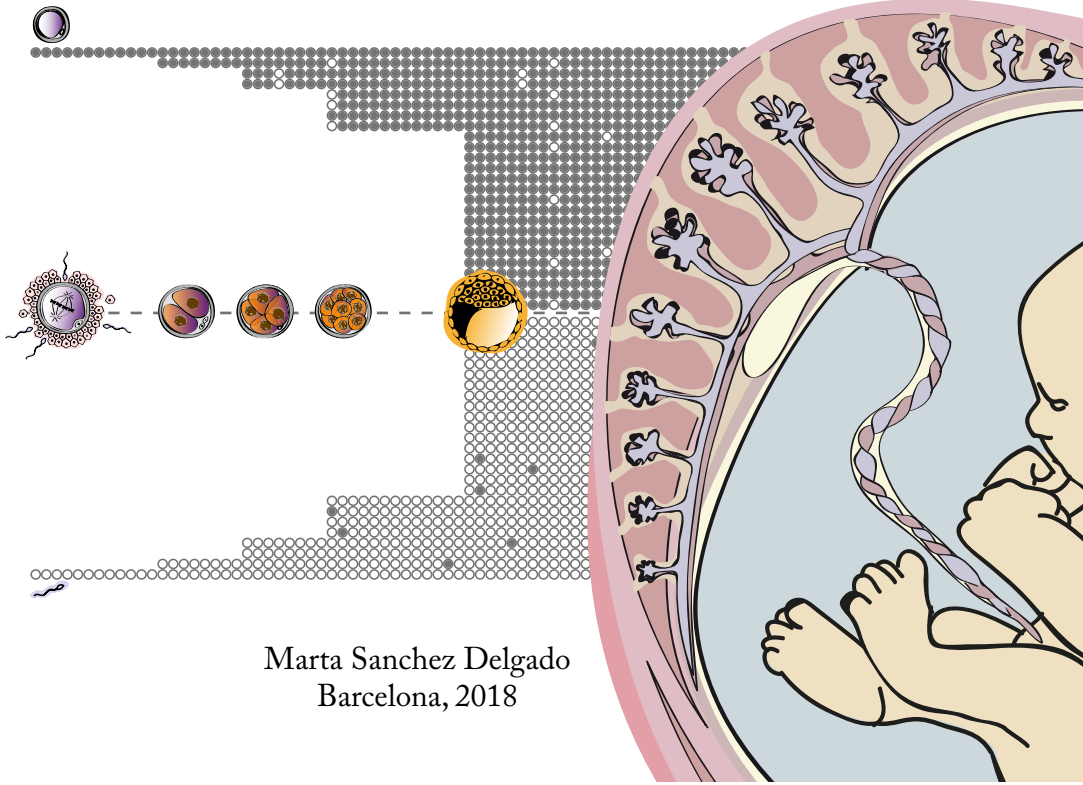
UNIVERSITAT DE BARCELONA

# Acquisition and maintenance of transient imprinted DMRs in the human placenta and their role in development

Doctoral Thesis



Marta Sanchez Delgado
Barcelona, 2018

*Cover designed by Beatriz Cupillar Gimenez & Marta Sanchez Delgado*


# Acquisition and maintenance of transient imprinted DMRs in the human placenta and their role in development

Marta Sanchez Delgado

Doctoral Thesis

April 2018



BY NC ND

Universitat de Barcelona – Facultat de Medicina

Programa de Doctorat en Biomedicina

# Acquisition and maintenance of transient imprinted DMRs in the human placenta and their role in development

Memòria presentada per Marta Sanchez Delgado per optar al títol de Doctor per la Universitat de Barcelona

Aquest treball s'ha realitzat al grup d'Impressió Genètica i Càncer, dins del Programa d'Epigenètica i Biologia del Càncer (PEBC) de l'Institut d'Investigació Biomèdica de Bellvitge (IDIBELL)

Barcelona, Març 2018

**Dr. David Monk**
Director

**Dr. Manel Esteller**
Tutor

**Marta Sanchez Delgado**
Doctorant

# Abstract

**Genomic imprinting** is an epigenetic phenomenon resulting in the monoallelic expression of a subset of genes in a parent-of-origin-specific manner. In general, the promoters of these transcripts contain differentially methylated regions (DMRs), usually methylated on the non-expressed allele. In humans, genome-wide screening experiments indicate that most ubiquitous imprinted genes, associated with DMRs in all tissues, have already been identified. On the contrary, the existence of tissue-specific imprinted DMRs remains mostly uninvestigated.

This thesis aims to determine the extent of imprinting in the human placenta and how these genes can influence intrauterine growth and development. During this dissertation, a total of 72 human placenta-specific DMRs were confirmed. All of these regions inherit methylation from the oocyte and are stable through embryonic reprogramming, being lost after implantation in somatic tissues. Furthermore, we described imprinted monoallelic expression for 20 genes associated with these novel placenta-specific DMRs. Imprinted expression also occurs in the pre-implantation embryo as highlighted by the paternal expression of *ZHX3* in cleavage stage embryos.

The aberrant expression from the maternal allele of some placenta-specific genes is likely to play an essential role in the hydatidiform mole phenotype since they include crucial genes involved in different biological processes including epigenetic modifications and metabolic processes. In our placenta cohort, we observe polymorphic placenta-specific imprinting, with biallelic expression correlating with biallelic permissive histone marks which can be independent of the allelic methylation state in some cases. Although we have not observed a higher frequency of polymorphic imprinted methylation in placentas from complicated pregnancies or those conceived using assisted reproductive technologies, further characterisation, including extensive quantitative expression studies, are needed to ascertain the role of placenta-specific imprinted genes in development.

# Acknowledge

First of all, I would like to thank my thesis director, Dr Monk, to give me the opportunity to be part of his group and for all his mentoring, support, trust and help. Thanks, Dave, I really enjoy and learn a lot during all these years! También agradecer a la Dra Iglesias-Platas su colaboración e inicial obtención de todas las muestras de placenta con las que ha sido posible llevar a cabo este trabajo. Muchas gracias Isa.

Y llegó el fin de otra etapa. Parece que fue ayer cuando me reuní por primera vez con mi tutora del *Treball Fí de Grau*, la Dra Camprubí, para introducirme en el mundo del *Imprinting*. Moltes gràcies Cristina per acompanyar-me en aquelles primeres passes! Sense la teva tutoria inicial, possiblement no m'hauria enamorat tant d'aquest món! I és que tot va començar el curs 2012/2013, acabant la carrera i començant les pràctiques d'estiu IDIBELL al *Imprinting Grup*. Amy, moltes gràcies per la teva acollida inicial, vam coincidir molt poquet però em vas ajudar moltíssim amb tot, tant amb l'anglès com a donar les primeres passes al lab... i també gràcies pels cafès de bon matí, algun dia tornarem a coincidir a la mateixa ciutat i els repetim :).

Pero el PEBC5 en el verano de 2013 era muy diferente al actual. Pude coincidir con gente fantástica, aunque durante poquitos meses. Thanks, Franck, for making the start so fun, it was a pleasure to meet you. Y exactamente por las mismas razones, también fue un placer coincidir contigo José, el otro lado del PEBC5, que una vez Alex volvió de escribir su tesis, pasó a ser todo como un mismo grupo. Gracias Alex por compartir esos primeros años en el lab. Y ahí entráis vosotras: Rita y Soraya. Rita, la primera persona en auto-presentarse jajajajaja gracias por transmitir esa paz y por el apoyo moral que siempre das con tus buenas palabras. Y Sojaji, ¡por dios! ¡creo que eres con quien más tiempo he coincidido en el lab! He aprendido tanto todos estos años contigo... Aun me queda un viaje pendiente a Cenizate... ¡Te prometo que lo haré! I per aquesta primera part al *Imprinting and cell cycle lab* no m'oblido de vosaltres dues, Judith i Miriam! Judith, la felicitat i dolçor personalitzada, et trobo a faltar molt! I Miriam!, tot i que únicament vas estar uns mesos al lab, en realitat hi continues formant part, t'agraeixo molt que m'hagis encomanat una miqueta d'aquesta tranquil · litat que reboses :)

Mientras tanto, entre ese PEBC5 inicial al actual, han ido haciendo pequeñas estancias personas fantásticas de las que siempre se sacan experiencias maravillosas: Aitor, Iris, Lara, Anne, Alex, Sonia, Núria... ¡Muchas gracias a todos y cada uno de vosotros!

Y llegó septiembre de 2015, el ecuador de mi etapa en las poyatas del *Imprinting Group* y con él, un enorme cambio: la mejor compañera de viaje que se pueda tener. Ana, madre mía, pero que te puedo decir... Mil gracias por todo, por aparecer en mi vida (de la que espero, ¡no salgas nunca!), por el día a día, por tu sonrisa, por tu ayuda, por tu paciencia, por escuchar, por confiar... Por ser como eres!! Contigo todo cambió, y cuando digo todo, ¡es todo! Llegaste con una nueva etapa, que junto con José y Paolo, formamos el mejor y más unido *Imprinting Group* que ha existido. ¡Paolo! Ya te conocía de antes por ser *amiguito* de Alex pero no era lo mismo... Muchas gracias por tu ayuda, tus bromas y tus conversaciones culinarias... Y hablando de conversaciones, José, está claro que sin ti, los descansos no habrían sido ni por asomo, tan y tan entretenidos ni con tanto debate polémico como han sido estos últimos años. Muchas gracias por eso y por tus ánimos y apoyo.

Pero de nuevo, la otra parte de este renovado PEBC5 también ha cambiado bastante y Yolanda, ¡tú has vivido los cambios de pleno! Ha sido un placer compartir estos 3 años contigo, a ti te queda todavía un poquito pero espero que sigas disfrutando y aprendiendo muchísimo. Y Patri, la última incorporación. Hemos coincidido muy poquito pero la verdad es que me ha encantado poder compartir este tiempo contigo. Y es que este último año, ¡ha sido realmente fantástico! Hemos compartido muchísimos momentos dentro y fuera del lab y está claro que esto es un hasta luego, porque este grupo seguirá haciendo sus cenas y cervecitas post-padel (aunque algunos nunca juguemos). ¡Cécile!, ¡que no me olvido de ti!, ¡mi francesita guapa!, muchas gracias por este año compartiendo poyata, por todos esos momentos irrepetibles aprendiendo castellano (increíblemente rápido y bien, por cierto) y ya sabes... *pi ja ma! & beaucoup de bisous!* Y bueno, este último año tampoco hubiera sido igual sin "las niñas", Lara y Paula, aunque vuestro paso por el PEBC5 ha sido breve, está claro que seguiréis formando parte de esta familia.

Pero fuera del arco metálico del PEBC5, también hay gente fantástica. ¡María y Jess! llegasteis con Ana, nuestras meriendas, cine, loterías, cenas... ¡Os estoy echando mucho de menos! ¡Espero no perder el contacto nunca! I en general, moltíssimes gràcies a tota la gent que forma part del PEBC: Alba, Olga, Marta Soler, Anna Martínez, Anna Marazuela, George, Nico, Fer, Laia Piqué, Laia Coll, Montse, Olga, Anaís, Humberto, Cátia, Carles, Tian, Clara, Xec, Sandra, Marina, Laura... i molts més! Moltes gràcies per tot! També en especial als organitzadors de tots els sopars/paintballs/antiretreats, i per tots els jocs ideats pel Pere, no hauria estat el mateix sense totes aquestes vivències. Tampoco habría sido lo mismo sin esa otra parte no científica: Rosa, Olga, Sonia, Mari, Isa, Porfi, Álvaro... Gracias por hacer más amenos todos esos tiempos muertos. Y ya terminando con el PEBC,

tampoco habría sido lo mismo sin ese equipo fantástico de fútbol: Guille, Manu, Jorge, Adrià, Adrian y Héctor. Gracias por todos los momentos compartidos fuera de estas cuatro paredes.

Pero está clarísimo que esta tesis no hubiera sido posible sin todo el apoyo que he recibido durante todos estos años de mi familia. En primer lugar, por haber convivido conmigo estos últimos 4 años, Carlos, cariño, gracias por aguantarme con paciencia todos estos años y por creer en mí cuando me entran las mil y una dudas... ¡Adri mi amor! Gracias por entender mis ausencias y estar siempre dispuesto a adaptarte a mi tiempo. Nos esperan muchos mini-viajes juntos como el de Oporto :)

Y por último, mil gracias a las dos personas que han hecho posible que yo haya podido estudiar y dedicarme a mi pasión, porque aunque no entendíais porque *la niña* quería estudiar Biología/Genética – *pero con eso tú luego entras en una buena empresa, ¿no?* – aun así, me habéis apoyado, animado, os habéis sacrificado en dinero y tiempo... – *baja la tele Manolo, que la niña está estudiando* –. Mama, papa, ¡GRACIAS! Os quiero con locura, esta tesis va enterita dedicada a vosotros.

# Contents

# List of Figures

# List of Tables

# Abreviations

Table 1: **General abreviations**

| | |
|---|---|
| **5-caC** | 5-carboxylcytosine |
| **5-fC** | 5-formylcytosine |
| **5-hmC** | 5-hydroxymethylcytosine |
| **5-mC** | 5-methylcytosine |
| **ART** | Assisted reproductive technology |
| **AS** | Angelman syndrome |
| **B6** | C57BL/6 mouse strain |
| **BS (DNA)** | Bisulphite treated DNA |
| **BWS** | Beckwith-Wiedemann syndrome |
| **C57BL/6J** | C57BL/6 mouse strain from Jackson Laboratory |
| **CAST** | *Mus musculus castaneous* |
| **CAST/EiJ** | CAST mouse strain from Jackson Laboratory |
| **cDNA** | Complementary deoxyribonucleic acid |
| **ChIP** | Chromatin immunoprecipitation |
| **CHM** | Complete hydatidiform mole |
| **di-dNTP** | modified deoxynucleosidetriphosphates (with dye terminators) |
| **DMR** | Differentially methylated region |
| **DNA** | Deoxyribonucleic acid |
| **DNMT** | DNA-methyltransferases |
| **dNTP** | Deoxynucleosidetriphosphates |
| **dpp** | *Days postpartum* |
| **E** | Embryonic day |
| **FIVI** | *Fundación del Instituto Valenciano de Infertilidad* / Valencian Infertility Institute Foundation |
| **FPLC** | Fast protein liquid chromatography |
| **gDMR** | Germline differentially methylated region |
| **GEO** | Gene Expression Omnibus |
| **GO** | Gene Ontology |

| | |
|---|---|
| **GOM** | Gain-of-methylation |
| **HM** | Hydatidiform mole |
| **ICM** | Inner cell mass |
| **ICR** | Imprinting control region |
| **ICSI** | Intracytoplasmic sperm injections |
| **IDIBELL** | *Institut d'Investigació Biomèdica de Bellvitge* / Bellvitge Biomedical Research Institute |
| **IUGR** | Intrauterine growth restriction |
| **IVF** | *In vitro* fertilisation |
| **JF1** | *Mus musculus molossus* |
| **LOI** | Loss-of-imprinting |
| **LOM** | Lack-of-methylation |
| **MC** | Mononuclear cell |
| **MLID** | Multi-locus imprinting disturbances |
| **MEG** | Maternally expressed gene |
| **methyl-seq** | Methylation sequencing |
| **MSRE** | Methylation-sensitive restriction enzyme |
| **NBDC** | National Bioscience Database Center |
| **ncRNA** | non-coding ribonucleic acid |
| **oxBS (DNA)** | Oxidated DNA followed by bisulphyte conversion |
| **PCR** | Polymerase chain reaction |
| **PEBC** | *Programa d'Epigenètica i Biologia del Càncer* / Cancer Epigenetics and Biology Program |
| **PEG** | Paternally expressed gene |
| **PMD** | Partially methylated domains |
| **RHM** | Recurrent hydatidiform moles |
| **RNA** | Ribonucleic acid |
| **RRBS** | Reduced-representation bisulfite sequencing |
| **RT-PCR** | Reverse transcription and polymerase chain reaction |
| **SCMC** | Subcortical maternal complex |
| **SRS** | Silver Russell syndrome |
| **TE**[1] | Trophectoderm |
| **TET** | Ten-eleven translocation |
| **TNDM** | Transient neonatal diabetes mellitus |
| **UPD** | Uniparental disomy |
| **WGBS** | Whole-genome bisulphite sequencing |
| **ZGA** | Zygote genome activation |

[1] **Exception**: in *Material and Methods* section, *TE* is also used for Tris-EDTA buffer solution.

Table 2: **Official amino acid nomenclature (IUPAC codes)**

| IUPAC amino acid code | Three letter code | Amino acid |
|---|---|---|
| A | Ala | Alanine |
| C | Cys | Cysteine |
| D | Asp | Aspartic Acid |
| E | Glu | Glutamic Acid |
| F | Phe | Phenylalanine |
| G | Gly | Glycine |
| H | His | Histidine |
| I | Ile | Isoleucine |
| K | Lys | Lysine |
| L | Leu | Leucine |
| M | Met | Methionine |
| N | Asn | Asparagine |
| P | Pro | Proline |
| Q | Gln | Glutamine |
| R | Arg | Arginine |
| S | Ser | Serine |
| T | Thr | Threonine |
| V | Val | Valine |
| W | Trp | Tryptophan |
| Y | Tyr | Tyrosine |

Table 3: **Official nucleotides nomenclature (IUPAC codes)**

| IUPAC nucleotide code | Base |
|---|---|
| A | Adenine |
| C | Cytosine |
| G | Guanine |
| T (or U) | Thymine (or Uracil) |
| M | A or C |
| R | A or G |
| W | A or T |
| S | C or G |
| Y | C or T |
| K | G or T |
| V | A or C or G |
| H | A or C or T |

| IUPAC nucleotide code | Base |
|---|---|
| D | A or G or T |
| B | C or G or T |
| N | any base |
| . or - | gap |

Table 4: **Official nomenclature for epigenetic-related factors (Gene Cards)**

| Name | Aliases |
|---|---|
| AFF3 | AF4/FMR2 Family Member 3; LAF4; MLLT2-Like |
| CBX5 | Chromobox 5; HP1; HP1A; HEL25 |
| DNMT1 | DNA Methyltransferase 1; CXXC9; HSN1E; M.HsaI; ADCADN; AIM; MCMT1; DNA MTase HsaI; EC 2.1.1.37 |
| DNMT3A | DNA Methyltransferase 3 Alpha; TBRS; M.HsaIIIA; EC 2.1.1.37 |
| DNMT3L | DNA Methyltransferase 3 Like |
| DPPA3 | Developmental Pluripotency Associated 3; STELLAR; STELLA; Pgc7 |
| KDM1B | Lysine Demethylase 1B; AOF1; C6orf193; LSD2; BA204B7.3; DJ298J15.2 |
| KHDC3L | KH Domain Containing 3 Like, Subcortical Maternal Complex Membe; C6orf221; ECAT1; HYDM2; FILIA |
| NLRP2 | NLR Family Pyrin Domain Containing 2; NALP2; CLR19.9; PYPAF2; PAN1; NBS1 |
| NLRP5 | NLR Family Pyrin Domain Containing 5; NALP5; CLR19.8; PYPAF8; PAN11; MATER |
| NLRP7 | NLR Family Pyrin Domain Containing 7; NALP7; CLR19.4; PYPAF3; PAN7; NOD12; HYDM |
| OOEP | Oocyte Expressed Protein; KHDC2; C6orf156; OEP19; HOEP19; FLOPED |
| PADI6 | Peptidyl Arginine Deiminase 6; PAD6; PREMBL2; HPADVI; EC 3.5.3.15 |
| SETDB1 | SET Domain Bifurcated 1; ESET; KMT1E; KG1T; TDRD21; KIAA0067; EC 2.1.1.43; H3-K9-HMTase 4; Tudor Domain Containing 21 |
| TET1 | Tet Methylcytosine Dioxygenase 1; KIAA1676; CXXC6; LCX; BA119F7.1 |
| TET2 | Tet Methylcytosine Dioxygenase 2; KIAA1546; MDS |
| TET3 | Tet Methylcytosine Dioxygenase 3; KIAA0401 |
| TLE6 | Transducin Like Enhancer Of Split 6; PREMBL; GRG6 |

| Name | Aliases |
|---|---|
| TRIM28 | Tripartite Motif Containing 28; KAP1; KRIP-1; RNF96; TIF1B; PPP1R157; KRAB-Associated/Interacting Protein 1 |
| UHRF1 | Ubiquitin Like With PHD And Ring Finger Domains 1; HuNp95; ICBP90; RNF106; HUHRF1; HNP95; Np95; TDRD22 |
| ZFP57 | ZFP57 Zinc Finger Protein; C6orf40; ZNF698; TNDM1; BA145L22 |

# Summary

**Genomic imprinting** is an epigenetic phenomenon resulting in the monoallelic expression of a subset of genes in a parent-of-origin-specific manner. In general, mammalian imprinted genes are organised in clusters and regulated by imprinting control regions (ICRs). In most cases, ICRs manifest as differentially methylated regions (DMRs) where cytosine methylation marks one of the parental alleles, providing cis-acting regulatory elements that influence the allelic expression of neighbouring genes. The acquisition and maintenance of imprinted DMRs rely on the interplay between transcription and histone modifications, together with the action of specific epigenetic regulatory factors (e.g. DNMTs, TRIM28, DPPA3, ZFP57 and UHRF1). Additional factors may be involved as highlighted by their phenotypes when mutated, e.g. *KHDC3L* and *NLRP7* are associated with recurrent hydatidiform moles (RHMs). However, their role in the imprinting process remains to be defined. Until recently, all described imprinted DMRs were classified as germline or secondary DMRs. Germline DMRs, also known as primary DMRs, inherit methylation from the oocyte or sperm and, in general, are ubiquitously found in all tissues. On the other hand, secondary DMRs acquire methylation during development, in a tissue-specific and hierarchical fashion dictated by a neighbouring ICR. In humans, genome-wide screens for imprinted DMRs indicates that most ubiquitous imprinted genes have already been identified. However, due to the lack of targeted investigations, it is thought that most of the tissue-specific imprinted genes remain to be described.

I hypothesise that *compared to somatic tissues, the human placenta is enriched for genes subject to genomic imprinting that may be essential for intrauterine growth and development.* To determine the extent of imprinting in the human placenta and how these genes can influence human growth and development, the work carried out in this thesis mainly aimed to: (1) Study products of conception form *NLRP7*-mutated women; (2) Confirm additional placenta-specific DMRs identified by new bioinformatic pipelines; (3) Perform expression and epigenetic characterization of imprinted loci in the placenta (histone modifications and hydroximethylation); Finally, (4) assess the stability of allelic methylation at placenta-specific DMRs and the associated gene expression in placenta samples obtained from complicated pregnancies and those conceived using assisted reproductive technologies

(ART).

We initially thought that most maternally methylated placental-specific DMRs acquired their methylation pattern during early placental development. However, within the first year of this thesis, the publication of the human mature oocyte methylome revealed that all placenta-specific DMRs inherit their methylation from the female germline. During this dissertation, a total of **72** novel monoallelically methylated regions were confirmed in the human placenta, being 50 of them methylated on the maternal allele. Allele-specific expression studies confirmed the monoallelic expression of **20** genes (17 of them paternally expressed) whose promoters contained a placenta-specific DMR. As a result of comparing methylation profiles between RHMs from *NLRP7* -mutated women and control placentas, we obtained an initial list of loci subject to placenta-specific imprinting. Imprinting studies in RHMs revealed global lack-of-maternal methylation at imprinted DMRs, accompanied by biallelic expression. In-depth methylation and expression studies showed that placenta-specific DMRs do not orchestrate imprinted expression of neighbouring genes (and therefore do not imprint clusters but single genes) and are presented in the pre-implantation embryo following zygotic genome activation (exemplified by the *ZHX3* gene).

Through profiling a discovery set of placenta samples, we observe that placenta-specific imprinting is polymorphically regulated and that methylation associated with placenta-specific DMRs is not hydroxymethylation. Therefore, the methylation profile described in these regions is not an intermediate of demethylation. Further epigenetic characterisation revealed that polymorphic biallelic expression correlates with the presence of permissive histone marks, regardless of methylation state in some cases (i.e. a promoter may be a DMR, but H3K4me2/3 is present on both alleles). After assessing the stability of allelic methylation at both ubiquitous and placenta-specific DMRs in an extended placenta cohort, we did not observe a higher frequency of polymorphically low methylation in placentas from complicated pregnancies or those conceived using ARTs.

It seems that placenta-specific DMRs have appeared later in evolution, exclusively in the primate lineage. Due to the considerable number of regions described, their survival in the placenta is likely associated with the adaption of the oocyte, embryo and placenta epigenetic machinery rather the acquisition of a new genomic context. These observations are consistent with *NLRP7*, a primate-specific NLRP-family member, being a maternal-effect gene potentially involved in both oocyte establishment and zygotic maintenance of human imprinted DMRs (including placenta-specific regions). It has previously proposed that the phenotypes of hydatidiform moles, either the sporadic androgenetic and RHMs, are associated with aberrant genomic imprinting. However, no systematic analyses for imprinting defects had been reported before the studies in this PhD project. The aberrant expression from the maternal allele of some placenta-specific imprinted genes is likely to play an essential role in the hydatidiform mole phenotype since they include crucial genes

involved in different biological processes including epigenetic modifications and metabolism. Because of this, we favour the hypothesis that maternal silencing of these genes in early embryonic development and the placenta is a mechanism to prevent ovarian teratoma that arises from parthenogenetically activated oocytes. However, since we observe polymorphic biallelic expression in a minority of control placentas not associated with complicated pregnancies, this suggests that the severe phenotype associated with hydatidiform moles only occurs when genome-wide imprinting is disturbed and that milder developmental anomalies such as intrauterine growth restriction may occur when one or few genes are affected. Although we have not observed more regions with lack-of-methylation in our placentas from complicated pregnancies or conceived by ART when compared to control samples, further characterisation, including extensive quantitative expression studies, are needed to ascertain the role of placenta-specific imprinted genes in intrauterine growth and development.

Finally, taking into consideration the results of this PhD dissertation, we can conclude that there are more imprinted domains in the human placenta than in somatic tissues and that methylation at ubiquitous DMRs is more stable than placenta-specific DMRs in extra-embryonic tissues.

# Resumen

La **impronta genómica** es un fenómeno epigenético por el que un subconjunto de genes se expresa de manera monoalélica dependiendo del origen parental de cada alelo. En mamíferos, los genes regulados mediante este mecanismo se encuentran generalmente en grupos y regulados por regiones de control de la impronta (ICR, por sus siglas en inglés). Estas ICRs acostumbran a ser regiones diferencialmente metiladas (DMRs, por sus siglas en inglés) en las que la metilación del ADN marca uno de los alelos parentales, influyendo en la expresión de los genes circundantes. La adquisición y el mantenimiento de estas DMRs depende de la interacción entre procesos de transcripción, modificaciones de histonas y la acción de factores que influyen en la regulación epigenética (como por ejemplo, DNMTs, TRIM28, DPPA3, ZFP57 y UHRF1). Además, otro factores también podrían estar implicados en esta regulación epigenética, como por ejemplo *KHDC3L* y *NLRP7*, cuyas mutaciones en homocigosis están asociadas con molas hidatiformes recurrentes (MHR). No obstante, todavía no se ha definido su papel en la regulación por impronta genómica. Hasta hace poco, todas las DMRs asociadas a éste mecanismo se clasificaban como DMRs germinales o secundarias. Las DMRs germinales, también conocidas como DMRs primarias, heredan su metilación del ovocito o el espermatozoide y, en general, se encuentran de forma ubicua en todos los tejidos. Por otro lado, las DMRs secundarias adquieren su metilación durante el desarrollo, de forma jerárquica y específica de tejido, dictada por una ICR vecina. Cribados genómicos indican que en humanos ya han sido identificadas la mayoría de regiones reguladas por impronta genómica de manera ubicua. Por el contrario, debido a la escasez de investigaciones asociadas, se cree que siguen sin haber sido descritas la mayoría de regiones reguladas por este mecanismo de forma específica en un solo tejido.

Mi hipótesis inicial ha sido que, *en comparación con los tejidos somáticos, la placenta humana se encuentra enriquecida de genes regulados por impronta genómica y que ésta regulación podría tener un papel esencial en el crecimiento y desarrollo intrauterino.* Con el fin de determinar el alcance de la regulación por impronta genómica en la placenta humana y cómo estos genes pueden influir en nuestro crecimiento y desarrollo, el trabajo llevado a cabo en esta tesis se ha centrado principalmente en: (1) estudiar *conceptus* de

mujeres con ambos alelos del gen *NLRP7* mutados; (2) confirmar nuevas DMRs presentes específicamente en placenta identificadas mediante aproximaciones bioinformáticas; (3) realizar estudios de expresión y de otras modificaciones epigenéticas (como las marcas de histonas y hidoximetilación) en estas regiones; y finalmente, (4) evaluar la estabilidad de la metilación alélica en las DMRs específicas de placenta y su asociada expresión genética en placentas de embarazos complicados o asociados al uso de técnicas de reproducción asistida (TRA).

Inicialmente se pensó que la mayoría de las DMRs específicas de placenta adquirían la metilación en el alelo materno durante el desarrollo (es decir, que eran DMRs secundarias). Sin embargo, durante el inicio de esta tesis, la publicación del metiloma del ovocito humano maduro reveló que todas las DMRs específicas de placenta heredan su metilación de la línea germinal materna. Durante el transcurso de este trabajo hemos descrito **72** nuevas regiones monoalélicamente metiladas en placenta, y en 50 de las cuales se ha podido confirmar la metilación específica del alelo materno. En el presente trabajo también se identifica la expresión monoalélica de **20** genes (confirmándose la expresión del alelo paterno en 17 de ellos) cuyos promotores contienen una DMR. La primera lista de DMRs específicas de placenta confirmada en esta tesis se obtuvo como resultado de comparar los perfiles de metilación de MHRs (procedentes de mujeres con ambos alelos mutados del gen *NLRP7*) y placentas control. El estudio de estas MHRs mostró una falta de metilación global del alelo materno en las regiones reguladas por impronta genómica acompañada por la expresión bialélica de los genes asociados. Estudios posteriores de metilación y expresión mostraron que las DMRs específicas de placenta regulan la expresión monoalélica de un solo gen por dominio (en contraposición a los *clusters* de genes asociados a las DMRs ubicuas) y que ésta expresión exclusiva del alelo paterno ya se observa en embriones preimplantacionales tras la activación genómica del cigoto, siendo el gen *ZHX3* un ejemplo de ello.

Mediante el estudio de estas regiones en nuestra cohorte de placentas, hemos podido observar que la impronta genómica en la placenta humana es un fenómeno polimórfico y que la metilación asociada a estas DMRs no está formada en su mayoría por hidoximetilación. Por lo tanto, el perfil de metilación descrito en estas regiones no es un estado intermedio del proceso de demetilación. Estudios adicionales de otros mecanismos epigenéticos reveló que la expresión bialélica polimórfica se correlaciona con la presencia de marcas permisivas de histonas, en algunos casos, independiente del estado de metilación, es decir, aunque haya una DMRs en éste promotor, ambos alelos presentan H3K4me2/3. Después de evaluar la estabilidad de la metilación alélica tanto en DMRs ubicuas como específicas de placenta, no observamos una mayor frecuencia de metilación polimórfica en placentas de embarazos complicados o concebidos con TRAs.

Nuestros resultados sugieren que las DMRs propias de la placenta han surgido de manera específica en el linaje de los primates. Debido al gran número de regiones, su mantenimiento

como DMRs en placenta parece estar más asociada a una adaptación conjunta de la maquinaria epigenética del ovocito, embrión y placenta que no a un nuevo contexto genómico surgido de manera simultánea en todas estas regiones. Esta apreciación es consistente con que *NLRP7*, un miembro específico de primates de la familia proteica NLRP, sea un gen de efecto materno potencialmente implicado en el establecimiento en el ovocito de las DMRs reguladas por impronta genómica – incluyendo las DMRs específicas de placenta – y su mantenimiento en cigoto. Anteriormente se había propuesto que el fenotipo de las molas hidatiformes, tanto las esporádicas como las recurrentes, está asociado con una incorrecta regulación de la impronta genómica. Sin embargo, antes de los estudios presentados en esta tesis doctoral, no se había realizado ningún análisis sistemático de los defectos en la regulación de éste mecanismo. La expresión aberrante del alelo materno de algunas DMRs específicas de placenta podría jugar un papel importante en el fenotipo de las MHRs, ya que incluyen genes esenciales involucrados en diferentes procesos biológicos como las modificaciones de histonas o procesos metabólicos. Estas observaciones refuerzan la hipótesis de que el silenciamiento del alelo materno en estos genes durante el desarrollo embrionario temprano y en placenta, sería un mecanismo para prevenir el teratoma ovárico que surge de ovocitos activados partenogenéticamente. Sin embargo, el hecho de observar la expresión bialélica polimórfica en placentas control no asociadas con embarazos complicados ni con el uso de TRA, sugiere que el fenotipo severo asociado con molas hidatiformes sólo ocurre cuando se altera de manera global el mecanismo de impronta genómica. Por otro lado, anomalías del desarollo más leves como la restricción del crecimiento intrauterino, podrían ocurrir cuando sólo una o pocas DMRs se vieran afectadas. A pesar de que no hemos observado más regiones con falta de metilación en placentas de embarazos complicados o concebidas por TRAs en comparación con placentas control, se necesita una mayor caracterización, incluyendo estudios de expresión por métodos cuantitativos, para determinar el papel de los genes regulados por impronta genómica de específica en placenta en el desarrollo y crecimiento intrauterino.

Finalmente, teniendo en cuenta los resultados de esta tesis doctoral, podemos concluir que hay más dominios regulados por impronta genómica de forma exclusiva en la placenta que en tejidos somáticos y que la metilación en DMRs ubicuas es más estable que la de las DMRs presentes exclusivamente en placenta.

# Resum

La **impressió genòmica** és un fenomen epigenètic pel qual un subconjunt de gens s'expressa de manera monoal · lèlica depenent de l'origen parental de cada al · lel. En mamífers, els gens regulats mitjançant aquest mecanisme es troben generalment en grups, regulats per regions de control de la impressió (ICR, per les seves sigles en anglès). Aquestes ICRs acostumen ser regions diferencialment metilades (DMRs, per les seves sigles en anglès) en les quals la metilació de l'ADN marca un dels al · lels parentals, influint en l'expressió dels gens circumdants. L'adquisició i el manteniment d'aquestes DMRs depèn de la interacció entre processos de transcripció, modificacions d'histones i l'acció de factors que influeixen en la regulació epigenètica (com per exemple, DNMTs, TRIM28, DPPA3, ZFP57 i UHRF1). A més, es creu que altres factors també podrien estar implicats en aquesta regulació epigenètica, com per exemple *KHDC3L* i *NLRP7*, i que les seves mutacions en homozigosi estan associades amb moles hidatiformes recurrents (MHR). Tot i això, encara no s'ha definit el seu paper en la regulació per impressió genòmica. Fins fa poc, totes les DMRs associades a aquest mecanisme es classificaven com DMRs germinals o secundàries. Les DMRs germinals, també conegudes com DMRs primàries, hereten la seva metilació de l'oòcit o l'espermatozoide i, en general, es troben ubiquament en tots els teixits. D'altra banda, les DMRs secundàries adquireixen la seva metilació durant el desenvolupament, de forma jeràrquica i específica de teixit, i dictada per una ICR veïna. Cribatges genòmics indiquen que, en humans, ja s'han identificat la majoria de regions regulades per impressió genòmica de manera ubiqua. Per altra banda, a causa de la mancança de recerques associades, es creu que segueixen sense haver estat descrites la majoria de regions regulades per aquest mecanisme d'una manera específica de teixit.

La meva hipòtesi inicial ha estat que, *en comparació amb els teixits somàtics, la placenta humana es troba enriquida en gens regulats per impressió genòmica i que aquesta regulació podria tenir un paper essencial en el creixement i desenvolupament intrauterí.* Amb la finalitat de determinar la magnitud de la regulació per impressió genòmica en la placenta humana i com aquests gens poden influir en el nostre creixement i desenvolupament embrionari, el treball dut a terme en aquesta tesi es va centrar principalment en: (1) analitzar *conceptus* de dones amb tots dos al · lels del gen *NLRP7* mutats; (2) confirmar

noves DMRs presents específicament a la placenta i identificades mitjançant aproximacions bioinformatiques; (3) realitzar estudis d'expressió i d'altres modificacions epigenètiques (com les marques d'histones i hidoximetilació) en aquestes regions; i finalment, (4) avaluar l'estabilitat de la metilació al · lèlica en les DMRs específiques de placenta i la seva expressió genètica associada en placentes de gestacions complicades o associades a l'ús de tècniques de reproducció assistida (TRA).

Inicialment es va pensar que la majoria de les DMRs específiques de placenta adquirien la metilació de l'al · lel matern durant el desenvolupament, és a dir, que eren DMRs secundàries. No obstant això, durant l'inici d'aquesta tesi, la publicació del metiloma de l'oòcit humà madur va revelar que totes les DMRs específiques de placenta hereten la seva metilació de la línia germinal materna. Durant el transcurs d'aquest treball hem descrit **72** noves regions metilades de manera monoal · lèlica a la placenta, i en 50 de les quals s'ha pogut confirmar la metilació específica de l'al · lel matern. En el present treball també s'identifica l'expressió monoal · lèlica de **20** gens (confirmant-se l'expressió de l'al · lel patern en 17 d'ells), els promotors dels quals contenen una DMR. La primera llista de DMRs específiques de placenta confirmada en aquesta tesi es va obtenir com a resultat de comparar els perfils de metilació de MHRs (provinents de dones amb tots dos al · lels mutats del gen *NLRP7*) i placentes control. L'estudi d'aquestes MHRs va mostrar una falta de metilació global de l'al · lel matern a les regions regulades per impressió genòmica acompanyada per l'expressió bial · lèlica dels gens associats. Estudis posteriors de metilació i expressió van mostrar que les DMRs específiques de placenta regulen l'expressió monoal · lèlica d'un sol gen per domini i que aquesta expressió exclusiva de l'al · lel patern ja s'observa en embrions pre-implantacionals després de l'activació genòmica del zigot (exemplificat amb el gen *ZHX3*).

Mitjançant l'estudi d'aquestes regions en la nostra cohort de placentes, hem pogut observar que l'impresió genòmica a la placenta humana és un fenomen polimòrfic i que la metilació associada a aquestes DMRs no es troba formada en la seva majoria per hidoximetilació. Per tant, el perfil de metilació descrit en aquestes regions no és un estat intermedi del procés de demetilació. Estudis addicionals d'altres mecanismes epigenètics van revelar que l'expressió bial · lèlica polimòrfica es correlaciona amb la presència de marques permissives d'histones, i en alguns casos, independent de l'estat de metilació, és a dir, encara que estigui present la DMRs en aquest promotor, tots dos al · lels podrien presentar H3K4em2/3. Després d'avaluar l'estabilitat de la metilació al · lèlica tant en DMRs ubiqües com específiques de placenta, no vàrem observar una major freqüència de metilació polimòrfica a les placentes d'embarassos complicats o concebuts amb TRAs.

Sembla ser que les DMRs restringides a la placenta han sorgit de manera específica en el llinatge dels primats. Degut al gran nombre de regions, el seu manteniment com DMRs a la placenta sembla estar més associat a una adaptació conjunta de la maquinària

epigenètica de l'oòcit, l'embrió i la placenta que no pas a un nou context genòmic sorgit de manera simultània en totes aquestes regions. Aquestes observacions són consistents amb que *NLRP7*, un membre específic de primats de la família proteica NLRP, sigui un gen d'efecte matern potencialment implicat en l'establiment en l'oòcit de les DMRs regulades per impressió genòmica – incloent les específiques de placenta – i el seu manteniment al zigot. Anteriorment, s'havia proposat que el fenotip de les moles hidatiformes, tant les esporàdiques com les recurrents, estava associat amb una incorrecta regulació de l'impressió genòmica. Tot i això, abans dels estudis presentats en aquesta tesi doctoral, no s'havia realitzat cap anàlisi sistemàtica dels defectes en la regulació d'aquest mecanisme. L'expressió aberrant de l'al · lel matern d'algunes DMRs específiques de placenta podria jugar un paper important en el fenotip de les MHRs, ja que inclouen gens essencials involucrats en diferents processos biològics com les modificacions d'histones o processos metabòlics. Això reforça la hipòtesi de que el silenciamient de l'al · lel matern en aquests gens durant el desenvolupament embrionari primerenc i a la placenta, és un mecanisme per prevenir el teratoma ovàric que sorgeix d'oòcits activats partenogenèticament. No obstant això, el fet d'observar l'expressió bial · lèlica polimòrfica en placentes control no associades amb embarassos complicats ni l'ús de TRA suggereix que el fenotip sever associat amb moles hidatiformes només succeeix quan s'altera de manera global el mecanisme d'impressió genòmica. per l'altra banda, anomalies del desenvolupament més lleus, com la restricció del creixement intrauterí, podrien ocórrer quan una o poques DMRs es vegin afectades. Tot i que no hem observat més regions amb falta de metilació en placentes de gestacions complicades o concebudes mitjançant TRAs (en comparació amb placentes control), es necessita una major caracterització, incloent estudis d'expressió per mètodes quantitatius, per tal de determinar el paper dels gens regulats per impressió genòmica específicament a la placenta sobre el desenvolupament i creixement intrauterí.

Finalment, tenint en compte els resultats d'aquesta tesi doctoral, podem concloure que hi ha més dominis regulats per impressió genòmica exclusivament a la placenta que en teixits somàtics i que la metilació en DMRs ubiqües és més estable que la de les DMRs presents exclusivament a la placenta.

# Chapter 1

# Introduction

## 1.1 Epigenetic mechanisms and genomic imprinting

The term *epigenetics* was used for the first time in the early 1940s, when it was defined as:

> *the branch of biology, which studies the causal interactions between genes and their products which bring the phenotype into being* (Waddintong, 1942).

Over the years, the concept of epigenetics has changed and is now generally accepted as:

> *the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence* (C.-t. Wu & Morris, 2001).

Therefore, the principals of epigenetics have evolved from being the way to get a specific phenotype by interacting with the genotype, to a more complex concept, as an inherited process encompassing different mechanisms, by which the transcription of DNA is regulated without changing the nucleotide sequence. As will be discussed in depth in the following pages, epigenetic mechanisms are not only essential for the tissue-specific gene regulation and dynamic expression patterns at different stages of development but are also crucial for *genomic imprinting* and *X chromosome inactivation* (Fedoriw, Mugford, & Magnuson, 2012).

**Genomic imprinting** is an epigenetic phenomenon resulting in the monoallelic expression of a subset of genes in a parent-of-origin-specific manner. This phenomenon is described in the sex determination mechanism of some insects but also in autosomal chromosome genes in therian mammals and angiosperm plants.

In general, mammalian imprinted genes are organised in clusters and regulated by an imprinting control region (ICR). In most cases, ICRs manifest as a differentially methylated

region (DMR) where cytosine methylation at its $5^{th}$ position (5-methylcytosine; 5-mC) in CpG dinucleotides marks one of the parental alleles, providing cis-acting regulatory elements that influence the allelic expression of surrounding genes. As we will see in this section, transcription and histone modifications are also critically involved in the acquisition and maintenance of these imprinted DMRs.

Until recently, all described imprinted DMRs were classified as germline or secondary DMRs. Germline DMRs (gDMRs), also known as primary DMRs, inherit methylation from the oocyte or sperm and are, in general, ICRs. Therefore, these primary DMRs are presented in all tissues and developmental stages of the embryo. On the other hand, secondary DMRs acquire methylation during development, in a tissue-specific manner and in a hierarchical fashion dictated by a neighbouring ICR (see Figure 1.1) (John & Lefebvre, 2011).

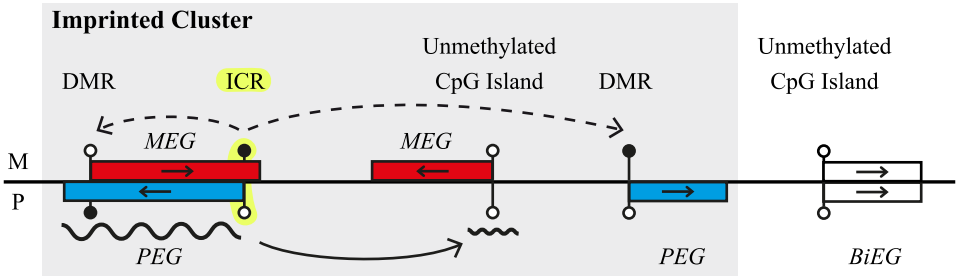



Figure 1.1: **General features of an imprinted cluster.** The maternally inherited chromosome (**M**, the upper part of the line) and the paternally inherited chromosomes (**P**, the lower part) of an imprinted locus are represented. The black lollipops indicate methylated CpG regions whereas white circles, unmethylated CpG regions. *MEG*: maternally expressed gene (red filed rectangle); *PEG*: paternally expressed gene (blue filled rectangle); *BiEG*: non-imprinted biallelically expressed genes (white filled rectangle). The arrows represent the direction of the transcripts.

### 1.1.1 The role of histone modifications in biological processes

DNA is not free in the nuclei of eukaryotic cells; it is associated with different protein factors and RNA forming the chromatin. The most compact state of chromatin is found when the cell is ready to divide and present as highly organised complexes: the metaphase chromosomes. The number of chromosomes is different in each eukaryotic species, but the organisation of the chromatin is similar in all of them (A. L. Olins & Olins, 1974). The chromatin is a dynamic structure allowing for DNA replication and gene expression/repression. In the chromatin, the DNA is wrapped around a core of eight histone proteins forming the nucleosome. This nucleosome octamer is comprised of two copies each of histone proteins H2A, H2B, H3 and H4 (Kornberg, 1974). These proteins undergo post-translational modifications of N-terminal tails, which influences

their interaction with DNA and other nuclear factors, including transcription factors or other DNA binding proteins (see Figure 1.2 for histone modifications mainly associated with active or repressive genes, extracted from: I. Iglesias-Platas & Monk (2016)).

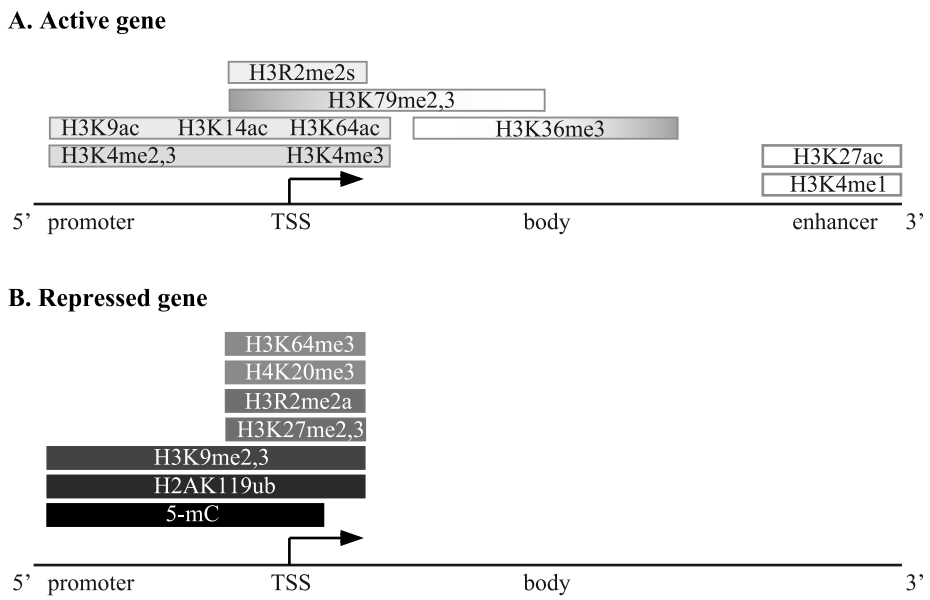



Figure 1.2: **Epigenetic modifications regulating gene expression.** Histone modifications mainly associated with (**A**) permissive and (**B**) repressive chromatin at expressed or silent genes, respectively. 5-mC also marks repressive regions. Colour code indicates the strongest (black) to weakest (white) degree of stability of epigenetic modification (Iglesias-Platas & Monk, 2016).

Histone modifications, globally called as histone code, are involved in different biological processes, including regulation of gene expression, DNA repair and chromosome condensation (see Table 1.1). The exact role of some histone tail modifications is unknown, but in the last years, the scientific community has made significant advances in this field. One of the best known is lysine (K) acetylation, which is a recurrent histone mark associated with active gene expression. This modification neutralises the positively charged lysine residues so that the binding with DNA becomes weak, thus making it more accessible. However, the nature of other histone modifications is unclear, while both repressive and permissive marks can co-localise nearby genes with poised transcription, called *bivalent domains* (B. E. Bernstein et al., 2006; Z. Wang et al., 2008).

Table 1.1: **The main histone modifications and its biological roles**.

| Histone | Modification | Role |
|---|---|---|
| H2A | H2AS1P | Mitosis; chromatin assembly |
| | H2AK4ac | Transcriptional activation |

| Histone | Modification | Role |
|---|---|---|
| | **H2AK5ac**[1] | Transcriptional activation; imprinting |
| | H2AK7ac | Transcriptional activation |
| | H2AK119P | Spermatogenesis |
| | **H2AR3me2** | Imprinting |
| | **H2AK119ub** | Transcriptional repression; X-inactivation; imprinting |
| H2B | H2BS14P | Apoptosis |
| | H2BS33P | Transcriptional activation |
| | H2BK5ac | Transcriptional activation |
| | H2BK11ac | Transcriptional activation |
| | **H2BK12ac** | Transcriptional activation; imprinting |
| | H2BK15ac | Transcriptional activation |
| | **H2BK16ac** | Transcriptional activation; imprinting |
| | **H2BK46ac** | Imprinting |
| | H2BK20ac | Transcriptional activation |
| | H2BK120ub | Spermatogenesis/meiosis |
| | H2BK123ub | Transcriptional activation |
| H3 | **H3K4me2** | Permissive euchromatin; imprinting |
| | **H3K4me3** | Transcriptional elongation; active euchromatin; imprinting |
| | **H3K9me2/3** | Transcriptional repression; DNA methylation; imprinting |
| | H3R17me | Transcriptional activation |
| | **H3K27me3** | Transcriptional silencing; bivalent genes; X-inactivation; imprinting |
| | **H3K36me3** | Transcriptional elongation; imprinting |
| | **H3K36ac** | Imprinting |
| | **H3K64ac** | Imprinting |
| | **H3K4ac** | Transcriptional activation; imprinting |
| | **H3K9ac** | Transcriptional activation; imprinting |
| | **H3K14ac** | Transcriptional activation; DNA repair; imprinting |
| | **H3K79me2** | Transcription activation; mitosis; imprinting |
| | **H3K79ac** | Imprinting |
| | **H1K18ac** | Transcriptional activation; DNA repair; DNA replication; imprinting |
| | H3K23ac | Transcriptional activation; DNA repair |
| | **H3K27ac** | Transcriptional activation; imprinting |
| | H3T3P | Mitosis |
| | H3S10P | Mitosis; meiosis; transcriptional activation |

[1] In bold, histone modifications observed at imprinted loci (J. Cao & Yan, 2012; Lawrence, Daujat, & Schneider, 2016; P. Singh et al., 2010)

| Histone | Modification | Role |
|---|---|---|
| | H3T11P | Mitosis |
| | H3S28P | Mitosis |
| H4 | **H4R3me** | Transcriptional activation; imprinting |
| | **H4K20me1** | Transcriptional silencing; imprinting |
| | **H4K20me3** | Heterochromatin; imprinting |
| | **H4K5ac** | Transcriptional activation; DNA repair; imprinting |
| | **H4K8ac** | Transcriptional activation and elongation; DNA repair; imprinting |
| | **H4K12ac** | Telomeric silencing; transcriptional activation; DNA repair; imprinting |
| | **H4K16ac** | Transcriptional activation; DNA repair; imprinting |
| | H4S1P | Mitosis |

Genomic imprinting is also associated with specific histone modifications. In the tissues or specific cell types in which they are monoallelically expressed, opposing histone modifications mark the active and repressed alleles at ICRs. The methylated allele is associated with H3K9me3 and H4K20me3, both generally linked with heterochromatin (Monk et al., 2011). The unmethylated and often expressed allele is mainly associated with H3K4me3, which is a permissive mark, in general, associated with gene expression (Regha et al., 2007). Although methylation at primary DMRs is present in a parent-of-origin manner in the entire organism, imprinted genes are not always expressed in all tissues. Studies in mice have revealed that at some imprinted DMRs, the unmethylated but not expressed allele is marked by monoallelic bivalent chromatin (H3K4me2/3 and H3K27me3) (Sanz et al., 2008). This chromatin state seems to protect the unmethylated region from de novo methylation and at the same time, evading expression.

### 1.1.2 The importance of DNA methylation

Although methylation at $6^{th}$ position of the adenine base has been described in some bacteria, fungi, unicellular eukaryotic organisms or nematodes, DNA methylation is found almost exclusively at the $5^{th}$ position of the cytosine base in most eukaryotes (Greer et al., 2015; Heithoff, Sinsheimer, Low, & Mahan, 1999; Law & Jacobsen, 2010; Mondo et al., 2017). Throughout the eukaryotic genome, DNA methylation is generally found in two different DNA contexts showing opposite function: silencing transposable elements and being positively correlated with gene expression when is present in gene bodies (Z. D. Smith & Meissner, 2013). Cytosine methylation is differently distributed throughout the genome, cell types and developmental stages in each eukaryote species, being more present

in plant genomes – due to the high transposable element content – and absent in model organisms such as the *Saccharomyces cerevisiae* yeast (Capuano, Muslleder, Kok, Blom, & Ralser, 2014; Proffitt, Davie, Swinton, & Hattman, 1984). In all vertebrates, 5-mC is predominantly present at CpG-dinucleotides. However, recent whole-genome studies in mammals have shown the existence of 5-mC at non-CG DNA methylation mainly in pluripotent cells and brain (W. Guo, Zhang, & Wu, 2016; Y. He & Ecker, 2015; Luo et al., 2017). Moreover, in 2017 a study was published revealing an allele-specific non-CG DNA methylation associated with active chromatin domains at X chromosome in female mouse neurones, which is consistent with the active X chromosome having higher 5-mC levels in gene bodies of expressed genes to stop cryptic initiation of expression (Keown et al., 2017; Sharp et al., 2011).

Vertebrate genomes are globally methylated, except for the short unmethylated CG-rich sequences termed *CpG islands*. These sequences are present in approximately the 70% of annotated vertebrate gene promoters (M. M. Suzuki & Bird, 2008). Although most CpG islands remain unmethylated throughout mammalian development regardless of the gene expression state, some of them become methylated during embryo development – e.g. X chromosome inactivation – or inherits its methylation states from the gametes – e.g. a genomic imprint –. In these contexts, the cytosine methylation at CG-rich promoters is generally correlated with stable gene silencing. But not all CpG islands are located in gene promoters. In both, the human and mouse genomes, approximately half of all CpG islands are distributed within gene bodies or between genes. Some of these CG-rich regions are associated with intragenic alternative promoters or with non-coding RNA (ncRNA) transcription (Deaton & Bird, 2011; M. M. Suzuki & Bird, 2008). These ncRNAs can be *in cis* regulators of gene expression. Some ICR, mainly the maternally methylated ones, are associated with the expression of long ncRNA, which in turn regulate the parent-of-origin expression of nearby genes. For example, in mice, the *Kv*DMR1 is a maternally methylated region which silences the maternal allele of a large 100 kb intergenic ncRNA transcript called *Kcnq1ot1*. This paternally expressed transcript ensures the silencing of flanking genes by recruiting the histone modifying machinery resulting in maternal expression (Figure 1.3) (Deaton & Bird, 2011).

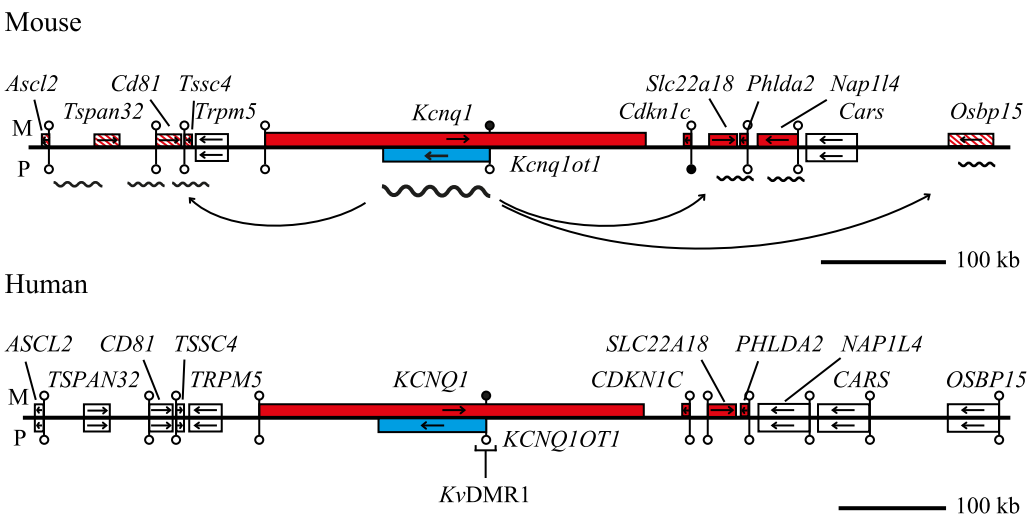



Figure 1.3: ***KCNQOT1* imprinted cluster in the mouse and human placenta.** Schematic representation showing the relative organisation of genes, CpG islands and DMRs for the *KCNQOT1* domain on mouse chr7qF5 and human chr11p15. The upper part of the line represents maternally inherited chromosome (**M**) and lower part paternally inherited chromosome (**P**). Black lollipops represent methylated CpG regions whereas white circles unmethylated CpG regions. Red filed rectangle represents the maternally expressed gene; red hatched rectangles represents placenta-specific imprinted transcripts. Blue filled rectangle paternally expressed gene and white filled rectangle biallelic expression of the non-imprinted gene. The arrows represent the direction of the transcripts. Note that the evolutionary differences within this domain also include the absence of somatically acquired (secondary DMR) maternal methylation at the *CDKN1C* promoter in humans. **Adapted from:** Monk (2015).

### 1.1.3 DNA methylation dynamics in mammals and its players

In mammals, DNA methylation at $5^{th}$ position of the cytosine base is catalysed by different family members of the DNA-methyltransferases proteins (DNMTs). The DNMT3 family (DNMT3A and DNMT3B) is involved in the *de novo* establishment of methylation and DNMT1 in the maintenance of the methylation pattern during cell division by having a high affinity for hemimethylated CpGs (T. Bestor, Laudano, Mattaliano, & Ingram, 1988; Okano, Xie, & Li, 1998). But DNA methylation is a dynamic process. Epigenetic reprogramming occurs at two specific stages of development: during primordial germ cells (PGC) determination and during mammalian early embryonic development, where genome-wide 5-mC demethylation and remethylation occurs (Seisenberger et al., 2013).

DNA demethylation can take place passively, when the loss of 5-mC occurs during successive rounds of replication in the absence of functional DNA methylation maintenance factors (mainly, DNMT family and related members); or mediated by the ten-eleven translocation (TET) enzymes, which can oxidize 5-mC to 5-hydroxymethylcytosine (5-

hmC), 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC) (Figure 1.4) (Kohli & Zhang, 2013). This second process, the oxidation of 5-mC, could involve an active erasure of 5-mC associated with DNA repair pathways or passive erasure by oxidising 5-mC to 5-hmC that is lost in each cell-replication since oxidative derivatives of 5-mC are not recognized by DNMT1 (A. Inoue, Shen, Dai, He, & Zhang, 2011; Messerschmidt, Knowles, & Solter, 2014).

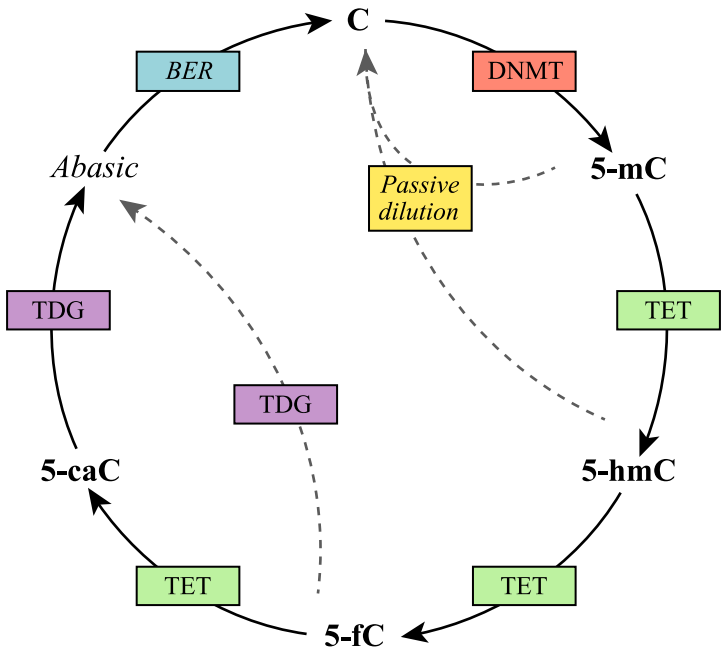



Figure 1.4: **Cytosine modification dynamics.** Cytosine methylation by DNMT enzymes and 5-mC, which can be oxidated to 5-hmC or passively demethylated in a replication-dependent manner. At the same time, 5-hmC can be passively diluted (since DNMT enzymes do not recognise 5-hmC modification) or oxidated to 5-fC and 5-caC. By the TDG action, 5-fC or 5-caC can be excised as part of the BER process that generates unmodified cytosines. *5-hmC*: 5-hydroxymethylcytosine; *5-fC*: 5-formylcytosine; *5-caC*: 5-carboxylcytosine; *TDG*: Thymine DNA glycosylase; *BER*: base excision repair. **Adapted from:** Kohli & Zhang (2013)

In mice, just after fertilisation, the zygote genome undergoes to DNA demethylation involving both, conversion to 5-hmC by the action of TET3 in the paternally-derived pronucleus (which will be lost in cell-replication), and the slower passive 5-mC demethylation due to a gradual non-recognition of DNMT proteins in the maternally-derived pronucleus (Figure 1.5) (Iqbal, Jin, Pfeifer, & Szabó, 2011; Smallwood et al., 2011). This 5-mC erasure culminates in a global DNA demethylation in the preimplantation embryo. Subsequently, at implantation, *de novo* remethylation takes place in the blastocyst by the action of DNMT3A and DNMT3B. As it will be described at the end of this section, imprinted DMRs escapes from embryonic epigenetic reprogramming, maintaining an intermediate

50% methylation level representing one fully methylated allele and one fully unmethylated (Wossidlo et al., 2011).

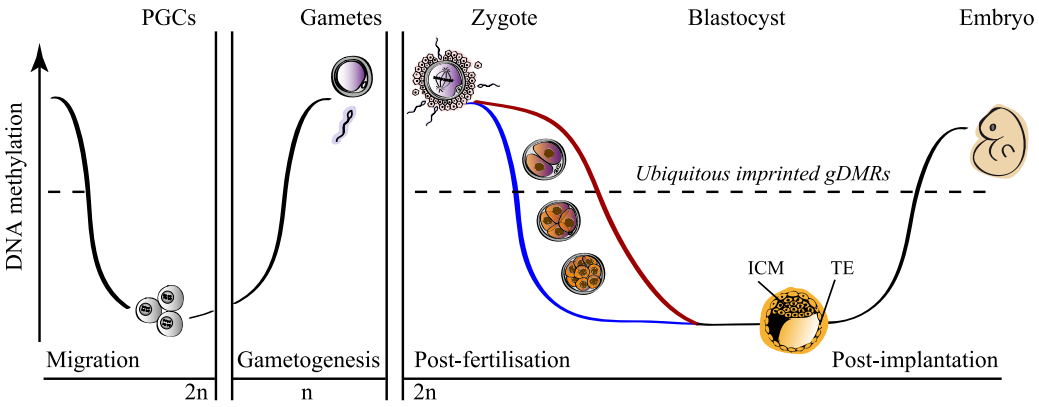



Figure 1.5: **Methylation dynamics during development.** Black (general) blue (sperm-derived) and red (oocyte-derived) lines represent the general methylation dynamics in the genome; the Y-axis represents a general gain- or loss-of-methylation, not a specific % of DNA methylation; the X-axis represents the general developmental stages in eutherian mammals, without a specific time-scale. The first part of the graph represents the general erasure of DNA methylation occurring during the migration of the PCGs (primordial germ cells) to the gonadal ridge during embryonic development, with the following methylation acquisition during gametogenesis, when mitosis occurs (**n**). The gametogenesis occurs in a different developmental-point and duration in the sperm and the oocyte and also, depending on the mammalian species. After fertilisation the embryo's genome (**2n**) start to lose methylation during the post-fertilisation epigenetic reprogramming: by the action of the TET3 protein in sperm-derived methylation (blue line) and depending on DNA replication for the oocyte-derived methylation (red line). After the blastocyst stage (*ICM*: inner cell mass; *TE*: trophectoderm), during embryo's implantation occurs the second epigenetic reprogramming, with a general DNA gain-of-methylation. The dashed line represents the DNA methylation state at ubiquitous imprinted gDMRs (differentially methylated regions presented as DMR in all embryonic and adult tissues), which is stable during all embryonic development and comparing with the other methylated regions of the genome, imprinted DMRs have an intermediate methylation state, which is also erased during PGC migration.

After implantation, a subset of cells is instructed to become PGCs. These PGCs undergo to a complex epigenetic reprogramming process which includes erasure of genome-wide DNA methylation patterns, to acquire the totipotency which enables differentiation to mature gametes – either the high methylated sperm or the partially methylated oocyte – (Hackett, Zylicz, & Surani, 2012; Smallwood et al., 2011). Uniquely, this demethylation includes imprinted regions and most repeat elements (Hajkova et al., 2002; Lane et al., 2003). In mice PCGs, demethylation occurs in a two-step process that starts with the rapid PGCs expansion with global passive demethylation since the DNMT enzymes are not expressed, followed by the conversion of 5-mC to 5-hmC by the action of TET1 and TET2 proteins (Hackett et al., 2013; Kagiwada, Kurimoto, Hirota, Yamaji, & Saitou,

2013). During this second step, 5-hmC is lost during replication since DNMT1 is not able to recognise and maintain this DNA modification (Hashimoto et al., 2012).

In August 2014, the dynamic pattern of the early human embryo methylome was published (H. Guo et al., 2014). In this work, it was shown that the major genome-wide demethylation in the human embryo is complete at the 2-cell stage, contrary to the previous observations in mice, in which the most marked demethylation occurs at the zygotic stage, with mild-gradual demethylation until the blastocyst stage (Z. D. Smith et al., 2012).

### 1.1.4 The life cycle of genomic imprints

Although genomic imprints encompass other epigenetic mechanisms, such as histone modifications or the action of ncRNAs, allelic DNA methylation at these loci is the most characterised feature. During mammalian gametogenesis, the critical DMRs for imprinting control are established. The majority of known gDMRs are maternally methylated and often associated with CpG island promoters, acquiring their methylation after birth in the growing oocyte. Only a few of gDMRs are paternally methylated (2 in humans and 3 in mice) and are intragenically located and relatively CpG poor (S.-P. Lin et al., 2003; Vu et al., 2000). This differentially methylated pattern is maintained during the embryonic development – and later in the adult tissues – except for the PGCs, where imprints are erased, and the new methylation pattern will be established depending on the embryo's gender.

The dynamic nature of methylation associated with the imprinting cycle has generally been determined in mouse, with different protein factors identified in each step of the process (Figure 1.6). Thus, the majority of DMRs present in somatic mouse tissues acquires their allelic methylation during gametogenesis, when the two parental genomes are separated, resulting from the cooperation of the *de novo* methyltransferase DNMT3A and its co-factor DNMT3L, which acts specifically during imprint acquisition and recognises H3K4 when unmodified (H3K4me0) (Figure 1.7) (D. Bourc'his, Xu, Lin, Bollman, & Bestor, 2001; Hata, Okano, Lei, & Li, 2002; Ooi et al., 2007). Besides the interaction between these factors, transcription and histone modifications are critically involved in this DNA methylation acquisition of both oocyte and spermatozoa (Gahurova et al., 2017; Henckel, Chebli, Kota, Arnaud, & Feil, 2012). For instance, in the growing mouse oocyte, the action of the H3K4 demethylase KDM1B is required to remove H3K4me3 before DNA methylation can be acquired (Figure 1.7B) (Ciccone et al., 2009; K. R. Stewart et al., 2015; Q. Zhang et al., 2013). In male PGC, there is an H3K4m3 enrichment at the sequences corresponding to maternally methylated gDMRs, protecting them from methylation in the sperm and also later, during embryo development (Henckel et al., 2012). Interestingly, on the unmethylated maternal allele of the sperm-derived methylated H19 DMR, both H3K4m3 enrichment and presence of CTCF binding protein acts as an insulator to protect

the region from *de novo* remethylation that takes place in the blastocyst after implantation (Rand, Ben-Porath, Keshet, & Cedar, 2004; Schoenherr, Levorse, & Tilghman, 2003).

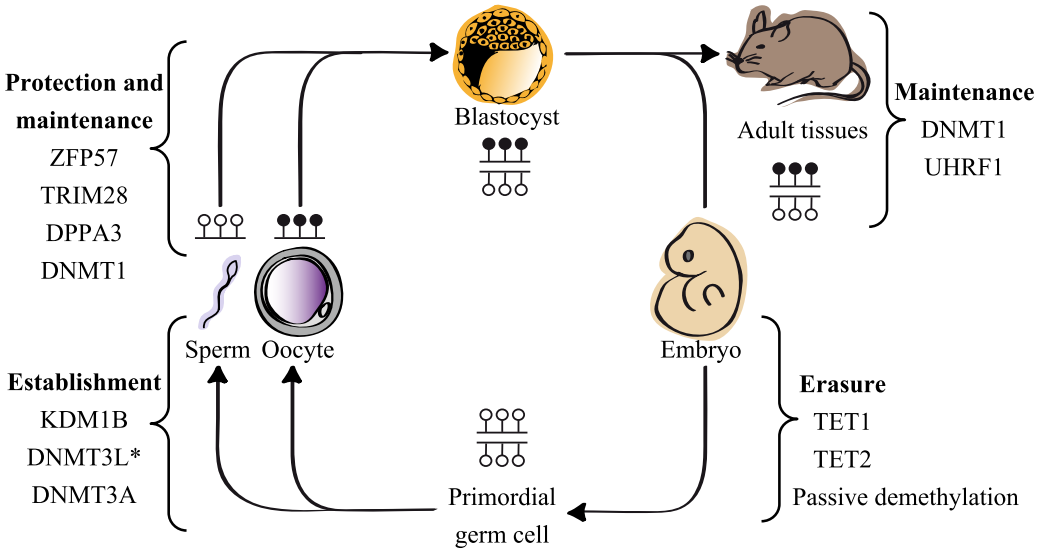



Figure 1.6: **Imprint methylation cycle in mice.** Schematic representation of DNA methylation dynamics at imprinted DMRs with the main factors involved in protection and maintenance during the early embryonic development (upper-left), maintenance during embryos development and adult tissue (upper-right), erasure in primordial germ cells (bottom-right) and establishment in gametes (bottom-left). Black lollipops represent methylated CpG whereas white circles unmethylated CpG. **Adapted from:** Sanchez-Delgado et al. (2016a).

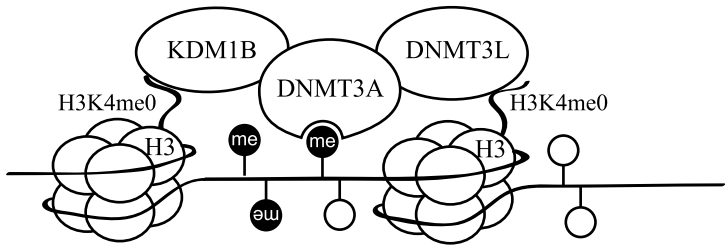
To date, all known gDMRs maintain their methylation status through embryonic development and later in life by a combination of generic and imprinting specific factors. Just after fertilisation, the zygote genome is not immediately transcribed, and maternal proteins originating from the fertilised oocyte have a key role following zygote formation. In 2006, the maternal factor DPPA3 was first described to prevent the DNA demethylation of all epigenetic asymmetries during this first stage of development, including the preservation of imprinted DMRs (Figure 1.8) (Nakamura et al., 2006). Predominantly observed at maternally derived imprints, DPPA3 selectively protects methylation from TET3-associated demethylation, specifically at sequences containing the $YYCAGSCTSS$ motif that is also marked by H3K9me2 (Bian & Yu, 2013; Nakamura et al., 2006, 2012).

Two years later, in 2008, it was also reported in mice that the maternal-zygotic effect factor ZFP57 maintains both sperm and oocyte-derived methylation imprints (X. Li et al., 2008). Likewise, the oocyte-derived DNMT1 and UHRF1 proteins have been recently described to have an essential role in the maintenance of both maternal and paternal imprints in the early stages of embryo development, as well as in the maintenance of other methylated non-imprinted regions (Maenohara et al., 2017; Quenneville et al., 2011). The

methylation protection at imprinted DMRs containing the $KGCC^{met}GC$ is associated with the recruitment of the ZFP57-TRIM28 complex. Additionally, this complex, together with the action of the hemimethylated-CG-binding protein UHRF1 and CBX5/HP1 recognition of H3K9me3, recruit DNMT1 for the DNA methylation maintenance and SETDB1 for the H3k9me3 (Figure 1.8B) (Bostick et al., 2007; X. Li et al., 2008; X. Liu et al., 2013; Rothbart et al., 2012; Sharif et al., 2007; Zuo et al., 2012).

The last step of the imprinting cycle is the erasure of allelic methylation, which occurs together with other epigenetic marks when the PGC enter into the germinal ridges. The depletion of DNMT proteins and the action of TET1 and TET2 eliminates imprint methylation together with the erasure of genome-wide DNA methylation patterns (explained in the previous subsection).

**A. Imprint methylation adquisition**




**B. Protection against methylation**

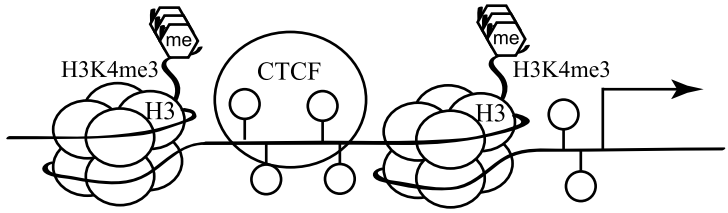



Figure 1.7: **Histone modifications and binding factors** in (**A**) imprint methylation adcquisition in both gametes and (**B**) protection against methylation in the unmethylated allele. Black lollipops represent methylated CpG whereas white circles unmethylated CpG.

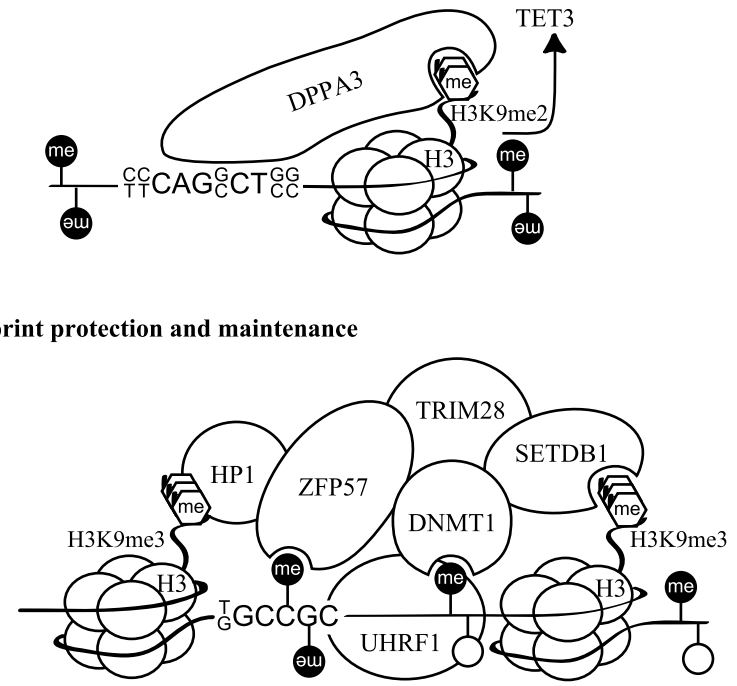



Figure 1.8: **Histone modifications and binding factors protecting methylation at imprinted DMRs during preimplantation reprogramming.** (**A**) Imprint protection from TET3 demethylation by DPPA3 protein and (**B**) imprint maintenance and protection from passive demethylation by the combined action of DNMT1 and the ZFP57-TRIM28 repressive complex which recognises the $TGCC^{met}GC$ hexanucleotide. Black lollipops represent methylated CpGs whereas white circles unmethylated CpGs.

## 1.2 Maternal-to-zygotic transition

In mammals, maternal-effect genes, such as *Zfp57* in mouse, play key roles in maintenance and protection of imprinted DMRs before the activation of embryonic genome. But it is also demonstrated that these factors and RNAs accumulated during oogenesis play essential roles in other biological processes during early embryonic development such as DNA replication, cell division, epigenetic reprogramming and even zygote genome activation (ZGA) itself (K.-H. Kim & Lee, 2014).

During oogenesis, there is a continual cross-talk between the oocyte nucleus and cytoplasm. All the proteins and RNAs accumulated in this process are essential not only for the zygote after fertilisation but also for all different oocyte maturation steps (Figure 1.9). During human folliculogenesis, the oocytes' growth is linked with the proliferation and differentiation of surrounding granulosa cells. Before birth, PGCs enter meiosis and remain arrested at the diplotene stage of meiotic prophase I, forming the primordial follicle. Each cycle from puberty to menopause, the primordial follicle enter to growth phase in response to the luteinizing hormone, until it reaches the antral stage, where the oocyte completes meiosis I and ovulation occurs. At this point the oocyte is arrested in the metaphase II, which only ends if it is fertilised by a sperm, unleashing a $Ca^{2+}$-induced cyclin degradation (Amleh & Dean, 2002). The acquisition of DNA methylation imprints occurs in parallel to oocyte growth and maturation, and it is fully established in the ovulated metaphase II oocytes – data from mice (Gahurova et al., 2017) –.

In 2008 the subcortical maternal complex (SCMC) was identified in mice (L. Li, Baibakov, & Dean, 2008). Assembled during oocyte growth, several maternal-effect proteins have been identified to be part of the SCMC, including OOEP, NLRP5, TLE6 and KHDC3L. Six years later, members of the same laboratory showed that the mice SCMC controls the zygotic symmetric division by regulating F-actin dynamics (X.-J. Yu et al., 2014). Subsequently, they identify human homologous genes to be expressed in the oocytes of foetal ovaries. The human OOEP, NLRP5, TLE6 and KHDC3L physically interact with each other and co-localize in the subcortex of human oocyte and early embryos (K. Zhu et al., 2014). As it will be shown throughout this thesis, NLRP family members and KHDC3L play a key role in germline methylation imprints, suggesting a possible link between SCMC and genomic imprinting regulation.

During the firsts cell cleavages of the zygote, there is a transition between exhaustion of maternal RNA and proteins (including the SCMC) and the transcription of the embryonic genome (see the lower part of Figure 1.9) (Schier, 2007). Genes expressed once ZGA occurs are essential for the early development and they are gradually activated during the different demanding developmental stages. The ZGA are coordinated with chromatin reorganisation and starts with a minor wave during the first cell cycle in mice – followed

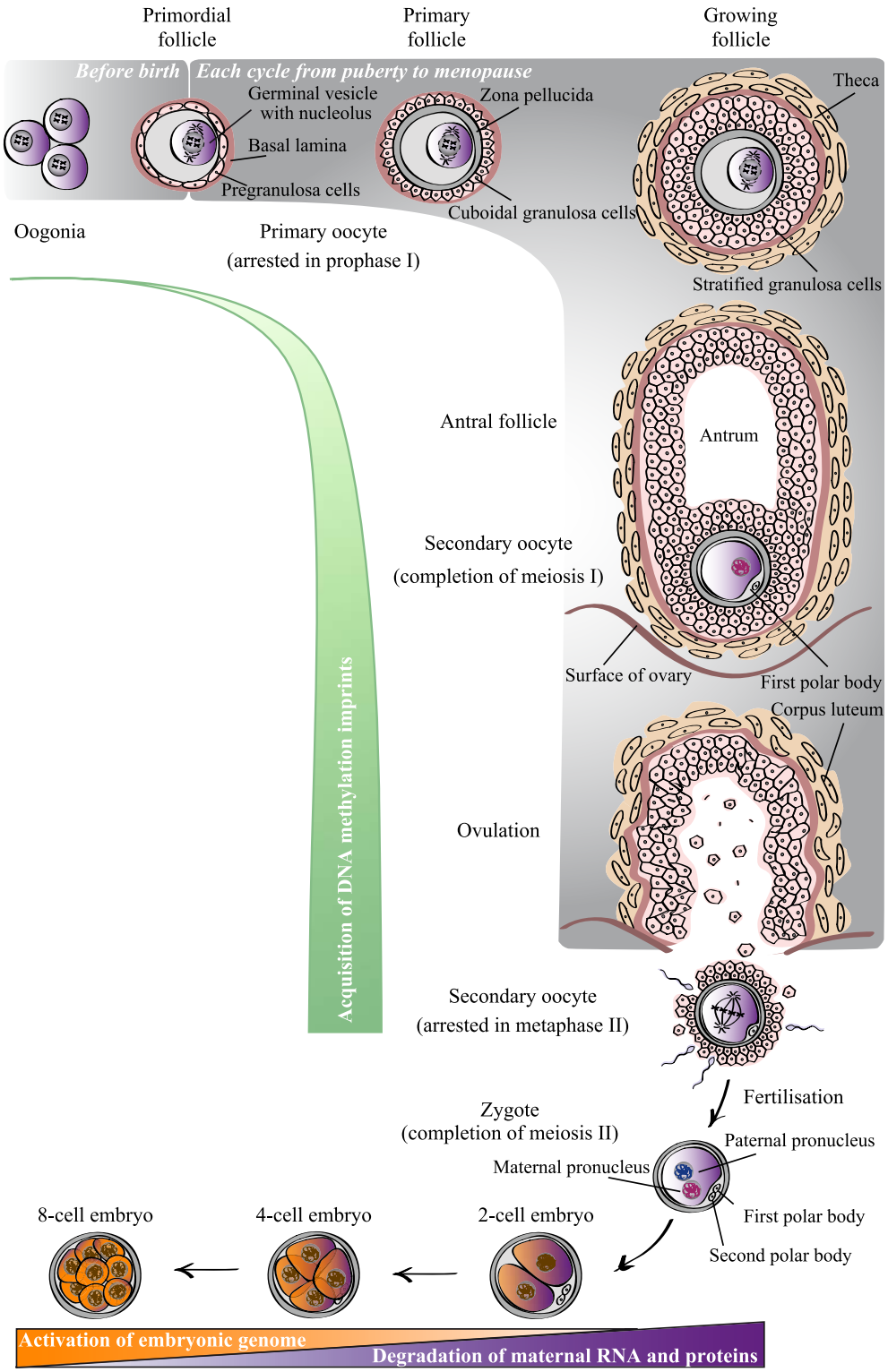



Figure 1.9: **Overview of oocyte maturation and early embryo development.** Schematic representation of both the oocyte and the follicle maturation in parallel (Monk, Sanchez-Delgado, & Fisher, 2017).

by a second wave of transcription during cycle 2 – and at 4-8 cell stage in human (Braude, Bolton, & Moore, 1988; Dobson et al., 2004).

## 1.3 Discovering genomic imprinting

Pronuclear transfer experiments in mice during the early 1980s showed that maternal and paternal genetic contributions were non-equivalent and that both pronuclei were indispensable for normal development in mammals (Figure 1.10) (McGrath & Solter, 1984; M. Surani, Barton, & Norris, 1984). The inheritance of two male pronuclei results in relatively good trophoblast development without an embryo, whereas the opposite, two maternal pronuclei result in embryonic tissues without a placenta. Thus, this revealed that somewhere in the mouse genome, exists a gene expression regulation dependent on parental origin and the concept of *genomic imprinting* as we know it today was conceived.

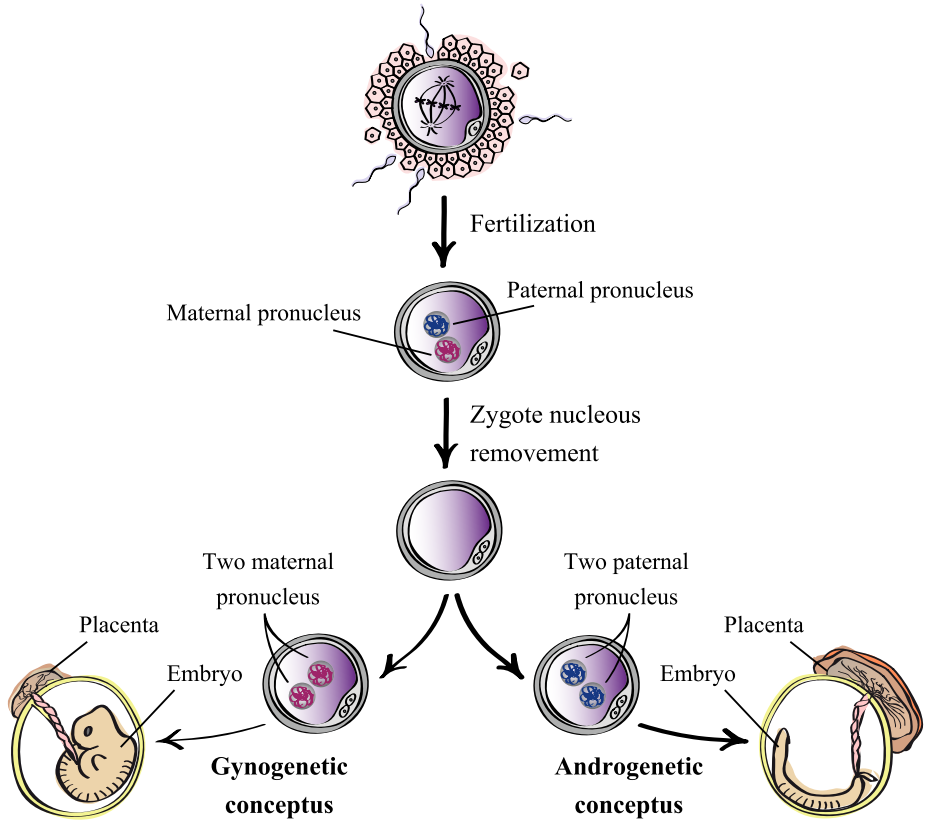



Figure 1.10: **Pronuclear transfer experiments in mice.** After fertilisation, zygote nuclei were removed and re-injected to generate embryos with two maternal pronuclei (*gynogenetic conceptus*, bottom-left), or two paternal pronuclei (*androgenetic conceptus*, bottom-right).

These two extreme developmental abnormalities also occur in human pregnancy: the parthenogenetic ovarian teratoma and androgenetically derived hydatidiform mole (HM) (Linder, McCaw, & Hecht, 1975; Wake, Takagi, & Sasaki, 1978). Because of that, part of the scientific community focused their effort to identify these potential *imprinted regions* using mouse model, as it was believed that imprinted genes would be evolutionary conserved.

### 1.3.1 Searching for imprinted loci

The search for chromosomes harbouring imprinted genes was initially based on reciprocal phenotype analysis of non-complementation using Robertsonian translocations in mice (B. M. Cattanach & Kirk, 1985). Later, the specific cytogenetic localisations responsible for the imprinting phenotypes were defined using reciprocal subchromosomal translocations (B. Cattanach & Beechey, 1997). However, the first imprinted genes with classic parental-origin-specific regulation were initially identified by accident, rather than through purpose-designed genetic screens. Thereby, the first pieces of evidence of the imprinting existence came from a report in 1974, when it was described a deletion at the *Tme* locus, which was lethal on maternal transmission (D. Johnson, 1974). In 1991, by studying the genes within the *Tme* locus, the first imprinted gene was described: the maternally expressed *Igf2r* transcript (Barlow, Stöger, Herrmann, Saito, & Schweifer, 1991). At the same time, two neighbouring genes were also reported to be imprinted: the paternally expressed *Igf2* and the maternally expressed gene *H19* (Bartolomei, Zemel, & Tilghman, 1991; DeChiara, Robertson, & Efstratiadis, 1991).

As a consequence of the continuous development of new technologies and techniques, more and more imprinted genes and regions have been identified. As a result of all this research, there is an updated online list of imprinted loci and parent-of-origin effects in mice, together with the localisation of these regions in chromosome maps: http://www.mousebook.org/mousebook-catalogs/imprinting-resource.

Since the wide application of next-generation sequencing technologies, the vast majority of mouse experiments designed to find new imprinted genes, especially those with a tissue-specific expression profile, have been carried out using reciprocal crosses between different mice strains. This approach allows for the discrimination of parental alleles by strain-specific polymorphisms.

#### 1.3.1.1 Hybrids and genome-wide approaches

The use of different mouse strains reveals essential genetic information for imprinting studies. Although imprinted genes are equally regulated in different mouse strains,

different polymorphisms that characterise each of these strains gives the unique parental information needed to distinguish if a transcript is expressed from the paternally or the maternally inherited chromosome. In 1991, the maternal expression of *H19* was described using different mouse subspecies, including the *Mus musculus* CAST/EiJ and C57BL/6J strains (Bartolomei et al., 1991). Seventeen years later, the first RNA-seq datasets using reciprocal crosses confirmed ~ 90 known imprinted genes and identified many novel imprinted transcripts (Babak et al., 2008). Nowadays, crosses between females and males of different strains are not only used for determining allelic expression but also for studying methylation and histone modifications (Figure 1.11).

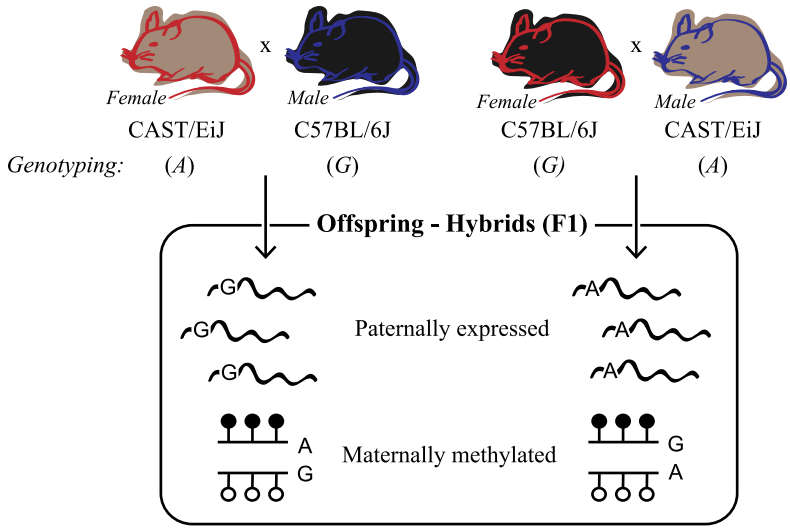



Figure 1.11: **Reciprocal crossing between mouse strains.** Grey mice represent CAST/EiJ strain (with *A* variant for a specific SNP) and black mice represent C57BL/6J strain (with *G* variant for a specific SNP).

#### 1.3.1.2 Transient and tissue-specific imprinting in mice

Early in 2000, a limited number of genes expressed in a parent-of-origin manner specifically in the mouse placenta were reported (reviewed in A. Wagschal & Feil (2006)). No DMRs were associated with these maternally expressed transcripts, but silencing was linked to specific histone marks, distinguishing the actively transcribed maternally inherited allele (H3K4me2 and H3K9ac) and the paternally silenced allele (H3K27me3 and H3K9me2) (Monk et al., 2006; Umlauf et al., 2004).

Subsequently, 28 transiently imprinted maternal germline DMRs were reported in mice twelve years later (Proudhon et al., 2012). All these regions were associated with ZFP57-binding motifs, being protected and maintained as a DMR during the epigenetic reprogramming that occurs following fertilisation. By studying the methylation profiles of

these 28 regions in hybrid mouse tissues, it was revealed that most of these transient gDMRs acquired a full DNA methylated state by the gain of methylation at the paternally derived allele at implantation, except for two regions maintained as maternal DMRs in the hypothalamus. During the limited time when these regions were monoallelically methylated, they were regulating the imprinted expression of nearby genes in a temporal and tissue-specific fashion (Proudhon et al., 2012).

#### 1.3.1.3 Human chromosomal regions associated with an imprinting disorder

In humans, the first suggestion of genomic imprinting was revealed by studying chromosomal regions associated with a non-Mendelian transmission that showed parent-of-origin inheritance patterns (Figure 1.12). In 1989, three years before the first imprinted genes were described in mouse, maternal uniparental disomy (UPD) was observed in several patients with Prader-Willi syndrome (PWS), leading the authors to speculate that the absence of paternally expressed genes could be responsible (R. D. Nicholls, Knoll, Butler, Karam, & Lalande, 1989). In the same year, similar sized and located 15q deletions were reported for PWS and Angelman syndrome (AS), the only difference being the parental origin of the deletions (J. Knoll et al., 1989). Four years later, *SNRPN* was mapped to the same interval (Glenn, Porter, Jong, Nicholls, & Driscoll, 1993). Despite this medical genetic breakthroughs, almost all imprinted genes in humans, whether they are associated with an imprinting disorder or not, were first described in mice.

#### 1.3.1.4 Human whole-genome strategies and placental imprinting tracing

Since high-throughput technologies began to be used, the search for new imprinted genes has been done from different perspectives. In humans, genome-wide experiments indicated that most ubiquitous imprinted genes had been identified, but this is not the case for the tissue-specific imprinted genes, where those confined to the placenta are proof of principle.

It was initially hypothesised that there was a general lack of placenta-specific imprinting in human since the majority of the mouse placenta-specific maternally expressed genes were biallelically expressed (Monk et al., 2006; A. Wagschal & Feil, 2006). But contrary to these initial reports, the human placenta harbours many more imprinted loci for which the mouse orthologous are not imprinted. The placenta is the organ with the most varied morphology between mammals because of the different reproductive strategies between species, which means that the best system to study human placenta development, or in this case, the placenta-specific genomic imprinting, is by directly investigating human placenta samples and not to rely on mouse models.

In 2012, seven placenta-specific imprinted genes in humans were identified by expression profiling with Affymetrix NspI chip (genotyping analysis), with only one showing conserved

imprinting in mouse placenta (Barbaux et al., 2012). In the same year, partially methylated domains (PMDs) were identified in the human placenta, revealing that PMDs cover 37% of placenta genome (Schroeder et al., 2013). However, parent-of-origin studies had not been performed for these PMDs. Subsequently, our laboratory utilised a different strategy relying on differential methylation (methyl-seq and high-density arrays) rather than the allelic expression to identify potentially new imprinted DMRs in the placenta. This revealed numerous novel DMRs of maternal methylation restricted to the placenta located far from the known imprinted clusters, potentially revealing the location of isolated imprinted regions not regulated by known ICRs (Court et al., 2014). Together with DMRs associated with *C19MC, DNMT1* and *AIM1*, which were previously reported to be exclusively present in human placenta, all human placenta-specific DMRs can be classified into two types. We classify placenta-specific DMRs depending if they are subject to *de novo* methylation in somatic tissues (Type 1, only DMRs associated with *C19MC, GPR1-AS* and *ZFAT*), or fully unmethylated in non-placenta cell types (Type 2, the remaining 13 DMRs) (Court et al., 2014; Das et al., 2013; Noguer-Dance et al., 2010).

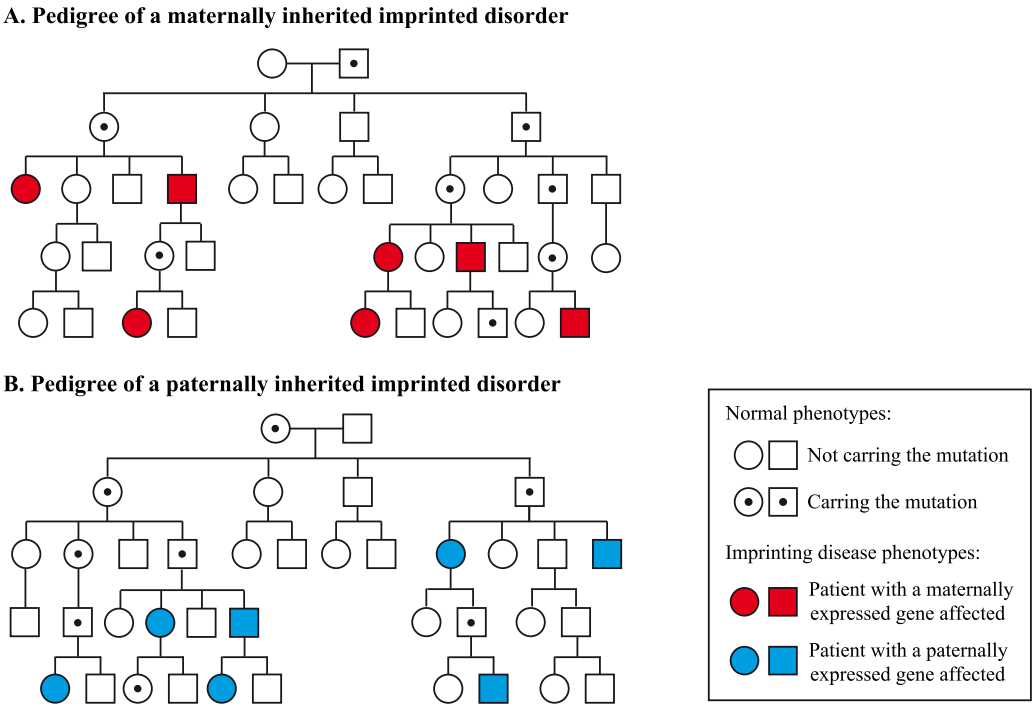



Figure 1.12: **Imprinting disorders' pedigree.** Non-mendelian inheritance for imprinting disorder affecting to (**A**) maternally expressed gene (red filled circle/square, only maternal transmission) or (**B**) paternally affected gene (red filled circle/square, only paternal transmission). Circles represent female individuals and square male individuals.

### 1.3.2 Classification of mouse and human DMRs

As described in the previous sections, there are different types of imprinted DMRs in the mouse and human genomes. In total, we can distinguish six imprinted DMR types depending on their acquisition and conservation:

i. **Germline DMRs**
   - **Ubiquitous gDMRs (primary):** Methylation acquired during gametogenesis and conserved as DMR in all somatic tissues. Most are ICR and some regulate secondary DMRs.
   - **Human placenta-specific gDMRs**:
     - **Type 1**: Methylation acquired during oogenesis. Conserved as DMR in placenta only and fully methylated in somatic tissues.
     - **Type 2:** Methylation acquired during oogenesis. Allelic methylation only present in placenta and devoid of methylation in somatic tissues.
   - **Mouse transient gDMRs:** Methylation acquired during oogenesis in a parent-of-origin manner and fully methylated following implantation.
   - **Mouse tissue-specific gDMRs:** Methylation acquired during oogenesis in a parent-of-origin manner with allelic methylation only maintained in specific tissues.

ii. **Secondary DMRs:** DMRs that acquired their methylation after fertilisation, often in a tissue-specific fashion, and reliant on a neighbouring ICR.

### 1.3.3 The causes of imprinting defects in humans

In humans, when the imprinted expression is disturbed, it often results in an imprinting disorder. These diseases can be caused by mutations of an imprinted transcript (e.g. *UBE3A* in Angelman syndrome) or by UPDs, deletions or epimutations that in general affects an entire imprinted cluster, even though the resulting disease is often associated with a single gene (D. J. Mackay & Temple, 2017; Sadikovic et al., 2014). It has been reported that imprinting disturbances can also be caused by mutations in a *trans*-acting transcription factor, resulting in a secondary epimutation (Sanchez-Delgado et al., 2016a).

When an epimutation causes an imprinting defect, this could result in either the reactivation of the silent allele or the silencing of the active allele (Figure 1.13). Once the epimutation has occurred, the aberrant methylation pattern will be maintained in all cells through embryo development. However, sometimes this abnormal methylation state occurs during the cleavage stages of preimplantation development and the consequence will be a mosaic for both non-affected and affected cells (Azzi et al., 2014). The origins of these primary epimutations are currently unknown, but in the last section of the introduction, we will discuss the possible environmental exposures that may influence epigenetic profiles.

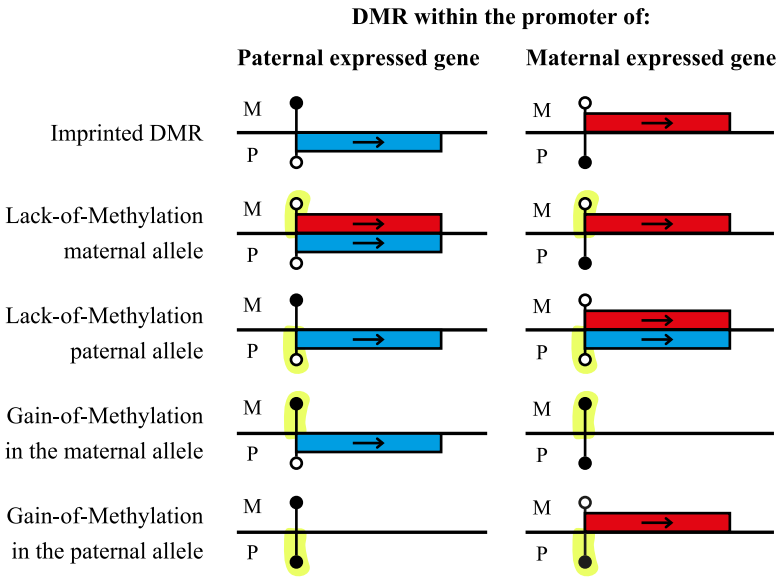



Figure 1.13: **Schematic representation of how epimutations at imprinted DMRs can affect genes.** In this general representation, imprinted genes are affected by either a lack-of-methylation resulting in biallelic expression or a gain-of-methylation so that the gene is not expressed. Black lollipops depict methylated CpG regions whereas white circles unmethylated CpG regions. Red filed rectangle represents the maternally expressed allele; blue filled rectangle paternally expressed allele. This is a general example, regulation of some imprinted genes are more complex and not directly affected by DNA methylation at its promoter.

#### 1.3.3.1 Imprinting syndromes

To date, 13 imprinting disorders have been described of which, eight are commonly studied (Table 1.2, ***Adapted from:*** Sanchez-Delgado, Martin-Trujillo, & Monk (2015)). Phenotypic analysis of the various imprinting syndromes has generally revealed that these genes are involved mainly in foetal growth and brain functions (since they shown a range of intellectual disabilities and behaviour abnormalities), as well as metabolic and hormonal alterations. For example, Beckwith-Wiedemann (BWS) and Silver Russell (SRS) syndromes are reciprocal phenotypes related with intrauterine growth, sharing a commonly affected chromosome region (11p15.5) in which the *IGF2* and *H19* genes are located (Eggermann, Eggermann, & Schönherr, 2008). A subset of BWS patients is caused by the over-expression of the *IGF2* growth factor, commonly related to paternal UPD11 and associated with the overgrowth phenotype of these patients (W. N. Cooper et al., 2005). On the contrary, SRS is characterised by severe intra-uterine growth restriction and is associated with epimutations that result in decreased *IGF2* expression (and in rare familial cases point mutations) (Gicquel et al., 2005; D. Liu, Wang, Yang, & Liu, 2017). Interestingly, both syndromes are also associated with defects of a neighbouring imprinting cluster on chromosome 11, the *KCNQ1OT1-CDKN1C* region (Demars et al., 2011). The majority of BWS cases (~50%) are caused by lack-of-methylation (LOM) at the *Kv*DMR1 (imprinted DMR in *KCNQ1OT1* promoter, see Figure 1.3) that results in the biallelic expression of the long ncRNA *KCNQ1OT1* and the concomitant silencing of the maternal expressed cell-cycle checkpoint gene *CDKN1C* (Diaz-Meyer et al., 2003). Point mutations in *CDKN1C* are associated with both BWS and SRS, with the latter caused by mutations in the PCNA domain that increases protein stability (F. Brioude et al., 2013). Higher expression levels of *CDKN1C* have also been reported in non-syndromic forms of intrauterine growth restriction (IUGR), suggesting that this gene is a key regulator of foetal growth (López-Abad, Iglesias-Platas, & Monk, 2016).

Table 1.2: **Different molecular aetiology of imprinted disorders and their frequencies**.

| Disorder | Chr. region | Molecular aberration | Estimated frequencies | Clinical features |
|---|---|---|---|---|
| Transient neonatal diabetes mellitus (TNDM) | 6q24 | pUPD6[2] | 40% | Transient diabetes, IUGR[3], hyperglycemia without ketoacidosis, macroglossia, |
| | | dup(6q)pat[4] | 30% | |
| | | *PLAGL1* LOM | 30% | |

[2] *pUPD#*: Paternal uniparental disomy of chromosome "#"
[3] *IUGR*: Intrauterine growth restriction
[4] *dup(#p-q)pat*: Paternal duplication of chromosome region "#p-q"

| Disorder | Chr. region | Molecular aberration | Estimated frequencies | Clinical features |
|---|---|---|---|---|
| (OMIM 601410) | | *ZFP57* mut.[5] | 50% | omphalocele. |
| Silver-Russell | 7 | mUPD7 | 7-10% | Severe pre- and postnatal |
| syndrome (SRS) | 11p15.5 | mUPD11 | Rare cases | growth restriction, |
| (OMIM 180860) | | dup(11p15)mat | Rare cases | macrocephaly, |
| | | H19 LOM | ~40% | hemihypotrophy, triangular |
| | | *CDKN1C* mut. | Rare families | face, feeding difficulties. |
| | | Maternal uniparental diploidy | n=1 | |
| Beckwith -Wiedemann syndrome (BWS) (OMIM 130650) | 11p15.5 | pUPD11 | 20% | Pre- and postnatal |
| | | Paternal uniparental diploidy | ~ 10% | overgrowth, organomegaly, macroglossia, omphalocele, neonatal hypoglycemia, |
| | | *H19* GOM | 5% | hemihypertrophy, increased |
| | | *Kv*DMR1 LOM | 50% | tumour risk. |
| | | *CDKN1C* mut. | 5% | |
| | | *NLRP2* mut. | One family | |
| Temple syndrome | 14q32 | mUPD14 | 80% | Pre- and postnatal growth |
| (OMIM 616222) | | del14q32 | 9% | restriction, hypotonia, |
| | | *MEG3* LOM | 11% | feeding difficulties in infancy, truncal obesity, scoliosis, precocious puberty. |

[5] *mut*: mutations.

| Disorder | Chr. region | Molecular aberration | Estimated frequencies | Clinical features |
|---|---|---|---|---|
| Kagami-Ogata syndrome (OMIM 608149) | 14q32 | pUPD14<br>del14q32<br>*MEG3* GOM | 65%<br>20%<br>15% | IUGR, polyhydramnion, abdominal and thoracal wall defects, bell-shaped thorax, coat-hanger ribs. |
| Angelman syndrome (AS)<br><br>(OMIM 105830) | 15q11-13 | pUPD15<br>del15q11q13<br><br>*SNRPN* LOM<br>*UBE3A* mut. | 1-2%<br>75%<br><br>~3%<br>5-10% | Severe motor and mental retardation, microcephaly, no speech, unmotivated laughing, ataxia, seizures. |
| Prader-Willi syndrome (PWS) (OMIM 176270) | 15q11-13 | mUPD15<br>del15q 11q13<br>*SNRPN* GOM | 25-30%<br>70-75%<br>~1% | Postnatal growth restriction, moderate mental retardation, neonatal hypotonia, hypogenitalism, hypopigmentation, hyperphagia. |
| Pseudohypopara<br><br>-thyroidism (PHP) (OMIM 603233) | 20q13 | pUPD20<br><br>*GNAS* methylation defects | Unknown | Resistance to parathyroid hormone and other hormones, albright hereditary osteodystrophy, subcutaneous ossifications, feeding behaviour anomalies, abnormal growth. |

#### 1.3.3.2 Multi-locus imprinting disturbances (MLID)

Despite the majority of imprinting syndromes are being caused by disturbances in one or contiguous imprinted transcripts within the same domain, recent cases have been described in which aberrant methylation is observed at many imprinted loci. It was originally reported that half of transient neonatal diabetes mellitus (TNDM) patients cases with hypomethylation at *PLAGL1* DMR have additional hypomethylation at other maternally methylated DMRs, which was termed *maternal hypomethylation syndrome* (D. Mackay et al., 2006). Subsequently, recessive mutations in the *ZFP57* gene were identified in TNDM pedigrees with multiple hypomethylation at maternally-derived methylated DMRs (D. J. Mackay et al., 2008). These findings represent the first demonstration of multi-locus imprinting disturbances (MLID) due to functional loss of a *trans*-acting protein factor protecting imprinted DMRs. Through screening for imprinting diseases other than TNDM, MLID cases have been identified in all patient cohorts with varying frequency (Figure 1.14) (reviewed in Sanchez-Delgado et al. (2016a)). However, the causal genes in the majority of MLID patients have not been identified, suggesting that not all *trans*-acting imprint factors have been identified (Caliebe et al., 2014).

The most severe case of MLID, which affects all maternally methylated imprinted DMRs, is the recurrent HM (RHM). As will be discussed in the next subsection, this type of mole which has a biparental contribution, are mostly associated with maternal-effect mutations in *NLRP7*, suggesting that this gene has an important role in the imprint regulation and oocyte methylation. Although its role in the epigenetic regulation and imprinting has not yet been elucidated, *NLRP7* is not the only NLRP family member associated with MLID cases. In 2009 a case of BWS associated to a maternal-effect mutation in *NLRP2* was described, and six years later, *NLRP5* was shown to be mutated in different MLID cases and reproductive wastage (Docherty et al., 2015; Meyer et al., 2009).

#### 1.3.3.3 Recurrent hydatidiform moles (RHM)

As explained earlier in the introduction, pronuclear transfer experiments in mice revealed that maternal and paternal genetic contributions were non-equivalent and both were indispensable for normal development in mammals (Figure 1.10) (McGrath & Solter, 1984; M. Surani et al., 1984). They show that mice inheriting two male pronuclei results in a trophoblastic mass with abnormal embryonic lineages and as was initially pointed out, this was similar to complete androgenetic HMs in humans (Figure 1.15) (Wake et al., 1978).

The HM is an abnormal human pregnancy with excessive trophoblast proliferation (hyperplasia of the placental villi) and abnormal or absent embryonic development (N. M. P. Nguyen & Slim, 2014). The molar tissue is classified according to their pathology, where it is possible to distinguish complete HM (CHM) with gross hyperplasia of villous and

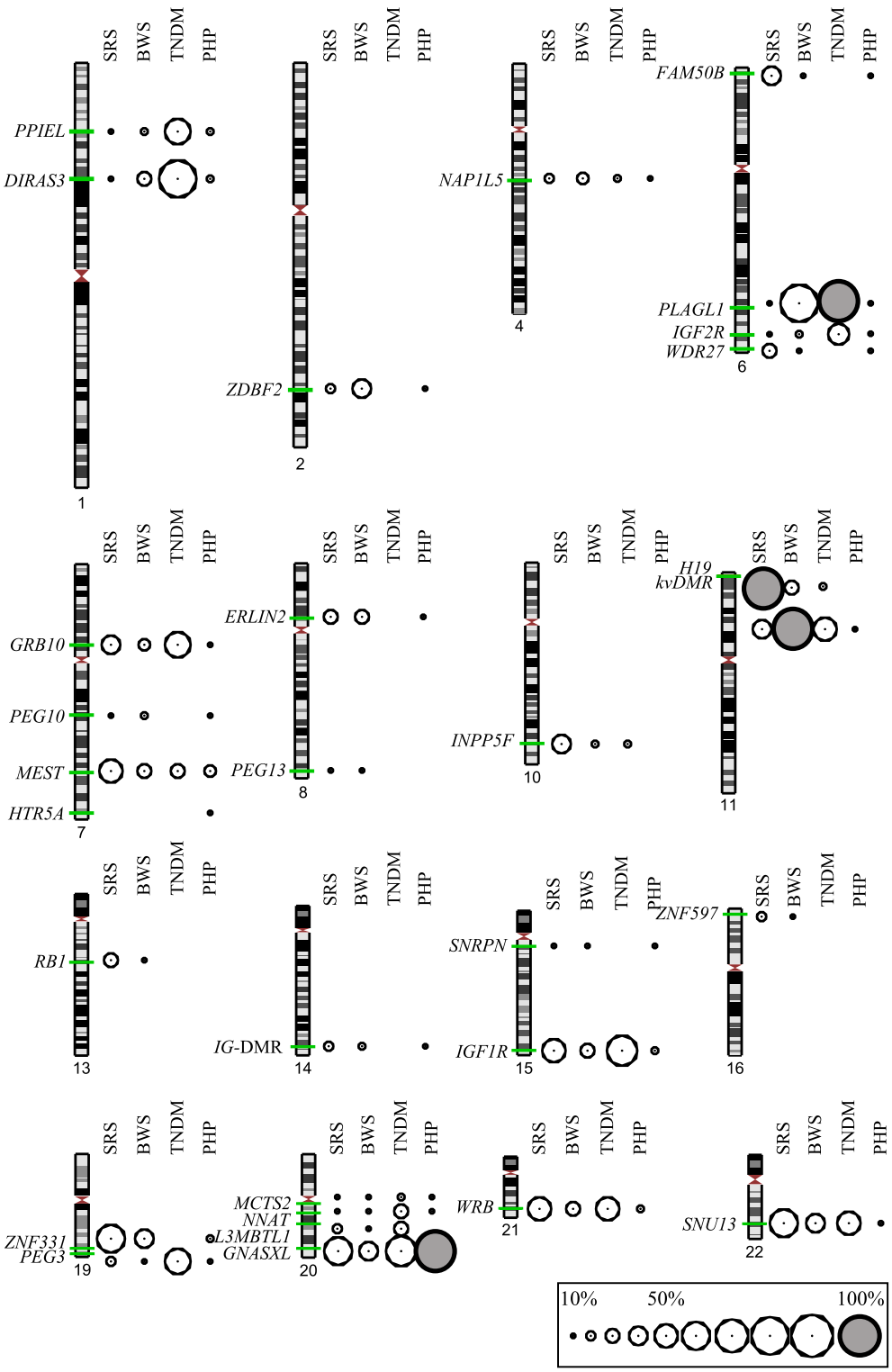



Figure 1.14: **Ideogram showing the positions of known imprinted domains and its frequency when hypomethylated in imprinting disorders with multi-locus disturbance.** Black circle indicate loci with LOM, with the grey circle indicating the primary disease-causing region. **Adapted from:** Sanchez-Delgado et al. (2016a).

no foetal development, and partial HM with more focal trophoblastic hyperplasia and greater but abnormal foetal development. Alternatively, HM can be categorised according to their genetic origin, whether there is a biparental contribution to their genome or they are androgenetic derived moles, formed by the fusion of two paternal pronuclei (Figure 1.15) (P. H. Dixon et al., 2012; N. M. P. Nguyen & Slim, 2014). Furthermore, HM may be sporadic or recurrent. While sporadic and CHMs are mostly diploid with an androgenetic origin, the sporadic and partial moles are triploid dispermic with two different copies of the paternal genome and one of the maternal genome. Moreover, RHMs are a rare conception in which molar tissue is diploid with a biparental contribution to their genome caused by maternal-recessive mutations, mostly in *NLRP7* and in a few cases in *KHDC3L* (Murdoch et al., 2006; N. M. P. Nguyen & Slim, 2014). Nevertheless, mutations in *NLPR7* and *KHDC3L* are not associated with androgenetic complete moles (P. H. Dixon et al., 2012; Mahadevan et al., 2013a).

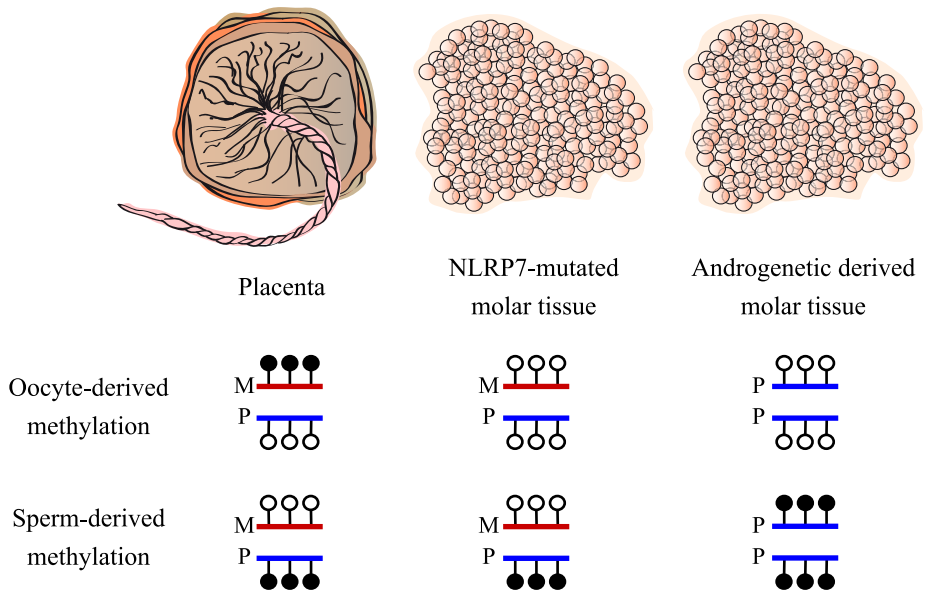



Figure 1.15: **Methylated profile of imprinted DMRs in placenta and hydatidiform moles.** Black lollipops represent methylated CpG regions whereas white circles unmethylated CpG regions. In normal placental tissue oocyte-derived methylation is maintained on the maternal allele (**M**) at imprinted DMRs, while reciprocally sperm-derived methylation is observed exclusively on the paternal allele (**P**). In both androgenetic and *NLRP7*-associated molar tissues, there is a complete LOM at the majority of maternal DMRs, but paternal DMRs can be differentiated. Methylation will appear hypermethylated in androgenetic moles due to the two copies of the paternal genome and imprinted in *NLRP7*-associated moles as a result of biparental fertilisation. **Adapted from:** Monk, Sanchez-Delgado & Fisher (2017)).

Before the identification of *NLRP7* maternal-effect mutations in 2006, biparental HMs were not associated with specific maternal mutation but were thought to result from

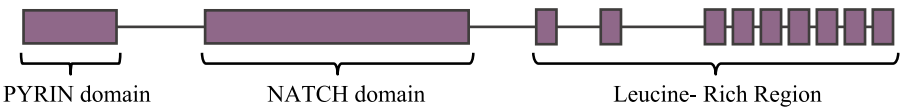



Figure 1.16: **Schematic representation of NLRP7 protein.** Schematic representation of the known protein domains NACHT, LRR and PYD domains of NLRP7.

an inherited global imprinting defect in the human female germ line due to a recessive maternal-effect mutation which disturbs the methylation of maternal imprinted DMRs (Judson, Hayward, Sheridan, & Bonthron, 2002; Murdoch et al., 2006). However, NLRP7 is a member of the CATERPILLER protein family involved in intracellular regulation of bacterial-induced inflammation, with some domain related to apoptosis, leucine-rich repeats and a protein-protein interaction domain conserved in all family members (Figure 1.16). The NLRP7 protein does not have a DNA binding motif, and to date, without a known role in imprint acquisition nor maintenance (Murdoch et al., 2006).

## 1.4 The origins of genomic imprinting

The imprinting phenomenon is not exclusive from mammals. In fact, the firsts evidence of different parental-origin inheritance was observed in insects and plants. Specifically, the term *imprint* was firstly used 24 years before Dr McGrath, Dr Surani and colleges first described the non-equivalence of parental genomes by pronuclear transfer experiments in mice. It was used to describe how the paternally-derived chromosomes (X chromosome and/or autosomal chromosomes) of some insects are selectively deleted or becomes heterochromatin for doses compensation and/or sex determination (S. W. Brown & Nur, 1964; Crouse, 1960).

Moreover, in 1970 it was described in the maize endosperm plant, defining the differential expression of specific genes depending on the parental origin of each allele (Kermicle, 1970). But once it was described the non-equivalence of both sets (maternally and paternally inherited) of autosomal chromosomes in mice, evolutionary biologists started to study genomic imprinting as a consequence of natural selection and its conservation in different species (McGrath & Solter, 1984; M. Surani et al., 1984)

### 1.4.1 X-inactivation, tandem repeats and retrotransposon silencing

In therian mammals, one of the two X chromosomes is inactivated in the female fetus to accomplish the dosage compensation, resulting in a similar dosage with a male embryo that carries a single active X (Lyon hypothesis of dosage compensation mechanisms (Lyon,

1961, 1962)). In most mammals, the determination of sex is given by the presence of a particular gene in the male-specific Y chromosome, which unlike the X chromosome present in both genders, contains only a few genes. It is believed that in mammals, X chromosome inactivation and genomic imprinting have coevolved with placentation (after the Prototheria-Theria divergence) since in monotremes, the egg-laying mammals, neither sex chromosomes inactivation nor genomic imprinting have been described (Figure 1.17) (Grützner et al., 2004; Killian et al., 2001). In marsupial (metatherian mammals) females, paternally derived X chromosome is selectively silenced in all tissues (Richardson, Czuppon, & Sharman, 1971; Sharman, 1971). On the other hand, in eutherian females, X chromosome inactivation occurs randomly in its somatic tissue (Lyon, 1972). Nevertheless, in the extra-embryonic lineages of some eutherian mammals, like mice or cows, the paternally derived X chromosome is selectively inactivated (Takagi & Sasaki, 1975; F. Xue et al., 2002).

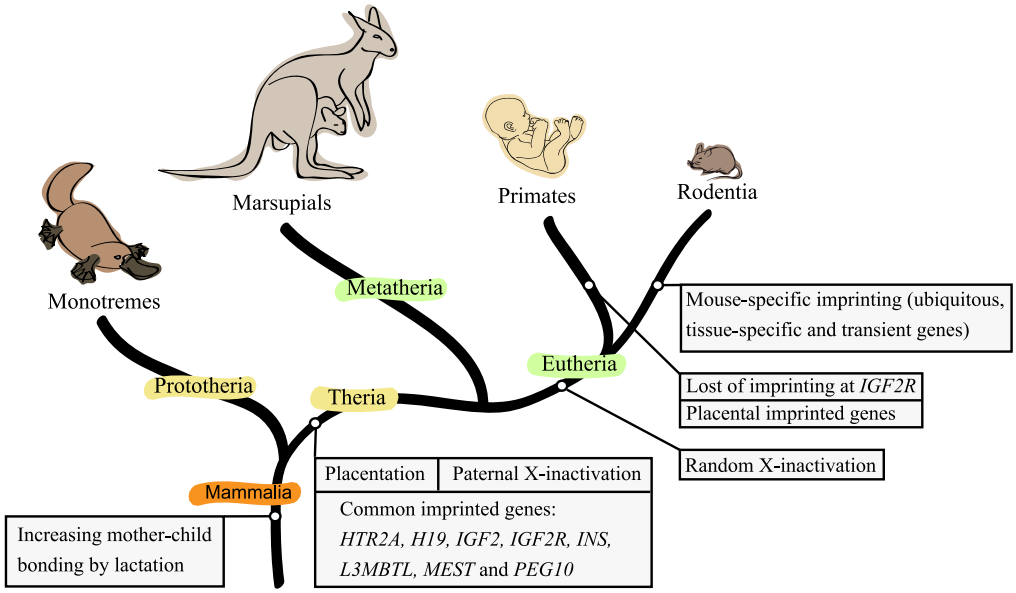



Figure 1.17: **Phylogenetic tree with the two living mammalian subclasses** Prototheria (no evidence of X-inactivation nor genomic imprinting in any of its species) and Theria. Both epigenetically regulated phenomena (X-inactivation and genomic imprinting) had been reported in Theria clades, being more developed and species-specific in Eutheria.

It seems that the dosage compensation by X-inactivation in mammals starts by acquiring a differential *imprint* in gametes to facilitate the silencing of the paternally derived X chromosome in female embryos. However, if the sole purpose of this X-inactivation is dose compensation and there is no other evolutionary advantage for the inactivation of the paternal X chromosome, the benefit of the random inactivation would exceed the higher evolutionary cost of this parent-of-origin strategy. If one of the two inherited X chromosomes of female zygote contains a recessive mutation and in all the cells of

the organism, the same X chromosome is inactivated, the chances to inactivate the non-mutated chromosome are 50%. In contrast, if inactivation of the X chromosome occurs randomly during the embryo development, the embryo will be a mosaic, with some cells transcribing the non-mutated X chromosome and other the mutated one, but the possible deleterious effect of the mutation will be diluted in this mosaic organism (T. Moore & Haig, 1991). However, despite having epigenetic features in common, the non-coding RNA associated with eutherian X-inactivation is not evolutionarily related to the one associated with metatherian X-inactivation (C. J. Brown et al., 1991; Grant et al., 2012). It seems that similar evolutionary pressures have been involved in the acquisition of both X-inactivation mechanisms, acting when both therian lineages had already separated. The fact that X chromosome carries some placenta-growth suppressors, e.g. the *ESX1* gene, it could have favoured the silencing of the paternally inherited copy of X chromosome in therian common ancestors and its conservation in marsupial and the placenta of many eutherians (Haig & Westoby, 1989; Reik & Lewis, 2005).

Therefore, the silencing mechanism of X chromosome in female mammals is reminiscent to imprinted paternally silencing observed at the distal chromosome 7 domain *Kcnq1ot1-Nap1l4* in mice (i.e. through the action of a paternally expressed non-coding RNA, which in the case of X-inactivation is *Xist* and *Kcnq1ot1* for chromosome 7). Another common feature is the existence of a gene cluster regulated by one specific region, ICR in imprinted clusters, and the X inactivation centre (XIC) associated with *Xist* and the antisense transcript *Tsix* (J. T. Lee, 2000; J. Lee, Davidow, & Warshawsky, 1999; Penny, Kay, Sheardown, Rastan, & Brockdorff, 1996; Sleutels, Zwart, & Barlow, 2002; H. Zhang et al., 2014). The epigenetic mechanisms controlling both X chromosome inactivation and genomic imprinting seems to have evolved under similar selective pressures, among other reasons, because both of them are related mainly with histone modification in marsupials and with histone modification and DNA methylation in eutherian mammals (Reik & Lewis, 2005). By mice experiments, it is shown that in absence of DNA methylation, X-inactivation and imprinting on distal chromosome 7 is unstable in somatic tissue but on the contrary, both regions are maintained and stable with the only regulation of histone modifications in the mouse placenta (Lewis et al., 2004; Sado et al., 2000). This observation suggests that the placenta has a key role in the evolution of both epigenetic mechanisms, and its epigenetic regulation is slightly different than in somatic tissues.

Moreover, both silencing phenomenon are associated with increased presence of tandem repeats and retrotransposons (Hutter, Helms, & Paulsen, 2006; Lyon, 1998; S. Suzuki et al., 2007). In fact, the genesis of some non-coding RNAs associated with X chromosome inactivation and some imprinting clusters have been directly related to miRNAs and other retroposed insertions early in mammalian evolution (Hore, Rapkins, & Graves, 2007). The fact that X chromosome is particularly rich in repetitive elements of LINE type and the high presence of tandem short repeats in *Xist* are thought to be involved in

its heterochromatinisation (Lyon, 1998; Migeon, 2016). Besides, it is reported that the CpG islands of eutherian imprinted genes contain tandem repeats more frequently than non-imprinted genes (Hutter et al., 2006). It remains possible that imprinting was initially the consequence of the same silencing mechanisms associated with repeat elements and retrotransposons. Evidence of it could be that protection of allelic methylation during preimplantation reprogramming utilises ZFP57-KAP1, a major repressor complex involved in the silencing of transposable elements (Ecco et al., 2016; Imbeault, Helleboid, & Trono, 2017). Furthermore, comparative sequences analysis in mammals has revealed that the paternally expressed Ty3/Gypsy family retrotransposon-derived gene *Peg10* is essential for mouse embryonic and placental development (Ono et al., 2001, 2006; S. Suzuki et al., 2007). This retrotransposon-derived gene has an orthologue gene in the marsupial tammar wallaby, but not in the monotreme platypus, and is the first evidence of a DMR associated to genomic imprinting in marsupials (where all imprinted genes founded were only associated with histone modification and not imprinted DMRs) (S. Suzuki et al., 2007). This finding suggests that the DNA methylation imprint could be linked to the silencing of retrotransposons, with the lethality of *Peg10*-knockout mouse revealing the acquired importance of this transcript in murine placentation (Ono et al., 2006).

### 1.4.2 Genomic imprinting and mammalian placentation

Mammalian evolution is marked by strong mother-child links, first, by the common characteristic of all three living mammals (monotremes, marsupials/metatherians and eutherians) which is lactation, and in the therian lineage, by the direct maternal-foetal physiological exchange during embryos development. This latter process has allowed for the development of a unique tissue: the placenta (Figure 1.18). This organ is essential for the intrauterine development of the foetus. The placenta is also directly responsible for bringing maternal and foetal blood supplies into contact, facilitating nutrients, gases and waste products exchange and determining resources allocation (Frost & Moore, 2010; Guernsey, Chuong, Cornelis, Renfree, & Baker, 2017).

As mentioned previously, it is believed that genomic imprinting in mammals, together with X chromosome inactivation, has coevolved with placentation, leading to the suggestion that the placenta is the main target organ driving the evolution of these two epigenetic phenomena. Although the biological functions of imprinted genes are varied and they often play a role in postnatal processes such as feeding adaptation, social behaviour or metabolism, the most significant function is observed during embryonic development, where most imprinted genes are expressed (Fountain, Tao, Chen, Yin, & Schaaf, 2017; A. Fowden, Coan, Angiolini, Burton, & Constancia, 2011; Stringer, Pask, Shaw, & Renfree, 2014). In eutherian mammals, imprinted genes are known to play essential roles in the growth and development of both the foetus and the placenta, as well as in the cell lineages

progresses such as neuronal differentiation (Jürgens et al., 2009; McGrath & Solter, 1984; M. Surani et al., 1984). Nevertheless, imprinted genes seem to have a more marked role in mother-child interaction during lactation in marsupial since the embryonic development *in utero* of these species is much shorter than in eutherians and the period of lactation is higher (Stringer et al., 2014).

Although there are some common biological and physiological characteristics in all species, the placenta is the organ that has evolved most differently in therian mammals. Maybe due to this diversity, genomic imprinting in placenta tissue is the most differentially regulated between eutherian species. However, the continuous discovery of new imprinted gene and DMRs in the human placenta, and their lack-of-conservation in other mammalian species including mice, could be due to the continuous coevolution of this epigenetic phenomenon with placenta tissue.

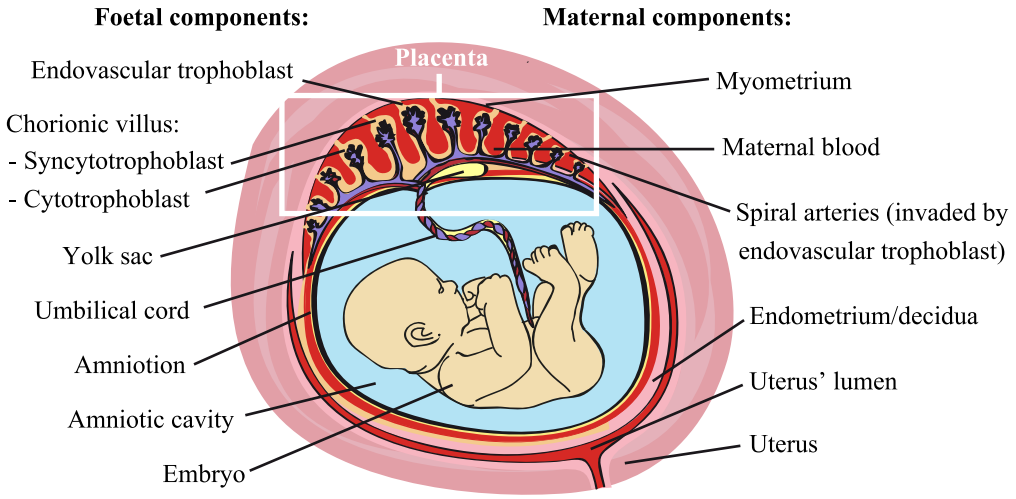



Figure 1.18: **Schematic representation of 12-13 weeks human gestation.** Main foetal (left) and maternal (right) placenta components. Human placenta is characterised by extensive invasion of trophoblast-lineage cells into the maternal uterus, allowing its direct contact with the maternal blood. **Adapted from:** Tarrade et al. (2001).

### 1.4.3 Evolutionary theories of genomic imprinting

One of the main common characteristics of sexual reproducing multi-cellular organisms is the alternation of its ploidy during its life cycle. This usually results in haploidy in the germline, with diploidy in somatic tissues. Higher ploidy has a cost, but there is also an evolutionary advantage by providing other copies when one gene allele becomes defective (Orr, 1995). A priori, mammals with genomic imprinting have the disadvantage of both haploidy and diploidy, i.e. they are diploid organisms, with its associated cost, but additionally, they accumulate the disadvantage of not having a back-up if the imprinted

gene copy expressed should become defective (Otto & Gerstein, 2008). However, the above conclusion is due to the fact that loss-of-function mutations (recessive mutations) are expected to be most common, but in the case of dominant mutation, genomic imprinting would be hiding its harmful effect in 50% of cases (Hurst, 1997). Nevertheless, genomic imprinting could not require an adaptive explanation. Some evolutionary biologists suggest that imprinting is an epiphenomenon that could just be a non-adaptive by-product of other functions, for example, an intermediate phase of silencing process by DNA methylation in germ cells to protect the genome from exogenous DNA like retroviruses or transposable elements (Barlow, 1993; S. Suzuki et al., 2007; Yoder, Walsh, & Bestor, 1997). But even though it may be the origin of genomic imprinting, it is clear that subsequently, there was a positive selection to conserve the dosage in which these imprinted genes are expressed, and for making some of them essential for the correct development in different species.

The most accepted evolutionary theory of genomic imprinting existence was actually acquired through the study of angiosperm plants, also known as flowering plants, by Dr David Haig and Dr Mark Westoby in 1989: *the kinship theory* (Haig & Westoby, 1989). This theory, also known as the *genetic conflict hypotheisis*, does not explain any specific epigenetic mechanism or genetic context that favour imprinting, but propose possible biological benefit of having some genes expressed in a parent-of-origin manner. Although the authors developed this theory using the endosperm tissue of angiosperm plants, which acquires resources from the mother to nourish the embryo, they extracted the idea of the kinship theory from the pronuclear transfer and the UPDs experiments in mice. For instance, when comparing mice with the same genetic contribution as the parents, offspring with two paternal-derived copies of chromosome 11 are bigger, and mice with two maternal-derived copies of chromosome 11 are smaller (B. M. Cattanach & Kirk, 1985). Likewise, the mouse conceptus from two male pronuclei has larger placenta than gynogenetic or normal conceptus (Figure 1.10) (McGrath & Solter, 1984; M. Surani et al., 1984). As the placenta is the tissue directly involved in the nutrient transfer from the mother to the developing embryo, they proposed that the paternally derived alleles have greater activity in acquiring resources than the maternally derived ones. They suggested that natural selection favours this parent-of-origin expression because there is an asymmetry between the interests of the parents when other offspring of the mother sometimes have different fathers. In this case, the alleles inherited from the mother (which will be present in 50% of all her offspring) will tend to reduce the resource consumption of each offspring to improve the reproductive success of all siblings. On the contrary, the alleles derived from the father favour the maximum growth of their offspring since siblings may come from other males. The authors also conclude that for the existence of a natural selection that favours a parent-of-origin expression, it is necessary that the genotype of the offspring will be active while the mother provides resources. They also suggest that if genomic imprinting disables parthenogenesis in mammals, the parent-of-origin expression

will not be present in groups of vertebrates that can reproduce by an asexual method. The typical example consistent with predictions of the kinship theory are the clusters associated with *Igf2* and *Igf2r*, containing both: paternally expressed genes (PEGs) acting as enhancers of fetal growth, and maternally expressed genes (MEGs) which inhibits this growth (Haig, 2004).

But not all imprinted genes expressed in the embryo and the placenta shows the parental-origin expression that kinship theory predicts according to its function. Furthermore, during embryo development, the transcripts *in contact* are the expressed from the maternal genome and the offspring genome. The *maternal-offspring coadaptation theory* predicts the genomic imprinting expression in offspring to increase its fitness, and addresses mainly, but not exclusive to the maternally expressed genes (Wolf & Hager, 2006). This theory predicts that selection often favours the coadaptation of the maternal and offspring traits by *pleiotropy* (i.e. matching or mismatching the offspring's expressed allele with its mother's allele in pleiotropic loci, affecting both, maternal and offspring traits, to increase offspring fitness) or by *linkage disequilibrium* (i.e. allowing to coadapted alleles to be associated, having one which determine a maternal trait that affects offspring fitness, and a second locus, which determine a offspring trait affecting its own fitness). However, in any case, it is assumed that selection favours coadaptation due to the fact that offspring fitness is determined by the combination of its own genome and its mother's genotype due to their close interaction during embryos development (Wolf & Hager, 2006).

Since genomic imprinting is present in species as different as plants and mammals, and that the imprinting status of some genes varies among mammalian species, imprinting is probably being gained and lost over evolutionary time. Consequently, imprinting may have evolved due to different selective and non-selective pressures at different loci and different species. Most imprinting hypotheses are not mutually exclusive. Imprinting might have first appeared as a non-adaptive phenomenon, a by-product of the retrotransposon silencing process, but once established, it could have evolved in response to genetic conflict or coadaptation. Because genomic imprinting has coevolved with mammalian placentation and most imprinted genes are expressed in the placenta, its invasive trophoblast lineages have been an important focus for additional selective pressures that may also have influenced to the imprinting evolution in mammals. For example, *The trophoblast defence theory* implies that the maternally expressed imprinted genes (which functions as growth repressors) and maternally methylated DMRs (which are repressing growth-promoter gene), have adapted to limit the growth of parthenogenetic concepti in the ovary (Solter, 1988; Varmuza & Mann, 1994).

## 1.5 Environment effect and pregnancy-related pathologies

Various studies have shown an association between the expression pattern of some imprinted genes in the placenta (e.g. *PHLDA2*) with foetal growth and newborn weight (Apostolidou et al., 2007). In this same way, aberrant methylation or expression of different imprinted genes in the placenta have been observed in pregnancy-associated pathologies, including IUGR or pre-eclampsia (A. I. Diplas et al., 2009; Yuen, Penaherrera, Von Dadelszen, McFadden, & Robinson, 2010).

Studies in mice have shown that deletion of some paternally expressed genes (e.g. *Igf2* or *Mest*) results in IUGR, whereas deletion of maternally expressed genes (*Igf2r* or *H19*) results in foetal overgrowth (M. Lau et al., 1994; Leighton, Ingram, Eggenschwiler, Efstratiadis, & Tilghman, 1995; F.-L. Sun, Dean, Kelsey, Allen, & Reik, 1997; Z.-Q. Wang, Fung, Barlow, & Wagner, 1994). Specifically, mouse models have demonstrated that imprinted genes could control foetal growth by acting in the placenta, due to its influence to the maternal nutrients' supply; by using chimaeras with *Igf2* deficient cells in the trophectoderm, it was shown growth restriction in both, the foetus and the placenta (Gardner, Squire, Zaina, Hills, & Graham, 1999). Additionally, the deficiency of a placenta-specific *Igf2* isoform, which is only expressed in a specific part of the mouse placenta, also alters the growth of both, foetus and placenta (Constância et al., 2002; Moore et al., 1997).

However, the cause-effect relationship in human disease is currently not clear. Are these pathologies caused by a general alteration of gene expression or is the alteration specific to imprinted genes? Are the changes observed the cause of the pregnancy complications or a response mechanism to ensure the pregnancy continues? Thus, different environmental factors have been proposed to disturb genomic imprinting and/or placenta and foetal growth and development, as has the use of assisted reproductive technologies (ART).

### 1.5.1 Maternal environmental exposures

Maternal exposure during pregnancy has a direct impact on the foetus and has been reported to subtly alter methylation at some imprinted genes (reviewed in Kappil, Lambertini, & Chen (2015)). Maternal nutrition is one of the most studied environmental factors, being maternal malnutrition central to the *foetal origin of adult diseases hypothesis* (Jimenez-Chillaron, Ramon-Krauel, Ribo, & Diaz, 2016; G. Wu, Bazer, Cudd, Meininger, & Spencer, 2004). Both maternal under- and over-nutrition can alter the weight and biochemical characteristics of the placenta, the consequence of which is foetal growth alterations, resulting in IUGR and overgrowth respectively (Godfrey, Robinson, Barker,

Osmond, & Cox, 1996; Lechtig et al., 1975; Wallace, Bourke, Aitken, Milne, & Hay, 2003; G. Wu, Pond, Flynn, Ott, & Bazer, 1998). The IUGR condition seems to be more influenced by the intrauterine environment than the genetic predisposition of the embryo since oocyte donation studies have shown a higher influence of recipient mother than donor on birth weight (Brooks, Johnson, Steer, Pawson, & Abdalla, 1995). Furthermore, the maternal environment can also be involved in more severe pregnancy-associated pathologies such as pre-eclampsia, which is associated with an aberrant placenta development, decreased trophoblast invasion and maternal artery remodelling. This severe pathology is characterised by pregnancy-associated hypertension presenting after 20 weeks of gestation with clinically relevant proteinuria (C. Zhang et al., 2002). However, there are discordant data of the possible epigenetic alteration associated with the maternal nutritional alteration. While some studies show stability at imprinted methylation in offspring from women with nutritional problems, other studies show low methylation patterns at certain imprinted genes (Heijmans et al., 2008; Ivanova, Chen, Segonds-Pichon, Ozanne, & Kelsey, 2012; Kappil et al., 2015). However, because aberrant patterns of imprinted genes' methylation or expression have been reported in both IUGR and pre-eclampsia without a consistent tendency, it is necessary to continue studying the possible relationship of all these factors (Bourque, Avila, Penaherrera, Von Dadelszen, & Robinson, 2010; A. I. Diplas et al., 2009; Kulkarni, Chavan-Gautam, Mehendale, Yadav, & Joshi, 2011; McMinn et al., 2006; Yuen et al., 2010; Zadora et al., 2017; X.-m. Zhu, Han, Sargent, Yin, & Yao, 2009).

Additional environmental maternal exposures have been reported to influence methylation pattern or imprinted gene expression in offspring. Exposure to air pollution, antibiotic, tobacco or alcohol consumption are all associated with changes in imprinting methylation or expression (Kappil et al., 2015; Kingsley et al., 2017; A. Vidal et al., 2013). Nevertheless, in most cases, these are single reports with limited sample sizes.

However, external environmental factors that influence the development and alter imprints in the offspring are not limited to maternal exposures. Alterations in male gametes have also been reported following paternal exposures, and it has been shown to be present following fertilisation (Soubry, Hoyo, Jirtle, & Murphy, 2014). For example, studies with obese parents have shown an increased LOM at some imprinted genes associated with paternal and not maternal over-nutrition (Soubry et al., 2015, 2013).

### 1.5.2 Assisted reproductive technologies (ART)

Due to the increasing use of ART to overcome subfertility, it is important to evaluate the possible health problems that may be associated with its use. Babies conceived through ART can have lower birth weights, a fact generally associated with its increased rate of multiple births (Schieve et al., 2002). For this reason, the possible effects of ART on influencing birth weight in human are restricted to studying singleton births, which still

suggests that the use of ART is associated with decreased birth weights (M. Hansen & Bower, 2014).

Numerous studies have suggested that children birth as a result of ART, including ovarian stimulation, *in vitro* fertilisation (IVF) and intracytoplasmic sperm injections (ICSI), have a higher risk of diseases with epigenetic aetiologies, including imprinting disorders (G. F. Cox et al., 2002; DeBaun, Niemitz, & Feinberg, 2003; Manipalviratn, DeCherney, & Segars, 2009; Sutcliffe et al., 2005). Specifically, the use of ART has increased the incidence of epigenetic disturbances causing BWS and AS by LOM at gDMRs that acquire methylation in the female germline (*Kv*DMR1 and *SNURF*, respectively) (M. Gomes, Gomes, Pinto, & Ramos, 2007; Manning et al., 2000).

Additional cohort studies have also shown significantly epigenetic changes in cord blood and placentas from babies conceived by ART, mainly, IVF/ICSI, which includes not only imprinted DMRs but also at transposon elements (Choux et al., 2017; M. Gomes, Huber, Ferriani, Amaral Neto, & Ramos, 2009; Nelissen et al., 2013).

There are various critical moments in the use of ART where the imprinting can be susceptible to change (e.g. during gametes maturation or when imprints are established). One of the most critical is during the first five days of development while the embryos are in culture media and not in the uterus, a period coinciding with epigenetic reprogramming and ZGA (El Hajj & Haaf, 2013). However, since people requiring ART have underlying fertility issues, the increased incidence of epigenetic variability may not be due to ART use itself, but to other factors that are causing the subfertility. To elucidate this issue, different mouse models have been described (Calle et al., 2012; El Hajj & Haaf, 2013). One of the first mouse studies linking *in vitro* culture with imprinting defects, compared IVF embryos maintained in different culture media with spontaneously conceived blastocysts. This revealed widespread changes to imprinted gene expression, but the alterations were restricted to the placenta (Doherty, Mann, Tremblay, Bartolomei, & Schultz, 2000; Mann et al., 2004). To assess if ICSI would affect male reproductive health, a study published in 2014 described a lower expression of some imprinted genes in testis of mice conceived by ICSI when compared with naturally conceived. Although all mice studied have a normal phenotype, the methylation at imprinted loci was also altered (X.-R. Xu et al., 2014). Additionally, a human study published in 2015, which used data from oocyte and sperm donors without any subfertility problem, suggests that there are an increase epigenetic disturbances due to the ART use and not only because of the underlying subfertility (Song et al., 2015).

However, not all published data suggest that epigenetic profiles and imprinting are altered following ART. Several studies comparing placentas from natural conceptions with ART have shown no significant changes at the level of allelic expression or methylation (Camprubí et al., 2013). Despite this, surveys with a high number of participants have shown an

increase in placenta-associated abnormalities, including pre-eclampsia, following ART treatments (Shevell et al., 2005). Similarly, mouse studies have revealed abnormal placenta development after ART (S. Chen et al., 2015). With this in mind, a comprehensive analysis of genome-wide imprinting, focusing on the newly described placenta-specific imprinted loci in a large cohort of placenta samples is necessary to determine if imprinted regions are prone to epigenetic instability during ART.

# Chapter 2

# Hypothesis and Objectives

**Hypothesis**: Compared to somatic tissues, the human placenta is enriched for genes subject to genomic imprinting that may, in turn, be essential for intrauterine growth and development.

The **overall aim** of this thesis was *to determine the extent of imprinting in the human placenta and how these genes can influence human growth and development.* Specifically, **we aimed to**:

1. Study products of conception from NLRP7-mutated women to expand our knowledge on the impact of these maternal-effect mutations on the different imprinted DMRs and to determine inter- and intra-individual variability that may exist between affected women.

2. Confirm by methylation-sensitive genotyping and bisulphite PCR the growing list of new placental-specific imprinted DMRs identified using our newly developed bioinformatics pipeline.

3. Determine if the placenta-specific maternally methylated regions result in allelic expression of nearby genes. In the cases for which imprinting is observed, do these genes influence the placenta development and fetal growth?

4. Ascertain if, like their ubiquitously imprinted counterparts, the placenta-specific DMRs dictate the allelic expression of neighbouring genes *in cis* (resulting in a cluster of imprinted genes).

5. Study the molecular mechanism maintaining placenta-specific imprinted DMRs during post-fertilisation reprogramming by determining the contribution of gamete-derived methylation difference to placenta-specific DMRs and the presence of other epigenetic modifications such as 5-hmC and histone modifications.

6. Assess the stability of allelic methylation at placenta-specific DMRs and the associated gene expression in samples obtained from complicated pregnancies (IUGR and pre-eclampsia) and those conceived using ART.

# Chapter 3

# Material and methods

## 3.1 Biological samples

Both, the originating Hospitals and Bellvitge Biomedical Research Institute (IDIBELL), granted ethical approval for the use of the following samples. Besides, for the human samples, written informed consent was obtained from all participants.

### 3.1.1 Human samples

#### 3.1.1.1 Recurrent hydatidiform moles and blood samples from *NLRP7*-mutated women

For the first part of this thesis, we used five molar biopsies from four different women with two mutated copies of the *NLRP7* gene who were referred to the Trophoblastic Tumour Screening and Treatment Centre, Charing Cross Hospital (London, UK). All women had presented with multiple RHMs (patients 1-4 had 3, 6, 4 and 3 previous RHMs, respectively) and provided informed consent to use their tissues for research.

The ethical approval was granted by the IDIBELL (PR096/10) and the Tissue Management Committee of the Imperial College Healthcare NHS Trust Research Tissue Bank (R15048), which is approved by the National Research Ethics Service to provide deemed ethics for projects accessing material and data stored within the Research Tissue Bank.

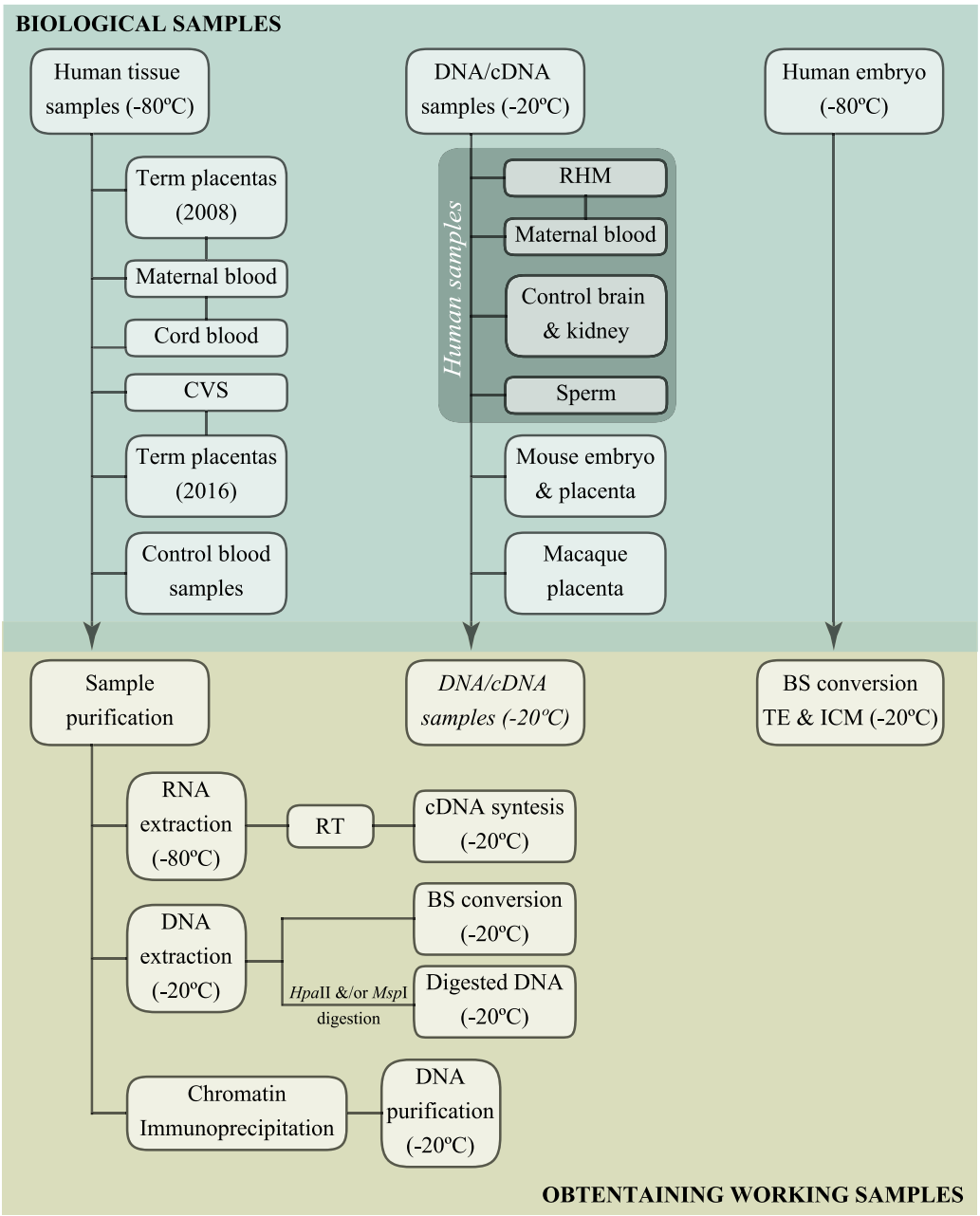



Figure 3.1: **Biological samples stored and DNA/RNA purification.** Paired samples: Term placentas (collected in 2008) with corresponding cord and maternal blood; Term placentas (collected in 2016) with corresponding chorionic villus sampling (*CVS*); Recurrent hydatidiform moles (*RHM*) and corresponding maternal blood. *BS*: bisulphite; *TE*: trophectoderm; *ICM*: inner cell mass; *RT*: reverse transcription reaction.

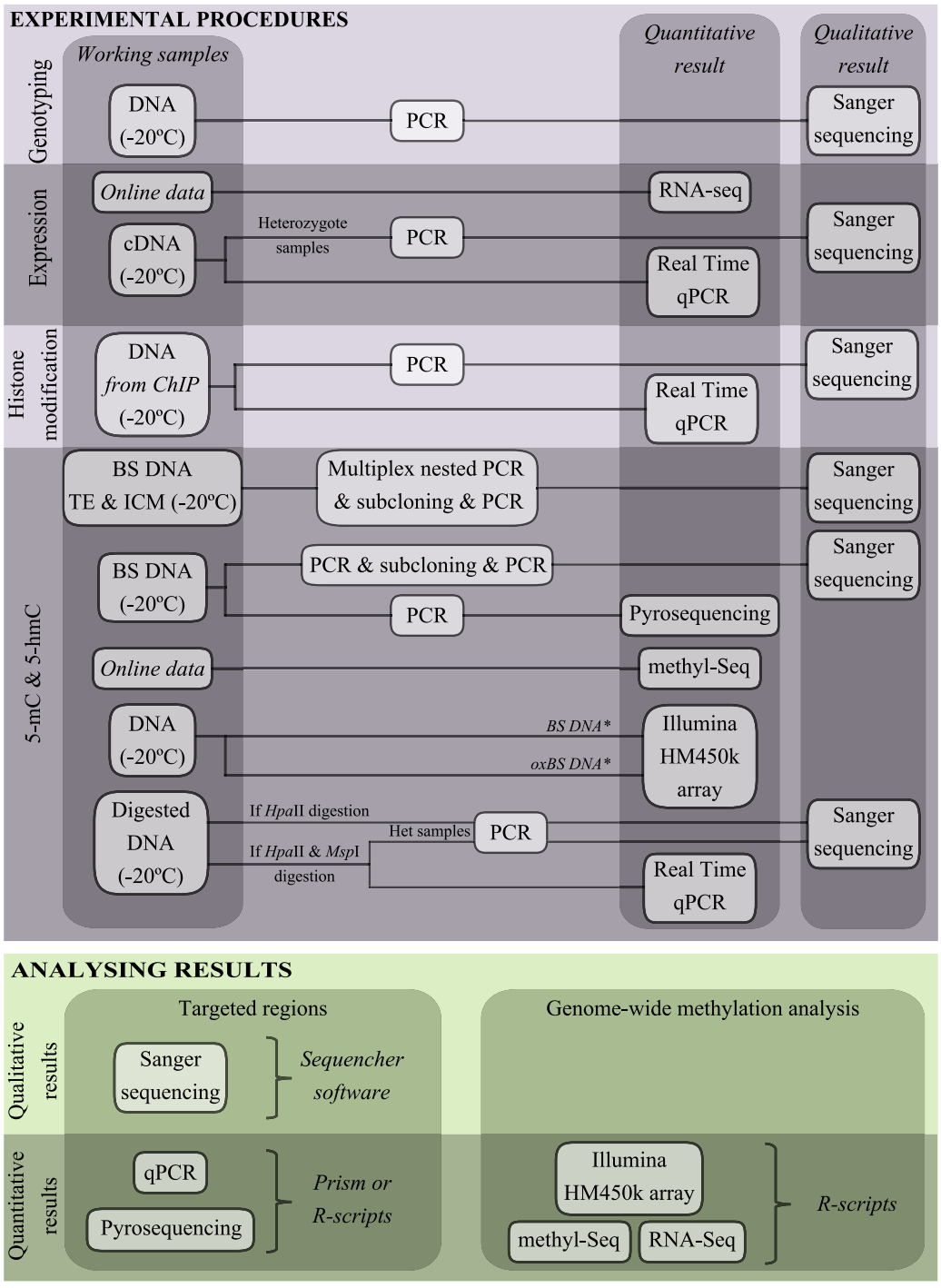



Figure 3.2: **Workflow of experimental procedures and analysis results.** *ChIP*: chromatin immunoprecipitation. *Het*: heterozygote; *PCR*: Polymerase Chain Reaction; *qPCR*: Quantitative PCR; * treated in the Genomic facility of the PEBC: oxidation treatment (*ox*) and bisulphite (*BS*) converted DNA.

#### 3.1.1.2 Placenta cohort, cord and maternal blood

A cohort of 142 placental biopsies (collected from 2008 – onwards), 83 with corresponding maternal blood samples were obtained at Hospital Sant Joan de Deu (Barcelona, Spain). Six of these placental samples presented pre-eclampsia, 54 presenting IUGR and 41 following ART. Our reference control placenta group are the 55 placentas that did not present pre-eclampsia, IUGR and are were naturally conceived (see Table S1 with clinical and anthropometric data associated with these samples). Additionally, for some qualitative analysis, placenta samples from the National Center for Child Health and Development (Tokyo, Japan) were also used. All placenta biopsies were collected from the foetal side around the cord insertion site. All placental-derived DNA samples were free of maternal DNA contamination based on microsatellite repeat and SNP analysis. To study methylation status during gestation, two placental biopsies (collected in 2016) with its corresponding first-trimester chorionic villus sampling DNA (CVS) were obtained at the Biobank of the Agusti Pi i Sunyer Institute for Biomedical Research (Barcelona, Spain).

All mothers provided written informed consent for themselves and their child before participating in the study. Ethical approval for collecting blood and placental samples was granted by the ethical committees of Hospital St Joan De Deu Ethics Committee (Study number 35/07), Bellvitge Institute for Biomedical Research (PR006/08) and the National Center for Child Health and Development (project 234).

#### 3.1.1.3 Control samples from other tissues

Control brain DNA samples were obtained from BrainNet Europe/Barcelona tissue bank. The dissection of frontal cortex was performed by an experienced pathologist on cadavers within 14 hours of death.

Peripheral blood samples, which were obtained from healthy volunteers from the Cancer Epigenetics and Biology Program (PEBC) at IDIBELL. Written consent was obtained from all individuals.

### 3.1.2 Human embryos

Three surplus human blastocysts were recruited at the Valencian Infertility Institute Foundation (FIVI). The blastocysts were thawed using the Cryotop method following manufacturer's instructions (Kuwayama, 2007) and incubated in CCM medium (Vitrolife, Göteborg, Sweden) for 6-12 hours before micro-dissection in order to allow their full expansion and the inner cell mass (ICM) and trophectoderm (TE), that were subsequently separated by micromanipulation using laser technology (OCTAX, Herborn, Germany). The

separated ICMs and TEs were individually placed in Polymerase Chain Reaction (PCR) tubes containing 2.5$\mu$l of PBS and immediately snap frozen at -80°C until processing.

The use of surplus human embryos and gametes for this study was evaluated and approved by the scientific and ethics committee of the FIVI (1310-FIVI-131-CS), Bellvitge Institute for Biomedical Research Ethics Committee (PR292/14), the National Committee for Human Reproduction (CNRHA) and the Regional Health Council of Valencia. Written informed consent was obtained from all participants.

### 3.1.3 Mouse DNA samples

Wild-type mouse embryos and placenta were produced by crossing C57BL/6 (B6) with *Mus musculus molossus* (JF1) or *Mus musculus castaneous* (CAST) mice were kind gifts from Dr Nakabayashi.The Institutional Review Board Committees approved mouse work at the National Centre for Child Health and Development, Japan (approval number A2010-002). Animal husbandry and breeding were conducted according to the institutional guidelines for the care, and the use of laboratory animals and samples were collected at embryonic day (E) 9.5 and E18.5.

### 3.1.4 Rhesus macaque DNA placental sample

A single placenta sample from rhesus macaque was obtained from the breeding colony of the Biomedical Primate Research Centre, Rijswijk, Netherlands using protocols (following a C-section procedure) approved by the Committee on the Ethics of Animal Tissue Collection at BPRC (Permit # 730). The EUPRIM-Net Bio-Bank is conducted and supervised by the scientific government board along with all lines of EU regulations and in harmonisation with Directive 2010/63/EU on the Protection of Animals Used for Scientific Purposes.

## 3.2 Obtaining working samples

To perform the different experiments for this thesis, we extracted DNA and RNA for all tissue samples used and subsequently synthesised complementary DNA (cDNA) from DNase-treated RNA for expression studies and treated DNA with bisulphite (BS) DNA samples to some methylation studies.

### 3.2.1 Mononuclear cell extraction from fresh blood

Mononuclear cells (MCs) were isolated from fresh peripheral blood (collected in a tube containing EDTA-anticoagulant) by density centrifugation by using Lymphoprep (Axis-Shield). The blood samples were diluted in an equal volume of PBS and gentle, this mix was layer on the same volume of Lymphoprep. Samples were centrifuged at 800g with no brake for 20 minutes at 4ºC. After centrifugation, MCs formed a distinct layer at the medium interface (see Figure 3.3). MCs were removed using a Pasteur pipette and transferred to another tube, where MCs were washed in PBS and centrifuged at 400g for 10 minutes. The resulting pellet was used for DNA or RNA extraction.

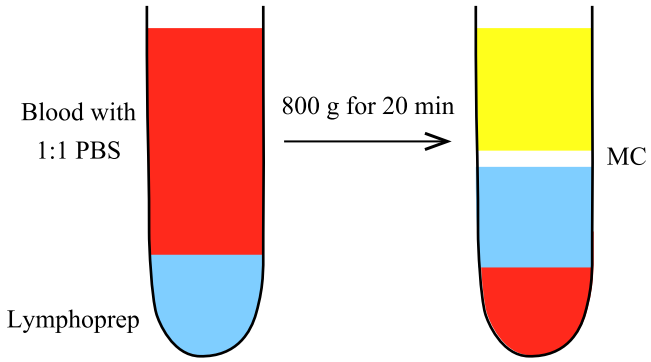



Figure 3.3: **Mononuclear cell extraction from fresh blood using Lymphoprep.**

### 3.2.2 DNA extraction from tissue samples and cells

Genomic DNA was isolated by the standard phenol/chloroform extraction procedure. Cells pellets were and washed with PBS and centrifuged at 1000 rpm for 5 minutes. After removing the PBS, cell pellets were resuspended in 2ml of lysis buffer (100mM NaCl, 10mM Tris-HCl pH = 8 and 25mM EDTA) supplemented with 25$\mu l$ of 20% SDS and 25$\mu l$ of Proteinase K (20mg/ml) (Sigma-Aldrich) and incubated overnight at 55ºC. The following day, the cell lysates were transferred to a Phase Lock gen tube (Prime5) and mixed with an equal volume of phenol/chloroform. To separate organic and the aqueous phases containing the DNA, centrifugation was performed at 1000 rpm for 5 minutes. The phenol/chloroform extraction was repeated 3-4 times and was followed by three extractions with chloroform only. Genomic DNA was precipitated by adding 1/10 volume of 3M NaOAc and 2.5 volumes of 100% EtOH. Genomic DNA was precipitated by inverting the tubes gently. High molecular weight DNA was transferred to a fresh tube and washed in 70% of EtOH and air-dried. Dried pellets were resuspended in TE (Tris EDTA) or $H_2O$. The quantity and purity of the DNA were determined by measuring absorbance at 260 nm (A260) and 280nm (A280) with a spectrophotometer (Nanodrop). An A260/280 ratio

of 1.8-2,0, indicating DNA free of contaminating phenol or protein, was obtained for all samples. Genomic DNA was stored at -20ºC until use.

#### 3.2.2.1 Sodium bisulphite DNA conversion

Different DNA Methylation Kits from ZYMO Research were used depending the number of DNA samples to convert and the quantity of starting material. In general, we used the EZ DNA Methylation kit (ZYMO, Orange CA) for BS DNA conversion, using $1\mu g$ of DNA as a template and following manufacturers' protocol. To $45\mu l$ of $1\mu g$ DNA, $5\mu l$ of M-Dilution Buffer were added and incubated at 37ºC for 15 minutes. Subsequently, $100\mu l$ of CT conversion Reagent ($750\mu l$ of $H_2O$ and $210\mu l$ of M-Dilution Buffer shaken for 10 minutes) was added to each sample and incubated for 16 cycles (96ºC for 30 sec and 50ºC for 60 min). After the incubation, we added $400\mu l$ of M-Binding Buffer to a Zymo-Spin IC Column with the corresponding collection tubes. Samples were loaded into the columns and mixed by inverting the column several times. After a centrifuge at 11000g for 30 seconds and discard the flow-through, washes were performed with $100\mu l$ of M-Wash Buffer and same centrifuge conditions. To finish the conversions in which all non-methylated cytosines are deaminated to uracils, we added $200\mu l$ of M-Desulphonation buffer to each column and incubate at room temperature for 15-20 minutes. After the incubation, we centrifuged at 1100g for 30 seconds and repeated twice the washing step, this time, with $200\mu l$ of M-Wash Buffer. Finally, we placed the columns into a 1.5ml tube and performed a double elution with $10\mu l$ of M-Elution Buffer by centrifuging 11000g for 30 seconds. BS-converted DNA was stored at -20ºC until use. Additionally, the EZ DNA Methylation-Direct kit (ZYMO, Orange, CA) for a small number of cells or low DNA concentration was used with surgically separated ICM and TE biopsies following suppliers instruction.

#### 3.2.2.2 Restriction enzyme digestion

##### 3.2.2.2.1 For DNA methylation analysis

The methylation-sensitive restriction enzymes (MSRE) used for our confirmation analysis was *Hpa*II (NEB). These endonucleases only cut DNA if the cytosine within the CpG dinucleotide in the recognition motif is unmethylated (*Hpa*II: C**CG**G). For this assay, the genomic DNA was first digested and then subjected to amplification by PCR using primers flanking the polymorphic variant and incorporates multiple restriction sites. The genotype obtained for the digested DNA was compared to that of genomic DNA. If the CpG is methylated, then a PCR product will be generated, however, in the absence of methylation, no amplification product will be generated. Therefore in the case of an imprinted DMR in which the genomic DNA is heterozygous, only one allele will be detected in the sequence

reaction obtained from the digested template. Alternatively, in some cases we also used *Hha*I (NEB) or *Bst*UI (NEB), which only cuts unmethylated sequence (*Hha*I: G**CG**C, *BstU*I: C**GC**G).

To $1\mu g$ of DNA (predetermined to be heterozygous for the studied region) we added $0.5\mu l$ of MSRE [$10U/\mu l$], $2.5\mu l$ of Buffer (NEB), and sterile $H_2O$ up to $22\mu l$. The mix was incubated for 4 hours at 37°C. The reaction was "boosted" with an additional 2U of the enzyme for a further 30 minutes to ensure full digestion. The enzyme was subsequently heated inactivation at 65°C for 20 minutes and cleaned with ethanol precipitation. The DNA was resuspended in $20\mu l$ with $1\mu l$ and used as a template for PCR reactions.

#### 3.2.2.2.2 For 5-hmC placenta study (5-hmC and 5-mC discrimination)

To quantify both 5-hmC and 5-mC, $5\mu g$ of heterozygous placenta DNA was subject to DNA Glucosylation using the Epimark kit (New England Biolabs). In brief, $267\mu l$ of DNA sample (diluted with nuclease-free $H_2O$) were mixed with $12.4\mu l$ of UDP-Glucose and $31\mu l$ of NEBuffer 4. The reaction was split into two tubes of $155\mu l$ each and 30U ($3\mu l$) of T4 Beta-glucosyltransferase (T4-BGT) was added to one tube and $3\mu l$ of $H_2O$ to the other. Both tubes were incubated during 18 hours at 37°C.

After incubation, the reactions were split again, this time in 3 tubes of $50\mu l$ reaction each. The 3 sets of tubes were subjected to different conditions (see Figure 3.4 for schematically representation). 100U ($1\mu l$) of MspI were added to one set of tubes and 50U ($1\mu l$) of *Hpa*II into the other set. No enzyme was added to the $3^{rd}$ set of tubes. All tubes were incubated at 37°C for 16 hours. The reaction was "boosted" with an additional 2U in the case of *Hpa*II and 4U for *Msp*I for a further 30 minutes to ensure full digestion. Finally, the DNA was subject to proteinase K digestion, clean-up and stored at -20°C until use.

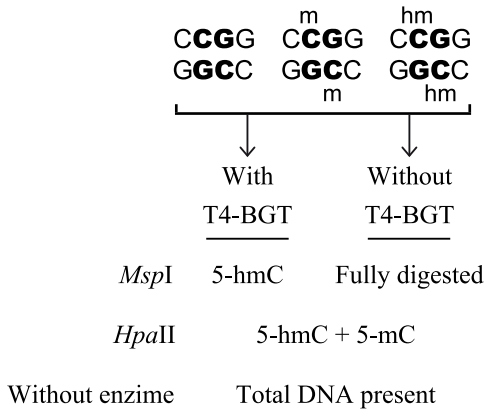



Figure 3.4: **The conditions of the EpiMark kit to discriminate 5-hmC and 5mC.**

### 3.2.3 RNA extraction from tissue samples and cells

Total RNA was isolated using TriZol reagent (Invitrogen). Cells were lysed in 1ml of TriZol, which is a commercial combination of guanidine isothiocyanate and phenol. Subsequently, $200\mu l$ of chloroform were added to each sample. After 30 minutes of centrifugation at 12000 rpm and 4ºC, each sample segregated into an organic phase containing DNA and proteins, leaving RNA in the aqueous supernatant layer. RNA was precipitated in 0.8 volumes of isopropanol and centrifuged for 1 hour at 12000 at 4ºC and washed with 70% ethanol. The RNA pellet was dissolved in $30\mu l$ RNase-free DEPC-treated water. The amount and purity of the RNA was determined by measuring absorbance at 260nm (A260) and 280nm (A280) with the spectrophotometer (Nanodrop). An A260/A280 ratio of 1.8-2.0, indicating RNA free of contaminating phenol or protein, was obtained for all RNA samples. RNA was stored at -80ºC until use.

#### 3.2.3.1 cDNA synthesis

To prepare RNA samples for the complementary DNA cDNA synthesis, we aliquot two tubes for each sample to generate a negative control set (to detect subsequent DNA contamination). Prior to the reverse transcription reaction (RT), we treated all RNA samples with DNAse I (Invotrogen) by adding to a $1\mu g$ of RNA, $8\mu l$ with DEPC-treated $H_2O$ (Invitrogen), $1\mu l$ of 10X DNase I reaction buffer (Invitrogen) and $1\mu l$ of DNase I (Amplification Grade, Invitrogen). After 15 minutes of incubation at room temperature, the enzyme was inactivated by adding $1\mu l$ of 25mM EDTA stop solution (Invitrogen) followed by 10 minutes incubation at 75ºC. The DNA-free RNA was subsequently used in a reverse transcription reaction to generate cDNA. Following incubation at 70ºC to remove any RNA secondary structure and a return to 37ºC, the RT-mix(+) (Table 3.1) was added to one sample set and the RT-mix(-) (without RT-MLV reverse transcriptase) to the other. The $20\mu l$ +/- reactions were incubate at 37ºC for 90 minutes and a final enzyme denaturation step of 10 minutes at 75ºC.

Table 3.1: **RT-mix (+) for reverse transcription reaction.**

| Volume ($\mu l$) | Reagent |
|---|---|
| 4 | reverse transcriptase buffer (Promega) |
| 1 | random hexamer primers (Promega) |
| 2 | 100nM dNTPs[1] (Promega) |
| 1 (40U) | M-MLV reverse transcriptase (Promega) |
| 0.25 (1U) | RNase inhibitor (RNasin, Promega) |

[1] *dNTP*: deoxynucleosidetriphosphates.

The presence of cDNA was confirmed by PCR for the housekeeping gene $\beta$-actin and the efficiently converted cDNA were stored at -20ºC until use.

### 3.2.4 Chromatin immunoprecipitation (ChIP)

To study histone modification at specific loci, we performed a native chromatin immunoprecipitation (ChIP). The next protocol is adapted from the provided by Dr Philippe Arnaud (GReD, Clermont-Ferrand, France) optimized by Cécile Choux (a visiting PhD student from University Hospital of Dijon) and Dr Paolo Petazzi, a post-doc in the laboratory.

All the buffers used in this protocol were prepared fresh each time, except buffer 2x. All working buffers remained in ice during the procedure except the elution and MNase digestion buffers (see Table S2 for all buffers used in the ChIP protocol).

#### 3.2.4.1 Nuclei purification

Approximately 100mg of frozen placenta tissue sample (-80ºC) was washed in cold PBS 1X (10ml in a 15ml tube) and centrifuge at 3220g during 4min (4°C). The supernatant was discarded and we placed the placenta to a 2ml microcentrifuge tube (conical, with screw cap clear; Thermo Scientific) with beads (from BeadBug prefilled tubes: 0.5mM Zirconium beads, triple-pure, high impact; Sigma-Aldrich) and 1ml of buffer 1. The tube was introduced into Precellys 24 tissue homogeniser (Bertin-Corp), for 3-4 repeated disruption cycles with a 90 seconds programme (between programs, the samples remained in ice during 5 minutes). We placed the tube containing the homogenise placenta and beads to a 15ml tube with 5ml of buffer 1 and 2ml of buffer 2 (NP40 0,8%). Slowly, we added the 8ml sample at the top of 25ml buffer 3 in a 50ml tube and in less than 5 minutes on the ice, the samples were centrifuged during 20 minutes at 8500 rpm at 4°C (acceleration 2, deceleration 1). The supernatant was removed and the nucleic pellet resuspended in MNase digestion buffer (usually 1 to 1.5ml).

The quality and quantity of the nuclei preparation was visualised under a microscope and subsequently quantified by Nanodrop with an aliquot diluted in 0.1% SDS final concentration. The sample was distributed in aliquots of $500\mu l$ in 1,5ml tube ($200\mu g$ of chromatin and $300\mu l$ of Mnase Digestion buffer). After 5 minutes at 37ºC, the samples were digested for 3 minutes at 37ºC with $1\mu l$ of Mnase ($15U/\mu l$) and then, the reaction were stopped by adding $20\mu l$ of EDTA 0.5 M. After a 15.8 g centrifuge of 10 minutes at 4ºC, the supernatant was retained in 1.5 tubes (corresponding to FRACTION S1). The pellets were resuspended in $500\mu l$ of lysis buffer and left 20-30 minutes in ice. A final centrifuge step was performed with and the supernatant, corresponding to the FRACTION S2, saved.

To confirm that the appropriately sized chromatin had been obtained following MNase digestion of fractions S1 and S2, we used the PCR clean-up manufactures' protocol from the Nucleospin gel and PCR clean-up kit (Macherey-Nagel) and extract DNA from 100$\mu l$ of each sample tube (for both, S1 and S1 fractions).

We checked the extracted DNA/nucleosome sizes on a ~1.5% agarose gel (2 g agarose and 125ml TE buffer). We loaded the samples (3$\mu l$ loading buffer with 15$\mu l$ DNA) and performed gel electrophoresis at 100 mV for approximately 45 minutes, followed by 2 hours agitating incubation in 200ml TE and 20$\mu l$ of Midori Green Advance (Nippon Genetics) to stain the DNA.

#### 3.2.4.2 Chromatin incubation with antibodies

The chromatin used for the ChIP assay is composed of a mix of an equal quantity of S1 and S2 fractions. Mixing these two fractions avoids the introduction of any bias because of a preferential MNAse digestion of one allele in the S1 "accessible" fraction.

We used 4$\mu g$ of chromatin (75% S1 and 25% S2 mixed) in 500$\mu l$ volume of incubation buffer for each immunoprecipitation. Since the buffer composition of S1 (in Mnase buffer) and S2 fraction (in Lysis buffer) were different, we adjusted the concentrations with taking into account the concentration of NaCl, Tris-Cl and EDTA already present in S1 (incubation buffer) and S2 (lysis buffer).

For bead washing, we washed three times the total of beads (Dynabeads Protein G; Invitrogen) required in 1ml of cold PBS with 5% BSA using the magnetic rack and resuspended in the same volume using PBS/BSA 5%. With five conditions (input 50%, IgG, H3K4me3, H3K4me2, H3K9ac), 20$\mu l$ clean beads/condition and 500$\mu l$ of chromatin/condition we mixed 100$\mu l$ of beads with 2.5ml of chromatin for pre-cleaning. Simultaneously, for each antibody – IgG (Millipore 12-371) 2.5$\mu l$; H3K4me3 (Diagenode C15410003-50) 1$\mu l$; H3K4me2 (Millipore 07-030) 2.5$\mu l$; H3K9ac (Cell Signaling 9649S) 2.5$\mu l$ –, we mixed with 40$\mu l$ beads and 410$\mu l$ PBS/BSA 5%. All tubes containing the samples with beads and antibodies with beads were incubated overnight at 4°C on a rotating machine.

For the chromatin sample tubes, we removed the beads with the magnetic rack and aliquot 520$\mu l$ of chromatin in 5 different 1.5ml siliconized tubes. We washed twice the beads with antibodies using 1ml of cold PBS with 5% BSA and resuspended in 40$\mu l$ of the same solution. We mixed the 40$\mu l$ of beads with antibodies and the 520$\mu l$ of cleaned chromatin, and it was incubated for 4 hours at 4°C on a rotating machine. Additionally, 260$\mu l$ of chromatin was stored at -20°C to use as 50% input.

After the incubation, a centrifuge step of 3 minutes at 3200 rpm was performed and the supernatants removed with a magnetic rack. We then performed 3 washes with each washing solution buffers (solution A, then B and finally C). In each wash, we briefly mixed

the samples and centrifuge 1 minute at 3200 rpm. After the last wash, we added 400$\mu l$ of elution buffer and the samples were placed into a rotating machine (20 rpm) during 30 min at room temperature. Finally, all samples, including input, were digested with 2.2$\mu l$ Proteinase K (100ug/ml) for 1 hour at 65°C. The beads were removed and the supernatant kept to its DNA extraction (BOUND FRACTION).

For the DNA extraction, we used the Nucleospin gel and PCR clean-up kit (Macherey-Nagel) but with a modified protocol. Since our samples contain SDS (except for the input), they were initially mixed with five volumes of NTB buffer (2ml) and 1 volume of isopropanol (400$\mu l$) instead of Buffer NT1. We followed the manufacturers' protocol for the rest of the procedure, and we performed two elution steps with 20$\mu l$ of $H_2O$ heated at 50°C. The DNA samples were stored at -20ºC until use.

## 3.3 Downstream experimental procedures

The main objective of this project was to characterise known imprinted regions in RHM samples and identify novel placenta-specific imprinted DMRs (see Figure 3.5 A-C for confirmation of maternal-derived methylation at placenta-specific DMRs). Additional experiments were performed to characterise the expression and epigenetic regulation of the identified regions in our placenta cohort.

In summary, for imprinting analysis performed in this thesis, qualitative results were obtained by direct sequencing (Sanger). Sequence traces were interrogated using Sequencher v4.6 (Gene Codes Corporation, MI). For all imprinting confirmations, "oppositely inherited genotypes" were needed, i.e. sample 1 with SNP variant A being the maternally methylated allele and SNP variant C the expressed allele, we also need additional samples with SNP variant C being the maternally methylated allele and SNP variant A expressed to discount mQTLs. These analysis include: *Samples genotyping* (Figure 3.5A) by PCR amplification of specific regions containing high polymorphic SNP (identified in the UCSC hg19 browser) within a candidate DMR, following by Sanger sequencing and selection of all heterozygous samples; *Methylation-sensitive genotyping* (Figure 3.5B) by *Hpa*II digestion, which cuts all unmethylated DNA and after a PCR, the only allele detected and amplified was the methylated allele, by comparing the HpaII digested samples with corresponding mothers, we will elucidate the parental-origin of the methylated allele; *Bisulphate PCR and subcloning* (Figure 3.5C) to describe the strand-specific methylation distribution and also, parental origin of the methylated allele by mothers' genotyping; *Allelic RT-PCR* (Figure 3.5D) allows us to detect if the transcript is biallelically or monoallelically expressed, and if monoallelically expressed, its parental origin by comparing with mothers' genotyping.

In the following subsections, the experimental strategies and procedures will be explained, indicating the samples used for each experiment.

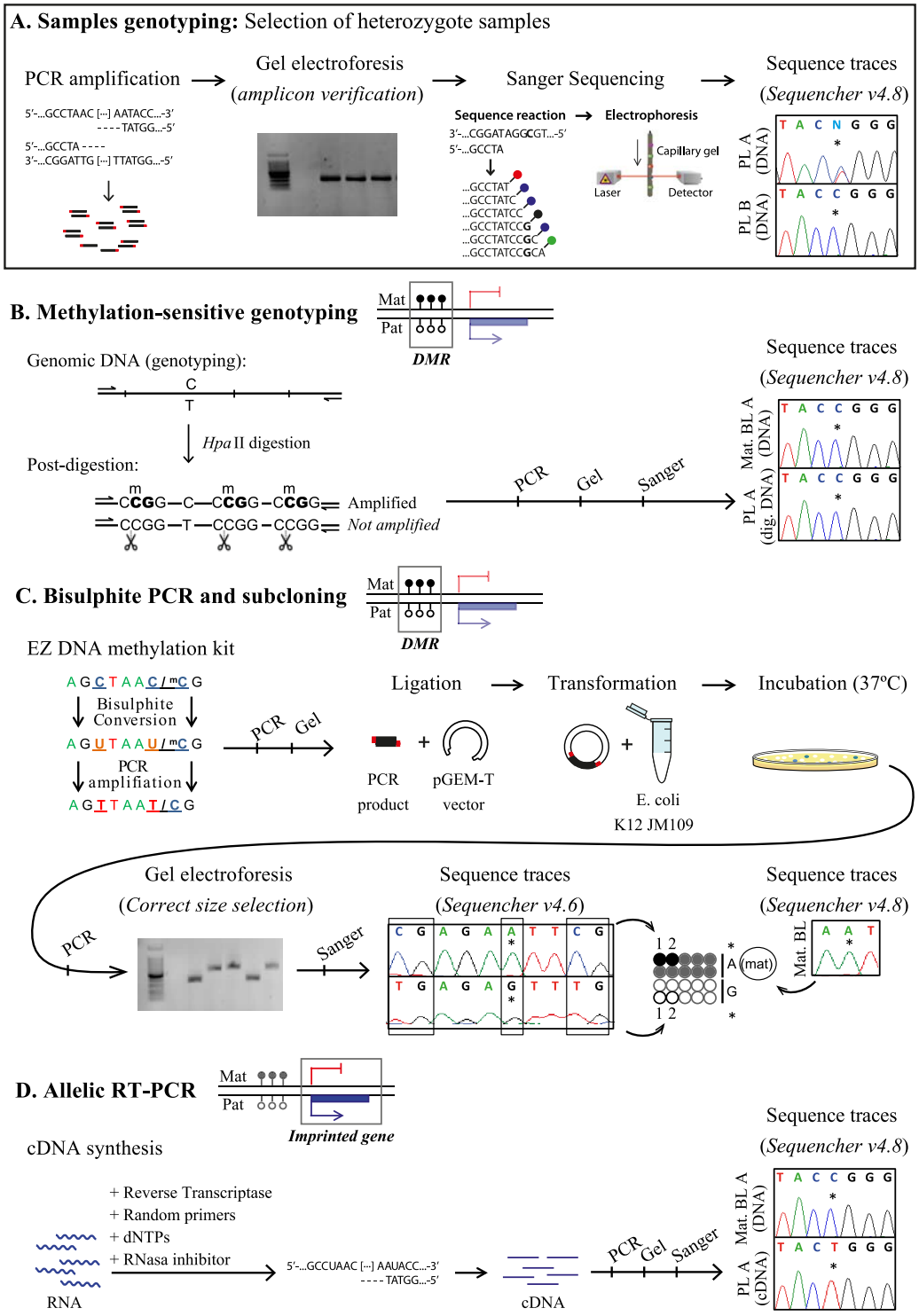



Figure 3.5: **Genotyping and imprinting analysis.** *Mat*: maternally derived allele; *pat*: paternally derived allele; *PL*: placenta sample; *BL*: blood sample (**A**) Samples genotyping; (**B**) Methylation-sensitive genotyping by *Hpa*II digestion; (**C**) Bisulphate PCR and subcloning; (**D**) Allelic RT-PCR.

### 3.3.1 The polymerase chain reaction (PCR)

The PCR was used for all DNA amplification procedures. The general PCR conditions were similar in all cases, with different primer annealing temperatures used depending on the primer pair used (See Tables S3 to S23 for the list of primers). Different templates were also used depending on the experimental procedure: BS-converted DNA for subcloning and methylation pyrosequencing procedures, DNA for genotyping, MSRE digested DNA for methylation-sensitive genotyping or cDNA for allelic expression analysis (see Figure 3.5 for examples).

#### 3.3.1.1 Mix preparation

For standard genotyping, we used 50ng of genomic DNA, digested DNA or cDNA as template in a $13\mu l$ reaction using Biotaq Taq-polymerase (Bioline). For all amplifications using BS-converted DNA or other difficult templates, we utilised hot-start Immolase Taq-polymerase (Bioline) in $25\mu l$ reactions (see Table 3.2 for PCR mix preparation). All PCRs performed included a non-template *Blank* to control for amplification from contamination. In the case of RT-PCR, the two sets of cDNA prepared were included.

Table 3.2: **PCR mix preparation.**

| Volume to $12\mu l$ | Volume to $24\mu l$ | Reagent |
|---|---|---|
| 5.775 | 11.55 | $H_2O$ (Braun) |
| 3.75 | 7.5 | 5mM Betaine (Sigma Aldrich) |
| 1.25 | 2.5 | 10x Buffer (Bioline) |
| 0.75 | 1.25 | 50mM MgCl (Bioline) |
| 0.25 | 0.5 | 2mM dNTPs[2] (Promega) |
| 0.25 | 0.5 | 100ng/$\mu l$ Primer F |
| 0.25 | 0.5 | 100ng/$\mu l$ Primer R |
| 0.1 | 0.2 | 5U/$\mu l$ Biotaq/Immolase (Bioline) |

#### 3.3.1.2 Thermo-cycling conditions

PCRs were performed in a Thermal cycler (Applied Biosystems). For Immolase enzyme-based reactions an initial denaturation step of 96ºC for 10 min was required for enzyme activation (see Figure 3.6 for cycling conditions) and if the PCR was performed for pyrosequencing, 45 cycles were required to exhaustion primers to optimise PCR product

[2] *dNTP*: deoxynucleosidetriphosphates.

isolation during the pre-pyrosequencing protocol. If large PCR products were amplified (<1kb), the extension step of each cycle was increased from 30 seconds to 45 seconds or 1-minute extension at 72ºC.

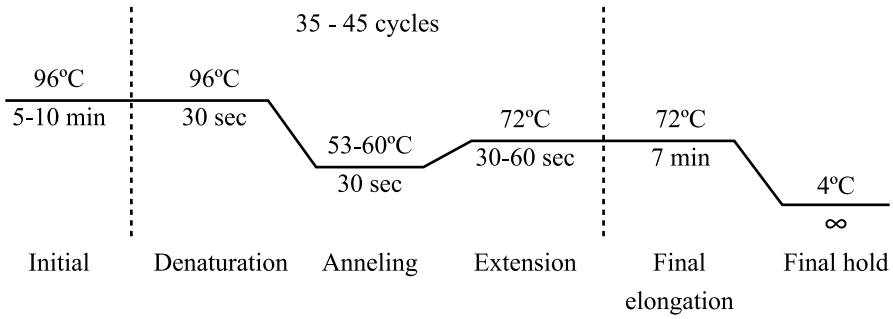



Figure 3.6: **Thermo-cycling PCR conditions.** If we used Biotaq enzyme, the initial step-time is 5 minutes, and if we used Immolase, 10 minutes.

#### 3.3.1.3 Agarose gel electrophoresis

To verify that PCR amplifications had generated the correct size amplicons, the PCR products were visualised upon gel electrophoresis of 1% – 2% agarose gels prepared with 0.5X TAE (diluted from 10X TAE stock solution. For 1l of 10X TAE, 48.4 g Tris base, 11.42ml of glacial acetic acid and 40ml of 0.5M EDTA with pH 8) and 4$\mu l$ of Midori Green Advance (Nippon Genetics). 0.5X TAE was also used as a gel electrophoresis buffer. In general 4$\mu l$ sample was mixed with 1$\mu l$ of loading dye (GeneRuler) and load into gel wells. A 100 bp or 1 kb DNA ladder (GeneRuler) was also loaded on the agarose gel to confirm PCR product size. Gel electrophoresis was carried out at 90-130V using PowerPack basic power supply (Bio-Rad) during 20-40 minutes. Finally, the PCR products were visualised using the Q-Box equipment and GeneSys automatic software. If the bands are sufficiently abundant, and they are of the correct size with no primer dimers, the products were used for a downstream procedure: direct sequencing when the template used was DNA, digested DNA or cDNA, and subcloning protocol or purification to pyrosequencing when BS-converted DNA was used as a template.

#### 3.3.1.4 PCR product purification from agarose gel

In some cases, for example, if primer dimers or multiple PCR products were observed on the agarose gel, the correctly sized amplicons were cut out from the gel with a clean blade and placed into a 1.5ml Eppendorf tube and subject to PCR clean-up. The extraction was performed using NucleoSpin Gel and PCR Clean-up kits (Macherey-Nagel) following the manufacturers' protocol. Gel fragments were weighed, and an appropriate volume of buffer

NT was added (for each 100mg agarose gel, 200$\mu l$ of Buffer NT) and incubate at 50ºC for 5-10 minutes with occasional vortexing to dissolve the gel. Samples were loaded onto a NucleoSpin Gel and PCR Clean-up Columns and centrifuged at 11000g for 30 seconds. The flow-through is discarded, and the columns washed twice with 700$\mu l$ of buffer NT3 and centrifuged at 11000g for 30 seconds for the first wash and 1 minute for the second to ensure that the membrane is dry. PCR product was eluted in 10-20$\mu l$ of NT or $H_2O$ and used for ligation or direct sequencing.

### 3.3.2 Multiplex nested PCR

We employed a multiplex nest PCR approach to maximise data generation from ICM and TE DNA samples which were subject to direct BS conversion. Two sets of primers (outer and inner pairs) were designed to each locus and robustly optimised in placenta-derived BS DNA to ensure efficient amplification of both methylated and unmethylated strands at a single annealing temperature without contamination or the formation of primer dimers.

For the multiplex step, outer primers (15 separate pairs targeting different loci) were co-amplified in the first reaction using Immolase Taq polymerase (Bioline) for 45 cycles and 1-minute extension at 72ºC (see Table 3.3 for the multiplex PCR-mix). The multiplex PCR was carried on in a total volume of 97$\mu l$, including the 30$\mu l$ of converted DNA (ICM and TE) we used as a template. We included three controls: blank with no sample, TE buffer from the TE and ICM extraction/conversion, and BS-converted DNA as a positive control.

Table 3.3: **Mix for the multiplex PCR.**

| Volume ($\mu l$) | Reagent |
|---|---|
| 13.7 | $H_2O$ (Braun) |
| 30 | 5mM Betaine (Sigma-Aldrich) |
| 10 | 10x Buffer (Bioline) |
| 3 | 50mM MgCl (Bioline) |
| 2 | 2mM dNTPs (Promega) |
| 0.25 (each primers) | 100ng/$\mu l$ **Outter Primes** |
| 0.8 | Immolase (Bioline) |

Nested second round amplifications utilised the internal primer pairs specific for each region with specific PCR performed on aliquots of the first round product. The nested reaction was performed in a total volume of 13$\mu l$ reaction (12$\mu l$ mix and 1$\mu l$ of first round PCR as template, with an additional blank for each nested PCR). These were amplified for 45 cycles at 53ºC (except *H19* region, that was amplified at 60ºC). All second round

amplification products were sub-cloned into the pGEM-T easy vector for direct sequencing.

### 3.3.3 Subcloning

To distinguish the methylation of each allele, PCR subcloning was performed by T-Vector System based on the fact that some thermostable DNA polymerases, like Taq DNA polymerase, adds a single A nucleotide to the 3' end of blunt DNA, producing "sticky-end" PCR products ideal for 3' T vector cloning. The A-tailed BS-PCR products were directly ligated to the pGEM-T easy vector (Promega) (see Figure 3.5C for the general process and Figure 3.7 for the pGEM-T Easy vector) by the action of T4 DNA ligase (Promega) following the standards protocols at 4ºC for 12-24 hours. Subsequently, the ligations were transformed into *E.coli* JM109 (Promega) by heat shock (30 min in ice, 45 seconds at 42ºC and 2 minutes in ice) followed by a growth in LB without antibiotic during 30 minutes shaking at 37ºC. The transformed cultures were then spread on LB-agar plate with ampicillin (Sigma Aldrich), X-Galactose (Promega) and IPTG (Sigma Aldrich) and grown at 37ºC overnight. The pGEM-T easy vector allows for blue/white selection based on hydrolysis of $\beta$-galactosides. Positive white clones were picked and grown in $50\mu l$ of LB-media without ampicillin for an hour. A selection PCR with primers designed to flank the T-cloning/multiple cloning sites of the pGEM-T easy vector were performed with $1\mu l$ of the culture acting as the template. All appropriately sized amplicons (the original BS PCR product plus ~ 330 bp of the pGEMT vector) were sequenced by Sanger sequencing using T7 primer located immediately internal to the PCR oligonucleotides (see Tables S7-14 for the list of primers).

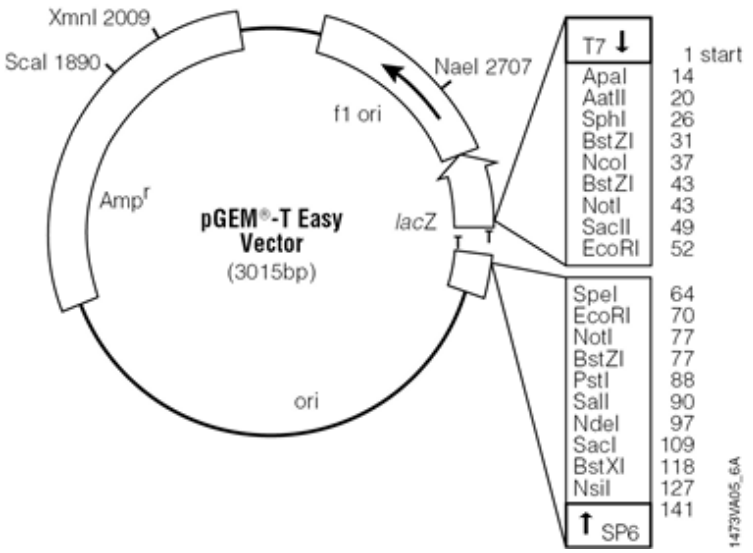



Figure 3.7: **The pGEM-T Easy vector** (Promega).

## 3.4 Obtaining and analysing results

The results of this thesis were obtained by both wet and dry experimental procedures. See Figure 3.2 for the general workflow.

### 3.4.1 Targeted regions

All the described experimental procedures in the previous section were performed for the propose of obtaining either, a **qualitative** – for heterozygote samples detection or to discriminate the epigenetic modified or expressed allele – or **quantitative** – to quantify the methylation, presence of histone modification at imprinted DMRs or quantify transcript abundance –.

#### 3.4.1.1 Qualitative analysis by Sanger sequencing

Sanger sequencing was used to genotype our samples. An inner primer was used for each target region to determine the DNA sequence of the region: genotyping DNA samples, discriminate the methylated/expressed allele or detect methylation after BS PCR and subcloning.

For the sequence reaction, in addition to dNTPs, fluorescent labelled modified di-dNTPs (dye terminators, different for each nucleotide) were also added. The DNA polymerase terminates the strand elongation when a di-dNTPs is incorporated. As a result, the DNA analyzer detects the nucleotide for each position by capillary electrophoresis. (see Figure 3.5A for a schematic representation of Sanger sequencing).

##### 3.4.1.1.1 PCR clean-up (first precipitation)

PCR products for sequencing were purified to remove unincorporated primers, dNTPs, salts and buffer. The sequencing templates were precipitated by the addition of 2.5 volumes of precipitation mix-1 (Precipitation mix1 preparation: 1ml of 3 M sodium acetate pH=4.6 and 250ml of 100% EtOH) and centrifuged at 2600g for 45 minutes at 4ºC or room temperature. The supernatant was removed, and the pellets washed with about 20$\mu l$ of 70% EtOH and 5 minutes at 3500 rpm. The pellets were air-dried and resuspended in 10-20$\mu l$ of $H_2O$. The final volume concentration of the clean PCR product was measured with a spectrophotometer (Nanodrop).

#### 3.4.1.1.2 Fluorescent-labelled cycle sequencing

Sequence reactions were carried out using the ABI PRISM BigDye Terminator Cycle Sequencing Kit, according to the manufacturer's protocol. Sequencing reactions were performed in a final volume of $10\mu l$ using specific ng of PCR product or DNA depending on the template size (see Table 3.4).

Table 3.4: **Amount of PCR product to use as a template for sequence reaction depending on its size.**

| PCR product size | ng of PCR product to use as template |
|---|---|
| 200pb | 6ng |
| 400pb | 12ng |
| 600pb | 18ng |
| 800pb | 25ng |
| 1kb | 30-50ng |

To $5.7\mu l$ of PCR product diluted with $H_2O$, $4\mu l$ of BigDye mix ($100\mu l$ of BigDye, $100\mu l$ of buffer and $300\mu l$ of $H_2O$) and $0.3\mu l$ of $100ng/\mu l$ sequence primer was added. The sequence reactions were carried out in a Thermal cycler (Applied Biosystems) by cycling conditions indicated in Figure 3.8.

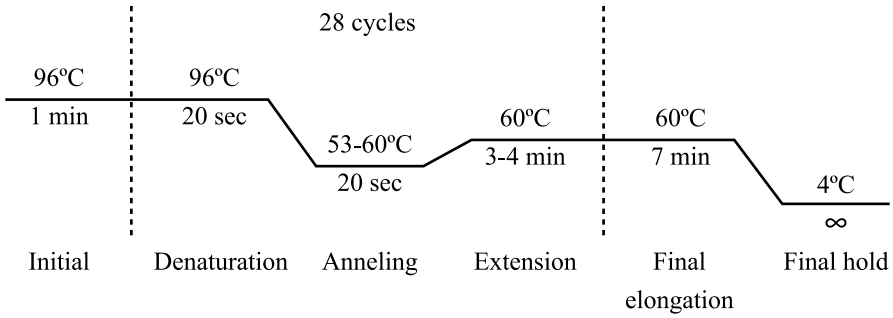



Figure 3.8: **Thermo-cycling conditions for fluorescent-labelled cycle sequencing.** For large sequence regions, 4 minutes extension step at 60ºC was required for each cycle.

#### 3.4.1.1.3 Sequencing clean-up (second precipitation)

The sequence reactions were precipitated by adding $170\mu l$ of the post-precipitation mix (1.6ml of 3 M sodium acetate pH=4.5, 13.3ml of $H_2O$ Mili-Q and 41.6ml of 100% EtOH) and centrifuge 3500 rpm (2600g) during 45 minutes at 4ºC. The supernatant was removed, and the pellet washed with $170\mu l$ of 70% EtOH for 5 minutes at 3500 rpm. After the pellet was air-dried, $10\mu l$ of Formamide (Thermo Fisher) and heat denatured at 96ºC for 1 minute.

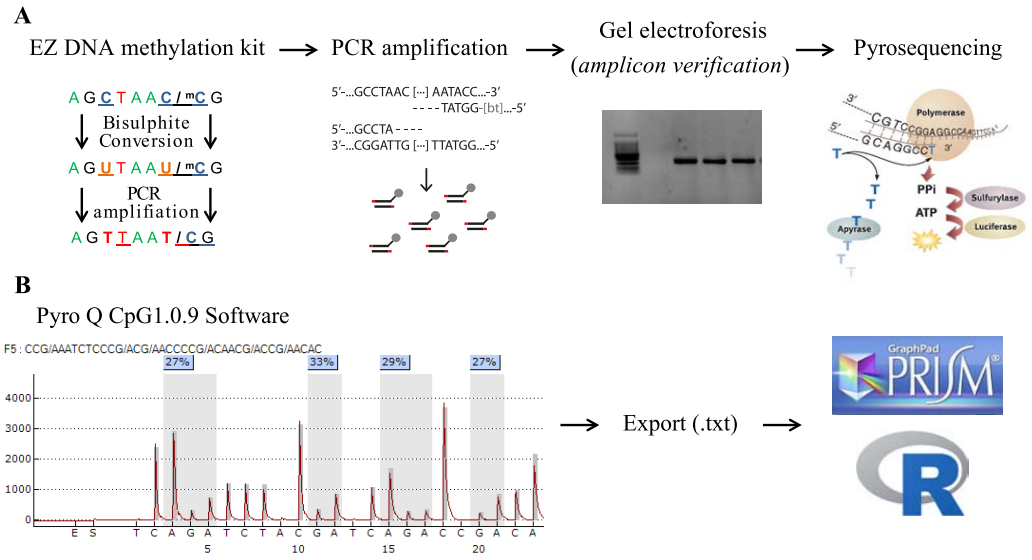



Figure 3.9: **Pyrosequencing procedure.** (**A**) Starting by BS DNA conversion as a template for PCR amplification by using one of the two primers biotinylated at 5' position. If we got the correct size amplicons and after a clean-up procedure, the pyrosequencing reactions were performed on a PyroMark Q96 instrument. (**B**) By using the Pyro Q CpG software, the % methylation of each CpG was obtained, and we analysed the results by statistical comparison between sample groups or studying % methylation for each region by using PRISM GraphPad or by R-scripts.

#### 3.4.1.1.4 DNA analyzer

Finally, the sequence reactions were run on an Electrophoresis ABI PRISM 3730 DNA analyzer (Applied Biosystems). The subsequent sequence electropherograms (sequence traces) were then analysed using Sequencher v4.8 program (see Figure 3.5 for examples).

### 3.4.1.2 Quantitative analysis by pyrosequencing

In pyrosequencing, in the presence of four enzymes (DNA polymerase, ATP sulfurylase, luciferase and apyrase) and two substrates (adenosine 5' phosphosulfate and luciferin), each incorporation event is accompanied by the release of pyrophosphate (PPi) in quantity equimolar to the amount of incorporated nucleotide. ATP sulfurylase converts PPi to ATP in the presence of adenosine 5' phosphosulfate and this ATP drives the luciferase-mediated conversion of luciferin to oxyluciferin that generates visible light in amounts that are proportional to the amount of ATP (see Figure 3.9).

For pyrosequencing, the same PCR amplification protocol was followed with the exception that one of the pair primers that was biotinylated and for sequencing, internal primers were designed being into the complementary strand of the biotinylated one. The entire

biotinylated PCR product (diluted to 40$\mu l$) was mixed with 38$\mu l$ of PyroMark Binding Buffer (Qiagen) and 2$\mu l$ (10mg/ml) streptavidin-coated Sepharose beads (Qiagen). Immobilization of the PCR products for purification was achieved by streptavidin-coated sepharose beads (Qiagen) with the use of the PyroMark Q96 Vacuum Prep Workstation. Purified samples were transferred and washed with ethanol 70%, denatured with sodium hydroxide and resuspended in commercial PyroMark Buffer (Qiagen). The single-stranded DNA was hybridised to 40 pmol sequencing primer dissolved in 11$\mu l$ of PyroMark Annealing Buffer (Qiagen) at 80$^{o}$C during 2 minutes. The pyrosequencing reaction was carried out on a PyroMark Q96 instrument.

#### 3.4.1.2.1 Quantitative methylation analysis

For methylation pyrosequencing, the peak height of C/T variants at CpG dinucleotides were analysed using Pyro Q CpG1.0.9 software (Biotage). The output of this program is the % of methylation for each CpG in the different regions investigated. The mean of methylation % were obtained in excel for each region. For the LOM first detection in control placentas, GraphPad Prism 6 software was used to generate the graph by using the Tukey method for plotting the whiskers and outliers – calculating the interquartile range (IQR; which is the difference between $25^{th}$ and $75^{th}$ *percentiles*) and outliers (dots being $75^{th}$ *percentile* + $1.5IQR$ *or* $25^{th}$ *percentile* − $1.5IQR$) – . For the groups' comparison (IUGR, ART and pre-eclampsia) R-script were used to generate violin plots and statistics. This time, the groups were compared by using the non-parametric test of Mann-Whitney-Wilcoxon since not all studied groups were normally distributed.

#### 3.4.1.2.2 Quantitative SNP analysis

To quantify the presence of each allele after an RT-PCR (i.e. the peak height of each SNP variant in heterozygote samples), PyroMark Q96 ID software version 2.5 (Quiagen). The output of this program is the SNP variant % for each interrogated SNP.

### 3.4.1.3 Quantitative analysis by real-time qPCR

During this thesis, different quantitative studies were performed by using qPCR with a fluorochrome assay (SYBR Green Master Mix; Applied Biosystems). Our general conditions were 5.5$\mu l$ of SYBR Green (Applied Biosystems), 5$\mu l$ of diluted sample (10ng of DNA or cDNA) and 0.5$\mu l$ of each primer dilution (100ng/$\mu l$). The assays were run in triplicate and with a negative control reaction with $H_2O$ for each region. The reaction was carried out in 384 well plates in 7900HT Fast Real-Time PCR System (Applied Biosystems). The PCR conditions were shown in Figure 3.10. Only samples with 2 or more valid readings per triplicate were included. Dissociation curves were obtained at the

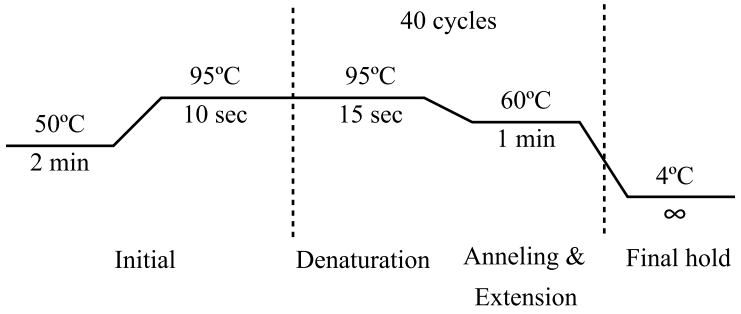



Figure 3.10: **Cycling conditions of qPCR** performed in 7900HT Fast Real-Time PCR System (Applied Biosystems).

end of each reaction to rule out the presence of primer dimers or unexpected DNA species in the reaction (list of primers in Table S15, S22 and S23). Non-template controls and a calibrator were included in each assay.

The expression analysis by qRT-PCR were normalised with *RPL19* and the different digested conditions for the enrichment 5-hmC study (see Figure 3.4) were estimated by normalising with non-digested DNA (input):

$$\Delta Ct = Ct_{studied\ gene} - Ct_{RPL19\ or\ input}$$

$$Final\ value = 2^{-\Delta Ct}$$

In the quantitative analysis of permissive histone modifications (H3K4me2m H3K4me2 and H3K9ac) for *LIN28B* and *R3HCC1* as a target DMRs, first, the output were normalised with its 50% input and then, were relative to its value in *GAPDH*:

$$\Delta Ct_{Target\ DMR} = Ct_{H3K4me2\ in\ Target\ DMR} - Ct_{input\ in\ Target\ DMR}$$

$$2^{-\Delta Ct}_{Target\ DMR}$$

$$\Delta Ct_{GAPDH} = Ct_{H3K4me2\ in\ GAPDH} - Ct_{input\ in\ GAPDH}$$

$$2^{-\Delta Ct}_{GAPDH}$$

$$Final\ value = 2^{-\Delta Ct}_{Target\ DMR} / 2^{-\Delta Ct}_{GAPDH}$$

All the results were analysed with the SDS 2.3 software (Applied Biosystems) and Excel. The final graphs were obtained by using GraphPad Prism 6 software or R-script.

### 3.4.2 Allelic RNA-seq analysis

The bioinformatic analysis to determine allelic expression in published RNA-seq datasets was performed by Dr Enrique Vidal. The abundance and genotypes of highly informative exonic SNPs within the transcripts flanking the gDMRs that maintained an intermediate methylation profile in blastocysts were called using Tophat v1.4.0 (Trapnell, Pachter, & Salzberg, 2009) (for the alignment) and Samtools v1.2 (H. Li et al., 2009) (for the filtering and allelic count) in two published single cell RNA-seq dataset for preimplantation embryos (GSE44183; GSE36552). For the purpose of this study, the data from individual cells were merged to reconstruct each embryo. In the case of the GSE44183 dataset, the embryonic genotypes were compared to the accompanying paternal exome-seq data from the sperm donor's blood sample.

### 3.4.3 Genome-wide methylation analysis

All array hybridisations were performed in the Genomic facility of the PEBC by Dr Sebastian Moran whereas the bioinformatics analyses were performed by Dr Alejandro Martin-Trujillo, Dr Franck Court, Dr Enrique Vidal and Dr Jose Hernandez-Mora.

#### 3.4.3.1 Searching for novel imprinted genes – first approach with Illumina Infinium HM450k datasets

##### 3.4.3.1.1 Sample preparation and array hybridization

We generated methylation datasets using the high-density Illumina Infinium HumanMethylation450 (HM450k) BeadChip array, which simultaneously quantify methylation at approximately 2% of all CpG dinucleotides in the human genome. BS conversion of 600ng of DNA was performed according to the manufacturer's recommendations for the Illumina Infinium Assay (EZ DNA methylation kit, ZYMO, Orange, CA). The BS-converted DNA was used for hybridisation following the Illumina Infinium HD methylation protocol. All the resulting data have been deposited in the Gene Expression Omnibus (GEO) database.

During this thesis we utilise methylation datasets previously generated and published by members of our laboratory or generated specific data: 1 maternal uniparental diploidy, 1 paternal uniparental diploidy, 4 androgenetic CHM, 67 term placenta samples (second or third trimester), 3 first trimester placenta samples, and 4 control blood samples (GSE52576); 4 peripheral blood of Fs with recessive *NLRP7* mutilations amd 5 RHM from women with *NLRP7* maternal-effect mutation (GSE66247).

#### 3.4.3.1.2 Filtering and statistical analysis

Before analysing the data, possible sources of technical biases that could influence results were excluded by applying signal background subtraction and inter-plate variation was normalised using default control probes in BeadStudio (version 2011.1_Infinium HD). The probes with a detection p-value >0.01 were discarded. The probes that lack signal values in one or more of the DNA samples analysed were also excluded. For the analysis of known imprinted domains, probes mapping to the DMRs identified by Dr Court and colleagues (Court et al., 2014) were directly analysed. Prior to screening for novel imprinted DMRs, all X chromosome CpG sites were excluded. The comparison between RHM and control placenta methylation values was performed by an in-house bioinformatic pipeline (using R-package). It was tested the difference of a minimum of 3 consecutive Infinium probes within 500bp windows via a linear model (empirical Bayes moderated p-value $< 0.01$) that provides a t-statistic, with an absolute methylation change of $> 20\%$ (beta 0.2). The result was 1232 probes (in 61 regions) which follow these criteria. All of them were hypomethylated. If we focus on the regions within CpG islands (56 regions following our criteria), most of them (48) were mapping to transcript promoters (as we can see in the Figure 3.11).

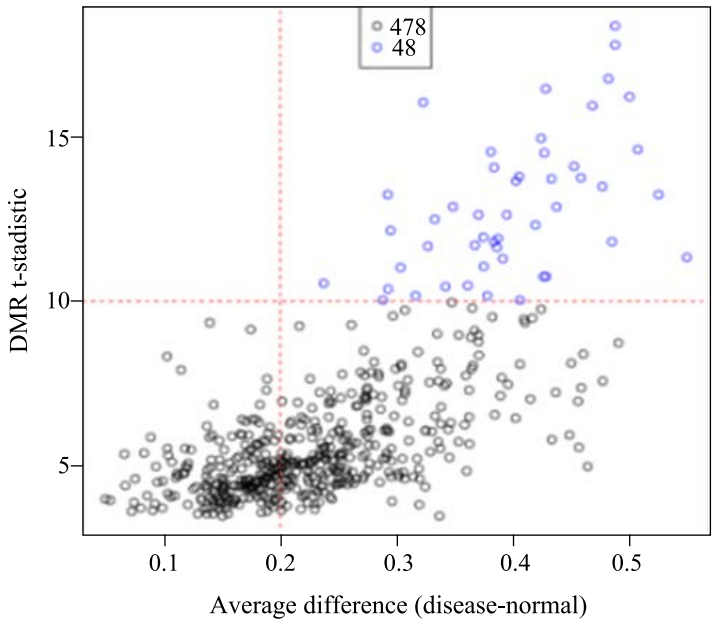



Figure 3.11: **Bioinformatic pipeline to identify new potential imprinted DMRs.** The blue circle represents the 48 regions located in a promoter and within a CpG island with an absolute methylation change of $> 20\%$ if when the methylation of HM and control placentas were compared (p-value = 0.01; t-stadistic = 10; DF = 2).

#### 3.4.3.2 Searching for novel imprinted DMRs with methyl-seq datasets

##### 3.4.3.2.1 Human methyl-seq data analysis

The methylation sequencing (methyl-seq) datasets were obtained in our laboratory and others by whole-genome bisulfite sequencing (WGBS). We analysed twenty-eight publicly available methylomes obtained from GEO or National Bioscience Database Center (NBDC) repositories. Two datasets were derived from human oocytes (JGAS00000000006), 5 from human sperm (JGAS00000000006 and GSE30340), 3 from brain (GSM913595, GSM916050, GSM1134680) 3 from CD4+ lymphocytes (GSE31263), 2 from liver (GSM916049, GSM1134681) and individual dataset from pre-implantation embryos (JGAS00000000006), placenta (GSM1134682), muscle (GSM1010986), CD34+ cells (GSM916052), sigmoid colon (GSM983645), lung (GSM983647), aorta (GSM983648), esophagus (GSM983649), small intestine (GSM983646), pancreas (GSM983651), spleen (GSM983652), adrenal (GSM1120325) and adipose tissue (GSM1010983).

Methylation calls were mapped to the hg19 genome. CpG methylation values were calculated using reads from both strands as (methylated) / (methylated + unmethylated). Only CpGs covered by at least 5 reads were considered for the analysis. For samples with duplicates, the average of methylation was used except for oocyte samples that present a low coverage. For this sample, the methylated and unmethylated calls of the two experiments were sum to calculate the methylation ratio. Using the cut off of 5 reads per CpG, the coverage of all analysis vary from 89.6% up to 96.9% of all the CpGs, except for the oocyte methylomes that cover 54.8% of CpGs sites.

##### 3.4.3.2.2 Identification of germline DMRs

The methylomes for oocyte and sperm were screened with a sliding windows approach to identify methylated and unmethylated intervals. Windows were defined as 25 consecutive CpGs and were only considered if the methylation levels were present for at least 10 CpG sites. This windows was classified methylated if $mean_{25CpGs} - 1SD_{25CpGs} > 0.75$ and unmethylated if $mean_{25CpGs} + 1SD_{25CpGs} < 0.25$. Overlapping windows with the same classification were merge and allowed us to identify 40025 unmethylated and 177787 methylated region in sperm and 118853 unmethylated and 102858 methylated regions in oocyte. A germline DMR was identified when opposite methylated regions in sperm and oocyte overlap for more than 25 CpGs and the position defined by the overlapping difference between methylated regions in sperm and oocyte.

#### 3.4.3.2.3 Identification of germline DMRs persisting in pre-implantation embryo

Intermediately methylation in blastocysts, placenta and somatic tissues were identified using the sliding windows approach with the following criteria $0.2 < mean_{25CpGs} + / - 1.5SD_{25CpGs} < 0.8$. Consecutive windows on each sample were fused to generate only a single region. A gDMR was considered to be conserved in pre-implantation embryos if the gDMR overlap with a partially methylated region in the blastocyst dataset. To identify the gDMR that persist in somatic tissues, all partially methylated region obtained in the 15 tissues were merged, and the number of samples partially methylated for each region was attributed to each region. Only regions $> 500$ bp were considered to generate the partial methylation region in tissues. To be considered as a ubiquitous gDMR, the partially methylated regions have to persist in the blastocyst, and in at least, 12 somatic tissues. Placenta-specific gDMR were identified when the partially methylated region is conserved at blastocyst stage but is not observed in additional tissues methylomes. All positional annotations (CpG islands, repeats and gene locations, etc) were obtained from UCSC web browser and genome build hg19.

#### 3.4.3.2.4 Methyl-seq analysis in other mammals

We used the methyl-seq datasets from GSE63330 that contains placenta methylation information from rhesus macaque, dog, horse, cow and mouse. The orthologous genomic intervals associated with the 551 human oocyte-derived gDMR that maintained an intermediate methylation profile throughout embryonic reprogramming and in placenta were extracted using the UCSC LiftOver function.

### 3.4.3.3 Profiling of oxBS-450K 5-hmC in the human placenta

#### 3.4.3.3.1 Sample preparation and array hybridization

For each placenta sample, $1\mu g$ of DNA was processed using the Cambridge Epigenetix TrueMethyl kit following manufacturer's instructions. Briefly, DNA was cleaned using provided magnetic beads and divided into two aliquots. Both aliquots were denatured and 1ml of oxidant solution was added to one of the aliquots (oxBS) while the other (BS) underwent a mock oxidation process by adding 1ml of ultrapure water. After incubating both aliquots at $40^{o}$C for 30 min, and centrifuged to eliminate precipitates, a BS conversion reaction was performed on the supernatant using the supplied reagents. Desulfonation and clean-up processes were applied using provided reagents before eluting the DNA in 10ml elution buffer. Then, the converted DNAs were hybridised to the HM450k array following kit instructions from Illumina. The resulting datasets have been uploaded to GEO data repository and are available under the accession number GSE93429.

#### 3.4.3.3.2 Bioinformatic analysis

Dr Jose Hernandez-Mora performed these bioinformatic procedures. Standard and TrueMethyl-CEGX oxBS HM450K datasets for cerebellum and frontal cortex were downloaded from the NCBI GEO repository (GSE74368). Enrichment of 5-hmC from in-house processed or publically available datasets was identified by subtracting the oxBS $\beta - value$ form the standard BS $\beta - value$ at each probe position on the HM450K array following quality control for probes with bad P > 0.01 in any of the samples, standard GenomeStudio normalization and removal of probes containing SNP (within the interrogation or extension base), cross-reactivity or mapping to the sex chromosomes. In total 387978 probes in common for cerebellum, frontal cortex and placenta datasets were analysed. Enriched region for 5-hmC were identified using Bumphunter with the maximum allowed gap between 2 consecutive probes being 800bp defined by 1000 permutations and a cut-off defined by the $95^{th}$ centile from the permutation distribution (see Figure 3.12 for distribution of $\Delta\beta$ values). Genomic data for MeDIP-seq and hMeDIP-seq in control placenta samples were retrieved from GSE63743. Genomic mapping of all reads and data analysis was done with Minfi22 and in-house R statistical package scripts for loci identified using oxBS-450K Bumphunter analysis.

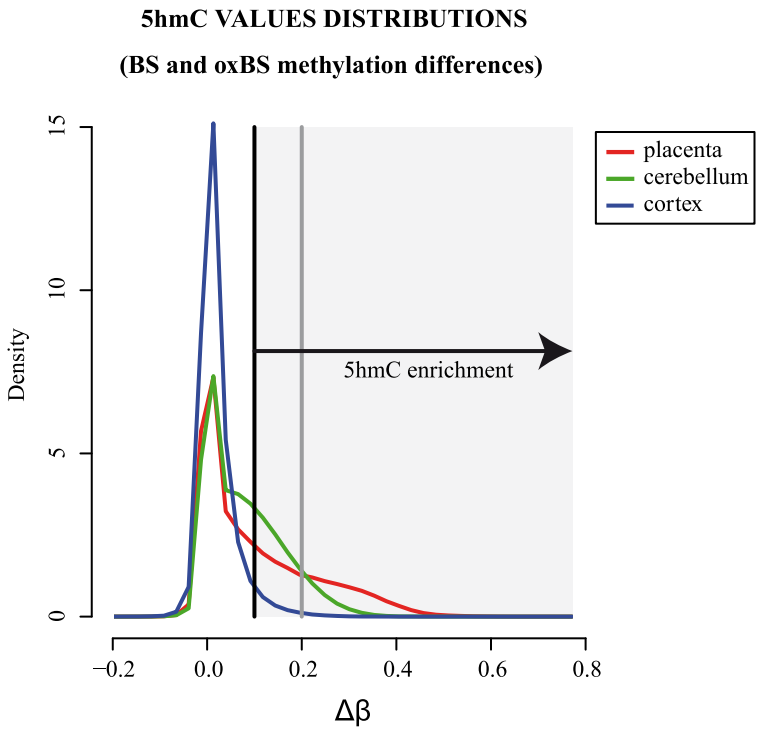



Figure 3.12: **5-hmC enrichment in biological samples.** The distribution of $\Delta\beta$ values in placenta, cerebellum and cortex samples revealing the positive skew towards 5-hmC enrichment.

# Chapter 4

# Results

## 4.1 RHMs and maternal-effect mutation in *NLRP7*

Phenotypically, RHMs are identical compared to spontaneous androgenetically derived CHMs, but with a biparental contribution to their genome. In this work, we present the genome-wide methylation profiles of both CHM of androgenetic origin and RHM from women with recessive *NLRP7* mutations. Collectively, the work described in this section was published in PLoS Genetics (2015).

### 4.1.1 RHMs from women carrying different pathogenic variants in *NLRP7*: homozygous and compound heterozygous mutations

Fresh RHM samples were initially used in this study collected from four different women (Figure 4.1): patients 1 and 2 were siblings carrying the same homozygous non-synonymous missense mutation (c.2078G>C; p.R693P), patient 3 was homozygous for a deletion that removes exons 2 to 5 (c.-39-1769_2129+228del), and finally, patient 4, was compound heterozygous for two mutations (c.2018C>G, c.2161X>T; p.S673X, p.R721W). During the course of this study patient 4 presented with second molar pregnancy that was also included in this work. Genetic analysis of the molar biopsies associated with maternal-effect *NLRP7* mutations revealed that mole RHM 4 and RHM 4.2 had inherited different mutated alleles (RHM 4 inherited the c.2161X>T at exon 6, and RHM 4.2 the c2018C>G at exon 5) (Figure 4.1B). It should be noted that the bioinformatic screening for novel imprinted DMRs was performed without the inclusion of RHM 4.2.

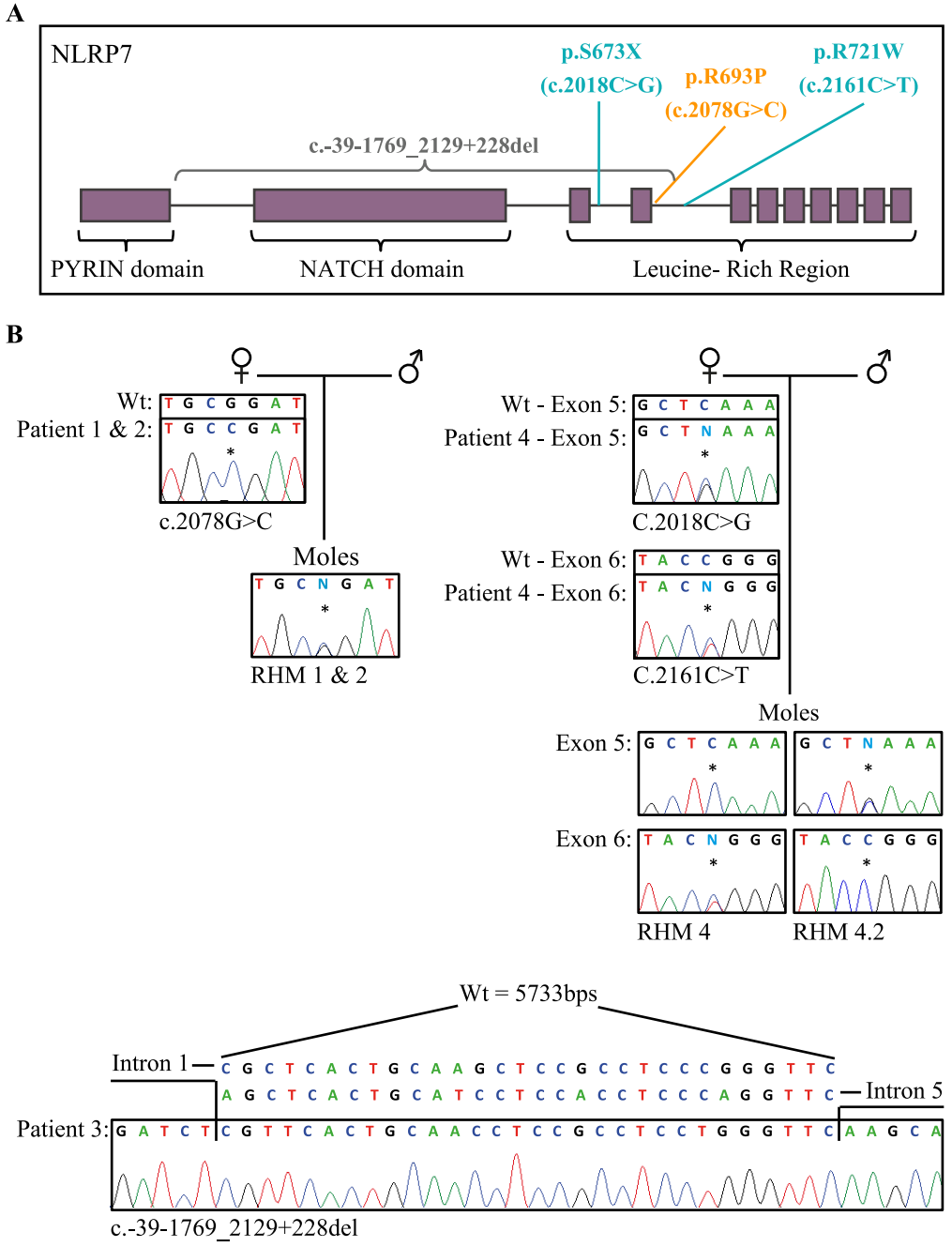



Figure 4.1: **Genotyping of *NLRP7* mutations described in our samples.** (**A**) Mapping the *NLRP7* mutations of our patients in a schematic representation of NLRP7 protein **B** recessive *NLRP7* mutations in female patients and heterozygous status in RHM samples. The asterisk* on the electropherogram highlights the position of the mutation. For patient 3, the position of the deletion is shown.

### 4.1.2 Normal DNA methylation profile in peripheral blood of females with *NLRP7* recessive mutations

Since NLRP7 protein plays a role in inflammation and other immune responses, and these women have deleterious mutations on both chromosomal copies, we tested if they had methylation anomalies in blood-derived DNA compared to healthy controls by using Illumina Infinium HM450k BeadChip array (Figure 4.2). The methylation analysis, both genome-wide and focusing on imprinted DMRs, failed to reveal any consistent differences between controls and samples from affected women.

### 4.1.3 DNA methylation profiling of imprinted DMRs in both androgenetic and biparental origin HMs

Next, we determined the genome-wide methylation profiles of the first four RHMs from females with *NLRP7* mutations using the Illumina Infinium HM450k platform. We compared the methylations status at both, ubiquitous imprinted DMRs (which are present in all embryo and adult tissues) and placenta-specific DMRs of four androgenetically derived CHMs, the four RHMs associated with maternal-effect *NLRP7* mutations and seven normal placental samples (three first trimester and four third trimester).

First, 36 known ubiquitous imprinted DMRs were assessed (Figure 4.3). This comprehensive analysis revealed that, while normal placental biopsies had partial methylation consistent with allelic methylation (except for *NNAT* and *GNAS*-AS1 promoters, already reported to be fully methylated in the placenta by Court et al. (2014)), the majority of maternally methylated DMRs presented with LOM in both androgenetic CHMs and RHMs. The only exceptions were for the *IGF1R* and *RB1* DMRs that maintain allelic methylation in both types of moles, and the *SNURF* DMR that retained partial methylation in RHM only. In addition, we observed some inter-individual differences. The *FAM50B* DMR maintained a partially methylated state in two androgenetically derived CHM and one RHM (from one of the sisters with the *NLRP7* p.R694P mutation), which also showed retained methylation at *PLAGL1* and *PEG10* DMRs.

The only paternally methylated ubiquitous DMR with probes present on this array is *H19*, that acquires methylation in the male germline and is paternally methylated in normal placenta samples. Consistent with the two copies of the sperm genome, the four androgenetic CHMs were fully methylated at this locus, whereas in RHM of biparental origin, this DMR had a profile comparable with controls. Additionally, the *ZDBF2* and *ZNF597* promoters were fully methylated in both androgenetic and biparental hydatidiform moles. This observation is consistent with the presumption that these regions acquire methylation on the paternal allele during early development under the hierarchical influence

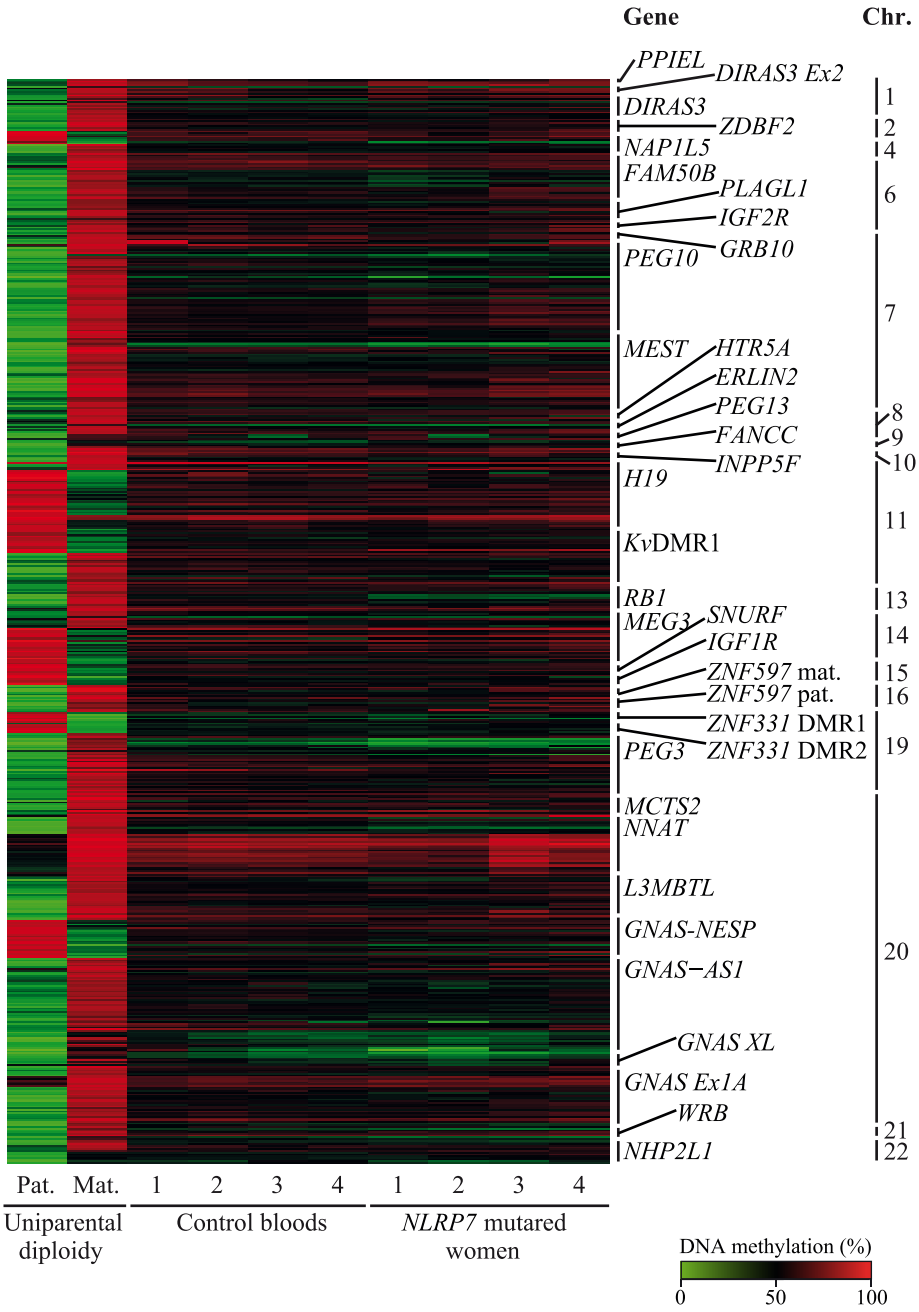



Figure 4.2: **Genome-wide methylation analysis in blood-derived DNA samples from the four females with recessive *NLRP7* mutations.** A Heatmap of the Infinium probes located within known imprinted DMRs. As controls for possible methylation changes, the profiles of reciprocal uniparental diploidy and four control blood samples are shown (*data analysis and graph generated by Dr Martin-Trujillo*).

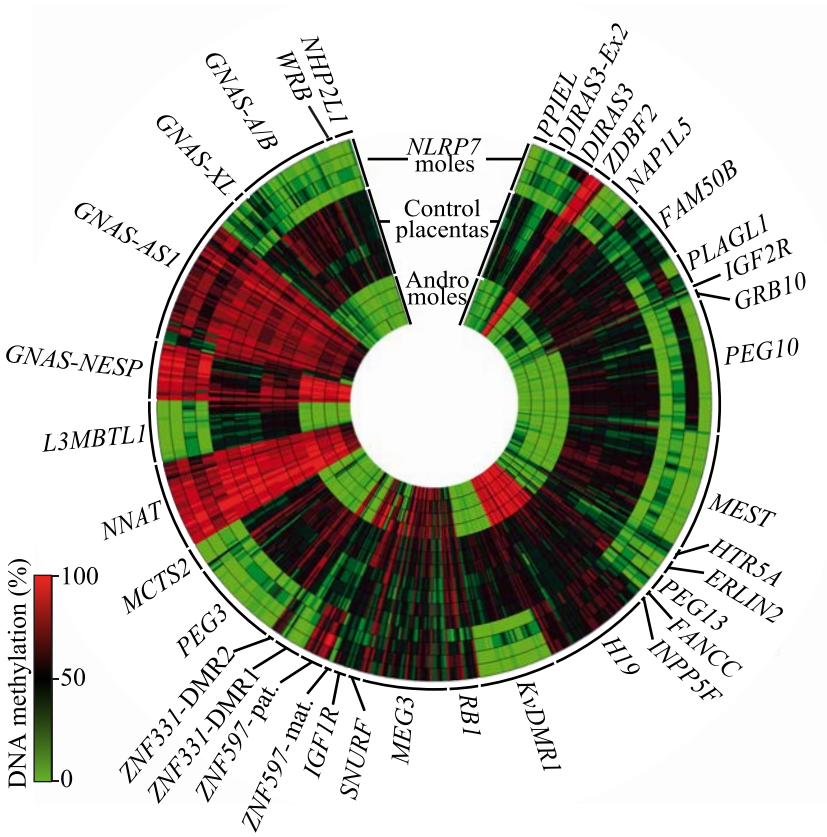



Figure 4.3: **Genome-wide methylation analysis in hydatidiform moles and control placenta samples at ubiquitous DMRs.** Circular Heatmap of the 616 Infinium array probes mapping to 36 ubiquitous imprinted DMRs. The inner circle represents the methylation values of androgenetic CHMs, the middle circles normal placental biopsies and the outer circles the RHMs (*data analysis and graph generated by Dr Martin-Trujillo*).

of the maternally methylated *GRP1*-AS and *ZNF597* DMRs, respectively (Court et al., 2014).

In addition, we analysed the methylation status of the 18 human placenta-specific imprinted DMRs that were reported prior to the start of this study (Court et al., 2014; Noguer-Dance et al., 2010; Okae et al., 2014). This revealed that, while first trimester and term placental biopsies had partial methylation, almost all androgenetically derived CHMs and RHMs presented with robust LOM, indicative of maternally methylated DMRs (Figure 4.4). The *ZFAT* DMR was the most remarkable exception, showing partial methylation for probes mapping to this interval in all androgenetic CHMs and RHMs. However, additional isolated placenta-specific DMRs in some samples showed residual methylation. This fact could be due to incompletely LOM but also to contamination of the villous trophoblast while preparing the sample.

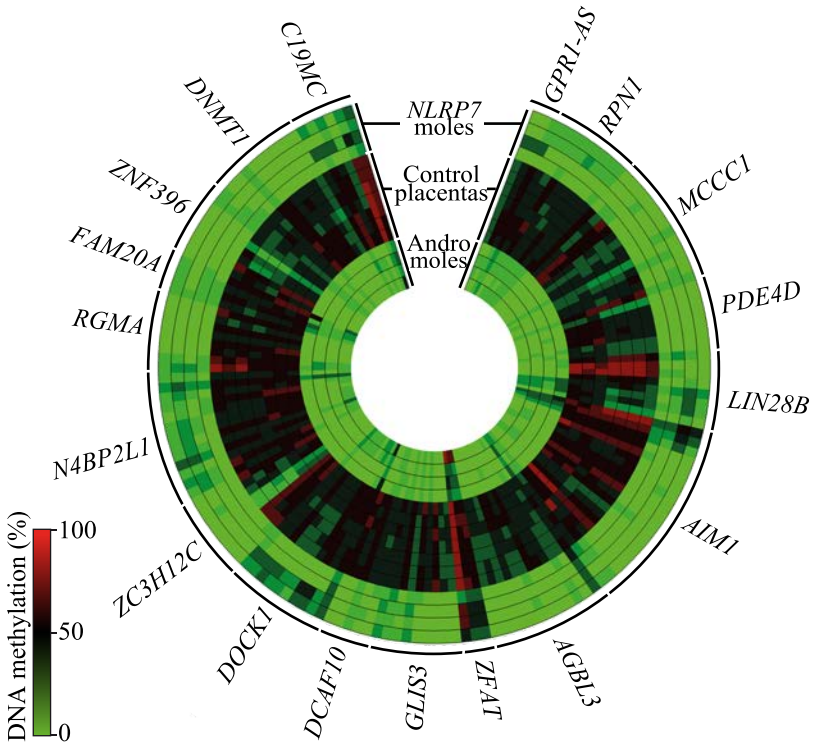



Figure 4.4: **Genome-wide methylation analysis in hydatidiform moles and control placenta samples at placenta-specific DMRs.** Circular Heatmap of the 153 Infinium array probes mapping to 18 placenta-specific imprinted DMRs. The inner circle represents the methylation values of androgenetic CHMs, the middle circles normal placental biopsies and the outer circles the RHMs (*data analysis and graph generated by Dr Martin-Trujillo*).

### 4.1.4 Confirmation of LOM at imprinted DMRs in RHMs by using targeted techniques

For some of the imprinted DMRs, the DNA methylation profile observed using the Illumina Infinium HM450k array were confirmed by both quantitative pyrosequencing, and allelic bisulphite PCR and subcloning. Additionally, DMRs not present on the array or with low probe coverage were also assessed.

Bisulphite PCRs followed by pyrosequencing and targeting the IG-DMR on chromosome 14, which acquires methylation from sperm but does not have probes on the Illumina Infinium HM450k array, revealed an imprint profile in all control placentas and RHMs (Figure 4.5A). Similarly, the *MEG3* DMR, which is regulated *in cis* by the IG-DMR, showed a similar partially methylated profile consistent with allelic methylation. Allelic bisulphite PCR at *H19* DMR revealed that the methylation observed in RHMs was on the paternal allele as all samples were informative for SNPs within the bisulphite PCR product (Figure 4.5B). The *PEG10* DMR, which was previously reported to be largely unaffected in three familiar RHM samples, shown 50% methylation at RHM 2 and RHM 4.2 (Hayward et al., 2009).

We performed similar confirmation experiments for the placenta-specific DMRs, which revealed that the methylation profile obtained from the Illumina Infinium HM450K array were reproducible, with the results from pyrosequencing being remarkably similar (Figure 4.6A). Importantly, the RHMs with almost normal methylation at *GRP1*-AS and *ZFAT* observed using the arrays was confirmed. Specifically, RHM 3 shown practically normal methylation in the three reported Type 1 placenta-specific DMRs (*GRP1*-AS, *ZFAT* and *C19MC*). The heterozygosity for SNPs in some RHM at specific regions also revealed that for several cases (*MCCC1* DMR in RHM 1, 2 and 3; *GLIS3* DMR in RHM 1 and 2; *DNMT1* DMR in RHM 2), methylation was completely absent from both parental alleles (Figure 4.6B).

Previous studies performed by members of our laboratory revealed that the *LINE-1* sequences are relatively hypomethylated in placenta compared to somatic tissues (Camprubí et al., 2013). These sequences elements contain ~12% of all CpG in the genome. Pyrosequencing analysis of these retrotransposable elements, as well as $\alpha - satellites$ and *Alu-Yb8* repeat sequences, revealed that RHMs had profile indistinguishable from normal placenta samples, suggesting that methylation at these abundant sites is normal and established/maintained independently from imprinted DMRs (Figure 4.7).

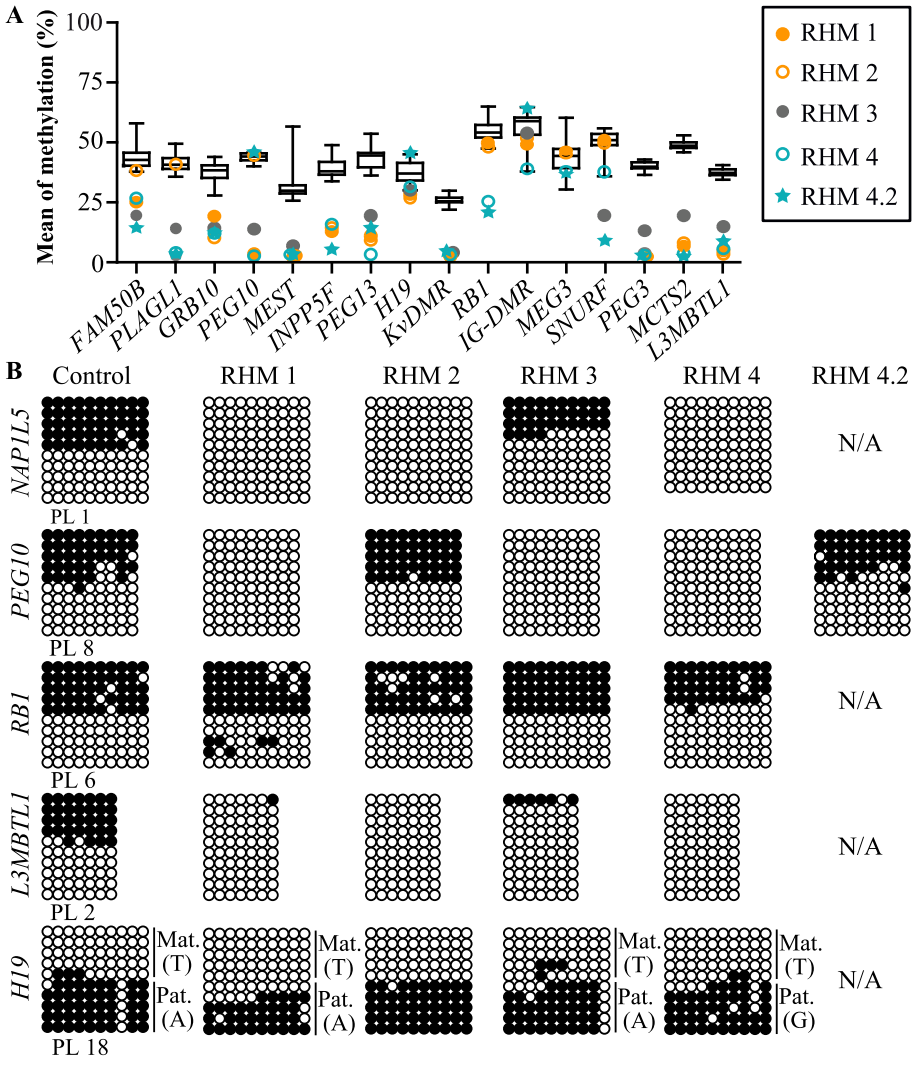



Figure 4.5: **Confirmation of the methylation profile in RHMs at ubiquitous DMRs.** (**A**) Quantitative pyrosequencing of 16 ubiquitous DMRs in the 5 RHMs. The boxplot shows the median methylation (whiskers 5-95% percentile) determined for 15 control placenta samples and the values of RHMs highlighted as different colour circles or a star. (**B**) Confirmation of LOM and allelic methylation by bisulphite PCR and subcloning in control placentas (PL) and all RHM at five ubiquitous DMRs. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

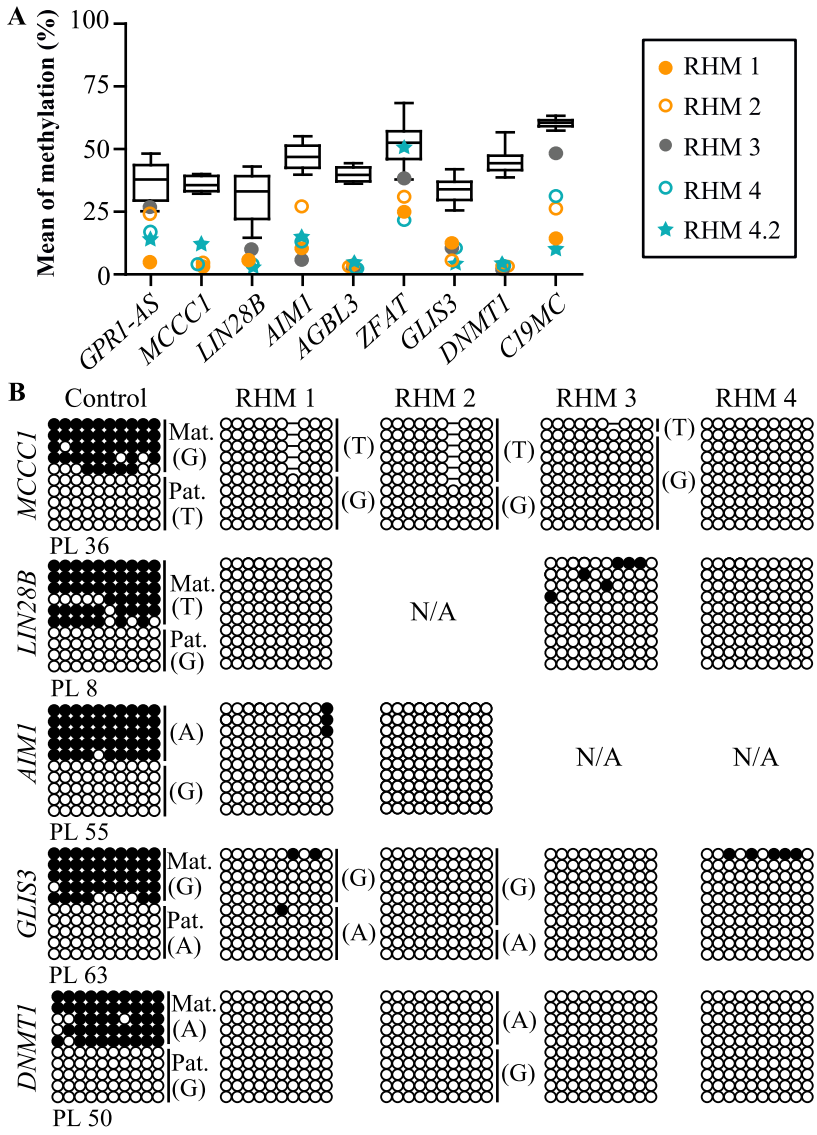



Figure 4.6: **Confirmation of the methylation profile in RHMs at placenta-specific.** (**A**) Quantitative pyrosequencing of 9 placenta-specific DMRs in the 5 RHMs. The boxplot shows the median methylation (whiskers 5-95% percentile) determined for 15 control placenta samples and the values of RHMs highlighted as different colour circles or a star. (**B**) Confirmation of LOM and allelic methylation by bisulphite PCR and subcloning in control placentas (PL) and all RHM at 5 placenta-specific DMRs. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

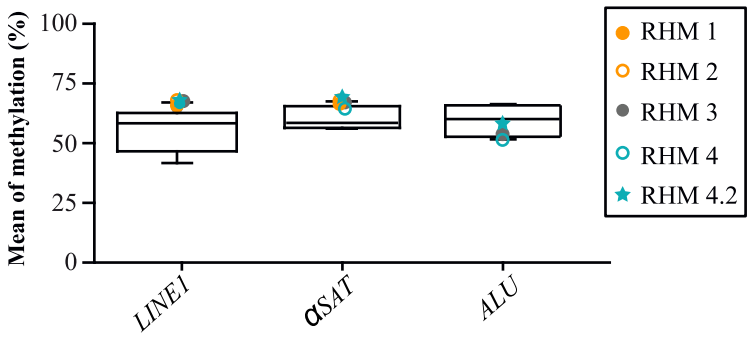



Figure 4.7: **Quantification of methylation at repeat elements in RHMs and control placentas.** Quantitative pyrosequencing of *LINE-1* $\alpha$ – *satellites* and *Alu-Yb8*, in the 5 RHMs. The boxplot shows the median methylation (whiskers 5-95% percentile) determined for 15 control placenta samples. The values of RHMs are highlighted as different colour circles or a star.

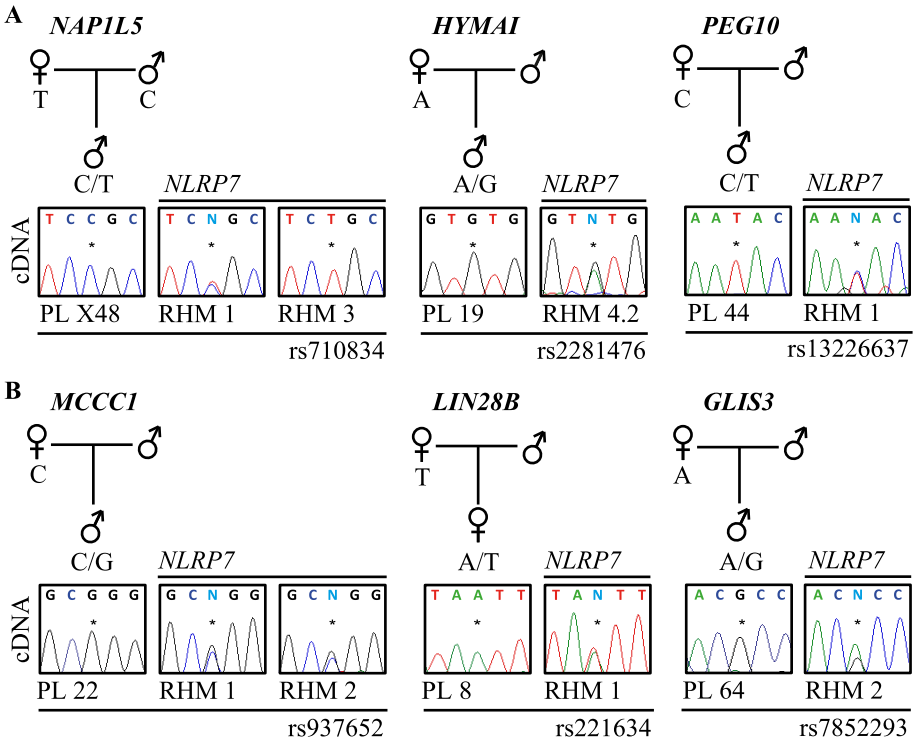



Figure 4.8: **Allelic expression profiling of RHMs at imprinted loci.** Allelic expression analysis of (**A**) Ubiquitous – *NAP1L5, HYMAI* and *PEG10* – and of (**B**) placenta-specific – *MCCC1, LIN28B* and *GLIS3* – imprinted genes in control placentas (PL) and molar tissue from females with *NLRP7* mutations (RHMs). The asterisk* on the electropherogram highlights the position of the SNP.

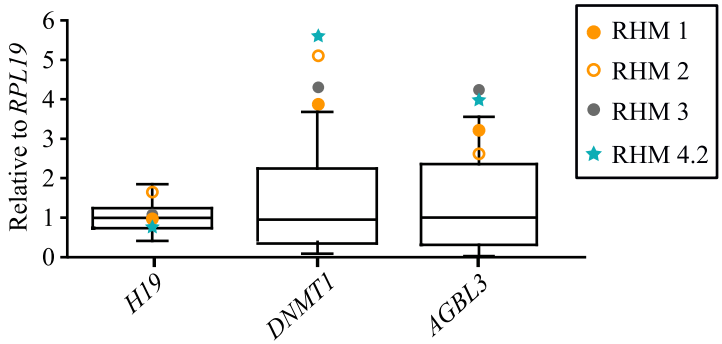



Figure 4.9: **Quantitative expression profiling of RHMs at imprinted loci.** Quantitative RT-PCR for *H19, DNMT1* and *AGBL3* in RHM samples. The boxplot shows the median expression (whiskers 5-95% percentile) determined for 15 control placenta samples. The values of RHMs are highlighted as different colour circles or a star – *experiment performed by Dr Martin-Trujillo* –.

### 4.1.5 Allele-specific expression analysis in RHM samples.

To determine if the LOM at imprinted DMRs were associated to an altered allelic expression, we performed RT-PCR that incorporated highly polymorphic SNPs in the PCR products on RNA/cDNA derived from RHM samples (Figure 4.8A). We confirmed that *NAP1L5, HYMAI* and *PEG10* transcripts were paternally expressed in control placenta samples but expressed from both alleles in the informative RHMs (i.e. the sample was heterozygous for the SNP) when they were fully unmethylated. Importantly, when stochastic allelic methylation was maintained, for example at the *NAP1L5* DMR in RHM 3, monoallelic expression was observed similar to controls. The association between allelic methylation and imprinting was also observed for the placenta-specific imprinted genes. Consistent with the lack of allelic methylation, the normally paternally expressed imprinted genes *MCCC1, LIN28B* and *GLIS3* were expressed from both parental alleles in RHMs (Figure 4.8B).

Furthermore, as an indirect indication of biallelic expression, qRT-PCR revealed an increased expression of *DNMT1* and *AGBL3* compared to normal placenta samples coherent with expression from both chromosomes (Figure 4.9). Expression was within the normal range for *H19* which is consistent with the maintained paternally derived methylation at this DMR.

## 4.2 Searching for novel imprinted genes

During the course of this doctoral thesis, two different high-throughput methylation screens have been performed with the aims of identifying novel placenta-specific maternally

methylated regions. The two different approaches, one based on comparing Illumina Infinium HM450k datasets for RHM and control placenta samples, and the second based on a screening to identify stable inherited methylation from the gametes, were published as separate works.

To respect the timeline of these discoveries, I will first focus on the analysis associated with RHMs from *NLRP7*-mutated women by using the Illumina Infinium HM450k platform, which was published in PLoS Genetics (2015) (Table 4.1 for the confirmed regions).

Table 4.1: **Imprinted DMRs and associated transcripts confirmed in our first analysis.**

| Chr | Start | End | DMR | CpGI | GenePos | Confirm. | Expres. |
|---|---|---|---|---|---|---|---|
| chr1[1] | 6685554 | 6686900 | *THAP3* | Y[2] | Prom[3] | *Hpa*II[4] & Cloning[5] | |
| chr1 | 51796019 | 51797237 | *TTC39A* | Y | Prom | Cloning | |
| chr1 | 60391609 | 60393337 | *CYP2J2* | Y | Prom | *Hpa*II | |
| chr2 | 62733299 | 62734454 | *TMEM17* | Y | Prom | *Hpa*II | |
| chr2 | 73519359 | 73521905 | *EGR4* | Y | Prom | *Hpa*II | |
| chr2 | 118943937 | 118944846 | *AC093901* | Y | Prom | *Hpa*II | |
| chr3 | 128336544 | 128337642 | *RPN1* | Y | Interg[6] | Cloning | |
| chr3 | 196728281 | 196731463 | *MFI2-AS1* | Y | Prom | *Hpa*II | |
| chr4 | 102711842 | 102713967 | *BANK1* | Y | Prom | *Hpa*II | |
| chr5 | 90976 | 92437 | *CTD-2231H16* | Y | Prom | *Hpa*II | |
| chr5 | 95066854 | 95069104 | *RHOBTB3* | Y | Prom | Cloning | 2 pat[7] 8 mono[8] |
| chr5 | 139493865 | 139494366 | *PURA* | Y | Prom | Cloning | |
| chr5 | 176055855 | 176058096 | *SNCB* | Y | Prom | Cloning | |
| chr6 | 14115948 | 14118251 | *CD83* | Y | Prom | *Hpa*II | 1 mono |
| chr7 | 11870826 | 11871746 | *THSD7A* | N[9] | Prom | *Hpa*II | |
| chr7 | 12610078 | 12611415 | *SCIN* | Y | Prom | *Hpa*II & Cloning | 6 pat 2 mono 3 bi[10] |

[1] *Chr*: Chromosome (Coordinates from methyl-seq profile information (which represents the complete DMR, including all Illumina Infinium HM450k probes detected for each region)
[2] *CpGI (Y)*: There is a CpG island within the DMR
[3] *GenePos (Prom)*: The DMR is located within a promoter region
[4] *Confirmation (*Hpa*II)*: The DMR was confirmed by *Hpa*II
[5] *Confirmation (Cloning)*: The DMR was confirmed by bisulhite PCR and subcloning
[6] *GenePos (Interg)*: The DMR is located in an intergenic region
[7] *Expression (pat)*: If the gene that gives the name to the DMR is paternally expressed.
[8] *Expression (mono)*: If the gene that gives the name to the DMR is monoallelically expressed.
[9] *CpGI (N)*: There is no CpG island within the DMR
[10] *Expression (bi)*: If the gene that gives the name to the DMR is biallelically expressed.

| Chr | Start | End | DMR | CpGI | GenePos | Confirm. | Expres. |
|---|---|---|---|---|---|---|---|
| chr7 | 43151168 | 43152684 | *HECW1* | Y | Prom | *Hpa*II | 1 pat<br>1 mono<br>1 bi |
| chr7 | 65877222 | 65879491 | *GS1-124K5* | Y | body[11] | *Hpa*II | |
| chr7 | 106297905 | 106303036 | *CCDC71L* | Y | Prom | Cloning | 2 pat<br>2 bi |
| chr10 | 23216688 | 23217883 | *ARMC3* | N | Prom | Cloning | |
| chr10 | 71891605 | 71893315 | *AIFM2* | Y | Prom | Cloning | |
| chr10 | 103533860 | 103537185 | *FGF8* | Y | Prom | Cloning | |
| chr11 | 7693877 | 7696363 | *CYB5R2* | Y | Prom | *Hpa*II &<br>Cloning | |
| chr11 | 10560091 | 10564029 | *RNF141* | Y | Prom | Cloning | 10 bi |
| chr12 | 22486690 | 22488547 | *ST8SIA1* | Y | Prom | *Hpa*II<br>Cloning | 3 mono |
| chr12 | 65218246 | 65219685 | *TBC1D30* | Y | Prom | *Hpa*II | 3 bi |
| chr12 | 65512862 | 65517731 | *WIF1* | Y | Prom | Cloning | |
| chr15 | 45314075 | 45316470 | *SORD* | Y | Prom | *Hpa*II | |
| chr15 | 79381827 | 79384742 | *RASGRF1* | Y | Prom | Cloning | 2 pat<br>2 bi |
| chr16 | 48397923 | 48400894 | *SIAH1* | Y | Prom | Cloning | 11 bi |
| chr16 | 66637337 | 66640186 | *CMTM3* | Y | Prom | *Hpa*II | 1 pat<br>3 mono<br>2 bi |
| chr16 | 68572176 | 68574654 | *ZFP90* | Y | Prom | – | 3 pat<br>2 bi |
| chr17 | 258676 | 261180 | *C17ORF97* | Y | Prom | *Hpa*II | |
| chr18 | 2905116 | 2906995 | *EMILIN2* | Y | Prom | *Hpa*II | |
| chr22 | 24551280 | 24553640 | *CABIN1* | Y | Prom | *Hpa*II &<br>Cloning | |

### 4.2.1 A comparison of RHM and control placenta samples identifies novel imprinted regions.

To expand our methylation analysis, Dr Alejandro Martin-Trujillo and Dr Enrique Vidal performed unbiased screens for additional loci with aberrant methylation in the RHM

[11] *GenePos (body)*: The DMR is located within a gene body

samples following the criteria detailed in the material and methods' section. In short, we were looking for loci with three consecutive probes that presented with a change in methylation of +/- 20% (0.2$\beta$). This analysis revealed 61 regions, 56 of which are CpG islands, 88% mapping to transcript promoters (Figure 4.10). Surprisingly, all candidate regions identified were partially methylated in normal placental biopsies, consistent with allelic imprinted methylation, and devoid of methylation in androgenetically derived CHMs which were used as an additional control. We did not observe any region with gains in methylation in molar samples compared to normal placenta biopsies.

We utilised two different molecular approaches to confirm if the observed partial methylation was present on one parental allele: standard allele-specific bisulphite PCR with subcloning and methylation-sensitive genotyping. This second assay compares the genotype profiles from genomic DNA with that subjected to *Hpa*II endonuclease digestion in which heterozygous samples would be reduced to homozygosity if allelic methylation is present. For both methods to be informative, heterozygous SNPs need to be identified, which involved a huge amount of genotyping. In total, ~8,100 individual genotype assays were performed to identify informative samples for following experiments.

For the 61 candidate regions selected for confirmation, 28 had highly informative SNPs that allowed parental origin of methylation to be determined. Some heterozygote placenta samples were selected for each regions and its genomic DNA were digested with the methylation-sensitive *Hpa*II endonuclease, which only cut its recognition sequence in absence of DNA methylation (Figure 4.11B and Figure 3.5B in material and methods' section for methylation-sensitive genotyping). Therefore, any PCR generated using digested DNA template would originate from methylated DNA.

One limitation of the methylation-sensitive genotyping assay is that only the CpG with *Hpa*II restriction site (C**CG**G) is assessed. To ensure that multiple neighbouring CpGs also show allelic methylation we also performed a standard allele-specific bisulphite PCR and subcloning for some regions, with the informative SNP incorporated in the PCR amplicon (Figure 4.11C and S2). Using this second method we confirmed allele-specific methylation for eight additional loci. However, we also confirmed strand-specific methylation for five regions for which we did not have heterozygous samples (*RHOBTB3, PURA, CCDC71L, FGF8* and *WIF1*) (Figure S2).

In total, we confirmed the presence of maternal methylation in multiple placenta samples in 22 DMRs, with a further seven regions being allelically methylated with parental genotypes being uninformative (Figure 4.11, Figure S1 and Figure S2).

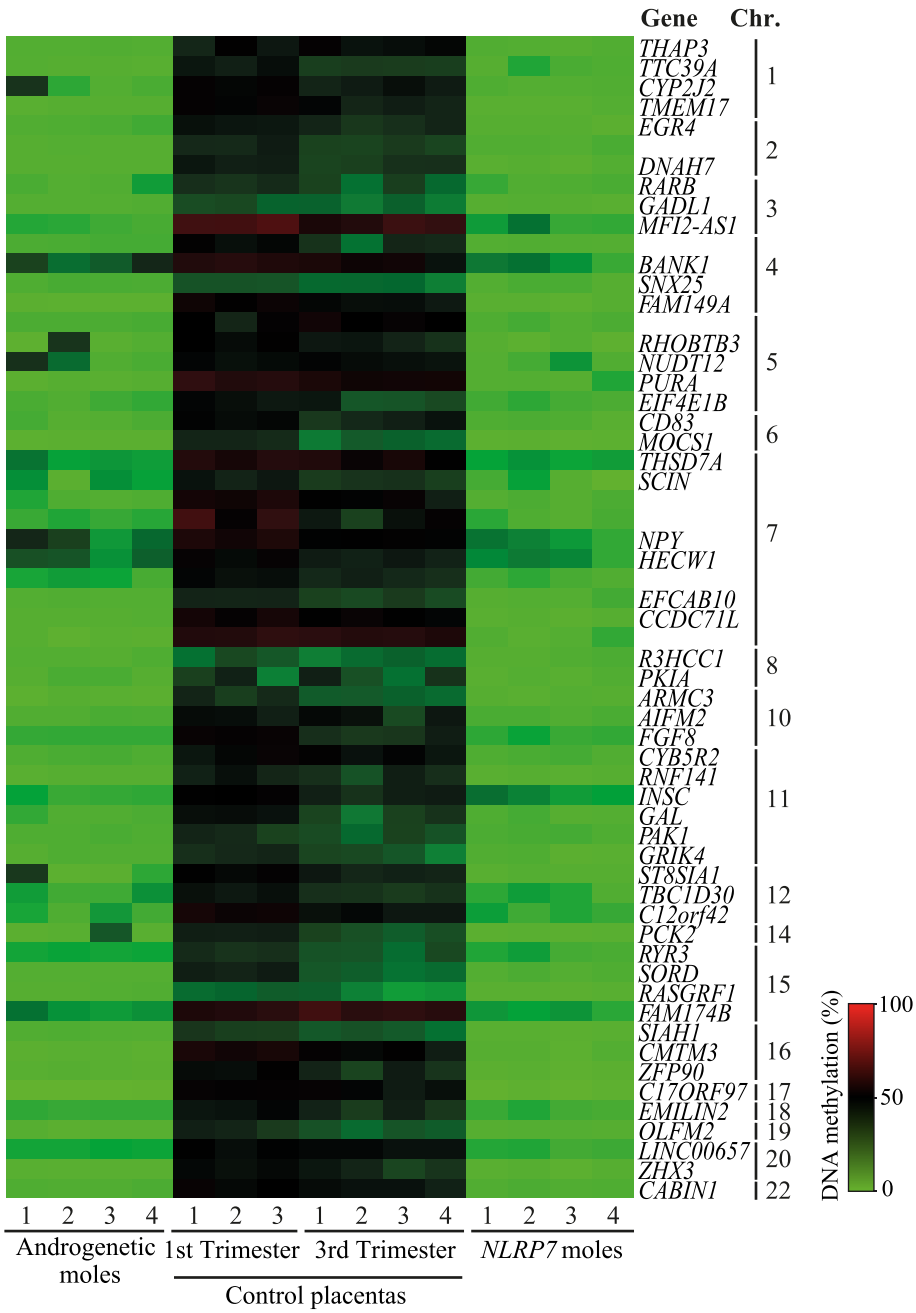



Figure 4.10: **Identification of novel placenta-specific DMRs by comparing RHMs and control placentas.** Heatmap for the $\beta_{mean}$ of the Infinium probes with a methylation difference (>20%, minimum of 3 consecutive probes) in RHMs compared to control placenta samples (*data analysis and graph generated by Dr Martin-Trujillo*).

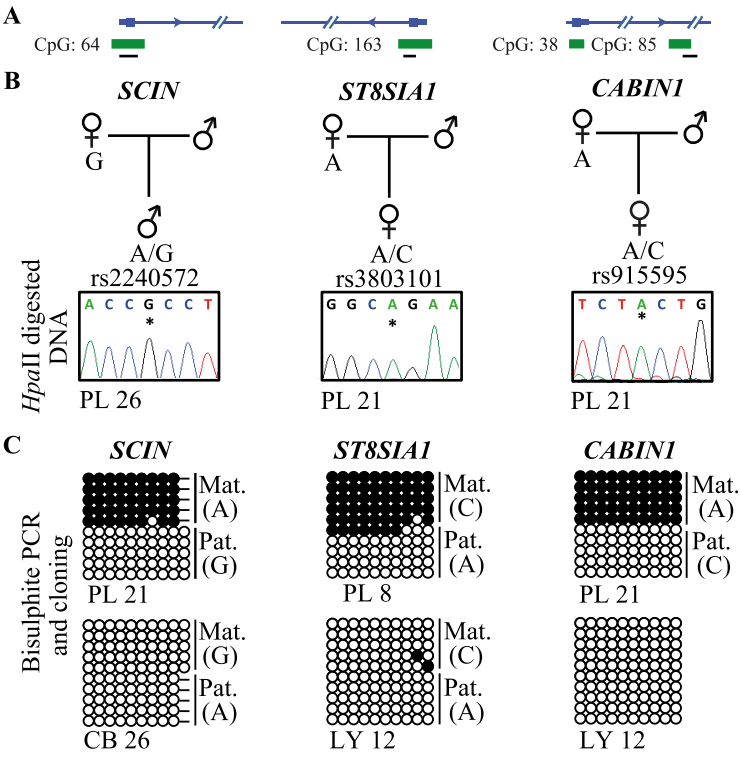



Figure 4.11: **Confirmation of allelic methylation at novel placenta-specific DMRs (array candidates).** (**A**) Schematic representation of *SCIN, ST8SIA1* and *CABIN1* transcripts (blue line) and the location of our PCR (black line) at placenta-specific DMR within CpG island (green rectangle) (**B**) Methylation profiles as determined by methylation-sensitive genotyping. The asterisk* on the electropherogram highlights the position of the SNP. (**C**) Confirmation of methylation profile in placenta (*PL*) and other tissues – cord blood (*CB*) and lymphocytes (*LY*) – by bisulphite PCR and sub-cloning at *SCIN, ST8SIA1* and *CABIN1* promoters. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

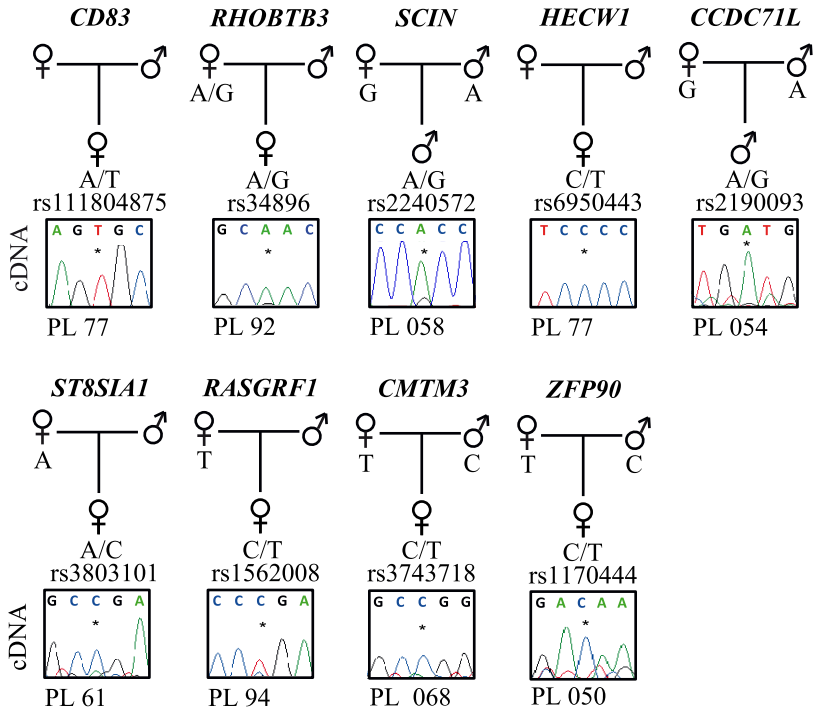



Figure 4.12: **Allele-specific RT-PCR analysis of array candidate placenta-specific imprinted genes.** Confirmation of monoallelic expression of *CD83, RHOBTB3* and *HECW1* and paternal expression of *SCIN, CCDC71L, ST8SIA1, RASGRF1, CMTM3* and *ZFP90* in term placenta samples (PL). The asterisk* on the electropherogram highlights the position of the SNP.

### 4.2.2 Imprinted paternal expression is associated with some, but not all placenta-specific DMRs.

The main biological significance of allele-specific methylation is allele-specific RNA expression, which in the case of maternally methylated DMRs at promoter regions, in general, is predicted to dictate paternal expression. We subsequently determined the allelic expression for a subset of transcripts (whose promoter contains a placenta-specific DMR) that contained highly polymorphic exonic SNPs. Allele-specific RT-PCR confirmed paternal expression or monoallelic expression in a polymorphic fashion in multiple placental biopsies of *CD83, RHOBTB3, SCIN, HECW1, CCDC71L, ST8SIA1, RASGRF1, CMTM3* and *ZFP90* (Figure 4.12). In situations where monoallelic expression was uninformative due to maternal DNA also being heterozygous or in the cases that we did not have maternal samples DNA, we tried to use the same SNP for the methylation and expression experiments. In many cases, the methylated allele was the repressed one, confirming an association between methylation and expression.

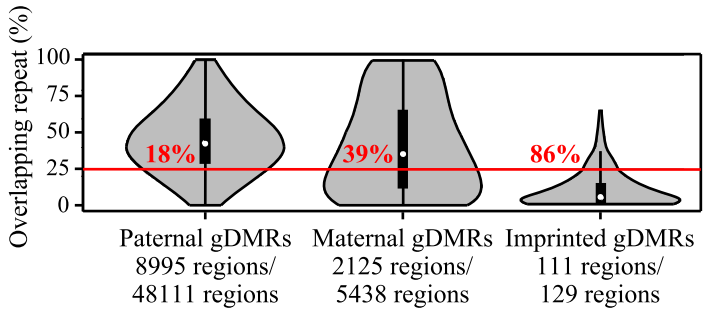



Figure 4.13: **Repetitive sequences at paternal gDMRs, maternal gDMRs and known imprinted gDMRs.** Violin plots classified as the % repetitive sequence for gDMRs paternally and maternally methylated and known imprinted gDMRs. In red, the % of unique sequences, defined as <25% of overlapping with repetitive sequences (*data analysis and graph generated by Dr Court and Dr Vidal*).

## 4.3 Novel imprinted DMRs from methyl-seq datasets

For the first two years of my thesis I had been confirming the extent placenta-specific imprinting in the human genome as a result of our screening using the Illumina Infinium HM450k array. This analysis has several limitations, most notably the coverage of the array used and the position of the probes. We therefore wished to perform an unbiased screen based on methyl-seq datasets generated in-house or obtained from GEO or NBDC databases.

Before fertilisation occurs, the oocyte and sperm have thousands of oppositely methylated regions. Most of this germline methylation is erased during the post-fertilisation epigenetic reprogramming. In a bioinformatic analysis performed by Dr Enrique Vidal and Dr Franck Court, we first focused on regions with opposing methylation between oocyte and sperm, which revealed 5438 regions exclusively methylated in the oocyte and 48111 in the sperm. A high proportion of regions methylated in sperm and hypomethylated in oocytes were intergenic or map to repeat elements, consistent with previous observations (Figure 4.13) (Okae et al., 2014). In contrast, the oocyte-derived specific methylated intervals were more uniformly distributed throughout the genome than the sperm-derived regions and often overlapped promoter CpG islands.

Since imprinted DMRs are known to be protected from post-fertilisation demethylation, we focussed on the methylation state of these regions at blastocyst. Of these regions, 80% of the oocyte-derived DMRs (4352/5438) remain partially methylated at the blastocyst stage – which from now on will be referred to as maternal gDMRs –, while only 1% of the sperm-derived DMRs (517/48111) – paternal gDMRs – survived to this reprogramming event (Figure 4.14). This result is consistent with the fact that post-fertilisation epigenetic reprogramming is thought to be mainly an active mechanism for the sperm-derived

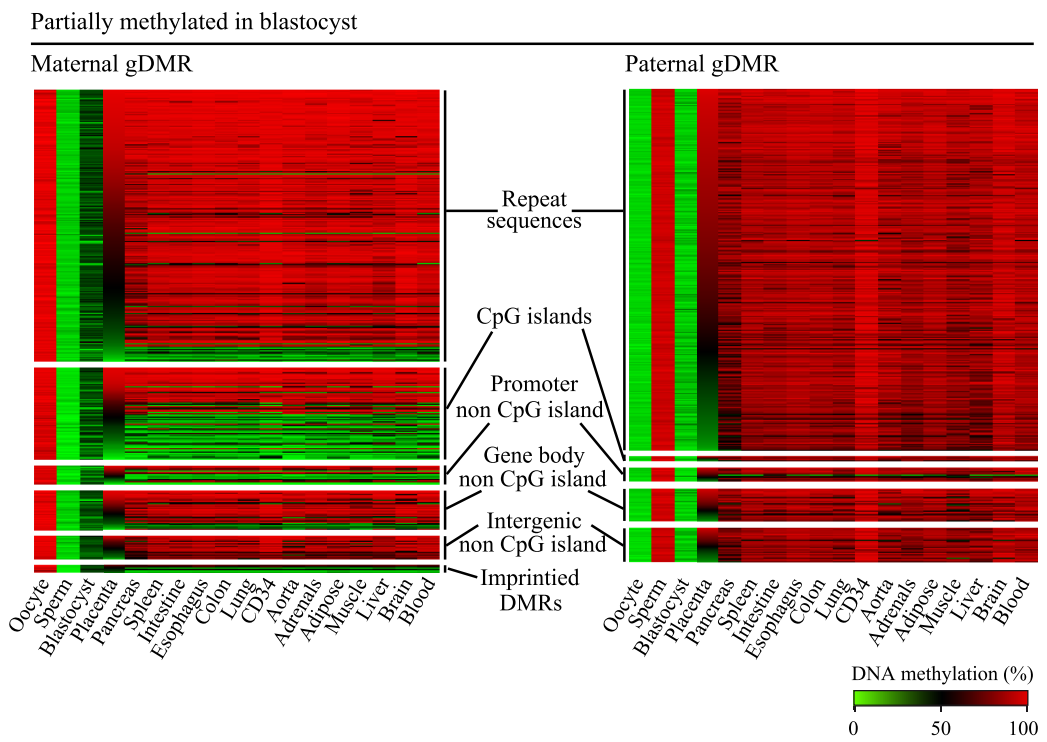



Figure 4.14: **Methylation profiling of human gametes, embryos and tissues.** Heatmap of germline methylation differences that remain partially methylated at the blastocyst stage (left, the 4352 maternal gDMRs; right, the 517 paternal gDMRs). Arranged in descending order according to their placenta methylation profile (from more to less methylated) (*data analysis and graph generated by Dr Court*).

methylation, whilst for the oocyte-derived methylation is largely subject to passive, replication-dependent demethylation.

This epigenetic reprogramming is particularly evident when the size of the gDMRs surviving to the blastocyst stage is taken into consideration. In total, ~7 Mb of the human genome encompasses oocyte-derived gDMRs of which 74% is partially methylated in preimplantation embryos, whereas 2.7 Mb is covered by sperm-derived gDMRs of which, only 11% is partially methylation at the same developmental stage (Figure 4.15). Therefore, maternal gDMRs are losing methylation after the blastocyst stage whereas the paternal gDMRs is largely reprogrammed during cleavage/preimplantation stages.

Numerous studies have shown that gDMR which persist uniformly in somatic tissues acts as ICRs. To determine if additional imprinted gDMRs, which may act as ICRs, are present in the human genome, we assessed the methylation profile of the oocyte and sperm-derived gDMRs in methyl-seq datasets in blastocysts, placenta and 14 different somatic tissues (Figure 4.16A). For this analysis, we searched for gDMRs persisting to the blastocysts stage and present in numerous somatic tissues, which would reveal possible

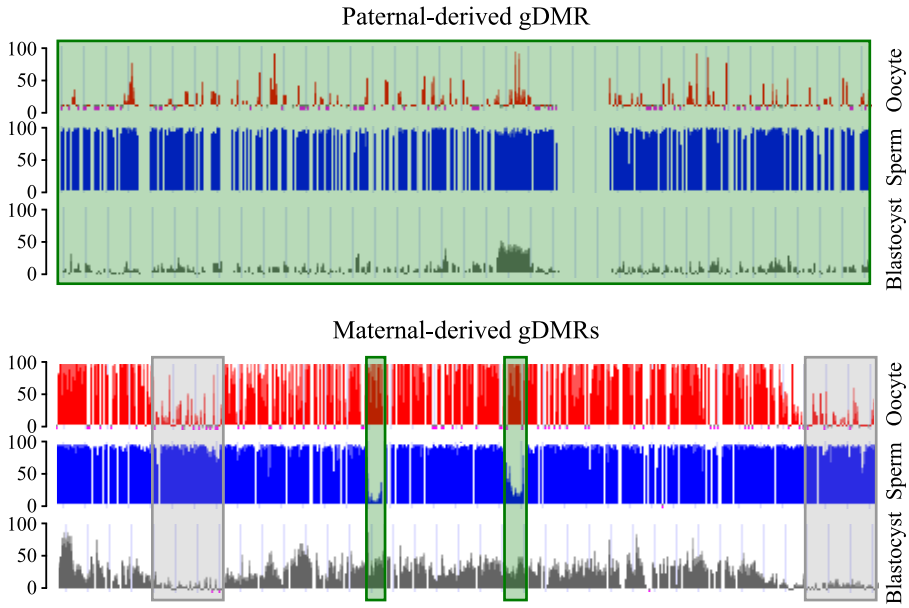



Figure 4.15: **Example of paternal and maternal derived gDMR.** Paternal-derived gDMRs (top, green box) shows large regions of fully methylated at sperm methyl-seq dataset, with no methylation at oocyte methyl-seq dataset and partially methylated in small CpG sites at the blastocyst methyl-seq dataset. In the case of the maternal-derived gDMRs (bottom, green boxes) are smaller regions fully methylated at oocyte methyl-seq dataset, unmethylated at sperm methyl-seq dataset and partially methylated in all the DMR at the blastocyst methyl-seq dataset. Most of the sperm-derived regions lose its methylation at blastocyst stage (bottom, grey boxes). The vertical black lines in the methyl-seq tracks represent the mean methylation value for individual CpG dinucleotides (blue for methylation at sperm, red for methylation at oocyte and grey for methylation at blastocyst stage).

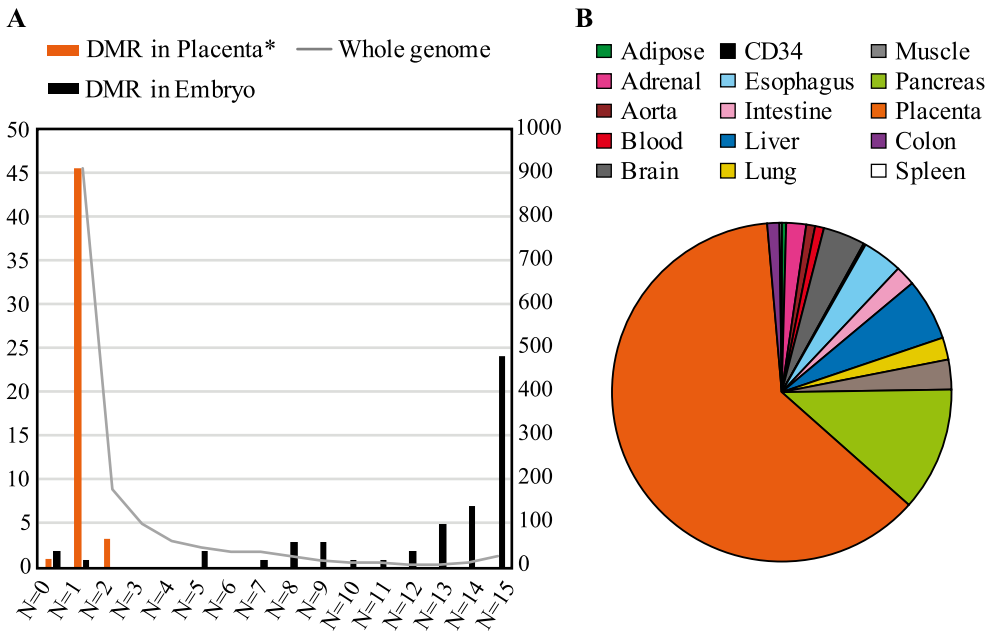



Figure 4.16: **Analysis of tissue-specific maintenance of germline methylation in different tissues.** (**A**) A bar graph showing maintaining as gDMR in different tissues. The bars represent the profiles of known ubiquitous (black) and placenta-specific (orange) DMRs with numbers corresponding to the left y-axis. The superimposed line graph represent the pattern of all remaining germline difference that is maintained to the blastocyst stage and corresponding to the right y-axis. DMR in placenta* corresponds to regions identified by Court et al. (2014) and the approach described in the previous results' section of this thesis by using Illumina Infinium HM450k platform (Sanchez-Delgado et al. 2015). (**B**) A pie graph showing the distribution of individual tissues maintaining a partially methylated profile (*data analysis and graph generated by Dr Court*).

novel ubiquitous DMRs. In addition, we also screened for loci partially methylated in only one tissue to determine if there are germline-derived DMRs maintained in a tissue-specific fashion (Figure 4.16B). The results of this analysis suggest that the majority of ubiquitous gDMRs have already been characterised, whilst pointing to the placenta as the main source of tissue-specific DMRs (Table 4.2 for the confirmed regions).

Table 4.2: **Imprinted DMRs and associated transcripts confirmed in our second analysis.**

| Chr | Start | End | DMR | CpGI | GenePos | Confirmation | Expression |
|---|---|---|---|---|---|---|---|
| chr7[12] | 138348118 | 138349069 | *SVOPL* | Y[13] | Promoter[14] | Cloning[15] | 2 mat[16] |
| | | | | | | | 4 mono[17] |
| chr9 | 98074294 | 98077105 | *FANCC* | Y | Gene_body | Cloning | |
| chr2 | 46655894 | 46657384 | *TMEM247* | N[18] | Promoter | *Hpa*II[19] & | |
| | | | | | | Cloning | |
| chr1 | 36348678 | 36351041 | *AGO1* | Y | Promoter | Cloning | 2 pat[20] |
| | | | | | | | 1 mono |
| chr1 | 181285914 | 181288964 | *CACNA1E* | Y | Intergenic[21] | *Hpa*II & | |
| | | | | | | Cloning | |
| chr2 | 62733299 | 62734454 | *TMEM17* | Y | Promoter | *Hpa*II | |
| chr2 | 74345587 | 74348298 | ncRNA | Y | Intergenic | *Hpa*II & | 2 pat |
| | | | (*TET3*) | | | Cloning | 1 mono |
| chr2 | 229043530 | 229047055 | *SPHKAP* | Y | Promoter | *Hpa*II | |
| chr3 | 21790873 | 21792537 | *ZNF385D* | N | Promoter | Cloning | |
| chr3 | 49312102 | 49314590 | *C3orf62/* | Y | Gene_body[22] | *Hpa*II | 1 pat |
| | | | *USP4* | | | | 1 mono |
| | | | | | | | 1 bi[23] |
| chr3 | 128718677 | 128720535 | *EFCC1* | Y | Promoter | *Hpa*II | |
| chr3 | 192124589 | 192127412 | *FGF12* | Y | Promoter | Cloning | |
| chr4 | 656576 | 658497 | *PDE6B* | Y | Promoter | *Hpa*II | |
| chr4 | 2819393 | 2820378 | *SH3BP2* | Y | Promoter | *Hpa*II & | 4 pat |
| | | | | | | Cloning | 2 mono |
| | | | | | | | 5 bi |
| chr4 | 4576220 | 4577911 | *STX18-AS1* | Y | Gene_body | *Hpa*II & | |
| | | | | | | Cloning | |
| chr4 | 8582615 | 8583642 | *GPR78* | Y | Promoter | *Hpa*II & | |
| | | | | | | Cloning | |

[12] *Chr*: Chromosome (Coordinates from methyl-seq profile information (which represents the complete DMR, including all Illumina Infinium HM450k probes detected for each region)
[13] *CpGI (Y)*: There is a CpG island within the DMR
[14] *GenePos (Prom)*: The DMR is located within a promoter region
[15] *Confirmation (Cloning)*: The DMR was confirmed by bisulhite PCR and subcloning
[16] *Expression (mat)*: If the gene that gives the name to the DMR is maternally expressed.
[17] *Expression (mono)*: If the gene that gives the name to the DMR is monoallelically expressed.
[18] *CpGI (N)*: There is no CpG island within the DMR
[19] *Confirmation (*Hpa*II)*: The DMR was confirmed by *Hpa*II
[20] *Expression (pat)*: If the gene that gives the name to the DMR is paternally expressed.
[21] *GenePos (Interg)*: The DMR is located in an intergenic region
[22] *GenePos (body)*: The DMR is located within a gene body
[23] *Expression (bi)*: If the gene that gives the name to the DMR is biallelically expressed.

| Chr | Start | End | DMR | CpGI | GenePos | Confirmation | Expression |
|---|---|---|---|---|---|---|---|
| chr4 | 187064415 | 187066853 | *FAM149A* | Y | Promoter | Cloning | 1 pat |
| chr4 | 93225276 | 93227244 | *GRID2* | Y | Promoter | *Hpa*II & Cloning | |
| chr4 | 154709200 | 154715220 | *SFRP2* | Y | Promoter | *Hpa*II | |
| chr6 | 129249737 | 129253478 | *LAMA2* | Y | Gene_body | Cloning | 8 bi |
| chr6 | 39901167 | 39902213 | *MOCS1* | Y | Promoter | Cloning | 1 pat<br>1 mono<br>1 bi |
| chr8 | 1321333 | 1322521 | *DLGAP2* | Y | Intergenic | *Hpa*II & Cloning | |
| chr8 | 23145610 | 23146931 | *R3HCC1* | Y | Promoter | *Hpa*II & Cloning | 3 pat<br>1 mono<br>4 bi |
| chr8 | 142137462 | 142140073 | *DENND3* | Y | Promoter | *Hpa*II | 6 bi |
| chr9 | 86151350 | 86154260 | *FRMD3* | Y | Promoter | *Hpa*II | |
| chr10 | 65221271 | 65225444 | *JMJD1C* | Y | Promoter | Cloning | 1 pat<br>2 mono |
| chr11 | 2812081 | 2814291 | *KCNQ1* | Y | Gene_body | *Hpa*II & Cloning | |
| chr11 | 77122000 | 77122997 | *PAK1* | Y | Promoter | Cloning | 2 mono<br>2 bi |
| chr11 | 132813023 | 132815361 | *OPCML* | Y | Promoter | Cloning | |
| chr12 | 2338254 | 2340879 | *CACNA1C* | Y | Gene_body | Cloning | |
| chr12 | 2800562 | 2800919 | *CACNA1C-AS1* | Y | Promoter | *Hpa*II | |
| chr13 | 102568126 | 102569981 | *FGF14* | Y | Promoter | *Hpa*II & Cloning | 2 bi |
| chr14 | 73712301 | 73712670 | *PAPLN-AS1* | Y | Promoter | *Hpa*II | 1 pat<br>3 mono<br>1 bi |
| chr15 | 33602942 | 33605250 | *RYR3* | Y | Promoter | *Hpa*II | |
| chr18 | 77376189 | 77378625 | *RP11-567M16.3* | Y | Intergenic | *Hpa*II | |
| chr19 | 13614882 | 13618186 | *CACNA1A* | Y | Promoter | *Hpa*II & Cloning | |
| chr20 | 32253872 | 32256071 | *ACTL10* | Y | Promoter | *Hpa*II & Cloning | 3 bi |
| chr20 | 39944425 | 39946301 | *ZHX3* | Y | Promoter | Cloning | 4 pat |

| Chr | Start | End | DMR | CpGI | GenePos | Confirmation | Expression |
|---|---|---|---|---|---|---|---|
| | | | | | | | 1 mono |
| | | | | | | | 3 bi |
| chr22 | 17089962 | 17090636 | *TPTEP1* | Y | Gene_body | *Hpa*II | |
| chr22 | 40057496 | 40061223 | *CACNA1I* | Y | Gene_body | *Hpa*II & Cloning | |

### 4.3.1 Novel ubiquitous imprinted DMRs

We observed only one sperm-derived region mapping to a known paternally methylated DMR in > 12 tissues, the *H19* DMR on chromosome 11. The only additional known paternally methylated DMR originating from sperm in humans, the *IG*-DMR on chromosome 14, was differently methylated between gametes and partially methylated in blastocysts, but had a DMR profile in five somatic tissues only. Using the same criteria, we observe 60 oocyte-derived DMRs in > 12 tissues, including 25 known maternally methylated imprinted DMRs (full list of regions, available online at https://doi.org/10.1371/journal.pgen.1006427 (Sanchez-Delgado et al., 2016b)).

Of the intervals not assigned as known imprinted loci, we performed confirmation experiments for 7 intervals and confirmed *FANCC* and *SVOPL* as being novel ubiquitous imprints (Figure 4.17A). Bisulphite PCR in multiple tissues, including placenta, brain and leukocytes revealed strand-specific methylation, which when informative was on the maternal allele (Figure 4.17B). The remaining candidates had mosaic methylation in all somatic tissues tested. Using allele-specific RT-PCR that incorporated a coding SNP within exon 5 of the *SVOPL* gene, we observed maternal expression in placenta and monoallelic expression in brain and leukocytes (Figure 4.17C). Unfortunately, we could not identify any informative samples to allow for the allelic expression of *FANCC* to be ascertained.

### 4.3.2 Methyl-seq revealed additional maternal placenta-specific DMRs

The screening of regions with partially methylated in blastocyst and conserved in only one tissue revealed 551 intervals inheriting methylation from the oocyte that survived only in the placenta, whereas only 38 regions inheriting methylation from the sperm were identified in this extra-embryonic tissue (full list of regions, available online at https://doi.org/10.1371/journal.pgen.1006427 (Sanchez-Delgado et al., 2016b)). We then selected both, maternally and paternally derived candidate regions for subsequent downstream

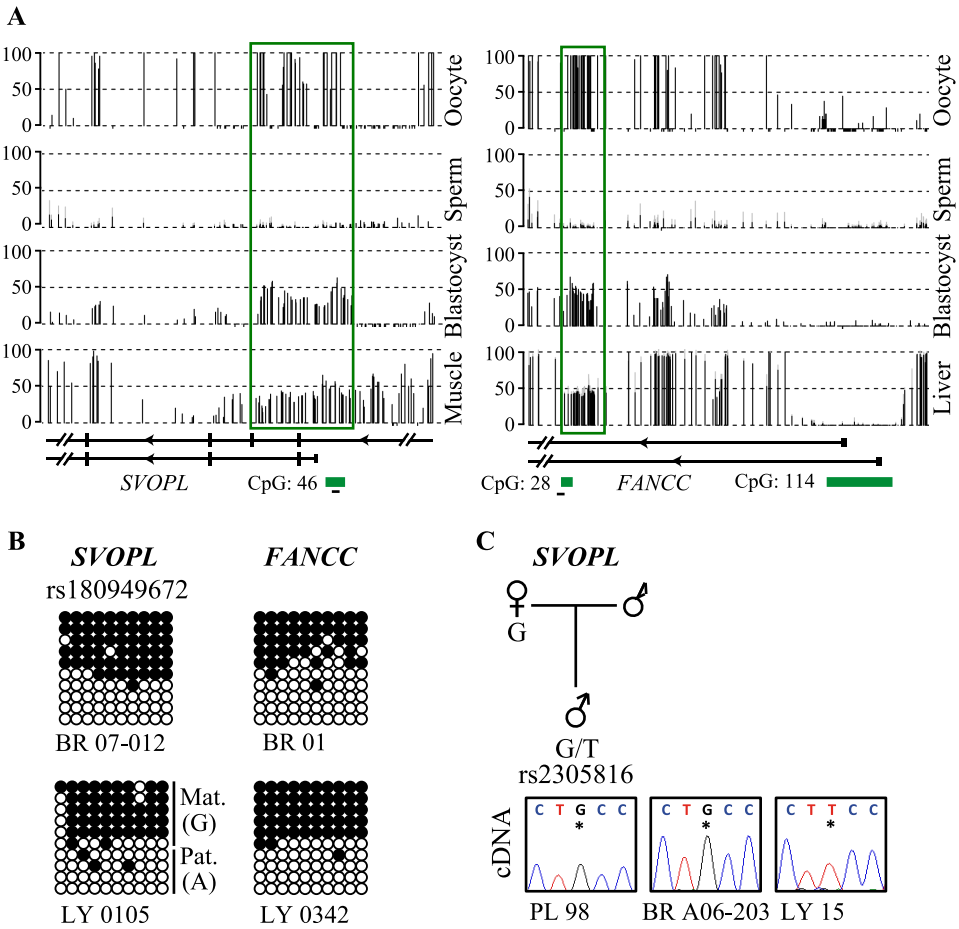



Figure 4.17: **Methylation and allelic expression profiling of novel ubiquitous imprinted loci.** (**A**) The two novel maternally methylated ubiquitous DMR associated with the *SVOPL* and *FANCC* genes exhibit promoters that are unmethylated in sperm, hypermethylated in oocytes and intermediately methylated in blastocysts, placenta and somatic tissue in methyl-seq datasets. The vertical black lines in the methyl-seq tracks represent the mean methylation value for individual CpG dinucleotides. Green boxes highlight the position of the gDMRs ( *data analysis and graph generated by Dr Court*). (**B**) Confirmation of allelic methylation ay *SVOPL* and *FANCC* DMRs by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative. (**C**) Allelic expression analysis of the novel ubiquitous *SVOPL* imprinted genes in control placentas (PL) and blood (LY) and brain (BR) samples. The asterisk* on the electropherogram highlights the position of the SNP. The gene is expressed from the maternally inherited allele in placenta samples.

confirmation. Using the same methylation-sensitive genotyping and standard bisulphite PCR and subcloning, we discarded all paternally-derived methylated candidates regions as candidates as they were all mosaically methylated (full list of regions available online at https://doi.org/10.1371/journal.pgen.1006427 (Sanchez-Delgado et al., 2016b)). On the other hand, we discovered that maternally methylated placenta-specific DMRs are highly abundant.

These candidates were from the two different classifications of placenta-specific DMRs, i.e. *TMEM247* were fully methylated in somatic tissues (Type 1 placenta-specific DMRs) whereas all the other regions were almost always high CG content intervals, robustly unmethylated in somatic tissues (Type 2 placenta-specific DMRs). The high validation frequency of maternal DMRs overlapping CpG islands compared to the low frequency of confirmation for low CpG-density regions is highlighted at the *GRID2* locus (Figure 4.18). This gene is associated with two maternally methylated gDMRs candidates with different genomic content. The promoter CpG island is robustly methylated on the maternal allele in placenta and is unmethylated in somatic tissues, whereas an intergenic region within intron 3, consisting of an Alu/SINE repeat, is a gDMR with a mosaic methylation profile in placenta that is fully methylated in all somatic tissues (Figure 4.18B). Overall we confirmed 38 new placenta-specific DMRs, with informative parental information for 28 (Figures S3, S4 and S5).

In addition to the methylation analysis, we also studied the expression of some transcripts whose promoter overlapped a placenta-specific DMRs by allele-specific RT-PCR. As an example, we observe paternal expression of a ~10 kb ncRNA overlapping a placenta-specific gDMR located 12 kb 3' to *TET3* (Figure 4.19B). To determine if this ncRNA influences expression *in cis*, we performed allelic RT-PCR for *TET3* (Figure 4.19C). We observe biallelic expression of *TET3* suggesting that the neighbouring ncRNA does not possess enhancer or repressive function in term placenta.

Paternal or monoallelic expression was also observed for several protein-coding genes, including *AGO1, USP4, SH3BP2, FAM149A, MOCS1, R3HCC1, JMJD1C, PAK1, PAPLN*-AS1 and *ZHX3* genes (Figure 4.20).

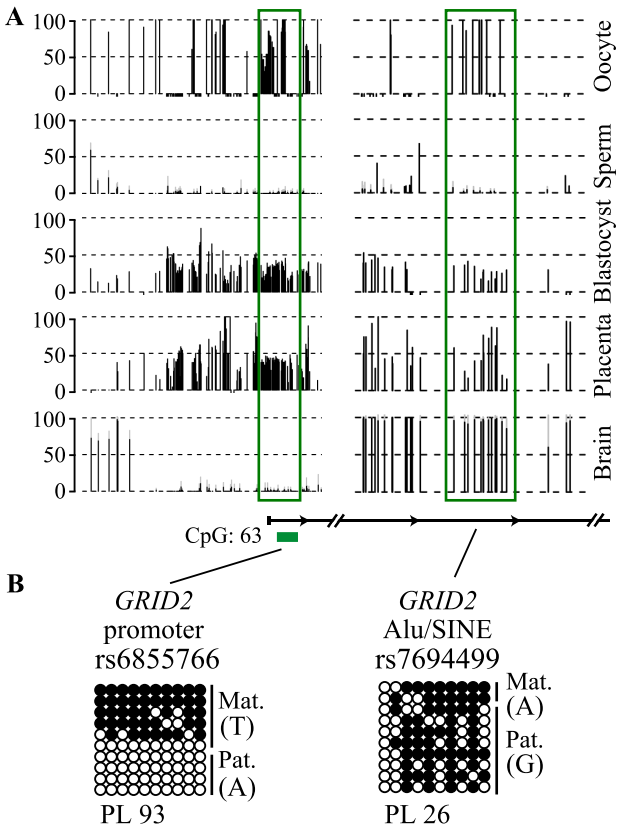



Figure 4.18: **Methylation analysis at *GRID2* locus.** (**A**) The *GRID2* gene exhibit several regions with oocyte-derived methylation. The methyl-seq data reveals that a 1.9 kb region overlapping the promoter remains an imprinted gDMR in placenta and blastocysts while it is demethylated in all other tissues. A second oocyte-derived gDMR, consisting mainly of an Alu/SINE repeat, becomes fully methylated in all tissues analysed. The vertical black lines in the methyl-seq tracks represent the mean methylation value for individual CpG dinucleotides. Green boxes highlight the position of the gDMRs (*data analysis and graph generated by Dr Court*). (**B**) Confirmation allelic and non-allelic methylation profile at two DMRs close to *GRID2* gene by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

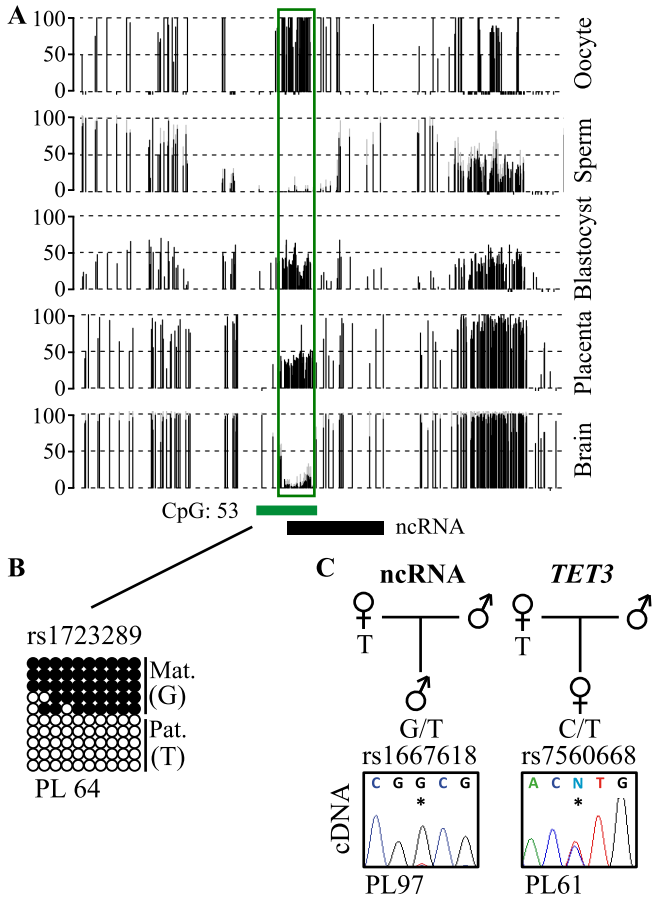



Figure 4.19: **Methylation and allelic expression analysis at *TET3* locus.** (**A**) The identification of a 10 kb ncRNA overlapping a placenta-specific gDMR 12 kb downstream of the *TET* gene. The vertical black lines in the methyl-seq tracks represent the mean methylation value for individual CpG dinucleotides. Green boxes highlight the position of the gDMRs (*data analysis and graph generated by Dr Court*).(**B**) Confirmation of allelic methylation at the 2.7 kb maternally methylated placenta-specific DMR confirmed by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative. (**C**) Allelic expression analysis of the novel placenta-specific ncRNA close to *TET* gene (paternally methylated) and the biallelic expression of *TET* transcript in control placenta samples (PL). The asterisk* on the electropherogram highlights the position of the SNP.

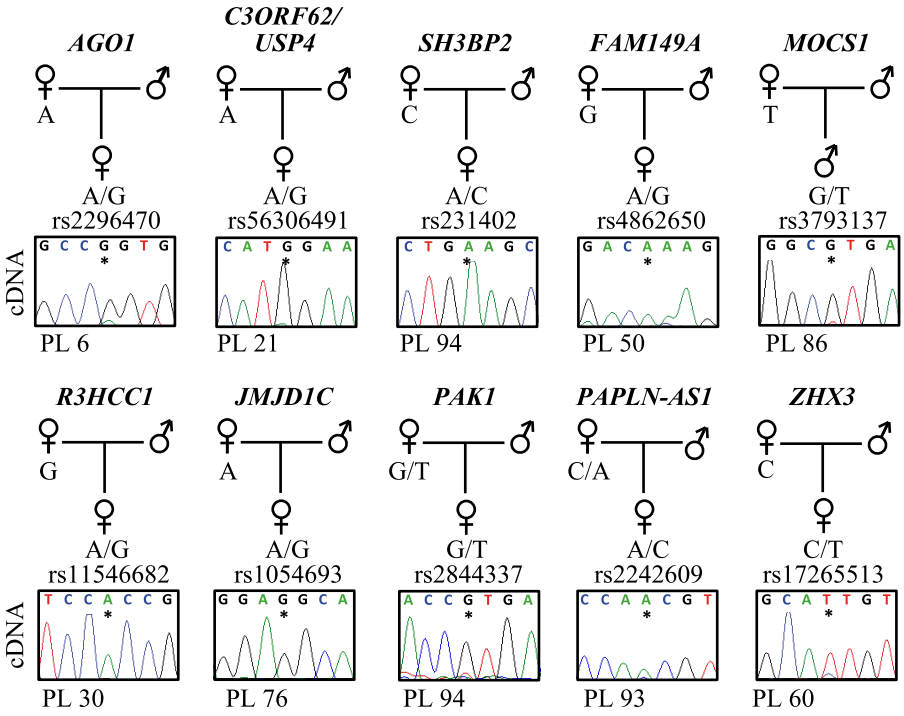



Figure 4.20: **Allele-specific RT-PCR analysis of methyl-seq candidate placenta-specific imprinted genes.** Confirmation of monoallelic expression of *PAK1* and *PAPLN-AS1* and paternal expression of *AGO1, C3ORF62/USP4, SH3BP2, FAM149A, MOCS1, R3HCC1, JMJD1C* and *ZHX3* in term placenta samples (PL). The asterisk* on the electropherogram highlights the position of the SNP.

## 4.4 Germline DMRs persisting in preimplantation embryo

By utilising a nested-multiplex bisulphite PCR and subcloning strategy, we confirmed the germline methylation asymmetries in human blastocyst of both ubiquitous and placenta-specific DMRs. Specifically, we confirm the methylation profile of two known ubiquitous imprinted gDMRs (*H19* and *MCTS2* DMRs) in sperm and blastocysts separated into the ICM and TE following microsurgery (Figure 4.21). Both regions showed ~50% methylation in both ICM and TE with the appropriate methylation profile in sperm, with no methylation at *MCTS2* DMR and fully methylated at *H19* DMR (Figure 4.21A). Within the same nested-multiplex reaction we also assessed the methylation of the two new ubiquitous gDMRs, *FANCC* and *SVOLPL*. The allelic methylation at these regions was less pronounced but significant enough to show that they are maternally methylated gDMR with no evidence of methylation in sperm-derived DNA samples (Figure 4.21B).

Next, I employed the same nested-multiplex PCR approach to assessing the methylation asymmetries of placenta-specific DMRs in a single human blastocyst. The *R3HCC1* loci on chromosome 8 exemplified the fate of opposing germline methylation difference as this gene has adjacent oocyte and sperm-derived gDMRs (Figure 4.22A). We showed that the Type 2 maternally methylated gDMR is maintained in both TE – which harbours the precursor cells of the placenta – and ICM – the precursor of the embryo itself –. On the other hand, the sperm-derived methylated region resolves to a state of mosaic methylation at the blastocysts stage and becomes fully methylated in placenta (Figure 4.22B).

Of this 15 regions assessed in human blastocyst by nest-multiplex bisulphite PCR, several were associated with developmentally interesting genes, including multiple members of two large gene families, the fibroblast growth factors (*FGF8, FGF13* and *FGF14*) – which had been reported to modulate the gradual segregation of different ICM derived lineages in mouse (Yamanaka, Lanner, & Rossant, 2010) – and calcium channel, voltage-dependent channel subunits (*CACNA1A, CACNA1C, CACNA1E* and *CACNA1I*) – associated with neurodevelopment and neuronal differentiation, which *CACNA1A* were identified to had a cortex-specific different methylation pattern (Davies et al., 2012) – (Figure 4.23A and Figure S6A). Additionally, we also check the methylation profile of genes involved in epigenetic regulation (*JMJD1C* and *DNMT1*) and microRNA processing (*AGO1* and *LIN28B*), all of which showed a consisted strand/allele-specific methylation pattern in both ICM and TE (Figure 4.23B and Figure S6B).

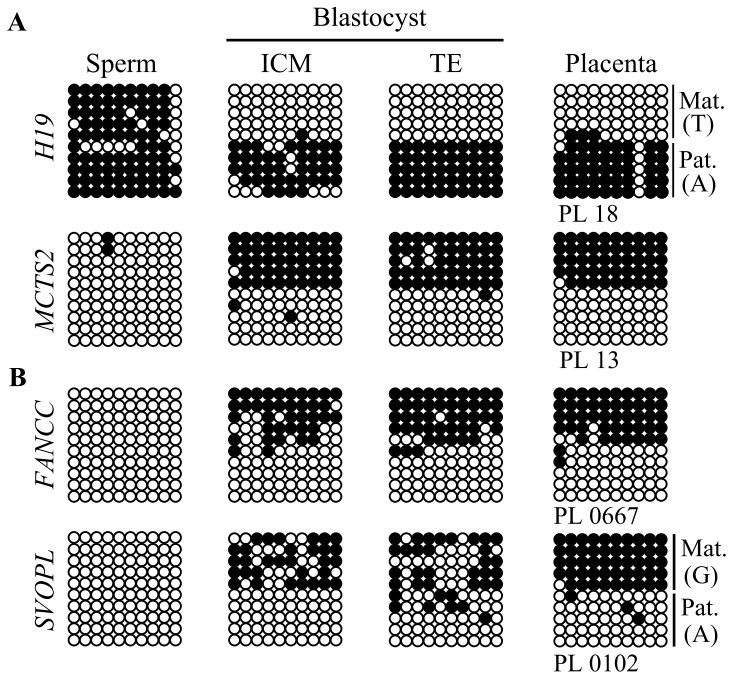



Figure 4.21: **Confirmation of allelic methylation at ubiquitous DMRs in blastocyst state.** Conformation of sperm methylation status and allelic methylation in placenta and blastocyst (ICM and TE) for (**A**) already reported and (**B**) new ubiquitous DMRs by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

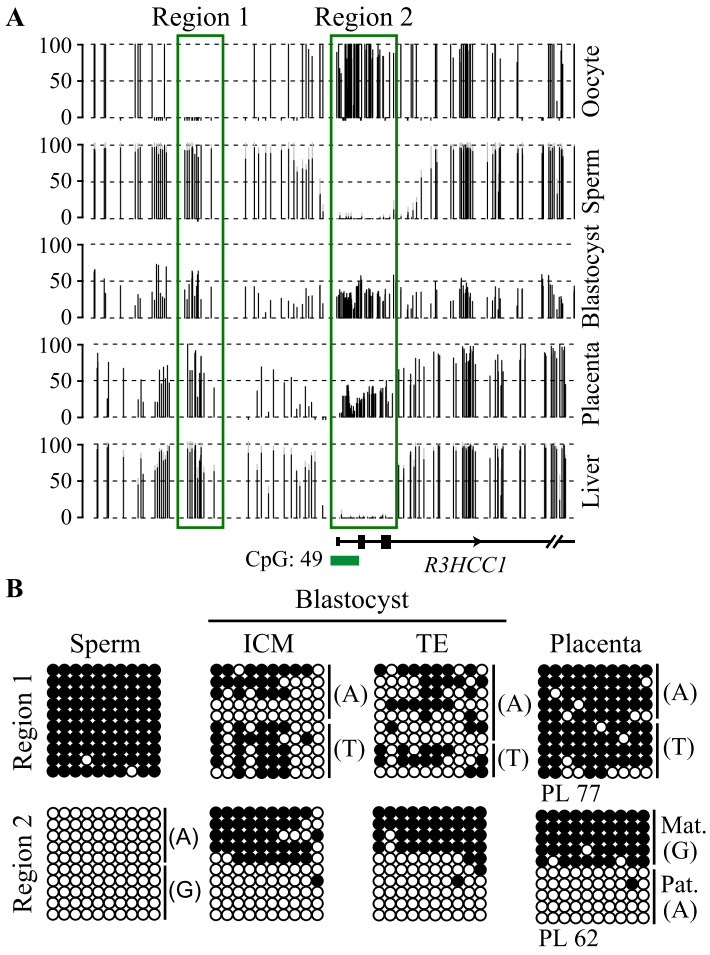



Figure 4.22: **Methylation profiling of opposing gDMRs using bisulphite PCR in human sperm and blastocysts.** (**A**) Methyl-seq datasets reveal that the *R3HCC1* gene has two adjacent gDMRs, an upstream paternal gDMR (region 1) that subsequently gains methylation on both alleles during the blastocyst stage and a placenta-specific maternally methylated promoter region (region 2). The vertical black lines in the methyl-seq tracks represent the mean methylation value for individual CpG dinucleotides. Green boxes highlight the position of the gDMRs (*data analysis and graph generated by Dr Court*). (**B**) Confirmation of methylation profile of both regions in sperm, blastocyst (ICM, TE) and placenta by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

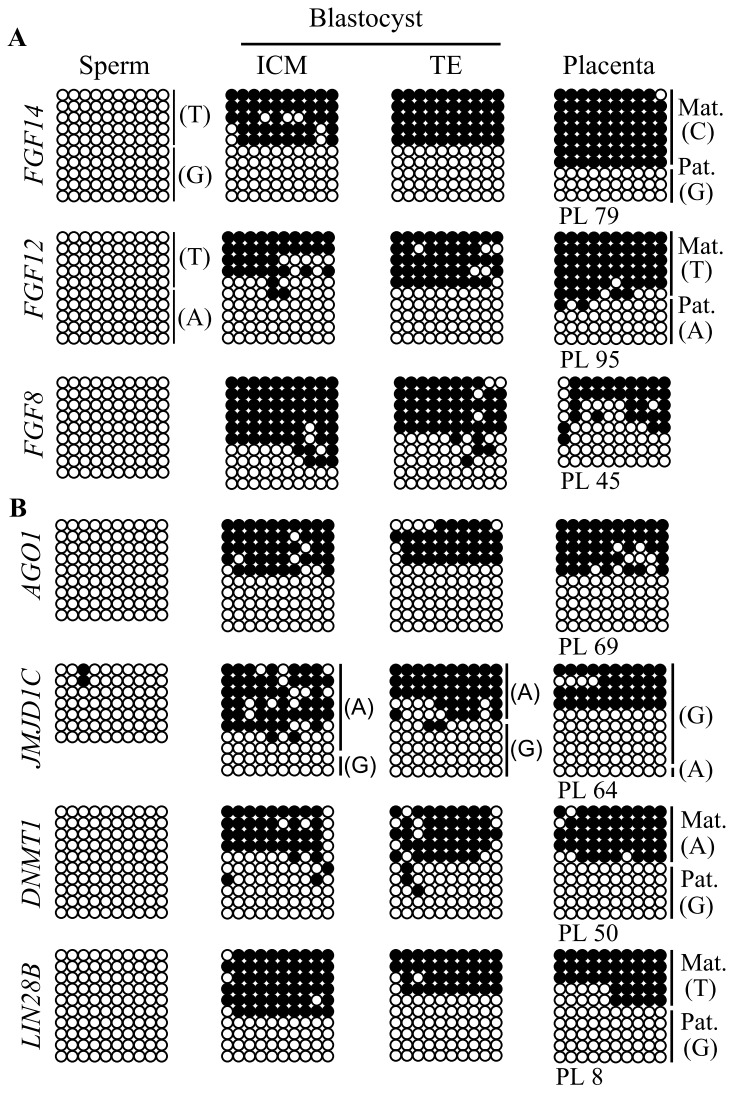



Figure 4.23: **Confirmation of allelic methylation at placenta-specific DMRs in blastocyst state.** Confirmation of unmethylated sperm and allelic methylation in blastocyst (ICM, TE) and placenta at placenta-specific DMRs. Confirmation at (**A**) DMRs within 3 FGFs family members and (**B**) other placenta-specific DMRs within other epigenetic regulator genes (*AGO1, JMJD1C, DNMT1* and *LIN28B*) by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

## 4.5 Imprinted expression of *ZHX3* in preimplantation embryos

In parallel, by using publically available single cell embryo RNA-seq datasets for which paternal genotypes were available (Z. Xue et al., 2013), gene expression profiles were analyzed in individual embryos to determine the temporal initiation of expression and their allelic origin.

To compare embryos at different stages it is important to take into consideration two events: ZGA and oocyte-derived transcript degradation. ZGA occurs soon after fertilisation (pre-major ZGA) and processed in successive waves of activation with the major changes reported at the 4-8 cell stage (Yan et al., 2013). Maternal transcript stores in the oocyte cytoplasm are diminished after fertilisation by a combination of degradation and recruitment to the polysome and translated prior to ZGA (Vassena et al., 2011) (See Figure 1.9 in introduction section). Transcripts highly abundant at the pronuclear stage and decreasing as developmental proceeds will not be expressed from the embryonic genome and will appear maternally derived. Embryonically transcribed genes that maintain high expression levels from the pronuclear stages would appear maternally expressed before 8-cell stage, switching to imprinted paternal expression with RNA synthesis from the unmethylated allele if the gDMRs are functional. Some instances of biallelic expression maybe wrongly classified since embryonic paternal expression and oocyte-derived transcripts may co-exist until late cleavage stage. Finally, genes that are activated during cleavage embryo development, but not originally expressed in the zygote are predicted to be from the paternal allele. Therefore functional paternal expression can only be categorized after genome activation (Figure 4.24A).

Using these criteria, Dr Enrique Vidal performed a bioinformatics screen of transcripts initiating from oocyte-derived gDMRs for imprinting. This informatic approach revealed that the *ZHX3* gene (one of the imprinted genes confirmed during this thesis, Figure 4.20) is paternally expressed in 4-8 cell embryos and the morula (see Figure 4.24B for the expression profile and Figure 4.25 for the methylation profile).

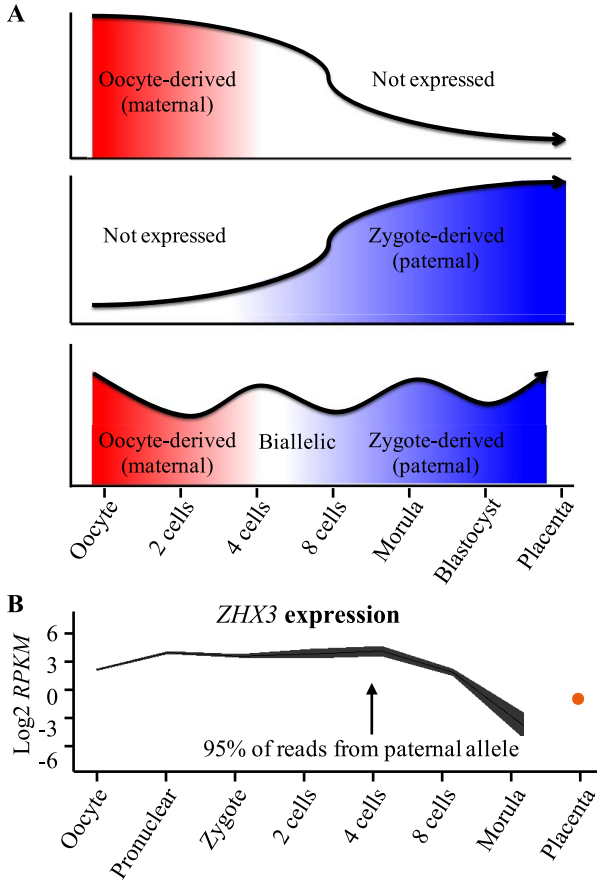



Figure 4.24: **Identification of novel imprinted genes in human embryos using allele-specific RNA-seq dataset.** (**A**) Schematic drawing of the sequential transcriptome switching from oocyte-derived transcripts to the embryonic expressed transcripts in human preimplantational embryos (**B**) The expression pattern of the *ZHX3* gene during human preimplantational development. High expression was observed from the zygote to the 8-cell stage, declining in the morula. Paternal expression as observed from the 4-cell stage onwards (*data analysis and graph generated by Dr Vidal).*

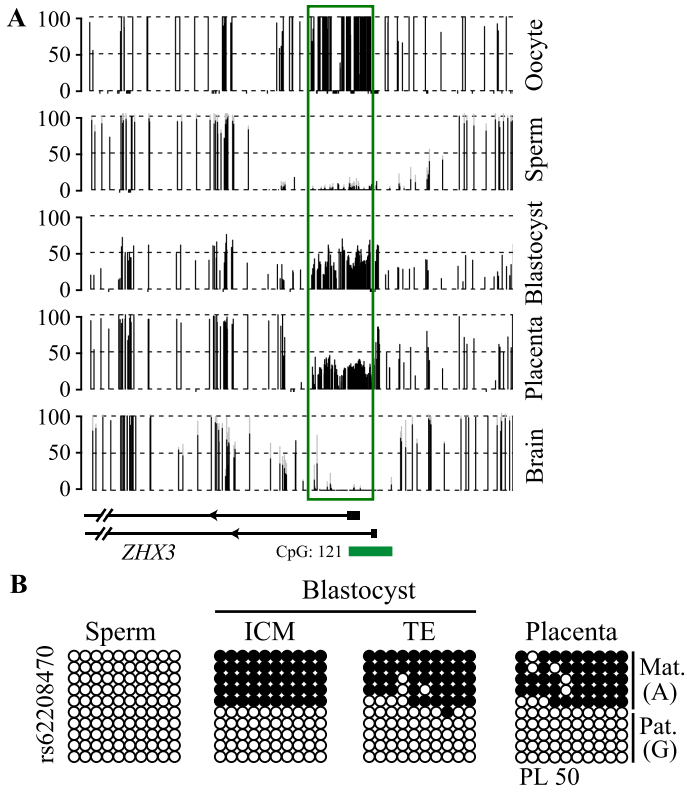



Figure 4.25: **Methylation analysis at placenta-specific *ZHX3* imprinted locus.** (**A**) Methyl-seq datasets reveal that the *ZHX3* gene has an adjacent placenta-specific DMR within its promoter. The vertical black lines in the methyl-seq tracks represent the mean methylation value for individual CpG dinucleotides. Green boxes highlight the position of the gDMRs (*data analysis and graph generated by Dr Court*).(**B**) Confirmation of absence of methylation at sperm and allelic methylation in blastocyst and placenta samples at placenta-specific *ZHX3* DMR by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

## 4.6 Placenta-specific DMRs regulates micro-imprinted domains

To determine whether placenta-specific DMRs can orchestrate allelic silencing of gene clusters similar to ubiquitous imprinted DMRs, we performed allele-specific RT-PCR for 20 flanking genes associated with loci containing imprinted transcripts (Figure 4.26 and Figure S7). Surprisingly, with the exception of *ADAM23* within the *GRP1-AS1/ZDBF2* domain on chromosome 2, we observe that the remaining 19 transcripts analysed are expressed equally from both parental chromosomes, indicating that placenta-specific DMRs do not possess the ability to regulate allelic expression of surrounding genes. Curiously, despite the evolutionary conservation of imprinting regulation at *GPR1-AS/ZDBF2* region (H. Kobayashi et al., 2013), paternal expression of *ADAM23* is not observed in mouse placenta (Figure 4.26A-B). Together, this suggests that this locus is regulated differently to the majority of placenta-specific imprinted loci identified and that subtle species differences exist (Duffié et al., 2014).

## 4.7 Placenta-specific DMRs are conserve in primates

Recently, placenta methyl-seq datasets have been produced from different mammalian species (Schroeder et al., 2015). Using this data, we showed that the orthologues of the vast majority of human placenta-specific DMRs do not show partial methylation in mouse, dog, cow or horse. However, some regions were conserved in rhesus macaque (*Macaca mulatta*) (Figure 4.27).

Previously, in a study performed by members of our laboratory, it was reported that the mouse orthologues of the initial set of placenta-specific DMRs were devoid of allelic methylation (Court et al., 2014). We extended this analysis using bisulphite PCR and subcloning to show that no human placenta-specific DMRs are conserved in mice (Figure 4.28).

Several studies have shown that maternally methylated gDMRs mark different loci in mouse compared to humans, suggesting that the mouse genome may possess a unique set of placenta-specific DMRs inherited from the female germline (Okae et al., 2014; Smallwood et al., 2011). Therefore, we determined the fate of oocyte-derived gDMRs in hybrid (B6 x JF1) mouse placenta. Consistent with our previous observations, no maternal gDMRs persist as placenta-specific DMRs, reinforcing that this phenomenon is not observed in mice (Figure S8).

Subsequently, to determine whether evolutionary closer mammals have a limited number of placenta-specific DMRs, we performed bisulphite PCR on a placenta sample from Rhesus

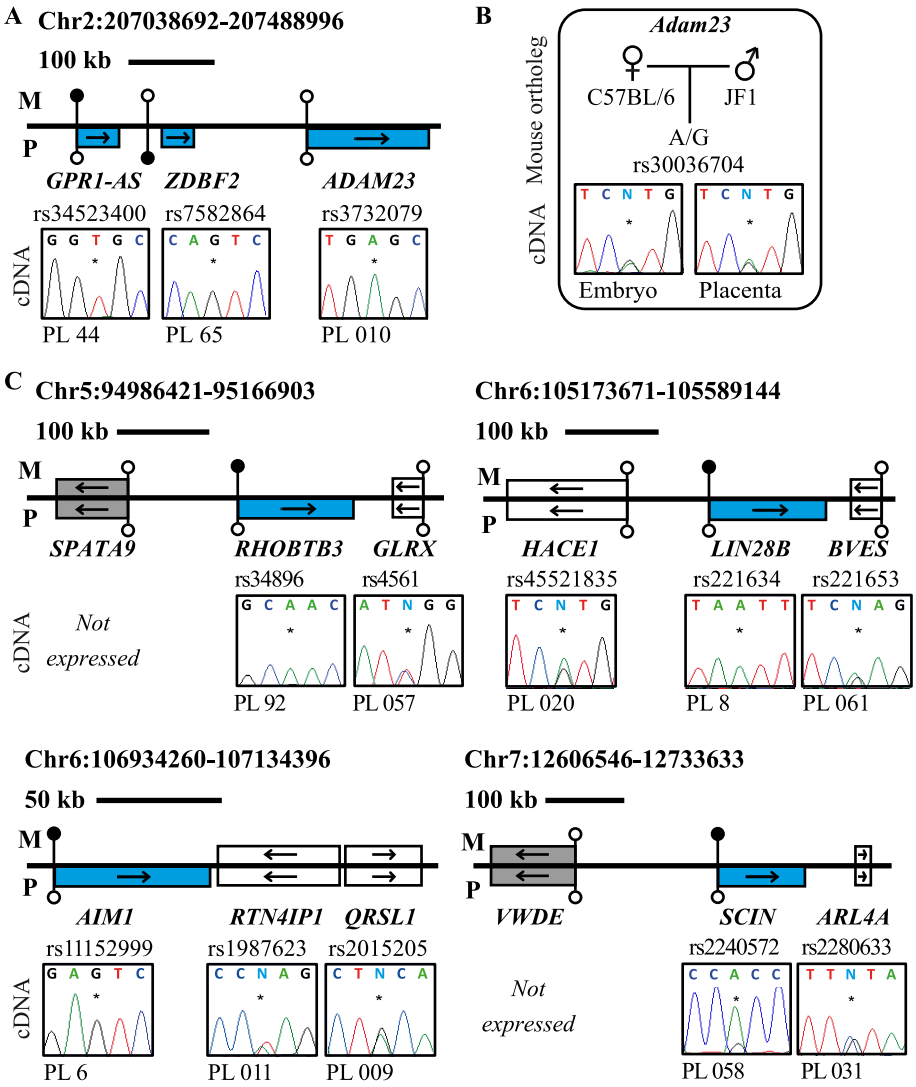



Figure 4.26: **Allele-specific RT-PCR analysis of candidate placenta-specific imprinted genes and its surrounding genes.** Confirmation of paternal expression of *GPR1-AS, ZDBF2, ADAM23, RHOBTB3, LIN28B, AIM1* and *SCIN* and no non-expressed or biallelically expressed surrounding genes in term placenta samples (PL). The asterisk* on the electropherogram highlights the position of the SNP. The maternally inherited chromosome (**M**, the upper part of the line) and the paternally inherited chromosomes (**P**, the lower part) of each placenta-specific imprinted locus are represented. The black lollipops indicate methylated CpG regions whereas white circles, unmethylated CpG regions. Blue filled rectangles represent paternally expressed genes; White filled rectangles represent non-imprinted ballelically; Grey filled rectangle represents non-expressed genes. The arrows represent the direction of the transcripts.

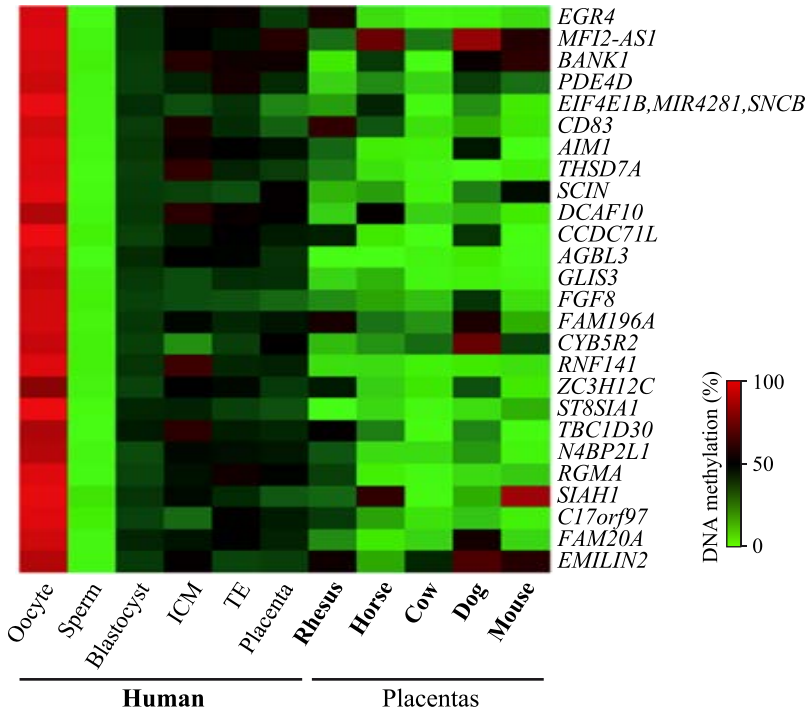



Figure 4.27: **Analysis of the orthologous sequences associated with human placenta-specific DMRs in different mammalian species.** Heatmap showing the methylation profiles of human placenta-specific gDMRs in methyl-seq datasets from placenta samples of the rhesus macaque, horse, cow, dog and mouse (*data analysis and graph generated by Dr Martin-Trujillo*).

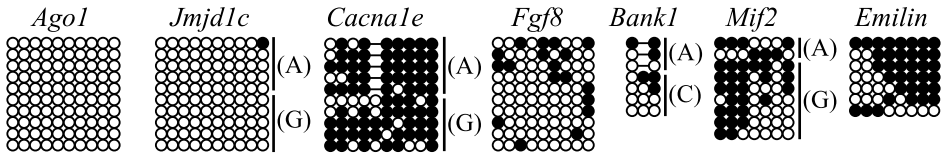



Figure 4.28: **Non-allelic methylation at human placenta-specific DMRs orthologous regions in *Mus musculus*.** Confirmation of non-allelic methylation profile at mouse orthologs for the human placenta-specific DMRs by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype.

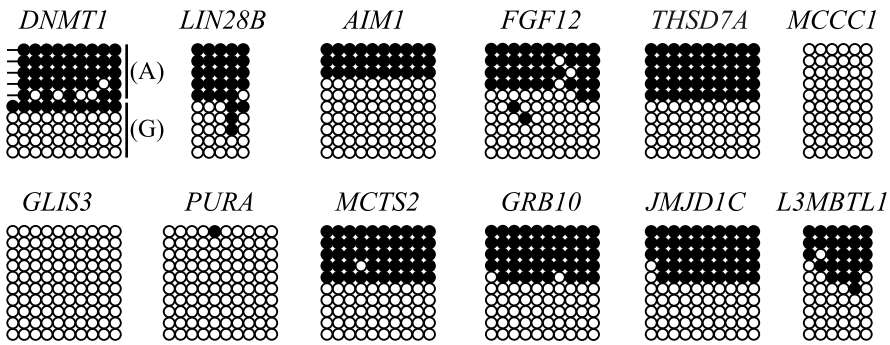



Figure 4.29: **Allelic methylation at human placenta-specific DMRs orthologous regions in *Macaca mulatta*.** Confirmation of LOM and allelic methylation by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype. Ubiquitous control regions: *MCTS2, GRB10* and *L3MBTL1*

macaque (Figure 4.29). This demonstrated the conservation of 6 of the 9 placenta-specific DMRs tested, as well as those associated with the ubiquitously imprinted *MCTS2, GRB10* and *L3MBTL1* genes.

## 4.8 The polymorphic placenta-specific imprinting

During the expression analysis of the new placenta-specific gene candidates, we observed that for certain genes, some placenta samples showed paternal expression, but others were biallelically expressed. When we checked the methylation profile of the corresponding DMRs, the placenta samples show maintained robust maternal methylation. The *MOCS1* gene is a clear example of loss-of-imprinting (LOI) but presence of allelic methylation (Figure 4.30A). This observation suggests that either these DMRs do not always influence allelic expression, or that the expression arises from a different promoter region. Since *in silco* database searches failed to identify transcripts originating from upstream promoters in placenta, we favour the first option.

To complicate the situation further, when analysing the total list of placenta-specific candidates identified in the two screens, we observed that the partial methylation might be absent in some samples. To examine this in greater depth, I initially performed confirmatory pyrosequencing for 29 placenta-specific DMRs (those with associated allelic expression) in a control placenta cohort of 55 term placenta samples from uncomplicated pregnancies (Figure 4.30C). We identified several loci presented with low/absence of methylation in a minority of samples and other regions like *THAP3*, with a higher frequency of LOM

samples. The placenta samples with LOM at specific DMR, shown biallelic expression to the associated gene. As an example, at *LIN28B* region, the LOM is associated with biallelic expression (Figure 4.30B).

Interestingly, when we extend the pyrosequence analysis to more samples from our placenta cohort (including IUGR, preeclampsia and ART), one of the samples lacking methylation at the *LIN28B* DMR (PL 216) comes from a multiple-gestation pregnancy with the dizygotic twin sibling (PL 215) having normal methylation at this locus (Figure 4.31).

Additionally, in our cohort, we have numerous samples coming from multiple-gestation pregnancy (not only PL 215 and PL 216). To assess inter-pregnancy/individual variation, we compared the methylation of imprinted DMRs in dizygotic twins and triplet pregnancies by hybridising corresponding placenta samples on Illumina Infinium HM450k array. These comparisons revealed variation at placenta-specific DMRs (Figure 4.32), but all ubiquitous DMRs had the appropriate methylation profile (Figure S9). In this approach, we analysed 112 regions incorporating 763 probes on the Illumina Infinium HM450k array associated with confirmed regions of maternal methylation specific to the placenta. This number represents the total number of confirmed placenta-specific DMRs from our work and from two studies that were subsequently published (Court et al., 2014; Hamada et al., 2016; Hanna et al., 2016; Sanchez-Delgado et al., 2015, 2016b). The siblings assessed from a triplet pregnancy shows different methylation pattern at 13 placenta-specific DMRs whereas the dizygotic twins at *THSD7A, EMID2, ZFP90* and *SEPT4*.

### 4.8.1 Histone tail modification signatures vary at polymorphic DMRs

To understand the role of additional epigenetic mechanisms in the polymorphic placenta-specific imprinting we studied post-translation-modifications of histones using ChIP. Due to the characteristics of imprinting regions, we focused our study on permissive histone modifications (H3K4me2, H3K4me3, H3K9ac), previously associated with non-methylated and expressed allele.

For the *LIN28B* gene, control samples (PL 8) with maintained imprinting were enriched for permissive histone marks (H3K4me2, H3K4me3, H3K9ac) in the paternally inherited chromosome, which is the non-methylated and expressed allele (Figure 4.33). However, in samples that were biallelically unmethylated (PL 216), we observed enrichment of permissive histone marks on both parental chromosomes and our quantitative analysis (normalised with *GAPDH*) showed a higher presence of these histone modifications respect control placenta.

In the case of *R3HCC1* DMR, we identify samples with both scenarios leading to polymorphic imprinted expression; one placenta with the expected 50% methylation at the DMR

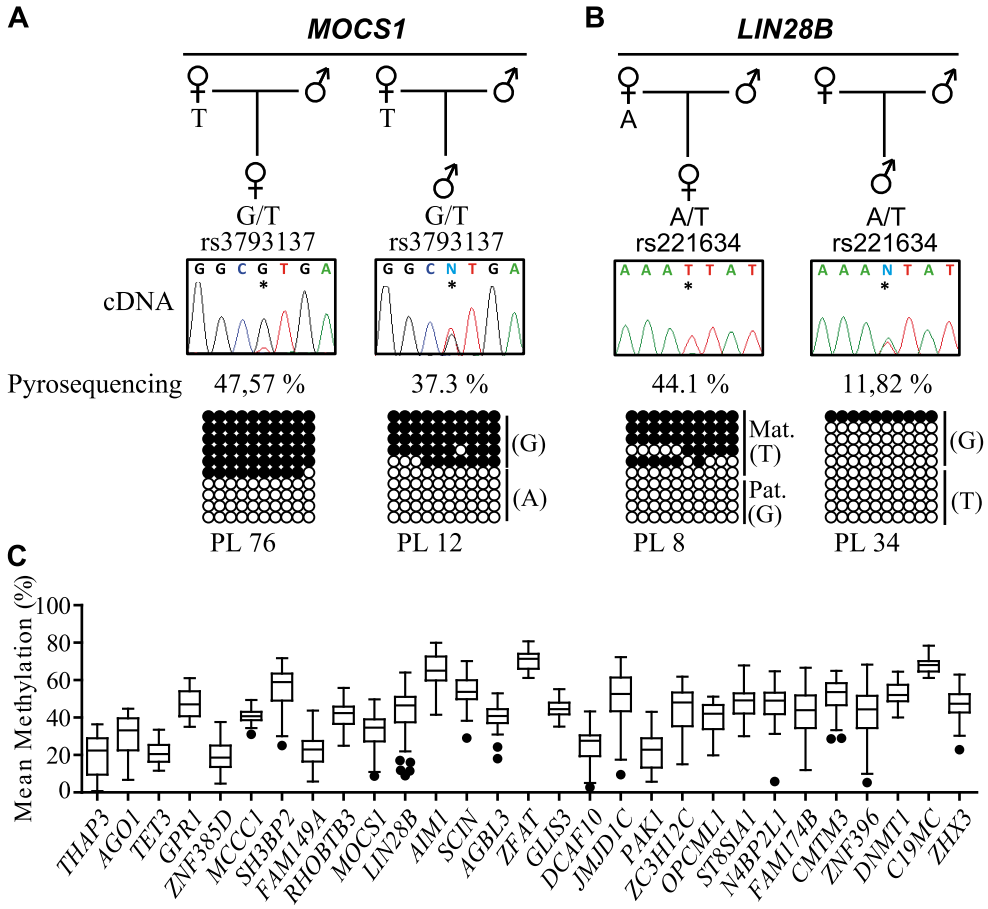



Figure 4.30: **Allele-specific expression and methylation analysis of genes with placenta-specific polymorphic imprinting.** Allelic methylation confirmed by bisulphite PCR and subcloning, where each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative. In the allelic expression analysis, the asterisk* on the electropherogram highlights the position of the SNP. (**A**) Confirmation of paternal expression of *MOCS1* gene in a term placenta sample (PL 76 respectively). However, other placenta samples showed biallelic expression for the same genes (e.g. PL 12). *MOCS1* had allelic maternal methylation at the DMR, confirmed by pyrosequencing and by bisulphite PCR and subcloning. (**B**) Confirmation of paternal *LIN28B* gene expression and maternally methylated DMR in PL 8 and non-methylated and associated with biallelic *LIN28B* expression in PL 34. (**C**) Quantitative pyrosequencing of 29 placenta-specific DMRs in 55 control placenta samples. The controls are represented as Turkey box-and-whisker plots with whiskers spanning from 25th to 75th percentiles +/- 1.5IQR to highlight outliers. Outliers were represented as black filled circles.

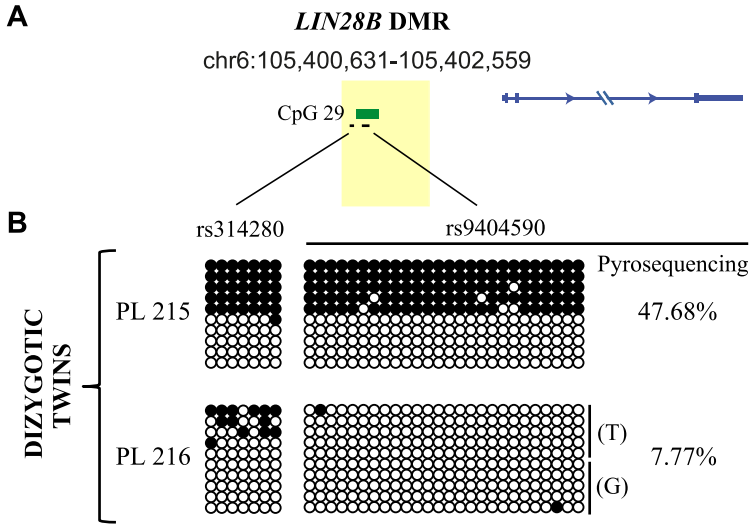



Figure 4.31: **Confirmation of DNA methylation pattern at *LIN28B* DMR in dizygotic twins.** (**A**) Schematic representation of *LIN28B* transcript (in blue) and the CpG island (green rectangle) where *LIN28B* DMR is located. The black line represents the location of our bisulphite PCRs and location of *LIN28B* DMR (indicated coordinates for the hg19 assembly of human UCSC Genome Browser) is highlighted as a yellow shaded region. (**B**) Confirmation of LOM and allelic methylation by bisulphite PCR and subcloning of placenta samples from dizygotic twins (PL 215 and PL 216). Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. The group of amplicons from the same allele are shown with the letters in the parentheses indicating SNP genotype. Quantification of methylation at this region by pyrosequencing is indicated as % of methylation.

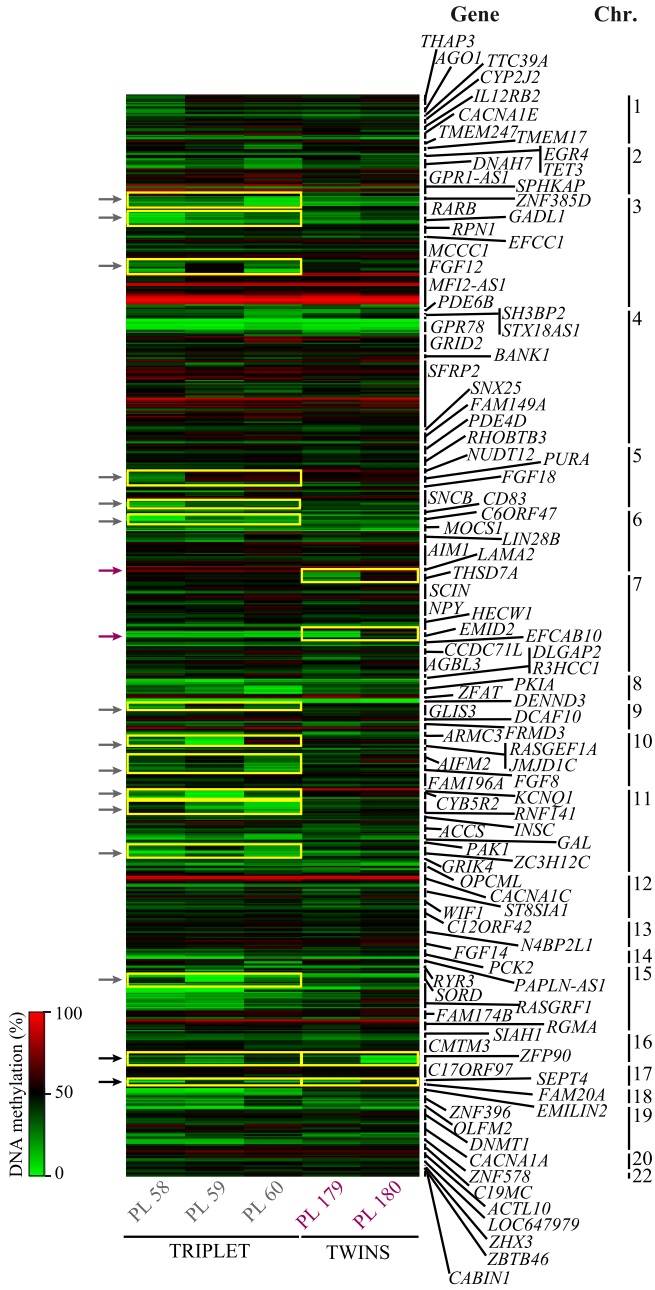



Figure 4.32: **Genome-wide methylation analysis in placentas from dizygotic twins and triplet pregnancy at placenta-specific DMRs.** A Heatmap of the Infinium probes located within known placenta-specific imprinted DMRs. The samples hybridisated were placentas from a triplet pregnancy (PL 58, PL 59 and PL 60) and dizygotic twins (PL 179 and PL 180). Regions with changes between siblings were highlighted with a yellow rectangle and pointed with coloured narrow (grey for differences between triplets: *FGF12, PURA, SCNCB, C6ORF48, GLIS3, RASGF1A, AIFM2 CYB5R2, RNF141, ZC3HC12C* and *SORD* DMRs; purple for differences between twins: *TH5D7A* and *EMID2* DMRs; black for differences found between two sets of siblings: *ZFP90* and *SEPT4* DMRs) (*data analysis and graph generated by Dr Hernandez-Mora*).

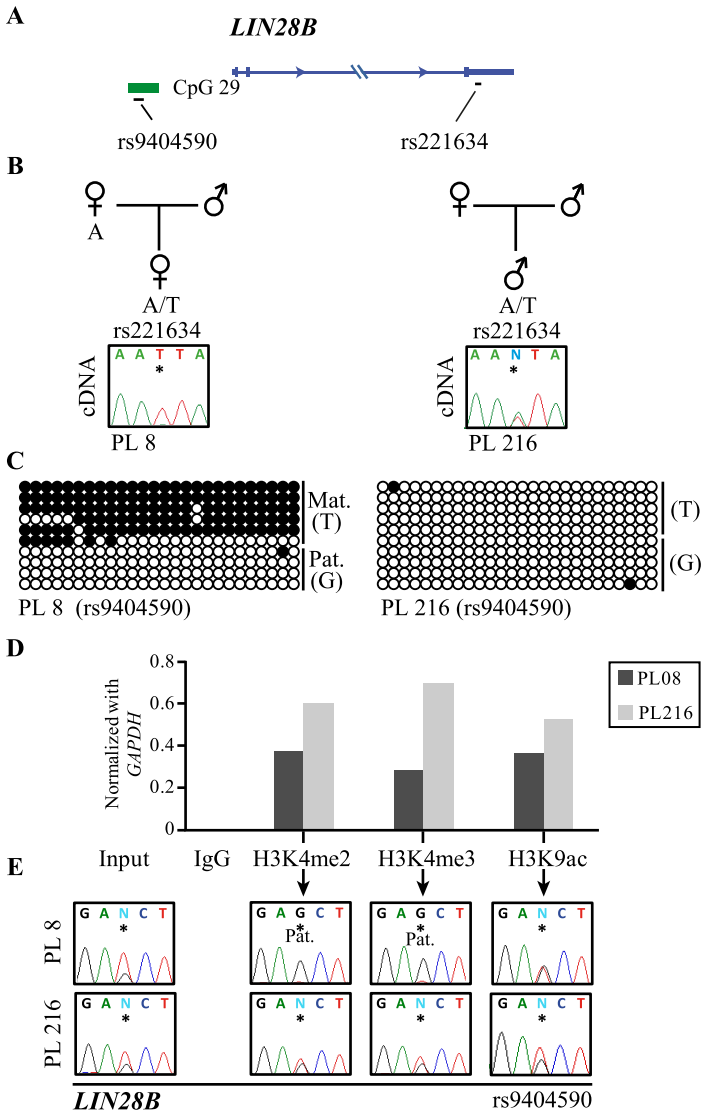



Figure 4.33: **Expression, methylation and histone modification analysis of *LIN28B* region with polymorphic imprinting.** The asterisk* on the electropherogram highlights the position of the SNP. (**A**) Schematic representation of *LIN28B* transcript (in blue) and the CpG island (green rectangle) where *LIN28B* DMR is located. The black line represents the location of our bisulphite, allelic RT and quantitative PCRs. (**B**) Allelic expression analysis at control placenta (PL 8) and LOM case (PL 216) for *LIN28B* DMR. (**C**) Allelic methylation analysis at control placenta (PL 8) and LOM case for *LIN28B* DMR (PL 216) by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. Each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative. (**D**) Quantitative PCR for *LIN28B* DMR using ChIP (for H3K4me2, H3K4me3, H3K9ac) extracted DNA as template and normalised with *GAPDH* region. (**E**) Allelic analysis for permissive histone marks for *LIN28B* DMR.

(PL 62) and another lacking allelic methylation (PL 98). qPCR on ChIP material from control placenta samples with maternal methylation and monoallelic expression (PL 30) revealed paternal enrichment for the permessive histone modifications, as expected (Figure 4.34A-E). However, in the samples with biallelic expression, this permessive marks were observed decrised but equally on both parental chromosomes. For all samples, control allelic precipitation at ubiquitous DMRs (*MEST* and *SNURF*) confirmed the quality of the ChIP material (Figure 4.35).

### 4.8.2 Placenta-specific DMRs maintain its methylation level during gestation

One possible mechanism leading to the apparent polymorphic placenta-specific DMRs is the failure to maintain allelic methylation during gestation. To address this, we performed methylation analysis using the Illumina Infinium HM450K arrays on different placenta samples collected from the same pregnancies but at different time points (bioinformatics analysis performed by Dr Jose R Hernandez). For the temporal comparison, first-trimester CVS were compared with corresponding biopsies from the term placentas (PL 4853477 and PL 4881886) (Figure 4.36A). This comparison revealed that DNA methylation level at the placenta-specific DMR is stable between time points with correlations higher (r=0.95) than between unrelated placenta samples (correlation range r=0.68-0.89). Detailed scrutiny of the paired samples revealed that the DMRs showing polymorphic LOM were the same in the first-term pairs, but differed between sets (sample 4853477 LOM for *TTC39A, GADL1, PURA, C6ORF47, AIFM2* and *GRIK4*; sample 4881886 LOM for *THSD7A*, and *PURA*).

In addition, to ensure that the methylation profiles observed were uniform across the placenta bed, we determined the placenta-specific DMR profiles from multiple biopsies from the same term placentas (PL 55 and 55b with r=0.96; PL 58 and 58b with r=0.95). Our analysis revealed that the methylation status of each placenta-specific DMR is maintained in different parts of the same placenta (PL 55 LOM for 14 placenta-specific DMRs including *THAP3, AGO1, FGF12, GPR78, R3HCC1, ZC3H12C* and *ZFP90*; PL 58 LOM for 12 placenta-specific DMRs including *GPR78, PURA, SCNCB, R3HCC1, GLIS3* or *EMILIN2*) (Figure 4.36B).

All ubiquitous DMRs for all these samples compared (both CVS vs term placentas and multiple biopsies from the same term placenta) had the appropriate methylation profile.

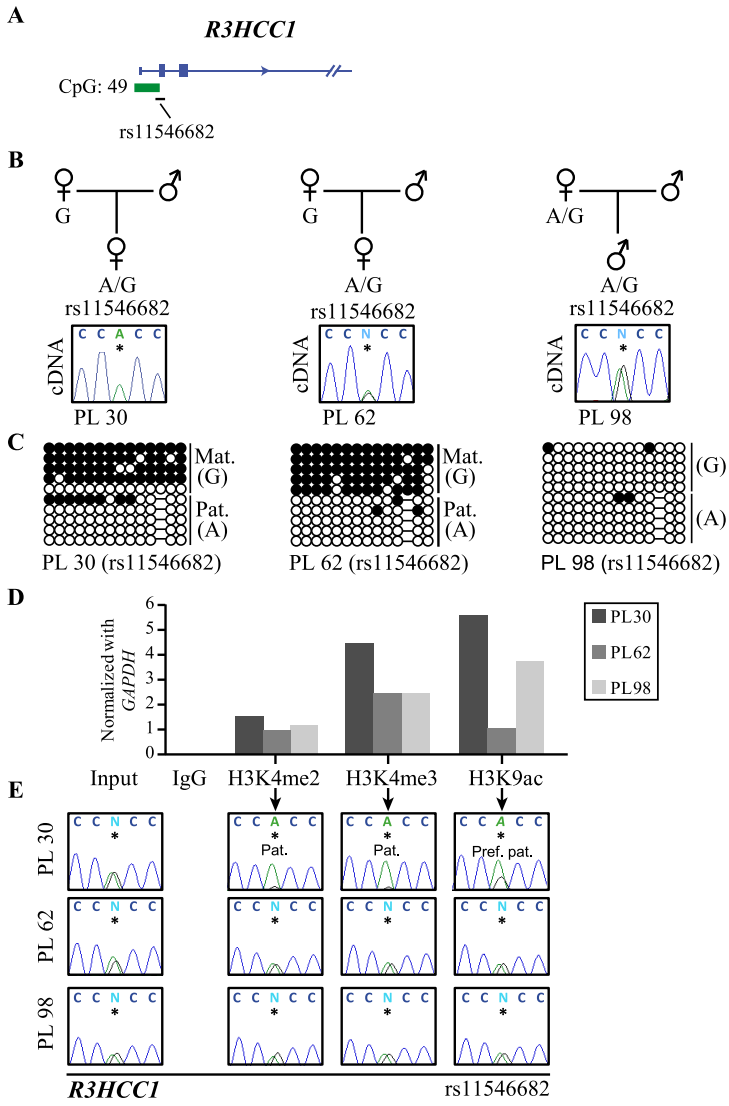



Figure 4.34: **Expression, methylation and histone modification analysis of *R3HCC1* region with polymorphic imprinting.** The asterisk* on the electropherogram highlights the position of the SNP. (**A**) Schematic representation of *R3HCC1* transcript (in blue) and the CpG island (green rectangle) where *R3HCC1* DMR is located. The black line represents the location of our bisulphite, allelic RT and quantitative PCRs. (**B**) Allelic expression analysis at control placenta (PL 30), LOI with normal methylation pattern (PL 62) and LOM and LOI case (PL 98) for *R3HCC1* DMR. (**C**) Allelic methylation analysis *R3HCC1* DMR by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. Each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative. (**D**) Quantitative PCR for *R3HCC1* DMR using ChIP (for H3K4me2, H3K4me3, H3K9ac) extracted DNA as template and normalised with *GAPDH* region. (**E**) Allelic analysis for permissive histone marks for *R3HCC1* DMR.

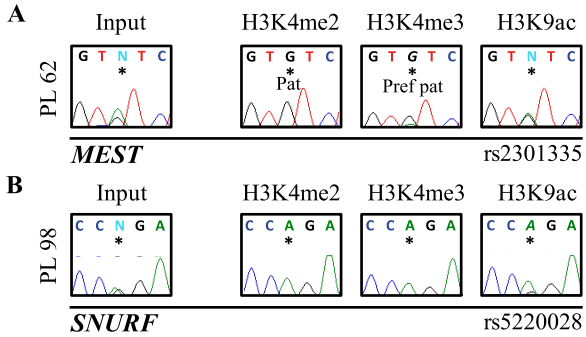



Figure 4.35: **Allele-specific presence of permissive histone marks in control imprinted DMRs for samples with LOI and LOM at *R3HCC1* region.** The case samples for *R3HCC1* DMR were also heterozygous for SNPs in other imprinted regions. The asterisk* on the electropherogram highlights the position of the SNP. Allelic presence of permissive histone marks are shown for (**A**) PL 62 at *MEST* DMR, and for (**B**) PL 98 at *SNURF* DMR.

## 4.9 Assesing imprinted DMRs in complicated pregnancies

Polymorphic events in placenta-specific DMRs suggest that these imprinted regions are maybe less stable than ubiquitous DMRs. Since it is a fact that higher rates of IUGR are observed in babies born following ART and initial studies in mice indicated that imprinting in the trophoblast is particularly susceptible to be altered *in vitro* embryo culture, we have considered important to determine if placental-specific DMRs is epigenetically vulnerable following ART and if an aberrant methylation profile of this regions is associated with IUGR. Nevertheless, an extensive analysis of the ubiquitous DMRs and its associated imprinted genes in IUGR placentas was also necessary to improve the knowledge about this concern in humans.

Using the Illumina Infinium HM450K array we profiled both ubiquitous and placenta-specific DMRs in 77 placenta samples. Additionally, we also performed quantitative pyrosequencing in 104 placenta samples at ubiquitous DMRs and 134 placenta samples at placenta-specific DMRs. The results obtained by these two methylation analysis are currently being processes taking in consideration not only the two classification criteria in the following subsections (IUGR and ART) but also other additional clinical or maternal information with a potential influence in the methylation status of imprinted DMRs (gestational age, tobacco consumption, maternal disease, previous miscarriages...). As a consequence, the next two subsections will be unreviewed preliminary data.

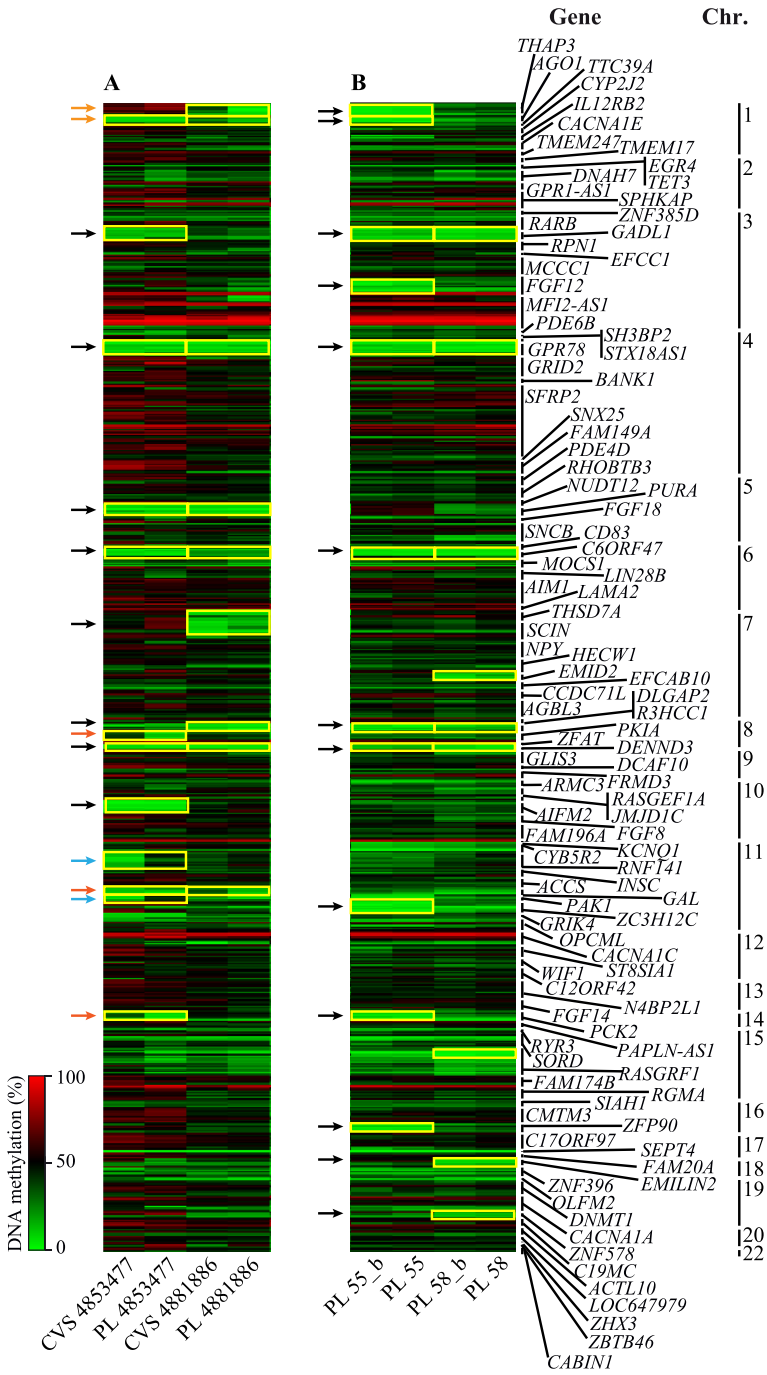



Figure 4.36: **Genome-wide methylation analysis in paired samples: CVS and term placentas.** A Heatmap of the Infinium probes located within all placenta-specific imprinted DMRs. (**A**) Paired samples: first-trimester CVS compared with corresponding biopsies from the term placentas (PL 4853477 and PL 4881886). (**B**) Multiple biopsies from the same term placentas (PL 55 and PL 58). Regions presenting LOM were highlighted with a yellow rectangle and pointed with coloured narrow (black for regions with LOM in all samples; orange for regions with LOM only present in term placentas; blue for regions with LOM only present in CVS) (*data analysis and graph generated by Dr Hernandez-Mora*).

### 4.9.1 Aberrant methylation at ubiquitous DMRs

Focusing initially on the 36 known ubiquitous DMRs with probes in the illumine Infinium HM450K array, we observed a gain-of-methylation (GOM) at two DMRs associated with *MEST* and *MCTS2* and additionally, we observed multiple cases with significant LOM associated with the *H19* and *NHP2L1* DMRs (Figure 4.37).

Subsequently, we tested by pyrosequencing a larger cohort of samples and confirmed these array observations (Figure 4.38A). Interestingly, all aberrant methylation profiles were restricted to the placenta as corresponding cord blood samples were in the normal range (Figure S10). In our pyrosequencing analysis, we also reported new LOM samples for *H19* DMRs and *NHP2L1* (Figure 4.38A). The 3 DMRs located in *H19* locus presented LOM in IUGR (n = 47) and control non-IUGR (n = 57) placentas. Two IUGR cases, samples PL 217 and PL 78 were LOM in the three regions and PL 67, PL 90 and PL 91 in two of them. In the case of control placentas, we also observe one sample being LOM in the three *H19* DMRs (PL 41) and two samples being LOM in two *H19* DMRs (PL 49 and PL 165). By using the Mann-Whitney-Wilcoxon non-parametric test (due to the non-normal distribution of analysed data), no significant differences were found in population means comparing IUGR vs non-IUGR groups at any tested region.

By RT-PCR followed by sequencing, we show the LOM of the *H19* DMR leads to the reactivation of the normally repressed allele *H19* gene in the three samples heterozygous for coding SNPs, accompanied by a concomitant decrease in *IGF2* expression as judged by qRT-PCR (Figure 4.38C). Furthermore, the samples with *NH2PL1* LOM were highly expressed compared with 50 normal samples. However, allelic studies could not be performed since these samples were not heterozygous for exonic SNPs. The single placenta sample with *MEST* GOM (PL 35) was the lowest expressed when compared to the same control set (Figure 4.38B).

### 4.9.2 Polymorphic events at placenta-specific DMRs are not associated with complicated pregnancies

We utilised the same Illumina Infinium HM450K array data to explore the profiles for the confirmed placenta-specific DMRs described from all published studies (Court et al., 2014; Hamada et al., 2016; Hanna et al., 2016; Sanchez-Delgado et al., 2015, 2016b) (Figure 4.39). This first analysis revealed that the majority of placenta-specific DMRs have some samples with low methylation, indicative of a polymorphic event. In extreme cases, such as *C6ORF47, MOCS1, R3HCC1, SEPT4* and *OLFM2*, there is an equal number, or even slightly more samples lacking methylation.

To confirm these observations, 31 placenta-specific DMRs were analysed by pyrosequencing.

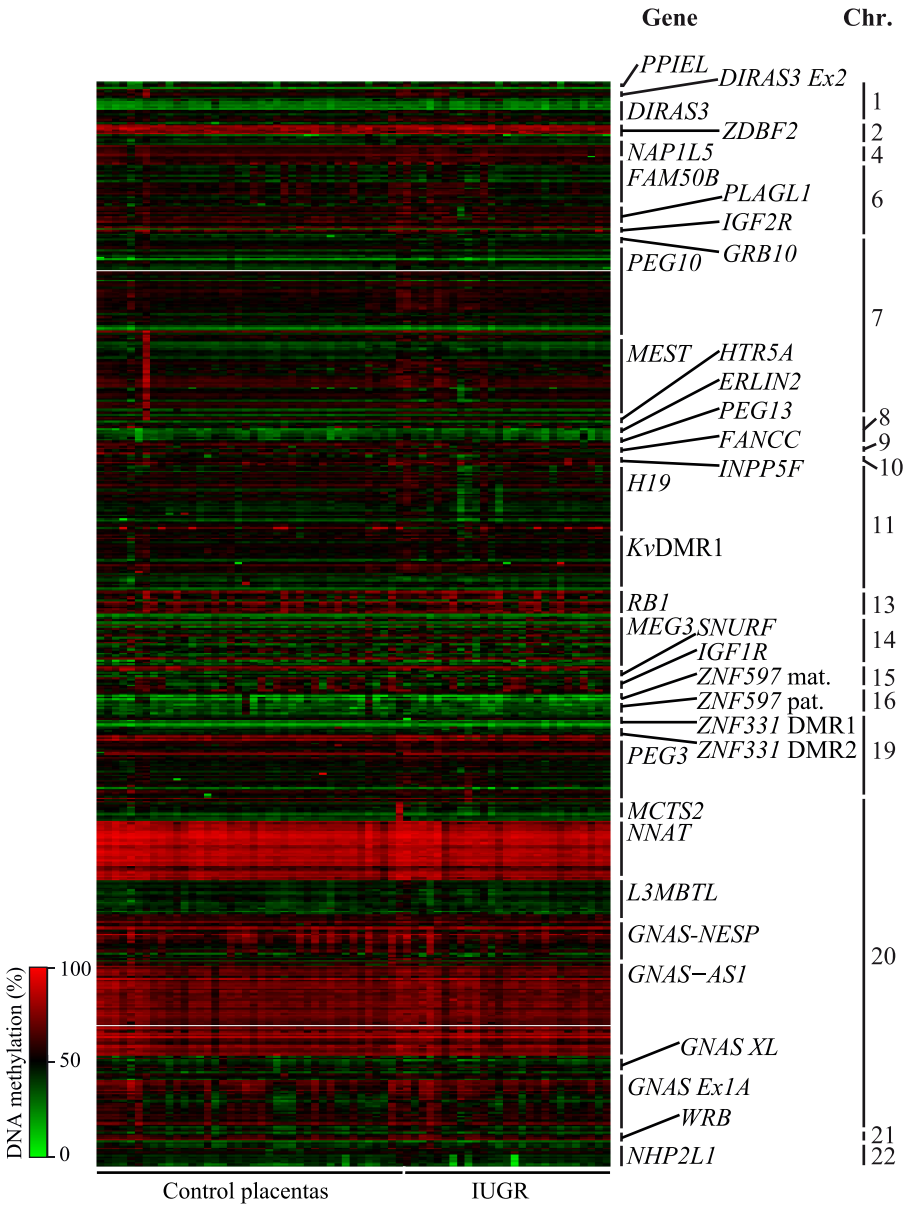



Figure 4.37: **Genome-wide methylation analysis in 77 placenta samples at ubiquitous DMRs** A Heatmap of 616 Infinium array probes mapping to 36 ubiquitous imprinted DMRs for 77 placentas, including IUGR and ART samples (*data analysis and graph generated by Dr Hernandez-Mora*).

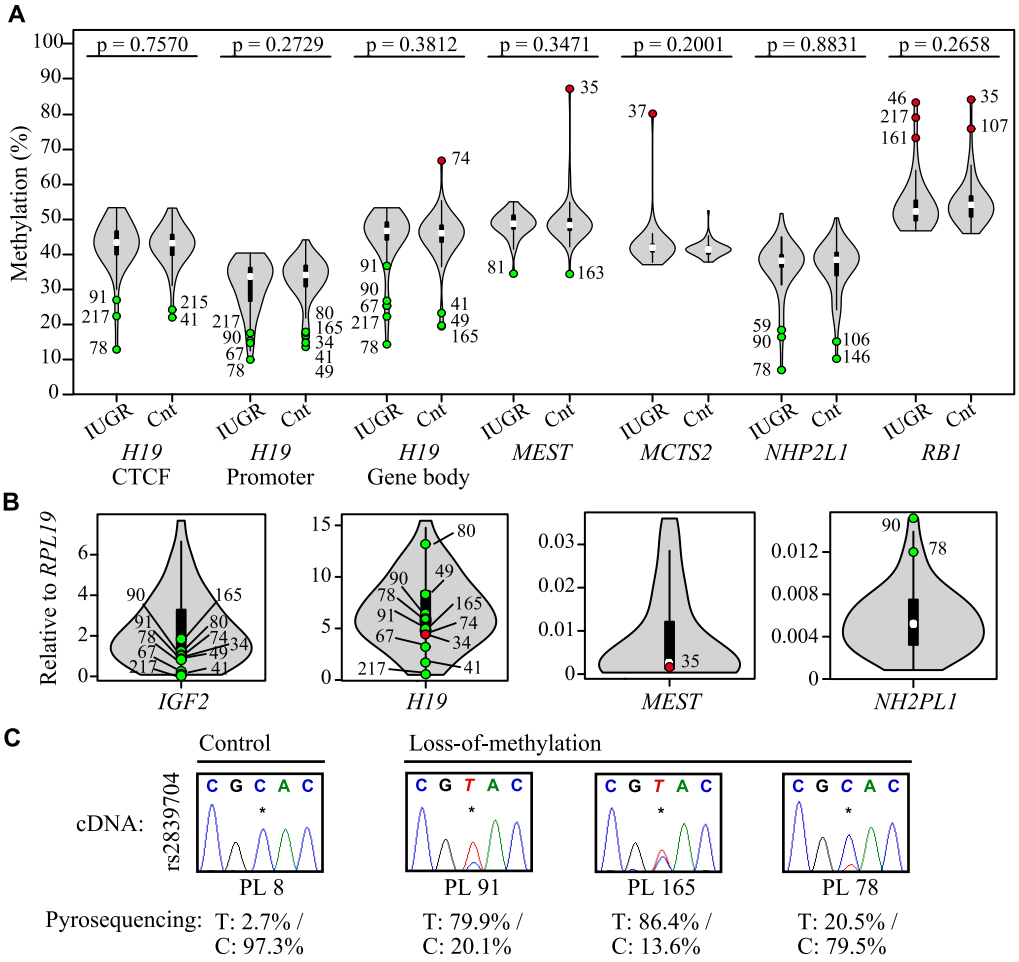



Figure 4.38: **Quantitative analysis at methylation and expression levels in ubiquitous imprinted regions.** Violine Plot used (**A** and **B**) include the mean (white dot) and the interquartile range (black box). Additionally, Violine plot also show a rotated kernel density plot on each side (non-parametric way to estimate the probability density of the data at different values) – *generated in R with the help of Dr Hernande-Mora* –. (**A**) Quantitative pyrosequencing of 5 ubiquitous regions (3 of them corresponding to *H19* region) in IUGR (n = 47) and non-IUGR (n = 57) placenta samples. Violin plot shows the distribution of DNA methylation values for all samples, green-filled dots corresponds to LOM placentas and red filled dots to GOM. At the top of two compared groups for each region, we can find the p-value of the Mann-Whitney-Wilcoxon non-parametric test. (**B**) Quantitative RT-PCR for *IGF2, H19, MEST* and *NH2PL1* in 50 placenta samples. Violin plot shows the distribution of expression value normalised with *RPL19* gene for all samples, green-filled dots corresponds to expression value for the corresponding LOM placentas and red filled dots for GOM. (**C**) On the top, allelic expression analysis for *H19* gene. The asterisk* on the electropherogram highlights the position of the SNP. On the bottom, its corresponding quantitative SNP analysis (% of presence of each allele in the RT-PCR) by pyrosequencing.

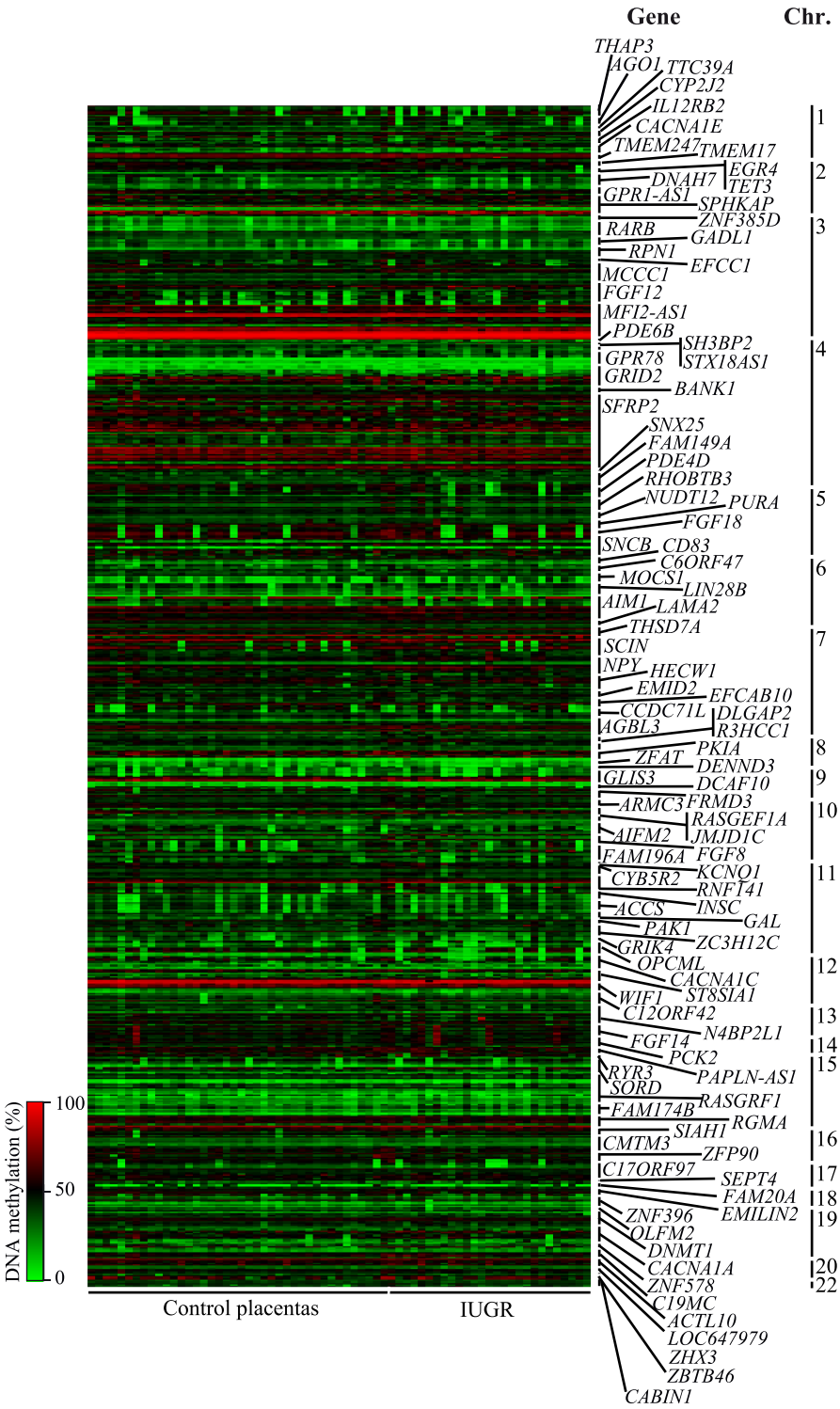



Figure 4.39: **Genome-wide methylation analysis in 77 placenta samples at all placenta-specific DMRs.** A Heatmap of 763 Infinium array probes mapping to 112 placenta-specific imprinted DMRs for 77 placentas, including IUGR and ART samples *(data analysis and graph generated by Dr Hernandez-Mora).*

In total 134 sample were assessed that included biopsies from complicated pregnancies (IUGR n = 51 ; ART n= 41) and controls (non-IUGR n = 77; non-ART n = 88, note groups are not mutually exclusive). By using the Mann-Whitney-Wilcoxon non-parametric test, we found no significant differences between the methylation levels in the IUGR group compared to non-IUGR samples for most of the regions, with only the exception of *AIM1* DMR (p-value = 0.038) and *N4BP2L1* DMR (p-value = 0.046) showing in both regions, an elevated mean of methylation in IUGR samples compared with non-IUGR (Figure 4.40A). These initial results and the observation of LOM placenta samples present in both complicated and control groups, suggest that methylation status at placenta-specific DMRs are not influencing in IUGR in our placenta cohort.

By using the same statistical test to compare ART and non-ART placenta samples, our results also suggested that the use of ART is not associated with changes in methylation of placenta-specific DMRs at our placenta cohort. Nevertheless, two regions, *DNMT1* (p value= 0.041) and *C19MC* DMRs (p-value = 0.043), shows statistically significant differences in population mean comparing ART samples vs non-ART, but this is due to an increase in variation of methylation in non-ART rather than a significant change in ART samples (Figure 4.40B). In addition to comparing the mean methylation for each group, we also compared whether there were differences in the frequency of samples with low methylation at each placenta-specific DMR. No single region presented with an increased frequency of LOM, with equal numbers of samples distributed between the IUGR, ART and control groups. The sample size was too small to observe any significant differences associated with pre-eclampsia (n = 6), but only one of these placenta samples had LOM at one single placenta-specific DMR (*SCIN*).

## 4.10 5-hmC in the placenta

We considered it necessary to confirm which methylated state was present at novel placenta-specific imprinted DMRs identified since standard bisulphite conversion and the methylation-sensitive genotyping using *Hpa*II cannot discriminate between 5-mC and 5-hmC. In a collaborative study with two other members of the laboratory (Dr Jose R. Hernandez and Dr Paolo Petazzi) we compared the abundance of 5-hmC in placenta and brain tissues using the TrueMethylCEGX oxBS kit combined with the Illumina Infinium HM450k platform. I was responsible for the confirmation of 5-hmC in the human placenta using the T4-BGT Epimark 5-hmC assay, and will only discuss the profiling of the placenta that is relevant for this thesis.

A bioinformatic comparison (performed by Dr Jose R. Hernandez-Mora) of four placenta samples and eight cerebellum biopsies, 5 of them with paired prefrontal cortex samples (Lunnon et al., 2016) revealed that the placenta had between 10 and 20 times less 5-hmC

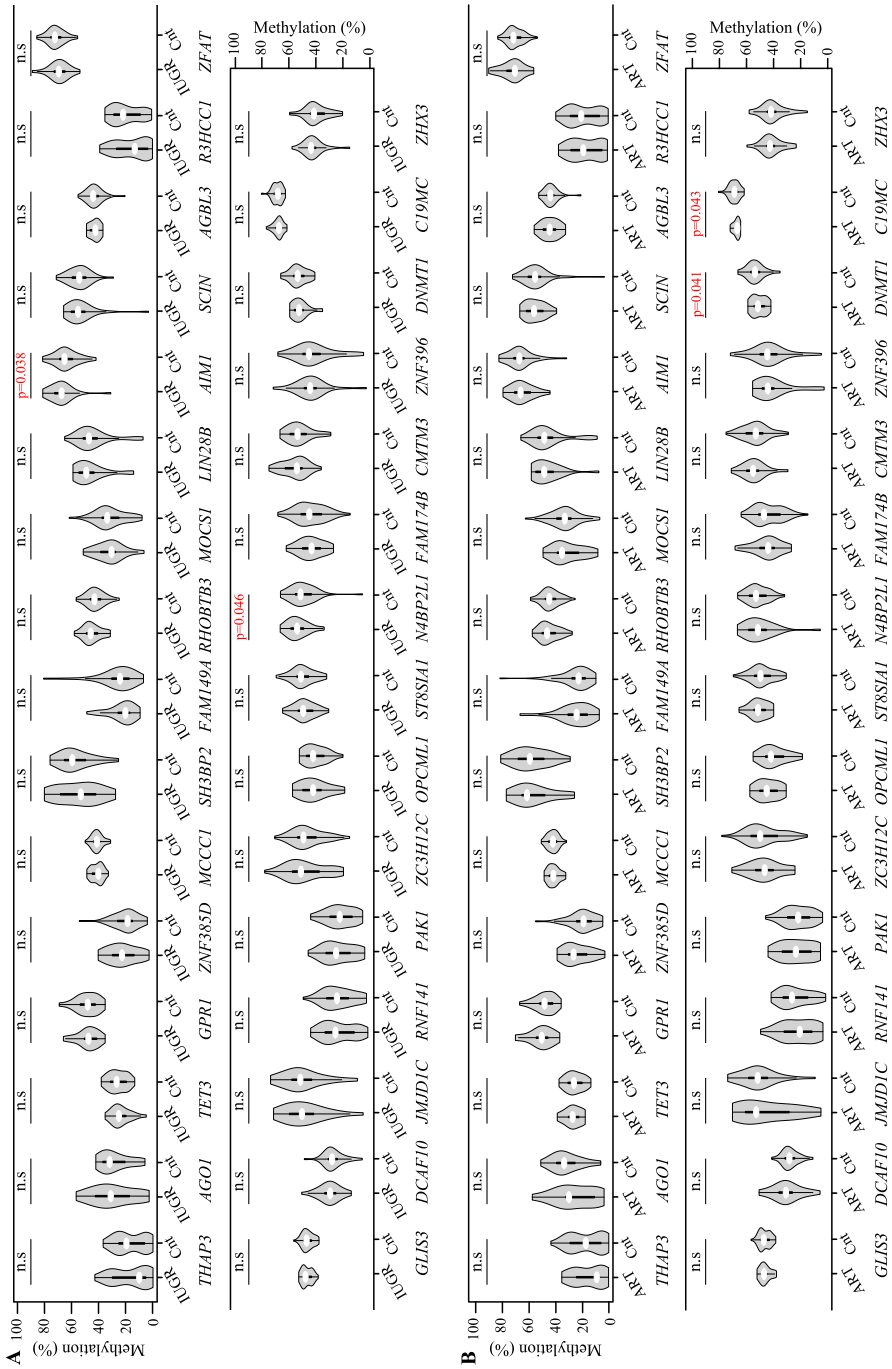



Figure 4.40: **Quantitative analysis at methylation levels in 31 placenta-specific DMRs.** Violine plot used (**A** and **B**) include the mean (white dot) and the interquartile range (black box). Additionally, Violine plot also show a rotated kernel density plot on each side (non-parametrtic way to estimate the probability density of the data at different values) – *generated in R with the help of Dr Hernande-Mora* –. Quantitative pyrosequencing of 31 placenta-specific DMRs were performed comparing (**A**) IUGR (n = 51) vs non-IUGR (n = 77), and (**B**) ART (n = 41) vs non-IUGR (n = 88) placenta samples.

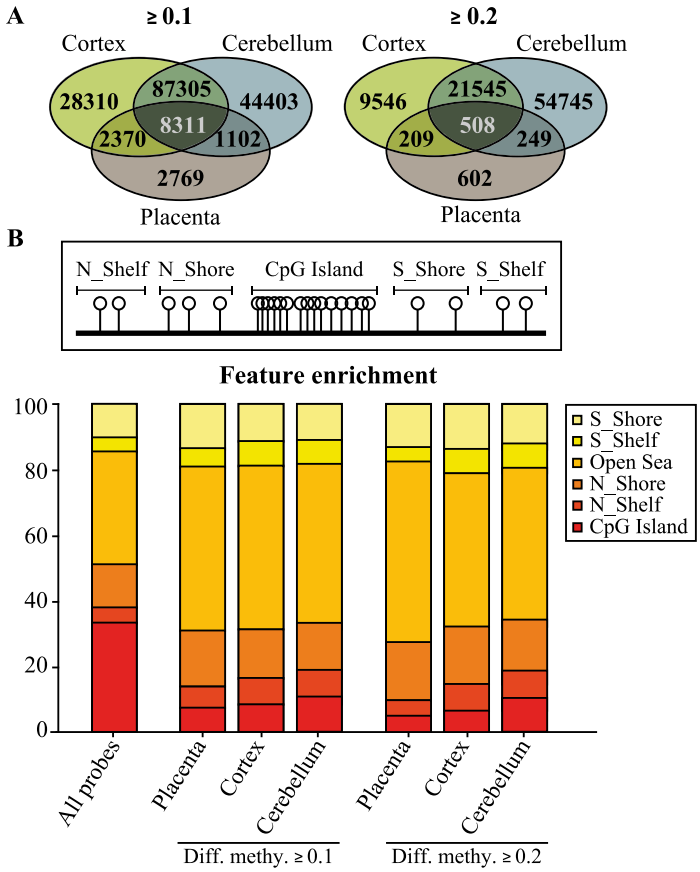



Figure 4.41: **Quantifying 5-hmC in human placenta and brain regions.** (**A**) Venn diagrams illustrating the number of probes detected using the threshold $\delta$ *beta* $> 0.1$ and $> 0.2$ after combined QC filtering for the three datasets. The figures represent the number of probes with an average greater than the threshold value in each tissue,. (**B**) All probes classified according to their genomic location. The bar chart illustrates probe enrichment classified by Illumina Infinium annotation in placenta, frontal cortex and cerebellu. (*data analysis and graph generated by Dr Hernandez-Mora*).

regions compared to brain (Figure 4.41A). The distribution of detected 5-hmC regions in the three different tissues is similar, being primarily located in "open sea" mapping outside defined gene and promoter intervals and CpG islands shelves, but depleted in CpG islands themselves (Figure 4.41B). To identify regions enriched for clusters of probes, a Bumphunter function was adapted for the Illumina Infinium HM450K array (Aryee et al., 2014). The top 500 candidate regions obtained from Bumphunter with a minimum of 2 probes for each tissue can be explored at our laboratory's website (http://www.humanimprints.net/). Interestingly only 48 are common in placenta and both brain datasets suggesting that the 5-hmC distribution is tissue-specific (Figure S11).

Focusing on placenta regions showing the highest levels of 5-hmC, we confirmed the

enrichment of 5-hmC in normal placenta biopsies by comparing our oxBS-450K data with published hMeDIP-seq datasets (L. Zhu et al., 2015). Ninety-seven of the top 100 placenta loci identified by oxBS-450K were also enriched for precipitated hMeDIP sequences (as an example, *SEMA3B* region on Figure 4.42A).

Subsequently, to confirm site-specific enrichment of 5-hmC, I utilised the T4-BGT Epimark 5-hmC assay based on *Msp*I enzyme digestion, which in combination with its methylation-sensitive isoschizomer *Hpa*II, allows for the discrimination of 5-mC from 5-hmC. *Msp*I is methylation-insensitive and can recognise its target C*CG*G irrespectively of any methylation, except in presence of a glucose moiety deposited by T4-BGT on 5-hmC. Following treatment with T4-BGT, the resulting DNA can be used for either qPCR and allele-specific PCR. Initial experiments confirmed the presence of 5-hmC at *SEMA3B* and mirR193B, two loci that revealed to have abundant 5-hmC/5-mC levels in placenta (Figure 4.42B)

Since some of the regions with the most abundant 5-hmC mapped to imprinted regions, I wished to determine if this epigenetic mark is located on one allele and present alongside 5-mC. In the placenta, the oxBS HM450k analysis revealed 5-hmC enrichment at *GNAS A/B* DMR, *H19* gene body – which is paternally methylated solely in placenta (Court et al., 2014) – and some placenta-specific DMRs (see Figure 4.43A for *MCCC1* region as an example). Quantitative assessments suggest that >20% of methylation in these regions is 5-hmC. However, quantification by using T4-BGT assay and qPCR shows that 5-hmC may be less abundant, but still reproducibly detectable and located on the maternally inherited allele at the *MCCC1, SCIN* and *ACTL10* and on the paternal allele at the *H19* gene body (Figure 4.43B-C).

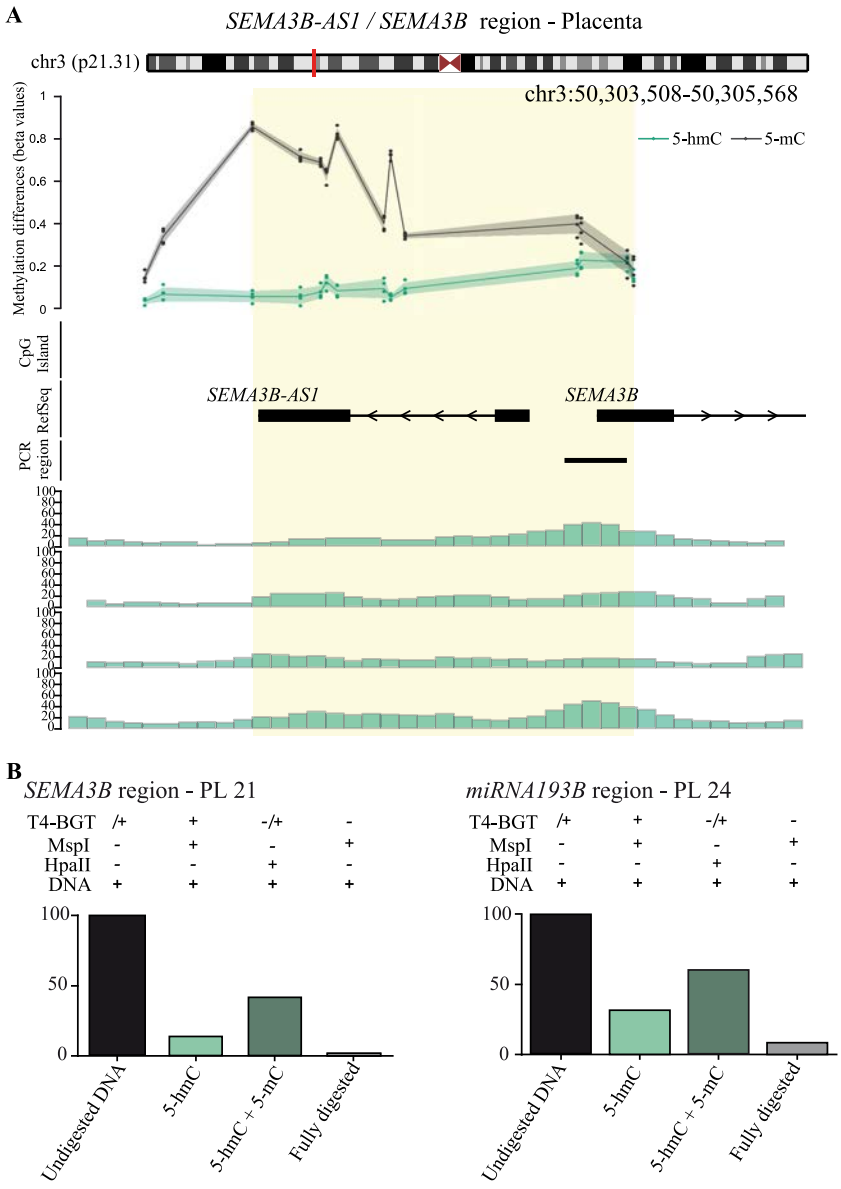



Figure 4.42: **Characterization of 5-hmC positive loci in placenta.** (**A**) Genomic interval overlapping the *SEMA3B* region on chromosome 3 showing the average distribution of 5-hmC (green) and 5-mC (black) in placenta samples. The upper panel represents the enrichment defined by oxBS-450K analysis with dots signifying data points for each probe (5-hmC $\delta\beta$; 5-mC oxBS) and the 95th confidence interval. The lower panel shows the corresponding hMeDIP enrichment in 4 control placenta samples. The Bumphunter-defined interval is highlighted as an orange shaded region (*data analysis and graph generated by Dr Hernandez-Mora*). (**B**) Quantitative PCR combined with T4-BGT assay targeting the *SEMA3B* and *miRNA193B* promoters confirm enrichment of both 5-hmC and 5-mC in placenta (PL).

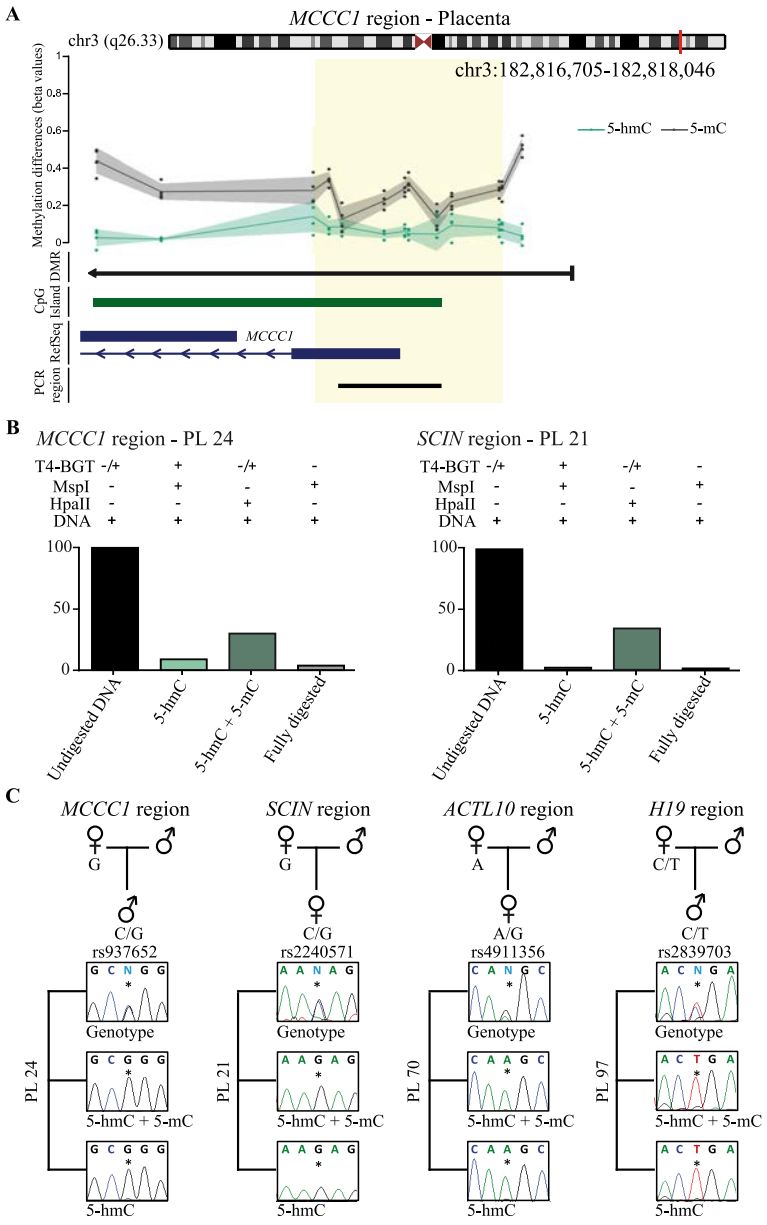



Figure 4.43: **Enrichment of 5-hmC overlapping placenta-specific imprinted DMRs.** (**A**) Genomic interval overlapping the *MCCC1* region showing the average distribution of 5-hmC (green) and 5-mC (black) in placenta samples. The panel represents the enrichment defined by oxBS-450K analysis with dots signifying data points for each probe (5-hmC $\delta\beta$; 5-mC oxBS) and the 95th confidence interval. The Bumphunter-defined interval is highlighted as an orange shaded region (*data analysis and graph generated by Dr Hernandez-Mora*). (**B**) Quantitative PCR combined with T4-BGT assay targeting the *MCCC1* and *SCIN* promoters confirm low levels of 5-hmC and higher of 5-mC in the placenta. (**C**) Analysis of allelic 5-hmC and 5-mC using T4-BGT genotyping in placenta-derived DNA (PL). The sequence traces of PCR amplicons generated using *Hpa*II digested DNA (representing 5-mC and 5-hmC) and *Msp*I in presence of T4-BGT (5-hmC) PCRs are shown. The asterisk* on the electropherogram highlights the position of the SNP.

# Chapter 5

# Discussion

## 5.1 NLRP family as members of the human SCMC

Studies in female mice that carry targeted mutations of SCMC members (*Ooep*, *Nlrp5*, *Tle6* or *Khdc3l*) revealed that these mutations act via maternal effect inducing cleavage-stage embryonic arrest and female sterility (L. Li et al., 2008). In humans, maternal loss-of-function of the genes encoded SCMC protein members results in reproductive problems: mutations in *KHDC3L* cause RHM (Parry et al., 2011); mutations in *NLRP5* are associated with reproductive wastage and MLID (Docherty et al., 2015); while maternal mutations in *PADI6* and *TLE6* cause early embryonic arrest and lethality by affecting progression of oocyte meiosis II through to zygote formation (Alazami et al., 2015; Y. Xu et al., 2016). Moreover, molecular characterisation of the SCMC has suggested that additional proteins may be associated with the complex as fast protein liquid chromatography (FPLC) analysis revealed that the molecular weight of human SCMC is higher than the sum of known components, with the NLRP family members being one of the best candidates (B. Kim, Kan, Anguish, Nelson, & Coonrod, 2010).

The *NLRP* genes have arisen from evolutionary duplications before the divergence of mammals, in addition to lineage-specific duplications in primates and rodents. Specifically, *NLRP7* is primate-specific and it probably originates from a direct duplication of *NLRP2*, the common ancestor in mouse-primate lineages (X. Tian, Pascal, & Monget, 2009). Recently, it was demonstrated in mice with *Nlrp2* over-expression, that this protein interacts with the SCMC proteins TLE6, OOEP, KHDC3L and NLRP5 (Mahadevan et al., 2017). In the same study, oocytes from *Nlrp2*-null females had altered the subcortical localisation of TLE6 SCMC protein. In addition, rare mid-gestation embryos derived from these targeted oocytes exhibit methylation abnormalities at some imprinted loci. In humans, NLRP2 and NLRP7 have similar protein localisations in oocytes and comparable

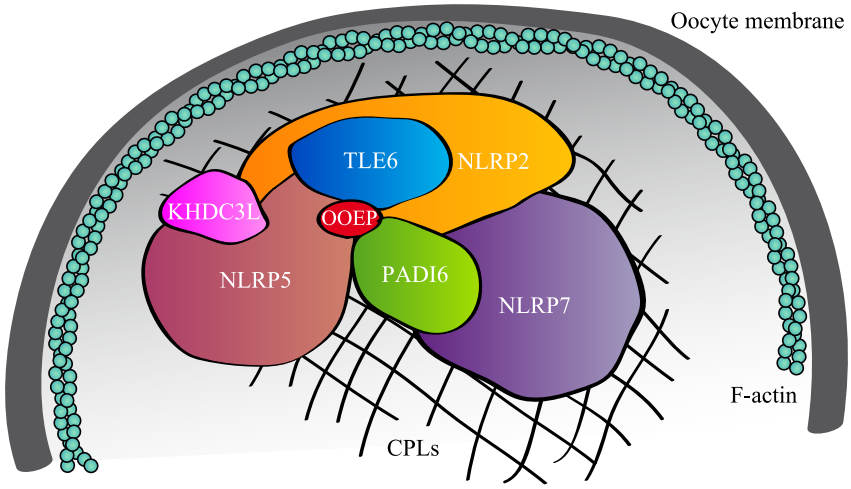



Figure 5.1: **Proposed members and structure of the human SCMC.** The SCMC interacts with both F-actin and the oocyte cytoplasmic lattices (CPLs). The schematic representation includes the five SCMC members already confirmed, and NLRP2 and NLRP7, which have been proposed as members of SCMC but addition co-immunoprecipitation experiments will be necessary to test its interactions with the other members of the complex (Monk, Sanchez-Delgado, & Fisher, 2017).

phenotypes when mutated (B. Kim et al., 2010). Additionally, *NLRP2*, *NLRP5* and *NLRP7* are highly expressed in human/macaque ovary, oocytes and preimplantational embryos, and NLRP7 co-localises with KHDC3L to the cortical region in human oocyte and 2-cell embryos (Akoury, Zhang, Ao, & Slim, 2014; P. Zhang et al., 2008). For these reasons, it was recently suggested that NLRP2 and NLRP7 may also be SCMC proteins (Figure 5.1) (Mahadevan et al., 2017; McDaniel & Wu, 2009; Monk, Sanchez-Delgado, & Fisher, 2017).

Current work in the laboratory supports the suggestion that NLRP7 could be part of the SCMC, and our observations that maternally methylated imprints are absent in moles inheriting maternal mutations, whilst paternally methylated DMRs are unaffected, implicate NLRP7 in oocyte-specific methylation establishment. Both NLRP7 and KHDC3L have been described to co-localise with *de novo* DNA-methyltransferases DNMT3A-B in the cytoskeleton of the oocyte and early human embryo (Akoury et al., 2014). In the same line, *in vitro* experiments have shown that loss of maternal *Nlrp2* in mice results in increased follicular atresia in ovaries of mutant females, indicating that *Nlrp2* mutation is affecting in the oocyte development (Mahadevan et al., 2017). This observation suggests that a possible function of SCMC is to regulate imprint acquisition (Alazami et al., 2015).

Nevertheless, the dynamic nature and possible SCMC involvement in imprinting establishment may require interactions with specific epigenetic regulator factors. It is possible that additional complexes with a different constitution of SCMC proteins ensure normal

methylation at some imprints. In our cohort of RHMs we observe that some DMRs are affected in some individuals and not others (e.g. *PEG10* region shows a normal imprinted pattern in one of the five RHM studied). Since the RHMs with opposing *PEG10* methylation is observed in consecutive pregnancies from the same patient, this observation cannot be explained by a simple inter-individual genetic effect. Searching in the literature, I found one possible explanation for this phenomenon. Studies of gene-specific timing and epigenetic memory in the mouse oocyte revealed that the different imprinted DMRs are established at different development time-points in the oocyte and that this timing was influenced depending on which grand-parental chromosome was inherited (Lucifero, Mann, Bartolomei, & Trasler, 2004). Indeed, this study focused on the timing of methylation acquisition at the *Snrpn* DMR in mouse oocytes by using strain-specific polymorphisms in 10 days postpartum (dpp) oocytes. This revealed that 67% of strands from oocytes with the grand-maternal inherited allele were methylated while no methylated strands were found from the oocytes with the grand-paternal inherited allele. These experiments suggest that the non-equivalence of the parental alleles may be due to the presence of a pre-existing epigenetic mark. If this phenomenon also exists in humans, it is a possible explanation for the incomplete imprinting defects in our studied RHM. If we hypothesise that NLRP7 protein plays a major role in the oocyte methylation, it would require the function of the other player(s) to distinguish the additional epigenetic marks on the grand-parental copies. Unfortunately, the *PEG10* DMR does not contain any SNPs to allow for the grand-parental alleles to be discriminated in our NLRP7 mutated female patients.

It is important not to forget that NLRP family members are involved in the *inflammasome response*, a component of the innate immune system (Latz, Xiao, & Stutz, 2013). This means that, despite observing no effect on imprinted DMRs in blood-derived DNA from women carrying a biallelic recessive mutation of *NLRP7* in our study, we cannot reject an indirect role of inflammation in genomic imprinting. The SCMC and NLRPs may be involved directly in ovarian function by regulating follicular and oocyte development, which is consistent with data generated in *Nlrp2*-null mice experiments (Bukovsky, 2006; Mahadevan et al., 2017). In addition, inflammasome activation results in the secretion of interleukin-1$\beta$ (IL-1$\beta$) that is a pro-survival factor required for oocyte nuclear maturation in many mammalian species (Caillaud, Duchamp, & Gérard, 2005; Messaed et al., 2011). Therefore, NLRP7 could influence maternally derived methylation by an indirect mechanism, regulating the process of oocyte selection during foetal development. There is a substantial reduction in oocyte number that occurs at the approximate time (14-20 weeks gestation) when oocytes presumably start to acquire the methylation signatures at imprinted regions (Gkountela et al., 2013). It is therefore plausible that disruption to the selection mechanism through defective NLRP7 may allow for the survival, and eventual dominant follicle recruitment, ovulation and fertilisation decades later, of an oocyte with an inappropriate methylation state.

However, a possible role in imprint maintenance it cannot be completely ruled out since it is likely that mutations in *NLRP2/Nlrp2* or *NLRP5/Nlrp5* are associated with early development problems in mice, human and macaque (Docherty et al., 2015; Mahadevan et al., 2017; Meyer et al., 2009; X. Wu, 2008). Furthermore, it must be noted that both maternally and paternally methylated imprinted DMRs are affected in MLID patients resulting from *NLRP5* mutations, which indicates a role in methylation maintenance of the SCMC or, a maintenance role specifically for NLRP5 protein, together with other non-SCMC factors (Docherty et al., 2015). Additionally, experimental evidence also suggests a potential role of murine NLRP2 in methylation maintenance, with the abnormal localisation of DNMT1 in oocytes from *Nlrp2*-null mouse females (Mahadevan et al., 2017). In summary, we hypothesise that the SCMC is an important complex for imprint acquisition in the oocyte with support functions in methylation imprint maintenance in the early embryo.

## 5.2 Abundance of placenta-specific DMRs in humans

Initially, it was thought that tissue-specific imprinting was only at the expression level since germline-derived DMRs are present in a parent-of-origin manner in all somatic tissues, but not all are associated with expression (Sanz et al., 2008). In 2005, complex imprinting of *Igf2r* was reported in mouse brain cell types (Yamasaki et al., 2005). In most tissues, this gene is maternally expressed and its promoter paternally methylated. On the contrary, in the brain, it is biallelically expressed and hypomethylated (DeChiara et al., 1991; Lui, Finkielstain, Barnes, & Baron, 2008). This observation suggested that tissue-specific imprinting was also regulated at the methylation level, leading researchers to screen for new imprinted genes by both expression and methylation strategies, mainly in the two tissues where imprinting was reported to have a key role: the brain and the placenta (Kelsey, 2011; A. Wagschal & Feil, 2006). Subsequently, it was reported that there was a general lack of conserved tissue-specific imprinting between eutherian species, with the first human placenta-specific imprinted locus, the *C19MC*, being absent in mice (Monk et al., 2006; Noguer-Dance et al., 2010). In 2011, a genome-wide DNA methylation profiling of human placentas triploids, CHM and control placentas, described many potential regions acting as DMRs in the human placenta, observations that were confirmed by two subsequent publications in 2013 and 2014 (Yuen, Jiang, Peñaherrera, McFadden, & Robinson, 2011). The first, reported two new maternally methylated regions and associated with the promoters of *DNMT1* and *AIM1*, resulting in paternal expression (Das et al., 2013). Months later, members of our laboratory identified 15 additional placenta-specific DMRs by methyl-seq and high-density methylation arrays comparing CHM and control placentas (Court et al., 2014).

When we initiated this project, we hypothesised that most of the placental-specific DMRs acquired their methylation pattern during early placental development. This was based on the unmethylated state of all placenta-specific DMRs in standard hES cells and phES, an *in vitro* parthenogenetically-activated cell model system which has previously been used to study aspects of oocyte biology (Court et al., 2014). As a consequence, we classified the placenta-specific regions as either placenta-specific DMRs Type 1 regions, which inherits methylation from the oocyte (showing methylation in phES) and becoming fully methylated in somatic tissues, or Type 2 placenta-specific DMRs, which potentially acquired methylation pattern during early placental development and remained unmethylated in somatic tissues (Figure 5.2) (Court et al., 2014). However, within the first year of my thesis, methyl-seq datasets obtained for mature human oocytes revealed that placental-specific DMRs inherit maternal methylation from the female germline (H. Guo et al., 2014). This conclusively showed that the phES/hES cell model does not faithfully maintain the epigenetic landscape of a mature oocyte-early embryo. Curiously, most placenta-specific DMRs are unmethylated not only in phES, but also in established placenta-derived cell lines (JEG3, JAR, SWAN71, TCL1), while ubiquitous DMRs are faithfully maintained (*personal communication – D. Monk and A. Monteagudo-Sanchez–*). This suggests that despite looking similar at the epigenetic level, specific maintenance factors are required for placenta-DMR that are absent in these cell lines. This makes it especially difficult to study these regions *in vitro*. However, it has been recently established a derivation of trophoblast stem cells from both blastocysts and first-trimester placenta samples that maintains placenta-specific DMRs (Okae et al., 2018).

During the course of this dissertation, I confirmed a total of **72** novel monoallelically methylated regions in placenta, of which **50** were confirmed methylated on the maternal allele. All candidate regions identified were devoid of methylation in somatic tissue except *TMEM247* DMR, which was fully methylated in non-placenta tissues. Moreover, our bioinformatics analysis reveals additional candidates that we were not able to analyse due to the lack of informative polymorphisms for allelic discrimination (a further **27** candidates in the first screening and **513** in the second one). Allele-specific RT-PCR confirmed the monoallelic expression of **20** genes (**17** of them paternally expressed) whose promoters contained a placenta-specific DMRs. One of the regions identified was associated with *TTC39A* gene. Whilst this gene has not previously been reported as imprinted, the adjacent transcript *EPS15*, was listed as a candidate imprinted gene in the placenta (Pozharny et al., 2010). However, our allelic RT-PCR expression analysis revealed biallelic expression of *TTC39A* transcript in all samples tested and 11 biallelic and one preferentially monoallelic sample for *ESP15*. In addition, we identified a maternally methylated CpG island not included in our list of candidates. This region overlap the promoter of *SNCB*, which was selected for analysis since it had previously been described as paternally expressed in the placenta (Metsalu et al., 2014).

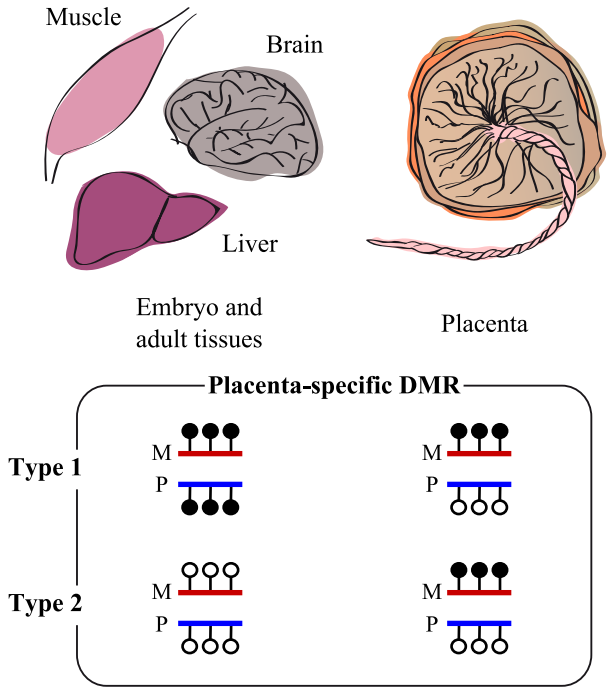



Figure 5.2: **Methylation status of placenta-specific DMRs in somatic tissues and the placenta.** Even though all placenta-specific DMRs inherits its methylated allele from the oocyte, we can classify these regions as Type 1, fully methylated in somatic tissues and Type 2, fully unmethylated in somatic tissues.

The samples and the whole genome methylation profiling strategies used to generate the data in results' sections 4.2 and 4.3 were not very different from those described by our laboratory in 2014 (Court et al., 2014). However, this raises the question why the screening for placenta-specific DMRs was more successful during my project (a total of **612** compared to **44** candidates by Court *et al.* ). The answer is due to the highly restrictive bioinformatic filtering analysis used in the first analysis. In the original study, a sliding window approach was used to identify regions containing more than 25 CpGs within the 0.25-0.75 range after filtering for repetitive sequences in methyl-seq samples (Court et al., 2014). Furthermore, regions were only selected if it contained 3 consecutive Illumina HM450k array probes with average was between 25% and 75% methylation, within 1.5 SD of the mean in control placenta samples and <0.2 in CHM. Using these parameters, it was not possible to identify polymorphic placenta-specific DMRs in our placenta cohort. Our informatic approaches, as described in material and methods section of this thesis, allowed us to find a more significant number of candidates and to describe the polymorphic phenomenon. These observations were confirmed by other groups performing similar relaxed bioinformatics analyses (Hamada et al., 2016; Hanna et al., 2016). Furthermore, the choice of molecular technique to confirm allelic methylation also allowed for the identification of polymorphically unmethylated samples. Whereas we first genotype a subset of placenta samples for each region and then, we performed the allelic methylation analysis only in the heterozygote cases, both Hanna et al. (2016) and Hamada et al. (2016) performed a targeted bisulfite PCR-amplicon sequencing (Illumina MiSeq or Hiseq system, depending on the study). They first interrogate the methylation status at a strand-specific level in all their samples and then selected the heterozygous ones to show parental-specific methylation. Using such strategies, they were able to largely describe the polymorphic methylation phenomenon (Hamada et al., 2016; Hanna et al., 2016).

The screening technique employed to identify whole genome DNA methylation was also different between publications. Comparing the candidate list of each study, those using the Illumina HM450K array (including the first analysis performed in the present work and Court et al. (2014)) or reduced-representation bisulfite sequencing (RRBS – used by Hanna et al. (2016)), the output was more limited than using WGBS (Hamada et al., 2016). Using WGBS, Hamada et al. (2016) reported approximately 18 times more maternal methylated gDMRs in the placenta than was reported using more targeted techniques. However, this discrepancy is highlight by the fact that most placenta-specific DMRs identified in our first analysis and Hanna et al. (2016) were Type 2, mapping to CpG-rich promoters (which have preferentially more probes on the HM450k array or enriched for *Msp*I sites for RRBS), whilst the vast majority of the candidates identified by WGBS by Hamada et al. (2016), were potentially, Type 1 placenta-specific DMRs, being intergenic and highly methylated in blood sample. This suggests that the Type 1 placenta-specific DMRs are located in non-CpG islands regions, which is consistent

with the fact that two of the four previously reported Type 1 placenta-specific DMRs (*TEMEM247* reported in this thesis and *GPR1-AS* (Court et al., 2014)) are not located within a CpG island (Figure 5.3). In our second approach we also utilised WGBS datasets, but we mainly got Type 2 placenta-specific DMRs candidates because our analysis was focussed on searching for new candidates in CpG islands and promoters regions, trying to obtain a more accurate list of DMRs regulating gene expression.

## 5.3 Isolated regulation of placenta-specific DMRs

Placenta-specific DMRs only regulate the monoallelic expression of the nearest transcript, and therefore, unlike the ubiquitous DMRs, these tissue-specific DMRs represent true micro-imprinted domains. This fact could indicate that placenta-specific imprinted genes and known ubiquitous imprinted transcripts are associated with different epigenetic mechanisms that influence the local genome architecture. The placenta is a unique organ where genes are differentially regulated. For example, during the present thesis, I was involved in a study to compare the 5-hmC states in placenta and brain tissues. Overall the 5-hmC distributions is different in both tissues, being very abundant in brain an relatively sparse in placenta. In brain, we described high 5-hmC in gene bodies of expressed imprinted transcript (consistent with previous studies which associates 5-hmC in gene body with expression), whilst in the placenta, we only where able to identify 5-hmC on the methylated allele of a few promoter DMRs (Hernandez-Mora et al., 2017). This may reveal a different general epigenetic regulation in the placenta, which would include the imprinted regions.

Of the more than twenty regions described in our study or previously reported, we did observe that three placenta-specific DMRs potential influenced allelic expressing of flanking genes, forming an imprinting cluster. First, the *C19MC* DMR which regulates the paternal expression of the largest human microRNA gene cluster and is ~100 kb from the paternally expressed *ZNF331* gene. Additionally, *ZNF331* gene is associated with two ubiquitous DMRs (Court et al., 2014; Noguer-Dance et al., 2010).

The location of one of the placenta-specific regions described in this thesis was surprising since it is within a known and highly conserved imprinted cluster associated with the *KCNQ1* gene. Within intron 8 of the *KCNQ1* gene is the KvDMR1, a ubiquitous maternally methylated DMR overlapping the promoter of the *KCQN1OT1* ncRNA. The placenta-specific DMR we identified is located in intron 11 of *KCNQ1*, ~92 kb downstream of the KvDMR1, and is not conserved in mouse. This new DMR is associated with an isolated transcript in placenta RNA-seq datasets that we could not link to the *KCNQ1* or *KCQ1DN* transcripts by RT-PCR, suggesting that it is an isolated ncRNA and not an alternative exon of *KCNQ1* or *KCQ1DN*. Specifically, this expression of this ncRNA

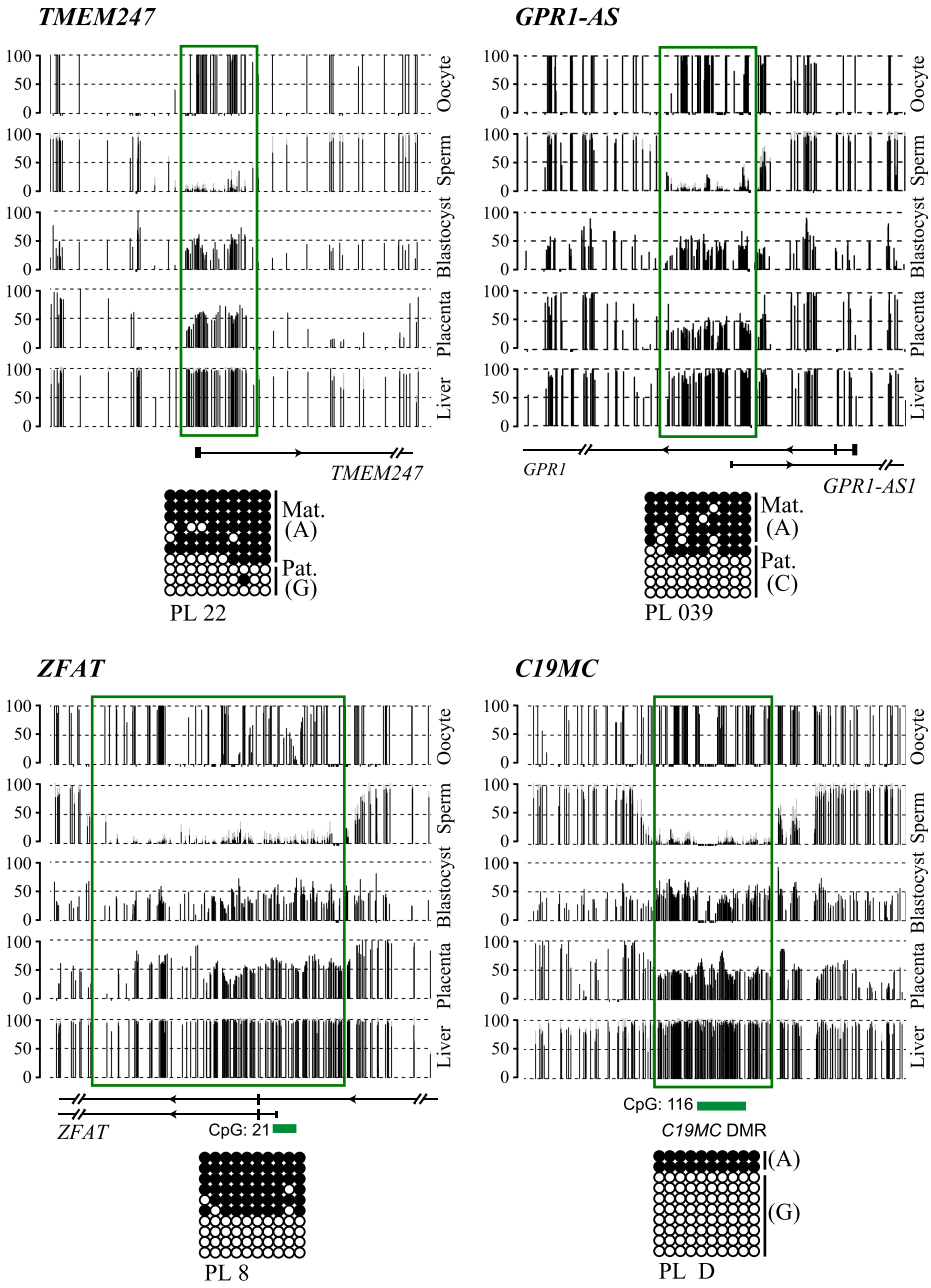



Figure 5.3: **Characterization of the four Type 1 placenta-specific DMRs reported before 2016.** The vertical black lines in the methyl-seq tracks represent the mean methylation value for individual CpG dinucleotides. The green boxes highlight the position of the gDMRs. Bisulphite PCRs on placenta derived-DNA were used for confirmation. Each circle represents a single CpG dinucleotide on a DNA strand, black circles for methylated cytosines and white for unmethylated cytosines. If informative, the parental-origin of methylation is indicated. For clarity, only the first CpG dinucleotides are shown.

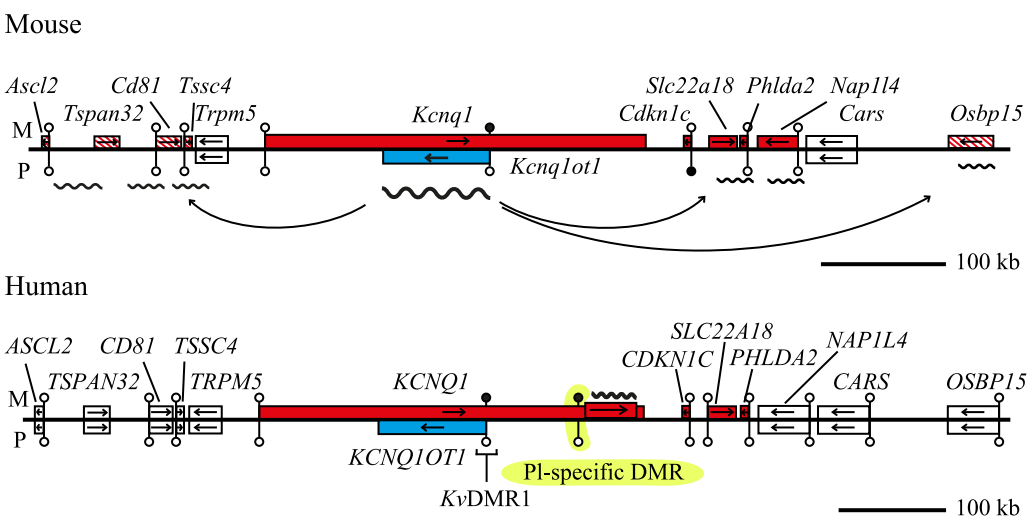



Figure 5.4: ***KCNQOT1* imprinted cluster in the mouse and human placenta.** Schematic representation showing the relative organisation of genes, its expression in placenta, CpG islands and DMRs for the *KCNQOT1* domain on mouse chr7qF5 and human chr11p15. The upper part of the line represents maternally inherited chromosome (**M**) and lower part paternally inherited chromosome (**P**). Black lollipops represent methylated CpG regions whereas white cycles unmethylated CpG regions. Red filed rectangle represents maternally expressed gene. Blue filled rectangle paternally expressed gene and white filled rectangle biallelic expression of anon-imprinted gene. The arrows represent the direction of the transcripts. Note that a placenta-specific DMR is present exclusively in the human placenta and associated with the paternal expression of a new transcript.

was maternally derived in five of our placenta samples, five additional samples showed monoallelic expression and two were biallelic (*unpublished data*) (Figure 5.4).

In addition, we also reported an imprinting cluster at the highly conserved *GPR1/ZDBF2* domain. The *Zdbf2* genes was initially reported to paternally expressed in different embryonic and adult mice tissues but not in the placenta (H. Kobayashi et al., 2009). Subsequent publications in mice described a maternal gDMR which regulates the paternal expression of *Liz* (a long isoform of *Zdbf2*, which some authors also named it as *Zdbf2linc*) in the early embryo. This specific gDMR maintains its imprinted methylation in extra-embryonic tissues but gains methylation on the paternally inherited allele in embryonic somatic tissues (Duffié et al., 2014). Until now, this is the only placenta-specific reported to be conserved in mammals as *GRP1-AS* is the human orthologue of *Liz* (despite having different size, see Figure 5.5) (Duffié et al., 2014; H. Kobayashi et al., 2013). Previously, it was reported that *ZDBF2* has a tissue-specific imprinted expression profile, being paternally expressed in human lymphocytes but not in the placenta (H. Kobayashi et al., 2009). However, several subsequent publications, including the results described in this thesis, have been shown that *ZDBF2* is paternally expressed in placenta, as is *GPR1-AS* (Hiura et al., 2010). In addition, we show that the gene adjacent to *ZDBF2, ADAM23*

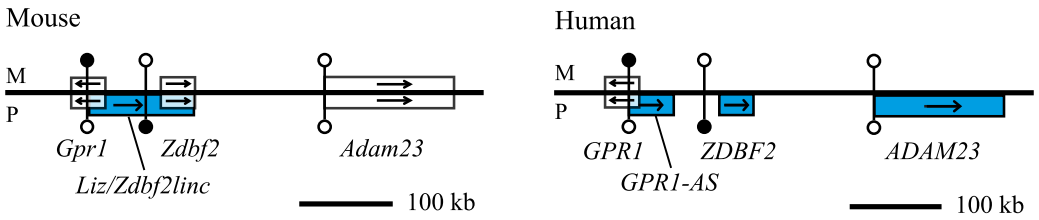



Figure 5.5: ***GPR1/ZDBF2* imprinted cluster in the mouse and human placenta.** Schematic representation showing the relative organisation of genes, CPG islands and DMRs for *Gpr1/Zdbf2* mouse domain on chr1qC2, and human *GPR1/ZDBF2* domain on chr2q33.3. The upper part of the line represents maternally inherited chromosome (**M**) and lower part paternally inherited chromosome (**P**). Black lollipops represent methylated CpG regions whereas white cycles unmethylated CpG regions. Blue filled rectangle paternally expressed gene and white filled rectangle biallelic expression of a non-imprinted gene. The arrows represent the direction of the transcripts. The region contains two conserved DMRs, but the paternally derived expression of *ZDBF2* and *ADAM23* present in the human placenta is not conserved in mouse placenta.

is also paternally expressed in human placenta, while is it biallelically expressed in B6 x JF1 mouse hybrid placenta samples. However, in spite of not having conserved parental expression for *ZDBF2* and *ADAM23* between human and mouse, there are many common genomic features in common between species: *GPR1* gene is paternally expressed in some somatic tissues but not the placenta; the cluster presents both maternally and paternally methylated DMRs; the paternally methylated DMRs are present in all tissues; and finally, the maternally methylated DMR is placenta-specific Type 1, fully methylated in somatic tissues (Duffié et al., 2014; Hiura et al., 2010) (Figure 5.5).

Consistent with our results, a recent methylation and expression study of placenta-specific gDMRs show the paternally-derived expression of 30 candidates, with only three regions potentially associated with the regulation of more than one imprinted gene (Hamada et al., 2016).

## 5.4 The potential role of placenta-specific imprinted genes

Most placenta-specific genes are involved in cellular and metabolic processes or the regulation of these processes (Gene Ontology (GO) database), with almost half of these genes being involved in the regulation of gene expression. Although GO analysis did not reveal significant over-representation in any category in placenta-specific genes, the two biological processes most represented were cell-cell signalling, its regulation, and the regulation of immune system process. Only one placenta-specific imprinted gene, *ZFAT*, is involved in placental development. However, we found two additional genes, *SFRP2* and

*LAMA2* involved in the regulation of embryonic development. Moreover, there are several strong candidates for influencing trophoblast development not directly recognised by the GO Consortium, including the cytochrome P450 subfamily member *CYP2J2*, that has previously been shown to be up-regulated in preeclampsia and *THSD7A*, a placental and endothelial protein that mediates cellular migration (F. Herse et al., 2012; C.-H. Wang et al., 2010). In addition, the deregulation and over-expression of the *C19MC* miRNA cluster regulated by a placenta-specific DMR would lead to the concomitantly increased abundance of 50 mature miRNAs that have recently been shown to regulate trophoblast invasion (L. Xie et al., 2014).

Other essential genes regulated by genomic imprinting specifically in the placenta are involved in epigenetic regulation including the DNA methyltransferase *DNMT1*; the Jumonji Domain Containing 1C (*JMJD1C*), which is a strong candidate histone demethylase and is thought to be a coactivator for key transcription factors; or the micro-RNA processors *LIN28B* and *AGO1*. Also, although its imprinted expression has not been shown, other placenta-specific DMRs are in the promoter of genes involved in key developmental processes, including the FGF family members (*FGF8, FGF12* and *FGF14*) that regulate trophoblast survival and placental angiogenesis, and different members of the calcium voltage-gated channel subunit alpha1 (*CACNA1A, CACNA1C, CACNA1I* and *CACNA1E*), which are essential genes for the cell-to-cell communication by forming the calcium channels of the cell membranes.

It is important to be noted that the placenta is a very dynamic tissue, that is continuously adapting to the maternal-foetal environment, changing its gene expression depending on not only the developmental stage of the pregnancy but also as a response of other environmental changes (Sandovici, Hoelle, Angiolini, & Constância, 2012). Since these expression changes that the placenta undergoes are more related to morphological adaptations, vascularization and other basic biological functions (like metabolic processes, fatty, folic acid and calcium transport) than to placenta or foetal growth, it might be that some genes associated with placenta-specific DMRs are actively involved in this placenta plasticity.

## 5.5 Placenta-specific imprinting during embryonic development

Our study has revealed that all placenta-specific DMRs have oocyte-derived methylation, with some of them coordinating paternal expression following ZGA in the blastocyst, as highlighted by *ZHX3*. Unfortunately, no additional paternally expressed genes were identified in this embryo study due to the lack of informative polymorphisms. Extrapolating these observations means that there are potentially thousands more transiently imprinted

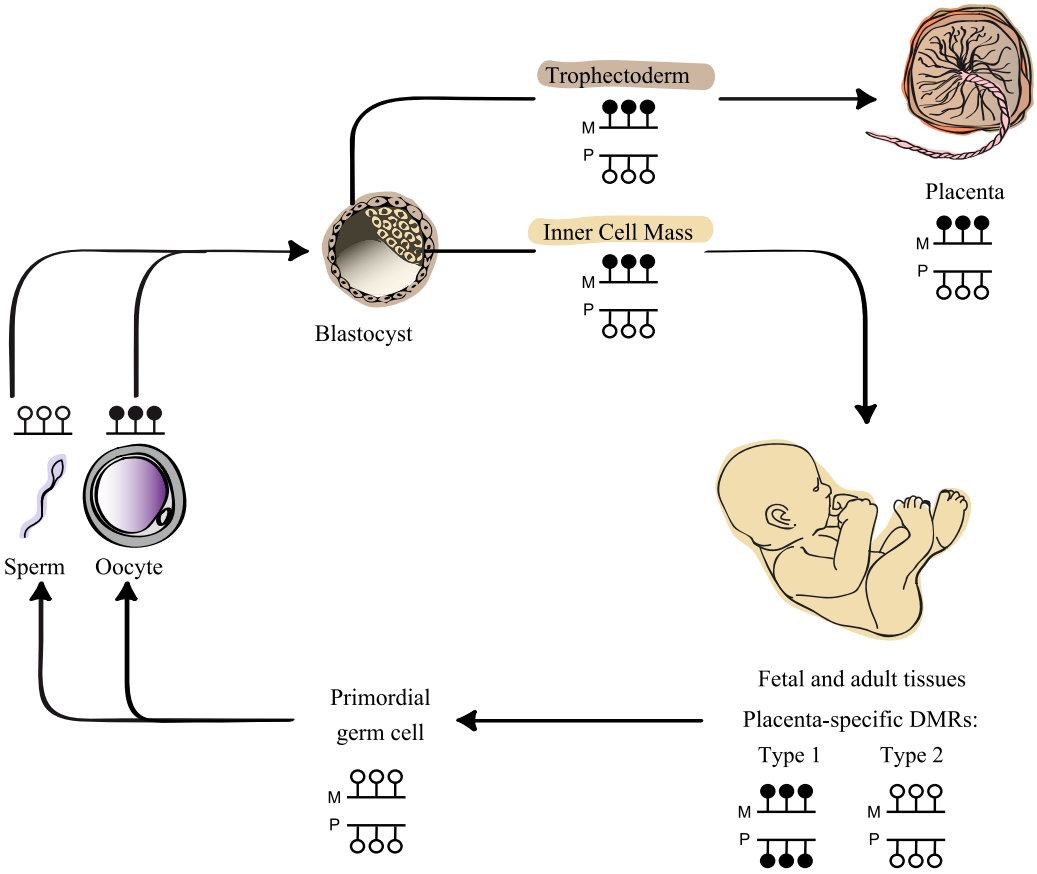



Figure 5.6: **Imprint placenta-specific methylation cycle.** Schematic representation of DNA methylation dynamics at placenta-specific imprinted DMRs from the gametes to the embryo, including blastocyst stage. Black lollipops represent methylated CpG whereas white cycles unmethylatyed CpG.

genes in the blastocysts that may have a physiological role in embryonic development.

We show that allelic methylation is present in both the ICM and TE of early preimplantation human blastocysts, revealing that 5-mC is selectively protected from embryonic reprogramming and that it is maintained following the first differentiation step (Figure 5.6). Furthermore, our data suggest that an additional small wave of targeted demethylation exists following implantation in cells specified for the somatic lineages that is absent during placenta differentiation. These observations have been noticed by others also. In 2016 it was reported the methylation state of these region in isolated extra-embryonic cell types, suggesting that placenta-specific DMRs remains ~50% methylated after blastocyst state in the first ICM derived cells, but are erased in the embryonic lineages take places before day 15 of gestation – data extrapolated from analyzing methylation status in different ICM derived extra-embryonic tissues: mesenchymal core of the placental villi, originated about

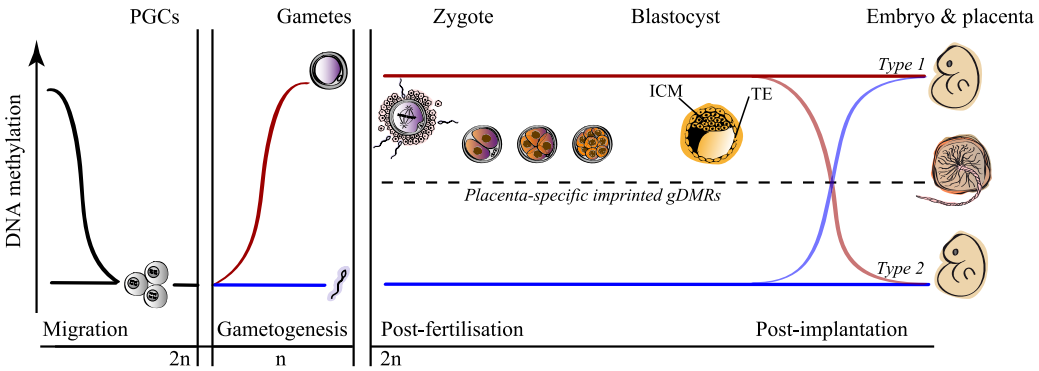



Figure 5.7: **Methylation dynamics at placenta-specific DMRs during development.** In the first two parts of the graph, the lines represent the general DNA methylation dynamics at placenta-specific imprinted DMRs regions in the gametes: black (general) blue (sperm) and red (oocyte), and after fertilisaion, it is represented in paralele, the methylation state at placenta-specific DMR regions in the genome embryo (2n, dashed black line), in the maternally inherited allele (red line) and the paternally inherited allele (blue line); the X axis represents the general stages of development in human, without a specific time-scale; the Y axis represents the DNA methylation state at the placenta-specific DMR, which are 50% methylated during early development (up to blastocyst stage) and during all placenta development (dashed black line), fully methylated at oocyte-derived copy (red line) during all embrionic development in the case of placenta-specific Type 1 DMRs or which a loss-of-methylation after blastocyste stage (the specific developmental time is unknown) for the Type 2, and fully unmethylated in sperm-derived copy (blue line) during all embryonic development for Type 2 DMRs or which a gain-of-methylation after blastociste stage (the specific developmental time is unknown) for the Type 1.

day 10 in human development, showing intermediate methylation level at placenta-specific DMRs (~46%); chorion (which is a mix of TE and ICM derived cells), originated before day 15 and presenting low methylation level at these regions (~33%); and finally, low methylation at placenta-specific DMRs (~8%) in amnion, fully derived from ICM cells and also originated before day 15 (Figure 5.7) (Hanna et al., 2016).

Techniques now allow for the complete methylomes of very small samples to be obtained, which has lead to the description of the 5-mC profiles in human sperm and individual oocytes. We analysed publically available dataset to reveal that a large number of regions are differentially methylated between gametes, with ~80% of oocyte-derived methylation present at the blastocyst stage. These regions are including the confirmed placenta-specific DMRs and potentially, other additional gDMRs that could dictate imprinting early in development. The regions we selected for confirmation in preimplantation embryos were chosen because we had already observed allelic methylation in placenta samples. Therefore a situation similar to mouse transient gDMR would not have been identified as these regions would be demethylated after implantation. Evidence for the existence of such an event was observed at the *NPAS3* gene in our embryos. This region was used as a control

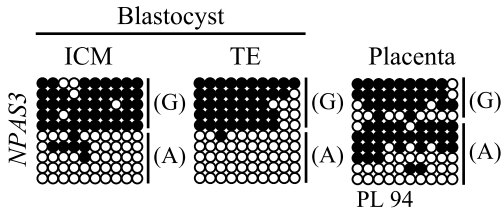



Figure 5.8: **Methylation profile of *NPAS3* region.** *NPAS3* was one of our candidate ubiquitous DMRs but its methylation profile after strand-specific bisulphite PCR in various tissues, including the placenta, revealed inconsistent allelic methylation profile. However, the region showed allelic methylation in both the ICM and TE. Each circle represents a single CpG dinucleotide on a DNA strand, black circles for methylated cytosines and white for unmethylated cytosines.

since it is methylated in oocytes, unmethylated in sperm and has an inconsistent mosaic methylation pattern in placenta and somatic tissues (Figure 5.8). We confirmed that the NPAS3 promoter is allelically methylated in ICM and TE of day 5 blastocysts. This result is consistent with reports suggesting that human blastocysts might contain thousand of transient maternal methylated gDMRs that lose the allelic methylation after implantation (Hamada et al., 2016; Okae et al., 2014).

In mice, ZGA initiated at the first cell division, where the limited number of mouse transient gDMR can orchestrate imprinting expression. However, in human embryos, genome-activation does not occur until day 2 of development, at the 4-cell stage (Braude et al., 1988; Dobson et al., 2004). Recent studies in mouse have shown that imprinted expression can occur in cleavage embryos independent of allelic DNA methylation. Chromatin profiling revealed the presence of H3K27me3 on the maternal allele at 77 genes, including *Xist* (A. Inoue, Jiang, Lu, & Zhang, 2017; A. Inoue, Jiang, Lu, Suzuki, & Zhang, 2017). Such studies are not currently possible in humans, although allelic RNA-seq profiling could identify imprinted genes not associated with methylation at this time.

To fully understand the transcription activity and hence the influence of gDMRs during early human development is difficult, complicated by not only scarcity of material to study (it is impossible to assess developmental dynamics between day 5/6 and week 8 of gestation) and in cell-type specificity, but also by the limitations of the techniques used. In our study, for assessment of allelic expression in blastocysts, we utilised publically deposited RNA-seq data which itself has problems (Z. Xue et al., 2013). Firstly, not all RNAs are annotated in the Refseq database and also small RNAs or those without poly-A tails are excluded. Furthermore, many imprinted transcripts originate from alternative transcription start-sites, and short read next-generation sequencing does not allow for full-length alternative transcripts to be assessed individually, that may mask imprinting if not all isoforms are monoallelically expressed. Finally, a major issue in most studies is the difficulty to avoid cross-contamination between cell-types.

A recent study at single-cell resolution compared the expression and methylation of mature and germinal vesicle oocytes, single sperm cells and preimplantation embryos (Zhu et al., 2018). Allelic profiling was possible as they compare results to the genotypes obtained from the sperm donors. The authors confirmed that there is higher methylation at gene promoter regions on the maternal alleles at all stages of development and that potentially, very few genes with parental-specific methylation events at their promoter regions show allele-specific gene expression. However, it must be stressed that these results were obtained from short read sequencing and did not take into account the influence of oocyte-derived transcripts that are highly abundant before/during ZGA.

## 5.6 Fitting evolutionary theories to placenta-specific imprinting

It is generally accepted that the genesis of ubiquitous gDMRs were closely related to retrotransposon insertions and other repetitive elements, which were recognised and methylated in a different manner in oocyte and sperm. The different conservation of these regions between mammalian species suggests that these DMRs appeared at a different evolutionary time point and are under the action of diverse natural selection pressures. In short, ubiquitous gDMRs have been hypothesised to be associated with the appearance of new genomic context (e.g. derived from the retrotransposon insertion) which are recognised differently by the epigenetic machinery present in the gametes responsible for *de novo* patterns. In parallel, factors that recognise and maintain these methylation differences have to co-evolved (such as ZFP57) ensuring the survival of allelic methylation during embryonic reprogramming (Hore et al., 2007).

It seems that placenta-specific gDMRs have appeared later in evolution, exclusively in the primate lineage. Due to the large number of regions, their survival in the placenta is more likely associated with the adaption of the oocyte, embryo and placenta epigenetic machinery instead of a new genomic context in all these regions. There are two protein families that are candidate factors: the NLRP and ZFP/ZNF families. Both are large multi-protein families with specific members in different mammalian species due to recent and continuous evolution. The NLRP-family members have no direct DNA binding motifs but our observations link them with genomic imprinting, presumably by interacting with other proteins with DNA-binding and methyltransferase activity in a multi-protein complex. Specifically, NLRP7 is an ideal candidate for a role in establishing/maintaining placenta-specific gDMRs since it is primate-specific and is known to interact with the YY1 transcription factor (Mahadevan et al., 2013b). Similarly, The ZFP/ZNFs are a large gene family which contain numerous zinc-finger domains that recognise specific DNA binding motifs and interact with the KAP1-corepressor complex. As has been described earlier,

ZFP57 is a key factor for genomic imprinting in both mouse and humans. Early in 2018, the binding sites of a large number of human ZFP/ZNF proteins was published, some being primate-specific (Coluccio et al., 2018). Motif searching in these placenta-specific DMRs will be needed to determine if ZFP/ZNF or other proteins are preferentially binding to these regions.

However, the fact that some orthologuous of human placenta-specific DMRs are also mouse transient gDMR, with one of them, the *Grp1/Zdbf2* locus, conserved also in the extraembryonic tissue in mice, suggests that maybe, these regions have also a differential *imprint* in the mouse gametes, but this *imprint* is not protected from epigenetic reprogramming in early intrauterine development (Court et al., 2014; Duffié et al., 2014; Hamada et al., 2016; Hanna et al., 2016; H. Kobayashi et al., 2009; Proudhon et al., 2012). In our comparison between different mammalian species, placenta-specific DMRs seems to be exclusive from the primate lineage. However, we will need to study the methylation status of these regions in other mammalian oocytes deeply to elucidate if the placenta-specific DMRs are the result of primate-specific factors acting acquisition and maintenance of these DMRs, or acting exclusively in their early embryonic and placenta-specific maintenance.

Additionally, to fully understand the regulation of these regions in the human placenta, it is necessary to study the genetic content and the regulation of placenta-specific DMRs. If they are regulating gene expression of key genes involved in intrauterine development, humans will have a huge disadvantage by being more sensitive to the effects of recessive mutations. However, since these regions are relatively new in our evolution and due to their large number, maybe they do not require an adaptive explanation. It is possible that they have arisen as a non-adaptive by-product of the slow erasure of maternally derived methylation during preimplantation development, combined the action of new primate-specific factor(s) involved their acquisition or early maintenance. It is also possible that we are describing paternally derived expression associated with genes involved in trophoblast development or metabolic processes that are expendable, being compensated by the expression of other genes with similar function. But exists some evidence suggesting a possible role of these epigenetically regulated regions.

We have shown that the phenotypes of both types of HM, androgenetically derived CHM and RHMs, are associated with aberrant genomic imprinting. However, before our publication, no systematic analyses for imprinting defects had been reported. Since the severe phenotype of human CHM, with no foetus (while the same phenomenon in mice is associated with poor but existent foetal development), we can partially associate this different degree of affectation to the existence of placenta-specific imprinted genes in human. It could be all placenta-specific DMRs and their associated imprinted genes are essential for development, but individually they are expendable. That is, the full effect of these genes will only be observed in combination (i.e. the LOM in HMs), but

individual variation may influence complex growth traits, but would not be lethal. This hypothesis would be compatible with the polymorphic nature of some placenta-specific DMRs in our control placenta cohort. Since the presence of LOM in some of these DMRs is not associated with complicate outcomes such as IUGR and pre-eclampsia, an extensive study of these regions is necessary to elucidate if placenta samples with greater number of placenta-specific DMRs showing LOM are associated with more extreme pregnancy complications.

If we hypothesise that some evolutionary pressures are acting in the maintenance of placenta-specific DMRs and imprinting, one theory that is compatible with the severe phenotype associated with aberrant imprinting in humans is the *Trophoblast defence hypothesis*. Since placenta-specific regions are all regulated by oocyte-derived maternal methylation, ovarian teratoma formation from parthenogenetically activated oocytes would be prevented due to the biallelic nature of the methylation and the lack of paternally expressed genes, many of which are involved in key early developmental processes such as cell-cell interaction and metabolism. Indeed to stop parthenogenetic activation only a few genes with essential developmental roles need to be paternally expressed (i.e. DNMT1), but since all these regions may appears at the same time, we cannot reject this hypothesis as one of the main evolutionary pressures favouring the maintenance of placenta-specific DMRs. Furthermore, the *kinship theory* is also compatible with the existence of placenta-specific DMRs since by regulating the expression of key genes in the placenta, although they are not directly related with growth, placental maternal silencing can restrict the development and cell proliferation of both, the foetus and the placenta.

## 5.7 What is the polymorphic placental imprinting telling us?

As a result of this thesis, we described two different polymorphic events: samples showing lack of both, methylation and imprinting, and other cases showing the presence of DMR but, with biallelic expression. Focussing on the first scenario, one possible explanation of the polymorphic methylation observed at some placenta-specific DMRs is the existence of SNP variants level which can influence the binding of transcription factors/imprinting regulators acting on oocyte methylation acquisition or maintenance. However, a comparison between control and LOM placenta samples for *LIN28B, ABGL2* and *ZC3H12C* DMRs fails to find any SNP genotypes/haplotypes associated with each group (*data not shown*). However, the polymorphic nature of placenta-specific imprinting could be due to completely different reasons.

We performed experiments to elucidate if polymorphic LOM reflects the lack of establish-

ment or a simple failure to maintain methylation during gestation. By comparing CVS sample (from 12-13 weeks gestation) with the term placenta from the same patient, our results strongly suggest that the LOM is already present in the first-trimester placenta and that the methylation profile does not change during pregnancy at these loci. Furthermore, by comparing placenta profiles of dizygotic twins and triplets who developed with separate independent placentas, we could rule out that LOM of specific regions has occurred due to adaption to the maternal uterine environment. This is unlikely since non-identical twins shown different methylation profiles for some placenta-specific DMRs. However, to ascertain if the genes associated with placenta-specific DMRs are involved in the adaption to nutritional differences it would be preferable to study rare identical twins that arose from an early event resulting in separate placenta and amniotic sacs (only ~20% of identical twins). The only way to truly determine if placenta-specific DMRs are polymorphically acquired in the female germline is to perform single oocyte-methylation analysis, an experiment currently being performed in the laboratory.

A factor that would influence the methylation state is the cell-specificity. Our placenta cohort includes all embryonic-derived placenta cell types, but the invasive trophoblast lineages are the directly responsible to the maternal and foetal blood contact (See Figure 1.18 at introduction section), what makes it the cell-type most relevant to our studies. However, trophoblast analysis performed by other groups has shown the presence of LOM in these isolated cells. However, a comparison between whole placenta villi (which includes all embryonic-derived and placenta-cell types) and trophoblasts from the same sample show slightly different methylation state in some of the regions (for example, *AIM1* DMR shows lower methylation in whole placenta villi than trophoblast for the same sample), but other regions present LOM in both paired samples (Hanna et al., 2016). Therefore, we cannot rule out the existence of subtle cell-type methylation for some placenta-specific loci.

The other polymorphic event observed in placenta-specific imprinting affecting only at the expression level is a phenomenon previously described in other imprinted regions. In humans, *IGF2R* and *SLC22A2* are adjacent biallelic expressed genes despite the presence of a DMR (Monk et al., 2006). In the orthologue region in mouse, imprinting requires the non-coding RNA *Airn*, which has not been detected in humans samples, even the rare samples with *IGF2R* imprinting (~5% of the population) (Cheong et al., 2015). Another example is the expression pattern of *GRB10* gene. The human *GRB10* has been previously shown to be maternally expressed only in cytokeratin 7 trophoblast cells, whilst it is expressed from both alleles in the other placenta cell-types and the majority of somatic tissues (Monk et al., 2009). Interestingly, in unsorted placenta biopsies, *GRB10* expression varied between maternal preferential to biallelic.

During this project, I have shown that 21 genes with a placenta-specific DMR in its

promoter are biallelically expressed – in all or some of the samples tested – using isoform-specific RT-PCR assays. Future work should focus on determining if genes associated with these placenta-specific DMRs are imprinted in specific placenta lineages. If this is the case, monoallelic expression in our samples could be masked by non-imprinted cell types. Other possible reason for the equal expression of both alleles could be that the expression arises from a different promoter region. However, my results suggest that additional epigenetic states may be involved in the polymorphic imprinted expression. We analysed two different placenta-specific regions, *LIN28B*, where we observe LOM and LOI in the same sample and *R3HCC1* where we observed both different polymorphic events, samples with LOM and LOI and others with LOI and presence of DMR. For the *LIN28B* region I have shown the presence of permissive histone marks exclusively on the non-methylated and expressed allele (especially for H3K4me2/3) paternal allele in control placentas, but on both alleles in LOM and LOI samples. Additionally, the quantitative analysis of permissive histone marks has shown an increase of these modifications when LOM and LOI are present. On the contrary, *R3HCC1* parent-of-origin expression seemed to be directly associated with the presence of permissive histone marks (biallelically distributed in both polymorphic scenarios) rather than with DNA methylation. However, the quantitative analysis showed a general loss of epigenetic regulation in this region associated to the LOI, since the presence of permissive histone marks were lower in LOI placentas than the control sample.

As a consequence of all these observations, the direct role of cytosine methylation at promoter regions and gene expression remains unclear. All the imprinting studies showing the gradual and initial DNA methylation pattern acquisition in gametes and its robust stability suggest an important role of this epigenetic mark to the allelic discrimination during embryonic development and adult tissues. However, since DNA methylation does act alone in influencing gene expression it is important to take into consideration the acquisition of other epigenetic marks such as histone modification. As a result, if a promoter of an imprinted gene loss any of these epigenetic marks, its allelic expression could be affected and LOI occurs.

## 5.8 Methylation changes at imprinted DMRs are not associated with IUGR

Unlike initial published studies, we report that there are more imprinted domains in the human placenta than in the somatic tissues with some loci remaining to be characterised. Nonetheless, placenta-specific DMRs seems to be less stable than ubiquitous DMRs, showing a higher methylation variation between individuals. Some of these regions are good candidates for the IUGR phenotype since they have been suggested to play a role in development, but we did not observe a higher frequency of low-level polymorphic

methylation at DMRs in IUGR samples versus control placentas in any of the placenta-specific DMRs tested. However, we did not quantify total expression of transcripts associated with these DMRs, so they may still influence pregnancy outcomes and fetal growth.

Interestingly we observe isolated changes in ubiquitous DMRs in our placenta cohort. Both GOM at *MEST* and LOM of *H19* DMRs are associated with growth restriction associated with Silver-Russell Syndrome (Kagami, Nagai, Fukami, Yamazawa, & Ogata, 2007). However, the abnormal methylation at these DMRs was restricted to the placenta, showing normal methylation range in blood samples from the same patient, suggesting that these "epimutations" maybe also be associated with a non-syndromic form of IUGR. Additionally, we also reported cases of LOM associated with higher expression levels compared to samples maintaining maternal methylation, and at least for *H19*, was due to reactivation of the normally silent allele. Additionally, we also observed some ubiquitous DMRs, with LOM and GOM in placentas from both IUGR and non-IUGR patients (i.e. GOM of *MEST* identified by HM450k arrays and LOM of *H19* and *NHP2L1* identified by pyrosequencing), suggesting that this variability is not associated with growth. Nevertheless, GOM in *MEG3* and *MEST* promoters have also been reported among normal individuals in blood and other somatic tissues (Haertle et al., 2017). However, other recent studies have shown a significant association between LOI at ubiquitous imprinted genes and IUGR pregnancies, including a monozygotic twins study, where increased *CDKN1C* expression and decreased *KCNQ1OT1* expression were detected in IUGR monozygotic twins compared with controls (Gou et al., 2017).

Although we cannot demonstrate this observation with statistical confidence due to the low cases of LOM reported, we do observe individual placenta samples from the IUGR group having a higher number of unmethylated imprinted DMRs. For example, the IUGR PL 217 sample has LOM at the *H19* and *RB1* ubiquitous DMRs and several placenta-specific DMR, e.g. for *THAP3, ZNF385D, MOCS, LIN28B, R3HCC1, ZC3H12C and SEPT4*. However, since we do not observe an increased number of LOM for any placenta-specific DMR in IUGR vs controls, this implies that these regions are not associated with complicated pregnancies in our cohort.

## 5.9 Polymorphic imprinting is not altered in ART

Overall the results of this thesis show that it is important to keep studying the possible disturbance of genomic imprinting in the use of ART. In our placenta cohort, there is no association between the use of ART and the methylation state of placenta-specific DMRs nor ubiquitous DMR tested. However, scientific studies are continually being published linking the use of ART with a higher frequency of imprinting disorders – last

cases described in pseudohypoparathyroidism Type 1B (Fernandez et al., 2017) –.

As it was mentioned in the introduction, the increasing incidence of epigenetic variability following ART use could be due to the subfertility problems of patients. For this propose, more imprinting analysis in gametes from subfertile vs fertile patients/donors are needed (L. Tang et al., 2017).

# Chapter 6

# Conclusions

The general conclusions of this PhD dissertation are:

1. Genomic imprinting is likely to play an important role in the HM phenotype since there is a genome-wide absence of maternal methylation associated with placenta-specific and ubiquitous DMRs.
2. There are more imprinted domains in the human placenta than in the somatic tissues.
3. There are two groups of placenta-specific DMRs: Type 1, which are fully methylated in somatic tissue and Type 2, which are fully unmethylated.
4. These placenta-specific DMRs are largely confined to primates.
5. All placenta-specific DMRs inherits the methylated allele from the oocyte.
6. The majority of placenta-specific DMRs are not regulating imprinted gene clusters.
7. The methylation at placenta-specific DMRs appears to be polymorphic between individuals, including non-identical twins.
8. Methylation is stable during gestation and in different anatomical sites, suggesting that this absence of methylation reflects a failure to establish rather than maintain methylation.
9. The polymorphic placenta imprinting is not associated with adverse pregnancy outcomes.
10. Allelic histone profiles are concordant with placenta-specific imprinted gene expression.
11. Methylation at ubiquitous DMRs is more stable in placenta compared to placenta-specific DMRs.

12. There is a generalised low level of 5-hmC in the placenta, which can be enriched on the methylated allele of some DMRs.

# References


Akoury, E., Zhang, L., Ao, A., & Slim, R. (2014). NLRP7 and khdc3l, the two maternal-effect proteins responsible for recurrent hydatidiform moles, co-localize to the oocyte cytoskeleton. *Human Reproduction*, deu291.

Alazami, A. M., Awad, S. M., Coskun, S., Al-Hassan, S., Hijazi, H., Abdulwahab, F. M., ... Alkuraya, F. S. (2015). TLE6 mutation causes the earliest known human embryonic lethality. *Genome Biology*, *16*(1), 240.

Amleh, A., & Dean, J. (2002). Mouse genetics provides insight into folliculogenesis, fertilization and early embryonic development. *Human Reproduction Update*, *8*(5), 395–403.

Apostolidou, S., Abu-Amero, S., O'donoghue, K., Frost, J., Olafsdottir, O., Chavele, K., ... Moore, G. (2007). Elevated placental expression of the imprinted phlda2 gene is associated with low birth weight. *Journal of Molecular Medicine*, *85*(4), 379–387.

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: A flexible and comprehensive bioconductor package for the analysis of infinium dna methylation microarrays. *Bioinformatics*, *30*(10), 1363–1369.

Azzi, S., Blaise, A., Steunou, V., Harbison, M. D., Salem, J., Brioude, F., ... others. (2014). Complex tissue-specific epigenotypes in russell–Silver syndrome associated with 11p15 icr1 hypomethylation. *Human Mutation*, *35*(10), 1211–1220.

Babak, T., DeVeale, B., Armour, C., Raymond, C., Cleary, M. A., Kooy, D. van der, ... Lim, L. P. (2008). Global survey of genomic imprinting by transcriptome sequencing. *Current Biology*, *18*(22), 1735–1741.

Barbaux, S., Gascoin-Lachambre, G., Buffat, C., Monnier, P., Mondon, F., Tonanny, M.-B., ... others. (2012). A genome-wide approach reveals novel imprinted genes expressed in the human placenta. *Epigenetics*, *7*(9), 1079–1090.

Barlow, D. P. (1993). Methylation and imprinting: From host defense to gene regulation?

*Science*, *260*(5106), 309–311.

Barlow, D. P., Stöger, R., Herrmann, B., Saito, K., & Schweifer, N. (1991). The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the tme locus. *Nature*, *349*(6304), 84–87.

Bartolomei, M. S., Zemel, S., & Tilghman, S. M. (1991). Parental imprinting of the mouse h19 gene. *Nature*, *351*(6322), 153–155.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., ... others. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, *125*(2), 315–326.

Bestor, T., Laudano, A., Mattaliano, R., & Ingram, V. (1988). Cloning and sequencing of a cDNA encoding dna methyltransferase of mouse cells: The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of Molecular Biology*, *203*(4), 971–983.

Bian, C., & Yu, X. (2013). PGC7 suppresses tet3 for protecting dna methylation. *Nucleic Acids Research*, *42*(5), 2893–2905.

Bostick, M., Kim, J. K., Estève, P.-O., Clark, A., Pradhan, S., & Jacobsen, S. E. (2007). UHRF1 plays a role in maintaining dna methylation in mammalian cells. *Science*, *317*(5845), 1760–1764.

Bourc'his, D., Xu, G.-L., Lin, C.-S., Bollman, B., & Bestor, T. H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science*, *294*(5551), 2536–2539.

Bourque, D., Avila, L., Penaherrera, M., Von Dadelszen, P., & Robinson, W. (2010). Decreased placental methylation at the h19/igf2 imprinting control region is associated with normotensive intrauterine growth restriction but not preeclampsia. *Placenta*, *31*(3), 197–202.

Braude, P., Bolton, V., & Moore, S. (1988). Human gene expression first occurs between the four-and eight-cell stages of preimplantation development. *Nature*, *332*(6163), 459–461.

Brioude, F., Oliver-Petit, I., Blaise, A., Praz, F., Rossignol, S., Le Jule, M., ... others. (2013). CDKN1C mutation affecting the pcna-binding domain as a cause of familial russell silver syndrome. *Journal of Medical Genetics*, jmedgenet–2013.

Brooks, A., Johnson, M., Steer, P., Pawson, M., & Abdalla, H. (1995). Birth weight: Nature or nurture? *Early Human Development*, *42*(1), 29–35.

Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R., & Willard, H. F. (1991). A gene from the region of the human x inactivation centre is

expressed exclusively from the inactive x chromosome. *Nature*, *349*(6304), 38.

Brown, S. W., & Nur, U. (1964). Heterochromatic chromosomes in the coccids. *Science*, *145*(3628), 130–136.

Bukovsky, A. (2006). Immune system involvement in the regulation of ovarian function and augmentation of cancer. *Microscopy Research and Technique*, *69*(6), 482–500.

Caillaud, M., Duchamp, G., & Gérard, N. (2005). In vivo effect of interleukin-1beta and interleukin-1RA on oocyte cytoplasmic maturation, ovulation, and early embryonic development in the mare. *Reproductive Biology and Endocrinology*, *3*(1), 26.

Caliebe, A., Richter, J., Ammerpohl, O., Kanber, D., Beygo, J., Bens, S., ... others. (2014). A familial disorder of altered dna-methylation. *Journal of Medical Genetics*, *51*(6), 407–412.

Calle, A., Fernandez-Gonzalez, R., Ramos-Ibeas, P., Laguna-Barraza, R., Perez-Cerezales, S., Bermejo-Alvarez, P., ... Gutierrez-Adan, A. (2012). Long-term and transgenerational effects of in vitro culture on mouse embryos. *Theriogenology*, *77*(4), 785–793.

Camprubí, C., Iglesias-Platas, I., Martin-Trujillo, A., Salvador-Alarcon, C., Rodriguez, M. A., Barredo, D. R., ... others. (2013). Stability of genomic imprinting and gestational-age dynamic methylation in complicated pregnancies conceived following assisted reproductive technologies. *Biology of Reproduction*, *89*(3).

Cao, J., & Yan, Q. (2012). Histone ubiquitination and deubiquitination in transcription, dna damage response, and cancer. *Frontiers in Oncology*, *2*.

Capuano, F., Muslleder, M., Kok, R., Blom, H. J., & Ralser, M. (2014). Cytosine dna methylation is found in drosophila melanogaster but absent in saccharomyces cerevisiae, schizosaccharomyces pombe, and other yeast species. *MAnalytical Chemistry*, *86*(8), 3697–3702.

Cattanach, B. M., & Kirk, M. (1985). Differential activity of maternally and paternally derived chromosome regions in mice. *Nature*, *315*(6019), 496–498.

Cattanach, B., & Beechey, C. (1997). Genomic imprinting in the mouse: Possible final analysis. *Genomic Imprinting: Frontiers in Molecular Biology*, *18*, 118–145.

Chen, S., Sun, F.-z., Huang, X., Wang, X., Tang, N., Zhu, B., & Li, B. (2015). Assisted reproduction causes placental maldevelopment and dysfunction linked to reduced fetal weight in mice. *Scientific Reports*, *5*.

Cheong, C. Y., Chng, K., Ng, S., Chew, S. B., Chan, L., & Ferguson-Smith, A. C. (2015). Germline and somatic imprinting in the nonhuman primate highlights species differences

in oocyte methylation. *Genome Research*, *25*(5), 611–623.

Choux, C., Binquet, C., Carmignac, V., Bruno, C., Chapusot, C., Barberet, J., ... Fauque, P. (2017). The epigenetic control of transposable elements and imprinted genes in newborns is affected by the mode of conception: ART versus spontaneous conception without underlying infertility. *Human Reproduction*.

Ciccone, D. N., Su, H., Hevi, S., Gay, F., Lei, H., Bajko, J., ... Chen, T. (2009). KDM1B is a histone h3k4 demethylase required to establish maternal genomic imprints. *Nature*, *461*(7262), 415–418.

Coluccio, A., Ecco, G., Duc, J., Offner, S., Turelli, P., & Trono, D. (2018). Individual retrotransposon integrants are differentially controlled by kzfp/kap1-dependent histone methylation, dna methylation and tet-mediated hydroxymethylation in naïve embryonic stem cells. *Epigenetics & Chromatin*, *11*(1), 7.

Constância, M., Hemberger, M., Hughes, J., Dean, W., Ferguson-Smith, A., Fundele, R., ... others. (2002). Placental-specific igf-ii is a major modulator of placental and fetal growth. *Nature*, *417*(6892), 945.

Cooper, W. N., Luharia, A., Evans, G. A., Raza, H., Haire, A. C., Grundy, R., ... others. (2005). Molecular subtypes and phenotypic expression of beckwith–Wiedemann syndrome. *European Journal of Human Genetics*, *13*(9), 1025–1032.

Court, F., Tayama, C., Romanelli, V., Martin-Trujillo, A., Iglesias-Platas, I., Okamura, K., ... others. (2014). Genome-wide parent-of-origin dna methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. *Genome Research*, *24*(4), 554.

Cox, G. F., Bürger, J., Lip, V., Mau, U. A., Sperling, K., Wu, B.-L., & Horsthemke, B. (2002). Intracytoplasmic sperm injection may increase the risk of imprinting defects. *The American Journal of Human Genetics*, *71*(1), 162–164.

Crouse, H. V. (1960). The controlling element in sex chromosome behavior in sciara. *Genetics*, *45*(10), 1429.

Das, R., Lee, Y. K., Strogantsev, R., Jin, S., Lim, Y. C., Ng, P. Y., ... others. (2013). DNMT1 and aim1 imprinting in human placenta revealed through a genome-wide screen for allele-specific dna methylation. *Bmc Genomics*, *14*(1), 685.

Davies, M. N., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., ... others. (2012). Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biology*, *13*(6), R43.

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes*

*& Development*, *25*(10), 1010–1022.

DeBaun, M. R., Niemitz, E. L., & Feinberg, A. P. (2003). Association of in vitro fertilization with beckwith-wiedemann syndrome and epigenetic alterations of lit1 and h19. *The American Journal of Human Genetics*, *72*(1), 156–160.

DeChiara, T. M., Robertson, E. J., & Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor ii gene. *Cell*, *64*(4), 849–859.

Demars, J., Rossignol, S., Netchine, I., Lee, K. S., Shmela, M., Faivre, L., ... others. (2011). New insights into the pathogenesis of beckwith–Wiedemann and silver–Russell syndromes: Contribution of small copy number variations to 11p15 imprinting defects. *Human Mutation*, *32*(10), 1171–1182.

Diaz-Meyer, N., Day, C., Khatod, K., Maher, E., Cooper, W., Reik, W., ... others. (2003). Silencing of cdkn1c (p57KIP2) is associated with hypomethylation at kvdmr1 in beckwith–Wiedemann syndrome. *Journal of Medical Genetics*, *40*(11), 797–801.

Diplas, A. I., Lambertini, L., Lee, M.-J., Sperling, R., Lee, Y. L., Wetmur, J. G., & Chen, J. (2009). Differential expression of imprinted genes in normal and iugr human placentas. *Epigenetics*, *4*(4), 235–240.

Dixon, P. H., Trongwongsa, P., Abu-Hayyah, S., Ng, S. H., Akbar, S. A., Khawaja, N. P., ... Fisher, R. A. (2012). Mutations in nlrp7 are associated with diploid biparental hydatidiform moles, but not androgenetic complete moles. *Journal of Medical Genetics*, jmedgenet–2011.

Dobson, A. T., Raja, R., Abeyta, M. J., Taylor, T., Shen, S., Haqq, C., & Pera, R. A. R. (2004). The unique transcriptome through day 3 of human preimplantation development. *Human Molecular Genetics*, *13*(14), 1461–1470.

Docherty, L. E., Rezwan, F. I., Poole, R. L., Turner, C. L., Kivuva, E., Maher, E. R., ... others. (2015). Mutations in nlrp5 are associated with reproductive wastage and multilocus imprinting disorders in humans. *Nature Communications*, *6*.

Doherty, A. S., Mann, M. R., Tremblay, K. D., Bartolomei, M. S., & Schultz, R. M. (2000). Differential effects of culture on imprinted h19 expression in the preimplantation mouse embryo. *Biology of Reproduction*, *62*(6), 1526–1535.

Duffié, R., Ajjan, S., Greenberg, M. V., Zamudio, N., Arenal, M. E. del, Iranzo, J., ... Bourc'his, D. (2014). The gpr1/zdbf2 locus provides new paradigms for transient and dynamic genomic imprinting in mammals. *Genes & Development*, *28*(5), 463–478.

Ecco, G., Cassano, M., Kauzlaric, A., Duc, J., Coluccio, A., Offner, S., ... Trono, D. (2016). Transposable elements and their krab-zfp controllers regulate gene expression in

adult tissues. *Developmental Cell*, *36*(6), 611–623.

Eggermann, T., Eggermann, K., & Schönherr, N. (2008). Growth retardation versus overgrowth: Silver-russell syndrome is genetically opposite to beckwith-wiedemann syndrome. *Trends in Genetics*, *24*(4), 195–204.

El Hajj, N., & Haaf, T. (2013). Epigenetic disturbances in in vitro cultured gametes and embryos: Implications for human assisted reproduction. *Fertility and Sterility*, *99*(3), 632–641.

Fedoriw, A., Mugford, J., & Magnuson, T. (2012). Genomic imprinting and epigenetic control of development. *Cold Spring Harbor Perspectives in Biology*, *4*(7), a008136.

Fernandez, M., Zambrano, M. J., Riquelme, J., Castiglioni, C., Kottler, M.-L., Jüppner, H., & Mericq, V. (2017). Pseudohypoparathyroidism type 1B associated with assisted reproductive technology. *Journal of Pediatric Endocrinology and Metabolism*, *30*(10), 1125–1132.

Fountain, M. D., Tao, H., Chen, C.-A., Yin, J., & Schaaf, C. P. (2017). Magel2 knockout mice manifest altered social phenotypes and a deficit in preference for social novelty. *Genes, Brain and Behavior*, *16*(6), 592–600.

Fowden, A., Coan, P., Angiolini, E., Burton, G., & Constancia, M. (2011). Imprinted genes and the epigenetic regulation of placental phenotype. *Progress in Biophysics and Molecular Biology*, *106*(1), 281–288.

Frost, J. M., & Moore, G. E. (2010). The importance of imprinting in the human placenta. *PLoS Genetics*, *6*(7), e1001015.

Gahurova, L., Tomizawa, S.-i., Smallwood, S. A., Stewart-Morgan, K. R., Saadeh, H., Kim, J., ... Kelsey, G. (2017). Transcription and chromatin determinants of de novo dna methylation timing in oocytes. *Epigenetics & Chromatin*, *10*(1), 25.

Gardner, R., Squire, S., Zaina, S., Hills, S., & Graham, C. (1999). Insulin-like growth factor-2 regulation of conceptus composition: Effects of the trophectoderm and inner cell mass genotypes in the mouse. *Biology of Reproduction*, *60*(1), 190–195.

Gicquel, C., Rossignol, S., Cabrol, S., Houang, M., Steunou, V., Barbu, V., ... others. (2005). Epimutation of the telomeric imprinting center region on chromosome 11p15 in silver-russell syndrome. *Nature Genetics*, *37*(9), 1003–1007.

Gkountela, S., Li, Z., Vincent, J. J., Zhang, K. X., Chen, A., Pellegrini, M., & Clark, A. T. (2013). The ontogeny of cKIT+ human primordial germ cells proves to be a resource for human germ line reprogramming, imprint erasure and in vitro differentiation. *Nature Cell Biology*, *15*(1), 113–122.

Glenn, C. C., Porter, K. A., Jong, M. T., Nicholls, R. D., & Driscoll, D. J. (1993).

Functional imprinting and epigenetic modification of the human snrpn gene. *Human Molecular Genetics*, *2*(12), 2001–2005.

Godfrey, K., Robinson, S., Barker, D., Osmond, C., & Cox, V. (1996). Maternal nutrition in early and late pregnancy in relation to placental and fetal growth. *Bmj*, *312*(7028), 410.

Gomes, M., Gomes, C., Pinto, W., & Ramos, E. (2007). Methylation pattern at the kvdmr in a child with beckwith–Wiedemann syndrome conceived by icsi. *American Journal of Medical Genetics Part A*, *143*(6), 625–629.

Gomes, M., Huber, J., Ferriani, R., Amaral Neto, A., & Ramos, E. (2009). Abnormal methylation at the kvdmr1 imprinting control region in clinically normal children conceived by assisted reproductive technologies. *Molecular Human Reproduction*, *15*(8), 471–477.

Gou, C., Li, M., Zhang, X., Liu, X., Huang, X., Zhou, Y., & Fang, Q. (2017). Placental characteristics in monochorionic twins with selective intrauterine growth restriction assessed by gradient angiography and three-dimensional reconstruction. *The Journal of Maternal-Fetal & Neonatal Medicine*, *30*(21), 2590–2595.

Grant, J., Mahadevaiah, S. K., Khil, P., Sangrithi, M. N., Royo, H., Duckworth, J., ... others. (2012). Rsx is a metatherian rna with xist-like properties in x-chromosome inactivation. *Nature*, *487*(7406), 254.

Greer, E. L., Blanco, M. A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., ... Shi, Y. (2015). DNA methylation on n 6-adenine in c. elegans. *Cell*, *161*(4), 868–878.

Grützner, F., Rens, W., Tsend-Ayush, E., El-Mogharbel, N., O'brien, P. C., Jones, R. C., ... Graves, J. A. M. (2004). In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird z and mammal x chromosomes. *Nature*, *432*(7019), 913–917.

Guernsey, M. W., Chuong, E. B., Cornelis, G., Renfree, M. B., & Baker, J. C. (2017). Molecular conservation of marsupial and eutherian placentation and lactation. *eLife*, *6*.

Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., ... others. (2014). The dna methylation landscape of human early embryos. *Nature*, *511*(7511), 606.

Guo, W., Zhang, M. Q., & Wu, H. (2016). Mammalian non-cg methylations are conserved and cell-type specific and may have been involved in the evolution of transposon elements. *Scientific Reports*, *6*, 32207.

Hackett, J. A., Sengupta, R., Zylicz, J. J., Murakami, K., Lee, C., Down, T. A., & Surani, M. A. (2013). Germline dna demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science*, *339*(6118), 448–452.

Hackett, J. A., Zylicz, J. J., & Surani, M. A. (2012). Parallel mechanisms of epigenetic

reprogramming in the germline. *Trends in Genetics*, *28*(4), 164–174.

Haertle, L., Maierhofer, A., Böck, J., Lehnen, H., Böttcher, Y., Blüher, M., ... others. (2017). Hypermethylation of the non-imprinted maternal meg3 and paternal mest alleles is highly variable among normal individuals. *PloS One*, *12*(8), e0184030.

Haig, D. (2004). Genomic imprinting and kinship: How good is the evidence? *Annu. Rev. Genet.*, *38*, 553–585.

Haig, D., & Westoby, M. (1989). Parent-specific gene expression and the triploid endosperm. *The American Naturalist*, *134*(1), 147–155.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., ... Surani, M. A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of Development*, *117*(1), 15–23.

Hamada, H., Okae, H., Toh, H., Chiba, H., Hiura, H., Shirane, K., ... others. (2016). Allele-specific methylome and transcriptome analysis reveals widespread imprinting in the human placenta. *The American Journal of Human Genetics*, *99*(5), 1045–1058.

Hanna, C. W., Peñaherrera, M. S., Saadeh, H., Andrews, S., McFadden, D. E., Kelsey, G., & Robinson, W. P. (2016). Pervasive polymorphic imprinted methylation in the human placenta. *Genome Research*, *26*(6), 756–767.

Hansen, M., & Bower, C. (2014). The impact of assisted reproductive technologies on intra-uterine growth and birth defects in singletons. In *Seminars in fetal and neonatal medicine* (Vol. 19, pp. 228–233). Elsevier.

Hashimoto, H., Liu, Y., Upadhyay, A. K., Chang, Y., Howerton, S. B., Vertino, P. M., ... Cheng, X. (2012). Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Research*, *40*(11), 4841–4849.

Hata, K., Okano, M., Lei, H., & Li, E. (2002). Dnmt3L cooperates with the dnmt3 family of de novo dna methyltransferases to establish maternal imprints in mice. *Development*, *129*(8), 1983–1993.

Hayward, B. E., De Vos, M., Talati, N., Abdollahi, M. R., Taylor, G. R., Meyer, E., ... others. (2009). Genetic and epigenetic analysis of recurrent hydatidiform mole. *Human Mutation*, *30*(5).

He, Y., & Ecker, J. R. (2015). Non-cg methylation in the human genome. *Annual Review of Genomics and Human Genetics*, *16*, 55–77.

Heijmans, B. T., Tobi, E. W., Stein, A. D., Putter, H., Blauw, G. J., Susser, E. S., ... Lumey, L. (2008). Persistent epigenetic differences associated with prenatal exposure to

famine in humans. *Proceedings of the National Academy of Sciences*, *105*(44), 17046–17049.

Heithoff, D. M., Sinsheimer, R. L., Low, D. A., & Mahan, M. J. (1999). An essential role for dna adenine methylation in bacterial virulence. *Science*, *284*(5416), 967–970.

Henckel, A., Chebli, K., Kota, S. K., Arnaud, P., & Feil, R. (2012). Transcription and histone methylation changes correlate with imprint acquisition in male germ cells. *The EMBO Journal*, *31*(3), 606–615.

Hernandez-Mora, J. R., Sanchez-Delgado, M., Petazzi, P., Moran, S., Esteller, M., Iglesias-Platas, I., & Monk, D. (2017). Profiling of oxBS-450K 5-hydroxymethylcytosine in human placenta and brain reveals enrichment at imprinted loci. *Epigenetics*, (just-accepted), 00–00.

Herse, F., LaMarca, B., Hubel, C., Laivuori, H., Huppertz, B., Verlohren, S., ... Dechend, R. (2012). OS066. intrauterine cyp2j2 expression and circulatingepoxyeicosatrienoic acid levels in preeclampsia. *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health*, *2*(3), 212–213.

Hiura, H., Sugawara, A., Ogawa, H., John, R. M., Miyauchi, N., Miyanari, Y., ... others. (2010). A tripartite paternally methylated region within the gpr1-zdbf2 imprinted domain on mouse chromosome 1 identified by meDIP-on-chip. *Nucleic Acids Research*, *38*(15), 4929–4945.

Hore, T. A., Rapkins, R. W., & Graves, J. A. M. (2007). Construction and evolution of imprinted loci in mammals. *TRENDS in Genetics*, *23*(9), 440–448.

Hurst, L. D. (1997). 12 l evolutionary theories of genomic imprinting. *Genomic Imprinting*, 211.

Hutter, B., Helms, V., & Paulsen, M. (2006). Tandem repeats in the cpg islands of imprinted genes. *Genomics*, *88*(3), 323–332.

Iglesias-Platas, I., & Monk, D. (2016). Nongenomic regulation of gene expression. *Current Opinion in Pediatrics*, *28*(4), 521–528.

Imbeault, M., Helleboid, P.-Y., & Trono, D. (2017). KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature*, *543*(7646), 550.

Inoue, A., Jiang, L., Lu, F., & Zhang, Y. (2017). Genomic imprinting of xist by maternal h3k27me3. *Genes & Development*, *31*(19), 1927–1932.

Inoue, A., Jiang, L., Lu, F., Suzuki, T., & Zhang, Y. (2017). Maternal h3k27me3 controls dna methylation-independent imprinting. *Nature*, *547*(7664), 419.

Inoue, A., Shen, L., Dai, Q., He, C., & Zhang, Y. (2011). Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. *Cell*

*Research*, *21*(12), 1670–1676.

Iqbal, K., Jin, S.-G., Pfeifer, G. P., & Szabó, P. E. (2011). Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proceedings of the National Academy of Sciences*, *108*(9), 3642–3647.

Ivanova, E., Chen, J.-H., Segonds-Pichon, A., Ozanne, S. E., & Kelsey, G. (2012). DNA methylation at differentially methylated regions of imprinted genes is resistant to developmental programming by maternal nutrition. *Epigenetics*, *7*(10), 1200–1210.

Jimenez-Chillaron, J. C., Ramon-Krauel, M., Ribo, S., & Diaz, R. (2016). Transgenerational epigenetic inheritance of diabetes risk as a consequence of early nutritional imbalances. *Proceedings of the Nutrition Society*, *75*(1), 78–89.

John, R. M., & Lefebvre, L. (2011). Developmental regulation of somatic imprints. *Differentiation*, *81*(5), 270–280.

Johnson, D. (1974). Hairpin-tail: A case of post-reductional gene action in the mouse egg? *Genetics*, *76*(4), 795–805.

Judson, H., Hayward, B. E., Sheridan, E., & Bonthron, D. T. (2002). A global disorder of imprinting in the human female germ line. *Nature*, *416*(6880), 539–542.

Jürgens, A. S., Kolanczyk, M., Moebest, D. C., Zemojtel, T., Lichtenauer, U., Duchniewicz, M., ... others. (2009). PBX1 is dispensable for neural commitment of ra-treated murine es cells. *In Vitro Cellular & Developmental Biology-Animal*, *45*(5-6), 252–263.

Kagami, M., Nagai, T., Fukami, M., Yamazawa, K., & Ogata, T. (2007). Silver-russell syndrome in a girl born after in vitro fertilization: Partial hypermethylation at the differentially methylated region of peg1/mest. *Journal of Assisted Reproduction and Genetics*, *24*(4), 131–136.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M., & Saitou, M. (2013). Replication-coupled passive dna demethylation for the erasure of genome imprints in mice. *The EMBO Journal*, *32*(3), 340–353.

Kappil, M., Lambertini, L., & Chen, J. (2015). Environmental influences on genomic imprinting. *Current Environmental Health Reports*, *2*(2), 155–162.

Kelsey, G. (2011). Epigenetics and the brain: Transcriptome sequencing reveals new depths to genomic imprinting. *Bioessays*, *33*(5), 362–367.

Keown, C. L., Berletch, J. B., Castanon, R., Nery, J. R., Disteche, C. M., Ecker, J. R., & Mukamel, E. A. (2017). Allele-specific non-cg dna methylation marks domains of active chromatin in female mouse brain. *Proceedings of the National Academy of Sciences*,

*114*(14), E2882–E2890.

Kermicle, J. (1970). Dependence of the r-mottled aleurone phenotype in maize on mode of sexual transmission. *Genetics*, *66*(1), 69.

Killian, J. K., Nolan, C. M., Stewart, N., Munday, B. L., Andersen, N. A., Nicol, S., & Jirtle, R. L. (2001). Monotreme igf2 expression and ancestral origin of genomic imprinting. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, *291*(2), 205–212.

Kim, B., Kan, R., Anguish, L., Nelson, L. M., & Coonrod, S. A. (2010). Potential role for mater in cytoplasmic lattice formation in murine oocytes. *PloS One*, *5*(9), e12587.

Kim, K.-H., & Lee, K.-A. (2014). Maternal effect genes: Findings and effects on mouse embryo development. *Clinical and Experimental Reproductive Medicine*, *41*(2), 47–61.

Kingsley, S. L., Deyssenroth, M. A., Kelsey, K. T., Awad, Y. A., Kloog, I., Schwartz, J. D., ... Wellenius, G. A. (2017). Maternal residential air pollution and placental imprinted gene expression. *Environment International*, *108*, 204–211.

Knoll, J., Nicholls, R., Magenis, R., Graham, J., Lalande, M., Latt, S., ... Reynolds, J. F. (1989). Angelman and prader-willi syndromes share a common chromosome 15 deletion but differ in parental origin of the deletion. *American Journal of Medical Genetics Part A*, *32*(2), 285–290.

Kobayashi, H., Yamada, K., Morita, S., Hiura, H., Fukuda, A., Kagami, M., ... Kono, T. (2009). Identification of the mouse paternally expressed imprinted gene zdbf2 on chromosome 1 and its imprinted human homolog zdbf2 on chromosome 2. *Genomics*, *93*(5), 461–472.

Kobayashi, H., Yanagisawa, E., Sakashita, A., Sugawara, N., Kumakura, S., Ogawa, H., ... Kono, T. (2013). Epigenetic and transcriptional features of the novel human imprinted lncRNA gpr1as suggest it is a functional ortholog to mouse zdbf2linc. *Epigenetics*, *8*(6), 635–645.

Kohli, R. M., & Zhang, Y. (2013). TET enzymes, tdg and the dynamics of dna demethylation. *Nature*, *502*(7472), 472–479.

Kornberg, R. D. (1974). Chromatin structure: A repeating unit of histones and dna. *Science*, *184*(4139), 868–871.

Kulkarni, A., Chavan-Gautam, P., Mehendale, S., Yadav, H., & Joshi, S. (2011). Global dna methylation patterns in placenta and its association with maternal hypertension in pre-eclampsia. *DNA and Cell Biology*, *30*(2), 79–84.

Kuwayama, M. (2007). Highly efficient vitrification for cryopreservation of human oocytes

and embryos: The cryotop method. *Theriogenology*, *67*(1), 73–80.

Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J., & Reik, W. (2003). Resistance of iaps to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis*, *35*(2), 88–93.

Latz, E., Xiao, T. S., & Stutz, A. (2013). Activation and regulation of the inflammasomes. *Nature Reviews Immunology*, *13*(6), 397–411.

Lau, M., Stewart, C., Liu, Z., Bhatt, H., Rotwein, P., & Stewart, C. L. (1994). Loss of the imprinted igf2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes & Development*, *8*(24), 2953–2963.

Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying dna methylation patterns in plants and animals. *Nature Reviews Genetics*, *11*(3), 204–220.

Lawrence, M., Daujat, S., & Schneider, R. (2016). Lateral thinking: How histone modifications regulate gene expression. *Trends in Genetics*, *32*(1), 42–56.

Lechtig, A., Yarbrough, C., Delgado, H., Martorell, R., Klein, R. E., & Béhar, M. (1975). Effect of moderate maternal malnutrition on the placenta. *American Journal of Obstetrics and Gynecology*, *123*(2), 191–201.

Lee, J. T. (2000). Disruption of imprinted x inactivation by parent-of-origin effects at tsix. *Cell*, *103*(1), 17–27.

Lee, J., Davidow, L. S., & Warshawsky, D. (1999). Tsix, a gene antisense to xist at the x-inactivation centre. *Nature Genetics*, *21*(4), 400–404.

Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A., & Tilghman, S. M. (1995). Disruption of imprinting caused by deletion of the h19 gene region in mice. *Nature*, *375*(6526), 34.

Lewis, A., Mitsuya, K., Umlauf, D., Smith, P., Dean, W., Walter, J., ... Reik, W. (2004). Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of dna methylation. *Nature Genetics*, *36*(12), 1291–1295.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The sequence alignment/map format and samtools. *Bioinformatics*, *25*(16), 2078–2079.

Li, L., Baibakov, B., & Dean, J. (2008). A subcortical maternal complex essential for preimplantation mouse embryogenesis. *Developmental Cell*, *15*(3), 416–425.

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., & Ferguson-Smith, A. C. (2008). A maternal-zygotic effect gene, zfp57, maintains both maternal and paternal

imprints. *Developmental Cell*, *15*(4), 547–557.

Lin, S.-P., Youngson, N., Takada, S., Seitz, H., Reik, W., Paulsen, M., ... Ferguson-Smith, A. C. (2003). Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the dlk1-gtl2 imprinted cluster on mouse chromosome 12. *Nature Genetics*, *35*(1), 97.

Linder, D., McCaw, B. K., & Hecht, F. (1975). Parthenogenic origin of benign ovarian teratomas. *New England Journal of Medicine*, *292*(2), 63–66.

Liu, D., Wang, Y., Yang, X.-A., & Liu, D. (2017). De novo mutation of paternal igf2 gene causing silver–Russell syndrome in a sporadic patient. *Frontiers in Genetics*, *8*, 105.

Liu, X., Gao, Q., Li, P., Zhao, Q., Zhang, J., Li, J., ... Wong, J. (2013). UHRF1 targets dnmt1 for dna methylation through cooperative binding of hemi-methylated dna and methylated h3k9. *Nature Communications*, *4*, 1563.

López-Abad, M., Iglesias-Platas, I., & Monk, D. (2016). Epigenetic characterization of cdkn1c in placenta samples from non-syndromic intrauterine growth restriction. *Frontiers in Genetics*, *7*.

Lucifero, D., Mann, M. R., Bartolomei, M. S., & Trasler, J. M. (2004). Gene-specific timing and epigenetic memory in oocyte imprinting. *Human Molecular Genetics*, *13*(8), 839–849.

Lui, J. C., Finkielstain, G. P., Barnes, K. M., & Baron, J. (2008). An imprinted gene network that controls mammalian somatic growth is down-regulated during postnatal growth deceleration in multiple organs. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *295*(1), R189–R196.

Lunnon, K., Hannon, E., Smith, R. G., Dempster, E., Wong, C., Burrage, J., ... others. (2016). Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. *Genome Biology*, *17*(1), 27.

Luo, C., Keown, C. L., Kurihara, L., Zhou, J., He, Y., Li, J., ... others. (2017). Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science*, *357*(6351), 600–604.

Lyon, M. F. (1961). Gene action in the x-chromosome of the mouse (mus musculus l.). *Nature*, *190*(4773), 372–373.

Lyon, M. F. (1962). Sex chromatin and gene action in the mammalian x-chromosome. *American Journal of Human Genetics*, *14*(2), 135.

Lyon, M. F. (1972). X-chromosome inactivation and developmental patterns in mammals.

*Biological Reviews*, *47*(1), 1–35.

Lyon, M. F. (1998). X-chromosome inactivation: A repeat hypothesis. *Cytogenetic and Genome Research*, *80*(1-4), 133–137.

Mackay, D. J., & Temple, I. K. (2017). Human imprinting disorders: Principles, practice, problems and progress. *European Journal of Medical Genetics*, *60*(11), 618–626.

Mackay, D. J., Callaway, J. L., Marks, S. M., White, H. E., Acerini, C. L., Boonen, S. E., ... others. (2008). Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in zfp57. *Nature Genetics*, *40*(8), 949–951.

Mackay, D., Boonen, S., Clayton-Smith, J., Goodship, J., Hahnemann, J., Kant, S., ... others. (2006). A maternal hypomethylation syndrome presenting as transient neonatal diabetes mellitus. *Human Genetics*, *120*(2), 262–269.

Maenohara, S., Unoki, M., Toh, H., Ohishi, H., Sharif, J., Koseki, H., & Sasaki, H. (2017). Role of uhrf1 in de novo dna methylation in oocytes and maintenance methylation in preimplantation embryos. *PLoS Genetics*, *13*(10), e1007042.

Mahadevan, S., Sathappan, V., Utama, B., Lorenzo, I., Kaskar, K., & Van den Veyver, I. B. (2017). Maternally expressed nlrp2 links the subcortical maternal complex (scmc) to fertility, embryogenesis and epigenetic reprogramming. *Scientific Reports*, *7*.

Mahadevan, S., Wen, S., Balasa, A., Fruhman, G., Mateus, J., Wagner, A., ... Van den Veyver, I. B. (2013a). No evidence for mutations in nlrp7 and khdc3l in women with androgenetic hydatidiform moles. *Prenatal Diagnosis*, *33*(13), 1242–1247.

Mahadevan, S., Wen, S., Wan, Y.-W., Peng, H.-H., Otta, S., Liu, Z., ... others. (2013b). NLRP7 affects trophoblast lineage differentiation, binds to overexpressed yy1 and alters cpg methylation. *Human Molecular Genetics*, *23*(3), 706–716.

Manipalviratn, S., DeCherney, A., & Segars, J. (2009). Imprinting disorders and assisted reproductive technology. *Fertility and Sterility*, *91*(2), 305–315.

Mann, M. R., Lee, S. S., Doherty, A. S., Verona, R. I., Nolen, L. D., Schultz, R. M., & Bartolomei, M. S. (2004). Selective loss of imprinting in the placenta following preimplantation development in culture. *Development*, *131*(15), 3727–3735.

Manning, M., Lissens, W., Bonduelle, M., Camus, M., De Rijcke, M., Liebaers, I., & Van Steirteghem, A. (2000). Study of dna-methylation patterns at chromosome 15q11-q13 in children born after icsi reveals no imprinting defects. *Molecular Human Reproduction*, *6*(11), 1049–1053.

McDaniel, P., & Wu, X. (2009). Identification of oocyte-selective nlrp genes in rhesus macaque monkeys (macaca mulatta). *Molecular Reproduction and Development*, *76*(2),

151–159.

McGrath, J., & Solter, D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell*, *37*(1), 179–183.

McMinn, J., Wei, M., Schupf, N., Cusmai, J., Johnson, E., Smith, A., ... Tycko, B. (2006). Unbalanced placental expression of imprinted genes in human intrauterine growth restriction. *Placenta*, *27*(6), 540–549.

Messaed, C., Akoury, E., Djuric, U., Zeng, J., Saleh, M., Gilbert, L., ... Slim, R. (2011). NLRP7, a nucleotide oligomerization domain-like receptor protein, is required for normal cytokine secretion and co-localizes with golgi and the microtubule-organizing center. *Journal of Biological Chemistry*, *286*(50), 43313–43323.

Messerschmidt, D. M., Knowles, B. B., & Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes & Development*, *28*(8), 812–828.

Metsalu, T., Viltrop, T., Tiirats, A., Rajashekar, B., Reimann, E., Kõks, S., ... others. (2014). Using rna sequencing for identifying gene imprinting and random monoallelic expression in human placenta. *Epigenetics*, *9*(10), 1397–1409.

Meyer, E., Lim, D., Pasha, S., Tee, L. J., Rahman, F., Yates, J. R., ... Maher, E. R. (2009). Germline mutation in nlrp2 (nalp2) in a familial imprinting disorder (beckwith-wiedemann syndrome). *PLoS Genetics*, *5*(3), e1000423.

Migeon, B. R. (2016). An overview of x inactivation based on species differences. In *Seminars in cell & developmental biology* (Vol. 56, pp. 111–116). Elsevier.

Mondo, S. J., Dannebaum, R. O., Kuo, R. C., Louie, K. B., Bewick, A. J., LaButti, K., ... others. (2017). Widespread adenine n6-methylation of active genes in fungi. *Nature Genetics*, *49*(6), 964.

Monk, D., Arnaud, P., Apostolidou, S., Hills, F., Kelsey, G., Stanier, P., ... Moore, G. (2006). Limited evolutionary conservation of imprinting in the human placenta. *Proceedings of the National Academy of Sciences*, *103*(17), 6623–6628.

Monk, D., Arnaud, P., Frost, J. M., Wood, A. J., Cowley, M., Martin-Trujillo, A., ... others. (2011). Human imprinted retrogenes exhibit non-canonical imprint chromatin signatures and reside in non-imprinted host genes. *Nucleic Acids Research*, *39*(11), 4577–4586.

Monk, D., Arnaud, P., Frost, J., Hills, F. A., Stanier, P., Feil, R., & Moore, G. E. (2009). Reciprocal imprinting of human grb10 in placental trophoblast and brain: Evolutionary conservation of reversed allelic expression. *Human Molecular Genetics*, *18*(16), 3066–3074.

Monk, D., Sanchez-Delgado, M., & Fisher, R. (2017). NLRPs, the subcortical maternal

complex and genomic imprinting. *Reproduction*, *154*(6), R161–R170.

Moore, G. E., Ali, Z., Khan, R. U., Blunt, S., Bennett, P. R., & Vaughan, J. I. (1997). The incidence of uniparental disomy associated with intrauterine growth retardation in a cohort of thirty-five severely affected babies. *American Journal of Obstetrics & Gynecology*, *176*(2), 294–299.

Moore, T., & Haig, D. (1991). Genomic imprinting in mammalian development: A parental tug-of-war. *Trends in Genetics*, *7*(2), 45–49.

Murdoch, S., Djuric, U., Mazhar, B., Seoud, M., Khan, R., Kuick, R., ... others. (2006). Mutations in nalp7 cause recurrent hydatidiform moles and reproductive wastage in humans. *Nature Genetics*, *38*(3), 300–302.

Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., ... others. (2006). PGC7/stella protects against dna demethylation in early embryogenesis. *Nature Cell Biology*, *9*(1), ncb1519.

Nakamura, T., Liu, Y.-J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S., ... Nakano, T. (2012). PGC7 binds histone h3k9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature*, *486*(7403), 415–419.

Nelissen, E. C., Dumoulin, J. C., Daunay, A., Evers, J. L., Tost, J., & Montfoort, A. P. van. (2013). Placentas from pregnancies conceived by ivf/icsi have a reduced dna methylation level at the h19 and mest differentially methylated regions. *Human Reproduction*, *28*(4), 1117–1126.

Nguyen, N. M. P., & Slim, R. (2014). Genetics and epigenetics of recurrent hydatidiform moles: Basic science and genetic counselling. *Current Obstetrics and Gynecology Reports*, *3*(1), 55–64.

Nicholls, R. D., Knoll, J. H., Butler, M. G., Karam, S., & Lalande, M. (1989). Genetic imprinting suggested by maternal heterodisomy in non-deletion prader-willi syndrome. *Nature*, *342*(6247), 281–285.

Noguer-Dance, M., Abu-Amero, S., Al-Khtib, M., Lefèvre, A., Coullin, P., Moore, G. E., & Cavaillé, J. (2010). The primate-specific microRNA gene cluster (c19mc) is imprinted in the placenta. *Human Molecular Genetics*, *19*(18), 3566–3582.

Okae, H., Chiba, H., Hiura, H., Hamada, H., Sato, A., Utsunomiya, T., ... others. (2014). Genome-wide analysis of dna methylation dynamics during early human development. *PLoS Genetics*, *10*(12), e1004868.

Okae, H., Toh, H., Sato, T., Hiura, H., Takahashi, S., Shirane, K., ... Arima, T. (2018). Derivation of human trophoblast stem cells. *Cell Stem Cell*, *22*(1), 50–63.

Okano, M., Xie, S., & Li, E. (1998). Cloning and characterization of a family of novel

mammalian dna (cytosine-5) methyltransferases. *Nature Genetics*, *19*(3), 219–220.

Olins, A. L., & Olins, D. E. (1974). Spheroid chromatin units ($\nu$ bodies). *Science*, *183*(4122), 330–332.

Ono, R., Kobayashi, S., Wagatsuma, H., Aisaka, K., Kohda, T., Kaneko-Ishino, T., & Ishino, F. (2001). A retrotransposon-derived gene, peg10, is a novel imprinted gene located on human chromosome 7q21. *Genomics*, *73*(2), 232–237.

Ono, R., Nakamura, K., Inoue, K., Naruse, M., Usami, T., Wakisaka-Saito, N., ... others. (2006). Deletion of peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nature Genetics*, *38*(1), 101–106.

Ooi, S. K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., ... others. (2007). DNMT3L connects unmethylated lysine 4 of histone h3 to de novo methylation of dna. *Nature*, *448*(7154), 714–717.

Orr, H. A. (1995). Somatic mutation favors the evolution of diploidy. *Genetics*, *139*(3), 1441–1447.

Otto, S. P., & Gerstein, A. C. (2008). The evolution of haploidy and diploidy. *Current Biology*, *18*(24), R1121–R1124.

Parry, D. A., Logan, C. V., Hayward, B. E., Shires, M., Landolsi, H., Diggle, C., ... others. (2011). Mutations causing familial biparental hydatidiform mole implicate c6orf221 as a possible regulator of genomic imprinting in the human oocyte. *The American Journal of Human Genetics*, *89*(3), 451–458.

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., & Brockdorff, N. (1996). Requirement for xist in x chromosome inactivation. *Nature*, *379*(6561), 131–137.

Pozharny, Y., Lambertini, L., Ma, Y., Ferrara, L., Litton, C. G., Diplas, A., ... others. (2010). Genomic loss of imprinting in first-trimester human placenta. *American Journal of Obstetrics and Gynecology*, *202*(4), 391–e1.

Proffitt, J. H., Davie, J. R., Swinton, D., & Hattman, S. (1984). 5-methylcytosine is not detectable in saccharomyces cerevisiae dna. *Molecular and Cellular Biology*, *4*(5), 985–988.

Proudhon, C., Duffié, R., Ajjan, S., Cowley, M., Iranzo, J., Carbajosa, G., ... others. (2012). Protection against de novo methylation is instrumental in maintaining parent-of-origin methylation inherited from the gametes. *Molecular Cell*, *47*(6), 909–920.

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., ... others. (2011). In embryonic stem cells, zfp57/kap1 recognize a methylated hexanucleotide to affect chromatin and dna methylation of imprinting control regions. *Molecular Cell*,

*44*(3), 361–372.

Rand, E., Ben-Porath, I., Keshet, I., & Cedar, H. (2004). CTCF elements direct allele-specific undermethylation at the imprinted h19 locus. *Current Biology*, *14*(11), 1007–1012.

Regha, K., Sloane, M. A., Huang, R., Pauler, F. M., Warczok, K. E., Melikant, B., ... others. (2007). Active and repressive chromatin are interspersed without spreading in an imprinted gene cluster in the mammalian genome. *Molecular Cell*, *27*(3), 353–366.

Reik, W., & Lewis, A. (2005). Co-evolution of x-chromosome inactivation and imprinting in mammals. *Nature Reviews Genetics*, *6*(5), 403–410.

Richardson, B., Czuppon, A., & Sharman, G. (1971). Inheritance of glucose-6-phosphate dehydrogenase variation in kangaroos. *Nature*, *230*(13), 154–155.

Rothbart, S. B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A. I., ... others. (2012). Association of uhrf1 with methylated h3k9 directs the maintenance of dna methylation. *Nature Structural & Molecular Biology*, *19*(11), 1155–1160.

Sadikovic, B., Fernandes, P., Zhang, V. W., Ward, P. A., Miloslavskaya, I., Rhead, W., ... Fang, P. (2014). Mutation update for ube3a variants in angelman syndrome. *Human Mutation*, *35*(12), 1407–1417.

Sado, T., Fenner, M. H., Tan, S.-S., Tam, P., Shioda, T., & Li, E. (2000). X inactivation in the mouse embryo deficient for dnmt1: Distinct effect of hypomethylation on imprinted and random x inactivation. *Developmental Biology*, *225*(2), 294–303.

Sanchez-Delgado, M., Martin-Trujillo, A., Tayama, C., Vidal, E., Esteller, M., Iglesias-Platas, I., ... others. (2015). Absence of maternal methylation in biparental hydatidiform moles from women with nlrp7 maternal-effect mutations reveals widespread placenta-specific imprinting. *PLoS Genetics*, *11*(11), e1005644.

Sanchez-Delgado, M., Martin-Trujillo, I.-P., & Monk, D. (2015). Causas y consecuencias de los defectos de metilación en múltiples loci en trastornos associados a la impronta genómica. In G. P. de Nanclares & P. Lapunzina (Eds.) (1st ed., pp. 223–258).

Sanchez-Delgado, M., Riccio, A., Eggermann, T., Maher, E. R., Lapunzina, P., Mackay, D., & Monk, D. (2016a). Causes and consequences of multi-locus imprinting disturbances in humans. *Trends in Genetics*, *32*(7), 444–455.

Sanchez-Delgado, M., Vidal, E., Medrano, J., Monteagudo-Sánchez, A., Martin-Trujillo, A., Tayama, C., ... others. (2016b). Human oocyte-derived methylation differences persist in the placenta revealing widespread transient imprinting. *PLoS Genetics*, *12*(11), e1006427.

Sandovici, I., Hoelle, K., Angiolini, E., & Constância, M. (2012). Placental adaptations to the maternal–fetal environment: Implications for fetal growth and developmental

programming. *Reproductive Biomedicine Online*, *25*(1), 68–89.

Sanz, L. A., Chamberlain, S., Sabourin, J.-C., Henckel, A., Magnuson, T., Hugnot, J.-P., ... Arnaud, P. (2008). A mono-allelic bivalent chromatin domain controls tissue-specific imprinting at grb10. *The EMBO Journal*, *27*(19), 2523–2532.

Schier, A. F. (2007). The maternal-zygotic transition: Death and birth of rnas. *Science*, *316*(5823), 406–407.

Schieve, L. A., Meikle, S. F., Ferre, C., Peterson, H. B., Jeng, G., & Wilcox, L. S. (2002). Low and very low birth weight in infants conceived with use of assisted reproductive technology. *N Engl J Med*, *2002*(346), 731–737.

Schoenherr, C. J., Levorse, J. M., & Tilghman, S. M. (2003). CTCF maintains differential methylation at the igf2/h19 locus. *Nature Genetics*, *33*(1), 66–69.

Schroeder, D. I., Blair, J. D., Lott, P., Yu, H. O. K., Hong, D., Crary, F., ... others. (2013). The human placenta methylome. *Proceedings of the National Academy of Sciences*, *110*(15), 6037–6042.

Schroeder, D. I., Jayashankar, K., Douglas, K. C., Thirkill, T. L., York, D., Dickinson, P. J., ... others. (2015). Early developmental and evolutionary origins of gene body dna methylation patterns in mammalian placentas. *PLoS Genetics*, *11*(8), e1005442.

Seisenberger, S., Peat, J. R., Hore, T. A., Santos, F., Dean, W., & Reik, W. (2013). Reprogramming dna methylation in the mammalian life cycle: Building and breaking epigenetic barriers. *Phil. Trans. R. Soc. B*, *368*(1609), 20110330.

Sharif, J., Muto, M., Takebayashi, S.-i., Suetake, I., Iwamatsu, A., Endo, T. A., ... others. (2007). The sra protein np95 mediates epigenetic inheritance by recruiting dnmt1 to methylated dna. *Nature*, *450*(7171), 908–912.

Sharman, G. (1971). Late dna replication in the paternally derived x chromosome of female kangaroos. *Nature*, *230*(5291), 231.

Sharp, A. J., Stathaki, E., Migliavacca, E., Brahmachary, M., Montgomery, S. B., Dupre, Y., & Antonarakis, S. E. (2011). DNA methylation profiles of human active and inactive x chromosomes. *Genome Research*, *21*(10), 1592–1600.

Shevell, T., Malone, F. D., Vidaver, J., Porter, T. F., Luthy, D. A., Comstock, C. H., ... others. (2005). Assisted reproductive technology and pregnancy outcome. *Obstetrics & Gynecology*, *106*(5, Part 1), 1039–1045.

Singh, P., Cho, J., Tsai, S. Y., Rivas, G. E., Larson, G. P., & Szabó, P. E. (2010). Coordinated allele-specific histone acetylation at the differentially methylated regions of

imprinted genes. *Nucleic Acids Research*, *38*(22), 7974–7990.

Sleutels, F., Zwart, R., & Barlow, D. P. (2002). The non-coding air rna is required for silencing autosomal imprinted genes. *Nature*, *415*(6873), 810–813.

Smallwood, S. A., Tomizawa, S.-i., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., ... Kelsey, G. (2011). Dynamic cpg island methylation landscape in oocytes and preimplantation embryos. *Nature Genetics*, *43*(8), 811–814.

Smith, Z. D., & Meissner, A. (2013). DNA methylation: Roles in mammalian development. *Nature Reviews Genetics*, *14*(3), 204–220.

Smith, Z. D., Chan, M. M., Mikkelsen, T. S., Gu, H., Gnirke, A., Regev, A., & Meissner, A. (2012). A unique regulatory phase of dna methylation in the early mammalian embryo. *Nature*, *484*(7394), 339.

Solter, D. (1988). Differential imprinting and expression of maternal and paternal genomes. *Annual Review of Genetics*, *22*(1), 127–146.

Song, S., Ghosh, J., Mainigi, M., Turan, N., Weinerman, R., Truongcao, M., ... Sapienza, C. (2015). DNA methylation differences between in vitro-and in vivo-conceived children are associated with art procedures rather than infertility. *Clinical Epigenetics*, *7*(1), 41.

Soubry, A., Hoyo, C., Jirtle, R. L., & Murphy, S. K. (2014). A paternal environmental legacy: Evidence for epigenetic inheritance through the male germ line. *Bioessays*, *36*(4), 359–371.

Soubry, A., Murphy, S., Wang, F., Huang, Z., Vidal, A., Fuemmeler, B., ... others. (2015). Newborns of obese parents have altered dna methylation patterns at imprinted genes. *International Journal of Obesity*, *39*(4), 650–657.

Soubry, A., Schildkraut, J. M., Murtha, A., Wang, F., Huang, Z., Bernal, A., ... Hoyo, C. (2013). Paternal obesity is associated with igf2 hypomethylation in newborns: Results from a newborn epigenetics study (nest) cohort. *BMC Medicine*, *11*(1), 29.

Stewart, K. R., Veselovska, L., Kim, J., Huang, J., Saadeh, H., Tomizawa, S.-i., ... Kelsey, G. (2015). Dynamic changes in histone modifications precede de novo dna methylation in oocytes. *Genes & Development*, *29*(23), 2449–2462.

Stringer, J., Pask, A., Shaw, G., & Renfree, M. (2014). Post-natal imprinting: Evidence from marsupials. *Heredity*, *113*(2), 145–155.

Sun, F.-L., Dean, W. L., Kelsey, G., Allen, N. D., & Reik, W. (1997). Transactivation of igf2 in a mouse model of beckwith–Wiedemann syndrome. *Nature*, *389*(6653), 809.

Surani, M., Barton, S. C., & Norris, M. (1984). Development of reconstituted mouse eggs

suggests imprinting of the genome during gametogenesis. *Nature*, *308*(5959), 548–550.

Sutcliffe, A., Peters, C., Bowdin, S., Temple, K., Reardon, W., Wilson, L., ... Maher, E. (2005). Assisted reproductive therapies and imprinting disorders: A preliminary british survey. *Human Reproduction*, *21*(4), 1009–1011.

Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: Provocative insights from epigenomics. *Nature Reviews Genetics*, *9*(6), 465–476.

Suzuki, S., Ono, R., Narita, T., Pask, A. J., Shaw, G., Wang, C., ... others. (2007). Retrotransposon silencing by dna methylation can drive mammalian genomic imprinting. *PLoS Genetics*, *3*(4), e55.

Takagi, N., & Sasaki, M. (1975). Preferential inactivation of the paternally derived x chromosome in the extraembryonic membranes of the mouse. *Nature*, *256*(5519), 640–642.

Tang, L., Liu, Z., Zhang, R., Su, C., Yang, W., Yao, Y., & Zhao, S. (2017). Imprinting alterations in sperm may not significantly influence art outcomes and imprinting patterns in the cord blood of offspring. *PloS One*, *12*(11), e0187869.

Tian, X., Pascal, G., & Monget, P. (2009). Evolution and functional divergence of nlrp genes in mammalian reproductive systems. *BMC Evolutionary Biology*, *9*(1), 202.

Trapnell, C., Pachter, L., & Salzberg, S. L. (2009). TopHat: Discovering splice junctions with rna-seq. *Bioinformatics*, *25*(9), 1105–1111.

Umlauf, D., Goto, Y., Cao, R., Cerqueira, F., Wagschal, A., Zhang, Y., & Feil, R. (2004). Imprinting along the kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of polycomb group complexes. *Nature Genetics*, *36*(12), 1296.

Varmuza, S., & Mann, M. (1994). Genomic imprinting—defusing the ovarian time bomb. *Trends in Genetics*, *10*(4), 118–123.

Vassena, R., Boué, S., González-Roca, E., Aran, B., Auer, H., Veiga, A., & Belmonte, J. C. I. (2011). Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. *Development*, *138*(17), 3699–3709.

Vidal, A., Murphy, S., Murtha, A., Schildkraut, J., Soubry, A., Huang, Z., ... others. (2013). Associations between antibiotic exposure during pregnancy, birth weight and aberrant methylation at imprinted genes among offspring. *International Journal of Obesity*, *37*(7), 907–913.

Vu, T. H., Li, T., Nguyen, D., Nguyen, B. T., Yao, X.-M., Hu, J.-F., & Hoffman, A. R. (2000). Symmetric and asymmetric dna methylation in the human igf2–H19 imprinted

region. *Genomics*, *64*(2), 132–143.

Waddintong, C. (1942). The epigenotype. *Endeavour*, *1*, 18–20.

Wagschal, A., & Feil, R. (2006). Genomic imprinting in the placenta. *Cytogenetic and Genome Research*, *113*(1-4), 90–98.

Wake, N., Takagi, N., & Sasaki, M. (1978). Androgenesis as a cause of hydatidiform mole 2. *Journal of the National Cancer Institute*, *60*(1), 51–57.

Wallace, J. M., Bourke, D. A., Aitken, R. P., Milne, J. S., & Hay, W. W. (2003). Placental glucose transport in growth-restricted pregnancies induced by overnourishing adolescent sheep. *The Journal of Physiology*, *547*(1), 85–94.

Wang, C.-H., Su, P.-T., Du, X.-Y., Kuo, M.-W., Lin, C.-Y., Yang, C.-C., ... others. (2010). Thrombospondin type i domain containing 7A (thsd7a) mediates endothelial cell migration and tube formation. *Journal of Cellular Physiology*, *222*(3), 685–694.

Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., ... others. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature Genetics*, *40*(7), 897–903.

Wang, Z.-Q., Fung, M. R., Barlow, D. P., & Wagner, E. F. (1994). Regulation of embryonic growth and lysosomal targeting by the imprinted igf2/mpr gene. *Nature*, *372*(6505), 464.

Wolf, J. B., & Hager, R. (2006). A maternal–offspring coadaptation theory for the evolution of genomic imprinting. *PLoS Biology*, *4*(12), e380.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C. J., Zakhartchenko, V., Boiani, M., ... Walter, J. (2011). 5-hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nature Communications*, *2*, 241.

Wu, C.-t., & Morris. (2001). Genes, genetics, and epigenetics: A correspondence. *Science*, *293*(5532), 1103–1105.

Wu, G., Bazer, F. W., Cudd, T. A., Meininger, C. J., & Spencer, T. E. (2004). Maternal nutrition and fetal development. *The Journal of Nutrition*, *134*(9), 2169–2172.

Wu, G., Pond, W. G., Flynn, S. P., Ott, T. L., & Bazer, F. W. (1998). Maternal dietary protein deficiency decreases nitric oxide synthase and ornithine decarboxylase activities in placenta and endometrium of pigs during early gestation. *The Journal of Nutrition*, *128*(12), 2395–2402.

Wu, X. (2008). Maternal depletion of nlrp5 blocks early embryogenesis in rhesus macaque monkeys (macaca mulatta). *Human Reproduction*, *24*(2), 415–424.

Xie, L., Mouillet, J.-F., Chu, T., Parks, W. T., Sadovsky, E., Knöfler, M., & Sadovsky, Y. (2014). C19MC microRNAs regulate the migration of human trophoblasts. *Endocrinology*,

*155*(12), 4975–4985.

Xu, X.-R., Fu, R.-g., Wang, L.-Y., Wang, N., Zhang, F., Le, F., ... others. (2014). Epigenetic inheritance of paternally expressed imprinted genes in the testes of icsi mice. *Current Pharmaceutical Design*, *20*(11), 1764–1771.

Xu, Y., Shi, Y., Fu, J., Yu, M., Feng, R., Sang, Q., ... others. (2016). Mutations in padi6 cause female infertility characterized by early embryonic arrest. *The American Journal of Human Genetics*, *99*(3), 744–752.

Xue, F., Tian, X. C., Du, F., Kubota, C., Taneja, M., Dinnyes, A., ... Yang, X. (2002). Aberrant patterns of x chromosome inactivation in bovine clones. *Nature Genetics*, *31*(2), 216–220.

Xue, Z., Huang, K., Cai, C., Cai, L., Jiang, C.-y., Feng, Y., ... others. (2013). Genetic programs in human and mouse early embryos revealed by single-cell rna [thinsp] sequencing. *Nature*, *500*(7464), 593–597.

Yamanaka, Y., Lanner, F., & Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development*, *137*(5), 715–724.

Yamasaki, Y., Kayashima, T., Soejima, H., Kinoshita, A., Yoshiura, K.-i., Matsumoto, N., ... others. (2005). Neuron-specific relaxation of igf2r imprinting is associated with neuron-specific histone modifications and lack of its antisense transcript air. *Human Molecular Genetics*, *14*(17), 2511–2520.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., ... others. (2013). Single-cell rna-seq profiling of human preimplantation embryos and embryonic stem cells. *Nature Structural & Molecular Biology*, *20*(9), 1131–1139.

Yoder, J. A., Walsh, C. P., & Bestor, T. H. (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics*, *13*(8), 335–340.

Yu, X.-J., Yi, Z., Gao, Z., Qin, D., Zhai, Y., Chen, X., ... others. (2014). The subcortical maternal complex controls symmetric division of mouse zygotes by regulating f-actin dynamics. *Nature Communications*, *5*.

Yuen, R. K., Jiang, R., Peñaherrera, M. S., McFadden, D. E., & Robinson, W. P. (2011). Genome-wide mapping of imprinted differentially methylated regions by dna methylation profiling of human placentas from triploidies. *Epigenetics & Chromatin*, *4*(1), 10.

Yuen, R. K., Penaherrera, M. S., Von Dadelszen, P., McFadden, D. E., & Robinson, W. P. (2010). DNA methylation profiling of human placentas reveals promoter hypomethylation of multiple genes in early-onset preeclampsia. *European Journal of Human Genetics*, *18*(9), 1006–1012.

Zadora, J., Singh, M., Herse, F., Przybyl, L., Haase, N., Golic, M., ... others. (2017).

Disturbed placental imprinting in preeclampsia leads to altered expression of dlx5, a human-specific early trophoblast marker. *Circulation*, *136*(19), 1824–1839.

Zhang, C., Williams, M. A., King, I. B., Dashow, E. E., Sorensen, T. K., Frederick, I. O., ... Luthy, D. A. (2002). Vitamin c and the risk of preeclampsia: Results from dietary questionnaire and plasma assay. *Epidemiology*, *13*(4), 409–416.

Zhang, H., Zeitz, M. J., Wang, H., Niu, B., Ge, S., Li, W., ... others. (2014). Long noncoding rna-mediated intrachromosomal interactions promote imprinting at the kcnq1 locus. *J Cell Biol*, *204*(1), 61–75.

Zhang, P., Dixon, M., Zucchelli, M., Hambiliki, F., Levkov, L., Hovatta, O., & Kere, J. (2008). Expression analysis of the nlrp gene family suggests a role in human preimplantation development. *PloS One*, *3*(7), e2755.

Zhang, Q., Qi, S., Xu, M., Yu, L., Tao, Y., Deng, Z., ... Wong, J. (2013). Structure-function analysis reveals a novel mechanism for regulation of histone demethylase lsd2/aof1/kdm1b. *Cell Research*, *23*(2), 225–241.

Zhu, K., Yan, L., Zhang, X., Lu, X., Wang, T., Yan, J., ... Li, L. (2014). Identification of a human subcortical maternal complex. *MHR: Basic Science of Reproductive Medicine*, *21*(4), 320–329.

Zhu, L., Lv, R., Kong, L., Cheng, H., Lan, F., & Li, X. (2015). Genome-wide mapping of 5mC and 5hmC identified differentially modified genomic regions in late-onset severe preeclampsia: A pilot study. *PloS One*, *10*(7), e0134119.

Zhu, P., Guo, H., Ren, Y., Hou, Y., Dong, J., Li, R., ... others. (2018). Single-cell dna methylome sequencing of human preimplantation embryos. *Nature Genetics*, *50*(1), 12.

Zhu, X.-m., Han, T., Sargent, I. L., Yin, G.-w., & Yao, Y.-q. (2009). Differential expression profile of microRNAs in human placentas from preeclamptic pregnancies vs normal pregnancies. *American Journal of Obstetrics & Gynecology*, *200*(6), 661–e1.

Zuo, X., Sheng, J., Lau, H.-T., McDonald, C. M., Andrade, M., Cullen, D. E., ... others. (2012). Zinc finger protein zfp57 requires its co-factor to recruit dna methyltransferases and maintains dna methylation imprint in embryonic stem cells via its transcriptional repression domain. *Journal of Biological Chemistry*, *287*(3), 2107–2118.

# Appendix I: Supplementary Figures and Tables

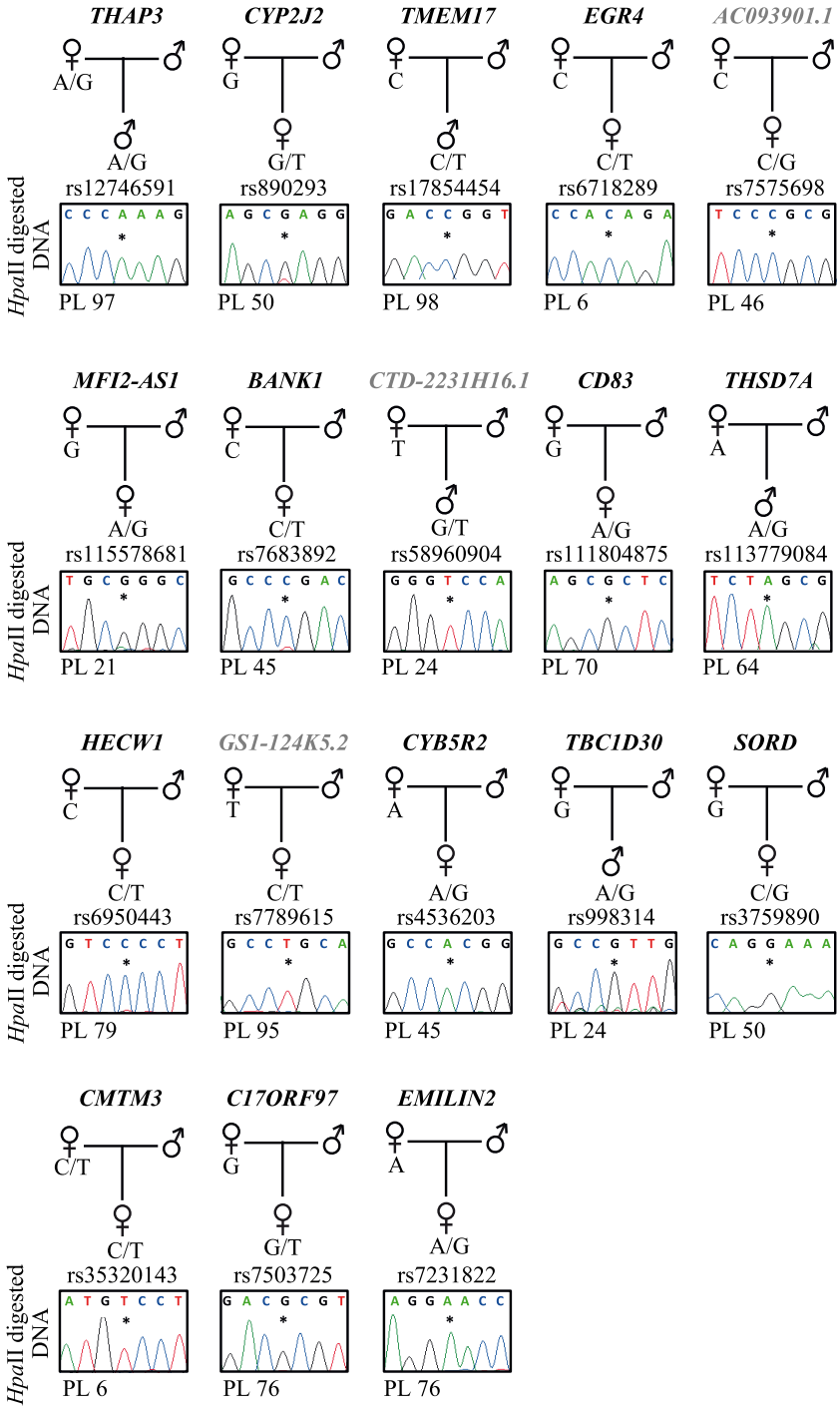



Figure S1: **Confirmation of allelic methylation at novel placenta-specific DMRs (array candidates) by *Hpa*II digestion.** Methylation profiles in placenta samples (*PL*) as determined by methylation-sensitive genotyping. The asterisk* on the electropherogram highlights the position of the SNP.

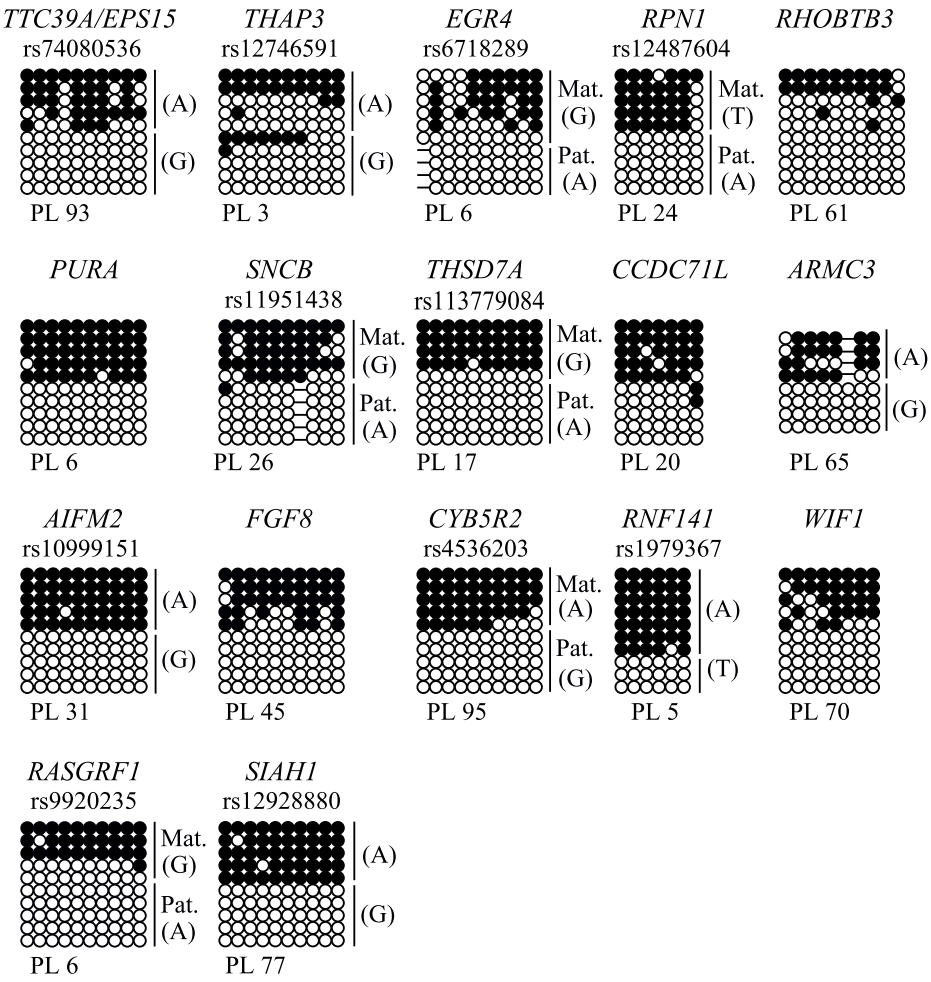



Figure S2: **Confirmation of allelic methylation at novel placenta-specific DMRs (array candidates) by bisulphite PCR and subcloning.** Confirmation of methylation profile in placenta (*PL*) by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

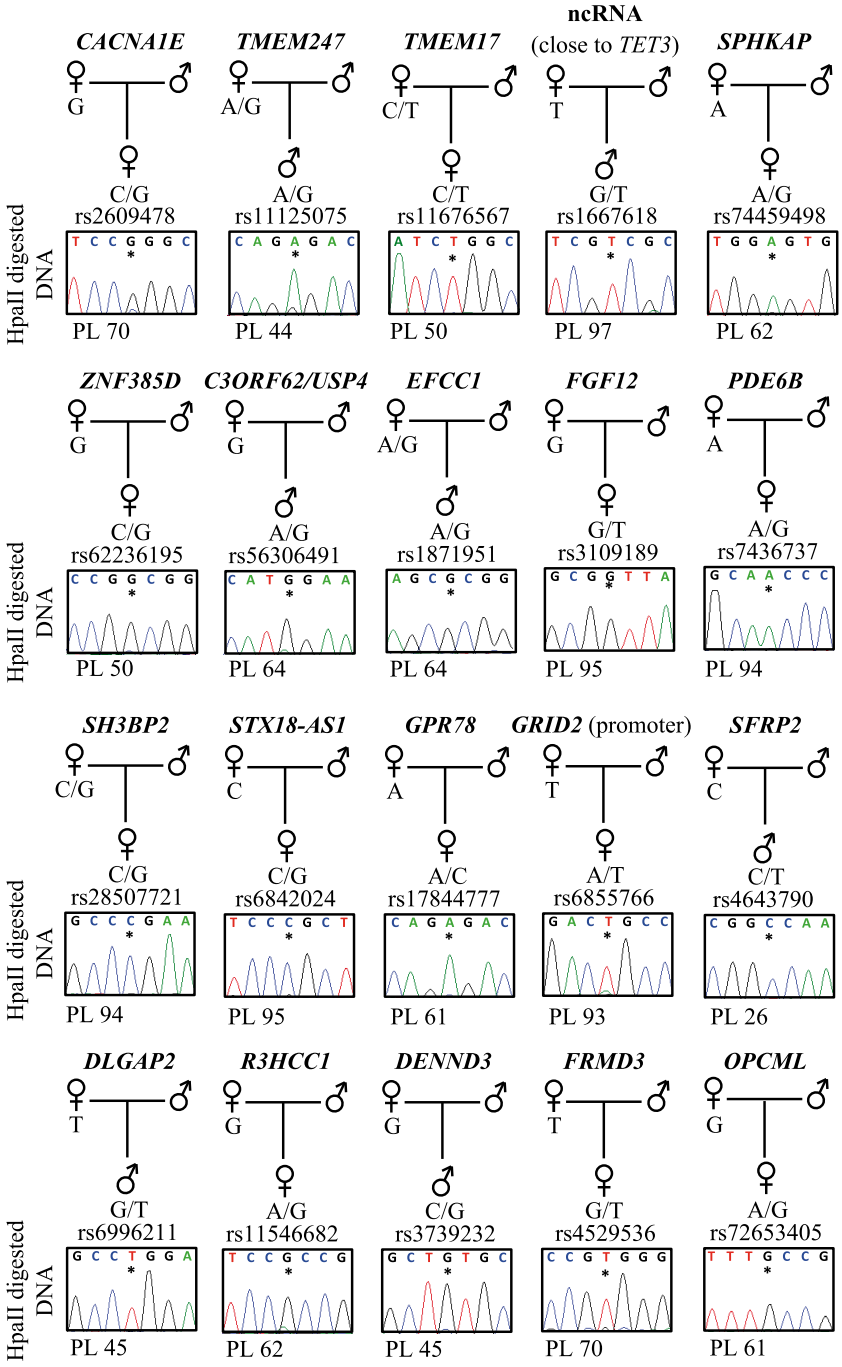



Figure S3: **Confirmation of allelic methylation at novel placenta-specific DMRs (methyl-seq candidates) by *Hpa*II digestion.** Methylation profiles in placenta samples (*PL*) as determined by methylation-sensitive genotyping. The asterisk* on the electropherogram highlights the position of the SNP.

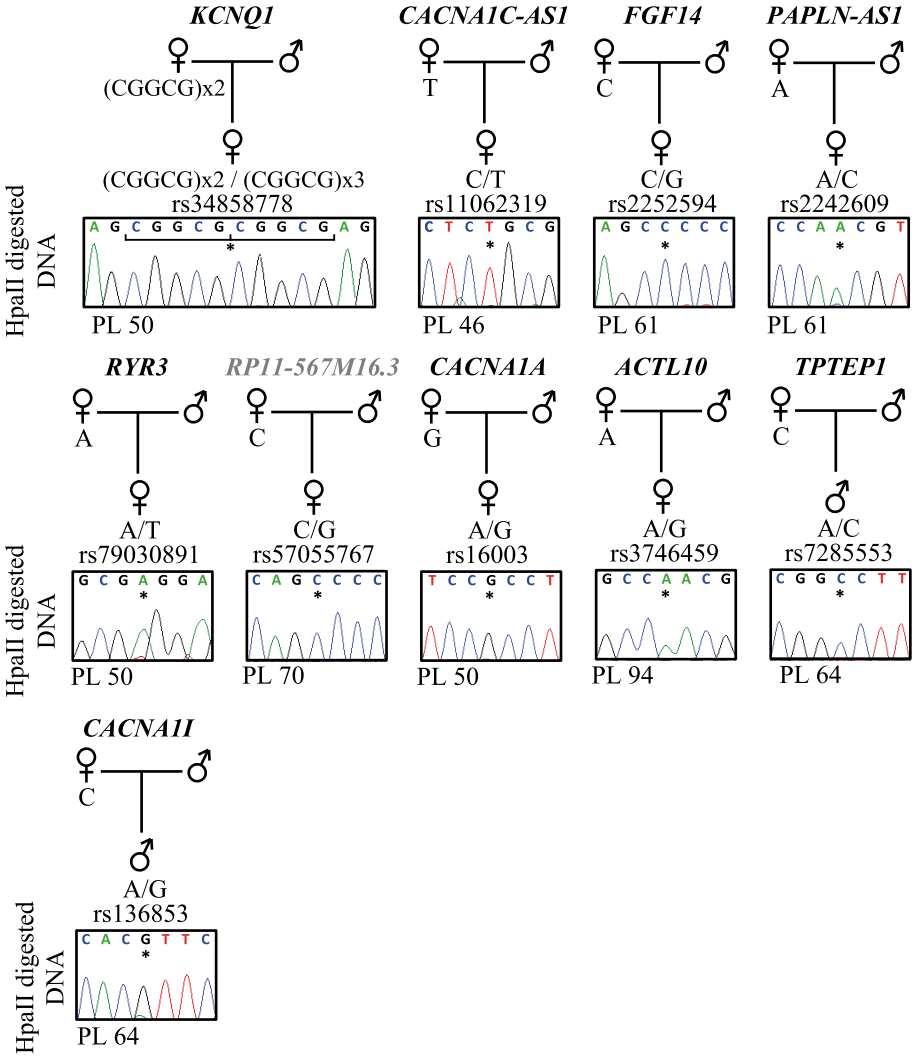



Figure S4: **Confirmation of allelic methylation at novel placenta-specific DMRs (methyl-seq candidates) by *Hpa*II digestion. (*cont*)** Methylation profiles placenta samples (*PL*) as determined by methylation-sensitive genotyping. The asterisk* on the electropherogram highlights the position of the SNP.

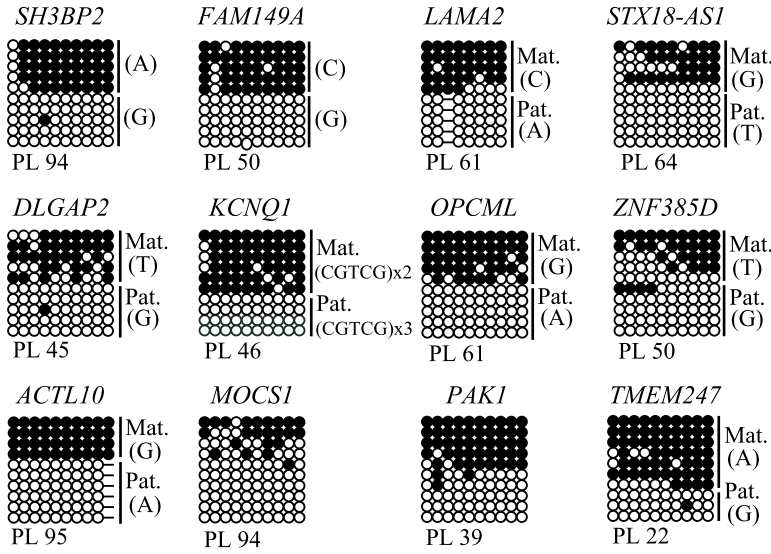



Figure S5: **Confirmation of allelic methylation at novel placenta-specific DMRs (methyl-seq candidates) by bisulphite PCR and subcloning.** Confirmation of methylation profile in placenta (*PL*) by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

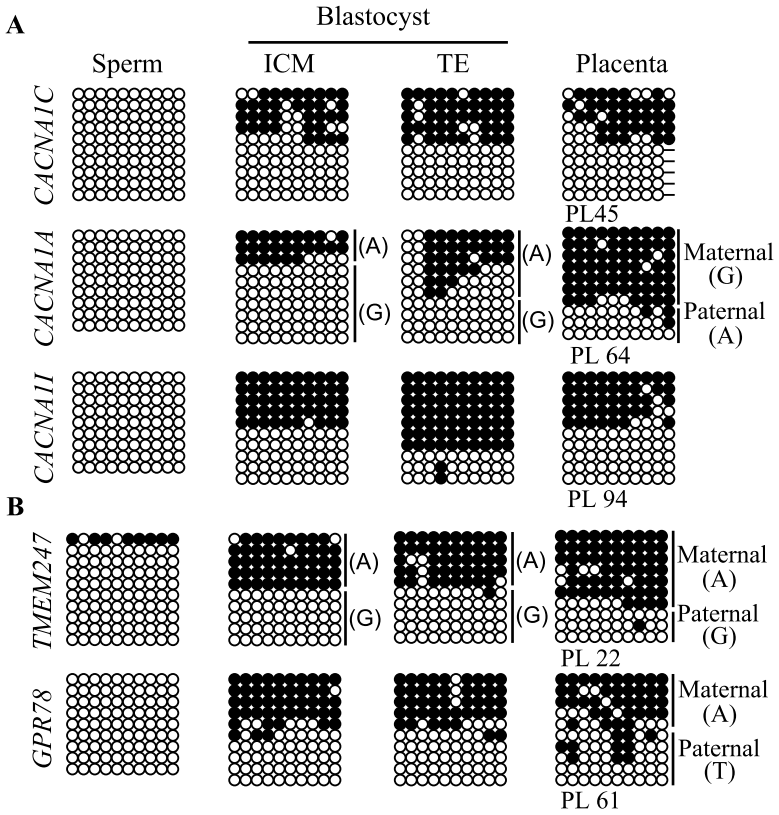



Figure S6: **Confirmation of unmethylated sperm and allelic methylation in blastocyst (ICM, TE) and placenta (PL) at placenta-specific DMRs.** Confirmation at (**A**) DMRs within 3 CACNA1 family members and (**B**) other placenta-specific DMRs within other epigenetic regulator genes (*GPR78* and *TMEM247*) by bisulphite PCR and subcloning. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype and parental-origin allele (*Mat.*: maternal allele; *Pat.*: paternal allele) if maternal blood samples were informative.

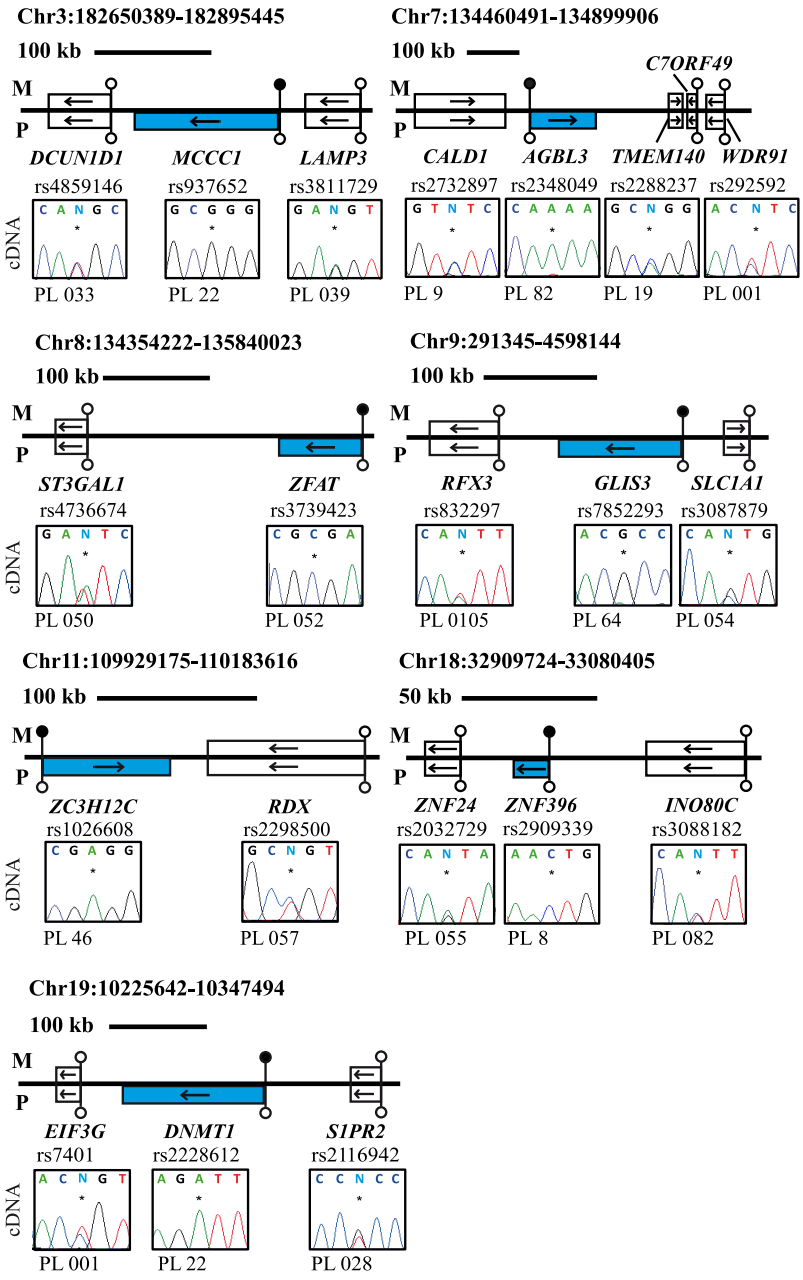



Figure S7: **Allele-specific RT-PCR analysis of candidate placenta-specific imprinted genes.** Confirmation of paternal expression of *MCCC1, AGBL3, ZFAT, GLIS3, ZC3HC12C, ZNF396* and *DNMT1* and no non-expressed or biallelically expressed surrounding genes in term placenta samples (PL). The asterisk* on the electropherogram highlights the position of the SNP. The maternally inherited chromosome (**M**, the upper part of the line) and the paternally inherited chromosomes (**P**, the lower part) of each placenta-specific imprinted locus are represented. The black lollipops indicate methylated CpG regions whereas white circles, unmethylated CpG regions. Blue filled rectangles represent paternally expressed genes; White filled rectangles represent non-imprinted ballelically; Grey filled rectangle represents non-expressed genes. The arrows represent the direction of the transcripts.

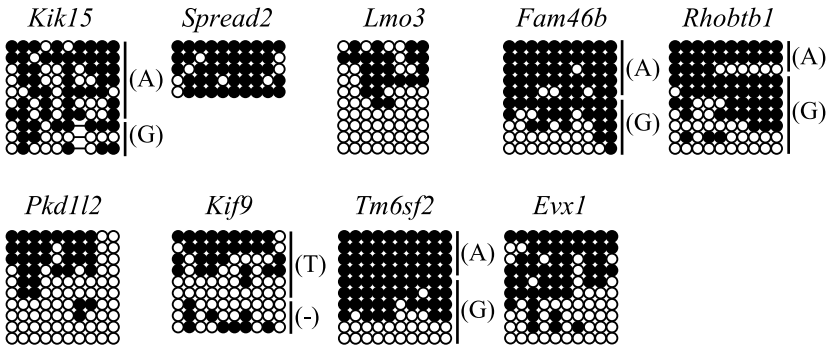



Figure S8: **Mouse placenta-specific gDMR cadidates (false positives).** Strand-specific methylation analysis in DNA-derived from mouse B6 x JF1 placenta samples for mouse gDMRs that maintain as partially methylated region in mouse placenta methyl-seq datasets. Each circle represents a single CpG dinucleotide on a DNA strand; the black circle indicate a methylated cytosine whereas white circles, an unmethylated cytosine. For clarity, only the first 10 CpG dinucleotides from each amplicon is shown, with the letters in the parentheses indicating SNP genotype.

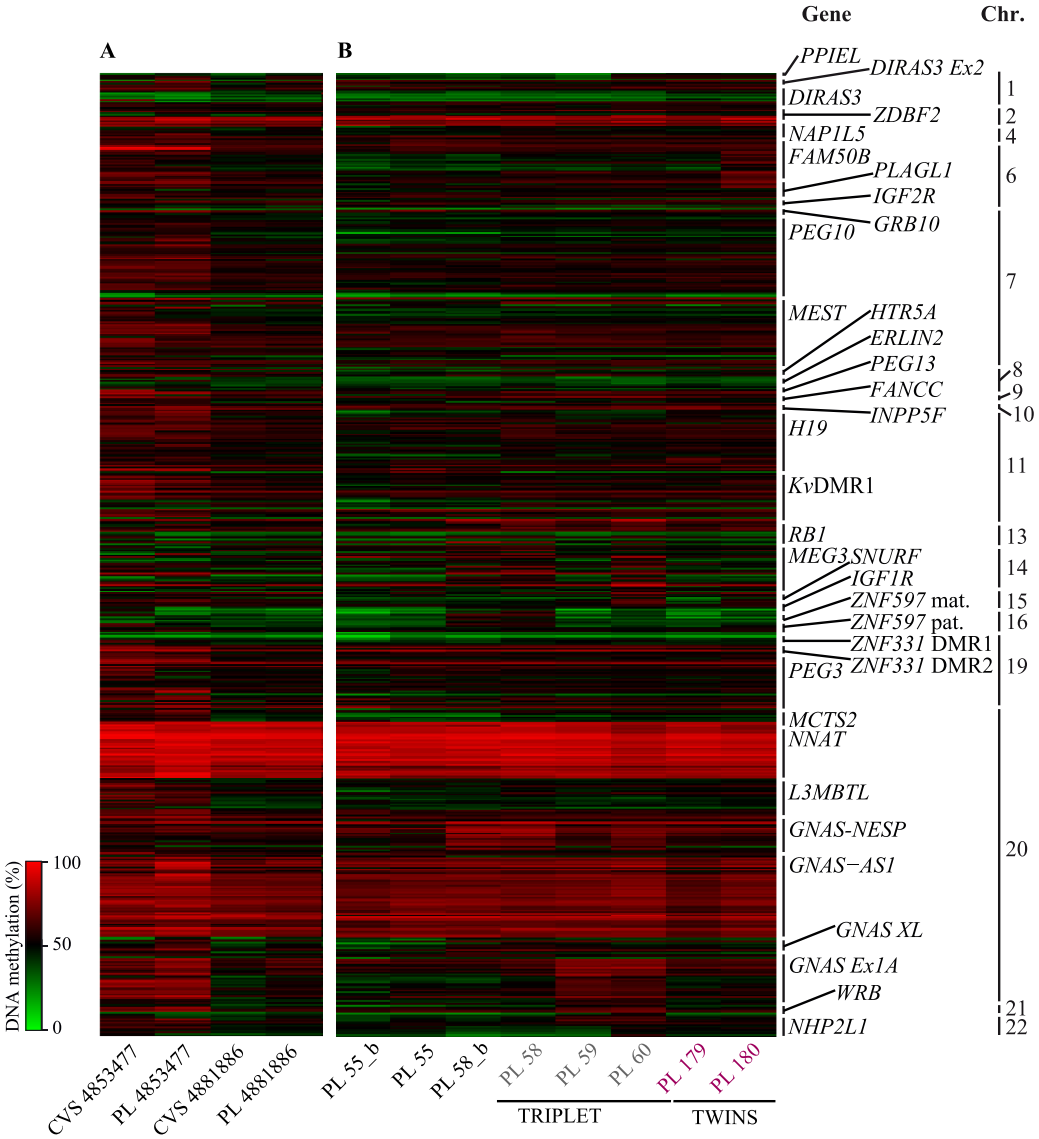



Figure S9: **Genome-wide methylation analysis at ubiquitous DMRs of paired samples (CVS vs corresponding term placentas), multiples biopsies from the same placenta and dizygotic twins or triplet pregnancy.** A Heatmap of the Infinium probes located within ubiquitous imprinted DMRs. (**A**) Paired samples: first-trimester CVS compared with corresponding biopsies from the term placentas (PL 4853477 and PL 4881886). (**B**) Multiple biopsies from the same term placentas (PL 55 and PL 58) and from a triplet pregnancy (PL 58, PL 59 and PL 60) and dizygotic twins (PL 179 and PL 180). – *data analysis and graph generated by Dr Hernandez-Mora* –.

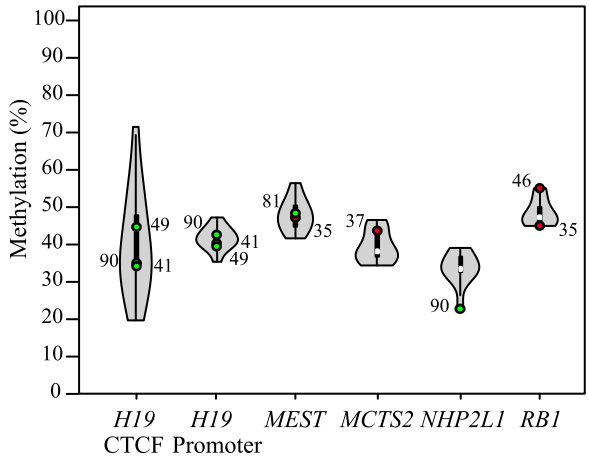



Figure S10: **Quantitative methylation analysis of cord blood samples in ubiquitous imprinted regions.** Violine Plot used include the mean (white dot) and the interquartile range (black box). Additionally, Violine plot also show a rotated kernel density plot on each side (non-parametric way to estimate the probability density of the data at different values) – *generated in R with the help of Dr Hernande-Mora* –. Quantitative pyrosequencing of 5 ubiquitous regions (3 of them corresponding to *H19* region) in cord blood samples (n = 20). Violin plot shows the distribution of DNA methylation values for all samples, green-filled dots are the cord blood methylation value corresponding to LOM placentas and red filled dots to GOM placentas.

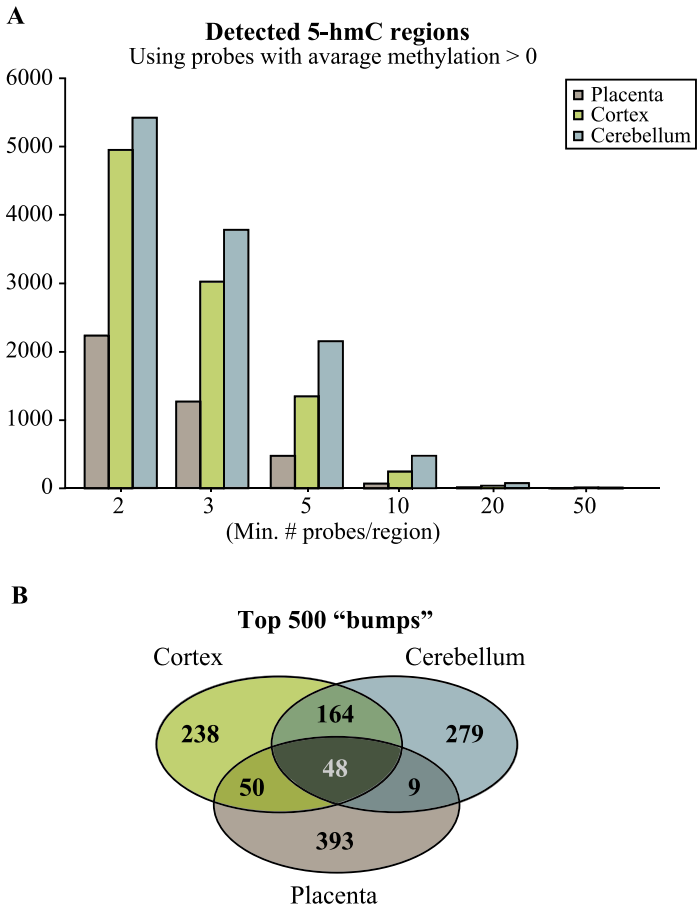



Figure S11: **Identified regions enriched for clusters of probes by Bumphunter function adapted for the Illumina Infinium HM450K array.** (**A**) A bar graph showing inclusion criteria for Bumphunter function with the number of probes contained within each 800 pb window (**B**) Venn diagrams illustrating the top 500 "bumps" obtained from oxBS-450k Bumphunter analysis with a minimum of two probes in placenta, frontal cortex and cerebellum – *data analysis and graph generated by Dr Hernandez-Mora* –.

## SUPPLEMENTARY TABLES

Table S1: **Clinical and anthropometric data associated with our placenta cohort samples.**

| ID | ART | Pre-ecl. | IUGR | Mat. Blood | NB sex | Mult. Gest. | Gest. Age | PL weight | NB weight | NB len | Mat Age |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PL 1[1] | No[2] | No[3] | No[4] | No[5] | F[6] | No[7] | 286[8] | 650[9] | 3390[10] | 50[11] | 33[12] |
| PL 2 | No | No | No | No | F | No | 273 | NA | 3180 | NA | 32 |
| PL 3 | No | No | No | No | M | No | 275 | 710 | 3760 | 51 | 25 |
| PL 5 | No | No | No | No | F | No | 266 | 670 | 3355 | 49 | 39 |
| PL 6 | No | No | No | No | F | No | 276 | NA | 3300 | 50 | 19 |
| PL 7 | No | No | No | No | F | No | 282 | 600 | 2800 | 50 | 26 |
| PL 8 | No | No | No | No | F | No | 289 | 730 | 3670 | 52,5 | 17 |
| PL 9 | No | No | No | No | M | No | 260 | 590 | 2790 | 46,5 | 35 |
| PL 10 | No | No | No | Yes | F | No | 240 | 600 | 2420 | 47 | 25 |
| PL 11 | No | Yes | No | No | F | No | 266 | 490 | 2330 | NA | 34 |
| PL 12 | No | No | No | No | M | No | 270 | 520 | 3250 | 49 | 26 |
| PL 13 | No | No | No | No | F | No | 271 | NA | 3250 | 49 | 36 |
| PL 14 | No | No | Yes | No | F | No | 277 | 630 | 2510 | 47 | 31 |
| PL 15 | No | No | No | No | F | No | 271 | 540 | 2800 | 48 | 36 |
| PL 16 | No | No | No | No | F | No | 280 | 600 | 3840 | NA | 35 |
| PL 17 | No | No | No | No | F | No | 287 | 590 | 3300 | NA | 33 |
| PL 18 | No | No | No | No | F | No | 269 | 690 | 3325 | 49 | 28 |
| PL 19 | No | No | No | No | M | No | 270 | 830 | 3490 | 51 | 27 |
| PL 20 | No | No | No | No | M | No | 278 | 980 | 3840 | 52 | 34 |
| PL 21 | No | No | No | No | F | No | 265 | 850 | 3280 | 51 | 25 |
| PL 22 | No | No | No | No | M | No | 283 | 670 | 3630 | NA | 25 |
| PL 23 | No | No | No | No | M | No | 283 | 750 | 3380 | 52 | 34 |
| PL 24 | No | No | Yes | No | M | No | 264 | NA | 2480 | 48 | 31 |
| PL 25 | No | No | Yes | Yes | M | Mon | 235 | 810 | 1410 | 42 | 35 |
| PL 26 | No | No | No | Yes | M | Mon | 235 | 810 | 1980 | 45 | 35 |

[1] Placenta sample ID.
[2] *ART*: Assisted reproductive technique.
[3] *Preecl.*: pre-eclampsia.
[4] *IUGR*: Intrauterine growth restriction.
[5] *Mat. Blood*: corresponding to maternal blood.
[6] *NB sex:* Newborn sex (*F*: fermale / *M*: male).
[7] *Mult. Gest.:* If it was a multiple-gestation pregnancy (*Mon*: monozygotic / *Diz*: dizygotic / *Trip*: triplet).
[8] *Gest. Age:* Gestational age (days).
[9] *PL weight*: placenta weight (g).
[10] *NB weight*: newborn weight (g).
[11] *NB len;* newborn length (cm).
[12] *Mat. Age:* maternal age.

| ID | ART | Pre-ecl. | IUGR | Mat. Blood | NB sex | Mult. Gest. | Gest. Age | PL weight | NB weight | NB len | Mat Age |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PL 27 | No | No | No | No | F | No | 266 | 630 | 3400 | 50 | 22 |
| PL 28 | Yes | No | No | Yes | M | Diz | 164 | NA | 575 | 32 | 38 |
| PL 29 | Yes | No | No | Yes | F | Diz | 165 | NA | 567 | 30 | 38 |
| PL 30 | No | No | No | No | F | No | 289 | 720 | 3000 | NA | 27 |
| PL 31 | No | No | No | Yes | F | No | 234 | 550 | 1730 | NA | 28 |
| PL 32 | No | NA | No | No | NA | No | 255 | 630 | 2430 | 47 | 30 |
| PL 33 | No | No | No | No | M | No | 272 | 880 | 3360 | 49 | 37 |
| PL 34 | No | No | No | No | M | No | 271 | 940 | 3670 | 50 | 24 |
| PL 35 | No | No | No | No | F | No | 281 | 820 | 3106 | 50 | 31 |
| PL 37 | No | No | Yes | No | F | No | 261 | 630 | 2520 | 46 | 34 |
| PL 38 | No | No | No | No | M | No | 271 | 700 | 3350 | 50 | 32 |
| PL 39 | No | No | No | No | M | No | 273 | 1450 | 4980 | 52 | 34 |
| PL 41 | Yes | No | No | Yes | M | No | 207 | 520 | 1535 | NA | 31 |
| PL 42 | No | No | No | Yes | M | Mon | 224 | 980 | 184 | 44 | 34 |
| PL 43 | No | No | No | Yes | M | Mon | 224 | 980 | 1830 | 43 | 34 |
| PL 44 | No | No | Yes | Yes | M | No | 259 | 480 | 1700 | 40 | 31 |
| PL 45 | No | No | Yes | No | F | No | 280 | 520 | 2740 | 51 | 22 |
| PL 46 | No | Yes | Yes | No | F | No | 278 | 530 | 2600 | 48 | 29 |
| PL 47 | No | Yes | Yes | Yes | M | No | 182 | 200 | 620 | 34 | 37 |
| PL 49 | No | No | No | Yes | M | No | 181 | 270 | 940 | 35 | 24 |
| PL 50 | No | No | Yes | Yes | F | No | 230 | 490 | 2105 | NA | 32 |
| PL 51 | No | No | Yes | No | F | No | 261 | NA | 2120 | 45 | 40 |
| PL 52 | No | NA | Yes | No | F | No | 256 | NA | 2240 | NA | 39 |
| PL 53 | No | No | Yes | Yes | M | No | 221 | NA | 1120 | 39 | 27 |
| PL 54 | No | No | Yes | No | M | No | 276 | 430 | 2740 | 48 | 33 |
| PL 55 | No | No | Yes | No | M | No | 271 | NA | 2320 | 45 | 30 |
| PL 56 | No | No | No | No | M | No | 277 | 550 | 3590 | 49 | 32 |
| PL 58 | Yes | No | No | Yes | M | Trip | 224 | 300 | 1700 | NA | 46 |
| PL 59 | Yes | No | Yes | Yes | M | Trip | 224 | 330 | 1530 | NA | 46 |
| PL 60 | Yes | No | Yes | Yes | F | Trip | 224 | 230 | 1150 | NA | 46 |
| PL 61 | No | No | Yes | No | F | No | 264 | 400 | 2330 | NA | 35 |
| PL 62 | No | No | No | Yes | F | Diz | 214 | 375 | 1425 | NA | 37 |
| PL 63 | No | No | No | Yes | F | Diz | 214 | 375 | 1290 | NA | 37 |
| PL 64 | No | No | Yes | Yes | M | No | 261 | 540 | 2160 | 47 | 34 |
| PL 65 | No | NA | Yes | Yes | M | No | 242 | 300 | 1620 | NA | 39 |
| PL 66 | NA | NA | Yes | No | M | No | NA | NA | 2370 | NA | 39 |
| PL 67 | No | No | Yes | Yes | M | No | 240 | NA | 1620 | 42 | 39 |

| ID | ART | Pre-ecl. | IUGR | Mat. Blood | NB sex | Mult. Gest. | Gest. Age | PL weight | NB weight | NB len | Mat Age |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PL 68 | No | No | Yes | Yes | F | No | 187 | 250 | 810 | NA | 26 |
| PL 69 | Yes | No | No | Yes | F | No | NA | 530 | 2380 | 48 | 40 |
| PL 70 | No | No | No | No | F | No | 284 | 530 | 3170 | 49 | 37 |
| PL 71 | No | No | Yes | Yes | F | No | 246 | 260 | 1345 | 40 | 25 |
| PL 72 | No | No | No | No | F | No | 238 | 660 | NA | NA | 28 |
| PL 73 | No | No | Yes | Yes | F | No | 258 | NA | 1900 | NA | 24 |
| PL 74 | No | No | No | Yes | M | No | 177 | 375 | 860 | NA | 34 |
| PL 76 | No | No | No | Yes | F | No | 239 | 530 | 2130 | 45 | 30 |
| PL 77 | Yes | No | No | Yes | F | Diz | 243 | 300 | 2070 | NA | 34 |
| PL 78 | Yes | No | Yes | Yes | F | Diz | 243 | 305 | 1460 | 39 | 34 |
| PL 79 | No | No | Yes | No | F | No | 282 | 470 | 2690 | 49 | 25 |
| PL 80 | Yes | No | No | Yes | F | Diz | 237 | 240 | 1800 | 44 | 28 |
| PL 81 | Yes | No | Yes | Yes | M | Diz | 237 | 350 | 1460 | 43 | 28 |
| PL 82 | No | No | Yes | No | F | No | 264 | 500 | 2460 | 46 | 30 |
| PL 86 | No | No | Yes | Yes | M | Diz | 207 | 150 | 800 | 32 | 31 |
| PL 87 | No | No | Yes | Yes | F | Diz | 207 | 150 | 960 | 37 | 31 |
| PL 88 | No | No | Yes | No | M | No | 275 | 530 | 2490 | 48 | 22 |
| PL 89 | No | No | No | Yes | M | Diz | 235 | 395 | 2410 | 47 | 34 |
| PL 90 | No | No | Yes | Yes | M | Diz | 235 | 195 | 1730 | 44 | 34 |
| PL 91 | No | No | Yes | Yes | F | No | 224 | 270 | 1330 | 40 | 34 |
| PL 92 | No | Yes | Yes | No | F | No | 261 | 450 | 2410 | 45 | 36 |
| PL 93 | No | No | No | Yes | F | No | 237 | 740 | 2600 | 48 | 35 |
| PL 94 | No | No | No | Yes | F | No | 230 | 530 | 2030 | 44 | 36 |
| PL 95 | No | Yes | No | Yes | F | No | 239 | 430 | 1815 | 42 | 35 |
| PL 97 | No | No | No | Yes | M | No | 233 | 530 | 1690 | NA | 43 |
| PL 98 | No | No | Yes | Yes | M | No | 251 | 400 | 1920 | 42 | 37 |
| PL 99 | No | No | Yes | Yes | M | No | 252 | 450 | 1940 | 40 | 29 |
| PL 100 | No | No | Yes | Yes | M | No | 260 | 470 | 2070 | NA | 25 |
| PL 101 | No | No | Yes | No | F | No | 257 | 650 | 2270 | 45 | 39 |
| PL 104 | No | No | Yes | No | M | No | 259 | 530 | 2080 | 44 | 27 |
| PL 106 | Yes | No | No | Yes | M | Trip | 229 | NA | 2005 | 44 | 29 |
| PL 107 | Yes | No | No | Yes | F | Trip | 229 | NA | 1880 | 44,5 | 29 |
| PL 108 | Yes | No | Yes | Yes | F | Trip | 229 | NA | 1825 | 43 | 29 |
| PL 109 | No | No | Yes | No | M | No | 260 | NA | 2230 | 45 | 27 |
| PL 126 | Yes | No | Yes | No | M | No | 273 | 480 | 2880 | NA | 31 |
| PL 127 | Yes | No | No | No | M | No | 250 | 520 | 2400 | 46 | 21 |
| PL 135 | No | No | No | Yes | M | No | 185 | NA | 970 | 36 | 16 |

| ID | ART | Pre-ecl. | IUGR | Mat. Blood | NB sex | Mult. Gest. | Gest. Age | PL weight | NB weight | NB len | Mat Age |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PL 136 | No | No | Yes | No | M | Diz | 259 | 320 | 1980 | 46,5 | 24 |
| PL 140 | Yes | No | No | No | M | No | 286 | 970 | 4090 | 51 | 35 |
| PL 141 | Yes | No | No | Yes | M | Diz | 257 | 900 | 3210 | 50,5 | 41 |
| PL 143 | Yes | No | Yes | Yes | F | Diz | 233 | 500 | 1980 | 44 | 32 |
| PL 144 | Yes | No | Yes | Yes | F | Diz | 233 | 250 | 1340 | 41 | 32 |
| PL 145 | Yes | No | No | Yes | F | No | 279 | 510 | 3350 | 52 | 30 |
| PL 146 | No | No | No | No | M | No | 271 | NA | 3100 | NA | 34 |
| PL 147 | No | Yes | No | Yes | F | No | 230 | 700 | 2000 | 44 | 34 |
| PL 148 | No | No | Yes | Yes | M | No | 263 | 500 | 1680 | 42,5 | 23 |
| PL 149 | No | No | No | Yes | F | Diz | 245 | 500 | 2350 | 45 | 38 |
| PL 151 | Yes | No | No | Yes | M | No | 244 | 650 | 2555 | 48 | 41 |
| PL 152 | Yes | No | Yes | Yes | F | No | 244 | 330 | 1770 | 43 | 41 |
| PL 154 | No | Yes | Yes | Yes | M | No | 249 | 255 | 1640 | 42 | 31 |
| PL 155 | No | No | Yes | Yes | F | Diz | 242 | NA | 1450 | 41 | 35 |
| PL 158 | Yes | No | Yes | Yes | F | Diz | 243 | 300 | 1670 | 41 | 32 |
| PL 159 | Yes | No | No | Yes | F | Diz | 243 | 300 | 2150 | 44,5 | 32 |
| PL 160 | No | No | Yes | No | F | No | 267 | 490 | 2210 | 46 | 38 |
| PL 161 | No | No | Yes | No | F | No | 276 | 430 | 2490 | 47,5 | 31 |
| PL 162 | Yes | No | No | Yes | F | No | 224 | 450 | 1885 | NA | 37 |
| PL 163 | Yes | No | No | Yes | F | Trip | 238 | 600 | 1870 | NA | 30 |
| PL 164 | Yes | No | No | Yes | M | Trip | 238 | 600 | 2200 | NA | 30 |
| PL 165 | Yes | No | No | Yes | F | Trip | 238 | 600 | 2200 | NA | 30 |
| PL 166 | No | NA | NA | Yes | NA | NA | NA | NA | NA | NA | NA |
| PL 167 | No | NA | NA | Yes | NA | NA | NA | NA | NA | NA | NA |
| PL 168 | NA | NA | No | Yes | NA | NA | NA | NA | NA | NA | NA |
| PL 170 | Yes | No | No | Yes | M | Diz | 255 | 465 | 2240 | 47 | 44 |
| PL 171 | Yes | No | Yes | Yes | M | Diz | 255 | 465 | 2030 | 45 | 44 |
| PL 172 | No | No | No | Yes | M | Diz | 178 | 220 | 890 | 30 | 31 |
| PL 175 | Yes | No | No | Yes | F | Diz | 203 | 285 | 1200 | 40 | 32 |
| PL 179 | No | No | No | No | F | Diz | 262 | NA | NA | NA | NA |
| PL 180 | No | No | No | No | F | Diz | 262 | NA | NA | NA | NA |
| PL 186 | Yes | No | No | Yes | F | Diz | 171 | 210 | 700 | 31 | 25 |
| PL 189 | Yes | NA | No | Yes | M | No | 285 | 585 | 3500 | NA | 36 |
| PL 190 | Yes | NA | No | Yes | F | No | 169 | 210 | 565 | 28,5 | 30 |
| PL 192 | Yes | No | No | Yes | M | Diz | 267 | 750 | 3300 | 50 | 41 |
| PL 212 | No | NA | NA | Yes | F | NA | NA | NA | NA | NA | NA |
| PL 215 | Yes | NA | No | Yes | F | Diz | 269 | 810 | 3630 | 51 | 33 |

| ID | ART | Pre-ecl. | IUGR | Mat. Blood | NB sex | Mult. Gest. | Gest. Age | PL weight | NB weight | NB len | Mat Age |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PL 216 | Yes | NA | No | Yes | M | Diz | 269 | 810 | 2720 | 49 | 33 |
| PL 217 | Yes | NA | Yes | Yes | NA | NA | NA | NA | NA | NA | NA |
| PL 222 | NA | NA | NA | Yes | NA | NA | NA | NA | NA | NA | NA |
| PL 225 | NA | NA | NA | Yes | NA | NA | NA | NA | NA | NA | NA |
| PL 226 | Yes | NA | No | No | NA | NA | NA | NA | NA | NA | NA |
| PL 230 | Yes | No | No | Yes | F | Trip | 234 | NA | 1675 | 44 | 47 |

Table S2: **Buffers used in the ChIP protocol**

| Buffers | (Working dilutions) |
|---|---|
| **BUFFER 2X:** | (Final concentration in 1X buffer): |
| 15 ml 1M KCl | 60 mM |
| 0,75ml 5M NaCl | 15mM |
| 1,25 ml 1M MgCl2 | 5mM |
| 0,5 ml 50mM EGTA | 0,1mM |
| 3,75 ml 1M Tris-HCl (pH 7,4-7,6) | 15mM |
| $H_2O$ to 125 ml | |
| **BUFFER 1** (0,3M sucrose): | |
| Sucrose 5,15 g | |
| 2X Buffer 25 ml | |
| $H_2O$ upto 50 ml | |
| 25$\mu l$ of 1M stock | DTT 0,5mM |
| 200$\mu l$ of 25mM stock | PMSF 0,1mM |
| 100$\mu l$ of 1,8 mg/ml stock | Aprotinin 3,6 ng/ml |
| 100$\mu l$ of 2.5M stock | 5mM Na-Butyrate |
| **BUFFER 2:** | |
| 10 ml buffer 1 | |
| 80$\mu l$ of NP40 (IGEPAL CA-630) | NP40 0,8% |
| **BUFFER 3** (1,2M sucrose): | |
| Sucrose 20,55g | |
| 2X Buffer 25ml | |
| $H_2O$ upto 50 ml | |
| 25$\mu l$ of 1M stock | DTT 0,5mM |
| 200$\mu l$ of 25mM stock, 50 $\mu l$ of 0,1M stock | PMSF 0,1mM |
| 100$\mu l$ of 1,8mg/ml stock | Aprotinin 3,6ng/ml |
| 100$\mu l$ of 2.5M stock | 5mM Na-Butyrate |
| **MNASE DIGESTION BUFFER:** | |
| For 10 ml: | |
| 1,1g sucrose | 0,32M Sucrose |
| 0,5ml Tris pH 7,5 1M | 50mM Tris pH 7,5 |
| 40$\mu l$ MgCl2 1M | 4mM MgCl2 |
| 10$\mu l$ CaCl2 1M | 1mM CaCl2 |
| 80$\mu l$ PMSF 25mM (20$\mu l$ PMSF 0,1M) | 0,1mM PMSF |
| 5mM Na-Butyrate | 20 $\mu l$ Na-Butyrate 2,5M |
| **LYSIS BUFFER:** | For 10 ml: |
| 1 mM Tris pH 7,5 | 10 ul tris pH 7,5 1M |

| Buffers | (Working dilutions) |
|---|---|
| 4$\mu l$ EDTA 0,5M | 0,2mM EDTA |
| 100$\mu l$ PMSF 25 mM (25 $\mu l$ of 0,1M stock) | 0,2mM PMSF |
| 100$\mu l$ Na-Butyrate 2,5M | 5mM Na butyrate |
| **INCUBATION BUFFER** | |
| | NaCL 50mM |
| | Tris pH 7,5 20mM |
| | Na-butyrate 20mM |
| | EDTA 5mM |
| | PMSF 0,1mM |
| **ELUTION BUFFER (for 5ml)** | |
| 50$\mu l$ NaCl 5M | |
| 100$\mu l$ Tris pH 7,5 1M | |
| 50$\mu l$ EDTA 0,5M | |
| 500$\mu l$ of 10% SDS | 1% final |
| **WASHING SOLUTIONS:** | |
| ***SOLUTION A:*** | |
| | 50mM Tris pH 7,5 |
| | 10mM EDTA |
| | 5mM Na-Butyrate |
| | 75mM NaCl |
| ***SOLUTION B:*** | |
| | 50mM Tris pH 7,5 |
| | 10mM EDTA |
| | 5mM Na-Butyrate |
| | 125mM NaCl |
| ***SOLUTION C:*** | |
| | 50mM Tris pH 7,5 |
| | 10mM EDTA |
| | 5mM Na-Butyrate |
| | 175mM NaCl |

**List of all primers – Forward (F) and Reverse (R) – used in this Thesis**

**Genotyping *NLRP7* mutations**

Table S3: ***NLRP7* primers – mutations in the recurrent hydatidiform moles**

| Gene | | | Sequence (5' -> 3') |
|---|---|---|---|
| NLRP7 | Exon 5 | F | CTTAAGATCTATACTGGTAGCAAAG |
| | | R | GAGATGAAACAGCACATTTCC |
| | Exon 6 | F | GTTTATACATGCCTCCACAC |
| | | R | GCTGGTTATGCAACACAGAAG |
| | Exon 10 | F | TACACGCTTGAGCCACTACC |
| | | R | GGACATGTTGGCATGCCTCTAG |
| | Exon 10 - Inner | F | TGGGGTTTGACCGTGTTGGC |
| | | R | ATGCCTCTAGGCCCAGCTAC |
| | Deletion - Out | F | GTCACAGCTGGGGAAAAAGATTTG |
| | | R | AGCCCACCACTGAATGACACA |
| | Deletion - In | F | CCTCCAAGGGTAAAGCTGTCC |
| | | R | ACAAATAGCAAGAGGACAGATA |

**Imprinted DMRs study**

***Methylation-sensitive genotyping PCRs***

Table S4: ***Hpa*II Primers – Placentas-specific DMRs in *NLRP7* HM450k array analysis**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| THAP3 | rs12746591 | F | TGGCTCGAGATGCCGAAGTCGT |
| | | R | CGTCCGCCACTCCCCGCATC |
| CYP2J2 | rs890293 | F | ACAAGGAAGAAGTTGCCAAGGA |
| | | R | CAGTAAACCCATGAGCTTTC |
| TMEM17 | rs17854454 | F | ACGCTCCAAATCCACGCCA |
| | | R | GGAACCACGGAGCCTCGGAGCG |
| EGR4 | rs6718289 | F | TGCCCAAACATGGACACAGGATG |
| | | R | CGTGAGCAAAAGAAATGAATGGC |
| Chr. 2 | rs7575698 | F | CACAAATATTTGGCTTTCCCCGGC |
| | | R | AGGGCGGGAAACTTGAGAACCA |
| MF12-AS1 | rs115578681 | F | GTTCGGAACCCAGAAGCATGG |
| | | R | ACAGTGGGGGACGAGGCAGGGCA |
| BANK1 | rs7683892 | F | GAGGAGGGAGCGCGCTGGGAG |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | R | TACCAATTCACCAGGCTGGGAT |
| Chr. 5 | rs58960904 | F | TCGGGGAGGAGCCGCCCAGGT |
| | | R | CGCCGTCTGCTGTGCTGAAAC |
| CD83 | rs111804875 | F | GCCCCGGCCTAAGCGGGACTAG |
| | | R | AGAGAGAGCGCGTCGTGAGCA |
| THSD7A | rs113779084 | F | AGCGCCCCGGCTCCCGGACGCCCA |
| | | R | CCCTGAGGACAGTTGCCTCC |
| SCIN | rs2240572 | F | TCCCGTTTCAGGGCGCCGCCT |
| | | R | CCCTTACCGAGCCAGAAGTGC |
| HECW1 | rs6950443 | F | AGGCGCTGTTGTTGAGCCGAA |
| | | R | CACGTGCCCCGCATAGTGCACC |
| Chr. 7 | rs7789615 | F | AGGTGCGCCCAATCGGAACTCGA |
| | | R | CAGGGGACGCTGCTGTGGTCC |
| CYB5R2 | rs4536203 | F | GCAGCGCTCCATCATGAGGCTG |
| | | R | AACTCGGCGACGCCCCTGCTT |
| ST8SIA1 | rs3803101 | F | TCTTACTTGCAAACGCACGCACGCT |
| | | R | GCCACATGTTTAGGAGGGAGCCGAG |
| TBC1D30 | rs998314 | F | GGCATGATGCCCAGCAGGTG |
| | | In - F | ATACAGGGAATCCTTGTA |
| | | R | CTTCAGGGCCATAACCATACTC |
| SORD | rs3759890 | F | GTAATGATGCACGAATTCATTTACT |
| | | R | TGGGAGCGAGAAGCCCGGGTGC |
| CMTM3 | rs35320143 | F | TGGCCGGGGACAGGAGGGGT |
| | | R | CCTTCTTCTTCACCCTGGTTC |
| C17ORF97 | rs7503725 | F | ACGCTCTGTTCTCGAGATCCCCGA |
| | | R | GTCGTTGTCCTTGTACCCCATTG |
| EMLIN2 | rs7231822 | F | GGATGGCGTTTCTTGGCGAG |
| | | R | ACCCCCGCACCGTCTCTGAACCA |
| CABIN1 | rs915595 | F | TCAGTTGCGCTCCTGGGCCCGCT |
| | | R | GCCAGCAAAGTGGAGACCCAGAC |

Table S5: ***Hpa*II Primers – Placentas-specific DMRs in methyl-seq analysis**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| CACNA1E | rs2609478 | F | ACCAAGCAGTGAGGGGGACTCA |
| | | R | CTCAAGCCGGATCAAAAAGCAAGTC |
| TMEM17 | rs11676567 | F | ACGCTCCAAATCCACGCCA |
| | | R | GGAACCACGGAGCCTCGGAGCG |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| FRMD3 | rs4529536 | F | TCCTGGCTGAGCGATTTGACGCT |
| | | R | ACGCTGCCGATGTCCCGGGGCTGA |
| KCNQ1 | rs34858778 | F | GCTTGGCGGCAGCTGCGGGTG |
| | | R | TCGGAACCCGGGCAACGAAAGGCT |
| TMEM247 | rs11125075 | F | TCCTCAACCCACCTCGGTGACCT |
| | | R | GTAGGTCTCCTTCAACTGTGACTG |
| TET3 | rs1667618 | F | CTTCCGGAGAGATGGCGGGCGAC |
| | | R | GTCCCTCCTCAACAAGATGG |
| SPHKAP | rs74459498 | F | GTCTTAAGATCTCGCTTTGGG |
| | | R | TCACTGCACTTGTAAATAAGT |
| ZNF385D | rs62236195 | F | GGTGGAAATGTCCCCGGCGTGG |
| | | R | TCTGCTTGTAGTACTCTCAGTCCT |
| C3ORF62/ | rs56306491 | F | CGGATTAGGTTGAAGGTCAGAC |
| USP4 | | R | TGTTGATCTGCCCTGTGTCGT |
| EFCC1 | rs1871951 | F | CTACCCCAGGCCCCTCCAAAGC |
| | | R | AGTGGTGCGCCAGGGCGCTC |
| FGF12 | rs3109189 | F | AGTGAATAAAACTTCCTTTAGA |
| | | R | CTCACCGCAGCTGGAAACAGCTGC |
| PDE6B | rs7436737 | F | GAGTACCTGTCCCTGGAGACG |
| | | R | TGCCGCGACGGGCTCTTCCTGCT |
| SH3BP2 | rs228507721 | F | GAAGCCGGCCATGCCCGCCGCGTG |
| | | R | CGTGCCGGCCTGAGTCCCCCATC |
| STX18-AS1 | rs6842024 | F | TGGCGGCGCTGGGAGGAGCGT |
| | | R | GTTGTGGTCGCCGTCGGCGTG |
| GRP78 | rs17844777 | F | CACTCGAGCCTCAGGCGTCCTCC |
| | | R | AGCCCAGCAGCAGGCCGGCATA |
| GRID2 | rs6855766 | F | CCAAGGACTGATTGTCGCGTGCC |
| | | R | TAGGTGGGGCCACCAGCATCT |
| SFRP2 | rs4643790 | F | CGCATGTTCTGGTATTCGATGC |
| | | R | TGTCCTCCCGGTGTCCCGCT |
| DLGAP2 | rs6996211 | F | GCGTCTTCATGGGTCTCATTTG |
| | | R | ATCCGGATTCGGGTCTCGGCGA |
| R3HCC1 | rs11546682 | F | GTTTCTAAGGCCCTTTCGGTG |
| | | R | AAGCGGTCCAGTTCCTCCTGGA |
| DENND3 | rs3739232 | F | TCTGCGGGCGACTGCGCGGCT |
| | | R | CTCGAGACTTCGGAGACTGC |
| CACNA1C | rs11062319 | F | ACTGTTCTCGTACCTGGAGTTA |
| | | R | GAGGCTCGTACAGAATATGTG |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| FGF14 | rs2252594 | F | GGCAGAGGAGGGGGTGCCAGG |
| | | R | AGGGCGAGAAGGCAGCGCGA |
| OPCML | rs72653405 | F | AAATGGTTTATTAGATCACA |
| | | R | CTGAGCGTGGCGTCCGCGCGTCC |
| PAPLN-AS1 | rs2242609 | F | GCCTGCAGAGAGCCCCAGAACG |
| | | R | AGGAGCGGGCATCCCAATGGGGA |
| RYR3 | rs7903891 | F | CAGTTTCTGAGGACTGTGAGTC |
| | | R | TGAAGTTTGGGCATCAAGGGTT |
| Chr 18 | rs57055767 | F | AGAGGACGCCCCAGGAATGA |
| | | R | GAGGGTCTCCGGCTTCTCGG |
| CACNA1A | rs16003 | F | GGGTCTCTCTCCCAGCCTGGAAGAG |
| | | R | GCAGGGCGGGCAGCCCGGGGCGCAAGG |
| ACTL10 | rs3746459 | F | AGTGGCCCGTGCTGGTGAGCGA |
| | | R | GCTGAAACTGGCCGGATGGTG |
| TPTEP1 | rs7285553 | F | AGAGAAGCGGCTAGAAAGCTGAGA |
| | | R | CTGGATTTATACTCAAGGCGCC |
| CACNA1I | rs136853 | F | CAGCCGTCGCTCCAGCTGGAACA |
| | | R | AGTCCCTGTTGTCGAGGGACA |

Table S6: ***Hpa*II Primers – False positives in methyl-seq analysis**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| Chr 1 | rs12061206 | F | GCTCCAGCCCTTTCCTGCCGG |
| | | R | CCCATCCGCTCCAGGAACCCAC |
| THSD7B | rs177727482 | F | GGACTGCAGCACCGGATGGTGCG |
| | | R | AGGACAGGAAATGGGAGCATCA |
| DPP6/ESYT3 | rs13095061 | F | TGGGTGCGGAGAGCGCACCTGT |
| | | R | GGGAGGACGAGGCTGGCATG |
| SLC2A2 | rs57966513 | F | TAGTGAATGGCGGGTCGCGCT |
| | | R | GTCTCCTTGGGGACCGCTAGACTG |
| RPS6KA2 | rs2984 | F | CTGACCAAACCGACCGGCTTC |
| | | R | AGAGTGACCACAAGTCCAGCA |
| OPRM1 | rs1799971 | F | TCCTGGCTACCTCGCACAGCGGT |
| | | R | CACATACATGACCAGGAAGTTTCC |
| RADIL | rs2292489 | F | TCGATCAACGTCATGATCCT |
| | | R | GCTATTCAAGGACCCCGCGCAG |
| NTNG2 | rs3824574 | F | CTGCACCGAGGAGTACTCGC |
| | | R | GACCTCGATGTTGGAGATGG |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| Chr 10 | rs60406779 | F | CACTCTGCGTGTGGAGGGAAGC |
| | | R | AGCTGGAAAACACATCCGAGTTA |
| UNC79 | rs490581 | F | ATACCAGAGGTTAGGTTAA |
| | | R | GCTGAGGACGGAGGGATGG |
| OCA2 | rs6497236 | F | AGATTTCCAACTGTTACGGA |
| | | R | GCTAGGGAAGGGCCCTCCTCCTG |
| FHOD3 | rs9946701 | F | ATGTGTCTGCAAACAAAACA |
| | | R | CTCCGCATAGCTCAGACCCTGC |
| GP6 | rs892089 | F | AGAAGCTGTAGCATCGGTAGG |
| | | R | TTTTGCCAAACCCTCGCTCT |

***BS and corresponding genotyping and pyrosequencing PCRs***

Table S7: **BS Primers – pGEM-T vector**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| pGEM-T vector | | F | GATGGTGCTGCAAGGCGATTAAGTTG |
| | | R | ATGTTGTGTGGAATTGTGTAGCGGA |
| | | Seq(S) T7 | TAATACGACTCACTATAGGG |

Table S8: **BS Primers – known ubiquitous DMRs**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| NAP1L5 | | F | GGGGTTTTTTAGTTATTTGATTAG |
| | | R | CAAAATCTCTCTAAACCAACTCT |
| FAM50B | | F | AGAGTGTTAGGTTTTTGGTAGGG |
| | | Bio - R | [btn]TCTCTAAATAACCACAACAACTTAC |
| | | Seq(P) - F | GGGAGGATTTTGGAGGAG |
| PLAGL1 | | F | GTGTTTAAATTAAGGTTYGGGG |
| | | Bio - R | [btn]ACCTTAACTTTACCCCCACCRATAA |
| | | Seq(P) - F | GTGGTAGGAGGAGGTTT |
| GRB10 | | F | GGGTTTTYGTGGGTATAGTTATTATT |
| | | Bio - R | [btn]CCCRCTCTCCAAATACTCAAATAAACT |
| | | Seq(P) - F | ATGGTTATATAATATTGTTTTATGG |
| PEG10 | | F | GTTTGGTTTAGGTGTGGGATTTT |
| | | Bio - R | [btn]AAACATTCTAAAATACTACTCCAT CTCCC |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | Seq(P) - F | TGGGATTTTATTTTTTTTGTT |
| MEST | | F | TYGTTGTTGGTTAGTTTTGTAYGGTT |
| | | Bio - R | [btn]AAAATAACACCCCCTCCTCAAAT |
| | | Seq(P) - F | GAGTTGGGGTTGTTTTTG |
| INPP5F | | F | TTTGTAAAAAGGAGAAAAATTTTGAAAGTTA |
| | | Bio - R | [btn]AAATCTATACCCTCRTTAATAAAC |
| | | Seq(P) - F | AAATAGTTGTTTGTATTTA |
| PEG13 | | F | AGTGTTTTAATTTATAGGGTATTGA |
| | | Bio - R | [btn]CACAACCCTACCAACAAAAAAAACC |
| | | Seq(P) - F | AGGTTTTTTATTGTTGTATGTA |
| H19 | *promoter* | F | GTTTGTTAGTAGAGTGYGTTYGYGAGTYG |
| | | Bio - R | [btn]ATAACACAAAAAACCCCTTCCTACCAC CATCAC |
| | | Seq(P) - F | GTTTTTAGATAGGAAAGTGG |
| H19 | *Gene body* | Bio - F | [btn]TTTTTTTTTTGAGAGATTTA |
| | | Bio - R | CCRAACCCTACACAAACACTTACC |
| | | Seq(P) - R | CTTACCAAAATAACTCACACTC |
| H19 | *CTCF* | F | AGGTGTTTTAGTTTTATGGATGATGG |
| | (BS) | Out - F | TTGATTTATTTTAGGGTGTATTGTTGAAG |
| | | In - F | AATAATGAGGTGTTTTAGTTTTATGGATG |
| | | R | CCATAAATATCCTATTCCCAAATAACC |
| | | Bio - R | [btn]TCCTATAAATATCCTATTCCCAAATAACC |
| | | Seq(P) - F | TAGTGTAGGTTTATATATTA |
| H19 | *CTCF* | F | TGTCGGTCGGAGCTTCCAGACT |
| | rs2839703/ | Bio - R | [btn]TTCCCCCGTCCCTTC |
| | rs2839704 | Seq(P) - F | CAAGGTGGCTCACATACGCAC |
| NHP2L1 | | F | GGAGATAATGGTTTAGTAGGAATATTTATT |
| | | Bio - R | [btn]TCTTCCTAATTATCCTAAAACCAAAAC |
| | | Seq(P) - F | TTTAAGAGTAGGAAG |
| KvDMR | | F | TGTTTTTGTAGTTTATATGGAAGGG |
| | | Bio - R | [btn]AAACATACCAAACCACCCACCTAACAAAA |
| | | Seq(P) - F | ATGGTAATGTTTGGTATTTAGAA |
| RB1 | | F | GGTAGGGTAGTTTTGGAAATGTTTAAG |
| | | Bio - R | [btn]AAACCACAAACCCTTACCC |
| | | Seq(P) - F | AGTTTTGGAAATGTTTAAG |
| IG-DMR | | Bio - F | [btn]ATTATTGAATTGGGTTTGTTAGTAG |
| | | R | CAAAACAACTCAAATCCTTTATAAC |
| | | Seq(P) - R | TTAAAATATATCAAAAAACC |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| SNURF | | F | TAGGTTGTTTTTTGAGAGAAGTTAT |
| | | Bio - R | [btn]AAAAAAAAAACTAAAACCCCTACACTAC |
| | | Seq(P) - F | GGTATAGTTGATTTTG |
| IGF1R | | F | GAAGTTTTTTATATGTAATGGGAAG |
| | | Bio - R | [btn]AAAAAACTCCCAAAATAAATACA |
| | | Seq(P) - F | TGAGATTTGGGTTTTGAGTT |
| PEG3 | | F | GGAGAAGTTTTGATAAGGAGG |
| | | Bio - R | [btn]CTCACTCACCTCAATACTAC |
| | | Seq(P) - F | TGTTTATTTTGGGTTGGT |
| MCTS2 | | Bio - F | [btn]GATTTTTATTGGAGAGGAATTAGTAG |
| | | In - F | GGAAGTTGTTAGATTTTA |
| | | Out - R | CTAAATACAATTAAACACACTTTCC |
| | | In - R | CATCAAACTTCTTAAACATAATCC |
| | | Seq(P) - R | TAAAATCTAACAACTTCC |
| L3MBTL1 | | F | TTTTAGGTTTTGAGTTGGGTTTT |
| | | Bio - R | [btn]CCCCCTCAAACCCTACCTCC |
| | | Seq(P) - F | AGGTTTGAGTTTTAGAAGGTT |

Table S9: **BS Primers – known placenta-specific DMRs**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| GPR1-AS | rs16838074 | F | GTTTAAGTTTGTGTTTATAAAGGG |
| | | Bio - R | [Btn]AAAAAAAAACTCTACTCCTCTCAA TATCTC |
| | | Seq(P) - F | GTTTAAGTTTGTGTTTATAAAGGG |
| MCCC1 | rs937652 | Bio - F | [Btn]GATTTTTTAGTATAGAGGTAGT |
| | | R | CTCTAATCTTTAATTAATATTC |
| | | Seq(P) - R | CAACCCAAAAATAAACTCAAACTC |
| LIN28B | rs9404590 | Out - F | TATTATTTTGGGGAAAGTGTGTTAAGGT |
| | | In - F | ATATTTTGAAGTGTTTTTGTTGTAA |
| | | In - R | CTAAACTATCTTCCAACAAACC |
| | | Bio- Out - R | [Btn]CATACCTAAAACTAAACTATCTTCC |
| | | Seq(P) - F | GTTTTGTATTTGTTTTTATGT |
| ZC3H12C | | Bio - F | [btn]TGGCGGTTTTTAGGTTTGTTTT |
| | | R | AAAAAATACGCTAAACTTTCC |
| | | Seq(P) - R | CCTAAACCTAAAAATCACAC |
| AIM1 | rs11152999 | F | TTTGAGTAAGATTGAGGGGGTATAA |
| | | Bio - R | [Btn]CCCTCTTAAATAAAAACTCCTTAAC |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | Seq(P) - F | GATATGGAGAAGAGGTTTAG |
| AGBL3 | | F | GATTGGTTTGTTAGAGGATT |
| | | Bio - R | [btn]CACTCACTCCTAAACACCTTA |
| | | Seq(P) - F | GAGGTATGTAGAAGTAAGG |
| ZFAT | rs12546315 | F | AAGGTTTTTATTTTTTAAAGTTGGAA |
| | | Bio - R | [Btn]CATAACCAAAAACCCTACTAAAAAC |
| | | Seq(P) - F | TTTAGATTTAGGGTAAGGGGGT |
| GLIS3 | rs7852293 | Bio - F | [Btn]TTGGATGGGGGTTTTTATTTT |
| | | R | CTCTAAACTTTAACTCCCCCACAC |
| | | Seq(P) - R | CTATAACATCCACTTACCAAATAAC |
| DNMT1 | rs8112895 | pyro Out - F | GAGATGTATAGTTTTGGGGGAAAGG |
| | | In - F | GGTATTGAGGGATTTTTTATTATAGAAG |
| | | In - R | TCTTCAAACTTCGATAAAC |
| | | Bio Out - R | [btn]AAATCCCCTTCTTCAAACAA |
| | | Seq(P) - F | GAGGGATTTTTTATTATAGAAG |
| C19MC | rs6509806 | F | GGTTGTTTATGTATTTTTTTA |
| | | Bio - R | [Btn]CAAATTCTAATCCCTCAAAAAAAAACC |
| | | Seq(P) - F | GTTTTTTGAAGTTTTTT |
| DCAF10 | | F | AATTTTCGTATTATGATTAGTTTTTA |
| | | Bio - R | [Btn]CTAAACTAAAAATTCCAAAAAC |
| | | Seq(P) - F | GGATTTGGTGTATTTTAGT |
| ZNF396 | | F | GTTTTTTTAGGGTTTTTAGG |
| | | Bio - R | [Btn]AATTCACCTAACAAAACAAA |
| | | Seq(P) - F | GGGTGGTGGAAGGGGGA |
| FAM174B | | Out - F | GTTAGTGTTTAGTATTAGTAGAAAG |
| | | pyro In - F | GTAGGGTAGGAGGTATATG |
| | | Bio - Out - R | [Btn]TTTACCACCCTCCTCAT |
| | | Seq(P) - F | TTAGGGTATTTGTGCGTGT |
| N4BP2L1 | | F | TAATTTTTAGAGGTTGGGTAT |
| | | Bio - R | [Btn]CTAACAACCCRCTTTAAAAATC |
| | | Seq(P) - F | GTAGTTGTTAAAGTAA |

Table S10: **BS Primers – repeat elements**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| LINE1 | | F | TTTTGAGTTAGGTGTGGGATATA |
| | | Bio - R | [btn]AAAATCAAAAAATTCCCT |
| | | Seq(P) - F | AGTTAGGTGTGGGATATAGT |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| ALUYb8 | | Bio - F | [btn]AGATTATTTTGGTTAATAAG |
| | | R | AACTACYAACTACAATAAC |
| | | Seq(P) - R | AATAACTAAATTACAAAC |
| a-SAT | | F | AGTTTATTTATAGAGTAGAGTAG |
| | | Bio - R | [btn]AAATCTTCACTTACAAATACCAC |
| | | Seq(P) - F | TGGGATTTTTTTGAGAATTT |

Table S11: **BS Primers – Placentas-specific DMRs in NLRP7 HM450k array analysis**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| TTC39A | rs74080536 | F | TAGTGAGTGTTAAATYGTGTTT |
| | | R | CTCCAACAAATAAAAACCC |
| THAP3 | rs12746591 | F | GTTATTTTTGTTGGGGGTAGTTA |
| | | Bio - R | [Btn]CCCTCTAAACCCAAAACCTC |
| | | Seq(P) - F | TAGGAAGTAGTTTATTTT |
| EGR4 | rs6718289 | F | GTGAGTAAAAGAAATGAATGG |
| | | R | ATACCCAAACATAAACACAAAATA |
| RPN1 | rs12487604 | F | GGTTTTTAATTTTTAGGATAGG |
| | | R | AATCCCTAAATAAAAATCCCAAA |
| RHOBTB3 | | F | AGAGTAGTTTGTTGTTGAA |
| | | Bio - R | [Btn]ATTTAAAAATTTAACCCATATA |
| | | Seq(P) - F | GTTTGGGAATGTTAAGTTGG |
| PURA | | F | GTTTTATTTTTTTTATGTTAGTGGTYG |
| | | R | TAAAACCCAAACCAAACCCCCR |
| SCIN | rs2240572 | pyro Out - F | TTGTAGGTTTGGAGGATTGAGAAGTT |
| | | In - F | TTGAGAAGTTGGAGTTGGTGT |
| | | Bio - Out - R | [Btn]CCTACAAATCTCAAACCCCTAC |
| | | Seq(P) - F | GGATGTTTATTTGGTGTTGT |
| SNCB | rs11951438 | F | GTTTTATTTGGATGYGGGG |
| | | R | AAAACRCTAAACTCAAAACTA |
| THSD7A | rs113779084 | F | TAGAAAGTAAAGTTTTTTTTGT |
| | | R | CCTAACAAACCTACTCAATCAAC |
| CCDC71L | | F | TAATTTTTTTTTGTGGATAGAAGT |
| | | R | CTATAACCAAAACCAATTCTTCA |
| ARMC3 | rs137988681 | Out - F | GAGGTATAATAAGGGATTGGG |
| | | In - F | TAATAAGGGATTGGGGGTT |
| | | In - R | TACCTCACTAACCCAAACAATCT |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | Out - R | CCCAACAAAATATTACCTCACTAACC |
| AIFM2 | rs10999151 | F | GGTTTAAGTTTATTTTTTTTAG |
| | | In - R | AAACCCRAAAAAAAAAC |
| | | Out - R | AACRAATTAAAAACCTAACTCA |
| FGF8 | | Out - F | TAGTTATTTGTTGTATGGTTAG |
| | | In - F | ATATATAGTAGGTGGTTYGGAGATTA |
| | | In - R | CTCCCCCTATCCCCGACCTATCCCCTC |
| | | Out - R | AAATTAAAAATCTAAACTAAAA |
| CYB5R2 | rs4536203 | F | GAGTTTAAGTAAGTTGAGGGG |
| | | R | AAAACTCTCACTCACRAA |
| RNF141 | rs1979367 | F | GAGTAGGGGTTTTATGAGGTGG |
| | | R | AACCTTTATATATCTATCATC |
| WIF1 | | F | AGAAGTTTATAAGTTTTTAGATG |
| | | R | CTCTAAAACATCCTCCTATACC |
| ST8SIA1 | rs3803101 | pyro Out - F | TAGTTTTTTATGTATATATATTTTTGGTT |
| | | In - F | ATTTTTGGTTTTTTTATTTGTAAA |
| | | Bio - Out - R | [Btn]CCACATATTTAAAAAAAAACC |
| | | Seq(P) - F | GATTTATGGTTATGGT |
| RASGRF1 | rs9920235 | F | GGGGCGAGTTCGTTGTTTTTTG |
| | | R | ACACCCTAAAATCTAATTAAC |
| SIAH1 | rs12928880 | F | GGGGTGGGATGGGGTTAAYGGYG |
| | | R | GGTTTTTTCGGTAGGTAGGTTTTTG |
| CMTM3 | | F | GGATAGGAGGGGTGGTTAAGAAAG |
| | | Bio - R | [Btn]TCCACCTTCTTCTTCACCCTAAT |
| | | Seq(P) - F | GATAGTTTTTTTGGATAGGG |
| CABIN1 | rs915595 | F | GTAGTGAGAGTTGTATTTATTTT |
| | | R | AACATCTATACATAAAACTAAA |

Table S12: **BS Primers – Placentas-specific DMRs in methyl-seq analysis**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| GRID2 | rs6855766 | Out - F | TAGGTTGTTTTYGAGGTTTGT |
| | | In - F | AGGTTTGTTTAGTAGATATT |
| | | In - R | AATAAAATTAAAACCTAAAACTA |
| | | Out - R | AAAAACTACAAAAACCTATCTCCC |
| | rs7694499 | Out - F | TGTATTATAAAGTTAGTGTGGAGTT |
| | | In - F | TTGTTTGGTTGGTTGGTTGGTTGT |
| | | In - R | TCCATACCAATATAAAAAAACAAT |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | Out - R | TTCAAAATATATAAAAAACAAAACT |
| R3HCC1 | rs4265198 | Out - F | GGTTAATAATTTTTATTTGAGAGGG |
| | | In - F | GTGTAATTTGGTTTTTTTTGTTTG |
| | | In - R | CTCATACCTATAATCCCAACACT |
| | | Out - R | ACTTACTAAACAAATTACACTTTA |
| | rs11546682 | Out - F | TTTGATTTGTYGGGGGTTTTTGGT |
| | | In - F | TGGATTAGAGGGTTTTTAAGGTTT |
| | | In - R | CAACTACTTCTACAACAAAAAAC |
| | | Out - R | CCCCTCACATAAACCTAC |
| TET3 | rs1667618/ | F | TTTTTYGGTAAGATTTTTYGT |
| | rs1723289 | Bio - R | [Btn]CTCCTCCTTCCRATCCCTCCTCAAC |
| | | Seq(P) - F | GAATGGTCGGTTATTT |
| ZHX3 | rs62208470 | pyro Out - F | GTTATTATTTTTCGGGTTTTAGG |
| | | Seq (P) - In - F | GGGGTTTTAGGGGTTAGGGTAG |
| | | Bio In - R | [Btn]TCCCAACAAAATACAAAAACCCT |
| | | Out - R | AAACATAAACTACCACATCCCAACA |
| TMEM247 | rs11125075 | Out - F | TATATTTAGGTTTTGAAGGAGTAT |
| | | In - F | TGAAGGAGTATYGATTATATTT |
| | | In - R | CTCAATAAATCTCCTTCAACTATAAC |
| | | Out - R | CCTCCAAAACTTTACTCACCTTC |
| FGF14 | rs2252594 | Out - F | GGGGTAGAGGAGGGGGTGTTAGG |
| | | In - F | GGGATTGGGGAGAGGGGAAGGG |
| | | In - R | AAAACCCRACACTAAATCCTACCCA |
| | | Out - R | ACTAATTTCCAAAACACTTAACAATCA |
| FGF12 | rs3109189 | Out - F | TTTTTTTTGTAGTGTTGTTGGT |
| | | In - F | TAGTTGGAAATAGTTGTTTT |
| | | In - R | CCCAACAAAATAAAATCTAC |
| | | Out - R | AAAAAAACAACAATTTAACTCCCTAC |
| FGF8 | | Out - F | ATATATAGTAGGTGGTTYGGAGATTA |
| | | In - F (meth) | GAGATTATTGCGTAGTCGGCG |
| | | In - F (unmeth) | GAGATTATTGTGTAGTTGGTG |
| | | In - R | ACRCCTAAAACTACAAACTCTA |
| | | Out - R | CTCCCCCTTACCCCRACCTATCCCCTC |
| CACNA1C | | Out - F | AGGGTTTTGGTTTTAGGGAGATA |
| | | In - F | TGTAAATGGTTTTTATAAGTAGAGT |
| | | In - R | CTTCATTCATTTCTATTAAAACC |
| | | Out - R | CTATTCTCATAACCTAAAATTAACC |
| CACNA1A | | Out - F | GTAAAGGATGTATAAGTAGTTAATG |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | In - F | GGATTATGGTATTTTATAATTTTA |
| | | In - R | AATCTCTCTCCCAACCTAAAAAAAA |
| | | Out - R | ACCCCCCTCCTCCCCACCTCCCATCAA |
| CACNA1AI | | Out - F | GTTGGAATAGTTTTAAGTATAAG |
| | | In - F | GAGTATGAGTTTTTGTTTTTTG |
| | | R | ACCAAATCCACCRAATCCCTATTA |
| AGO1. | | pyro Out - F | GTTTTGTTTTATTTATATTATTG |
| | | In - F | GGTTTTAAGGTATTTTTATTGG |
| | | In - R | AAAACCTAAATACCCCACACTCT |
| | | Bio Out - R | [Btn]TTCCCCCTCCTAAAAACCCTT |
| | | Seq(P) - F | TTTTTTTATTGGTTTTTGT |
| TMEM247 | rs11125075 | Out - F | TATATTTAGGTTTTGAAGGAGTAT |
| | | In - F | TGAAGGAGTATYGATTATATTT |
| | | In - R | CTCAATAAATCTCCTTCAACTATAAC |
| | | Out - R | CCTCCAAAACTTTACTCACCTTC |
| GPR78 | rs17844777 | Out - F | TAGGTATAGTTTAGTAGTAGGT |
| | | In - F | TAGTGGGAAGTTTATTGTTAGT |
| | | In - R | CACTAATACTACTTTATTAC |
| | | Out - R | CTACTATCCAACRCACTAATAC |
| JMJD1C | rs1054693 | Out - F | TGGGTTTTGTTTTAGTTTTTTTGTGT |
| | | pyro In - F | TTTGGGAGTTTAAAGGAGTGTTT |
| | | In - R | CTTACCCACCAACTCTACCC |
| | | Bio Out - R | [Btn]CTATCCCTATATAACACRACTC |
| | | Seq(P) - F | GGGGTAGAGCGGTCGGT |
| SH3BP2 | rs28562583 | F (unmeth) | TATTGAGTTTTGGTTTTTGAGTTTT |
| | | F (meth) | TATCGAGTTTCGGTTTTCGAGTTT |
| | | In - R | CCRACCTAAATCCCCCATCC |
| | | Out - R | [btn]CTACCCRCAAATCRACAAACCC |
| | | Seq(P) - F | GAGTTTTTGGATTAGGG |
| STX18-AS1 | rs6842024 | Out - F | GGTTGYGTTTTTTTTGAAGTG |
| | | In - F | GTGGGTTTTGGGATTAAGTGTTAG |
| | | In - R | AAAACTCTTAAAAACTCCRAAAA |
| | | Out - R | AACCCAAATCACACTATCACCCA |
| FAM149A | rs907438 | pyro Out - F | TYGTTTGAGTTGGGTTAGT |
| | | In - F | GGGTGGGGAGTTTTAGTTTYGG |
| | | bio In - R | [btn]AATACCTACAAAAAAACCC |
| | | Out - R | CTAACCCAACACCAACAAATCCC |
| | | Seq(P) - F | GGGGGTTGCGAGTATAG |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| LAMA2 | rs80287289 | Out - F | GATTTTGTAGTTTTGGAGTTTTAG |
| | | In - F | GTTTAAATATGGAATAGTGATG |
| | | In - R | TACACATCCCATACACTCAACT |
| | | Out - R | ATAAATCCAAATAAAACTATAAA |
| MOCS1 | rs575925610 | pyro Out - F | GGTAAGGTTGAAAGGTATTAAG |
| | | In - F | GGATTTAAAGTTTTAGTTTTTTG |
| | | Bio - In - R | [Btn]TACCCTTTTCCCCCAATAACT |
| | | Out - R | AAACAAAAATAACCCCRACA |
| | | Seq(P) - F | GGGATTTTTAGGAATTTT |
| DLGAP2 | rs6996211 | Out - F | TATGGGTTTTATTTGTTTTGGT |
| | | In - F | AGTAAGATTTTGTGTTGGAGAAA |
| | | In - R | CAAARCACACAAAACTCTAC |
| | | Out - R | CCTTATRGAACAAAAAACCC |
| KCNQ1 | rs34858778 | Out - F | TGGTTATTATTAAAAATAGATGT |
| | | In - F | TTTTYGTTAGGTGTGGATTT |
| | | In - R | CCATCTACCTCTAAAACACTCR |
| | | Out - R | TATTTAATCCACAAAAATTTCC |
| PAK1 | | pyro - F | TAAGGTAAGAGGAAAGTAAGATAT |
| | | Bio In - R | [Btn]CAAAAACCTTATAAATACCC |
| | | Out - R | CTAAACTACCAAAAACCCAACC |
| | | Seq(P) - F | TATTATTTGGTGGGGAAGGTTAGT |
| ACTL10 | rs3746459 | Out - F | GGTAGGGTTTGTTGTAAGAGGTTAG |
| | | In - F | GTTTAGTTGGAAGGTGGTTTGGTG |
| | | In - R | AACTCTAACTTAATACCCAACTA |
| | | Out - R | ATAAACCAAAACTCRAACTTCACCA |
| OPCML1 | rs72653405 | Out - F | GTTAGGGATGGAGTTGTTGTTATG |
| | | Seq(P) - In - F | GGGGTTGTTTTYGGGAGGAAGG |
| | | In - R | ATTTATTAAATCACACACATAAACAA |
| | | Bio Out - R | [Btn]ATACAAAAACTATCATAAAAAACTA |
| ZNF385D | rs62236195 | pyro Out - F | GTTGTTAATTTAGGGTGGGGTG |
| | | In - F | GTATTTTTAGTTTTGATGTTATAGG |
| | | Bio In - R | [Btn]AAACAAAACCCTTTCATACTACA |
| | | Out - R | ACATCAACTCTCACCCAAAACTAA |
| | | Seq(P) - F | TTTTAGTTTTGGAAAGGTAGTTT |
| BOD1L2 | rs11151997 | Out - F | TTTTTAGTTTATCGTTATTAT |
| | | In - R | CTATCAAACATATTCTTAAAAAATATCC |

Table S13: **BS Primers – New ubiquitous DMRs in methyl-seq analysis**

| Gene | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| SVOPL | rs180949672 | Out - F | AAGAGTTGGGGGTGAGTAGTAGTA |
| | | In - F | GTATATTTATGGAGGATTAGG |
| | | R | CCCTAATCCTCCACTAAACAAACCC |
| FANCC | | Out - F | TGGGGAAGATATATATTATAGAAGT |
| | | In - F | TAGAAGTATATATGTTAAGGATT |
| | | In - R | AAACTCCCATAATATATA |
| | | Out - R | AACACTAATTCCCAAACAC |

Table S14: **BS Primers – False positives in methyl-seq analysis**

| Gene | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| KIF26B | rs116793850 | Out - F | AGGGGTAGGGGTGAGGGAATGGA |
| | | In - F | AAGAGTTAGGTTGTTTTTAGA |
| | | In - R | CAAACCCCACTACCAAAATATCC |
| | | Out - R | CCTCAAAACCCCGAAAAAAC |
| LOC401312 | rs2905322 | Out - F | GTTTTTTTTTGGTTGGTTTTTTGG |
| | | In - F | ATAGTTTGTGAAGTATTGTTTGA |
| | | In - R | AAATAAATAAAACTCCTTCCTTAA |
| | | Out - R | ATTTCTCTAACTCTTCTAAAAAA |
| NPAS3 | rs10141940 | Out - F | TGTAATTTTTGTAGCGTTTTTAGGT |
| | | In - F | AGTGTATAGTAAGGAGTTGGA |
| | | In - R | CCCCAACTCTAAACACTTCC |
| | | Out - R | CCCTACAAACATAACCAAAC |
| PROSER2-AS1 | rs10906034 | F | TTATTAGATTGGGGTAGTTTGT |
| | | R | CAACAAAAAAAACAAAACAC |
| ZNF154 | rs528375562 | Out - F | GGGTGGGGGAAGGTGTGTGAGG |
| | | In - F | GTTATTAGATTTTGCGGGTAGAGTTG |
| | | In - R | ACTTTCCTTAAAAATCACAAAAA |
| | | Out - R | ACTAAACCAATAAAAACTCAAAA |
| PTCHD3 | rs7071851 | F | GAGGAATAGTAGGTGAAAAGTAGTT |
| | | R | CAAACACATTTCAAATTAATTAAAC |

**RT-PCR and qRT-PCR Primers**

Table S15: **RT Primers – Known ubiquitous and placenta-specific DMRs**

| Gene | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| RPL19 | Reference gene | qRT - F | GCGGAAGGGTACAGCCAAT |
| | | qRT - R | GCAGCCGGCGCAAA |
| H19 | | qRT - F | AACCCACAACATGAAAGAAATGG |
| | | qRT - R | AGAGGGTTTTGTGTCCGGATT |
| DNMT1 | | qRT - F | AGATGGAGACGAGAAAGATGA |
| | | qRT - R | ACTGAATGCACTTGGGAGGGTGGG |
| AGBL3 | | qRT - F | TGTTAAAAGTGCACACCGATGAAGAT |
| | | qRT - R | CTGGAAAAGATGTCAGAAGAATC |
| NAP1L5 | rs710834 | Geno/RT - F | GGTGAGCTCTTGGATCTTGG |
| | | Geno/RT - R | GCGGCTTCTCCTCTAACATG |
| HYMAI | rs2281476 | Geno/RT - F | GCCTACGTGCGGGTCCGGG |
| | | Geno/RT - R | GTTGGCGAGGTTAGAGCGGCC |
| PEG10 | rs13073 | Geno/RT - F | ACAGAGATGTAAGAGGCAGGC |
| | | Geno/RT - R | CTAGTCACCACTTCAAAACACAC |
| PEG3 | rs4801386 | Geno/RT - F | ACAATCATTCTCTTGTTTACCA |
| | | Geno/RT - R | CATAGGAAGGACTGAGGTTGGAAC |
| MCCC1 | rs937652 | Geno/RT - F | GTTCCTCTCCGCCGCCACCAG |
| | | Geno/RT - R | TAGGCTCAGGCTCCGACGGT |
| LIN28B | rs221634 | Geno/RT - F | GTATTGGTCCTGTTAGGTTTCGG |
| | | Geno/RT - R | TGTCTCATTTGAGTCATGCTATT |
| GLIS3 | rs7852293 | Geno/RT - F | ATGGGGGTTTTCATTTCCGA |
| | | Geno/RT - R | CCGGGATTTCCACAACAAAGCC |
| ZNF396 | rs2909339 | Geno - F | AGGCAGGAAAATTGCTTGAA |
| | | Geno - R | GGACTGCAGTGGACCCTTTA |
| | | RT - F | TGGGTGCTTCTGTAAGATGG |
| | | RT - R | TGAAAAGGTTGTCTTGTGATGC |
| DNMT1 | rs2228612 | Geno - F | GTGCGAGTTGGCGATGTG |
| | | Geno - R | CAGTCTTCTTTTTCCTAAGACCAG |
| | | RT - F | CGAGGACGAAGATGGAGACG |
| | | RT - R | CTGAATGCACTTGGGAGGGT |
| AIM1 | rs11152999 | Geno - F | GCAGAAATCCACCGACTCC |
| | | Geno - R | CTTGGGTGGGGACTCCTT |
| | | RT - F | GGTGTTCGACGACGAGGT |
| | | RT - R | GAACTGGCTGGCTAAATGGA |
| GRP1-AS | rs34523400 | Geno - F | TGCATCTGAGAATTTTGTTCCTTT |

| Gene | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | RT - F | TGCCTGGCCACTTGGAAAAAG |
| | | Geno/RT - R | CCCTGACCCCTGGAAAATTTTG |
| AGBL3 | rs2348049 | Geno - F | CACATTGAATAGTGTTGAAAGGAAA |
| | | Geno - R | ACCGTGTTGGGCTCAATG |
| | | RT - F | TCAAACACTGGAAAAGATGTCAG |
| | | RT - R | GGGCTCAATGGACCAGAAG |
| ZFAT | rs3739423 | Geno/RT - F | CTCATCAAGCACATCCGAGA |
| | | Geno/RT - R | GGAGCTGACGAACTTCTTGC |
| ZC3H12C | rs10266807 | Geno - F | AAGAAGTGTACCTATGGACACA |
| | | Geno - R | ACCCTCTTGGCAGACATGGCCCCA |
| | | Geno/RT - R | GGTGGGAAGCTTTTCTTCTAGG |

Table S16: **RT Primers – *NLRP7* HM450k array analysis (new studied genes)**

| Gene | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| RHOBTB3 | rs34896 | Geno - F | ATTTGTTTTCTCAGAAGAGTTA |
| | | RT - F | ACGAGAGCAGCTTGTTGCTGAA |
| | | Geno/RT - R | CTTTCCTATGTAGAAGTCATC |
| CD83 | rs111804875 | RT - F | GGGCATAAAAGGGCAGCCGGCG |
| | | RT - R | TCGAAAGAACCATTTTGCCCCT |
| SCIN | rs2240572 | Geno/RT - F | CGGGAGCTATACCACGAAGA |
| | | Geno - R | AGTTTGCGGTCCCCTTTACT |
| | | RT - R | GAACAGTGAAGATGGCAGCA |
| CCD71L | rs2190093 | Geno/RT - F | GGTGGAGGCATTTCTTGGTA |
| | | Geno/RT - R | CCTGTGGTCTTCAGCTGTCA |
| HECW1 | rs6950443 | RT - F | AGGCGCTGTTGTTGGAGCCGGAA |
| | | RT - Revers | GAGTTTCTAGAAGGAGACGCCATGG |
| AIFM2 | rs1053495 | Geno - F | GGTTCAGTGAGGTTCTGGTTG |
| | | Geno - R | CTGGCCTCTCCATCAAGGT |
| | | RT - F | CCAACCTGGTGATTCTCTGC |
| | | RT - R | CCATCAAGGTGGAGACTGCC |
| CYB5R2 | rs13172 | Geno - F | GAGCACCTTCCTCCTCCAG |
| | | Geno - R | GGTACATGGCAAGGCACTTT |
| | | RT - F | TGGTCAGAAAAGAGCTTGAAGA |
| | | RT - R | CACCCAAAAGAGGAGAACCA |
| ST8SIA1 | rs3803101 | RT - F | TCTTACTTGCAAACGCACGCACGCT |
| | | RT - R | GCCACATGTTTAGGAGGGAGCCGAG |
| RASGRF1 | rs1562008 | Geno - F | TCTGTCCATCACCAAGACATCGT |

| Gene | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | RT - F | TGGATAACATCCGATGCAATGGGCT |
| | | Geno/RT - R | TGCTTGCTGAGCGCTGAAGGGT |
| CMTM3 | rs3743718 | Geno - F | CGAACCAGGGTGAAGAAGAA |
| | | Geno - R | CAAAGTTTGGGGACCTCGAC |
| | | RT - F | CGGGGTCTTTGCTGTTTAGA |
| | | RT - R | GCAGAGTCCCCTTGTTTGAG |
| ZFP90 | rs1170444 | Geno - F | CAGGCGCACTTCCAGTTCT |
| | | Geno - R | ACCCGCGCACTCAGAATC |
| | | RT - F | AGGCGCACTTCCAGTTCT |
| | | RT - R | AGGTGGCTATAGTTCTCCAGC |
| CABIN1 | rs17854874 | Geno - F | GAAGGCTTACCCTGCTGTGA |
| | | Geno - R | GGACACTTCTGCCTCTGTCC |
| | | RT - F | GAGTGTTCCGATGTGGCTCT |
| | | RT - R | CAGGACCTTCTGCCCAAATA |
| TBC1D30 | rs939875 | Geno - F | TGCCAAGAATGCTGTCATCCACA |
| | | RT - F | AGCAGCAGGTTCATCAGGTGTACA |
| | | Geno/RT - R | CTGCCCCCAGCCCATTGGTCTTGC |

Table S17: **RT Primers – methyl-seq analysis analysis (new studied genes)**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| SVOLP | rs2305816 | Geno - F | TCTTGAAATTGATGACAACGT |
| | | RT - F | GGACCACATTACAGATCTGGG |
| | | Geno/RT - R | CTTGAAGTGCAAATGTTGAGGAG |
| AGO1. | rs2296470 | Geno - F | ACACTCCTAGTCTAATTCCTACA |
| | | RT - F | ATGGAAGCGGGACCCTCGGGAGCA |
| | | Geno/RT - R | GATCTCCTTGGTGAAGCGAACG |
| C3ORF62/ USP4 | rs56306491 | Geno/RT - F | GATGTTCAATATACCCAGTACCCATG |
| | | Geno/RT - R | TCTTATTAGCTCAGGAGGTGGGCACT |
| SH3BP2 | rs231402 | Geno - F | GCATTGGCCTGTCCCTATGAAGG |
| | | Geno - R | ATCCAGAGGGCACACCATGTGA |
| | | RT - F | CGTGGATCGCCCCGGGGAAGC |
| | | RT - R | ACGCAGCGTTTGTGGATGATGACA |
| FAM149A | rs48562650 | Geno - F | TGGAAGAACAGTCTGGCCTAAACT |
| | | RT - F | ACAGGCCCGGGGCTGATCGTA |
| | | Geno/RT - R | GTGCTGCCCTGGAAATGCTGG |
| MOCS1 | rs3793137 | Geno/RT - F | CCCGACTTCTACCAGGGATGCC |
| | | Geno/RT - R | TCATAGGAAGAAGGCACAAGT |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| R3HCC1 | rs11546682 | RT - F | GACGCCGAGGGCGGCTGCGACG |
| | | Geno/RT - R | AAGCGGTCCAGTTCCTCCTGGA |
| JMJD1C | rs1054693 | Geno/RT - F | GGGAGGGCCCGAGGTCGCTG |
| | | Geno - R | TACCCACCAGCTCTGCCCGCGT |
| | | RT - R | AGTGGTATTCCACCAAGAAAGTTGA |
| PAK1 | rs2844337 | Geno - F | GGAACAGCTTGGGAATGCCATGG |
| | | RT - F | GAGAGGTTCAGCTAAAGAGCTG |
| | | Geno/RT - R | AGCCGAGAGCATTTCACAAGAAACA |
| OPCML | rs72653405 | RT - F | GAGTCGCCGACCGGGCTGCAGAG |
| | | RT - R | ACCCGGGTTACCCGGTCATCTA |
| PAPLN-AS1 | rs2242609 | RT - F | GGGCACAGAAGGAGTAGTGAGGTG |
| | | Geno/RT - R | AGGAGCGGGCATCCCAATGGGGA |
| TET3 | rs1667618 | Geno/RT - F | CTTCCGGAGAGATGGCGGGCGAC |
| | | Geno/RT - R | GTCCCTCCTCAACAAGATGG |
| | rs7560668 | Geno - F | TGCCTATTAACTTTCCTGTTGGT |
| | | RT - F | AGGCCACACCCACCAAGGCTGA |
| | | Geno/RT - R | GAGTTCCCGGATAGAGGCGACCG |
| ZHX3 | rs17265513 | Geno - F | AGTTCCAGTCAGCCAGGCATCTGCCA |
| | | RT - F | ACCGCCGCCGCTGCCAGGACTGAAA |
| | | Geno/RT - R | GGAACTTGTGGAAGGAGTTCTTCAG |

Table S18: **RT Primers – flanking genes (and corresponding genotyping primers)**

| Gene | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| ZNF24 | rs2032729 | Geno - F | TGATGGTAGGACTGAAAATGGA |
| | | Geno - R | TGAGCCCTGACTGAAGTGTTT |
| | | RT - F | GCTCAGGGATTACCAAGTTCTG |
| | | RT - R | TGTTTCTTTCGAGAGGGATTTC |
| INO80C | rs3088182 | Geno - F | CCAGGCCAACTACACAGACC |
| | | Geno - R | ACAGCACTGGCATTTTCAGA |
| | | RT - F | TGCTCCTCCATCCTTTAAGC |
| | | RT - R | GCACTGGCATTTTCAGATTG |
| VWDE | rs2192828 | Geno/RT - F | TGGGGAACACTGTCAGAATG |
| | | Geno/RT - R | AGGTGTGTTGCATCGTTTCC |
| ARL4A | rs2280633 | Geno - F | GGTTTGTTTAGCGCAGCAAG |
| | | Geno - R | ACAGCCCATTCCCCATTT |
| | | RT - F | GCTTGGAGAGCTAAGCTGGA |

| Gene | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | RT - R | TGTGGAAAGACTGAAATGAAGG |
| GLRX | rs4561 | Geno - F | GAAGTCCTGTCCTGTGAGCA |
| | | Geno - R | AGCTCCAATCTGCTTTAGCC |
| | | RT - F | TCAGTCAATTGCCCATCAAA |
| | | RT - R | GCAGAGCTCCAATCTGCTTT |
| SPATA9 | rs213509 | Geno/RT - F | TCATTATGGACTGAAACTGGAGAA |
| | | GenoRT - R | GGACACTCTACTGATCTTGCCTCT |
| S1PR2 | rs2116942 | Geno - F | CGATCCTCTACAAAGCCCACT |
| | | Geno - R | CAGAAACGTGGGTGACGTG |
| | | RT - F | CCTTCAGCATCCTCCTTCTG |
| | | RT - R | CTATCTGGGGTCACCCAGTG |
| LAMP3 | rs3811729 | Geno - F | TTTTCGGCTGCTTCTCATAA |
| | | Geno - R | GGGACTCCAGAGTGCGTTTA |
| | | RT - F | AGCAGGTGGTTTCGTTTCTC |
| | | RT - R | AGTTTGGTGAGGTGCTTGCT |
| DCUN1D1 | rs4859146 | Geno - F | AATATCGTTTCATGCCTTCCA |
| | | Geno - R | TGTGGGCCAGCAAATCTTTA |
| | | RT - F | CGCAGAAGGATAAAGTTCGTCA |
| | | RT - R | TTTGTCCTGGATTCTTTGCAAAA |
| RFX3 | rs832297 | Geno - F | CGGAAGCTTGCATTGTTAACA |
| | | Geno - R | AGGTGCTTTTCATTGGACAGA |
| | | RT - F | CGTGTCTCCTGGAAATCTGG |
| | | RT - R | AACGGAAGCTTGCATTGTTA |
| SLC1A1 | rs3087879 | Geno - F | AAGCCACGTTGCAGAAAAAG |
| | | Geno - R | GCAGAATGACAAGCAGAGGA |
| | | RT - F | CACCCTGATCATTGCTGTCG |
| | | RT - R | AGCAGAATGACAAGCAGAGG |
| RDX | rs2298500 | Geno - F | TCATAAGCTGATGCCTTCCA |
| | | Geno - R | CTGGCTCATGGTTCTCTCCT |
| | | RT - F | TCTTGCCAAAGCAGCATAAA |
| | | RT - R | AGATGGGGTGCAGTAGATGC |
| ADAM23 | rs3732079 | Geno - F | ATGCTAGGAGAAATCCACATCCA |
| | | RT - F | TCCACGTATTGGTCAACTTCAGGGT |
| | | Geno/RT - R | CATGGCCCGAACAGACTTTACC |
| ZDBF2 | rs7582864 | Geno/RT - F | ACTTTTGATTCTGAACAACTTCA |
| | | Geno/RT - R | CTACTTTGGGTTGATCAGCCAC |
| WDR91 | rs292592 | Geno - F | TAGCCCGGGAGGTCTATGTT |
| | | Geno - R | TTACCTGACCACCGAAGGAC |

| Gene | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | RT - F | CCTGAGCGTCCTGTTTCAGT |
| | | RT - R | TTTCTTGGCCGAGCTCTGAG |
| EIF3G | rs7401 | Geno - F | GCTTCGTCCTCACATTACCC |
| | | Geno - R | AGGACTGGAGACGCAAGAAC |
| | | RT - F | AGACCACTGGCCAATCCAAG |
| | | RT - R | GCTTTATTGCCCTTGGAGCC |
| RTN4IP1 | rs1987623 | Geno - F | CAGTGTTTACAATGGAATTTCTGA |
| | | Geno - R | GCGTGAACTTTGACAATGACTT |
| | | RT - F | ATGCTCCGTTTTGCATGCTT |
| | | RT - R | ATCACCCAAGCAGGCATGAC |
| QRSL1 | rs2015205 | Geno - F | GCACTGGAAAGTGGAAGAGG |
| | | Geno - R | TCAAGGACTGCTGAACAATCA |
| | | RT - F | GCCATGTATGCTGCAACCAG |
| | | RT - R | CTGGTCACAAAACGCACGTC |
| HACE1 | rs45521835 | Geno - F | GCCTAATATCACTTGGATGTTGGT |
| | | Geno - R | CACAATGACTGCTTATGAGAAAATC |
| | | RT - F | TCACTTGGATGTTGGTGGGT |
| | | RT - R | TGGCAAACACAATGACTGCT |
| BVES | rs221653 | Geno - F | AGCAATGAGGTTGTCCATTTG |
| | | Geno - R | TACTTCTGGCCTGGGAGAAA |
| | | RT - F | GGATTCACGCCACTCTCCT |
| | | RT - R | AGCAGATCTTTCAGTGGCCT |
| ST3GAL1 | rs4736674 | Geno - F | TGTGTTAGCCAAAAGGCACA |
| | | Geno - R | CACCTCATCAGCCCATCTTT |
| | | RT - F | GGAGAATGAGGAGCACTGGT |
| | | RT - R | CTTCAGGTTCAGCTCTGGGT |
| CALD1 | rs2732897 | Geno/RT - F | TGGTACTGATTTTTTAGGTTGGT |
| | | Geno/RT - R | CAGTTCTAATCTCTCTTCTTTTCAATAC |
| TMEM140 | rs2288237 | Geno/RT - F | GCAACCTCACTGACCTGCCCAACCTG |
| | | Geno/RT - R | AGAGCCTGACCCACTTCCACACA |

**Other species studies**

Table S19: **BS Primers – *Macaca mulatta* methylation study**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| DNMT1 | | F | AGGTATTGAGGGATTTTTTATTAGAGAAG |
| | | R | TCTTCAAACAACRATAAAG |
| LIN28B | | F | TGAAATATAAAGGTAATTTGT |
| | | R | ACATAAAAACAAATACAAAAC |
| AIM1 | | F | TTTGAGTAAGATTGAGGGGGTATAA |
| | | R | CCCTCTTAAATAAAAACTCCTTAAC |
| FGF12 | | F | GTTTTTTTTGTAGTGTTGTTGGT |
| | | R | AAAAAAACAACAATTTAACTCCCTAC |
| THSD7A | | F | TAGAAAGTAAAGTTTTTTTTGT |
| | | R | CCTAACAAACCTACTCAATCAAC |
| MCCC1 | | F | GATTTTTTAGTATAGAGGTAGT |
| | | R | CTCTAATCTTTAATTAATATTC |
| GLIS3 | | F | TtTGGATGGGGGTTTTTATTTT |
| | | R | CTCTAAACTTTAACTCCCCCACAC |
| PURA | | F | GGTGGATATTTAGAATAAG |
| | | R | AACCCTACRTAAAACCCAAACCAAACCCC |
| MCTS2 | | F | GGAAGTTGTTAGATTTTA |
| | | R | CTAAATACAATTAAACACACTTTCC |
| GRB10 | | F | GGGTTTTYGTGGGTATAGTTATTATT |
| | | R | CCCRCTCTCCAAATACTCAAATAAACT |
| L3MBTL1 | | F | TTTTAGGTTTTGAGTTGGGTTTT |
| | | R | CCCCCTCAAACCCTACCTCC |

Table S20: **BS Primers – *Mus musculus* methylation study**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| Ago1. | | Out - F | AGATTGAGGAGTTTAGYGTGA |
| | | In - F | TTTAGGTATTTCGTATTTGTAT |
| | | In - R | CCTCTCCATTAACCTTTATTAC |
| | | Out - R | CGCCCAACCCCTCCCCTCTTAC |
| Jmjd1c | | Out - F | GATTTTYGGGAGTTGGTTTT |
| | | In - F | TAGAAGTAGTTTGGTTGTTGTTT |
| | | In - R | CCACAAACAAACTAACAAAAATAAAC |
| | | Out - R | CTTCCCACCAACTCRAACC |
| Cacna1e | | Out - F | AGGAGGTTATTGATAGTTTTATA |
| | | In - F | ATGTTAGGTTTAGGTTTTTTA |
| | | In - R | CCTTCCCCACAACACTACATC |
| | | Out - R | CCTTACACCCCCAACAAAACCAAC |
| Fgf8 | | Out - F | GGTTGGGTGGGGAGGAGTTGGG |
| | | In - F | GGGGTTTAGTTTTTTGGTTG |
| | | In - R | ACTAAAATACCCACCTACTATCTA |
| | | Out - R | ATATCACTACTACCCAAAAAAACA |
| Bank1 | rs30780713 | Out - F | AGGGAGTTGGTTTTATATAGATTGA |
| | | In - F | AGGGGAATGGGTAAGGATGTTTTA |
| | | In - R | CACTACAACTCCCAAACTCCAAC |
| | | Out - R | CTACAAAAAAAATTACAAACAAAAC |
| Mif2 | rs32546086 | Out - F | AGAGTTAAGGTTTAGGTTGTTA |
| | | In - F | AGAATTAGAAAAAGTTTTTTAAA |
| | | In - R | CTAAAACACTACCAAAACCTTCTC |
| | | Out - R | CCTCAAAATTATATAACTCTAAC |
| Emilin | rs33061851 | Out - F | GGGTTTTTAGGTAGGATGGG |
| | | In - F | GATGGGGTAGGGGTTTATTTATAG |
| | | In - R | TTCCAAACCTCTTTCTCCCTT |
| | | Out - R | AATAAAATATCAAATCCTACACAA |
| Svopl | | Out - F | TAGGTAGGGAGTATAGGTAGGT |
| | | In - F | TTGAGGTTAGTGTTTATATTT |
| | | In - R | CCAAACTAAATCTATAATTCTC |
| | | Out - R | CCAATACCTATCAAAACAAC |
| Fancc | | Out - F | TTAGAGTTGTTTTAGTAGAT |
| | | In - F | TTGAAAATTTGGTATAGTTTT |
| | | In - R | TCAAAAACCACATTCCAAAAC |
| | | Out - R | TAAAATTTAACCTATCTAACAACTC |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| Kik15 | | Out - F | GATTTTTGGTTTTTTGTGGTTTTAGGG |
| | | In - F | GTTGTTTGATAATAGATTTTAGTTTG |
| | | In - R | TTATCTTCTCTAACCACTTTATCT |
| | | Out - R | TCAAAAACTTAAAAAAAAAAAAAAATTCT |
| Spread2 | | Out - F | GTTTTAAGATTTGTAGTTTTGGG |
| | | In - F | GTGAGGATGGGTAGAAGTAG |
| | | In - R | TACTCTACTACTAAAAAAAATTTTTAT |
| | | Out - R | TAAACCAATACATATCAAACTAATACT |
| Lmo3 | | Out - F | GTTGGGTTTATTGAATATTTTTTTTTG |
| | | In - F | GTTTAATTGAGAAGGTAAGTTAGGG |
| | | In - R | ACTTTATACCTTCTTAAAAAAAAATAAAAAA |
| | | Out - R | AAAACCCTCTCTACAAACCTACA |
| Fam46b | | Out - F | GAGTTTAAGATGATTTGGAAGGAAG |
| | | In - F | GTTGGATAGAGAGATGAGGTTTTAG |
| | | In - R | ACTTAATATTCCAAATAAACCTA |
| | | Out - R | ACCTAAACTACAAAAACCTAA |
| Rhobtb1 | | Out - F | GGTTTAGTTGGTAGTTAATTAGG |
| | | In - F | GGTTTTTATATGATTTAGGGAG |
| | | In - R | ACTCATATAAATAAATTTAATATCCA |
| | | Out - R | AAAAACTTAACAAACTATCTACA |
| Pkd1l2 | | Out - F | GATAATTTTGTATAGGATTTTTGTTG |
| | | In - F | GTAGTTTGGGATGTTGGAATGAG |
| | | In - R | ATATACCCTACAAAAAAACCCA |
| | | Out - R | TTCAAAACTTTAAACAATAAACT |
| Kif9 | | Out - F | AGGAATTATAAAGAATTTAGAGA |
| | | In - F | TTAGTTTAAAGGTTGTTTTAGAGTTTT |
| | | In - R | CACTATTAACTAATAAACTCCC |
| | | Out - R | ATAACAAACAATACCTCAATAAA |
| Tm6sf2 | | Out - F | GTTTTTGGTAATTTAGAGTTTTG |
| | | In - R | AACCCCATAATCCTCATAA |
| | | Out - R | TTACTACCCCTCATCCTCCCT |
| Evx1 | | Out - F | GTATGTGTATTTGTGTGTTTTGGG |
| | | In - F | GTTTTGGGTAGTAGTTTTATTAGTGG |
| | | In - R | AACCTACCTTAATAATAATTTCA |
| | | Out - R | TTAACATCTCCTTCCTACTAACCT |
| Stk10 | | F | GTTGAAGGTTAGATAGTAATAG |
| | | R | ATCCCCTACTATATACAACCAATCCA |
| Scm12 | rs29306878 | F | TAGTTTGTTGTTAGGATTTTTGAT |

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| | | R | CTCCACTTCCCAAAAAAAACC |
| Spry1 | | F | AGTTGTTGTTGAAGGTAT |
| | | R | CATACAAATACAATCCTAAAAAC |

Table S21: **RT Primers – *Mus musculus* gene expression study**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| Svolp | | RT - F | CTTGCTGGCTCCCTGCTCATC |
| | | RT - R | CCGCATCGCCTCCATCCCTGGC |
| Fancc | | RT - F | AGCTGCTTCCTGCAGATCAGTA |
| | | RT - R | GATGGAGTGCTGTATAAGAGAG |
| Adam23 | | RT . F | TGGCTCAAAACCTTGGAATC |
| | | RT- R | AGCGACCCTGATTTTGATTG |

**Histone modifications and hydromethylation studies**

Table S22: **5-hmC qPCR Primers**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| SEMA3B | | 5hmC - F | GTCCCTACCAGCCCCTCCCTTG |
| | | 5hmC - R | ACCCTCTGACGGCTCCTTTAAGA |
| MCCC1 | rs937652 | 5hmC - F | CCACTCCGTGACTCCCCAGTAC |
| | | 5hmC - R | ACTGCCCAATCGTCGCTACTGA |
| SCIN | rs2240571 | 5hmC - F | GCCAGGCACAACTGAGATCCA |
| | | 5hmC - R | TTGCCCGCCCGGGCGAACTCTT |
| ACTL10 | rs4911356 | 5hmC - F | GCCGCGCATAGTGCTGAAGAG |
| | | 5hmC - R | TCCCACAGCCCTTCCAGCGCCT |

Table S23: **ChIP qPCR Primers**

| Region | SNP | | Sequence (5' -> 3') |
|---|---|---|---|
| GAPDH1 | | ChIP - F | GACCCCTTCATTGACCT |
| | | ChIP - R | CCACCACCCTGTTGCTG |
| LIN28B | | ChIP(SNP) - F | AGTGGTTTACCAAAACGCGGCGGA |
| | | qChIP - F | ACGTTTGCCATCGCGGCTTCCCGA |
| | | ChIP - R | CAAGAAAATGAGCACCCCCACC |
| R3HCC1 | | ChIP(SNP) - F | GTTTCTAAGGCCCTTTCGGTG |
| | | qChIP - F | CTCTCCCACCTGTCACCCTGGC |
| | | ChIP - R | AAGCGGTCCAGTTCCTCCTGGAT |
| MEST | | ChIP(SNP) - F | GACTCCGGCTTCCCTCTG |
| | | ChIP(SNP) - R | GGTCACATCGCCGTTCTCACTC |
| SNURF | | ChIP(SNP) - F | AGCACCACCACCAAGTAGGCAA |
| | | ChIP(SNP) - R | CGGCGGTGGGCATTGGCAGC |

# Appendix II: Publications

During the course of this doctoral thesis and as a result of it, four scientific publications have been published. In addition, I have contributed to several review article discussing the issues covered in this thesis:

i. Sanchez-Delgado M, Martin-Trujillo A, Tayama C, Vidal E, Esteller M, Iglesias-Platas I, Deo N, Barney O, Maclean K, Hata K, Nakabayashi K, Fisher R,, Monk D. (2015) **Absence of Maternal Methylation in Biparental Hydatidiform Moles from Women with *NLRP7* Maternal-Effect Mutations Reveals Widespread Placenta-Specific Imprinting.** *PLoS Genet* 11(11): e1005644. https://doi.org/10.1371/journal.pgen.1005644.

ii. Sanchez-Delgado M, Court F, Vidal E, Medrano J, Monteagudo-Sánchez A, Martin-Trujillo A, Tayama C, Iglesias-Platas I, Kondova I, Bontrop R, Poo-Llanillo ME, Marques-Bonet T, Nakabayashi K, Simón C, Monk D. (2016) **Human Oocyte-Derived Methylation Differences Persist in the Placenta Revealing Widespread Transient Imprinting.** *PLoS Genet* 12(11): e1006427. https://doi.org/10.1371/journal.pgen.1006427.

iii. Hernandez Mora, JR*, Sanchez-Delgado, M*, Petazzi, P*, Moran, S, Esteller, M, Iglesias-Platas, I, & Monk, D. (2017) **Profiling of oxBS-450K 5-hydroxymethylcytosine in human placenta and brain reveals enrichment at imprinted loci.** *Epigenetics*, (ahead of print), 00-00. (* These authors contributed equally to this work) http://dx.doi.org/10.1080/15592294.2017.1344803.

iv. Hernandez Mora, J R, Tayama, C., Sánchez-Delgado, M, Monteagudo-Sánchez, A, Hata, K, Ogata, T., Medrano, J, Poo-Llanillo, ME, Simón, S., Moran, S, Esteller, M, Tenorio, J, Lapunzina, P, Kagami, M, & Monk, D, Nakabayashi, K. **Characterization of parent-of-origin methylation using the Illumina Infinium MethylationEPIC array platform.** *Accepted Epigenomics.*

v. Sanchez-Delgado M, Riccio A, Eggermann T, Maher ER, Lapunzina P, Mackay D, Monk D. (2016) **Causes and Consequences of Multi-Locus Imprinting Disturbances in Humans.** *Trends Genet.*32(7):444-55. http://dx.doi.org/10.1016/j.tig.2016.05.001 **Review.**

vi. Monk D, Sanchez-Delgado M, Fisher R. (2017) **NLRPs, the subcortical maternal complex and genomic imprinting.** *Reproduction.* 154(6):R161-R170. http://dx.doi.org/10.1530/REP-17-0465 – **Review.**

*The data about* ***assesing imprinted DMRs in complicated pregnancies and the histone modifications analysis*** *is currently unpublished.*

# Appendix III: Director's report

I certify that the PhD student Marta Sanchez-Delgado will defend her PhD thesis by article publication compendium, 3 of which are already published in scientific journals. Her contribution to each of the articles is specified below:

Article 1:
Authors: **Sanchez-Delgado M**, Martin-Trujillo A, Tayama C, Vidal E, Esteller M, Iglesias-Platas I, Deo N, Barney O, Maclean K, Hata K, Nakabayashi K, Fisher R, Monk D.

Title: Absence of Maternal Methylation in Biparental Hydatidiform Moles from Women with NLRP7 Maternal-Effect Mutations Reveals Widespread Placenta-Specific Imprinting.

Journal: PLoS Genet. 2015;11: e1005644. Impact Factor (7.1)
doi: 10.1371/journal.pgen.1005644.
PMID: 26544189

In this article Marta Sanchez-Delgado was the main person responsible, along with Dr Rosemary Fisher and myself, for conceiving the experimental design. She was solely responsible for all confirmation experiments involving bisulphite PCR, pyrosequencing and allelic expression following initial bioinformatic data screening by Alex Martin-Trujillo. She was also instrumental in the design and optimization of methylation-sensitive genotyping used to initially screen candidate DMRs. Marta, along with other team members interpreted all results. She was also responsible for the assembly of all figures in the publication and was directly involved in the writing and revision of the manuscript.

Article 2:
Authors: **Sanchez-Delgado M**, Court F, Vidal E, Medrano J, Monteagudo-Sánchez A, Martin-Trujillo A, Tayama C, Iglesias-Platas I, Kondova I, Bontrop R, Poo-Llanillo ME, Marques-Bonet T, Nakabayashi K, Simón C, Monk D.

Title: Human Oocyte-Derived Methylation Differences Persist in the Placenta Revealing Widespread Transient Imprinting.

Journal: PLoS Genet. 2016. 12: e1006427. (Impact Factor 6.1)
doi: 10.1371/journal.pgen.1006427
PMID: 27835649

This work was performed in collaboration with the fertility team of Professor Carlos Simon at IVI Valencia who supplied all blastocysts samples used in this project. In this article Marta Sanchez Delgado was the sole person responsible for generating all the molecular data. All members of the team devised the experimental strategies, with bioinformatics analysis performed by Alex Martin-Trujillo and Enrique Vidal. Finally Marta, along with Alex and myself, interpreted all results. Marta was responsible for cataloguing all results, which now appear on our laboratories webpage. She was also responsible for the

compilation of all figures in the publication and was directly involved in the writing and revision of the manuscript.

Authors: Hernandez Mora JR*, **Sanchez-Delgado M***, Petazzi P*, Moran S, Esteller M, Iglesias-Platas I, Monk D.

Title: Profiling of oxBS-450K 5-hydroxymethylcytosine in human placenta and brain reveals enrichment at imprinted loci.

Journal Epigenetics. [Epub ahead of print]. (Impact Factor 4.4)
doi: 10.1080/15592294.2017.1344803.
PMID: 28678681

In this article Marta Sanchez Delgado was involved in the conception and performed all validation experiments involving methylation (both quantitative and allelic T4-BGT assays) following the exhaustive data analysis by Jose Hernandez, the team's bioinformatican. She and Paolo Petazzi were in charge of generating the TET-expression profiles in placenta and brain biopsies. She was responsible for the compilation of all figures in the publication and was directly involved in the writing and revision of the manuscript.

In witness of the candidate and the PhD tutor, I hereby sign the present doctoral thesis,

20th March 2018, L'Hospitalet de Llobregat (Barcelona)

Dr David Monk
Group leader of Imprinting & Cancer Group,
Cancer Epigenetics and Biology Program (PEBC),
Bellvitge Biomedical Research Institute (IDIBELL)
Avda. Gran Via 199-203 08908 L''Hospitalet de Llobregat, Barcelona, Spain
Email: dmonk@idibell.cat
http://www.humanimprints.net/

# Appendix IV – Article 1: *Absence of Maternal Methylation in Biparental Hydatidiform Moles from Women with NLRP7 Maternal-Effect Mutations Reveals Widespread Placenta-Specific Imprinting*

http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1005644

# Appendix V – Article 2: *Human Oocyte-Derived Methylation Differences Persist in the Placenta Revealing Widespread Transient Imprinting*

http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1006427

# Appendix VI – Article 3: *Profiling of oxBS-450K 5-hydroxymethylcytosine in human placenta and brain revealsenrichment at imprinted loci*

https://www.tandfonline.com/eprint/IIrE4mvGMkk76tdVjFpY/full