lichti et al., 2014) .

Employing two different mass-spectrophotometry approaches, a labeled method called Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) (Ross et al., 2004; Zieske, 2006) and a label-free strategy (Liquid Chromatography- Mass Spectrophotometry or LC-MS (Lee and Kerns, 1999), the change in protein abundance in post-natal mouse cortex after a month of exposure was investigated.

5.3 Methods: workflow and exploratory analysis

In order to identify global changes in protein abundance induced by EE an initial exploratory analysis was performed. Fold changes in the data produced by the labeled iTRAQ method and the label–free, LC-MS approach were measured (Figure 5.1 A). For iTRAQ two biological replicates consisting of five pooled mouse cortices per condition (10 animals per condition) were used. The same pooling strategy was used for LC-MS but with three biological replicates (15 animals in total per condition).

Although iTRAQ allows a fast and accurate measure of differences between conditions it also reflects batch biases introduced by various factors, i.e. intrinsic biological variability, processing biases, and technical biases (Figure 5.1C). Different methods can be applied in order to remove batch effects, normalize data and obtain differentially abundant proteins. As iTRAQ only reports fold changes, a threshold of discrepancy between the replicates was established. In this way only protein changes that had the same direction of fold change were taken into account. The premise selected to choose the proper threshold of discrepancy involved comparing the fold change conservation between each biological comparison (EE_bio-rep1/CTL_bio-rep1 versus EE_bio-rep2/CTL_bio-rep2). If a similar direction is evident between two samples, the value is closer to 1. To probe the consistency of this methodology a second strategy was followed by assuming that fold changes of both replicates follow a normal distribution. Then, the method tests if fold changes differ significantly from 1 (p =0.05, T-test two sided). For the second approach we use the tool limma(Ritchie et al., 2015) for both correcting the batch effect (Figure 5.4) and for measuring the differential abundance of proteins.

For further details see Methods Chapter.

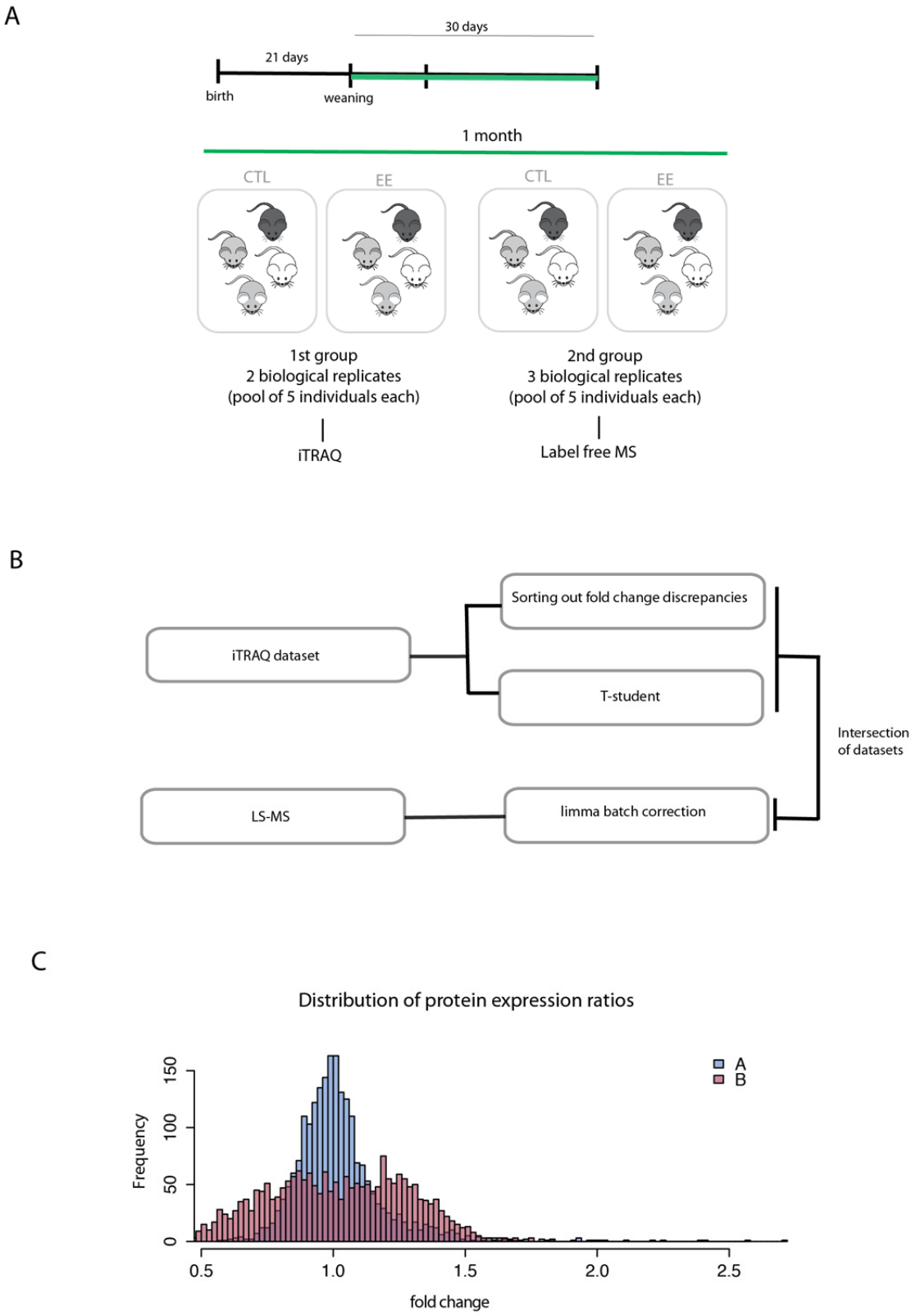


Figure 5.1. Proteome workflow and pipeline
A) Overview of the study. After weaning mice were exposed to a month of EE. The first group of animals bred consisted in two biological replicates per condition, each corresponding to a pooled sample of 5 different mouse cortices, after protein extraction the samples were injected into the Mass Spectrophotometer (MS) following the labeled iTRAQ method. The second group of animals consisted in 3 biological replicates equally pooled that were injected into the MS following a label free method MC/LS. **B)** Computational analysis pipeline. For the first group a discrepancy and a T-student (two sides, p-value<0.05) was conducted and for the second an analysis with limma with loess correction was performed. Finally both methods were intersected. **C)** Distribution frequency of the fold changes per each replicate contrast: A(EE1/CTL1), B(EE2/CTL2). The figure shows a clear different fold distribution for each replicate contrast.

5.4 Results

Using the aforementioned threshold of discrepancy the iTRAQ method detected 2170 proteins of which 447 were altered due to EE exposure. These differences were approximately equally distributed between 211 gains and 236 losses of protein abundance in EE compared to untreated samples. However the second strategy using the T-test in iTRAQ identified only 33 gains and 39 losses. LC-MS detected a total of 1779 proteins, of which 79 were up-regulated and 65 were down-regulated due to EE.

Described below are the results from the iTRAQ method, followed by the LC-MS method.

5.4.1 iTRAQ reveals that EE exposure induces changes in abundance of proteins related to mitochondrial activity, metabolism and neurotransmitter-associated proteins

In order to categorize the global changes in protein abundance induced by EE, a gene ontology analysis was performed on GO biological processes, revealing the main functions affected by exposure to EE. Global changes in mitochondrial activity, metabolic processes and neurotransmission were found. These molecular events are well interconnected, as ion homeostasis and energy in the form of ATP and GTP boost synaptic function (Vos et al., 2010).

GO analysis reveals that changes due to EE in both protein gains and losses share similar terms (Figure 5.2A). Hence a GO network was generated using ClueGO and Cluepedia(Bindea et al., 2013; Bindea et al., 2009) to try to determine interesting hubs (Figure 5.2B). For instance, gains seem to have more protein contributors in metabolic processes by elements such as small molecule catabolic process (p = 92.0E-6), fatty acid oxidation (p = 2.9E-3) or coenzyme metabolic process (p = 510.0E-6, Figure 5.2A). This is also illustrated in Figure 5.2.B where the color of the nodes and the edges shows a tendency towards gains (in blue). In this particular case, the nodes are more enriched by up-regulated proteins that promote a higher activity in metabolic pathways, with gained contributors such as the enolase Eno1 and the Etfdh (Figure 5.2B). Enolases are enzymes important for gluconeogenesis, pointing to greater production of glucose and therefore energy. Moreover it has been shown that they have an important anti-apoptotic role in the context of Alzheimer Disease (AD)(Butterfield and Lange, 2009). Etfdh is part of the respiratory chain reaction in the mitochondria. It has been reported that *Etfdh* mutants in zebrafish (called it *xavier*) lead to fatty acid disruption and mitochondrial dysfunction inducing an altered neurogenesis(Song et al., 2009). Next, there is a global contribution of proteins related with synaptic function and activity, especially enriched by terms both in gains and losses related to vesicle docking and neurotransmission. It seems that there is a slightly higher fraction of up-regulated proteins in terms related with docking the vesicles (p =4.3E-3) but a lower fraction in the exocytosis process (p = 120.0E-6), pointing to a general loss of neurotransmission (Figure 5.2 B, vesicle docking cluster). Concerning the vesicle localization we found gained proteins belonging to the SNARE complex, such as: Snap25, Napb, Napa and Clpx1. Regarding exocytosis proteins that are down-regulated we found: Exoc4, Exoc8, Nrxn1, Rab15, Stx1a, Stx1b and Scfd1. Moreover there is an enrichment of proteins with decreased abundance related with negative regulation of cytoskeletal organization (p = 11.0E-3). Here we find proteins such as Pick1, Add1, Add2, Coro1a, Hdac6,

Mtap1a and Capzb. Interestingly Pick1 regulates the trafficking and internalization of AMPA receptors (Xu et al., 2014a). It has also been shown that loss of function of Hdac6 ameliorates cognitive deficits in a mouse model of AD (Govindarajan et al., 2013). In addition, we find telomere maintenance by a cluster of chaperones that are both gained (Cct2 and Cct8) and lost (Cct3, Cct4, Cct5 and Cct6a) and that may be important in telomerase biogenesis(Freund et al., 2014). Finally, we can find only one group exclusively with down-regulated proteins. These correspond to transmembrane receptors in the mitochondria, such as Atp6v1c1, which are important for the homeostasis of ionic balance (Vos et al., 2010).

Globally EE induces changes in metabolic processes, especially targeting energy production in the mitochondria. Also, due to the correct tuning of synapsis number during postnatal development, the increased neuronal activity induced by EE exposure may trigger the gaining of strength of some synapses. Meanwhile other synapses are being lost. An important discovery is the involvement of telomere maintenance. The research of Romano et al. highlights the importance of this process during the induction of environmental stress in yeast (Romano et al., 2013). This points to other environmental factors such as EE that may induce the positive control of the telomeres.

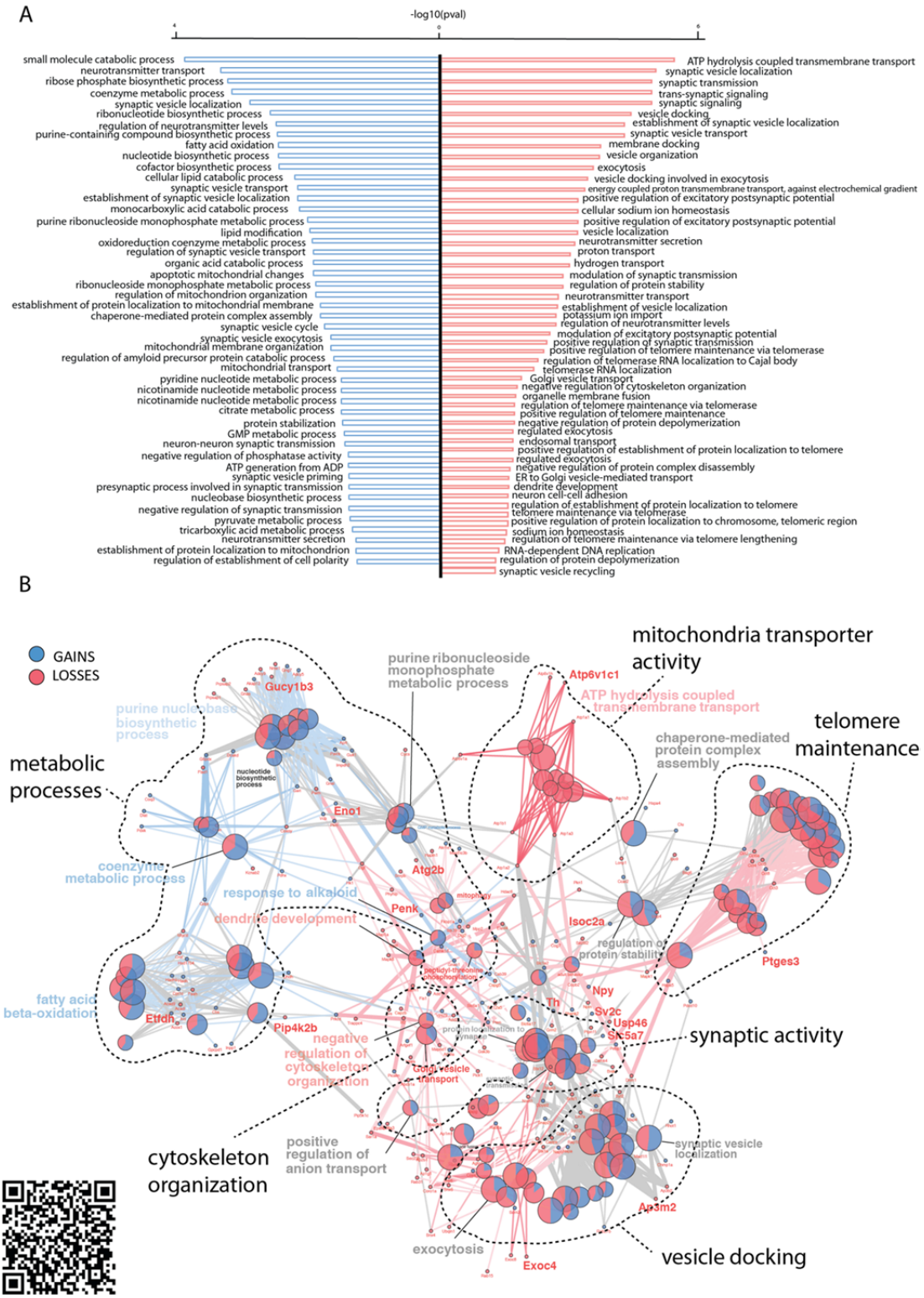


Figure 5.2. Gene ontology analysis of iTRAQ results
A) Biological processes GO analysis plot. Blue bars represent gains and red losses. Selected GO terms are ranked by the score given by -log10(p-value) **B)** Biological processes network of both gains (blue) and losses (red). Proteins clustered for each term are also represented in blue if gained and red is lost higlighting the top changes. QR code gives access to complementary analysis.

5.4.2 iTRAQ: top protein abundance changes induced by EE

The class of proteins with the highest increase in abundance within EE consists of neuropeptides such as the Neuropeptide Y ($\overline{FC}$= 2.3, Npy) and Proenkephalin ($\overline{FC}$= 1.94, Penk) (Figure 5.3). Npy is normally highly expressed in the limbic and cortical areas of the forebrain(Allen et al., 1983). Npy shows increased expression in environmentally enriched animals(Ragu Varman and Rajan, 2015) and plays an important role in resilience to stress factors (Cohen et al., 2012; Hendriksen et al., 2012; Sah and Geracioti, 2013). Penk is widely expressed in different tissues(Keshet et al., 1989), and is thought to play roles in tissue protection and pain control by mimicking opioid release (Denning et al., 2008; Konig et al., 1996). Although its expression is not exclusive in neuronal populations it seems that it behaves as a neuronal activity-dependent protein. This is evidenced by its increased expression regulated by glutamate agonists release (Just et al., 1998). These results highlight the role of EE triggering anxiolytic effects against stress factors, also supported by recent publications (Reichmann et al., 2016).

The second most important category of up-regulated proteins are transporters that are especially enriched in the mitochondria, playing important roles in cholinergic and glutamatergic circuitry. The most abundant of these was found to be the mitochondrial Solute Carrier Family 25, Member 40 ($\overline{FC}$=2.318, Slc25a40), which was recently reported as being important for neuronal survival and synaptic transmission in *Drosophila* (Slabbaert et al., 2016). Also found to be in high abundance was the Translocase of Inner Mitochondrial Membrane 10 ($\overline{FC}$=1,23, Timm10) protein which has been reported to be down-regulated in Schizophrenia (Arion et al., 2015) and Alzheimer's disease (Liang et al., 2008), as well as the Sec14-like Lipid Binding 2 ($\overline{FC}$= 1.538, Sec14l2) which is a phospholipid transfer protein that mediates vitamin E processing (Zingg, 2015). Interestingly Sec14l2 has been reported as a regulator of phosphatidylinositol 3-kinase activity and shows low GTPase activity that is important in signaling pathways upon neuronal activation (Habermehl et al., 2005; Kempná et al., 2003). Notably its partner, Sec14l1, has been reported to interact with cholinergic transporters (Ribeiro et al., 2007). Moreover, the amino acid transporter ADP-Ribosylation-like Factor 6-Interacting protein 5 ($\overline{FC}$=1.32, Arl6ip5) likely interacts with EAAC1, a protein regulating glutamate transport (Lin et al., 2001). Furthermore the Sodium-carrier choline transporter ($\overline{FC}$=1.29, Slc5a7) is involved in the re-uptake of the precursor choline from the synaptic space to the presynaptic region (Dobransky and Rylett, 2005) and plays a key role in cholinergic circuitry in cerebral cortex (Berry et al., 2015). In summary, these protein transporters are important for glutamatergic and cholinergic transmission, highlighting the importance of functional ion homeostasis at the synaptic buttons mediated in part by the mitochondria.

As described earlier, metabolic processes play an important part in the up- and down-regulation of proteins upon EE exposure. We find important reductases such as the Tyrosine 3-monooxygenase ($\overline{FC}$=1.634, Th) involved in synthesis of catecholamines such as dopamine, noradrenaline, and adrenalin. Th is essential for development and mice unable to synthesize dopamine are hypoactive (Zhou and Palmiter, 1995). The second most up-regulated reductase is the Aldo-keto Reductase Family 1, Member E2 ($\overline{FC}$=1.5, Akr1e2) involved in reduction of 1,5-anhydro- D-fructose to 1,5-anhydro-D-glucitol indicating a faster mono-saccharide processing due to the enhanced activity,

supporting its role in gluconeogenesis identified earlier in the GO enrichment analysis. Finally the increased abundance of Electron Transfer Flavoprotein Dehydrogenase ($\overline{FC}$=1.29, Etfdh) points to the same direction, being important for metabolic process and a proper respiratory chain reaction (Song et al., 2009). Related with proper calcium homeostasis mediated by the mitochondria, we found a gain in the Tumor Suppressor Candidate 2 Tusc2 ($\overline{FC}$=1.35, Tusc2)(Uzhachenko et al., 2014).

Other up-regulated proteins due to EE treatment that are important for signaling pathways are cAMP and the cAMP-inhibited cGMP 3', 5'-cyclic phosphodiesterase 10A ($\overline{FC}$=1.34, Pde10a). This protein eliminates cAMP- and cGMP-mediated intracellular signaling important for synaptic transducing signaling in dopaminergic neurons at the striatum (Russwurm et al., 2015). Even though it is highly expressed at the striatum it can be found in other regions including the cerebral cortex(Coskran et al., 2006). As well as that, enrichment of Heteromeric G proteins such as Gng3 and Gng5 can be activated upon the activation of the G-coupled receptors playing a key role in signaling transduction (Betty et al., 1998; Gainetdinov et al., 2004).

Concerning the proteins with reduced abundance upon EE exposure related with metabolic processes, we found a remarkable loss of 3-Hydroxyacyl-CoA dehydratase 3 ($\overline{FC}$= 0.677, Hacd3) involved in the dehydration step of very long-chain fatty acid (VLCFA) synthesis (Ikeda et al., 2008). Other losses included the Ca^{2+}-binding protein Calumenin ($\overline{FC}$= 0.678, Calu) which is involved in protein folding (Vasiljevic et al., 2012) and Autophagy S. ceriveseae Homolog 2 ($\overline{FC}$=0.738, Atg2b) which is important for folate and lipid droplets regulation(Sun et al., 2016).

Notably, as previously explained, there is a loss in exocytosis proteins. Among the top down-regulated proteins is the loss of Exocyst Complex Component 4 ($\overline{FC}$=0.696, Exoc4) involved in vesicle trafficking at neuronal synapses (Hsu et al., 1996). Furthermore we found the Ubiquitin-specific peptidase 46 ($\overline{FC}$= 0.705, Usp46) to be important in the regulation of AMPA receptors (Huo et al., 2015) and the Vacuolar type ATPase G2 ($\overline{FC}$=0.794, Atp6v1g2) that is important for the accumulation of neurotransmitters into secretory vesicles (Kawamura et al., 2015).

Other important proteins found to be down-regulated are post-translational modifiers such as the previously mentioned histone deacetylase ($\overline{FC}$= 0.776, Hdac6) and the arginine methyltranferase ($\overline{FC}$= 0.783, Prmt8) that are normally highly expressed in brain tissues (Lee et al., 2005). Moreover, other proteins involved in plasticity are the loss of Guanylate Cyclase 1 Beta-3 ($\overline{FC}$= 0.765, Gucy1b3), a main receptor for nitric oxide (NO)(Sagi et al., 2014).

Overall, gains show a higher enrichment for neurotransmission based on the neuropeptides, Npy and Penk. However the top losses show important proteins for controlling neurotransmittor release. Both in gains and losses we found an enrichment of proteins related with metabolic process, paying special attention to the gain of gluconeogenesis, enhancing energy supply to the system. Moreover these findings highlight the importance of the mitochondria, which have a substantial role not only in providing energy but also controlling ion homeostasis for normal synaptic function. Losses also show some chromatin posttranslational modifiers especially the histone deacetylase Hdac6.

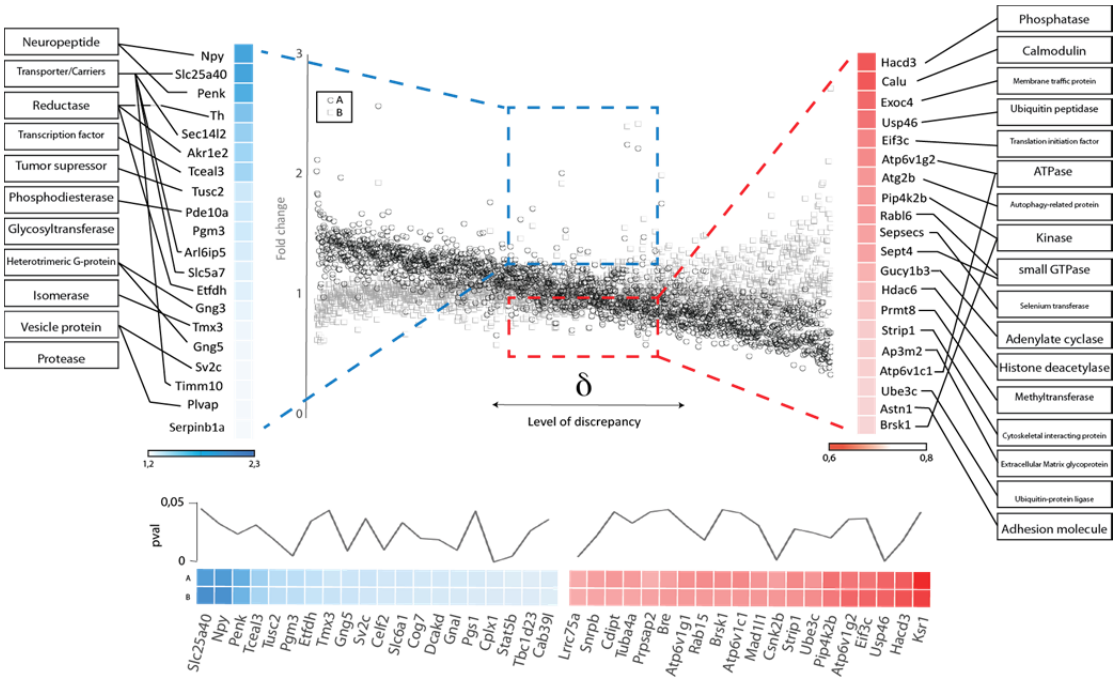


Figure 5.3. Top proteomic changes induced by EE
Graph shows the ratio of each biological replicate (EE/CTL) and sorted by discrepancy. Central values are the most similar (values around 1+/- 0.1). Selected gains and losses by this method are selected and plotted in the extremes with its own Panther protein classification (www.pantherdb.org,) (Thomas et al., 2003)

5.4.3 Label-free mass spectrophotometry overview

In order to validate the results obtained by the iTRAQ method a second mass spectrophotometry injection was conducted using a label-free method with three new replicates. In this case results show a clear bias (batch effect) resulting in a general under-representation of proteins in EE samples (Figure 5.4 A and B). Although losses can represent a biological result of EE, it is clear that this lower distribution of values in the experimental condition can be explained by technical problems during the experimental procedure (e.g. miscalculation of protein concentration before injection to the MS). Based on the distribution of fold-change values and the PCA analysis (Figure 5.4 C and D) it is probable that a batch effect occurred. In order to remove the batch effect that skewed the distribution of fold changes between EE and non-EE samples, we used limma (Ritchie et al., 2015) with a loess regression normalization method (Smyth and Speed, 2003) .

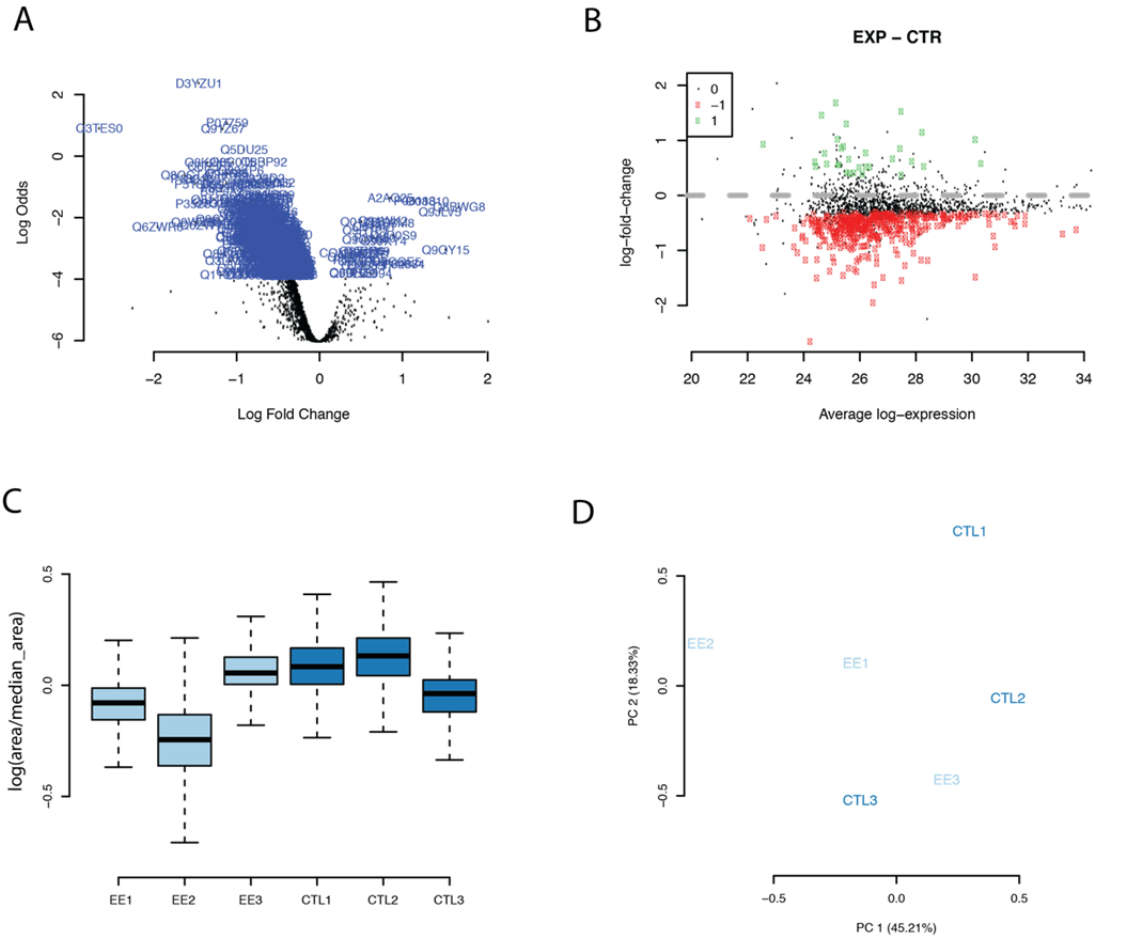


Figure 5.4 LC/MS data overview
A) Volcano plot showing the logOdds vs logFC. The majority of the fold changes due to EE fall into the losses. **B)** Plot showing the differential analysis distribution of logFC vs log(expression) **C)** Boxplot of the log ratio of the area signal and the median. **D)** PCA plot of all area values of each sample.

5.4.4 Label free mass spectrophotometry reveals a retraction in synaptic plasticity

LC/MS gene ontology analysis reveals some terms shared with the iTRAQ dataset that are related with metabolic and synaptic transmission elements. However interestingly it reveals other key factors for normal synaptic function. For example, gains revealed signaling pathways such as cAMP-dependent protein kinase complex (p = 74.0E-6), Hedgehog signaling by the negative regulation of smoothened (p = 32.0E-3), both sharing protein kinases such as Prkaca, Prkacb and Prkar1b (Figure 5.6A). Notably there are terms such as small ribosomal subunit (p = 10.0E-3) with proteins such as Gja1, Rps12, Rps4x and Rps5. This term denotes the substantial role of ribosomes in protein changes at the synapsis (Schuman et al., 2006). In addition the appearance of the endoplasmic reticulum term (p = 18.0E-3) is of interest, as it has also been highlighted in recent research (Ramirez and Couve, 2011).

The most important aspect of the down-regulated proteins is a clear loss of synaptic function and neurotransmission. There is an enrichment of NMDA (p = 280.0E-6), AMPA (p = 37.0E-3) and GABA (p = 36.0E-3) receptor loss, pointing to retraction of synapses or long-term depression processes (LTD, reviewed here (Collingridge et al., 2010)). This is of relevance during post-natal development, as the extra amount of synapses created during early development needs to be depleted while the animals are

gaining experience. This process of fine-tuning, or synaptic scaling, is a key step of synaptic plasticity (Turrigiano, 2008) and it has been demonstrated in the visual cortex (Wiesel, 1982) and the neuromuscular junction (Sanes and Lichtman, 1999). Moreover there is a loss of proteasome-related proteins (p = 14.0E-3), specifically for ubiquitin-mediated protein degradation at the synapse(Tsai, 2014). Signaling pathways are also enriched in losses, such as the nerve growth factor (p = 49.0E-3), gonadotropin (p = 6.9E-3) and ephrin signaling (p = 46.0E-3).

Regarding the top gain changes we can also find proteasomal-related proteins that were previously found in the GO losses (Figure 5.6B). We find Cul3 that can induce targets to be degraded by the organelle (Zhang et al., 2004) and also the proteasome subunit Psmb3. On top of that gains also show enrichment for neurotransmitter transporters such as Slc6a11 that transport GABA (Son et al., 2014), or Slc17a6 that mediates the uptake of glutamate into synaptic vesicles at the presynaptic terminal (Wallen-Mackenzie et al., 2010). Notably we find a histone linker gain such as Hist1h1b, pointing to a potential effect in chromatin remodeling or compaction. Furthermore there is enrichment of cell adhesion molecules such as Cadm1 and signaling pathways related to cAMP signaling by the enrichment of the protein Prkar1b and the Ras superfamily members, Rab2a and Rhog.

Concerning the top protein losses it is important to highlight the enrichment of postsynaptic proteins, e.g. Shank1 adapter, responsible for the proper clustering of receptors at the postsynaptic membrane (Hayashi et al., 2009). Also found is an enrichment of the NMDA receptor mediating synaptic plasticity Grin1 protein and the kinase Camk2b. Equally to gains we find proteasome contributions, such as Psmd3 subunit.

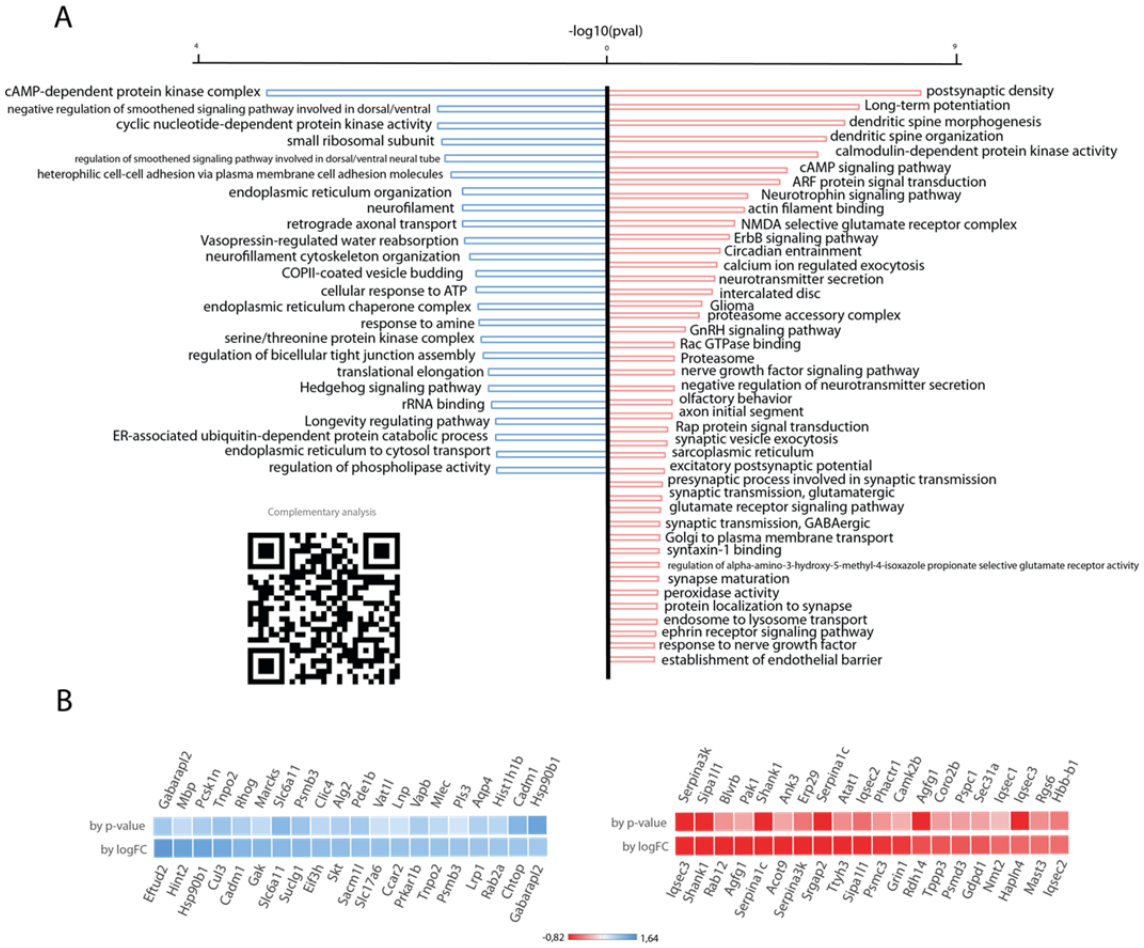


Figure 5.6 Gene ontology analysis and top changes in protein abundance induced by EE in LC/MS
A) Gene ontology analysis of biological process, cellular components, molecular function and KEGG pathways showing elements with p-value<0.05. QR gives access to complementary analysis **B)** Heatmap of the top changes in LC/MS dataset sorted by logFC and by p-value.

5.4.5 Label free mass spectrophotometry and iTRAQ method: two different techniques with different detection power

In order to know if using different methods for protein detection could mutually validate the results we first interrogate the power of detection of each technique by intersecting both datasets (Figure 5.7 A). As a result the iTRAQ strategy was able to detect a higher number of proteins (2170) compared to the LC/MS method (1779). The number of proteins mutually detected by both methods was 1370 proteins (~63% of the total iTRAQ dataset; ~77% of LC/MS dataset). These results are in agreement with previous publications showing similar sensibility for both methods. For instance, the research of Sandberg et al. found similar observation when intersecting both datasets. They found that around 79% of the proteins detected by LC/MS were shared with the iTRAQ method, meanwhile only 60% of iTRAQ results were also found with the label free strategy (Sandberg et al., 2014). This fact allows to partly claim that they have different detection range and affirm that iTRAQ has greater sensitivity compared to the label-free methodology. This could explain that given a small number of protein level changes due to EE exposure we could get an even lesser number of mutually shared elements.

After the unwanted variation removal by the limma-loess pipeline for the LC/MS dataset, results showed 79 gains and 65 losses in protein abundance due to EE exposure.

However the differentially abundant proteins that were detected show an extremely low intersection with the results from iTRAQ. Only five proteins were mutually lost: Agfg1 (G-protein modulator), Omg (extracellular matrix receptor), Coro1a, Coro2b (non-motor binding protein) and Gdpd1 (phosphodiesterase); and four were mutually gained: Hnrnpa1 (nuclear ribonucleoprotein), Ccar2 (calmodulin), Hrsp12 (ribonuclease) and Th (tyrosine 3-monooxygenase) (Figure 5.7 B).

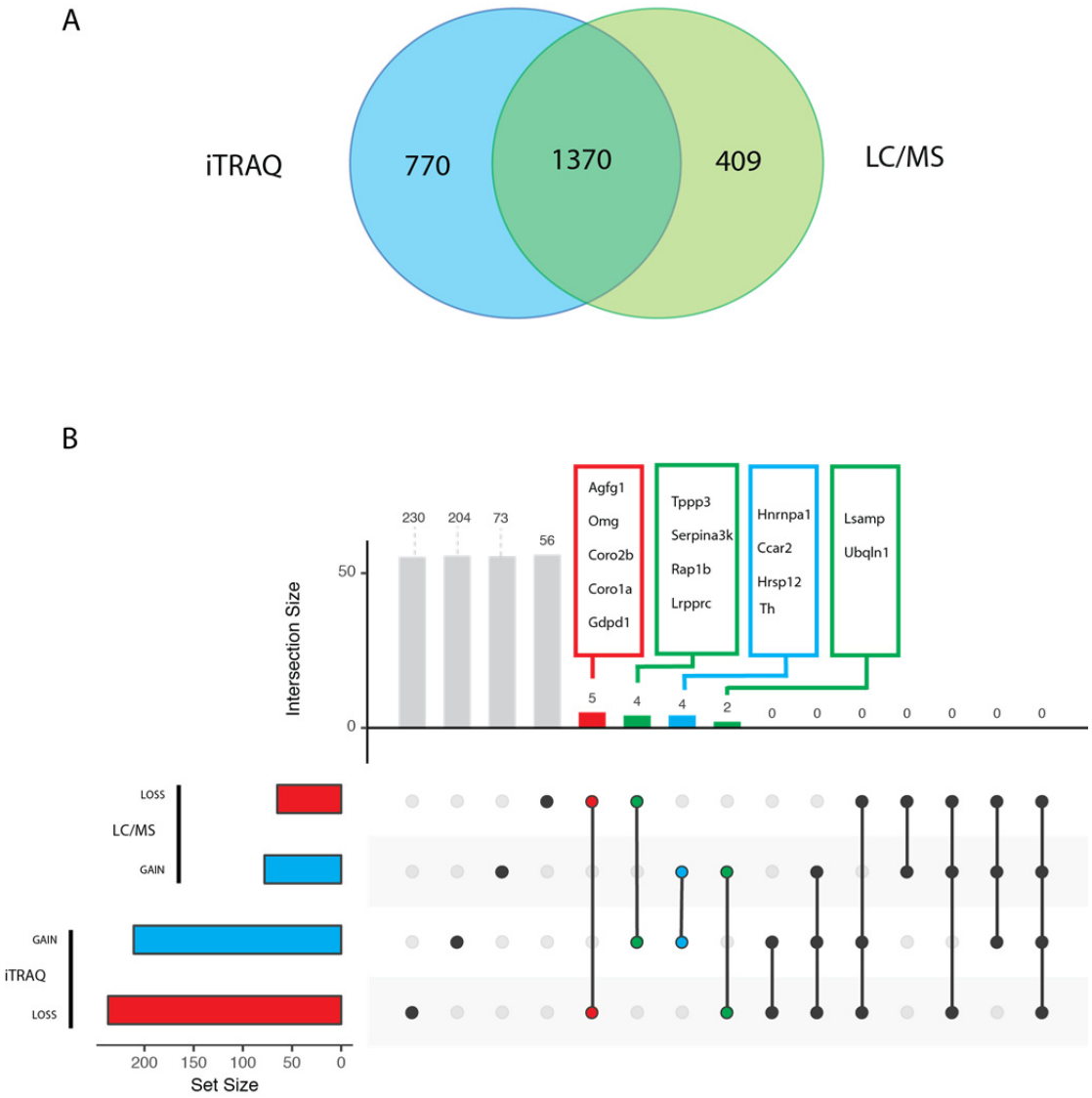


Figure 5.7 Label free and iTRAQ intersection analysis
A) Venn diagram showing how the total number of proteins detected by each method intersects with each other. **B)** Intersection plot of the differential protein analysis for each method. Left plot illustrates the total number of proteins in each dataset. Shared proteins between dataset are represented by dots linked with a black bar (blue if gains, red if losses, green if it is a discrepancy).

5.4.6 Proteome and RNA expression

When compared with RNA-seq data, we found almost no intersection of differential expressed genes and differentially abundant proteins (DAPs) (Figure 5.7). Only the resulting output of the iTRAQ procedure shows some shared terms. Specifically 10 proteins commonly found as gained due to EE by RNAseq and iTRAQ analysis are

Nrcam (neuronal cell adhesion), Hk1 (hexokinase-1), Napb (membrane trafficking protein), Vcp (endoplasmic reticulum ATPase), Grb2 (growth factor receptor-bound protein 2), Tsc22d3 (transcription factor), Eno1 (enolase, involved in gluconeogenesis), Slc6a17 (cation transporter), Sv2c (synaptic vesicle) and Dpp6 (protease). And mutually lost: Gng13 (heterotrimeric G protein), Pdia6 (isomerase), Snx6 (membrane trafficking regulatory protein), Slc39a12 (transporter), Gnas (heterotrimeric G protein), Pisd (decarboxylase), Hapln1 (extracellular matrix glycoprotein) and Acadl (mitochondrial dehydrogenae). Although not much intersection of shared up- and down-regulated proteins is evident, the importance of the organelles found to be involved, namely the mitochondria or the endoplasmic reticulum, is highlighted by some of the proteins found to be shared.

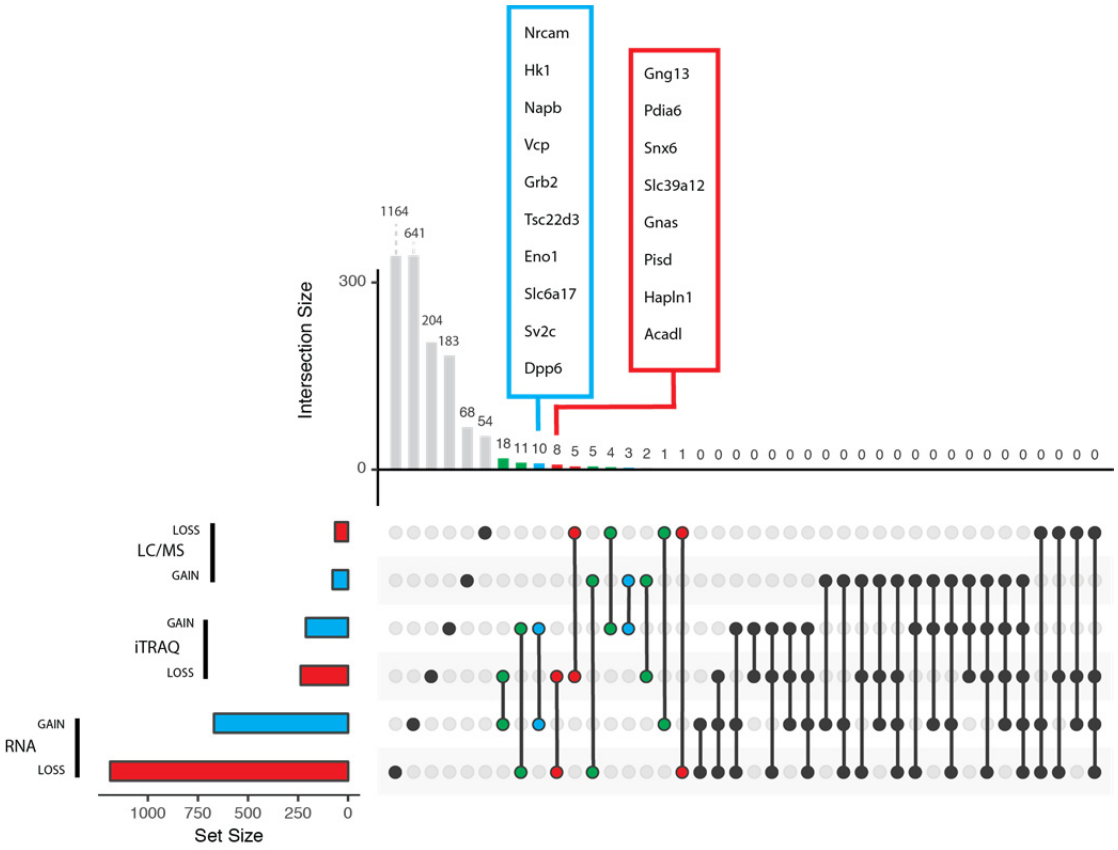


Figure 5.8 Relationship between the proteome and the transcriptome
Graph showing the intersection between the different proteomic experimental methods and RNA-seq expression. Left plot shows size of each dataset. Intersections are highlighted by connecting wires, blue: mutually gained; red: mutually lost; green denotes a discrepancy. Upper plot shows the size of the intersection and the mutually protein names shared in gains and losses.

5.4.7 Discussion

This work reflects important concerns about different experimental techniques in proteomic research. In light of these current results it can be argued that both techniques are complementary in protein discovery. They both point to a shrinkage of synapses due to a physiological need of synaptic scaling during early experience-dependent neuronal activity (Flavell and Greenberg, 2008; Turrigiano, 2008). For example, in auditory discrimination tasks in mice where fine tuning of the synaptic connectome is also needed (Kahne et al., 2012; Kahne et al., 2016). The same happens in spatial learning processes that induce a similar effect in rat hippocampus (Henninger et al., 2007). However complementary results from each of them highlight the importance of

some organelles and functions such as the mitochondria(Vos et al., 2010), the translational machinery (Hernandez et al., 2015; Schuman et al., 2006), the endoplasmic reticulum (Ramirez and Couve, 2011) and the proteasome (Tsai, 2014) in synaptic function.

In the context of EE, McNair et al. also compared their proteomics findings with the transcript expression analysis of Rampon et al. They found little intersection, supporting our current results (McNair et al., 2007; Rampon et al., 2000). This unexpected result could be explained by different biological dynamics of transcription and translation. Also spatiotemporal constrains might play a role, as translation and degradation dynamics at the synapse takes place locally where the first signaling cascade steps occurs, having an intimate space of action-reaction. Whereas transcription needs for the transducing signal to travel to the nucleus, trigger the mechanism for the transcript production. Later the transcript needs to travel back again to the synaptic button. We further found that transcriptome analysis (see previous chapter) is in much higher agreement with epigenome and chromatin accessibility analysis with respect to enriched GO terms that are biased to an increase of synaptic plasticity. However, the proteome results seem to be contradictory, pointing to a retraction of the synapses. Indeed, we should see both processes as complementary, as both are taking place due to the on going fine-tuning. The transcripts that are related with learning and memory processes might be created to travel to specific synapsis to gain strength. Meanwhile the overall synaptosome is being lost due to the removal of silent synapses. This is a well-known physiological step during postnatal development. Because of that, our results clearly indicate that EE enhances the induction of synaptic scaling.

Moreover, it is also important to take into account that we have proteomic signals from the whole cortex. Thus the results are likely influenced by cell type heterogeneity and cell type composition of the cortex samples of cases and controls. It is therefore important to establish whether these current results could be extended to specific neuronal populations by doing FAC-sorting or by studying synaptosome preparations.

Chapter 6
Discussion

The main aim of this project was to study not only the changes that EE could trigger to the different layers of mouse cortex cell regulation, but also to create a broad spectrum of epigenomic, transcriptomic and proteomic data to study how the different layers of information could be integrated. Hence, our primary objective was to better understand common dynamics and changes conserved across the different datasets. To achieve our goals we had to tackle diverse problems and to reach multiple important milestones. First, we had to choose from a large number of sequencing protocols available to date for the analysis of the epigenome, each of them having different power of detection. Second, the brain harbors many different cell types, each having a different epigenomic state. And finally, the whole cortex is not a homogenous tissue; it has different structures, with different cell type composition and function. All this points are important to take into considerations during data analysis in order to improve our understanding of the brain regulome and to advance the state-of-the-art methods in Neuroscience research.

Despite these issues and challenges in this work we have tried to overcome some of the difficulties by de-convoluting the studied signals such as the epigenome, the chromatin accessibility or even the 3D conformation (work in progress). To this end we have studied single cell types, for which we had to develop FAC-sorting and sequencing protocols able to deal with low numbers of cells. We have already achieved cell-type specific ChIP-seq profiles of histones and open chromatin ATAC-seq in sorted neuronal populations. To complete the full cell-type specific profile, future directions of this research will focus to produce data for the transcriptome, the proteome and chromatin conformation from sorted neuronal populations. Moreover, specific areas of the cerebral cortex could be selected to better understand gene regulation at the level of sensorial, motor and cognitive perception and to reduce the complexity of the data.

On top of that we have performed the first attempts to computationally integrate the results taking into account a Fisher method like procedure in which significant changes could be integrated. We have created the basement for a tool that we called Multiple Signal Integrator (MSI, work in progress). Despite some success concerning the integration of various epigenomic marks several issues remain, i.e. discordant signals from different marks or disagreement with expression data. More testing is required probably using datasets with stronger differences than induced by EE, e.g. two different tissue types.

This thesis underpins some important facts to be consider in future research:

1) The need of cell-type specific epigenetic studies to avoid a complex mix of signals from different cell types with varying epigenomic state.
2) A need for comprehensive computational methods to integrate complex, genome-wide epigenomic and transcriptomic data tracks.
3) Establish better strategies to find the correct balance between synthetic and disperse information of NGS results. My current work has tried to look for different methods to weight the results in order to provide the most interesting changes induced by the treatment. The general pipeline has been looking firstly

to the gene ontology to obtain a general overview of the main functions changing, followed by the ranking of the top results. The main issue with this strategy is that GO analysis is sometimes unspecific and also top changes are not the most biological relevant. Regarding this problem, I am currently working on using other strategies in order to extract more valuable information using different methods, e.g. looking for interaction pathways, by cell localization, etc.

4) A better definition of the molecular phenotypes is needed. The reason behind that is the weak correlation between the proteome and transcriptome. A large number of publications use the transcriptome as a molecular readout but it is obvious that transcripts are not the main effectors in function. Thus, empowering the research in proteomics field is necessary. Moreover, finding accurate ways to study other regulome modulators, such as non-coding elements, ribosome protein production, and protein degradation by the proteasome.

6.1 Intersecting cerebral gene regulome datasets

Up to now we have been intersecting each epigenomic or proteomic dataset with differentially expressed genes obtained by RNA-seq. Hence, we have used expression as the main molecular phenotypic readout. This generates several issues, as a) the level of transcripts is influenced by transcription, post-processing and degradation efficiency and does not always correlate well with epigenomic or proteomic measures, b) RNA-seq data is often as noisy as any of the other quantitative sequencing types and c) other relationships that might important are ignored, e.g. correlations between histone marks. In this section I will provide a brief overview of the gene accessibility and epigenome intersection together with a global perspective view of the all datasets intersection.

6.1.1 Gene activity: the epigenome and chromatin accessibility

Both chromatin accessibility and gene body marks are related with gene activity (Buenrostro et al., 2013; Dong and Weng, 2013; Huff et al., 2010; Lepoivre et al., 2013; Tsompana and Buck, 2014) As discussed in Chapter 1 and 2, chromatin accessibility might be more related with transcription factor binding activity in the promoter or enhancers of genes whereas gene body marks with the transcription process (Dong and Weng, 2013; Jonkers and Lis, 2015; Lepoivre et al., 2013). However, by intersecting chromatin accessibility results with activating histone marks we found some correlation (Figure 6.1B).

Around 12% of gains found in ATAC-seq both in promoters and enhancers, are also found in both gains of H3K36me3 and H3K79me2 (blue arrows, Figure 6.1B). Looking briefly how these genes intersect with RNA seq (Figure 6.1C and D), a small part could be translated into up-regulated transcripts. In addition a small number of ATAC-seq losses is also found in H3K79me2 down-regulated genes (red arrows, Figure 6.1B). This result is interesting because it reproduces the small number of losses found in ATAC-seq data. However, we can also find a big number of discrepancies (green arrows Figure 6.1B). This result implies that differential promoter accessibility and epigenomic features found in the gene body profile might have different dynamic features. A possible explanation is that genes could be ready for transcription, being in a kind of resting state. Thus, promoters show higher accessibility for transcription factors to bind, but the transcription process is not taking place or is slowed down due the effect

of the treatment. Other possible explanations have been published recently, e.g. the absence of canonical epigenomic marks in developmentally regulated genes (Perez-Lluch et al., 2015).

Similarly, we address the question of how ATAC-seq data from total neurons and pyramidal neurons is related with the neuronal gene body mark H3K79me2 (Figure 6.1A). In both cases there is small intersection and the biggest corresponds to both NeuN and pyramidal neuron gains and H3K79me2 losses, pointing again a mechanism of transcription resting phase.

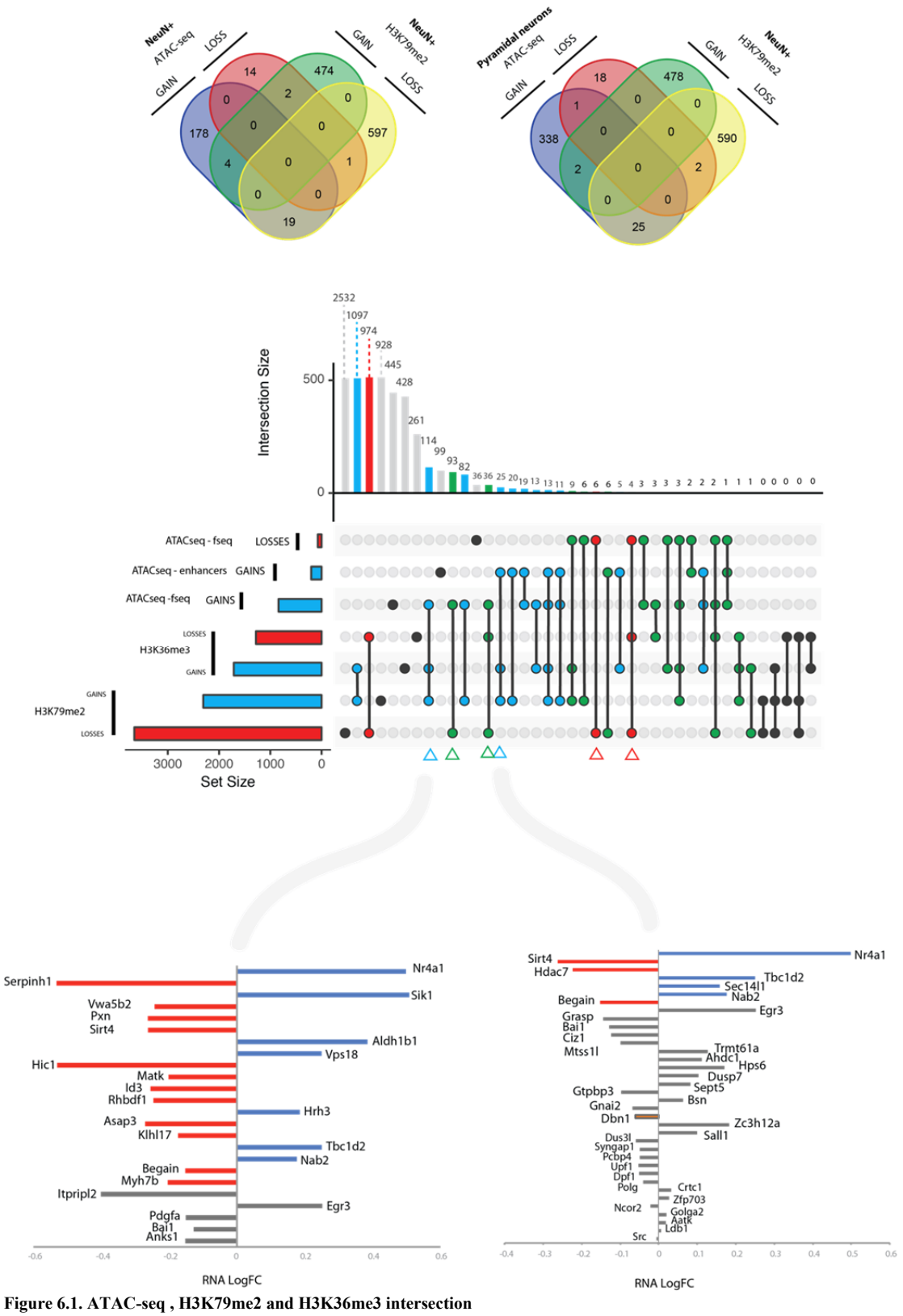


Figure 6.1. ATAC-seq , H3K79me2 and H3K36me3 intersection
A) Intersection between NeuN+ and pyramidal neurons with NeuN+ H3K79me2 dataset. **B)** Interesetion of whole cortex gene accessibility and whole cortex broad gene bony marks. The plot shows the intersection of ATAC-seq fseq DAS, ATAC-seq enhancer prediction DAS, H3K79me2 and H3K36me3. Left bar plot shows the set size of gains (blue bar) and losses (red bar) per dataset. Dots linked by bar represents common genes found in the datastets. Blue dots are gains, red losses, and green represents a discrepancy in the direction of the fold change. Upper bar plot show the size of the intersection with the corresponding colors if is a common gain (blue), loss (red) and green (discrepancy). **C)** and **D)** RNA-seq fold change values for selected intersection (blue arrows).

6.1.2 A goblal perspective of integration

As the integration of multi-dimensional omics data by computational methods is a field of active development in the lab, this analysis has given us a first glimpse on how the epigenomic marks might be related (bearing in mind the caveat that statistically significant shared terms might not have biological relevance). Hence, in order to have a full perspective of how all differences induced by EE might be related, we intersect here whole cortex datasets (with some simplifications: for ATAC-seq only fseq results, for Proteome only iTRAQ results, Figure 6.2). The plot shows that effectively H3K79me2 and H3K36me3 are the marks most closely related as previously discussed in Chapter 1. Moreover the differential activity induced by EE in these marks is most correlated with altered RNA expression (red arrow and blue arrow B). Gains in chromatin accessibility (ATAC-seq) are second most correlated with differential expression (blue arrow C). However, accessibility and epigenetic changes also show a high level of discrepancy (green arrows), as described in previous section.

Interestingly we find that there are 67 targets of up-regulated microRNAs that are shared with increased coverage of H3K79me2 and H3K36me3 due EE exposure (blue arrow A). This intersection could elicit some link between microRNA expression and histones modifications.

Finally as previously discussed in Chapter 4, the proteome does not show much intersection with any of the other datasets, despite some losses that are commonly shared with down-regulated genes in H3K79me2 dataset.

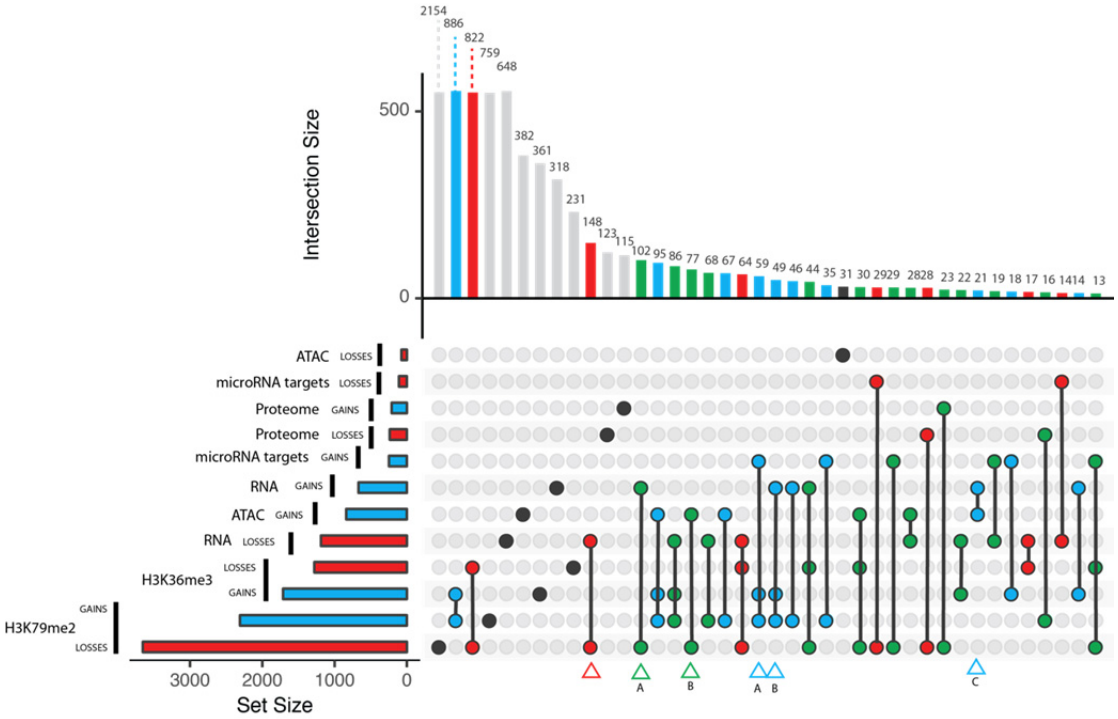


Figure 6.2. A global perspective
Global intersection of whole cortex datasets.

6.2 A model to explain discrepancies.

We have found a considerable amount of discrepancies that showed interesting terms related with known effects triggered by EE. Then, it arises the question how it is possible that biological relevant elements can be discrepant between different datasets, when expecting the contrary results.

The reality might be complex enough hiding important pieces of information modulated by other important biological effectors not studied here. For instance, in a situation of cell heterogeneity, discrepancies between epigenetic and expression data could be explained by a scenario in which a small number of cells are enough to induce changes in the total transcriptome output. Hence, the levels of a transcript in total tissue sample might be only marginally influenced by the active transcription in a specific cell type of that tissue. As illustrated by the Figure 6.3 a pure population of cells is producing a particular transcript showing high coverage for gene body active marks such as H3K79me2. However, when looking at the results for that particular histone mark in total tissue we might observe a down-regulation if the rest of the cell-types have a strong down-regulation of that particular gene (Figure 6.3 A). Thus, as they are not producing the transcript it is possible that the effect of the treatment in the pure cell population might be high enough to see up-regulated results in the total transcriptomic data (Figure 6.3 B).

In light of previous claim, histone modifications might work as effectors modulating the process of transcription and having then, an impact in the spatiotemporal constrains that might occur upon the treatment. In other words the process of transcription as measured by histone marks is not the same as the total yield of transcripts and might not always show a direct relationship. But other scenarios can also occur if the epigenomic state of different cell types changes over time, adding additional complexity. This can potentially mask then real effects of a given treatment. Hence, remarking the importance of developing methodologies to sort cell-type specific cells in epigenetic studies.

On top of that, other molecular effectors can modulate the picture. For example the total yield of a particular transcript can be influenced by the action of non-coding elements such as microRNAs (Figure 6.3.B). On top of that the proteome depends on other mechanisms of regulation, such as the rate of production of proteins by the ribosomes (translation efficiency) and protein degradation by the proteasome (Figure 6.3.C). These two effectors also modulate the final picture of the total yield of proteins, and importantly the total yield of molecular phenotypes.

For that reason to better understand cell regulation, on top of using FAC-sorting strategies, it should also addressed the analysis of all possible effectors that can modulate cell regulation. Especially when studding effects of a particular treatment.

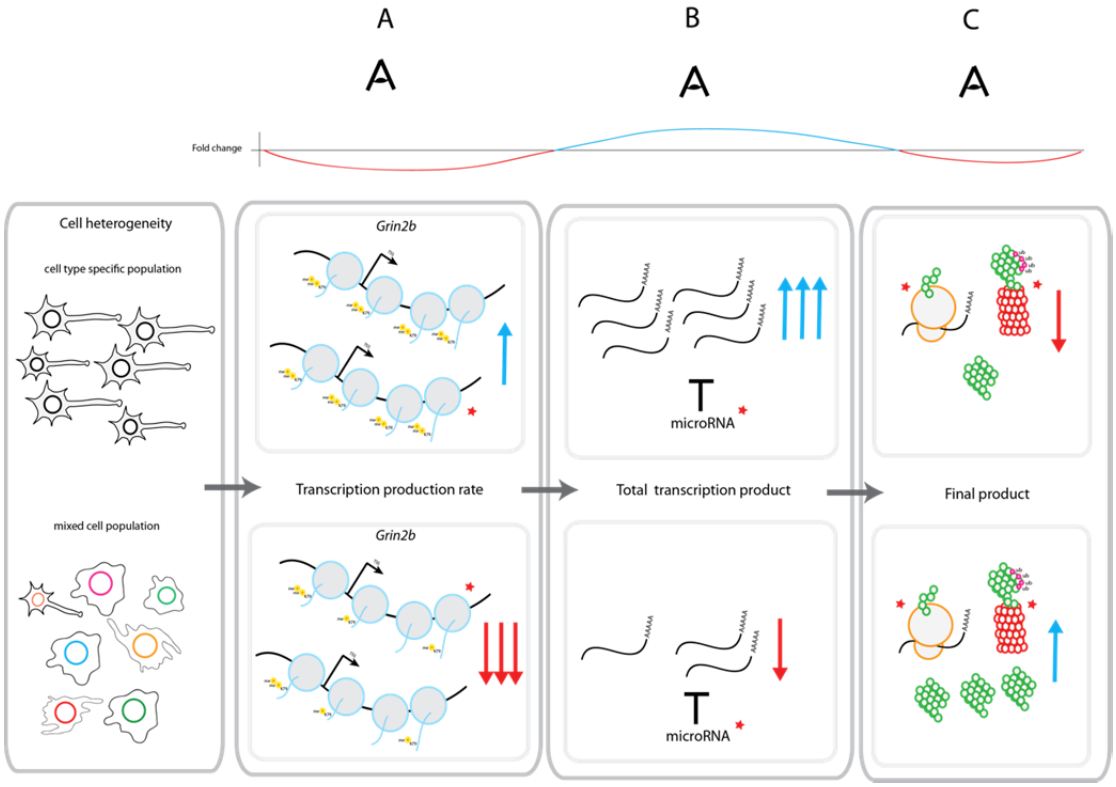


Figure 6.3. Discrepancy model
Figure showing a cell heterogeneity context in which a gene (*Grin2b*) has increased histone coverage for a specific type of cells, meanwhile the rest of the tissue shows down-regulation for that mark. This results in a higher number of transcripts been produced coming from that specific cell-type. Depending of protein production and degradation rate we can find a scenario in which the total yield of protein is decreased. The different modulators of the process are highlight by a red start

6.3 General overview of the results.

This research highlights the importance of studying all possible layers of regulation to better understand the changes of molecular patterns that a treatment can trigger. Despite some lack of regulation modulators, we haven seen that gene body histone marks are the epigenetic marks that change the most among other epigenetic modifications and correlate best with transcriptional activity. In future studies this could help to dissect the contributions of transcription, post-processing and degradation to observed RNA levels, as histone marks are mostly related to transcription and, to some extend, splicing.

Interestingly we have seen that by sorting neuronal cells the coverage profile of H3K79me2 is different compared to the whole cortex dataset, suggesting a possible change of transcription speed dynamics. Notably we have seen that a large majority of genes found up-regulated in neuronal cells are related with synaptic transmission, such as *Gria2* and *Gria4*, genes have been reported to show experience-dependent up-regulation (Naka et al., 2005). In general the neuronal population shows changes in genes related with an increased Rho signaling, CamK and MAPK and BMP signaling. It also shows higher activity of ionic and receptor genes playing important roles in synaptic function, such as: AMPA, GABA and Kainate receptors, together with potassium ionic channels and metabotropic glutamate receptors. This points to an increase of neurotransmission and synaptic reinforcement. But losses in Wnt signaling associated genes, Notch and MAPK signaling could also point to processes of synaptic turnover, evidencing the previous reported results of synaptic scaling during postnatal development (Bednarek and Caroni, 2011).

In contrast the chromatin shows a general gain of accessibility with low numbers of open chromatin regions being down-regulated. Notably, this bias is equally found in whole cortex, NeuN population and pyramidal neurons. For each of them we found that enriched increased developmental pathways such as Wnt and BMP can modulate morphological changes. However, those gains in accessibility are not always translated into changes in up-regulated transcripts, but rather significantly found in both directions of change. Hence, transcriptome data also reports differentially expressed genes that support synaptic scaling. However it is remarkable that the proteome data mostly indicates increased synaptic turnover.

In general, we were unable to identify a particular direction in synaptic changes. Instead we found that that both synaptic reinforcement and turnover might occur in response to environmental enrichment during postnatal development. The discrepancies between the different epigenomic features and different experimental designs could be explained by a model, in which a single snapshot of a complex tissue does not allow to disentangle changing epigenomic signals over time and space. A simplified schematic view for the results is illustrated in Figure 6.4, but a deep study needs to be done to fully understand all the molecular experience-dependent changes that might occur upon EE exposure.

In summary, we found that the analysis of EE treatment effects on the developing mouse cortex using multi-dimensional epigenomic, transcriptomic and proteomic data has led to interesting hypotheses about the molecular functions underlying learning. However, the picture is blurred due to the complex cell-type composition of cortex tissue. Based on a small number of single cell-type sequencing experiments we can conclude, that follow up studies should focus on the epigenome and transcriptome of sorted cell-types. To further reduce the impact of noise on the analysis, we would also suggest using an increased number of replicates.

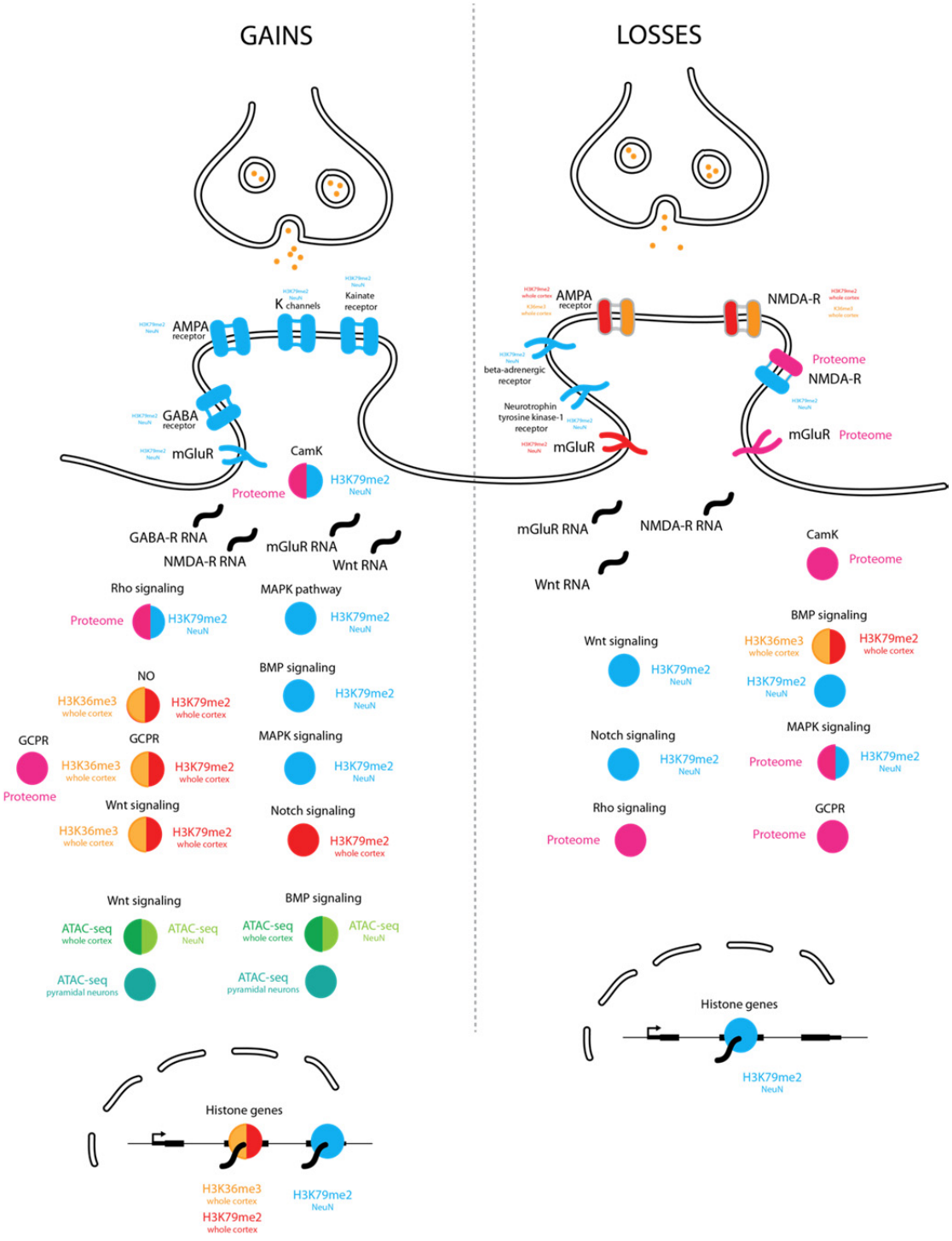


Figure 6.3. Molecular overview of induced changes by EE exposure
The figure illustrates in a schematic way the gains (left) and losses (right) of the main datasets.

6.3 Conclusions

Here I conclude the most important discoveries found in my PhD:

- Differential activity analysis using a ChIPseq peak dependent method allows determining small changes in confined ('peaky') histone mark regions, but shows fractionated results in long peaks, e.g. gene body marks.

- Differential activity analysis using a ChIP-seq peak-independent methodology using pre-defined gene/promoter annotations (developed in-house) improves the identification of activity changes in gene body marks.

- H3K4me3 differential binding induced by EE indicates increased activity of norepinephrine stimulus, glucocorticoid metabolic processes, astrocyte fate commitment, neurotrophine TRK receptors, mitochondrial permeabilization and L-glutamate transport. Losses of H3K4me3 coverage shows decrease activity on genes related with neuronal fate, oligodendrocite differentiation, BMP signaling, G protein gated potassium receptors, inhibition of calcium voltage gated channels and Notch signaling.

- Differential binding of H3K79me2 shows distinct patterns upon EE in each cell population, insinuating a possible role in modulating transcription rates.

- Gene body marks, and especially H3K79me2 are the epigenetic modifications most related with gene expression levels confirming previous reports.

- Both whole cortex H3K79me2 and H3K36me3 differential analysis revealed up-regulation of neurodevelopmental genes such as Wnt ligand, nitric oxide production, increased activity in histone genes and G-protein coupled receptors. H3K79me2 alone could show increased activity of GTPases, Ras, MAPK, ephrin and Notch signaling due EE.

- Regarding the down-regulated genes in the whole cortex, we found a decrease of BMP and importantly learning associated genes such as *Gria1*, *Grin2a*, and *Grin2b* in both gene body marks. As this is unexpected we hypothesized that the observation could be caused by either 1) cell-type heterogeneity/composition (an unwanted bias) or by 2) synaptic scaling-associated changes (a biologically relevant effect).

- FAC-sorting allowed dissecting whole cortex epigenetic signals, evidencing a greater contribution of NeuN negative population in the down-regulation of learning-associated genes previously reported in the whole cortex datasets. Interestingly, NeuN+ neuronal cells showed the expected increase of activity of *Gria1*, *Grin2a*, and *Grin2b,* supporting our first hypothesis that the unexpected decrease in activity in whole-tissue is caused by the complex cell-type composition of cortex.

- Neuronal H3K79me2 shows a significant increase of neurotransmission evidenced by the enrichment of up-regulated ionic channels such as AMPA

receptor subunits, Kainate ionic receptor, potassium channels, GABA receptors and nicotin acetylcholine receptor, pointing to an enhanced long term potentiation processes. Together with gain activity of BMP, of Rho signaling and a loss of Wnt signaling, also modulators of synaptic morphological changes.

- EE induces a general gene chromatin accessibility gain both in promoter and enhancers enriching for associated genes related with Wnt and BMP signaling in both whole cortex, total neurons and pyramidal neurons.

- Motif discovery in the differential chromatin accessibility regions allows to predicted transcription factors that were significantly expressed. Highlighting up-regulated effectors related with the early-expression genes important for synaptic plasticity.

- Transcriptomic analysis revealed that EE induces the up-regulation of learning and memory-associated genes, consistent with previous results.

- Induced changes of microRNA expression were highly related with the control of cytoskeletal organization associated genes, mediated in part by the increased expression of miR-125b-5p. Differential expressed miRNA also show decreased expression of let7 family that mediates control over developmental processes.

- Proteomic techniques used in the present study (iTRAQ and LC/MS) show complementary results pointing to synaptic turnover due EE.

- Chromatin accessibility and gene body epigenomic marks are the most related datasets analyzed in the present work. Together they explain best the transcriptomic changes induced by EE exposure. Chromatin accessibility is related to transcription factor binding and transcription initiation, while gene body histone modifications are related to transcription efficiency (transcription rate). This could be an explanation for discrepant results, as a single snapshot cannot unravel the dynamics of regulation over time and space.

Chapter 7
Methods

7.1. Animals, treatments, sample collection and pooling

Wild type mice (*C57BL/6J*) and Tg(Thy1-YFP) (strain *B6.Cg-Tg(Thy1-YFPH)2Jrs/J* No. 003782; The Jackson Laboratories) were kept and bred according to local (Catalan law 5/1995 and Decrees 214/97, 32/2007) and European regulations (EU directives 86/609 and 2001-486). All experimental procedures were approved by the local ethical committee (Procedure Code: ISA-11-1358).

Upon weaning at three weeks of age, female pups were separated into two experimental groups: control and enriched environment. Control mice were raised in standard conditions with two individuals per cage, while environmental enrichment consisted of raising 8 mice in a large cage with toys, houses and tunnels, which were changed every three days for novelty. Both groups had unlimited access to food and water at all times.

At eight weeks of age, mice were euthanized by carbon dioxide, and the cerebral cortex was dissected within one minute of death. The tissue was immediately flash frozen in liquid nitrogen. In the case of Tg(Thy1-YFP) animals, the tissue was immersed in Hank's Balanced Salt Solution (HBSS 1X, Gibco 14065-049) before proceeding with the sample preparation for the FAC-sorting (see Section 2)

For each condition and replicate, we pooled the cortices of five mice. The frozen cortices were ground together in a frozen mortar containing liquid nitrogen, to obtain a fine powder of pooled cortex tissue. The powder was aliquoted and stored at -80°C until further use.

7.1.1 Nucleic acid extraction

DNA was extracted using Phenol:chroloform:iso-amyl alcohol (25:24:1) according to manufacturer guidelines (Sigma 77617-500ml). RNA was extracted using Qiazol total RNA (Qiagen Cat No:79306) kit according to the manufacturer's instructions. The RNA was quantified by Qubit. ® 2.0 Fluorometer (Life Technologies) and the quality has been assessed using a Nanodrop 2000c (Thermo Scientific) and a 2100 Bioanalyzer (Agilent Technologies, CA, USA).

7.1.2 Nuclei isolation

To obtain fresh nuclei, ground frozen tissue was resuspended in tissue lysis buffer (1x PBS containing 0.1% Triton X-100, 1x Complete Protease inhibitor cocktail tablette (cOMPLETE mini EDTA-free, Roche Cat No.11836170001) and 1 mg/ml AEBSF (Pefabloc, Roche Cat No.11585916001)) and dissociated by 60-90 strokes in a glass dounce homogenizer (7 ml tissue grinder Tenbroek, Wheaton Cat No. 357424). Nuclei

were counted using a haemocytometer and constantly checked under a microscope (Leica DM-IL).

7.2. FAC-sorting

7.2.1 Sorting total neurons using NeuN marker

After the nuclei preparation (Section 1.2), nuclei were resuspended in 1ml of PBS-PI 1X (1X-PBS, 1x Complete Protease inhibitor cocktail (cOMPLETE mini EDTA-free, Roche Cat No.11836170001**)**, 1 mg/ml AEBSF Pefabloc (Pefabloc, Roche Cat No.11585916001) and 0,1 mg/ml of BSA (BSA, Molecular Biology Grade NEB, Cat No.B9000S)). It is added to the solution 1,5 µl of anti-NeuN, clone A60, Alexa Fluor ® 555 Conjugate (Merk Millipore Cat No. MAB377A5) and incubated at 4°C during 1,5h in dark conditions. Sample was spun during 10 mins, 500g at 4°C and washed with 1 ml of PBS-PI 1X. After 1 µl of 4′,6-Diamidine-2′-phenylindole dihydrochloride was added (DAPI, Roche Cat. No. 10236276001). Immediately after the sample was given to the FAC-sorting facility. Samples were sorted at 12PSI in cold condition in a INFLUX sorter (BD Biosciences INFLUX^{TM}). After, the sorted samples were spun during 40 mins at 700g at 4°C to collect the nuclei and proceed with desired technique to follow.

7.2.2 Sorting pyramidal neurons

Tg(Thy1-YFP) animals were dissected and immediately submerged in Hank's Balanced Salt Solution (HBSS 1X, Gibco 14065-049). To dissociate the brain sample a commercial kit was used using the manufacturer guidelines (Neural Dissociation Kit (P) MACS Milteny Biotec Cat.No. 130-092-628; LS Columns Cat.No 130-042-401; Myelin removal beads Cat No. 130-096-731; MidiMACSTM separator Cat.No. 130-042-302). Cells were sorted using an INFLUX^{TM} sorter (BD Biosciences INFLUX^{TM}). After sorting, samples were spun for 40 mins at 700g at 4°C to collect the nuclei.

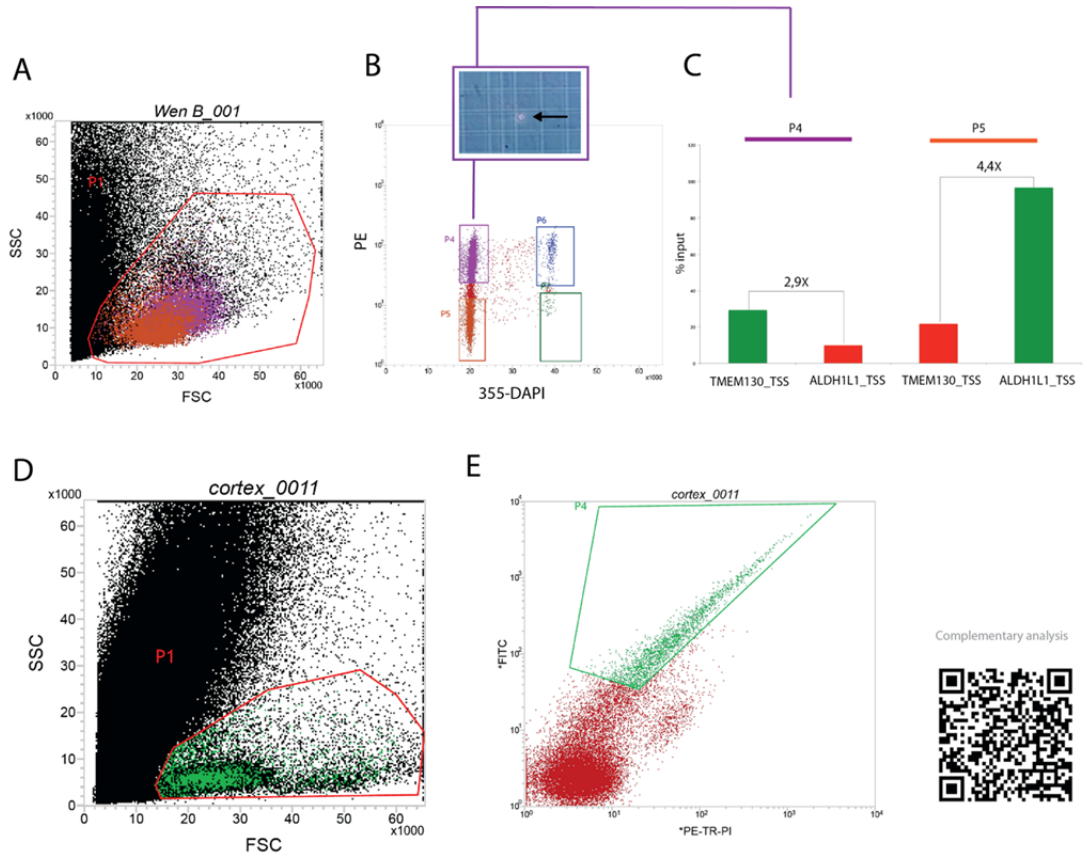


Figure 7.1. FAC-sorting procedure and preliminar QC analysis
A) Side-scattered light (SSC) versus Forward-scattered light (FSC) plot of NeuN+ nuclei preparation. The SSC value reports the complexity of the sample measured by the refraction of the laser light beamed in the surface of the nuclei or cells in the flow. The FSC reports the intensity of the signal given by the conjugated antibody (Alexa Fluor-555). **B)** Intensity of the signal of the conjugated antibody (Alexa-555) versus DAPI intensity. P4 and P6 represent the nuclei positive for the neural marker NeuN+, highlighted a nuclei at 40X after sorting in a Leica DM-Il **C)** Quality control to test the accuracy of the neuronal sorting. The plot illustrates a qPCR of an anti-H3K27ac ChIP library using specific primers to detect neuronal specific TSS H3k27ac activity (*Tmem130*) and glial activity (*Aldh1l1*), based in (Cahoy et al., 2008). **D)** Side-scattered light (SSC) versus Forward-scattered light (FSC) plot of the YFP pyramidal neuron preparation **E)** FITC signal of the YFP neuronal cells. QR at the right section of the Figure 7.1 gives access to complementary analysis.

7.3. Chromatin Immunoprecipitation sequencing: ChIP-seq

7.3.1 Chromatin preparation

Nuclei were obtained as described in section 1.2 and cross-linked with 0,5% formaldehyde (Sigma F8775-25ml) for 5 minutes at room temperature. Residual formaldehyde was quenched by addition of glycine (MAGnify™ Glycine P/N 100006373) to a final concentration of 0.125M and incubation for 5 minutes at room temperature. Nuclei were pelleted by centrifugation at 500g during 10 mins at 4C and suspended in 300 ul lysis buffer (MAGnify™ Chromatin Immunoprecipitation System, Cat no.49-2024). Chromatin was fragmented by sonication in a Covaris S2 (For histone marks: Duty Cycle: 20, Intensity: 8, Cycles per Burst: 200, Time: 15mins, for FACS-sorted nuclei: Duty Cycle: 20, Intensity: 8, Cycles per Burst: 200, Time: 25mins) to a median size of 200 bp, aliquoted and stored at -80°C until further use.

For non-histone proteins, such as CTCF, no nuclei preparation was performed. Homogenized tissue was cross-linked with 0,5% formaldehyde for 10 minutes at room

temperature. Sonication conditions used were: Duty Cycle: 5, Intensity: 2, Cycles per Burst: 200, Time: 25mins.

7.3.2 Chromatin immunoprecipitation

Chromatin immunoprecipitation was performed using antibodies against histone modifications (H3K27ac, H3K4me3, H3K4me1, H3K79me2, H3K36me3, H3K27me3, H3K9me2 and CTCF, Appendix – Table 1) with the MAGnify™ Chromatin Immunoprecipitation System, (Invitrogen Cat no.49-2024). For whole cortex and NeuN histonic ChIP-seq a total amount of 50k nuclei was used per ChIP (~330 ng). For non-histonic ChIP-seq a total yield of 700k nuclei was used (~4 μg). The immunoprecipitation was performed according to manufacturer's instructions. Recovered ChIP DNA was used to construct sequencing libraries, using NEBNext Ultra (New England Cat.No. E7370L) kit according to the manufacturer's protocol and sequenced 50bp single end (SE) on a HiSeq2000 sequencer (Illumina) yielding 30-40 million reads per sample. Four samples were pooled together in a lane for the 1st group of breedings. For the FACS isolated samples, 6 libraries were pooled per each lane.

7.3.3 qPCR

Quality of the ChIP-seq was determined by qPCR, using positive and negative primers to detect the regions where the histones should be placed in the genome (Appendix - Table 2 and 3). Primers were diluted to a final concentration of 300 ng in Power SYBR Green PCR MM (Applied Biosystems Cat.No 4367659). qPCR cycles consisted in: 50°C-2min, 95°C-10min, 40 cycles of 95°C-15s 60°C-1min, 95°C-15s, 60°C-15s and 95°C-15s. The samples were loaded in 96-well tray (MicroAmp® Optical 384-Well Reaction Plate with Barcode Cat. No. 4309849) and run in a Applied Biosystem qPCR system (7900 HT Fast Real-Time PCR System).

7.3.4Computational analysis of ChIP-seq data

ChIP-seq single-end reads were aligned to mouse mm9 Gencode reference genome assembly using Bowtie v1.0.0 with 8 cores parallelized and 20GB of RAM (Langmead et al., 2009). Then the resulting BAM files were sorted by genomic position and indexed using samtools (Li et al., 2009). We set up two pipelines implementing two strategies for differential binding activity analysis, a) peak-calling dependent and b) peak-calling independent. For the peak-dependent strategy we have predicted peaks using 6 different peak callers. We used the recommend parameters for each histone mark following the ENCODE guidelines (Landt et al., 2012)and the manual of each peak-caller: MACS v1.4.2, MACS2 v2.1.0 (Zhang et al., 2008), SICER v1.1(Xu et al., 2014b), HOMER v4.7 (Heinz et al., 2010), MUSIC (Harmanci et al., 2014) and BCP v1.1(Xing et al., 2012). Next, each of the resulting peak annotations was loaded together with the mapped reads BAM file into the differential analysis script DiffBind (Ross-Innes et al., 2012), a wrapper tool for the commonly employed differential analysis tools edgeR(Robinson et al., 2010) and DeSeq2 (Love et al., 2014). As a result we obtained 12 lists of differential binding sites (DBS), i.e. the results of edgeR and DeSeq2 per each peak annotation.

The peak-independent strategy consisted in the identification of DBS in pre-defined regions representing the gene bodies and the promoters. First, we calculated the average coverage profiles of each histone mark around the TSSs and the gene bodies. Based on these plots we have selected the most informative region per histone mark and CTCF, i.e. the region with the highest coverage density. The full gene body was selected for the broad marks H3K36me3 and H3K79me2; 500bp upstream and 1500bp downstream the TSS for H3K4me3, H3K27ac and CTCF; and 700bp upstream the TSS and the full gene body for the repressive marks H3K27me3 and H3K9me3 H3K4me1 was excluded, as it showed strong enrichment only in enhancers, which have not been considered here. Reads falling in each region were counted using featurecounts v1.4.6 (Liao et al., 2014).Then the data was loaded into edgeR and Deseq2 to perform differential binding analysis.

Gene Ontology analysis was performed using the Cytoscape (Shannon et al., 2003) tools ClueGo and Cluepedia (Bindea et al., 2013; Bindea et al., 2009) always with term fusion option to avoid having similar terms in the output.

7.4. ATAC-seq

ATAC-seq was performed with a few minor modifications from (Buenrostro et al., 2013). One hundred thousand nuclei were treated with 2.5 µl Tn5 at 37°C for 30 minutes, followed by cleanup on a Qiagen Minelute column. Fragments >1kb in size were removed using AmpureXP beads (Agencourt AMPure XP, Beckman Coulter Cat.No. A63881) as follows: DNA volume was adjusted to 50µl, 25µl AmpureXP beads were added and incubated at room temperature for 10 minutes. Supernatant was recovered and incubated with one volume of AmpureXP beads for another 10 minutes at room temperature. Beads were washed twice in 200 µl of 80% ethanol and dried, and the DNA fragments were eluted in 20µl 10mM Tris-HCl pH8.0 after 20 minutes incubation at room temperature. DNA fragments were subjected to 11 cycles of PCR amplification with custom adapter primers from Buenostro et al. (2014). PCR reactions were cleaned up with one volume AmpureXP beads and quality controlled by Bioanalyzer (Agilent Technologies). Libraries were sequenced 50bp paired-end (PE) on a HiSeq2000 sequencer yielding 50M of reads per library.

The ATAC-seq control with mouse cortex DNA digested overnight at 65°C using proteinase K at 0.01µg/µl final concentration (Thermo Fisher, Cat.No. EO0491). Next, DNA was purified using Phenol:chroloform:iso-amyl alcohol (25:24:1) according to manufacturer guidelines (Sigma 77617-500ml). Then, purified DNA was loaded to the Tn5 reaction following (Buenrostro et al., 2013) as previously described.

7.4.1 ATAC-seq Computational analysis

Reads were aligned to the mouse mm9 Gencode reference genome assembly using bwa-mem v0.0.7 (Li and Durbin, 2009). Peaks were called using f-seq v1.81(Boyle et al., 2008) and then loaded to DiffBind as previously explained for ChIP-seq. Different background noise profiles were provided for benchmarking: input ChIP-seq DNA, ATAC-seq control background and no control. The same strategy was also used with the annotation of predicted enhancers (Jhanwar et al. *in preparation*) loaded to DiffBind for the differential analysis. Gene ontology analysis was performed similarly to ChIP-seq analysis.

7.5 RNA-seq

7.5.1 Poly-A RNA-seq

Illumina TruSeq™ RNA sequencing library was prepared from total RNA using the TruSeq™ RNA sample preparation kit (Illumina Inc., Cat.No. RS-122-2001) according to the manufacturer's protocol. Briefly, starting with 0.5 μg of total RNA, the mRNA fraction was first purified using poly-A based enrichment selection by oligo-dT magnetic beads, followed by heat digestion in the presence of divalent cations, resulting in fragments of 80-250nt, with the major peak at 130nt. The fragments were converted into first strand cDNA using random hexamers and followed by second strand cDNA synthesis primed by RNaseH. Double stranded cDNA was end-repaired, 3´adenylated and the 3´-"T" nucleotide at the Illumina adaptor was used for the adaptor ligation. The ligation product was amplified with 15 cycles of PCR.

7.5.2 Directional RNA sequencing library preparation

Directional RNA libraries were prepared using the ScriptSeqTM Complete Gold Kit (Human/Mouse/Rat) (Epicentre Biotechnologies), according to the manufacturer's protocol. Briefly, 3 μg of total RNA were depleted of both cytoplasmic and mitochondrial rRNAs using the Ribo-Zero™ Gold rRNA Removal Reagents. The total rRNA depletion of the samples was confirmed on a 2100 Bioanalyzer RNA 6000 Pico Chip. For the library preparation we used 50 ng of Ribo-Zero-treated RNA as starting material for the ScriptSeqTM v2 RNA-Seq Library Preparation Kit. This library preparation consists of randomly primed first-strand cDNA synthesis and incorporation of an Illumina platform-specific 3' sequencing tag. Multiplexing indices were added through 12 cycles of PCR, using the FailSafeTM PCR Enzyme Mix (Epicentre Biotechnologies) followed by purification using the AMPure XP purification system (Agencourt, Beckman Coulter Cat.No. A63881).

Both the directional mRNA and the Poly-A RNA libraries were sequenced in paired end mode with read length 2x101bp on a HiSeq2000 (Illumina, Inc) following the manufacturer's protocol.

7.5.3 miRNA-seq

RNA was extracted using miRNeasy Mini Kit (Qiagen Cat.No. 217004). Small RNA libraries were generated using the TruSeq (Illimuna, Cat.No. RS-122-2001) kit according to the manufacturer's protocol. The resulting ~22bp insert libraries were sequenced on a HiSeq2000 sequencer (Illumina), yielding 15-20 million reads per sample.

7.5.4 Computational analysis and validation

Reads were aligned to mouse mm9 Gencode reference genome assembly using Tophat 2.0.6 version (Kim et al., 2013). Number of reads overlapping exons in the Gencode mm9 annotation was counted and summarized at the gene-level using featureCounts v1.4.6 (Liao et al.2014). mRNA had a batch effect that was corrected by using RUVg strategy of the tool RUV-seq (Risso et al., 2014). RUVg was used to calculate sample-specific weights, which were then incorporated into the design model during differential expression analysis in edgeR. A subset of the genes differentially expressed were validated by RT-qPCR, using SuperScript III First-Strand (Invitrogen, Cat. No. 18080-051) for the reverse transcription following manufacturer guidelines and then performing a qPCR using the primers listed in Appendix – Table 4.

Small RNA reads with homo-polymer and low PHRED scores were removed using FASTX-Toolkit and a custom script. Ligation adapters were clipped using the AdRec.jar program from the seqBuster (Pantano et al., 2010) suite with the following options: java -jar AdRec.jar 1 8 0.3. A custom search subsequently clipped shorter adapters: if there were no matches to the first 8 nts, then matches to the first 7 nts of the adapter were searched in the last 7 nts of the read, then matches of the first 6 to the last 6 positions and so on. Reads that had no matches were retained, but not clipped. Last, reads shorter than 18 nts were discarded. miRNA quantifications for analysis were calculated as read counts using miraligner.jar from the seqBuster suite using the following options: java --jar miraligner.jar 1 3 1, and using the mouse miRBase version 18 annotations as the reference. Reads that mapped equally well to two or more miRNAs are counted fully towards each miRNA. As miRNA dataset presented a strong batch bias and any tool tested was able to remove the unwanted variation. So another strategy was carried out by simple filtering out microRNAs with low read coverage (<50 counts), normalizing the libraries by read counts per million (rpm) and then applying the contrast of EE against CTLs, looking exclusively of conserved direction of the fold change values. A subset of microRNA differentially expressed was validated by RT-qPCR, using SuperScript III First-Strand (Invitrogen, Cat. No. 18080-051). A custom reverse transcription protocol of each of the selected microRNA to be validated was performed, modified from (Chen et al., 2005). The process consists in designed RT-primers that with a constant 5' primer sequence and 8 nucleotides at the 3' end that are complementary to the 3' of the miRNA to be validated. For the qPCR an universal reverse primer is used together with a forward primer that has a constant 5' sequence and 12-15 nucleotides at the 3' with the same sequence as the 5' end of the miRNA to be validated. (Appendix- Table 5).

Gene Ontology analysis was performed similarly to ChIP-seq computational analysis (section 3.5)

7.6. Proteomics

7.6.1 Protein extraction and mass spectrophotometry

Samples were minced with RIPA-M buffer (1% NP40, 1% Sodium deoxycholate, 0.15M NaCl, 0.001M EDTA, 0.05 TrisHCl pH=7.5, 1X cOMPLETE Mini EDTA free, 0.01M NaF, 0.01M Sodium pyrophosphate, 0.005M β-glycerolphophate) and sonicated with Diagenode Bioruptor (0,5 ON-0,5-OFF medium intensity during 5mins), then spun during 10 min 16000 rmp at 4C and precipitated with acetone at -20C during 1 hour.

After spinning the samples during 10 mins 16000 rpm at 4C the pellet was dried and resuspended in Urea/200 mM ABC, resonicated during 10 mins (0,5 ON-0,5-OFF medium intensity) and quantified prior the mass spectrophotometry injection following Isobaric tags for relative and absolute quantitation (iTRAQ) or Liquid Chromatography/Mass-Spectophotometry (LC/MS).

7.6.2 Computional analysis

The analysis of iTRAQ data consisted in sorting the discrepancy ($\delta = \varphi/\beta$) between both biological contrasts (φ = EE1/CTL1 and β = EE2/CTL2), following the premise that consistent results will be indicated by:

$$\lim \delta = 1$$
$$\varphi \leftrightarrow \beta$$

Discrepant results then will be considered as values far from $\delta = 1$. Based on this consideration we select all the proteins that follow the condition and we established a threshold of ±0.1.

The protein expression values from the LC/MS were log2-transformed and loess normalized using normalize.ExpressionSet.loess function from BioConductor package AffyPML. Differential expression analysis was conducted with a standard limma(Smyth and Speed, 2003) pipeline by calculating sample weights, fitting a linear model for each gene across all samples and calculating moderated F-statistics. Unadjusted p-values were used to rank the proteins.

Gene Ontology analysis was performed similarly to ChIP-seq computational analysis (section 3.5)

Appendix

Table 1. Antibodies used in the present study

Antibody	Provider	Cat.No	Lot #	Concentration stock	Concentration used
CTCF polyclonal antibody	Diagenode	C15410210	A2359-0010	2,5µg/µl	1µg per IP
H3K4me3 clone MC315	Merck Millipore	04-745	2276411	undetermined	~1/20
H3K27ac polyclonal antibody	Diagenode	C15410196(pAb-196-050)	A1723-0041D	2,8µg/µl	1µg per IP
H3K4me1 polyclonal antibody	Diagenode	C15410194(pAb-194-050)	A1862D	1,5µg/µl	1µg per IP
H3K36me3 polyclonal antobody	Diagenode	C15410192(pAb-192-050)	A1847-001P	1,2µg/µl	1µg per IP
H3K79me2 polyclonal antibody	Diagenode	C15410051(pAb-051-050)	A85-0011	1,1µg/µl	1µg per IP
H3K9me3 polyclonal antibody	Diagenode	C15410196(pAb-193-050)	A1671-001P	2,7µg/µl	1µg per IP
H3K27me3 polyclonal antibody	Diagenode	C15410195(pAb-195-050)	A1811-001P	1,9µg/µl	1µg per IP
Anti-NeuN clone A60 Alexa Fluor ® 555 conjugate	Merck Millipore	MAB377A5	2649280	undetermined	~1/1000

Table 2. ChIP qPCR mouse primers

Primers	Forward Sequence 5' - 3'	Reverse Sequence 5'-3'	CTCF	K4me1	K4me3	K27ac	K36me3	K79me2	K27me3	K9me3
Baat_TSS	AGGCATTTATTCAAAGTCCAAG	AGTATTTCCCTCCTCAACCAG								
Calr_TSS	AACTGGGAAGCAATGGAAAG	CTACCTCTCACCCGAACCTG								
Dppa2_TSS	GGCCTCAAGACAACAGGAAG	TCACCTTCTTGTCTCCAGCA								
Duoxa1_-1kb	CCCAACTCATTATTCCATCTATC	ATACCCAACACACCACCTTC								
Duoxa1_+1700bp	CAGTGAGAAAGTGCCCAACTC	GGACAATCAAACCCACACC								
Duoxa1_TSS	CGCTCCTCCGTTCTGG	GCTTCTATTCCGCCCTACTG								
Eef1G_+1500bp	GAACTTGACTGCCCAAAGG	TAGGCTTGCTTCACACAATCTC								
Eef1G_500bp_upst	ATTTGGGTTTAGGGTTTGTC	CAGAGGGTCGGGTAAGG								
Eef1G_END	GCAAAGCCGTCAATCAAG	AATCACAAAGGAGGCAGGAC								
Eef1G_TSS	TATGCCGTCCTCGCTTTC	GGGCTGTCGGTGCTG								
Hoxa1_TSS	CATTGGCTGGTAGAGTCACG	CAACTTTCTCACTTCCTCCATAGG								
Hprt_TSS	GGCTTACCTCACTGCTTTCC	CCTTGGCTCACCACGAC								
Notch4_END	GGAAGGAAGGAAGCGACAC	TCAACCGGACATCCTAAACC								
Nrgn_TSS	GTCCTCGCTCCAGTTCTCC	AACCAGCCTCTCTCCTCTTG								
Rpl4_TSS	CGGAAGACGAGTGGAAGGAT	GGGAAGCCCAGTGTAGAAGA								
Satb1_END	TGAGCCAAAGATGGGTATGA	CCAAGTCAAGTGGTTTAGTTACAG								
Satb1_LESS_PROX	CTACCCTTGCCACTGAGAGA	ACGCTTCATTTCCTGCTGA								
Vnn3_TSS_	TTCCTCAATGCCGTGTTC	CAGACGCAAGTCGTGATG								
B2m	GGGAAAGTCCCTTTGTAACCT	GCGCGCGCTCTTATATAGTT								
Hbb	GAGCCCTGGTCATTTTCACT	AGTCAGCTGGCATTGTTCCT								
Grinb2b	GAGACTGGGTTAGCGGAGAG	GTGGAACGCATCCTTGAAGT								
Gadl	TGTCTCACCAAAGTCCCTGTC	CACGTCTGGTTCGGTGTCT								
Intergenic	CAGGAAAGCTGGAAGTGGAA	CCCTGCTCTGAACACTGTGA								
Intgenic_+6632Eef1g_	CATTCCTGGAAGCTGGGTTA	CACTACCAGATGCAGGAGCA								
Enhancer region_chr11	TACTGGGCTTTGCTTCCCTA	AGATCTTTCTGGCAGCCTCA								
Enhancer region_chr8	CCAGAGCGATCACAGAAACA	CAGGGCGCTGACTTTATAGC								
chr14_69935989	TGCCTCAAAGAAGGGTATGG	GAGTAAAGGGCTGGTGCTTG								
chr18_76381681	AGGGCAGAGAGGCTGTGTTA	CACTCTGGAATGCGTTAGCA								
Calr_Intronic	GCCTCTATTCCCACCTTTCC	GGCTTGGACCTCCACAAGTA								
Rpl30_TSS	CTCTTCTAGGCCGACAGAGG	CCATAGCAACCAACTACCGC								

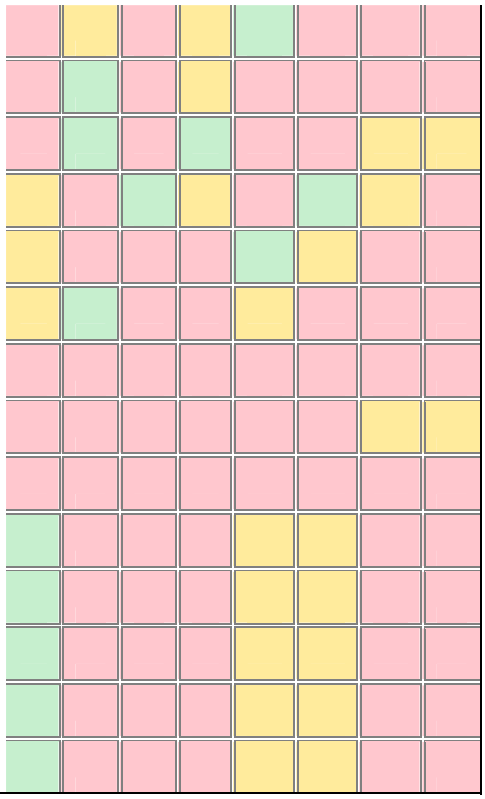
PP1A_END	CAGCTCCCAACTACTGTGGA	GTTGAGCCAGGCGTCTATTG	
Enhancer region_2_chr11	CTCTCCATCTGCCTGAGTCC	GGAGAAATGAGCTGCTGAGG	
Enhancer_region_2_chr8	CACCTCAGATCCAAGCAGAAATC	TGGAGGTACCCAGATAAATG	
Rpl26L1_TSS	TCCGTGTCTTTCCTAGCCAG	GTCCCCTTAACCCTCTCCAG	
Rpl26L1_intronic	TGCTTTCCACCACTAGCTCA	TTGCTGCACTAAGGACCAGA	
Bahcc1_CTCF	CCCAATATATGTGACCTCCCAGCCATGCAC	AGGAAACGTGGGCTGCACACATGGTAGAG	
Tox_CTCF	TTGCCTTTCTTTGTCTTTCTCCAGTGTCTC	ATGCCTCTATGAAAATTATCTCAAGGG	
Intergenic_CTCF	CAAGGCCAACCCAACCTACACTTATGTCTG	CAAGATGGAGCTGTTCCCGTTTGAC	
Art1_CTCF	TGCTCACACCGCCTCTTAGCCTGTTTACCA	CACCTCTTCTCTTGTCCAGACAGCGCACCC	
Eef1g_upsEND1	GGTGGCTGTTCTCCGAATTC	CCTCTCCTGGACAATGCTCA	
Eef1g_upsEND2	ATGCTGGCTAGCTTTCTTGG	CTGGAACCAGAATTCGGAGA	
chr10_interg	TCTTGTCCGCAAGTCACAAG	CGGATTTGCCTCTCTCTCAG	
Nkx2-1_END_downs	CGCTTTCTAGGTGTGGAAGC	TCCCCAAGAGGCTGTAGAGA	
chr4_interg_murc	GGTACTCCAAGTCCCAAGGG	AGACCAGAATCAGAGGGCAA	

Table 3. NeuN+ FACs-ChIP qPCR mouse primers

name	primers	comments
01_RXFOX3-NeuN_F	AAGTTTGGGCTAGGGTTGTG	neuron-specific
02_RXFOX3-NeuN_R	CAAGGATCGGAGGTTACTGC	neuron-specific
09_Tmem130_F	GTACCAGTCTTGAGCATCGG	neuron-specific
10_Tmem130_R	GTCCAGAACCTAAGCTACCG	neuron-specific
13_Nov_F	GCGCACTATAAAACCTGCAC	neuron-specific
14_Nov_R	GGAAGAGGCTCATGTTCTGC	neuron-specific
19_Aldh1L1_F	ACAATCCAAGGGTCTGCTTG	astrocyte-specific
20_Aldh1L1_R	CATTAACCACCCTGCCTCTC	astrocyte-specific
21_Gfap_F	AGGGAGTGGAGGAGTCATTC	astrocyte-specific
22_Gfap_R	CATAAAGGCCCTGACATCCC	astrocyte-specific
29_Ntsr2_F	TCGTGGCTATTACTTCGTGC	astrocyte-specific
30_Ntsr2_R	ACCCAGACTAGTGACAGCAG	astrocyte-specific
35_Fa2h_F	CCGGTTACCTGGAGAGAAAC	oligodendrocyte-specific
36_Fa2h_R	CGCCACTTTAGAGAGCTCAG	oligodendrocyte-specific
37_Tmem125_F	ATCTGGGAGTCTCAGGGAAG	oligodendrocyte-specific
38_Tmem125_R	GCAAGGGGATTCTCTGAAGG	oligodendrocyte-specific
43_Gpr62_F	TGCGCTAAGTAGAAGGCATC	oligodendrocyte-specific
44_Gpr62_R	TCTAGCAAACCCCAGTGAAG	oligodendrocyte-specific

Table 4. mRNA validation primers

	Sequence
Grin2b_F	CCTTAATTGCCGGAGACTCC
Grin2b_R	CATCTTTTATGGCCACTTCG
Bdnf_F	TGCTTAGGCCTTCTGCTTTA
Bdnf_R	AAGATGAAGCGATTGTTTGC
Grin2a_F	GCTGACAAGGATCCGACATC
Grin2a_R	CACCTGATAGCCTTCCTCAG
Arc_F	CAGCAGTGATTCATACCAGTG
Arc_R	AGCCAAATGAGTATGGGTCA
Fos_F	GAGCAGCTATCTCCTGAAGAG
Fos_R	CTGCATAGAAGGAACCGGAC
Nr4a1_F	GCTGGAGATGCCCTGTATTC
Nr4a1_R	GAGCAGGGACTGCCATAGTA
Nos3_F	AGGACAACCTCATCCCTGTG
Nos3_R	AGAATGGTTGCCTTCACACG
Nos1_F	TCTGTCCGTCTCTTCAAACG
Nos1_R	GACGGCGAGAATGATGTC
Ngf_F	GCAGTGAGGTGCATAGCG
Ngf_R	CAGAGTCTCCTTCTGGGACA
Egr2_F	CGGGAGATGGCATGATCAAC
Egr2_R	GGATGGCGGCGATAAGAATG
Dusp5_F	CTCCCACTTTCAAGAAGCAA
Dusp5_R	GGTACGGAATGTGCAGTAGG
Dusp1_F	TCTCCAAGGAGGATATGAAGC
Dusp1_R	TGGCTGAGACGTTGATCAAG
Slc6a12_F	TGCTTCCTGAACAGTGCTAC
Slc6a12_R	TCACCTCAAACAAGGAGAGG
Mapk12_F	AGAACGTCATTGGGCTACTG
Mapk12_R	CTTCACTCAGGGTCTCATGC
Camk2n1_F	CAAGCGCGTTGTTATTGAAG
Camk2n1_R	ATTGGGAACAGCCAAATCAG
Gh_F	AACTGAAGGACCTGGAAGAGG
Gh_R	GGCGTCAAACTTGTCATAGG
Prl_F	GGAGTAGGTGGTATTCAAGAAGC
Prl_R	TGTAGGGATGGAAGTTGTGACC
Omp_F	GAAGCAGGATGGTGAGAAGC
Omp_R	GTGTCATGAGGTTGGTGAGG
Pmch_F	TCCAATGCACTCTTGTTTGG
Pmch_R	CGACAACGGATCTTTCTGC
Btg2_F	TGAAGTGTCTTACCGCATCG
Btg2_R	ACAGTTCCCCAGGTTGAGG

Sox18_F	AAATCCGGATCTGCACAACG
Sox18_R	AGGCCGGTACTTGTAGTTGG
Dgkh_F	CAGAAACAGATGAAGCTAAAGAGG
Dgkh_R	TGATTCTAGTGTTGAGCACAGG
Bdnf_F	ACTGCAGTGGACATGTCTGG
Bdnf_R	AGTAAGGGCCCGAACATACG
Grin2b_F	GTGGATATGCAAGCGAGAAG
Grin2b_R	CCAGAACTTGGGAGAACAGC
Arc_F	GAAGTGGTGGGAGTTCAAGC
Arc_R	TCCTCAGCGTCCACATACAG
Fos_F	CTGTCCGTCTCTAGTGCCAAC
Fos_R	TGACACGGTCTTCACCATTC
Egr2_F	AGCTGCCTGACAGCCTCTAC
Egr2_R	TTGCCCATGTAAGTGAAGGTC

Table 5. microRNA validation primers

Universal_primer	TGGTGTCGTGGAGTCGGCAATTCAGTTG
RT-mmu-miR-7a-5p	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGACAACAAA
RT-mmu-miR-598-3p	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGTGACGATG
RT-mmu-miR-98-5p	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAACAATAC
RT-mmu-miR-34a-5p	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGACAACCAG
RT-mmu-miR-451a	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAACTCAGT
RT-mmu-miR-8114	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAGGCGGAG
RT-mmu-miR-6769b-3p	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGCTATGGGT
RT-mmu-miR-7046-3p	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGCTGTAGGG
RT-mmu-miR-1306-5p	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGGGACGTTT
RT-mmu-miR-702-3p	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGGGAGCGGG
F-mmu-miR-7a-5p	ACACTCCAGCTGGGTGGAAGACTAGTGATT
F-mmu-miR-598-3p	ACACTCCAGCTGGGTACGTCATCGTTGT
F-mmu-miR-98-5p	ACACTCCAGCTGGGTGAGGTAGTAAGTTG
F-mmu-miR-34a-5p	ACACTCCAGCTGGGTGGCAGTGTCTTAGC
F-mmu-miR-451a	ACACTCCAGCTGGGAAACCGTTACCATTA
F-mmu-miR-8114	ACACTCCAGCTGGGTCACCCATCTCCTC
F-mmu-miR-6769b-3p	ACACTCCAGCTGGGCCCTCTCTGTCCCA
F-mmu-miR-7046-3p	ACACTCCAGCTGGGTGAACCACCCATCCC
F-mmu-miR-1306-5p	TGAACCACCCATCCCCACCACCTCCCCTGCA
F-mmu-miR-702-3p	ACACTCCAGCTGGGTGCCCACCCTTTACC

Abdo, H., Li, L., Lallemend, F., Bachy, I., Xu, X.J., Rice, F.L., and Ernfors, P. (2011). Dependence on the transcription factor Shox2 for specification of sensory neurons conveying discriminative touch. Eur J Neurosci *34*, 1529-1541.

Agostini, M., Tucci, P., Killick, R., Candi, E., Sayan, B.S., Rivetti di Val Cervo, P., Nicotera, P., McKeon, F., Knight, R.A., Mak, T.W.*, et al.* (2011). Neuronal differentiation by TAp73 is mediated by microRNA-34a regulation of synaptic protein targets. Proc Natl Acad Sci U S A *108*, 21093-21098.

Alarcon, J.M., Malleret, G., Touzani, K., Vronskaya, S., Ishii, S., Kandel, E.R., and Barco, A. (2004). Chromatin acetylation, memory, and LTP are impaired in CBP+/- mice: a model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. Neuron *42*, 947-959.

Alberini, C.M. (2009). Transcription factors in long-term memory and synaptic plasticity. Physiol Rev *89*, 121-145.

Alfano, G., Conte, I., Caramico, T., Avellino, R., Arno, B., Pizzo, M.T., Tanimoto, N., Beck, S.C., Huber, G., Dolle, P.*, et al.* (2011). Vax2 regulates retinoic acid distribution and cone opsin expression in the vertebrate eye. Development *138*, 261-271.

Allen, Y.S., Adrian, T.E., Allen, J.M., Tatemoto, K., Crow, T.J., Bloom, S.R., and Polak, J.M. (1983). Neuropeptide Y distribution in the rat brain. Science *221*, 877-879.

Allis, C.D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. Nat Rev Genet *17*, 487-500.

Altafaj, X., Dierssen, M., Baamonde, C., Marti, E., Visa, J., Guimera, J., Oset, M., Gonzalez, J.R., Florez, J., Fillat, C.*, et al.* (2001). Neurodevelopmental delay, motor abnormalities and cognitive deficits in transgenic mice overexpressing Dyrk1A (minibrain), a murine model of Down's syndrome. Hum Mol Genet *10*, 1915-1923.

Ander, B.P., Barger, N., Stamova, B., Sharp, F.R., and Schumann, C.M. (2015). Atypical miRNA expression in temporal cortex associated with dysregulation of immune, cell cycle, and other pathways in autism spectrum disorders. Mol Autism *6*, 37.

Andin, J., Hallbeck, M., Mohammed, A.H., and Marcusson, J. (2007). Influence of environmental enrichment on steady-state mRNA levels for EAAC1, AMPA1 and NMDA2A receptor subunits in rat hippocampus. Brain Res *1174*, 18-27.

Arendash, G.W., Garcia, M.F., Costa, D.A., Cracchiolo, J.R., Wefes, I.M., and Potter, H. (2004). Environmental enrichment improves cognition in aged Alzheimer's transgenic mice despite stable beta-amyloid deposition. Neuroreport *15*, 1751-1754.

Arion, D., Corradi, J.P., Tang, S., Datta, D., Boothe, F., He, A., Cacace, A.M., Zaczek, R., Albright, C.F., Tseng, G.*, et al.* (2015). Distinctive transcriptome alterations of prefrontal pyramidal neurons in schizophrenia and schizoaffective disorder. Mol Psychiatry *20*, 1397-1405.

Artegiani, B., de Jesus Domingues, A.M., Bragado Alonso, S., Brandl, E., Massalini, S., Dahl, A., and Calegari, F. (2015). Tox: a multifunctional transcription factor and novel regulator of mammalian corticogenesis. EMBO J *34*, 896-910.

Aruga, J., Yokota, N., Hashimoto, M., Furuichi, T., Fukuda, M., and Mikoshiba, K. (1994). A novel zinc finger protein, zic, is involved in neurogenesis, especially in the cell lineage of cerebellar granule cells. J Neurochem *63*, 1880-1890.

Atkins, C.M., Selcher, J.C., Petraitis, J.J., Trzaskos, J.M., and Sweatt, J.D. (1998). The MAPK cascade is required for mammalian associative learning. Nat Neurosci *1*, 602-609.

Auer, M., Hausott, B., and Klimaschewski, L. (2011). Rho GTPases as regulators of morphological neuroplasticity. Ann Anat *193*, 259-266.

Bailey, T., Krajewski, P., Ladunga, I., Lefebvre, C., Li, Q., Liu, T., Madrigal, P., Taslim, C., and Zhang, J. (2013). Practical guidelines for the comprehensive analysis of ChIP-seq data. PLoS Comput Biol *9*, e1003326.

Banday, A.R., Baumgartner, M., Al Seesi, S., Karunakaran, D.K., Venkatesh, A., Congdon, S., Lemoine, C., Kilcollins, A.M., Mandoiu, I., Punzo, C.*, et al.* (2014). Replication-dependent histone genes are actively transcribed in differentiating and aging retinal neurons. Cell Cycle *13*, 2526-2541.

Barbosa, A.C., Kim, M.S., Ertunc, M., Adachi, M., Nelson, E.D., McAnally, J., Richardson, J.A., Kavalali, E.T., Monteggia, L.M., Bassel-Duby, R.*, et al.* (2008). MEF2C, a transcription factor that facilitates learning and memory by negative regulation of synapse numbers and function. Proc Natl Acad Sci U S A *105*, 9391-9396.
Barrett, R.M., and Wood, M.A. (2008). Beyond transcription factors: the role of chromatin modifying enzymes in regulating transcription required for memory. Learn Mem *15*, 460-467.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Bednarek, E., and Caroni, P. (2011). beta-Adducin is required for stable assembly of new synapses and improved memory upon environmental enrichment. Neuron *69*, 1132-1146.

Begenisic, T., Spolidoro, M., Braschi, C., Baroncelli, L., Milanese, M., Pietra, G., Fabbri, M.E., Bonanno, G., Cioni, G., Maffei, L.*, et al.* (2011). Environmental enrichment decreases GABAergic inhibition and improves cognitive abilities, synaptic plasticity, and visual functions in a mouse model of Down syndrome. Front Cell Neurosci *5*, 29.

Beglopoulos, V., Montag-Sallaz, M., Rohlmann, A., Piechotta, K., Ahmad, M., Montag, D., and Missler, M. (2005). Neurexophilin 3 is highly localized in cortical and cerebellar regions and is functionally important for sensorimotor gating and motor coordination. Mol Cell Biol *25*, 7278-7288.

Bellou, S., Hink, M.A., Bagli, E., Panopoulou, E., Bastiaens, P.I., Murphy, C., and Fotsis, T. (2009). VEGF autoregulates its proliferative and migratory ERK1/2 and p38 cascades by enhancing the expression of DUSP1 and DUSP5 phosphatases in endothelial cells. Am J Physiol Cell Physiol *297*, C1477-1489.

Bennett, E.L., Rosenzweig, M.R., and Diamond, M.C. (1969). Rat brain: effects of environmental enrichment on wet and dry weights. Science *163*, 825-826.

Berke, B., Wittnam, J., McNeill, E., Van Vactor, D.L., and Keshishian, H. (2013). Retrograde BMP signaling at the synapse: a permissive signal for synapse maturation and activity-dependent plasticity. J Neurosci *33*, 17937-17950.

Berry, A.S., Blakely, R.D., Sarter, M., and Lustig, C. (2015). Cholinergic capacity mediates prefrontal engagement during challenges to attention: evidence from imaging genetics. Neuroimage *108*, 386-395.

Betty, M., Harnish, S.W., Rhodes, K.J., and Cockett, M.I. (1998). Distribution of heterotrimeric G-protein beta and gamma subunits in the rat brain. Neuroscience *85*, 475-486.

Bindea, G., Galon, J., and Mlecnik, B. (2013). CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics *29*, 661-663.
Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pages, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091-1093.

Bintu, L., Ishibashi, T., Dangkulwanich, M., Wu, Y.Y., Lubkowska, L., Kashlev, M., and Bustamante, C. (2012). Nucleosomal elements that control the topography of the barrier to transcription. Cell *151*, 738-749.

Blaschke, R.J., Monaghan, A.P., Schiller, S., Schechinger, B., Rao, E., Padilla-Nash, H., Ried, T., and Rappold, G.A. (1998). SHOT, a SHOX-related homeobox gene, is implicated in craniofacial, brain, heart, and limb development. Proc Natl Acad Sci U S A *95*, 2406-2411.

Blazquez, G., Canete, T., Tobena, A., Gimenez-Llort, L., and Fernandez-Teruel, A. (2014). Cognitive and emotional profiles of aged Alzheimer's disease (3xTgAD) mice: effects of environmental enrichment and sexual dimorphism. Behav Brain Res *268*, 185-201.

Bose, M., Munoz-Llancao, P., Roychowdhury, S., Nichols, J.A., Jakkamsetti, V., Porter, B., Byrapureddy, R., Salgado, H., Kilgard, M.P., Aboitiz, F*., et al.* (2010). Effect of the environment on the dendritic morphology of the rat auditory cortex. Synapse *64*, 97-110.

Bot, A.M., Debski, K.J., and Lukasiuk, K. (2013). Alterations in miRNA levels in the dentate gyrus in epileptic rats. PLoS One *8*, e76051.

Bousiges, O., Vasconcelos, A.P., Neidl, R., Cosquer, B., Herbeaux, K., Panteleeva, I., Loeffler, J.P., Cassel, J.C., and Boutillier, A.L. (2010). Spatial memory consolidation is associated with induction of several lysine-acetyltransferase (histone acetyltransferase) expression levels and H2B/H4 acetylation-dependent transcriptional events in the rat hippocampus. Neuropsychopharmacology *35*, 2521-2537.

Boyle, A.P., Guinney, J., Crawford, G.E., and Furey, T.S. (2008). F-Seq: a feature density estimator for high-throughput sequence tags. Bioinformatics *24*, 2537-2538.
Bramham, C.R., and Messaoudi, E. (2005). BDNF function in adult synaptic plasticity: the synaptic consolidation hypothesis. Prog Neurobiol *76*, 99-125.

Bruses, J.L. (2010). Identification of gene transcripts expressed by postsynaptic neurons during synapse formation encoding cell surface proteins with presumptive synaptogenic activity. Synapse *64*, 47-60.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods *10*, 1213-1218.

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. Nature *523*, 486-490.

Butterfield, D.A., and Lange, M.L. (2009). Multifunctional roles of enolase in Alzheimer's disease brain: beyond altered glucose metabolism. J Neurochem *111*, 915-933.

Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A.*, et al.* (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. J Neurosci *28*, 264-278.

Caleo, M., Tropea, D., Rossi, C., Gianfranceschi, L., and Maffei, L. (2009). Environmental enrichment promotes fiber sprouting after deafferentation of the superior colliculus in the adult rat brain. Exp Neurol *216*, 515-519.

Cao, L., Jiao, X., Zuzga, D.S., Liu, Y., Fong, D.M., Young, D., and During, M.J. (2004). VEGF links hippocampal activity with neurogenesis, learning and memory. Nat Genet *36*, 827-835.

Cao, L., Liu, X., Lin, E.J., Wang, C., Choi, E.Y., Riban, V., Lin, B., and During, M.J. (2010). Environmental and genetic activation of a brain-adipocyte BDNF/leptin axis causes cancer remission and inhibition. Cell *142*, 52-64.

Caunt, C.J., and Keyse, S.M. (2013). Dual-specificity MAP kinase phosphatases (MKPs): shaping the outcome of MAP kinase signalling. FEBS J *280*, 489-504.
Caygill, E.E., and Johnston, L.A. (2008). Temporal regulation of metamorphic processes in Drosophila by the let-7 and miR-125 heterochronic microRNAs. Curr Biol *18*, 943-950.

Cech, T.R., and Steitz, J.A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. Cell *157*, 77-94.

Cermenati, S., Moleri, S., Cimbro, S., Corti, P., Del Giacco, L., Amodeo, R., Dejana, E., Koopman, P., Cotelli, F., and Beltrame, M. (2008). Sox18 and Sox7 play redundant roles in vascular development. Blood *111*, 2657-2666.

Chen, C., Ridzon, D.A., Broomer, A.J., Zhou, Z., Lee, D.H., Nguyen, J.T., Barbisin, M., Xu, N.L., Mahuvakar, V.R., Andersen, M.R.*, et al.* (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res *33*, e179.

Chen, C., Wirth, A., and Ponimaskin, E. (2012). Cdc42: an important regulator of neuronal morphology. Int J Biochem Cell Biol *44*, 447-451.

Cho, E.G., Zaremba, J.D., McKercher, S.R., Talantova, M., Tu, S., Masliah, E., Chan, S.F., Nakanishi, N., Terskikh, A., and Lipton, S.A. (2011). MEF2C enhances dopaminergic neuron differentiation of human embryonic stem cells in a parkinsonian rat model. PLoS One *6*, e24027.

Cho, H.H., Cargnin, F., Kim, Y., Lee, B., Kwon, R.J., Nam, H., Shen, R., Barnes, A.P., Lee, J.W., Lee, S.*, et al.* (2014). Isl1 directly controls a cholinergic neuronal identity in the developing forebrain and spinal cord by forming cell type-specific complexes. PLoS Genet *10*, e1004280.

Chwang, W.B., O'Riordan, K.J., Levenson, J.M., and Sweatt, J.D. (2006). ERK/MAPK regulates hippocampal histone phosphorylation following contextual fear conditioning. Learn Mem *13*, 322-328.

Ciani, L., Boyle, K.A., Dickins, E., Sahores, M., Anane, D., Lopes, D.M., Gibb, A.J., and Salinas, P.C. (2011). Wnt7a signaling promotes dendritic spine growth and synaptic strength through Ca(2)(+)/Calmodulin-dependent protein kinase II. Proc Natl Acad Sci U S A *108*, 10732-10737.

Cohen, H., Liu, T., Kozlovsky, N., Kaplan, Z., Zohar, J., and Mathe, A.A. (2012). The neuropeptide Y (NPY)-ergic system is associated with behavioral resilience to stress exposure in an animal model of post-traumatic stress disorder. Neuropsychopharmacology *37*, 350-363.

Collingridge, G.L., Peineau, S., Howland, J.G., and Wang, Y.T. (2010). Long-term depression in the CNS. Nat Rev Neurosci *11*, 459-473.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szczesniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X.*, et al.* (2016). A survey of best practices for RNA-seq data analysis. Genome Biol *17*, 13.

Cosin-Tomas, M., Alvarez-Lopez, M.J., Sanchez-Roige, S., Lalanza, J.F., Bayod, S., Sanfeliu, C., Pallas, M., Escorihuela, R.M., and Kaliman, P. (2014). Epigenetic alterations in hippocampus of SAMP8 senescent mice and modulation by voluntary physical exercise. Front Aging Neurosci *6*, 51.

Coskran, T.M., Morton, D., Menniti, F.S., Adamowicz, W.O., Kleiman, R.J., Ryan, A.M., Strick, C.A., Schmidt, C.J., and Stephenson, D.T. (2006). Immunohistochemical localization of phosphodiesterase 10A in multiple mammalian species. J Histochem Cytochem *54*, 1205-1213.

Costa, D.A., Cracchiolo, J.R., Bachstetter, A.D., Hughes, T.F., Bales, K.R., Paul, S.M., Mervis, R.F., Arendash, G.W., and Potter, H. (2007). Enrichment improves cognition in AD mice by amyloid-related and unrelated mechanisms. Neurobiol Aging *28*, 831-844.

Cross, S.H., Meehan, R.R., Nan, X., and Bird, A. (1997). A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. Nat Genet *16*, 256-259.

Cui, X., Kawashima, H., Barclay, T.B., Peters, J.M., Gonzalez, F.J., Morgan, E.T., and Strobel, H.W. (2001). Molecular cloning and regulation of expression of two novel mouse CYP4F genes: expression in peroxisome proliferator-activated receptor alpha-deficient mice upon lipopolysaccharide and clofibrate challenges. J Pharmacol Exp Ther *296*, 542-550.

D'Alessio, A.C., Weaver, I.C., and Szyf, M. (2007). Acetylation-induced transcription is required for active DNA demethylation in methylation-silenced genes. Mol Cell Biol *27*, 7462-7474.

Davis, S., Bozon, B., and Laroche, S. (2003). How necessary is the activation of the immediate early gene zif268 in synaptic plasticity and learning? Behavioural Brain Research *142*, 17-30.

Day, J.J., and Sweatt, J.D. (2011). Epigenetic mechanisms in cognition. Neuron *70*, 813-829.

de la Torre-Ubieta, L., and Bonni, A. (2011). Transcriptional regulation of neuronal polarity and morphogenesis in the mammalian brain. Neuron *72*, 22-40.
Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev *25*, 1010-1022.

Del Arco, A., Segovia, G., Canales, J.J., Garrido, P., de Blas, M., Garcia-Verdugo, J.M., and Mora, F. (2007a). Environmental enrichment reduces the function of D1 dopamine receptors in the prefrontal cortex of the rat. J Neural Transm (Vienna) *114*, 43-48.

Del Arco, A., Segovia, G., Garrido, P., de Blas, M., and Mora, F. (2007b). Stress, prefrontal cortex and environmental enrichment: studies on dopamine and acetylcholine release and working memory performance in rats. Behav Brain Res *176*, 267-273.

Denning, G.M., Ackermann, L.W., Barna, T.J., Armstrong, J.G., Stoll, L.L., Weintraub, N.L., and Dickson, E.W. (2008). Proenkephalin expression and enkephalin release are widely observed in non-neuronal tissues. Peptides *29*, 83-92.

Dobransky, T., and Rylett, R.J. (2005). A model for dynamic regulation of choline acetyltransferase by phosphorylation. J Neurochem *95*, 305-313.

Dong, X., and Weng, Z. (2013). The correlation between histone modifications and gene expression. Epigenomics *5*, 113-116.

Dottori, M., Gross, M.K., Labosky, P., and Goulding, M. (2001). The winged-helix transcription factor Foxd3 suppresses interneuron differentiation and promotes neural crest cell fate. Development *128*, 4127-4138.

Dufour, A., Seibt, J., Passante, L., Depaepe, V., Ciossek, T., Frisén, J., Kullander, K., Flanagan, J.G., Polleux, F., and Vanderhaeghen, P. (2003). Area Specificity and Topography of Thalamocortical Projections Are Controlled by ephrin/Eph Genes. Neuron *39*, 453-465.

Durairaj, R.V., and Koilmani, E.R. (2014). Environmental enrichment modulates glucocorticoid receptor expression and reduces anxiety in Indian field male mouse Mus booduga through up-regulation of microRNA-124a. Gen Comp Endocrinol *199*, 26-32.

Dzamba, D., Honsa, P., Valny, M., Kriska, J., Valihrach, L., Novosadova, V., Kubista, M., and Anderova, M. (2015). Quantitative Analysis of Glutamate Receptors in Glial Cells from the Cortex of GFAP/EGFP Mice Following Ischemic Injury: Focus on NMDA Receptors. Cell Mol Neurobiol *35*, 1187-1202.

Eacker, S.M., Dawson, T.M., and Dawson, V.L. (2013). The interplay of microRNA and neuronal activity in health and disease. Front Cell Neurosci *7*, 136.

Edbauer, D., Neilson, J.R., Foster, K.A., Wang, C.F., Seeburg, D.P., Batterton, M.N., Tada, T., Dolan, B.M., Sharp, P.A., and Sheng, M. (2010). Regulation of synaptic structure and function by FMRP-associated microRNAs miR-125b and miR-132. Neuron *65*, 373-384.

English, J.D., and Sweatt, J.D. (1996). Activation of p42 mitogen-activated protein kinase in hippocampal long term potentiation. J Biol Chem *271*, 24329-24332.

English, J.D., and Sweatt, J.D. (1997). A requirement for the mitogen-activated protein kinase cascade in hippocampal long term potentiation. J Biol Chem *272*, 19103-19106.
Escalante, A., Murillo, B., Morenilla-Palao, C., Klar, A., and Herrera, E. (2013). Zic2-dependent axon midline avoidance controls the formation of major ipsilateral tracts in the CNS. Neuron *80*, 1392-1406.

Fabel, K., Wolf, S.A., Ehninger, D., Babu, H., Leal-Galicia, P., and Kempermann, G. (2009). Additive effects of physical exercise and environmental enrichment on adult hippocampal neurogenesis in mice. Front Neurosci *3*, 50.

Faherty, C.J., Raviie Shepherd, K., Herasimtschuk, A., and Smeyne, R.J. (2005). Environmental enrichment in adulthood eliminates neuronal death in experimental Parkinsonism. Brain Res Mol Brain Res *134*, 170-179.

Falkenberg, T., Mohammed, A.K., Henriksson, B., Persson, H., Winblad, B., and Lindefors, N. (1992). Increased expression of brain-derived neurotrophic factor mRNA in rat hippocampus is associated with improved spatial memory and enriched environment. Neurosci Lett *138*, 153-156.

Fang, L., Cai, J., Chen, B., Wu, S., Li, R., Xu, X., Yang, Y., Guan, H., Zhu, X., Zhang, L.*, et al.* (2015). Aberrantly expressed miR-582-3p maintains lung cancer stem cell-like traits by activating Wnt/beta-catenin signalling. Nat Commun *6*, 8640.

Feng, J., Zhou, Y., Campbell, S.L., Le, T., Li, E., Sweatt, J.D., Silva, A.J., and Fan, G. (2010). Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. Nat Neurosci *13*, 423-430.

Fischer, A., Sananbenesi, F., Wang, X., Dobbin, M., and Tsai, L.H. (2007). Recovery of learning and memory is associated with chromatin remodelling. Nature *447*, 178-182.
Flavell, S.W., and Greenberg, M.E. (2008). Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system. Annu Rev Neurosci *31*, 563-590.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suner, D., Cigudosa, J.C., Urioste, M., Benitez, J.*, et al.* (2005). Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A *102*, 10604-10609.

Francis, D.D., Diorio, J., Plotsky, P.M., and Meaney, M.J. (2002). Environmental enrichment reverses the effects of maternal separation on stress reactivity. J Neurosci *22*, 7840-7843.

Freeman, A., Franciscovich, A., Bowers, M., Sandstrom, D.J., and Sanyal, S. (2011). NFAT regulates pre-synaptic development and activity-dependent plasticity in Drosophila. Mol Cell Neurosci *46*, 535-547.

Freund, A., Zhong, F.L., Venteicher, A.S., Meng, Z., Veenstra, T.D., Frydman, J., and Artandi, S.E. (2014). Proteostatic control of telomerase function through TRiC-mediated folding of TCAB1. Cell *159*, 1389-1403.

Friedman, Y., Karsenty, S., and Linial, M. (2014). miRror-Suite: decoding coordinated regulation by microRNAs. Database (Oxford) *2014*.

Gabel, H.W., Kinde, B., Stroud, H., Gilbert, C.S., Harmin, D.A., Kastan, N.R., Hemberg, M., Ebert, D.H., and Greenberg, M.E. (2015). Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. Nature *522*, 89-93.

Gagne, J., Gelinas, S., Martinoli, M.G., Foster, T.C., Ohayon, M., Thompson, R.F., Baudry, M., and Massicotte, G. (1998). AMPA receptor properties in adult rat hippocampus following environmental enrichment. Brain Res *799*, 16-25.

Gainetdinov, R.R., Premont, R.T., Bohn, L.M., Lefkowitz, R.J., and Caron, M.G. (2004). Desensitization of G protein-coupled receptors and neuronal functions. Annu Rev Neurosci *27*, 107-144.

Gangemi, R.M., Daga, A., Marubbi, D., Rosatto, N., Capra, M.C., and Corte, G. (2001). Emx2 in adult neural precursor cells. Mech Dev *109*, 323-329.

Gao, J., Kang, X.Y., Sun, S., Li, L., Zhang, B.L., Li, Y.Q., and Gao, D.S. (2016). Transcription factor Six2 mediates the protection of GDNF on 6-OHDA lesioned dopaminergic neurons by regulating Smurf1 expression. Cell Death Dis *7*, e2217.

Garofalo, S., D'Alessandro, G., Chece, G., Brau, F., Maggi, L., Rosa, A., Porzia, A., Mainiero, F., Esposito, V., Lauro, C.*, et al.* (2015). Enriched environment reduces glioma growth through immune and non-immune mechanisms in mice. Nat Commun *6*, 6623.

Giunta, M., Edvardson, S., Xu, Y., Schuelke, M., Gomez-Duran, A., Boczonadi, V., Elpeleg, O., Muller, J.S., and Horvath, R. (2016). Altered RNA metabolism due to a homozygous RBM7 mutation in a patient with spinal motor neuropathy. Hum Mol Genet.

Gobeske, K.T., Das, S., Bonaguidi, M.A., Weiss, C., Radulovic, J., Disterhoft, J.F., and Kessler, J.A. (2009). BMP signaling mediates effects of exercise on hippocampal neurogenesis and cognition in mice. PLoS One *4*, e7506.

Gogolla, N., Galimberti, I., Deguchi, Y., and Caroni, P. (2009). Wnt signaling mediates experience-related regulation of synapse numbers and mossy fiber connectivities in the adult hippocampus. Neuron *62*, 510-525.

Gomez, A.M., Altomare, D., Sun, W.L., Midde, N.M., Ji, H., Shtutman, M., Turner, J.R., Creek, K.E., and Zhu, J. (2016). Prefrontal microRNA-221 Mediates Environmental Enrichment-Induced Increase of Locomotor Sensitivity to Nicotine. Int J Neuropsychopharmacol *19*.

Govindarajan, N., Rao, P., Burkhardt, S., Sananbenesi, F., Schluter, O.M., Bradke, F., Lu, J., and Fischer, A. (2013). Reducing HDAC6 ameliorates cognitive deficits in a mouse model for Alzheimer's disease. EMBO Mol Med *5*, 52-63.

Graff, J., and Tsai, L.H. (2013). Histone acetylation: molecular mnemonics on the chromatin. Nat Rev Neurosci *14*, 97-111.

Graff, J., Woldemichael, B.T., Berchtold, D., Dewarrat, G., and Mansuy, I.M. (2012). Dynamic histone marks in the hippocampus and cortex facilitate memory consolidation. Nat Commun *3*, 991.

Greenough, W.T., and Volkmar, F.R. (1973). Pattern of dendritic branching in occipital cortex of rats reared in complex environments. Exp Neurol *40*, 491-504.

Greer, P.L., and Greenberg, M.E. (2008). From synapse to nucleus: calcium-dependent gene transcription in the control of synapse development and function. Neuron *59*, 846-860.

Grippo, A.J., Ihm, E., Wardwell, J., McNeal, N., Scotti, M.A., Moenk, D.A., Chandler, D.L., LaRocca, M.A., and Preihs, K. (2014). The effects of environmental enrichment on depressive and anxiety-relevant behaviors in socially isolated prairie voles. Psychosom Med *76*, 277-284.

Gupta, S., Kim, S.Y., Artis, S., Molfese, D.L., Schumacher, A., Sweatt, J.D., Paylor, R.E., and Lubin, F.D. (2010). Histone methylation regulates memory formation. J Neurosci *30*, 3589-3599.

Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol *15*, 509-524.

Habermehl, D., Kempna, P., Azzi, A., and Zingg, J.M. (2005). Recombinant SEC14-like proteins (TAP) possess GTPase activity. Biochem Biophys Res Commun *326*, 254-259.

Halder, R., Hennion, M., Vidal, R.O., Shomroni, O., Rahman, R.U., Rajput, A., Centeno, T.P., van Bebber, F., Capece, V., Garcia Vizcaino, J.C*., et al.* (2016). DNA methylation changes in plasticity genes accompany the formation and maintenance of memory. Nat Neurosci *19*, 102-110.

Hamasaki, T., Leingartner, A., Ringstedt, T., and O'Leary, D.D. (2004). EMX2 regulates sizes and positioning of the primary sensory and motor areas in neocortex by direct specification of cortical progenitors. Neuron *43*, 359-372.

Hardingham, N., Dachtler, J., and Fox, K. (2013). The role of nitric oxide in pre-synaptic plasticity and homeostasis. Front Cell Neurosci *7*, 190.

Harmanci, A., Rozowsky, J., and Gerstein, M. (2014). MUSIC: identification of enriched regions in ChIP-Seq experiments using a mappability-corrected multiscale signal processing framework. Genome Biol *15*, 474.

Hayashi, M.K., Tang, C., Verpelli, C., Narayanan, R., Stearns, M.H., Xu, R.M., Li, H., Sala, C., and Hayashi, Y. (2009). The postsynaptic density proteins Homer and Shank form a polymeric network structure. Cell *137*, 159-171.

He, Y., Yang, C., Kirkmire, C.M., and Wang, Z.J. (2010). Regulation of opioid tolerance by let-7 family microRNA targeting the mu opioid receptor. J Neurosci *30*, 10251-10258.

Hebb DO (1947). "The effects of early experience on problem solving at maturity". American Psychologist. **2**: 306–7.

Hebb, D.O. (1949). The Organization of Behavior. New York: Wiley & Sons.

Heinla, I., Leidmaa, E., Kongi, K., Pennert, A., Innos, J., Nurk, K., Tekko, T., Singh, K., Vanaveski, T., Reimets, R*., et al.* (2015). Gene expression patterns and environmental enrichment-induced effects in the hippocampi of mice suggest importance of Lsamp in plasticity. Front Neurosci *9*, 205.

Heinrich, J., Proepper, C., Schmidt, T., Linta, L., Liebau, S., and Boeckers, T.M. (2012). The postsynaptic density protein Abelson interactor protein 1 interacts with the motor protein Kinesin family member 26B in hippocampal neurons. Neuroscience *221*, 86-95.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.
Hendriksen, H., Bink, D.I., Daniels, E.G., Pandit, R., Piriou, C., Slieker, R., Westphal, K.G., Olivier, B., and Oosting, R.S. (2012). Re-exposure and environmental enrichment reveal NPY-Y1 as a possible target for post-traumatic stress disorder. Neuropharmacology *63*, 733-742.

Henninger, N., Feldmann, R.E., Jr., Futterer, C.D., Schrempp, C., Maurer, M.H., Waschke, K.F., Kuschinsky, W., and Schwab, S. (2007). Spatial learning induces predominant downregulation of cytosolic proteins in the rat hippocampus. Genes Brain Behav *6*, 128-140.

Hernandez, A.I., Alarcon, J.M., and Allen, K.D. (2015). New ribosomes for new memories? Commun Integr Biol *8*, e1017163.

Herring, B.E., and Nicoll, R.A. (2016). Long-Term Potentiation: From CaMKII to AMPA Receptor Trafficking. Annu Rev Physiol *78*, 351-365.

Ho, N., Liauw, J.A., Blaeser, F., Wei, F., Hanissian, S., Muglia, L.M., Wozniak, D.F., Nardi, A., Arvin, K.L., Holtzman, D.M.*, et al.* (2000). Impaired synaptic plasticity and cAMP response element-binding protein activation in Ca2+/calmodulin-dependent protein kinase type IV/Gr-deficient mice. J Neurosci *20*, 6459-6472.

Hockly, E., Cordery, P.M., Woodman, B., Mahal, A., van Dellen, A., Blakemore, C., Lewis, C.M., Hannan, A.J., and Bates, G.P. (2002). Environmental enrichment slows disease progression in R6/2 Huntington's disease mice. Ann Neurol *51*, 235-242.
Holloway, R.L., Jr. (1966). Dendritic branching: some preliminary results of training and complexity in rat visual cortex. Brain Res *2*, 393-396.

Hsu, S.C., Ting, A.E., Hazuka, C.D., Davanger, S., Kenny, J.W., Kee, Y., and Scheller, R.H. (1996). The mammalian brain rsec6/8 complex. Neuron *17*, 1209-1219.
Hu, Y.S., Long, N., Pigino, G., Brady, S.T., and Lazarov, O. (2013). Molecular mechanisms of environmental enrichment: impairments in Akt/GSK3beta, neurotrophin-3 and CREB signaling. PLoS One *8*, e64460.

Hu, Z., Zhao, J., Hu, T., Luo, Y., Zhu, J., and Li, Z. (2015). miR-501-3p mediates the activity-dependent regulation of the expression of AMPA receptor subunit GluA1. J Cell Biol *208*, 949-959.

Huff, J.T., Plocik, A.M., Guthrie, C., and Yamamoto, K.R. (2010). Reciprocal intronic and exonic histone modification regions in humans. Nat Struct Mol Biol *17*, 1495-1499.
Huo, Y., Khatri, N., Hou, Q., Gilbert, J., Wang, G., and Man, H.Y. (2015). The deubiquitinating enzyme USP46 regulates AMPA receptor ubiquitination and trafficking. J Neurochem *134*, 1067-1080.

Ickes, B.R., Pham, T.M., Sanders, L.A., Albeck, D.S., Mohammed, A.H., and Granholm, A.C. (2000). Long-term environmental enrichment leads to regional increases in neurotrophin levels in rat brain. Exp Neurol *164*, 45-52.

Ikeda, M., Kanao, Y., Yamanaka, M., Sakuraba, H., Mizutani, Y., Igarashi, Y., and Kihara, A. (2008). Characterization of four mammalian 3-hydroxyacyl-CoA dehydratases involved in very long-chain fatty acid synthesis. FEBS Lett *582*, 2435-2440.

Irier, H., Street, R.C., Dave, R., Lin, L., Cai, C., Davis, T.H., Yao, B., Cheng, Y., and Jin, P. (2014). Environmental enrichment modulates 5-hydroxymethylcytosine dynamics in hippocampus. Genomics *104*, 376-382.

Jadavji, N.M., Kolb, B., and Metz, G.A. (2006). Enriched environment improves motor function in intact and unilateral dopamine-depleted rats. Neuroscience *140*, 1127-1138.
Jain, P., and Bhalla, U.S. (2014). Transcription control pathways decode patterned synaptic inputs into diverse mRNA expression profiles. PLoS One *9*, e95154.

Jankowsky, J.L., Melnikova, T., Fadale, D.J., Xu, G.M., Slunt, H.H., Gonzales, V., Younkin, L.H., Younkin, S.G., Borchelt, D.R., and Savonenko, A.V. (2005).

Environmental enrichment mitigates cognitive deficits in a mouse model of Alzheimer's disease. J Neurosci *25*, 5217-5224.

Janky, R., Verfaillie, A., Imrichova, H., Van de Sande, B., Standaert, L., Christiaens, V., Hulselmans, G., Herten, K., Naval Sanchez, M., Potier, D.*, et al.* (2014). iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput Biol *10*, e1003731.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Ji, D., Lape, R., and Dani, J.A. (2001). Timing and location of nicotinic activity enhances or depresses hippocampal synaptic plasticity. Neuron *31*, 131-141.
Jin, S.G., Wu, X., Li, A.X., and Pfeifer, G.P. (2011). Genomic mapping of 5-hydroxymethylcytosine in the human brain. Nucleic Acids Res *39*, 5015-5024.
Jonas, S., and Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet *16*, 421-433.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet *19*, 187-191.

Jonkers, I., and Lis, J.T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. Nat Rev Mol Cell Biol *16*, 167-177.

Just, L., Olenik, C., Heimrich, B., and Meyer, D.K. (1998). Glutamatergic control of the expression of the proenkephalin gene in rat frontoparietal cortical slice cultures. Cereb Cortex *8*, 702-709.

Kahne, T., Kolodziej, A., Smalla, K.H., Eisenschmidt, E., Haus, U.U., Weismantel, R., Kropf, S., Wetzel, W., Ohl, F.W., Tischmeyer, W.*, et al.* (2012). Synaptic proteome

changes in mouse brain regions upon auditory discrimination learning. Proteomics *12*, 2433-2444.

Kahne, T., Richter, S., Kolodziej, A., Smalla, K.H., Pielot, R., Engler, A., Ohl, F.W., Dieterich, D.C., Seidenbecher, C., Tischmeyer, W.*, et al.* (2016). Proteome rearrangements after auditory learning: high-resolution profiling of synapse-enriched protein fractions from mouse brain. J Neurochem *138*, 124-138.

Kapranov, P., Willingham, A.T., and Gingeras, T.R. (2007). Genome-wide transcription and the implications for genomic organization. Nat Rev Genet *8*, 413-423.
Karow, M., and Berninger, B. (2015). In-TOX-icating neurogenesis. EMBO J *34*, 832-834.

Karpova, N.N. (2014). Role of BDNF epigenetics in activity-dependent neuronal plasticity. Neuropharmacology *76 Pt C*, 709-718.

Kawamura, N., Sun-Wada, G.H., and Wada, Y. (2015). Loss of G2 subunit of vacuolar-type proton transporting ATPase leads to G1 subunit upregulation in the brain. Sci Rep *5*, 14027.

Kempermann, G., Kuhn, H.G., and Gage, F.H. (1997). More hippocampal neurons in adult mice living in an enriched environment. Nature *386*, 493-495.

Kempná, P., Zingg, J.-M., Ricciarelli, R., Hierl, M., Saxena, S., and Azzi, A. (2003). Cloning of novel human SEC14p-like proteins: ligand binding and functional properties. Free Radical Biology and Medicine *34*, 1458-1472.

Keshet, E., Polakiewicz, R.D., Itin, A., Ornoy, A., and Rosen, H. (1989). Proenkephalin A is expressed in mesodermal lineages during organogenesis. EMBO J *8*, 2917-2923.
Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Konig, M., Zimmer, A.M., Steiner, H., Holmes, P.V., Crawley, J.N., Brownstein, M.J., and Zimmer, A. (1996). Pain responses, anxiety and aggression in mice deficient in pre-proenkephalin. Nature *383*, 535-538.

Korzus, E., Rosenfeld, M.G., and Mayford, M. (2004). CBP histone acetyltransferase activity is a critical component of memory consolidation. Neuron *42*, 961-972.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.
Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science *324*, 929-930.

Kronenberg, G., Reuter, K., Steiner, B., Brandt, M.D., Jessberger, S., Yamaguchi, M., and Kempermann, G. (2003). Subpopulations of proliferating cells of the adult hippocampus respond differently to physiologic neurogenic stimuli. J Comp Neurol *467*, 455-463.

Kurshan, P.T., Phan, A.Q., Wang, G.J., Crane, M.M., Lu, H., and Shen, K. (2014). Regulation of synaptic extracellular matrix composition is critical for proper synapse morphology. J Neurosci *34*, 12678-12689.
Kuzumaki, N., Ikegami, D., Tamura, R., Hareyama, N., Imai, S., Narita, M., Torigoe, K., Niikura, K., Takeshima, H., Ando, T.*, et al.* (2011). Hippocampal epigenetic modification at the brain-derived neurotrophic factor gene induced by an enriched environment. Hippocampus *21*, 127-132.

Landt, S.G., Marinov, G.K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B.E., Bickel, P., Brown, J.B., Cayting, P.*, et al.* (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res *22*, 1813-1831.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

LaPlant, Q., Vialou, V., Covington, H.E., 3rd, Dumitriu, D., Feng, J., Warren, B.L., Maze, I., Dietz, D.M., Watts, E.L., Iniguez, S.D.*, et al.* (2010). Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens. Nat Neurosci *13*, 1137-1143.

Larance, M., and Lamond, A.I. (2015). Multidimensional proteomics for cell biology. Nat Rev Mol Cell Biol *16*, 269-280.

Lazarov, O., Robinson, J., Tang, Y.P., Hairston, I.S., Korade-Mirnics, Z., Lee, V.M., Hersh, L.B., Sapolsky, R.M., Mirnics, K., and Sisodia, S.S. (2005). Environmental enrichment reduces Abeta levels and amyloid deposition in transgenic mice. Cell *120*, 701-713.

Le, M.T., Xie, H., Zhou, B., Chia, P.H., Rizk, P., Um, M., Udolph, G., Yang, H., Lim, B., and Lodish, H.F. (2009). MicroRNA-125b promotes neuronal differentiation in human cells by repressing multiple targets. Mol Cell Biol *29*, 5290-5305.

Leal, G., Comprido, D., and Duarte, C.B. (2014). BDNF-induced local protein synthesis and synaptic plasticity. Neuropharmacology *76 Pt C*, 639-656.

Lee, J., Sayegh, J., Daniel, J., Clarke, S., and Bedford, M.T. (2005). PRMT8, a new membrane-bound tissue-specific member of the protein arginine methyltransferase family. J Biol Chem *280*, 32890-32896.

Lee, M.S., and Kerns, E.H. (1999). LC/MS applications in drug development. Mass Spectrom Rev *18*, 187-279.

Leidi, M., Mariotti, M., and Maier, J.A. (2010). EDF-1 contributes to the regulation of nitric oxide release in VEGF-treated human endothelial cells. Eur J Cell Biol *89*, 654-660.

Lepoivre, C., Belhocine, M., Bergon, A., Griffon, A., Yammine, M., Vanhille, L., Zacarias-Cabeza, J., Garibal, M.A., Koch, F., Maqbool, M.A.*, et al.* (2013). Divergent

transcription is associated with promoters of transcriptional regulators. BMC Genomics *14*, 914.
Levenson, J.M., O'Riordan, K.J., Brown, K.D., Trinh, M.A., Molfese, D.L., and Sweatt, J.D. (2004). Regulation of histone acetylation during memory formation in the hippocampus. J Biol Chem *279*, 40545-40559.

Levi, O., Jongen-Relo, A.L., Feldon, J., Roses, A.D., and Michaelson, D.M. (2003). ApoE4 impairs hippocampal plasticity isoform-specifically and blocks the environmental stimulation of synaptogenesis and memory. Neurobiology of Disease *13*, 273-282.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078-2079.

Li, Q., Li, X., Wang, L., Zhang, Y., and Chen, L. (2016). miR-98-5p Acts as a Target for Alzheimer's Disease by Regulating Abeta Production Through Modulating SNX6 Expression. J Mol Neurosci.

Liang, W.S., Reiman, E.M., Valla, J., Dunckley, T., Beach, T.G., Grover, A., Niedzielko, T.L., Schneider, L.E., Mastroeni, D., Caselli, R.*, et al.* (2008). Alzheimer's disease is associated with reduced expression of energy metabolism genes in posterior cingulate neurons. Proc Natl Acad Sci U S A *105*, 4441-4446.

Liang, X., Song, M.R., Xu, Z., Lanuza, G.M., Liu, Y., Zhuang, T., Chen, Y., Pfaff, S.L., Evans, S.M., and Sun, Y. (2011). Isl1 is required for multiple aspects of motor neuron development. Mol Cell Neurosci *47*, 215-222.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

Lichti, C.F., Fan, X., English, R.D., Zhang, Y., Li, D., Kong, F., Sinha, M., Andersen, C.R., Spratt, H., Luxon, B.A.*, et al.* (2014). Environmental enrichment alters protein expression as well as the proteomic response to cocaine in rat nucleus accumbens. Front Behav Neurosci *8*, 246.

Lin, C.I., Orlov, I., Ruggiero, A.M., Dykes-Hoberg, M., Lee, A., Jackson, M., and Rothstein, J.D. (2001). Modulation of the neuronal glutamate transporter EAAC1 by the interacting protein GTRAP3-18. Nature *410*, 84-88.

Lindvall, O., Kokaia, Z., Bengzon, J., Elmer, E., and Kokaia, M. (1994). Neurotrophins and brain insults. Trends Neurosci *17*, 490-496.

Lisman, J., Schulman, H., and Cline, H. (2002). The molecular basis of CaMKII function in synaptic and behavioural memory. Nat Rev Neurosci *3*, 175-190.
Lopez-Atalaya, J.P., and Barco, A. (2014). Can changes in histone acetylation contribute to memory formation? Trends Genet *30*, 529-539.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

Lu, Y., Christian, K., and Lu, B. (2008). BDNF: a key regulator for protein synthesis-dependent LTP and long-term memory? Neurobiol Learn Mem *89*, 312-323.

Lun, A.T., and Smyth, G.K. (2016). csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. Nucleic Acids Res *44*, e45.

Luscher, C., and Malenka, R.C. (2012). NMDA receptor-dependent long-term potentiation and long-term depression (LTP/LTD). Cold Spring Harb Perspect Biol *4*.

Luscher, C., and Slesinger, P.A. (2010). Emerging roles for G protein-gated inwardly rectifying potassium (GIRK) channels in health and disease. Nat Rev Neurosci *11*, 301-315.

Ma, P., Zhao, S., Zeng, W., Yang, Q., Li, C., Lv, X., Zhou, Q., and Mao, B. (2011). Xenopus Dbx2 is involved in primary neurogenesis and early neural plate patterning. Biochem Biophys Res Commun *412*, 170-174.

Marat, A.L., Dokainish, H., and McPherson, P.S. (2011). DENN domain proteins: regulators of Rab GTPases. J Biol Chem *286*, 13791-13800.

Marzluff, W.F., Gongidi, P., Woods, K.R., Jin, J., and Maltais, L.J. (2002). The Human and Mouse Replication-Dependent Histone Genes. Genomics *80*, 487-498.

Matevossian, A., and Akbarian, S. (2008). Neuronal nuclei isolation from human postmortem brain tissue. J Vis Exp.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol *28*, 495-501.

McNair, K., Broad, J., Riedel, G., Davies, C.H., and Cobb, S.R. (2007). Global changes in the hippocampal proteome following exposure to an enriched environment. Neuroscience *145*, 413-422.

Melendez, R.I., Gregory, M.L., Bardo, M.T., and Kalivas, P.W. (2004). Impoverished rearing environment alters metabotropic glutamate receptor expression and function in the prefrontal cortex. Neuropsychopharmacology *29*, 1980-1987.

Mellios, N., Huang, H.S., Grigorenko, A., Rogaev, E., and Akbarian, S. (2008). A set of differentially expressed miRNAs, including miR-30a-5p, act as post-transcriptional inhibitors of BDNF in prefrontal cortex. Hum Mol Genet *17*, 3030-3042.

Meyers, E.A., Gobeske, K.T., Bond, A.M., Jarrett, J.C., Peng, C.Y., and Kessler, J.A. (2016). Increased bone morphogenetic protein signaling contributes to age-related declines in neurogenesis and cognition. Neurobiol Aging *38*, 164-175.

Miller, C.A., Gavin, C.F., White, J.A., Parrish, R.R., Honasoge, A., Yancey, C.R., Rivera, I.M., Rubio, M.D., Rumbaugh, G., and Sweatt, J.D. (2010). Cortical DNA methylation maintains remote memory. Nat Neurosci *13*, 664-666.

Miller, C.A., and Sweatt, J.D. (2007). Covalent modification of DNA regulates memory formation. Neuron *53*, 857-869.

Minatohara, K., Akiyoshi, M., and Okuno, H. (2015). Role of Immediate-Early Genes in Synaptic Plasticity and Neuronal Ensembles Underlying the Memory Trace. Front Mol Neurosci *8*, 78.

Missler, M., Zhang, W., Rohlmann, A., Kattenstroth, G., Hammer, R.E., Gottmann, K., and Sudhof, T.C. (2003). Alpha-neurexins couple Ca2+ channels to synaptic vesicle exocytosis. Nature *423*, 939-948.

Misteli, T., Gunjan, A., Hock, R., Bustin, M., and Brown, D.T. (2000). Dynamic binding of histone H1 to chromatin in living cells. Nature *408*, 877-881.

Mohammed, A.H., Henriksson, B.G., Soderstrom, S., Ebendal, T., Olsson, T., and Seckl, J.R. (1993). Environmental influences on the central nervous system and their implications for the aging rat. Behav Brain Res *57*, 183-191.

Mohanan, S., Cherrington, B.D., Horibata, S., McElwee, J.L., Thompson, P.R., and Coonrod, S.A. (2012). Potential role of peptidylarginine deiminase enzymes and protein citrullination in cancer pathogenesis. Biochem Res Int *2012*, 895343.

Moraes, F., and Goes, A. (2016). A decade of human genome project conclusion: Scientific diffusion about our genome knowledge. Biochem Mol Biol Educ *44*, 215-223.

Mori Sequeiros Garcia, M., Gorostizaga, A., Brion, L., Gonzalez-Calvar, S.I., and Paz, C. (2015). cAMP-activated Nr4a1 expression requires ERK activity and is modulated by MAPK phosphatase-1 in MA-10 Leydig cells. Mol Cell Endocrinol *408*, 45-52.

Morse, S.J., Butler, A.A., Davis, R.L., Soller, I.J., and Lubin, F.D. (2015). Environmental enrichment reverses histone methylation changes in the aged hippocampus and restores age-related memory deficits. Biology (Basel) *4*, 298-313.

Mosaferi, B., Babri, S., Mohaddes, G., Khamnei, S., and Mesgari, M. (2015). Post-weaning environmental enrichment improves BDNF response of adult male rats. Int J Dev Neurosci *46*, 108-114.

Mullen, R.J., Buck, C.R., and Smith, A.M. (1992). NeuN, a neuronal specific nuclear protein in vertebrates. Development *116*, 201-211.

Muller, S. (2014). In silico analysis of regulatory networks underlines the role of miR-10b-5p and its target BDNF in huntington's disease. Transl Neurodegener *3*, 17.

Munzel, M., Globisch, D., Bruckl, T., Wagner, M., Welzmiller, V., Michalakis, S., Muller, M., Biel, M., and Carell, T. (2010). Quantification of the sixth DNA base hydroxymethylcytosine in the brain. Angew Chem Int Ed Engl *49*, 5375-5377.

Murillo, B., Ruiz-Reig, N., Herrera, M., Fairen, A., and Herrera, E. (2015). Zic2 Controls the Migration of Specific Neuronal Populations in the Developing Forebrain. J Neurosci *35*, 11266-11280.

Nachat-Kappes, R., Pinel, A., Combe, K., Lamas, B., Farges, M.C., Rossary, A., Goncalves-Mendes, N., Caldefie-Chezet, F., Vasson, M.P., and Basu, S. (2012). Effects of enriched environment on COX-2, leptin and eicosanoids in a mouse model of breast cancer. PLoS One *7*, e51525.

Naka, F., Narita, N., Okado, N., and Narita, M. (2005). Modification of AMPA receptor properties following environmental enrichment. Brain Dev *27*, 275-278.

Naka, F., Shiga, T., Yaguchi, M., and Okado, N. (2002). An enriched environment increases noradrenaline concentration in the mouse brain. Brain Res *924*, 124-126.

Nakatani, T., Kumai, M., Mizuhara, E., Minaki, Y., and Ono, Y. (2010). Lmx1a and Lmx1b cooperate with Foxa2 to coordinate the specification of dopaminergic neurons and control of floor plate cell differentiation in the developing mesencephalon. Dev Biol *339*, 101-113.

Nakayama, A.Y., Harms, M.B., and Luo, L. (2000). Small GTPases Rac and Rho in the maintenance of dendritic spines and branches in hippocampal pyramidal neurons. J Neurosci *20*, 5329-5338.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Nefzger, C.M., Su, C.T., Fabb, S.A., Hartley, B.J., Beh, S.J., Zeng, W.R., Haynes, J.M., and Pouton, C.W. (2012). Lmx1a allows context-specific isolation of progenitors of GABAergic or dopaminergic neurons during neural differentiation of embryonic stem cells. Stem Cells *30*, 1349-1361.

Nishimura, K., Murayama, S., and Takahashi, J. (2015). Identification of Neurexophilin 3 as a Novel Supportive Factor for Survival of Induced Pluripotent Stem Cell-Derived Dopaminergic Progenitors. Stem Cells Transl Med *4*, 932-944.

Nithianantharajah, J., and Hannan, A.J. (2006). Enriched environments, experience-dependent plasticity and disorders of the nervous system. Nat Rev Neurosci *7*, 697-709.
Novkovic, T., Mittmann, T., and Manahan-Vaughan, D. (2015). BDNF contributes to the facilitation of hippocampal synaptic plasticity and learning enabled by environmental enrichment. Hippocampus *25*, 1-15.

Oliva, C.A., Vargas, J.Y., and Inestrosa, N.C. (2013a). Wnt signaling: role in LTP, neural networks and memory. Ageing Res Rev *12*, 786-800.

Oliva, C.A., Vargas, J.Y., and Inestrosa, N.C. (2013b). Wnts in adult brain: from synaptic plasticity to cognitive deficiencies. Front Cell Neurosci *7*, 224.

Pajoro, A., Madrigal, P., Muino, J.M., Matus, J.T., Jin, J., Mecchia, M.A., Debernardi, J.M., Palatnik, J.F., Balazadeh, S., Arif, M.*, et al.* (2014). Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. Genome Biol *15*, R41.

Pantano, L., Estivill, X., and Marti, E. (2010). SeqBuster, a bioinformatic tool for the processing and analysis of small RNAs datasets, reveals ubiquitous miRNA modifications in human embryonic cells. Nucleic Acids Res *38*, e34.

Paoletti, P., Bellone, C., and Zhou, Q. (2013). NMDA receptor subunit diversity: impact on receptor properties, synaptic plasticity and disease. Nat Rev Neurosci *14*, 383-400.

Peleg, S., Sananbenesi, F., Zovoilis, A., Burkhardt, S., Bahari-Javan, S., Agis-Balboa, R.C., Cota, P., Wittnam, J.L., Gogol-Doering, A., Opitz, L.*, et al.* (2010). Altered histone acetylation is associated with age-dependent memory impairment in mice. Science *328*, 753-756.

Perez-Lluch, S., Blanco, E., Tilgner, H., Curado, J., Ruiz-Romero, M., Corominas, M., and Guigo, R. (2015). Absence of canonical marks of active chromatin in developmentally regulated genes. Nat Genet *47*, 1158-1167.

Pham, T.M., Soderstrom, S., Winblad, B., and Mohammed, A.H. (1999). Effects of environmental enrichment on cognitive function and hippocampal NGF in the non-handled rats. Behav Brain Res *103*, 63-70.

Pierani, A., Moran-Rivard, L., Sunshine, M.J., Littman, D.R., Goulding, M., and Jessell, T.M. (2001). Control of interneuron fate in the developing spinal cord by the progenitor homeodomain protein Dbx1. Neuron *29*, 367-384.

Plath, N., Ohana, O., Dammermann, B., Errington, M.L., Schmitz, D., Gross, C., Mao, X., Engelsberg, A., Mahlke, C., Welzl, H.*, et al.* (2006). Arc/Arg3.1 is essential for the consolidation of synaptic plasticity and memories. Neuron *52*, 437-444.

Prast, H., and Philippu, A. (2001). Nitric oxide as modulator of neuronal function. Prog Neurobiol *64*, 51-68.

Priel, A., Tuszynski, J.A., and Woolf, N.J. (2010). Neural cytoskeleton capabilities for learning and memory. J Biol Phys *36*, 3-21.

Purro, S.A., Ciani, L., Hoyos-Flight, M., Stamatakou, E., Siomou, E., and Salinas, P.C. (2008). Wnt regulates axon behavior through changes in microtubule growth directionality: a new role for adenomatous polyposis coli. J Neurosci *28*, 8644-8654.

Rabani, M., Raychowdhury, R., Jovanovic, M., Rooney, M., Stumpo, D.J., Pauli, A., Hacohen, N., Schier, A.F., Blackshear, P.J., Friedman, N.*, et al.* (2014). High-resolution sequencing and modeling identifies distinct dynamic RNA regulatory strategies. Cell *159*, 1698-1710.

Ragu Varman, D., and Rajan, K.E. (2015). Environmental Enrichment Reduces Anxiety by Differentially Activating Serotonergic and Neuropeptide Y (NPY)-Ergic System in Indian Field Mouse (Mus booduga): An Animal Model of Post-Traumatic Stress Disorder. PLoS One *10*, e0127945.

Ramirez, O.A., and Couve, A. (2011). The endoplasmic reticulum and protein trafficking in dendrites and axons. Trends Cell Biol *21*, 219-227.

Rampon, C., Jiang, C.H., Dong, H., Tang, Y.P., Lockhart, D.J., Schultz, P.G., Tsien, J.Z., and Hu, Y. (2000). Effects of environmental enrichment on gene expression in the brain. Proc Natl Acad Sci U S A *97*, 12880-12884.

Ramsahoye, B.H. (2002). Measurement of genome wide DNA methylation by reversed-phase high-performance liquid chromatography. Methods *27*, 156-161.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

Rasmuson, S., Olsson, T., Henriksson, B.G., Kelly, P.A., Holmes, M.C., Seckl, J.R., and Mohammed, A.H. (1998). Environmental enrichment selectively increases 5-HT1A receptor mRNA expression and binding in the rat hippocampus. Brain Res Mol Brain Res *53*, 285-290.

Reichmann, F., Wegerer, V., Jain, P., Mayerhofer, R., Hassan, A.M., Frohlich, E.E., Bock, E., Pritz, E., Herzog, H., Holzer, P.*, et al.* (2016). Environmental enrichment induces behavioural disturbances in neuropeptide Y knockout mice. Sci Rep *6*, 28182.

Reul, J.M., Hesketh, S.A., Collins, A., and Mecinas, M.G. (2009). Epigenetic mechanisms in the dentate gyrus act as a molecular switch in hippocampus-associated memory formation. Epigenetics *4*, 434-439.

Ribeiro, F.M., Ferreira, L.T., Marion, S., Fontes, S., Gomez, M., Ferguson, S.S., Prado, M.A., and Prado, V.F. (2007). SEC14-like protein 1 interacts with cholinergic transporters. Neurochem Int *50*, 356-364.

Risso, D., Ngai, J., Speed, T.P., and Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotechnol *32*, 896-902.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res *43*, e47.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Rodriguez-Paredes, M., and Esteller, M. (2011). Cancer epigenetics reaches mainstream oncology. Nat Med *17*, 330-339.

Rogers, J., Vo, U., Buret, L.S., Pang, T.Y., Meiklejohn, H., Zeleznikow-Johnston, A., Churilov, L., van den Buuse, M., Hannan, A.J., and Renoir, T. (2016). Dissociating the therapeutic effects of environmental enrichment and exercise in a mouse model of anxiety with cognitive impairment. Transl Psychiatry *6*, e794.

Romano, G.H., Harari, Y., Yehuda, T., Podhorzer, A., Rubinstein, L., Shamir, R., Gottlieb, A., Silberberg, Y., Pe'er, D., Ruppin, E.*, et al.* (2013). Environmental stresses disrupt telomere length homeostasis. PLoS Genet *9*, e1003721.

Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S.*, et al.* (2004). Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics *3*, 1154-1169.

Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R.*, et al.* (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature *481*, 389-393.

Rosso, S.B., and Inestrosa, N.C. (2013). WNT signaling in neuronal maturation and synaptogenesis. Front Cell Neurosci *7*, 103.

Rosso, S.B., Sussman, D., Wynshaw-Boris, A., and Salinas, P.C. (2005). Wnt signaling through Dishevelled, Rac and JNK regulates dendritic development. Nat Neurosci *8*, 34-42.

Russwurm, C., Koesling, D., and Russwurm, M. (2015). Phosphodiesterase 10A Is Tethered to a Synaptic Signaling Complex in Striatum. J Biol Chem *290*, 11936-11947.

Sagi, Y., Heiman, M., Peterson, J.D., Musatov, S., Scarduzio, M., Logan, S.M., Kaplitt, M.G., Surmeier, D.J., Heintz, N., and Greengard, P. (2014). Nitric oxide regulates synaptic transmission between spiny projection neurons. Proc Natl Acad Sci U S A *111*, 17636-17641.

Sah, R., and Geracioti, T.D. (2013). Neuropeptide Y and posttraumatic stress disorder. Mol Psychiatry *18*, 646-655.

Sakurada, M., Shichiri, M., Imamura, M., Azuma, H., and Hirata, Y. (2008). Nitric oxide upregulates dimethylarginine dimethylaminohydrolase-2 via cyclic GMP induction in endothelial cells. Hypertension *52*, 903-909.

Salero, E., Perez-Sen, R., Aruga, J., Gimenez, C., and Zafra, F. (2001). Transcription factors Zic1 and Zic2 bind and transactivate the apolipoprotein E gene promoter. J Biol Chem *276*, 1881-1888.

Sandberg, A., Branca, R.M., Lehtio, J., and Forshed, J. (2014). Quantitative accuracy in mass spectrometry based proteomics of complex samples: the impact of labeling and precursor interference. J Proteomics *96*, 133-144.

Sanes, J.R., and Lichtman, J.W. (1999). Development of the vertebrate neuromuscular junction. Annu Rev Neurosci *22*, 389-442.

Santiago, M., Antunes, C., Guedes, M., Sousa, N., and Marques, C.J. (2014). TET enzymes and DNA hydroxymethylation in neural development and function - how critical are they? Genomics *104*, 334-340.

Santos, A.R., Comprido, D., and Duarte, C.B. (2010). Regulation of local translation at the synapse by BDNF. Prog Neurobiol *92*, 505-516.

Sarnat, H.B., Nochlin, D., and Born, D.E. (1998). Neuronal nuclear antigen (NeuN): a marker of neuronal maturation in early human fetal nervous system. Brain Dev *20*, 88-94.

Saxena, A., and Carninci, P. (2011). Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs. Bioessays *33*, 830-839.

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell *132*, 887-898.

Schuman, E.M., Dynes, J.L., and Steward, O. (2006). Synaptic regulation of translation of dendritic mRNAs. J Neurosci *26*, 7143-7146.

Scott, A., Hasegawa, H., Sakurai, K., Yaron, A., Cobb, J., and Wang, F. (2011). Transcription factor short stature homeobox 2 is required for proper development of tropomyosin-related kinase B-expressing mechanosensory neurons. J Neurosci *31*, 6741-6749.

Segovia, G., Yague, A.G., Garcia-Verdugo, J.M., and Mora, F. (2006). Environmental enrichment promotes neurogenesis and changes the extracellular concentrations of glutamate and GABA in the hippocampus of aged rats. Brain Res Bull *70*, 8-14.

Shahbazian, M.D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. Annu Rev Biochem *76*, 75-100.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.

Shimozaki, K., Clemenson, G.D., Jr., and Gage, F.H. (2013). Paired related homeobox protein 1 is a regulator of stemness in adult neural stem/progenitor cells. J Neurosci *33*, 4066-4075.

Shonesy, B.C., Jalan-Sakrikar, N., Cavener, V.S., and Colbran, R.J. (2014). CaMKII: a molecular substrate for synaptic plasticity and memory. Prog Mol Biol Transl Sci *122*, 61-87.

Sim, S.E., Bakes, J., and Kaang, B.K. (2014). Neuronal activity-dependent regulation of MicroRNAs. Mol Cells *37*, 511-517.

Simpson, J., and Kelly, J.P. (2011). The impact of environmental enrichment in laboratory rats--behavioural and neurochemical aspects. Behav Brain Res *222*, 246-264.

Slabbaert, J.R., Kuenen, S., Swerts, J., Maes, I., Uytterhoeven, V., Kasprowicz, J., Fernandes, A.C., Blust, R., and Verstreken, P. (2016). Shawn, the Drosophila Homolog of SLC25A39/40, Is a Mitochondrial Carrier That Promotes Neuronal Survival. J Neurosci *36*, 1914-1929.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat Rev Genet *14*, 204-220.

Smyth, G.K., and Speed, T. (2003). Normalization of cDNA microarray data. Methods *31*, 265-273.

Son, W.Y., Lee, H.J., Yoon, H.K., Kang, S.G., Park, Y.M., Yang, H.J., Choi, J.E., An, H., Seo, H.K., and Kim, L. (2014). Gaba transporter SLC6A11 gene polymorphism associated with tardive dyskinesia. Nord J Psychiatry *68*, 123-128.

Song, Y., Selak, M.A., Watson, C.T., Coutts, C., Scherer, P.C., Panzer, J.A., Gibbs, S., Scott, M.O., Willer, G., Gregg, R.G.*, et al.* (2009). Mechanisms underlying metabolic and neural defects in zebrafish and human multiple acyl-CoA dehydrogenase deficiency (MADD). PLoS One *4*, e8329.

Sousa, K.M., Villaescusa, J.C., Cajanek, L., Ondr, J.K., Castelo-Branco, G., Hofstra, W., Bryja, V., Palmberg, C., Bergman, T., Wainwright, B.*, et al.* (2010). Wnt2 regulates progenitor proliferation in the developing ventral midbrain. J Biol Chem *285*, 7246-7253.

Steiner, B., Zurborg, S., Horster, H., Fabel, K., and Kempermann, G. (2008). Differential 24 h responsiveness of Prox1-expressing precursor cells in adult hippocampal neurogenesis to physical activity, environmental enrichment, and kainic acid-induced seizures. Neuroscience *154*, 521-529.

Stiles, J., and Jernigan, T.L. (2010). The basics of brain development. Neuropsychol Rev *20*, 327-348.

Sun, Q., Yang, Y., Li, X., He, B., Jia, Y., Zhang, N., and Zhao, R. (2016). Folate deprivation modulates the expression of autophagy- and circadian-related genes in HT-22 hippocampal neuron cells through GR-mediated pathway. Steroids *112*, 12-19.

Sweatt, J.D. (2001). The neuronal MAP kinase cascade: a biochemical signal integration system subserving synaptic plasticity and memory. J Neurochem *76*, 1-10.

Szeligo, F., and Leblond, C.P. (1977). Response of the three main types of glial cells of cortex and corpus callosum in rats handled during suckling or exposed to enriched, control and impoverished environments following weaning. J Comp Neurol *172*, 247-263.

Tai, H.C., Besche, H., Goldberg, A.L., and Schuman, E.M. (2010). Characterization of the Brain 26S Proteasome and its Interacting Proteins. Front Mol Neurosci *3*.

Takamiya, K., Kostourou, V., Adams, S., Jadeja, S., Chalepakis, G., Scambler, P.J., Huganir, R.L., and Adams, R.H. (2004). A direct functional link between the multi-PDZ domain protein GRIP1 and the Fraser syndrome protein Fras1. Nat Genet *36*, 172-177.

Takao, K., Tanda, K., Nakamura, K., Kasahara, J., Nakao, K., Katsuki, M., Nakanishi, K., Yamasaki, N., Toyama, K., Adachi, M.*, et al.* (2010). Comprehensive behavioral analysis of calcium/calmodulin-dependent protein kinase IV knockout mice. PLoS One *5*, e9460.

Tao, X., Finkbeiner, S., Arnold, D.B., Shaywitz, A.J., and Greenberg, M.E. (1998). Ca2+ influx regulates BDNF transcription by a CREB family transcription factor-dependent mechanism. Neuron *20*, 709-726.

Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., and Narechania, A. (2003). PANTHER: a library of protein families and subfamilies indexed by function. Genome Res *13*, 2129-2141.

Tirone, F., Farioli-Vecchioli, S., Micheli, L., Ceccarelli, M., and Leonardi, L. (2013). Genetic control of adult neurogenesis: interplay of differentiation, proliferation and survival modulates new neurons function, and memory circuits. Front Cell Neurosci *7*, 59.

Tolias, K.F., Duman, J.G., and Um, K. (2011). Control of synapse development and plasticity by Rho GTPase regulatory proteins. Prog Neurobiol *94*, 133-148.
Tsai, N.P. (2014). Ubiquitin proteasome system-mediated degradation of synaptic proteins: An update from the postsynaptic side. Biochim Biophys Acta *1843*, 2838-2842.

Tsompana, M., and Buck, M.J. (2014). Chromatin accessibility: a window into the genome. Epigenetics Chromatin *7*, 33.

Turner, B.M. (2007). Defining an epigenetic code. Nat Cell Biol *9*, 2-6.
Turrigiano, G.G. (2008). The self-tuning neuron: synaptic scaling of excitatory synapses. Cell *135*, 422-435.

Unemura, K., Kume, T., Kondo, M., Maeda, Y., Izumi, Y., and Akaike, A. (2012). Glucocorticoids Decrease Astrocyte Numbers by Reducing Glucocorticoid Receptor Expression In Vitro and In Vivo. Journal of Pharmacological Sciences *119*, 30-39.

Uzhachenko, R., Ivanov, S.V., Yarbrough, W.G., Shanker, A., Medzhitov, R., and Ivanova, A.V. (2014). Fus1/Tusc2 is a novel regulator of mitochondrial calcium handling, Ca2+-coupled mitochondrial processes, and Ca2+-dependent NFAT and NF-kappaB pathways in CD4+ T cells. Antioxid Redox Signal *20*, 1533-1547.

van Praag, H., Kempermann, G., and Gage, F.H. (1999). Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. Nat Neurosci *2*, 266-270.

van Praag, H., Kempermann, G., and Gage, F.H. (2000). Neural consequences of environmental enrichment. Nat Rev Neurosci *1*, 191-198.

Varjosalo, M., and Taipale, J. (2008). Hedgehog: functions and mechanisms. Genes Dev *22*, 2454-2472.

Vasiljevic, M., Heisler, F.F., Hausrat, T.J., Fehr, S., Milenkovic, I., Kneussel, M., and Sieghart, W. (2012). Spatio-temporal expression analysis of the calcium-binding protein calumenin in the rodent brain. Neuroscience *202*, 29-41.

Vecsey, C.G., Hawk, J.D., Lattal, K.M., Stein, J.M., Fabian, S.A., Attner, M.A., Cabrera, S.M., McDonough, C.B., Brindle, P.K., Abel, T.*, et al.* (2007). Histone deacetylase inhibitors enhance memory and synaptic plasticity via CREB:CBP-dependent transcriptional activation. J Neurosci *27*, 6128-6140.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A.*, et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351.

Vincent, S.R. (2010). Nitric oxide neurons and neurotransmission. Prog Neurobiol *90*, 246-255.

Voglis, G., and Tavernarakis, N. (2006). The role of synaptic ion channels in synaptic plasticity. EMBO Rep *7*, 1104-1110.

Vos, M., Lauwers, E., and Verstreken, P. (2010). Synaptic mitochondria in synaptic transmission and organization of vesicle pools in health and disease. Front Synaptic Neurosci *2*, 139.

Wagner, A.K., Chen, X., Kline, A.E., Li, Y., Zafonte, R.D., and Dixon, C.E. (2005). Gender and environmental enrichment impact dopamine transporter expression after experimental traumatic brain injury. Exp Neurol *195*, 475-483.

Wallen-Mackenzie, A., Wootz, H., and Englund, H. (2010). Genetic inactivation of the vesicular glutamate transporter 2 (VGLUT2) in the mouse: what have we learnt about functional glutamatergic neurotransmission? Ups J Med Sci *115*, 11-20.

Wang, X., Cao, L., Wang, Y., Wang, X., Liu, N., and You, Y. (2012). Regulation of let-7 and its target oncogenes (Review). Oncol Lett *3*, 955-960.

Wang, Y., Li, M., Stadler, S., Correll, S., Li, P., Wang, D., Hayama, R., Leonelli, L., Han, H., Grigoryev, S.A.*, et al.* (2009). Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. J Cell Biol *184*, 205-213.
Wen, L., and Tang, F. (2014). Genomic distribution and possible functions of DNA hydroxymethylation in the brain. Genomics *104*, 341-346.

West, A.E., and Greenberg, M.E. (2011). Neuronal activity-regulated gene transcription in synapse development and cognitive function. Cold Spring Harb Perspect Biol *3*.
Westwood, J.A., Darcy, P.K., and Kershaw, M.H. (2013). Environmental enrichment does not impact on tumor growth in mice. F1000Res *2*, 140.

Widman, D.R., Abrahamsen, G.C., and Rosellini, R.A. (1992). Environmental enrichment: the influences of restricted daily exposure and subsequent exposure to uncontrollable stress. Physiol Behav *51*, 309-318.

Wiesel, T.N. (1982). Postnatal development of the visual cortex and the influence of environment. Nature *299*, 583-591.

Withers, G.S., Higgins, D., Charette, M., and Banker, G. (2000). Bone morphogenetic protein-7 enhances dendritic growth and receptivity to innervation in cultured hippocampal neurons. Eur J Neurosci *12*, 106-116.

Wolf, S.A., Kronenberg, G., Lehmann, K., Blankenship, A., Overall, R., Staufenbiel, M., and Kempermann, G. (2006). Cognitive and physical activity differently modulate disease progression in the amyloid precursor protein (APP)-23 model of Alzheimer's disease. Biol Psychiatry *60*, 1314-1323.

Woo, C.C., Donnelly, J.H., Steinberg-Epstein, R., and Leon, M. (2015). Environmental enrichment as a therapy for autism: A clinical trial replication and extension. Behav Neurosci *129*, 412-422.

Wu, C., Qiu, R., Wang, J., Zhang, H., Murai, K., and Lu, Q. (2009). ZHX2 Interacts with Ephrin-B and regulates neural progenitor maintenance in the developing cerebral cortex. J Neurosci *29*, 7404-7412.

Xing, H., Mo, Y., Liao, W., and Zhang, M.Q. (2012). Genome-wide localization of protein-DNA binding and histone modification by a Bayesian change-point method with ChIP-seq data. PLoS Comput Biol *8*, e1002613.

Xu, H., Gelyana, E., Rajsombath, M., Yang, T., Li, S., and Selkoe, D. (2016). Environmental Enrichment Potently Prevents Microglia-Mediated Neuroinflammation by Human Amyloid beta-Protein Oligomers. J Neurosci *36*, 9041-9056.

Xu, J., Kam, C., Luo, J.H., and Xia, J. (2014a). PICK1 mediates synaptic recruitment of AMPA receptors at neurexin-induced postsynaptic sites. J Neurosci *34*, 15415-15424.
Xu, S., Grullon, S., Ge, K., and Peng, W. (2014b). Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. Methods Mol Biol *1150*, 97-111.

Ye, X., Luo, H., Chen, Y., Wu, Q., Xiong, Y., Zhu, J., Diao, Y., Wu, Z., Miao, J., and Wan, J. (2015). MicroRNAs 99b-5p/100-5p Regulated by Endoplasmic Reticulum Stress are Involved in Abeta-Induced Pathologies. Front Aging Neurosci *7*, 210.

Young, D., Lawlor, P.A., Leone, P., Dragunow, M., and During, M.J. (1999). Environmental enrichment inhibits spontaneous apoptosis, prevents seizures and is neuroprotective. Nat Med *5*, 448-453.

Yu, H.C., Wu, J., Zhang, H.X., Zhang, G.L., Sui, J., Tong, W.W., Zhang, X.Y., Nie, L.L., Duan, J.H., Zhang, L.R.*, et al.* (2015). Alterations of miR-132 are novel diagnostic biomarkers in peripheral blood of schizophrenia patients. Prog Neuropsychopharmacol Biol Psychiatry *63*, 23-29.

Zanca, R.M., Braren, S.H., Maloney, B., Schrott, L.M., Luine, V.N., and Serrano, P.A. (2015). Environmental Enrichment Increases Glucocorticoid Receptors and Decreases GluA2 and Protein Kinase M Zeta (PKMzeta) Trafficking During Chronic Stress: A Protective Mechanism? Front Behav Neurosci *9*, 303.

Zhang, D.D., Lo, S.C., Cross, J.V., Templeton, D.J., and Hannink, M. (2004). Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex. Mol Cell Biol *24*, 10941-10953.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zheng, J.J., Li, S.J., Zhang, X.D., Miao, W.Y., Zhang, D., Yao, H., and Yu, X. (2014). Oxytocin mediates early experience-dependent cross-modal plasticity in the sensory cortices. Nat Neurosci *17*, 391-399.

Zhou, Q.Y., and Palmiter, R.D. (1995). Dopamine-deficient mice are severely hypoactive, adipsic, and aphagic. Cell *83*, 1197-1209.

Zieske, L.R. (2006). A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. J Exp Bot *57*, 1501-1508.

Ziller, M.J., Gu, H., Muller, F., Donaghey, J., Tsai, L.T., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E.*, et al.* (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*, 477-481.

Zingg, J.M. (2015). Vitamin E: A Role in Signal Transduction. Annu Rev Nutr *35*, 135-173.