



# Characterisation of monoallelic regulation variation during human development

Dagne Daskeviciute

A thesis submitted for the degree of Doctor of Philosophy (PhD)

University of East Anglia  
School of Biological Sciences

September 2024

© This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived therefrom must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

# Abstract

Genomic imprinting is the parent-of-origin monoallelic expression of genes. It is an epigenetic process in which chromosomal regions from both parents become differentially marked, primarily by DNA methylation. Several research groups, including ours, have previously found that the human placenta contains many unique differentially methylated regions (DMRs) that are not present in other somatic tissues. A more extensive characterisation of these placenta-specific DMRs revealed that they are derived from human oocytes and are maintained throughout pre-implantation development. I refer to these regions as placenta-specific maternal DMRs (mDMRs). Many of these placenta-specific mDMRs were identified by screening whole-genome bisulphite sequencing (WGBS) datasets from human gametes, blastocysts, and somatic tissues, including the term placenta. Curiously, some mDMRs were found to be highly polymorphic in the human population, and only some regulate monoallelic expression. The role of these placenta-specific mDMRs during development remains unclear, and many of the previously identified regions have yet to be fully characterised.

In addition, several groups have identified a novel form of imprinting, initially mediated by histone post-translational modifications (PTMs) in rodent pre-implantation embryos. These histone PTMs are later replaced by secondary differentially methylated regions (sDMRs), often at endogenous retroviral elements (ERVs) in rodent extra-embryonic tissues. This type of imprinting is referred to as non-canonical imprinting. Non-canonical imprinting has been shown to be critical for imprinted X chromosome inactivation (XCI) in rodent embryos and plays an important role in normal placental development. A few studies have attempted to investigate whether non-canonical imprinting is conserved in human embryos, but the findings have been inconsistent. Their status in the human placenta remains uninvestigated.

During my PhD, I revisited placenta-specific mDMRs discovered by our group and others, which led to the identification of two promising placenta-specific mDMRs located at the CpG island promoters of G0/G1 Switch Regulatory Protein 2 (*GOS2*) and Phosphoinositide-3-Kinase Regulatory Subunit 1 (*PIK3R1*) isoform 3. I applied various molecular biology techniques, including methylation-sensitive genotyping, bisulphite PCR, and allelic RT-PCR, followed by Sanger sequencing, in a large placental cohort to characterise their allelic usage. I demonstrated that the placenta-specific mDMRs of *GOS2* and *PIK3R1* isoform 3 are highly polymorphic, exhibiting maternal allele-specific methylation and monoallelic expression. Bisulphite-converted DNA from placental trophoblast and stromal cells, isolated using magnetic cell separation, revealed cell type-

specific imprinting of these two genes.

Additionally, I applied the same techniques to investigate non-canonical imprinted genes, primarily on the human term placenta and human pre-implantation embryos. I screened human orthologs of mouse and rat non-canonical imprinted genes, non-canonical imprinted genes previously reported in human embryos, genes with primate-specific ERV long terminal repeat (LTR) elements, and genes harbouring potential placental sDMRs. The results provided no evidence of non-canonical imprinting in the human placenta. However, further research is needed to investigate non-canonical imprints in human pre-implantation embryos. During this screen, I also demonstrated that *XIST* ncRNA, which is required for XCI in females, is not imprinted in human placental samples. Additionally, I identified several novel placenta-specific mDMRs that may regulate placenta-specific imprinting in the human placenta.

## **Access Condition and Agreement**

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

# Table of Contents

<b>Abstract</b> .....	<b>II</b>
<b>Table of Contents</b> .....	<b>IV</b>
<b>List of Tables</b> .....	<b>X</b>
<b>List of Figures</b> .....	<b>XI</b>
<b>Abbreviations</b> .....	<b>XIII</b>
<b>Declaration</b> .....	<b>XVIII</b>
<b>Acknowledgments</b> .....	<b>XIX</b>
<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1. Epigenetic modifications .....	5
1.1.1. 5-methylcytosine (5mC) .....	5
1.1.2. DNA methyltransferases (DNMTs) .....	7
1.1.3. DNA demethylation .....	8
1.1.4. Non-CG methylation.....	8
1.1.5. Histone proteins .....	11
1.1.6. Histone post-translation modifications (PTMs) and their role in gene regulation .....	14
1.1.7. Permissive histone marks.....	16
1.1.7.1. Histone 3 lysine 4 trimethylation (H3K4me3).....	16
1.1.7.2. Histone 3 lysine 4 monomethylation (H3K4me1) and histone 3 lysine 27 acetylation (H3K27ac) .....	17
1.1.7.3. Histone 3 lysine 36 di- and tri-methylation (H3K36me2 and H3K36me3) .....	17
1.1.8. Repressive histone marks.....	18
1.1.8.1. Histone 3 lysine 9 di- and tri-methylation (H3K9me2, H3K9me3) .....	19
1.1.8.2. Histone 3 lysine 27 trimethylation (H3K27me3) .....	20
1.1.8.3. Histone 2 A lysine 119 monoubiquitination (H2AK119ub1).....	21
1.2. Epigenetic landscapes of male and female gametes.....	21
1.3. Embryonic genome activation (EGA).....	22
1.4. Epigenetic modifications and their role during embryonic development .....	24
1.5. Endogenous retroviruses (ERVs) .....	28
1.6. Genomic imprinting .....	31
1.7. The life cycle of imprints .....	32

1.7.1.	Pre-implantation epigenetic reprogramming.....	32
1.7.2.	Post-implantation maintenance of imprinted regions .....	33
1.7.3.	Genome-wide epigenetic reprogramming in primordial germ cells (PGCs) ... .....	33
1.7.4.	Establishment of sex-specific imprints.....	34
1.8.	Classical imprinted genes .....	34
1.9.	DNA methylation-dependent imprinting .....	36
1.10.	DNA methylation-independent imprinting.....	38
1.11.	X chromosome inactivation (XCI).....	40
1.12.	Use of uniparental cell lines and embryos to study parental epigenomes .....	41
1.13.	Human placenta.....	43
1.13.1.	Early development of the human placenta .....	45
1.13.2.	Spiral artery remodelling and other placental cell types .....	47
1.13.3.	Placental epigenome .....	49
1.13.4.	Placental transcriptome .....	52
1.13.5.	Placental pathologies .....	53
1.13.5.1.	Pre-eclampsia (PE) .....	53
1.13.5.2.	Mechanism of PE .....	55
1.13.5.3.	Intrauterine growth restriction (IUGR).....	56
1.13.6.	Human trophoblast stem cells (hTSCs).....	58
1.14.	PhD aims .....	61
<b>Chapter 2: Methods .....</b>		<b>65</b>
2.1.	Ethics approval for human samples .....	66
2.2.	Samples .....	66
2.2.1.	Placental cohort and parental samples .....	66
2.2.1.1.	Placental samples.....	66
2.2.1.2.	Blood and saliva samples .....	67
2.2.2.	Human embryos.....	67
2.2.3.	Mouse placenta .....	68
2.2.4.	Cell lines .....	68
2.2.4.1.	Cytotrophoblast stem cells 27 and 30 (CT27 and CT30).....	68

2.2.4.2.	Mole 1 and Mole 2.....	68
2.3.	Molecular Biology techniques .....	69
2.3.1.	Mononuclear cell extraction from the cord and maternal blood .....	69
2.3.2.	DNA extraction .....	70
2.3.2.1.	DNA isolation from placental biopsies and blood.....	70
2.3.2.2.	DNA extraction from saliva .....	71
2.3.3.	RNA extraction from the placental samples and cell lines.....	72
2.3.4.	DNA digestion with methylation-sensitive restriction enzymes.....	72
2.3.5.	cDNA synthesis.....	73
2.3.6.	Sodium Bisulphite DNA conversion.....	74
2.3.7.	Polymerase chain reaction (PCR).....	75
2.3.7.1.	Different types of PCR for investigating candidate genes .....	77
2.3.7.1.1.	Nested PCR.....	77
2.3.7.1.2.	Genotyping PCR .....	78
2.3.7.1.3.	Methylation-sensitive genotyping.....	78
2.3.7.1.4.	Bisulphite PCR or Nested Bisulphite PCR .....	81
2.3.7.1.5.	Allelic RT-PCR or Nested allelic RT-PCR.....	81
2.3.8.	PCR product purification.....	81
2.3.8.1.	Agarose gel electrophoresis .....	81
2.3.8.2.	Gel extraction.....	82
2.3.8.3.	PCR clean-up by ethanol precipitation.....	82
2.3.9.	Sub-cloning.....	83
2.3.10.	Sanger sequencing .....	85
2.3.10.1.	Sample preparation for sequencing using the BigDye Terminator (BDT) .....	85
2.3.10.2.	Post-sequencing clean-up and capillary electrophoresis .....	86
2.3.10.3.	Sanger sequencing data analysis .....	86
2.4.	Cell isolation and culture techniques .....	88
2.4.1.	Cell isolation from the human placenta by Magnetic-Activated Cell Sorting (MACS) .....	88
2.4.1.1.	Protocol overview .....	88

2.4.1.2.	Percoll gradient .....	88
2.4.1.3.	Magnetic-activated cell sorting (MACS) .....	89
2.4.1.4.	Isolation of placental cell types .....	90
2.4.1.4.1.	Placenta dissection .....	90
2.4.1.4.2.	Percoll gradient.....	92
2.4.1.4.3.	MACS .....	93
2.4.2.	Cell culture conditions for CT and Mole cell lines .....	94
2.4.2.1.	Cryogenic storage and cell recovery .....	94
2.4.2.1.1.	Preparation of cell culture plates for cell lines .....	94
2.4.2.2.	Cell maintenance and harvesting.....	95
2.5.	Quantitative techniques .....	95
2.5.1.	Placenta-specific imprinted genes with mDMR .....	95
2.5.2.	Non-canonical imprinting in the human placenta and embryos.....	96
2.5.3.	Quantitative real-time reverse transcription PCR (qRT-PCR).....	96
2.5.3.1.	Comparative $\Delta\Delta\text{Ct}$ method .....	97
2.5.4.	Pyrosequencing.....	99
2.6.	Single-cell methylation and transcriptome sequencing (scM&T-seq) of human pre-implantation embryos .....	101
2.6.1.	Overview of scM&T-seq .....	101
2.6.2.	Preparation of single-cell RNA sequencing (scRNA-seq) libraries .....	102
2.7.	Publicly available methyl-seq datasets .....	107
2.8.	Publicly available Illumina methylation array datasets.....	108
2.9.	Bioinformatics analysis.....	109
2.9.1.	Gene synteny analysis .....	109
2.9.2.	ScRNA-seq processing .....	110
2.9.2.1.	Overview of scRNA-seq analysis.....	110
2.9.2.2.	ScRNA-seq analysis.....	111
2.9.2.3.	Bioinformatic programs and tools .....	112
2.9.2.3.1.	STAR aligner .....	112
2.9.2.3.2.	GATK.....	113
2.9.2.3.3.	GATK pipeline for RNA-seq: .....	113

<b>Chapter 3: <i>PIK3R1</i> and <i>GoS2</i> are human placenta-specific imprinted genes associated with germline-inherited maternal DNA methylation .....</b>	<b>116</b>
3.1. Introduction.....	117
3.2. The datasets used for identifying candidate genes with placenta-specific maternal gDMRs.....	118
3.3. Gene selection criteria and analysis .....	119
3.4. Discovery of novel placenta-specific maternal DMRs (mDMRs) in the human placenta.....	120
3.5. Allelic methylation at <i>GoS2</i> mDMR in the human placenta.....	121
3.6. Allelic methylation at placenta-specific mDMR of <i>PIK3R1</i> isoform 3 in the human placenta.....	126
3.7. Cell-type specific methylation of <i>GoS2</i> and <i>PIK3R1</i> .....	130
3.8. Polymorphic imprinting events in the placental cohort.....	133
3.9. Placenta-specific mDMRs are not conserved in the mouse placenta.....	138
<b>Chapter 4: Non-canonical imprinting, manifesting as post-fertilisation placenta-specific parent-of-origin methylation, is not conserved in humans ...</b>	<b>142</b>
4.1. Introduction.....	143
4.2. Selection of human candidate genes for non-canonical imprinting .....	145
4.3. Human orthologs of mouse non-canonical imprints are not conserved in the human placentae.....	147
4.3.1. <i>SFMBT2</i> .....	149
4.3.2. <i>JADE1</i> .....	150
4.3.3. <i>SMOC1</i> .....	150
4.3.4. <i>GAB1</i> .....	151
4.3.5. <i>SLC38A4</i> .....	152
4.3.6. <i>SALL1</i> .....	152
4.4. Human orthologs of rat non-canonical imprints are not conserved in the human placenta.....	159
4.5. Mouse and rat non-canonical imprinted genes demonstrate biallelic expression in human pre-implantation embryos.....	162
4.6. X chromosome demonstrates random inactivation in the human placenta and embryos.....	168

4.7.	LTR-driven human transcripts are not imprinted in the placenta.....	172
4.8.	Previously identified human candidates of non-canonical imprinting demonstrate biallelic expression in the human placenta.....	177
4.9.	Human genes with placenta sDMR regions are not imprinted in the human placenta .....	180
4.10.	Novel candidate imprinted genes with placenta-specific mDMRs.....	187
4.11.	Candidate gene expression in different placental cell lines.....	191
4.12.	Gene synteny .....	192
<b>Chapter 5:</b>	<b>Discussion .....</b>	<b>195</b>
5.1.	Polymorphic imprinting of <i>GOS2</i> and <i>PIK3R1</i> in the human placenta .....	196
5.1.1.	Functional role of <i>GOS2</i> during gestation and the placenta .....	205
5.1.2.	Functional role of <i>PIK3R1</i> during gestation and the placenta.....	207
5.1.3.	Study limitations.....	211
5.1.4.	Future research .....	214
5.1.5.	Conclusions .....	216
5.2.	Non-canonical imprinting might not be conserved in humans.....	217
5.2.1.	Study limitations.....	226
5.2.2.	Future research .....	227
5.2.3.	Conclusions .....	227
<b>References.....</b>		<b>229</b>
<b>Appendix.....</b>		<b>264</b>

# List of Tables

Table 2.1. cDNA synthesis Master Mix for one reaction .....	74
Table 2.2. The reagents used for a standard PCR Master Mix (for one sample) .....	76
Table 2.3. LB broth .....	85
Table 2.4. LB agar Petri dishes .....	85
Table 2.5. The amount of a PCR product required for sequencing with BDT.....	86
Table 2.6. Reagents and equipment required for trophoblast and stromal cell enrichment from placental cell suspensions by MACS.....	89
Table 2.7. Enzyme solutions .....	91
Table 2.8. Buffers.....	91
Table 2.9. Preparation of a 90% percoll density gradient for cell separation. ....	92
Table 2.10. Bead Mix .....	103
Table 2.11. G&T-seq wash buffer .....	104
Table 2.12. RT Master Mix.....	104
Table 2.13. PCR Master Mix .....	105
Table 3.1. Result summary for the <i>GoS2</i> placenta-specific mDMR .....	124
Table 3.2. Result summary of <i>PIK3R1</i> isoform 3 placenta-specific mDMR .....	129
Table 4.1. The list of mouse non-canonical imprinted genes and their status in humans.....	149
Table 4.2. Result summary for the human orthologues of mouse and rat non-canonical imprinted genes.....	156
Table 4.3. The list of rat non-canonical imprinted genes and their status in humans .....	162
Table 4.4. Human pre-implantation embryos collected for scM&T-seq and used for single-cell transcriptome sequencing.....	164
Table 4.5. Summary of allelic expression of the orthologs of mouse and rat non-canonical imprinted genes in human pre-implantation embryos .....	168
Table 4.6. Result summary for human genes with LTR-associated promoters .....	176
Table 4.7. Result summary for human genes with placental sDMRs.....	183
Table 4.8. Result summary for human genes with placenta-specific mDMRs .....	189

# List of Figures

Figure 1.1. Changes in global DNA methylation levels throughout the lifecycle of imprinted genes.....	3
Figure 1.2. Genes located in repeat-rich regions near centromeres. ....	13
Figure 1.3. Distributions of diverse permissive and repressive histone marks in IMR90 cells and human hippocampus. ....	15
Figure 1.4. Schematic representation of changes in transcriptome during embryonic genome activation (EGA).....	23
Figure 1.5. Schematic representation of DNA methylation-dependent transient imprinting. ....	37
Figure 1.6. Model of H3K27me3-dependent imprinting. ....	40
Figure 1.7. Different stages of human placental development.....	44
Figure 2.1. Mononuclear cell isolation from maternal peripheral blood or cord blood with the Lymphoprep™ density gradient. ....	70
Figure 2.2. PCR standard amplification conditions with the BIOTAQ or IMMOLASE DNA polymerase. ....	75
Figure 2.3. Methylation-sensitive genotyping assay.....	80
Figure 2.4. pGEM®-T easy vector with the insertion site within <i>lacZ</i> .....	84
Figure 2.5. Cycling conditions for sequencing with BDT.....	86
Figure 2.6. Dissection of the human placenta from several places. ....	90
Figure 2.7. Placental cell separation using a percoll gradient. ....	93
Figure 2.8. Standard qRT-PCR cycling conditions for the Power SYBR™ Green PCR Master Mix. ....	96
Figure 2.9. Standard TaqMan cycling conditions.....	98
Figure 2.10. Pyrograms of two placental samples showing the level of DNA methylation (%) at seven CpG sites (grey highlight) in a PCR amplicon. ....	101
Figure 2.11. Reverse transcription conditions with RT Master Mix. ....	105
Figure 2.12. PCR cycling conditions for KAPA HiFi HotStart PCR Master Mix.....	105
Figure 2.13. PCR cycling conditions for Nextera PCR Master (NPM) Mix to incorporate indexes .....	106
Figure 2.14. Summary of variant calling analysis in pseudo bulk scRNA-seq datasets. ....	111
Figure 3.1. Applied strategy to characterise the methylation and expression patterns of candidate genes with placenta-specific mDMRs. ....	120
Figure 3.2. DNA methylation and allelic expression patterns of <i>GoS2</i> .....	123
Figure 3.3. Investigation of residual maternal contamination in placental samples. ....	125
Figure 3.4. Characterisation of DNA methylation and allelic expression for <i>PIK3R1</i> isoform 3. ....	129
Figure 3.5. Characterisation of allelic methylation of <i>GoS2</i> and <i>PIK3R1</i> placenta-specific	

mDMRs in different placental cell types. ....	132
Figure 3.6. Polymorphic placenta-specific mDMRs of <i>PIK3R1</i> isoform 3 and <i>GOS2</i> . ....	135
Figure 3.7. Quantified expression of <i>GOS2</i> and two <i>PIK3R1</i> isoforms in normal and disease-affected placental samples. ....	137
Figure 3.8. Correlation between <i>PIK3R1</i> isoform 3 expression and DNA methylation levels at the placenta-specific mDMR in normal placental samples. ....	138
Figure 3.9. The mouse orthologs of human <i>PIK3R1</i> and <i>GOS2</i> are not conserved in the mouse placenta. ....	140
Figure 4.1. Applied strategy to investigate the methylation and expression patterns of candidate genes with placental sDMRs. ....	147
Figure 4.2. Characterisation of DNA methylation and allelic expression for the human orthologs of mouse non-canonical imprinted genes. ....	155
Figure 4.3. Characterisation of DNA methylation and allelic expression for the human orthologs of rat non-canonical imprinted genes. ....	161
Figure 4.4. Allelic expression of human orthologs of mouse and rat non-canonical imprinted genes in human pre-implantation embryos. ....	167
Figure 4.5. Expression and methylation patterns of <i>XIST</i> in human pre-implantation embryos and term placental samples. ....	171
Figure 4.6. Characterisation of allelic methylation at gene promoters containing primate-specific LTRs. ....	174
Figure 4.7. Investigation of previously reported human candidate non-canonical imprinted genes. ....	179
Figure 4.8. Systematic screening of human loci with placental sDMRs. ....	182
Figure 4.9. Expression profiling of human canonical imprinted genes, orthologs of mouse and rat non-canonical imprinted genes, and human genes with placenta-specific mDMRs in Mole 1 cells. ....	192
Figure 4.10. Macro-synteny maps showing highly conserved genomic regions between mouse, rat, and human chromosomes. ....	194
Figure 5.1. Representation of the <i>Kcnq1ot1</i> imprinting cluster controlled by <i>KvDMR</i> in the mouse (chr 7qF5) and human (chr 11p15) genomes. ....	203
Figure 5.2. PI3K functions in different signalling pathways. ....	209

# Abbreviations

5caC	5-carboxylcytosine	CS	Cytotrophoblastic shell
5fC	5-formylcytosine	CT27	Cytotrophoblast stem cells 27
5hmC	5-hydroxymethylcytosine	CT30	Cytotrophoblast stem cells 30
5mC	5-methylcytosine	CTBs	Primary placental cytotrophoblasts
6mA	N6-methyladenine	CTCF	CCCTC-binding factor
A	Arginine	CTD	C-terminal domain
AC	Amniotic cavity	CTs	Placental cytotrophoblast stem cells
aESCs	Androgenetic embryonic stem cells	CUT&RUN	Cleavage Under Targets and Release Using Nuclease
AG	Androgenetic	CUT&Tag	Cleavage Under Targets and Tagmentation
AGA	Appropriate for gestational age	ddH <sub>2</sub> O	Double-distilled water
AKT	Protein kinase	DHS	DNase I hypersensitive sites
ALPP	Placental alkaline phosphatase	DKO	Double knockout
ART	Assisted reproduction technologies	DNMTs	DNA methyltransferases
ASE	Allele-specific expression	DOHaD	Developmental Origin of Health and Diseases
ATGL	Adipose triglyceride lipase	Dox	Doxycycline
B2m	Beta-2 microglobulin	dpf	Days post-fertilisation
BDT	BigDye Terminator	DRS	Nanopore direct RNA sequencing
BER	Base excision repair	DRS	Direct RNA sequencing
BH	Brc homology	E	Embryonic day
BWS	Beckwith-Wiedemann syndrome	ECM	Extracellular matrix components
CBX	Chromodomain proteins	EEC	Extra-embryonic coelom
CG	CpG sites	eEVTs	Endovascular extravillous trophoblast cells
ChIP-seq	Chromatin immunoprecipitation followed by sequencing	EGA	Embryonic genome activation
CHM	Complete hydatidiform moles	EMT	Epithelial-to-mesenchymal transition
CL	Cleavage stage	EPC	Ectoplacental cone
CNV	Copy number variant		

Epi	Epiblast	Gαq	G protein alpha q subunit
ERVs	Endogenous retroviral elements	H2AK119ub1	Histone 2 A lysine 119 monoubiquitination
ESCs	Embryonic stem cells	H3K18ub	Histone 3 lysine 18 ubiquitination
EVTs	Extravillous trophoblast cells	H3K23ub	Histone 3 lysine 23 ubiquitination
ExE	Extra-embryonic ectoderm	H3K27ac	Histone 3 lysine 27 acetylation
ExM	Extra-embryonic mesoderm	H3K27me3	Histone 3 lysine 27 trimethylation
FACS	Fluorescence-activated cell sorting	H3K36me2	Histone 3 lysine 36 dimethylation
FAs	Fatty acids	H3K36me3	Histone 3 lysine 36 trimethylation
FBs	Fibroblasts	H3K4me1	Histone 3 lysine 4 monomethylation
FcRn	Neonatal Fc receptor	H3K4me2	Histone 3 lysine 4 dimethylation
FGOs	Fully grown oocytes	H3K4me3	Histone 3 lysine 4 trimethylation
FGR	Foetal growth restriction	H3K56me3	Histone 3 lysine 56 trimethylation
FISH	Fluorescent in situ hybridisation	H3K64me3	Histone 3 lysine 64 trimethylation
GoS2	Go/G1 Switch Regulatory Protein 2	H3K9ac	Histone 3 lysine 9 acetylation
GATK	Genome Analysis Toolkit	H3K9me1	Histone 3 lysine 9 monomethylation
GDM	Gestational diabetes mellitus	H3K9me2	Histone 3 lysine 9 dimethylation
gDMR	Germline differentially methylated regions	H3K9me3	Histone 3 lysine 9 trimethylation
gDNA	Genomic DNA	H4K20me3	Histone 4 lysine 20 trimethylation
GDP	Guanosine diphosphate	HBEGF	Heparin-binding EGF-like growth factor
GE	Glandular epithelium	HBs	Hofbauer cells
GEO	Gene Expression Omnibus	hCG	Human chorionic gonadotrophin
GG	Gynogenetic	hESCs	Human embryonic stem cells
GLUT4	Glucose transporter type 4	HESCs	Human endometrial stromal cells
GTP	Guanosine triphosphate	hiPSCs	Human induced pluripotent stem cells
GV	Germinal vesicle	hiPSCs	Human induced pluripotent stem cells
HMDs	Highly methylated domains	LE	Luminal epithelium

HMs	Human androgenetic hydatidiform moles	liDNase-seq	Low-input DNase I-sequencing
HP1	Heterochromatin protein 1	LINEs	Long interspersed nuclear elements
HSL	Hormone-sensitive lipase	lncRNAs	Long non-coding RNAs
hTSCs	Human trophoblast stem cells	LOI	Loss of imprinting
HX1	HuaXia1	LPL	Lipoprotein lipase
IC1	Imprinting control centre 1 of <i>H19/IGF2</i>	LSH	Lymphocyte-specific helicase
ICM	Inner cell mass	LTRs	Long terminal repeats
ICRs	Imprinting control regions	MACS	Magnetic-Activated Cell Sorting
ICSI	Intracytoplasmic sperm injection	MAF	Minor allele frequency
iEVTs	Interstitial extravillous trophoblast cells	matKO	Maternal knockout
IGF1R	Insulin-like growth factor 1 receptor	MBD	Methyl-CpG-binding domains
IGF2R	Insulin-like growth factor 2 receptor	MCscan	Multiple collinearity scan
iPSCs	Induced pluripotent stem cells	MEC	Maternal endothelial cells
IR	Insulin receptor	MeCP2	Methyl-CpG-binding protein 2
IRS	Insulin receptor substrate	mESCs	Mouse embryonic stem cells
IUGR	Intrauterine growth restriction	MGL	Monoacylglycerol lipase
IUI	Intrauterine insemination	MII	Metaphase II
IVF	In vitro fertilisation	miRNA	microRNA
K	Lysine	mTORC1	Mammalian target of rapamycin complex 1
KAP1	KRAB domain-associated protein 1	mTORC2	Mammalian target of rapamycin complex 2
KO	Knockout	mTSCs	Mouse trophoblast stem cells
KOS	Kagami-Ogata syndrome	MZT	Maternal-to-zygote transition
KRAB-ZFPs	Krüppel-associated box domain zinc-finger proteins	N6-Mtases	DNA-(adenine N6)-methyltransferase
KvDMR1	<i>Kcnq1/KCNQ1</i> differentially methylated region 1	ncRNA	Non-coding RNA
LAC	Lacunae	NKs	Natural killer cells

OHSS	Ovarian hyperstimulation syndrome	PWS	Prader-Willi syndrome
PADs	Polycomb-associated domains	qRT-PCR	Quantitative real-time reverse transcription PCR
PBAT	Post-bisulphite adaptor tagging	RFTS	Replication foci targeting sequence
PCGF	Polycomb group ring-finger protein	RNAi	RNA interference
PE	Pre-eclampsia	RNA-seq	RNA sequencing
pESCs	Parthenogenetic embryonic stem cells	RSA	Recurrent spontaneous abortions
PG	Parthenogenetic	rTSCs	Rat trophoblast stem cells
PGCs	Primordial germ cells	SAM	S-Adenosylmethionine
PHMs	Partial hydatidiform moles	scBS-seq	Single-cell bisulphite sequencing
PI3K	Phosphatidylinositol 3-kinase	scM&T-seq	Single-cell methylation and transcriptome sequencing
<i>PIK3R1</i>	Phosphoinositide-3-Kinase Regulatory Subunit 1	SCNT	Somatic cell nuclear transfer
PIP2	Phosphatidylinositol (4,5)-bisphosphate	scRNA-seq	Single cell RNA sequencing
PIP3	Phosphatidylinositol (3,4,5)-trisphosphate	sDMRs	Secondary or somatic differentially methylated regions
PlGF	Placental growth factor	SGA	Small for gestational age
PMDs	Partially methylated domains	SHORT	Short stature, Hyperextensibility of joints/hernia, Ocular depression, Rieger anomaly and Teething delay
PRC1	Polycomb repressive complex 1	SINEs	Short interspersed nuclear elements
PRC2	Polycomb repressive complex 2	SMART-seq2	Switching Mechanism At the end of the 5'-end of the RNA Template sequencing 2
PrE	Primitive endoderm	SMC	Smooth muscle cells
<i>PSG</i>	Pregnancy-specific glycoprotein gene cluster	SNPs	Single-nucleotide polymorphisms
PSGs	Prespermatogonia cells	SRA	SET and RING finger-associated domain
PTMs	Post-translational modifications	SRS	Silver-Russell syndrome
PVP	Polyvinylpyrrolidone	STBs	Syncytiotrophoblasts

STR	Short Tandem Repeat	ZGA	Zygotic genome activation
STRT-seq	Single-cell tagged reverse transcription sequencing	ZNFs	Zinc finger proteins
SVA	SINE-VNTR-Alu retrotransposons		
T2DM	Type 2 diabetes mellitus		
TADs	Topologically associated domains		
TDG	Thymine DNA glycosylase		
tDMRs	Transient differentially methylated regions		
TE	Trophectoderm		
TET	Ten-Eleven-Translocation		
TFs	Transcription factors		
TGs	Triacylglycerols		
TNDM	Transient neonatal diabetes mellitus		
TS14	Temple syndrome		
TSSs	Transcription start sites		
TTD-PHD	Tudor and Plant Homeodomain		
UBL	Ubiquitin-like		
UHRF1	Ubiquitin-like plant homeodomain and RING finger domain 1		
uNK	Uterine natural killer cells		
UPW	UltraPure™ DEPC-Treated Water		
WES	Whole-exome sequencing		
WGBS	Whole-genome bisulphite sequencing		
WGS	Whole-genome sequencing		
WHO	World Health Organization		
XCI	X chromosome inactivation		
Xp	Paternal X chromosome		
YS	Yolk sac		

# Declaration

I declare that the work presented in this thesis has not been submitted to support any other application for a degree in this or another learning institution. The work conducted by other authors has been fully cited and referenced.

Some portions of this thesis were used for co-authored publications:

Daskeviciute D, Chappell-Maor L, Sainty B, Arnaud P, Iglesias-Platas I, Simon C, et al. Non-canonical imprinting, manifesting as post-fertilization placenta-specific parent-of-origin dependent methylation, is not conserved in humans. *Hum Mol Genet* [Internet]. 2025 Mar 20;34(7):626–38. Available from: <https://academic.oup.com/hmg/article/34/7/626/7959540>

Daskeviciute D, Sainty B, Chappell-Maor L, Bone C, Russell S, Iglesias-Platas I, et al. PIK3R1 and GOS2 are human placenta-specific imprinted genes associated with germline-inherited maternal DNA methylation. *Epigenetics* [Internet]. 2025 Dec 31;20(1). Available from: <https://www.tandfonline.com/doi/full/10.1080/15592294.2025.2523191>

# Acknowledgments

I would like to thank Prof David Monk for the opportunity to carry out my PhD project in his group. Thank you for your support, guidance, and patience during these four years. I would also like to thank my secondary project supervisor, Dr Wilfried Haerty, for his insights and valuable feedback throughout this project.

A special thanks to Dr Louise Chappell-Maor for training me in the laboratory and for answering all the burning questions that arose during this project. You are brilliant! I also appreciate your help in generating the scRNA-seq datasets of human embryos.

I would like to thank Dr Claudia Buhigas for her assistance with bioinformatic analyses and for the insightful conversations we shared.

I would like to thank Dr Iain Macaulay and his group members for allowing Dr Louise Chappell-Maor and me to use their laboratory equipment and facilities to generate the scRNA-seq datasets of human pre-implantation embryos.

I also want to thank past and present members of the Monk group, including Dr Sarah Russell, Caitlin Bone and Dr Leighton Folkes, for their help during this PhD project, as well as for the interesting and exciting conversations we had.

A special thanks to fellow Monk group students Becky Sainty and Kelly Chen for the stimulating discussions about science and beyond, as well as for their support both during this PhD journey and outside of it.

I would also like to thank Prof Andrea Münsterberg for giving me the opportunity to present my work during the developmental biology group lab meetings.

Furthermore, I would like to extend my gratitude to the other PhD students at BMRC and beyond for their inspiring and interesting conversations.

Finally, I want to express my deepest gratitude to my family for their continuous support, and especially to my partner for his patience and constant encouragement. I would not have made it without you.

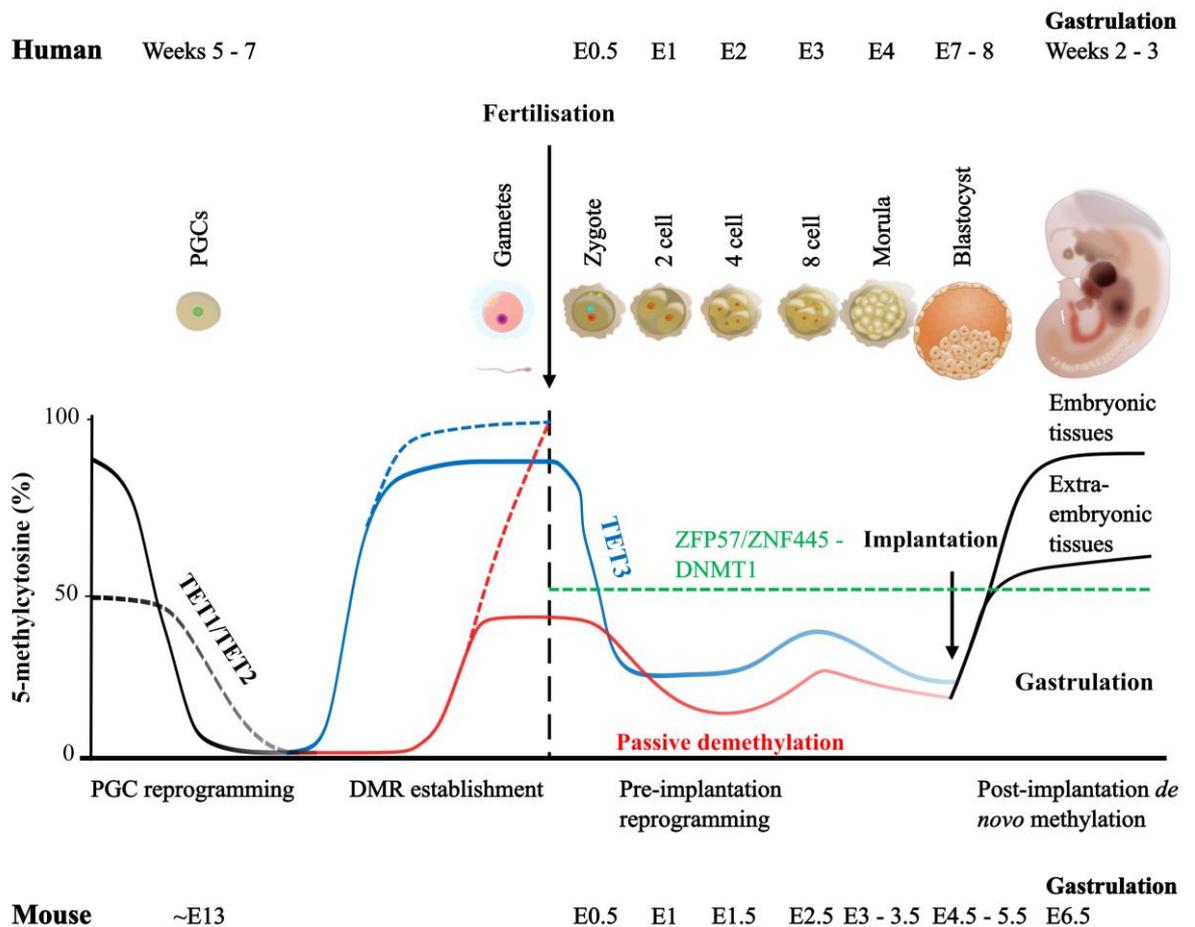
Last but not least, I am grateful to my beloved dog, Kyuubi, for patiently waiting for me to return from work and for lifting my spirits during challenging times.



# **Chapter 1: Introduction**

Epigenetic modifications such as DNA methylation and histone modifications play a key role in regulating gene expression, lineage commitment, cellular differentiation and maintenance of genome stability (1,2). These epigenetic marks are modifications of DNA molecules and associated proteins that do not alter the nucleotide sequence (3). Recent advances in single-cell sequencing technologies provided new insights into the human epigenetic landscape, which is highly dynamic and undergoes dramatic changes during gametogenesis and early embryogenesis (4).

Following fertilisation, the sperm nucleus is rapidly decondensed, and all protamines are replaced with canonical histone variants derived from the oocyte (5,6). Within a few hours of fertilisation, both maternal and paternal pronuclei fuse and form a diploid zygote through a process known as syngamy (7). Around this time, parental genomes present distinct epigenetic landscapes with global DNA methylation levels slightly higher in sperm than in the oocyte. These differences become largely equalised during epigenetic reprogramming, ensuring a totipotent state crucial for embryo development (4,8,9) (**Figure 1.1**). Concomitantly, maternal transcripts accumulated during oogenesis are gradually depleted, which stimulates transcription from the embryonic genome in a process known as embryonic genome activation (EGA), detectable at the 4- to 8-cell stage in human embryos (10–12). All these processes define the maternal-to-zygote transition (MZT)(12).



**Figure 1.1. Changes in global DNA methylation levels throughout the lifecycle of imprinted genes.**

Following fertilisation, the maternal genome (red line) undergoes passive demethylation, while the paternal genome (blue line) is actively demethylated by the Ten-Eleven-Translocation 3 (TET3) enzyme, resulting in the lowest DNA methylation levels at the blastocyst stage. After implantation, *de novo* methylation is established by DNA methyltransferases (DNMTs), and the genome of the post-implantation embryo becomes gradually hypermethylated (black line). However, imprinted regions (dashed green line) are protected from demethylation by the ZFP57/ZNF445-DNMT1 complex and such regions maintain approximately 50% methylation throughout the organism's life. Primordial germ cells (PGCs) of the post-implantation embryo undergo genome-wide epigenetic reprogramming driven by passive and active demethylation mediated by TET1 and TET2 (black line). Parent-specific imprints experience slower reprogramming (dashed black line). In males, new methylation patterns at germline differentially methylated regions (gDMR) or paternal imprinting control regions (ICRs) (paternal ICRs, dashed blue line; whole genome, blue line) are established earlier, while female-specific methylation at gDMRs/ICRs (maternal ICRs, dashed red line; whole genome, red line) is fully deposited by the time oocytes reach metaphase II (MII). These parent-specific methylation marks are established by *de novo* DNMTs. Light blue circle – paternal pronucleus, red circle – maternal pronucleus, orange circle – embryonic nucleus.

Immediately prior to embryo implantation (early blastocyst stage), the parental genomes reach their lowest DNA methylation levels (**Figure 1**) (4,8,9). This is followed by progressive global remethylation observed in post-implantation embryos, leading to a

gradual loss of cellular potency as the embryo has undergone committed differentiation, resulting in the formation of the epiblast (Epi), primitive endoderm (PrE) and trophoctoderm (TE) (13,14). The Epi will give rise to the embryo proper, while the PrE and TE will contribute to extra-embryonic tissues, including the placenta. Once cell-type-specific epigenetic marks are established, they are steadily maintained throughout the life course of an adult, with some variation observed during human ageing (15). The exception to this is primordial germ cells (PGCs) located in the gonadal ridges of the developing embryo (16). These cells must undergo genome-wide epigenetic reprogramming so that new sex-specific epigenetic marks can be established in the developing oocyte and sperm, which will be passed on to the succeeding generation (**Figure 1**) (7,17).

In nature, nearly all sexually reproducing diploid organisms, including humans, inherit two alleles of a gene, one from each parent, and demonstrate biallelic expression. Genomic imprinting is an epigenetic phenomenon that results in a subset of genes being monoallelically expressed based on their parent-of-origin. Therefore, these genes bypass classical Mendelian inheritance laws (7,17,18). Imprinted genes usually exist in clusters, known as imprinting domains, which are regulated by imprinting control regions (ICRs), i.e., genomic regions, which demonstrate allele-specific DNA methylation derived from either maternal or paternal chromosomes. These differentially methylated regions (DMRs), found at ICRs, are acquired during gametogenesis. ICRs coordinate the expression of proximal genes in a way that one allele will be expressed while the other allele will be permanently silenced or imprinted. Genomic imprints are not affected by the epigenetic reprogramming event during pre-implantation development and are maintained throughout an organism's lifespan, but they are erased in PGCs to set new sex-specific imprints (**Figure 1**). Thus, genomic imprints are representative examples of intergenerational epigenetic inheritance (19).

Recent studies have shown that some genes may demonstrate transient imprinting, which exists temporarily until the blastocyst implantation in human embryos (20–23). These genes are enriched with maternal germline differentially methylated regions (gDMRs) inherited from the oocyte and act as ICRs that coordinate paternal allele-biased expression. Such transiently imprinted regions survive pre-implantation reprogramming but are mainly lost in the post-implantation embryo by gaining or losing DNA methylation on one of the parental alleles. Several studies have reported that some of these transiently imprinted genes maintain their imprinting marks in the human placenta (20–22). Moreover, it was recently noticed that histone modifications, such as histone 3 lysine 27 trimethylation (H3K27me3), could mediate imprinted monoallelic expression in mouse morulae (24–26). These 'non-canonical' imprinting marks are deposited in oocytes and persist after fertilisation, resulting in paternal expression. However, these marks are

restricted to pre-implantation stages. Interestingly, it was found that a few genes maintain paternal-specific expression in extra-embryonic lineages. Additionally, there is some promising evidence that this DNA methylation-independent imprinting mechanism may exist in human embryos (27). To date, it remains unclear how many transiently imprinted genes and other monoallelically expressed transcripts exist in human embryos. Also, it is currently unknown whether transient imprinting influences the early embryonic transcriptome and affects lifelong genome regulation. In this literature review, I will describe the major types of epigenetic modifications, and then I will overview the initial events during human pre-implantation development, followed by the lifecycle of canonical imprinted genes. Finally, I will discuss some of the more recent discoveries in the field of genomic imprinting.

## 1.1. Epigenetic modifications

### 1.1.1. 5-methylcytosine (5mC)

DNA methylation is a stable chemical modification of the DNA nucleotide sequence, which is important for many processes within mammalian cells (28,29). It is vital for the regulation of gene expression and, therefore, for cell lineage specification, X chromosome dosage compensation, and the repression of retroviral elements, and plays a central role in genomic imprinting (1,2,30). A methyl group is attached to the 5-carbon atom of the cytosine ring, forming 5-methylcytosine (5mC). The methylation reaction is carried out by a family of enzymes known as DNA methyltransferases (DNMTs), which use the universal methyl donor, S-Adenosylmethionine (SAM), to transfer the methyl group to DNA (31).

These 5mCs are mainly located upstream of guanine nucleotides and are referred to as CpG sites (CG). CpG sites are frequently clustered in the human genome and form so-called CpG islands, which are rich in C and G nucleotides (GC content  $\geq 50\%$ ) and are typically 200 to 500 bp in length. Other genomic regions, more distant from CpG islands, can also be methylated. Such regions include CpG island shores within 2 kb upstream and downstream of islands (referred to as CpG island north shore and CpG island south shore), CpG island shelves within 2-4 kb upstream and downstream of islands (termed as CpG island north shelf and CpG island south shelf), and finally the sea (regions with low-density CpG sites in the genome) (32–34).

It is estimated that there are over 28 million CpG sites, which constitute approximately 1%

of the human genome (35–37). The majority of CpG islands are located near or within gene promoters or other regulatory elements and are predominantly in the hypomethylated state. Some CpG dinucleotides are found within gene bodies (intragenic) and intergenic regions but at lower densities and are typically hypermethylated. These differences in the CpG dinucleotide distribution across the mammalian genome can be explained by the higher mutation burden observed at methylated CpGs, as 5mC is prone to spontaneous deamination, resulting in a cytosine-to-thymine substitution (CG to TG transition) (38–40). For instance, intragenic CpG islands, compared to CpG sites in other genomic regions, are more rapidly depleted in the human genome due to the spontaneous deamination of 5mC (36,40). Consequently, hypermethylated CpG sites may be gradually lost from intergenic and intragenic regions over evolutionary time in the human genome, while highly populated and unmethylated regions, such as CpG island promoters, are preserved. Moreover, it is suggested that around 75% of mammalian CpG dinucleotides are methylated (41). The highest global methylation levels have been reported in naïve T cells and neurons in the adult human cortex, with more than 85% methylation at CpG sites, whereas bladder smooth muscle and heart fibroblast cells are among the most hypomethylated somatic cell types, with less than 60% methylation (42).

Hypermethylated regions are enriched in satellite DNAs, repetitive elements (centromeric, pericentromeric (43) and sub-telomeric repeats (41,44), transposable elements (45,46)), non-repetitive intergenic DNA, and gene bodies to ensure genome integrity (41). As noted earlier, CpG islands at gene promoters are typically depleted of DNA methylation. Methylation at such CpG islands was found to show a negative correlation with gene transcription (47,48). Therefore, 5mC is generally considered to be a repressive mark, especially when methylated CpG sites are abundant in heterochromatic regions (49).

It is still not fully understood how DNA methylation represses gene transcription, but several mechanisms have been proposed to explain how this DNA modification can directly or indirectly inhibit gene expression. Firstly, some transcription factors (TFs) are sensitive to methylated CpG sites within their binding motifs and, therefore, unable to bind to methylated DNA. A recent study conducted by Yin *et al.* (2017) investigated the binding of 542 human TFs to either methylated or unmethylated DNA. Authors reported that the binding of 117 (23%) TFs, including bHLH-, bZIP- and ETS- families of TFs, was reduced by the presence of 5mC within their binding motifs (50). Conversely, the other 175 (34%) TFs, including multiple homeodomain TFs, such as OCT4, HOXB13, HOXC11 or CDX1, showed a preference for 5mC-containing binding motifs, possibly due to hydrophobic interactions formed between 5mC and amino acids of TFs. Secondly, DNMTs, which establish and maintain DNA methylation in mammals (discussed in more detail in the following **Section 1.1.2**), can interact with repressive chromatin remodelers such as lymphocyte-specific helicase (LSH (51)), as well as other epigenetic modifiers,

including histone 3 lysine 9 (H3K9) methyltransferases and histone deacetylases, all of which are involved in heterochromatin formation (52–54). Finally, proteins containing methyl-CpG-binding domains (MBD), such as MBD1, MBD2, MBD4 and methyl-CpG-binding protein 2 (MeCP2) (55–57), as well as zinc finger proteins (ZNFs) (58), recognise methylated DNA regions and can recruit nucleosome remodelers (59), histone deacetylases (59,60), or H3K9 methyltransferases (53) that promote chromatin condensation and inhibit gene expression (61,62).

### 1.1.2. DNA methyltransferases (DNMTs)

As mentioned previously, DNMTs establish and maintain DNA methylation marks. There are four major DNMTs in humans: DNMT3A, DNMT3B, DNMT3L and DNMT1. New DNA methylation marks are established by the *de novo* DNMT3A and DNMT3B in embryos and germ cells (63,64). A third family member - DNMT3L, is also important for *de novo* methylation (46,65). DNMT3L lacks the N-terminal catalytic domain and, therefore, has no enzymatic function, but it acts as a co-factor and works in combination with DNMT3A and DNMT3B. In mice, DNMT3A and DNMT3L are primarily present in oocytes and early embryos. These two DNMTs are responsible for the establishment of imprints in female and male germ cells. Similarly, the enzymatic function of DNMT3B becomes more important during later stages of development. In humans, it has been suggested that DNMT3B, rather than DNMT3A, plays a more important role during global DNA remethylation in the early blastocysts (66). Also, the *DNMT3L* transcripts were not detected in human oocytes, indicating species-specific differences, but data generated by Monk lab reveals a sharp increase in expression from the morula stage, suggesting DNMT3L may have acquired a role in *de novo* methylation post-implantation (66–68). DNMT1 is the maintenance DNMT, whose primary role is to establish DNA methylation marks on hemimethylated DNA immediately after replication (69,70). It is recruited by the ubiquitin-like plant homeodomain and RING finger domain 1 (UHRF1) protein to replication sites to establish a normal level of methylation on a newly synthesised DNA strand. DNMT1 is initially detected in the nucleus of germinal vesicles (GV), but later, it is removed from the nucleus to the cytoplasm, where it remains throughout pre-implantation development (66–68,71). Several studies have demonstrated that mouse mutants carrying mutations in any of these DNMTs die at early stages of pregnancy, illustrating the importance of these enzymes for development (46,64,65,72).

### 1.1.3. DNA demethylation

DNA methylation marks can also be removed from the DNA strand. This process is essential for genome-wide epigenetic reprogramming events (7,17). DNA can be demethylated in two ways: passive and active demethylation. During passive demethylation, *de novo* methylation marks are not established on a newly synthesised DNA strand, resulting in replication-coupled dilution of DNA methylation. The active demethylation process is carried out by a family of enzymes known as Ten-Eleven-Translocation (TET) proteins, which include TET1, TET2 and TET3 (73,74). 5mC is oxidized by TET proteins, leading to the formation of 5-hydroxymethylcytosine (5hmC) (75), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) intermediates. Such intermediates then can be removed either passively, by replication-dependent demethylation, or actively, by thymine DNA glycosylase (TDG) via base excision repair (BER) (74). It has been proposed that these 5mC intermediates may possess a regulatory function. For instance, a higher accumulation of 5hmC was discovered in neuronal tissues and embryonic stem cells (ESCs) (76,77). In ESCs, this mark is enriched in active distal regulatory regions, particularly enhancers (52), whereas in nervous tissue, it is located within neuron-specific gene bodies (78). Similarly, in mouse embryonic stem cells (mESCs), 5fC and 5caC densities were shown to be the highest at poised enhancers (74,79,80). Also, a positive correlation was identified between the binding frequency of transcriptional coactivator p300 and the density of 5fC and 5caC marks (79). These findings indicate that 5mC intermediates may play an active role in the demethylation process by recruiting TFs and other proteins that can interact with p300, and thus, these marks may indirectly initiate transcription.

### 1.1.4. Non-CG methylation

For a long time, it was assumed that cytosine is the only nucleotide that can be modified with a methyl group. Recently, it has been shown that the N-6 position of adenine can also be methylated and is referred to as N6-methyladenine (6mA) (81). The role of 6mA modification in the mammalian genome is currently unclear, but it has been shown to be abundant in the human genome. One study detected 881,240 6mA sites, accounting for 0.051% of total adenines, with levels ranging from 0.023% to 0.064% across all human chromosomes and the highest level observed in the mitochondrial genome (0.184%) (82). However, this mark was found to be more prevalent in other kingdoms, including fungi (budding yeast- *Saccharomyces cerevisiae*) and plants (*Arabidopsis thaliana*) (83,84). In

one study, mESCs were used to investigate the function of 6mA (85). The authors provided some evidence that 6mA is enriched at the 5' UTR and ORF1 regions of young (<1.5 million years old) but not old (>6 million years old) long interspersed nuclear element 1 (LINE-1) transposons located on the X chromosome. Additionally, a negative correlation was detected between the expression of genes near young L1 elements and 6mA deposition. Thus suggesting that 6mA plays a protective role against young retrotransposons. In contrast, other study used HuaXia1 (HX1) human blood cells and found that 6mA is abundant around exonic regions and is positively correlated with gene expression. The authors concluded that 6mA is an active gene signature in human cells (82).

Instead of trying to detect endogenous levels of 6mA in the human genome, Broche and colleagues chose to artificially introduce two bacterial DNA-(adenine N6)-methyltransferase (N6-MTases), such as EcoDam and CcrM into human HEK293 cells to induce high levels of this modification (86). These enzymes show a high affinity for the CATC and GANTC motifs in DNA. It was found that the expression of these N6-MTases had an additive effect, leading to elevated levels of 6mA, which in turn caused a significant decrease in cell viability. Further analysis of these cells revealed changes in the expression of 99 genes (66 upregulated and 33 downregulated). Upregulated genes exhibited reduced levels of H3K27me3 (repressive histone mark; discussed in greater detail later in this literature review), particularly at the *CUX2* and *PAPPA* genes, suggesting that 6mA might inhibit the binding of Polycomb Repressive Complex 2 (PRC2 catalyses the methylation of H3K27). In contrast, downregulated genes such as *EGF* and *EMILIN2* were found to be enriched for JUN family TF binding sites, indicating that 6mA within JUN TF binding motifs may interfere with TF binding to their target genes, especially since the expression of these genes returned to normal after 6mA depletion. Overall, this study (86) suggests that a 6mA modification in the human genome can affect two molecular pathways, which can further alter downstream gene expression, leading to phenotypic changes in human cells.

Methyl groups can be attached to several other nucleotide sequences of DNA, including mCpHpG and mCpHpH, where H represents A, C, or T nucleotides. Non-CpG methylation has been observed in oocytes, induced pluripotent stem cells (iPSCs) and neurons, and it has been suggested as a hallmark of ESCs (8,9,41,87–91). Among the different non-CpG methylation marks detected across various cell types, mCpA is the most abundant in human oocytes (mean CpA methylation = 5.6%, mean CpG methylation = 53%) (8), neurons of the adult human cortex (CpA methylation = ~10% and CpG methylation = ~84%) (42) or in ESCs (41), whereas methylation at other non-CpG sites is relatively rare (41,42,91), particularly in somatic cells. The functional relevance of non-CpG methylation

in mammals remains unclear. Some argue that it may be a by-product of hyperactive *de novo* DNMTs, as these modifications lack the symmetry required for maintenance by DNMT1 (62,92,93). Others propose that non-CpG methylation plays a crucial role in maintaining pluripotency and regulating lineage-specific gene expression (90,94,95).

Lister *et al.* (2009) (41) and Ziller *et al.* (2011) (91) reported that non-CpG methylation is common in human embryonic stem cells (hESCs), where approximately 25% of all methylated cytosines occur at non-CpG sites. They also found that this methylation pattern is lost upon differentiation, potentially due to the global downregulation of *de novo* DNMTs. This was further supported by the deletion of either DNMT3A or DNMT3B, which led to a genome-wide loss of non-CpG methylation in hESCs, indicating that *de novo* DNMTs are required for non-CpG methylation (91). Ziller *et al.* (2011) observed that mCpA distribution was positively correlated with CpG methylation and was slightly more abundant in introns and short interspersed nuclear elements (SINEs) in human iPSCs and hESCs (91). In contrast, other non-CpG methylation marks were randomly distributed across different genomic regions. Similarly, Lister *et al.* (2009) showed that non-CpG methylation was depleted at TF binding sites and enhancers but enriched in gene bodies, suggesting that mCpH may play an important role in maintaining pluripotency (41).

Guo and colleagues found that mCpH (~25% of all methylated cytosines) was abundant in granule neurons derived from the adult mouse dentate gyrus but absent in the mouse spleen (90). Many mCpH-rich regions identified in mouse neurons had orthologs in the human brain, showing the same enrichment for mCpH. Additionally, mCpH levels gradually increased during neuronal maturation in both the mouse and human brain. The same study reported that mCpH located away from CpG sites was negatively associated with proximal gene expression, suggesting a role in gene repression in the mouse brain. Furthermore, MeCP2, a protein highly expressed and important in the brain, was found to interact with mCpH (90). The presence of both mCpG and mCpH enhances MeCP2 binding, which can recruit histone deacetylases and other complexes, further supporting the role of non-CpG methylation in gene repression in neurons (56,57,90). For instance, in type 2 diabetes mellitus (T2DM) patients, the major promoter of the peroxisome proliferator-activated receptor  $\gamma$  coactivator 1 $\alpha$  (*PGC-1 $\alpha$* ) gene exhibited higher non-CpG methylation than CpG methylation in skeletal muscle compared to healthy individuals (96). This increase in non-CpG methylation was associated with *PGC-1 $\alpha$*  downregulation, which, in turn, was linked to reduced mitochondrial density and mitochondrial dysfunction in T2DM patients. Therefore, mCpH may play an important role in gene repression.

### 1.1.5. Histone proteins

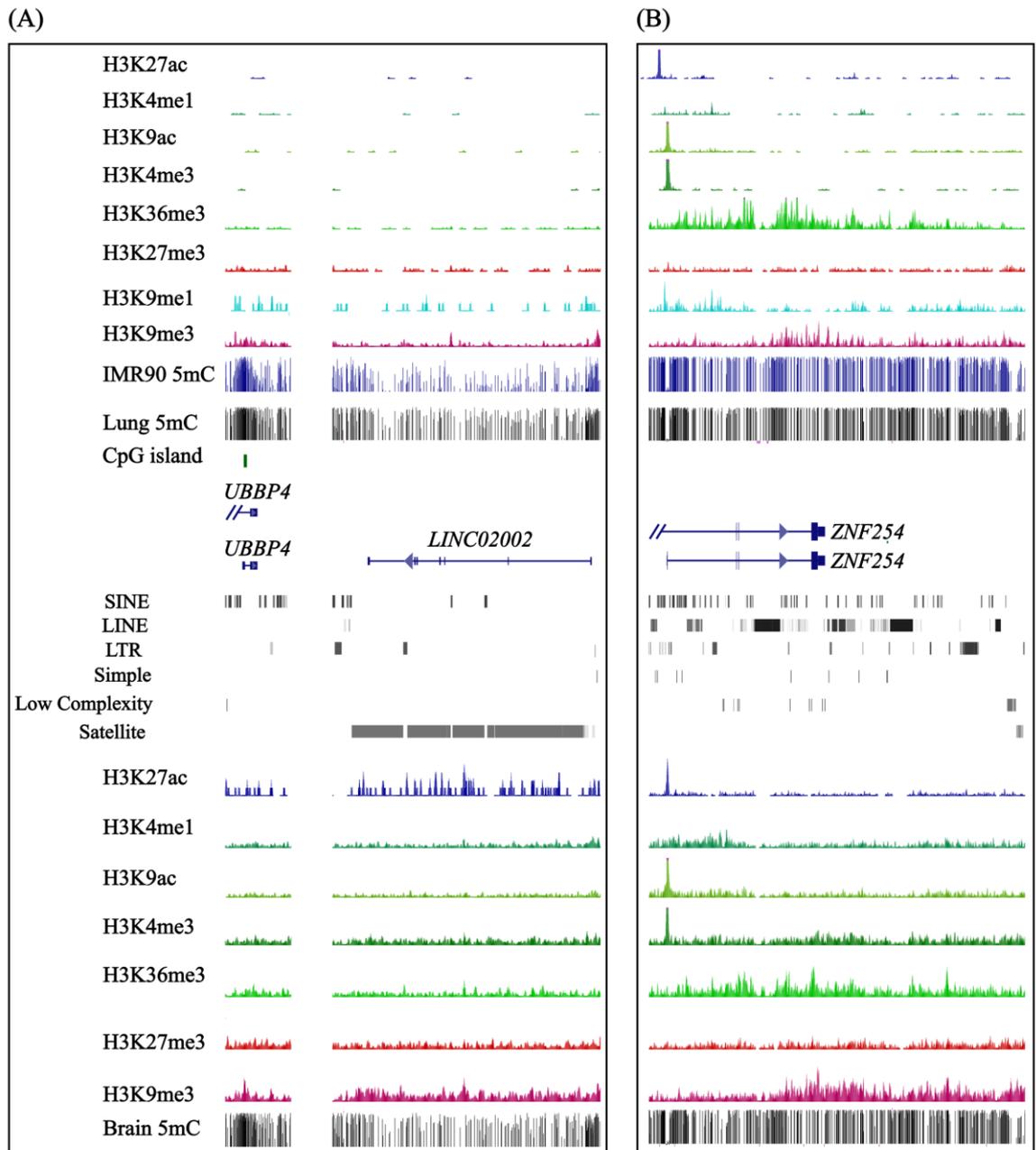
To fit around a 2-meter length of DNA into the tiny nucleus of a mammalian cell, DNA must be tightly packed, which is achieved through a series of steps. At the smallest scale, a DNA strand (~146 bp) is tightly coiled 1.75 times around histone octamers (97) and forms nucleosome core particles, which, together with the linker heterochromatic adaptor protein HP1, form a chromatin fiber often characterised as “beads on a string” (98).

Histones are small proteins (approximately 100-140 amino acid residues) highly conserved between eukaryotes (62). The negatively charged DNA double-helix is tightly wound around canonical histone proteins that contain many positively charged arginine (A) and lysine (K) residues, which help to pack a large DNA macromolecule into a tiny cell nucleus (approximately 10- 23  $\mu\text{m}$  in diameter) (63,64). There are four core histones: H2A, H2B, H3 and H4. Two copies of each histone are assembled in a so-called octamer structure onto which 146-147 bp of the DNA is wrapped (62).

Histone proteins can be modified with either covalent or non-covalent post-translational modifications (PTMs). Here, I will focus on several major covalent modifications catalysed by enzymes, generally referred to as writers (99,100). Histones have variable-length protruding N-terminal tails that are subject to a plethora of PTMs (review in (99,100)), which play an important role in regulating chromatin accessibility. These PTMs can be instructive and change chromatin conformation by recruiting reader proteins (99–102) such as CHD1, BAHD1 and UHRF1, or they can be established as a consequence of other cellular processes, such as transcription (103). In some cases, large multi-subunit writer complexes contain reader domains capable of recognising their histone marks, thereby reinforcing the deposition of the same modification. Such a positive feedback loop is important for constitutive heterochromatin formation (highly condensed genomic regions that are gene-poor, contain HP1, and are mainly composed of repetitive elements such as telomeric and pericentromeric repeats (**Figure 1.2**) (104), and are also characterised by late replication) or X chromosome inactivation (XCI) (105,106). For instance, SUV39H1 and SUV39H2, histone methyltransferases that establish histone 3 lysine 9 trimethylation (H3K9me<sub>3</sub>; discussed in more detail in **Section 1.1.8.1**), contain a chromodomain (CBX) that recognises H3K9me<sub>3</sub> and promotes the spreading of this mark at pericentromeric regions found in constitutive heterochromatin (**Figure 1.2**) (107–110). In addition, the globular core domains of histones can also be decorated with several modifications that are suggested to have a more pronounced effect on chromatin conformation or DNA accessibility (101). These modifications can interfere with histone-DNA interactions, leading to nucleosome destabilisation. Overall, histone PTMs are highly dynamic, as they can be removed by eraser proteins and re-established by writers depending on cellular

needs (99,100). However, this turnover seen in histone PTMs has to be delicately balanced, as imbalances in some PTMs can lead to various diseases, such as cancer (111).

Chromatin can further form loops involving genes with enhancers that are bound by cohesin rings and the CCCTC-binding factor (CTCF) (112–115). Such loops can create self-interacting topologically associated domains (TADs) bordered by CTCF (112). Within a TAD, genomic regions can physically interact, whereas interactions between genomic regions from different TADs (inter-TAD contacts) are less likely to occur. At the higher organisational level, multiple TADs can be partitioned into “A” and “B” compartments that occupy different regions in the nucleus (116,117). The “A” compartment includes euchromatin, which is decorated by active histone PTMs and is usually located closer to the centre of the nucleus. At the same time, the “B” compartment is present at the nuclear periphery and contains heterochromatin harbouring repressive histone PTMs. Finally, all chromosomes have their territories in the cell nucleus (116,118).



**Figure 1.2. Genes located in repeat-rich regions near centromeres.**

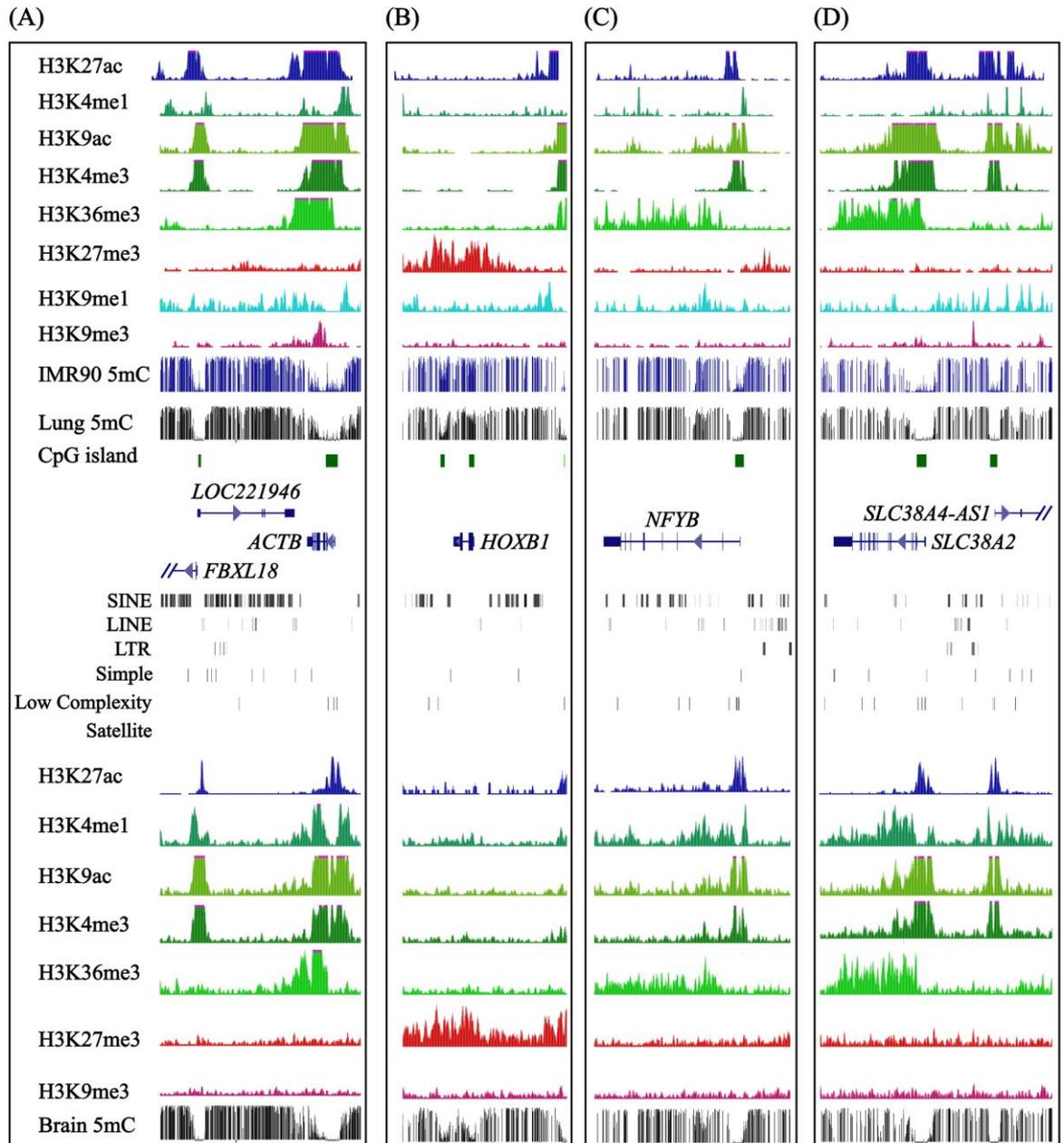
(A) Repressed *UBBP4*, located between satellite repeats on chromosome 17 (119). (B) *ZNF254*, located near the centromere of chromosome 19 in a repeat-rich region containing short interspersed nuclear elements (SINEs), long interspersed nuclear elements (LINEs), long terminal repeats (LTRs) and satellite repeats. Upper histone ChIP-seq tracks represent data from IMR90 cells, while lower tracks are from the brain hippocampus. Vertical lines in the DNA methylation tracks indicate the mean methylation levels at individual CpG dinucleotides. H3K9ac marks transcriptionally active chromatin and is enriched at active promoters. H3K9me1, found at intergenic regions and occasionally overlapping promoters, acts as a substrate for SUV39H1/2 and is therefore considered a repressive histone PTM. Histone modification and IMR90 methylation data were obtained from the NIH Roadmap Epigenomics Consortium (120), while lung and brain methylation tracks were generated by Monk group and are described in **Section 2.7**.

### 1.1.6. Histone post-translation modifications (PTMs) and their role in gene regulation

Acetylation of lysine (K) residues present in histones can reduce the positive charge of K residues, and therefore, the chromatin can become more accessible for TFs or other accessory proteins (121–123). Thus, several acetylated histone residues are often found at active gene promoters and enhancers (**Figure 1.3A, C, D**). It is also believed that this modification acts in an accumulative fashion, as removing a single acetyl group mildly affects transcription (124).

Methylation can occur at K and A residues of histones, and one, two or three methyl groups can be transferred to these residues (100,125). The effect of this modification is highly context-dependent, as it can be repressive and help establish heterochromatin (**Figure 1.2**) or have an activating effect and promote gene transcription (**Figure 1.3A, C, D**).

Other modifications can also be observed on histone proteins, such as crotonylation (126) and phosphorylation (127), which has a similar effect as acetylation, or lactylation (128), which are all associated with active gene transcription. Glutarylation was shown to play a role in chromatin accessibility (129), while ubiquitination is involved in repressing developmental genes (130–132).



**Figure 1.3. Distributions of diverse permissive and repressive histone marks in IMR90 cells and human hippocampus.**

(A) Housekeeping  $\beta$ -actin gene (*ACTB*); (B) Developmental gene (*HOXB1*) repressed by the polycomb repressive complex 2 (PRC2); (C) *NFYB* - a transcription factor highly expressed in IMR90 cells, according to Harmonizome (133). (D) *SLC38A4* - a gene expressed in the brain and associated with Alzheimer's disease (134). Upper histone ChIP-seq tracks represent data from IMR90 cells, while lower histone ChIP-seq tracks are from the brain hippocampus. Vertical lines in the DNA methylation tracks represent the mean methylation levels at individual CpG dinucleotides. H3K9ac marks transcriptionally active chromatin and is enriched at active promoters. H3K9me1, found at intergenic regions and occasionally overlapping promoters, acts as a substrate for SUV39H1/2 and is therefore considered a repressive histone PTM. Histone modification and IMR90 methylation tracks were obtained from the NIH Roadmap Epigenomics Consortium (120), while lung and brain methylation tracks were generated by Monk group and are described in Section 2.7.

## 1.1.7. Permissive histone marks

### 1.1.7.1. Histone 3 lysine 4 trimethylation (H3K4me3)

This modification is written down by SETD1A and SETD1B and established by KMT2B, KMT2C, KMT2D, KMT2A and ASH1L in humans (100,111,135). It is often found at gene promoters near transcription start sites (TSSs) and forms narrow domains (**Figure 1.3A, C, D**) that can be captured by Chromatin Immunoprecipitation followed by sequencing (ChIP-seq), Cleavage Under Targets and Release Using Nuclease (CUT&RUN) or Cleavage Under Targets and Tagmentation (CUT&Tag) (136–138). However, broad non-canonical domains can be found in early human embryos and especially in early mouse pre-implantation embryos, which are reduced to canonical domains by the late two-cell stage (136,137,139–142). The breadth or width of the peak is associated with the strength of transcription, as shown in human placental trophoblasts, where broader histone 3 lysine 4 trimethylation (H3K4me3) domains were responsible for higher trophoblast-specific gene expression (143). Therefore, this modification is important for gene transcription.

Although it is believed that the H3K4me3 modification itself does not instruct gene transcription, but rather it can recruit readers or other proteins that bring transcription-required machinery (122,141). In ESCs, this modification is prevalent at silent gene promoters together with other histone modifications, such as H3K27me3 (termed bivalent domains) (142,144). Such promoters are poised and can be easily activated for expression. Moreover, this modification is found in CpG-rich regions such as CpG islands, which comprise a large proportion of all mammalian gene promoters (48,145). Thus, it is suggested that H3K4me3 protects the CpG islands from DNA methylation, which is highly mutagenic (40). Methylation at lysine 4 residue inhibits the binding of the ADD domain found in *de novo* DNMTs (146). As a result, many regions decorated by H3K4me3 in the male mouse germline remain hypomethylated compared to other regions that are hypermethylated (140,147). In mature mouse oocytes, broad non-canonical H3K4me3 domains (137,139,140) have to be removed prior to zygotic genome activation (ZGA), as it can impair the activity of *de novo* DNMTs (148), which is vital during pre-implantation development. It was found that KMT2B is responsible for establishing non-canonical H3K4me3 peaks in developing mouse oocytes, as the loss of *Kmt2b* led to an 80% decrease in H3K4me3 and a complete loss of non-canonical H3K4me3 domains, while canonical domains remained unaffected (149). In addition, *Kmt2b* knockout (KO) oocytes failed to develop or ovulate fully. These findings suggested that canonical and non-canonical H3K4me3 domains are established by different methyltransferases. Furthermore, canonical and non-canonical H3K4me3 domains observed in mouse oocytes are inherited by the zygote and observed in early pre-implantation embryos. However,

non-canonical H3K4me3 domains are absent by the late 2-cell stage, possibly due to the upregulation of the *Kdm5a* and *Kdm5b* demethylases in 2-cell stage embryos (139,142). Depletion of both KDM5A and KDM5B led to the retention of broad H3K4me3 domains and failure to reach the blastocyst stage due to impaired activation of ZGA genes. After the 2-cell stage, H3K4me3 becomes restricted to canonical regions such as active gene promoters, enhancers and bivalent domains (137,139,140). Therefore, the timely removal of non-canonical H3K4me3 domains before the 2-cell stage is crucial for normal mouse embryonic progression.

#### 1.1.7.2. Histone 3 lysine 4 monomethylation (H3K4me1) and histone 3 lysine 27 acetylation (H3K27ac)

Histone 3 lysine 4 monomethylation (H3K4me1) is catalysed by KMT2C (also known as MLL3) or KMT2D (also known as MLL4), and it is located at active enhancers, as well as in intergenic regions (**Figure 1.3D**) (135,150). Histone 3 lysine 27 acetylation (H3K27ac) is written by histone acetylases, such as p300 and CBP, and many other enzymes, and it marks active enhancers and promoters (**Figure 1.2B, Figure 1.3A, C, D**) (123,135). In human trophoblasts or human trophoblast stem cells (hTSCs), some active ERVs can be decorated by both of these marks and some harbour only one of these marks (138). Both of these histone PTMs can mark super-enhancers (genomic regions populated by several clustering enhancers) (151).

#### 1.1.7.3. Histone 3 lysine 36 di- and tri-methylation (H3K36me2 and H3K36me3)

SETD2 establishes histone 3 lysine 36 trimethylation (H3K36me3), while NSD1-3 can add mono- and di-methylation to lysine 36 of histone 3 in humans and mice (103,135,152). Histone 3 lysine 36 dimethylation (H3K36me2) decorates large intergenic regions (152). On the other hand, H3K36me3 is correlated with gene expression, and therefore, it is found in the bodies of actively transcribed genes (**Figure 1.3A, C, D**). H3K36me3 inhibits transcription from hidden promoters present within genes (“cryptic transcription”) (153) as during transcription elongation (103), it is laid down over gene bodies by SETD2 (SETD2 binds to RNA polymerase II C-terminal domain (CTD)), it can regulate splicing (154) and plays a role in DNA damage repair (155).

H3K36me2/3 are recognised by the PWWP (Pro-Trp-Trp-Pro) domain that is present in *de novo* DNMTs, including DNMT3A and DNMT3B (156,157). Consequently, the bodies of actively transcribed gene bodies are usually hypermethylated (**Figure 1.3C, D**). The

oocytes of *Setd2*-deficient female mice showed not only a global loss of H3K36me3 but also other epigenetic aberrations (158). Genomic regions previously marked by H3K36me3 were overtaken by H3K4me3 and H3K27me3, while other regions gained *de novo* DNA methylation. Interestingly, imprinted ICRs were also decorated by H3K4me3 and remained hypomethylated. Another study investigated the association between *de novo* methylation and H3K36me2, H3K36me3, or both marks in mouse fully grown oocytes (FGOs) (159). The authors found that the loss of H3K36me2 resulted in genome-wide loss of DNA methylation in intermediately methylated regions, especially on the X chromosome, with only slight changes in other histone PTM distributions. On the other hand, *Setd2* KO FGOs, as reported previously, showed a global loss of H3K36me3 and DNA methylation. In *Setd2* KO FGOs, some regions gained methylation at sites lacking H3K36me3, and other regions showed an increase in H3K36me2, while the distribution of H3K27me3 was largely unaffected and H3K4me3 was slightly reduced. The loss of both H3K36me2 and H3K36me3 led to a global DNA methylation loss, resembling the phenotype observed in *Dnmt3a* KO FGOs (160). Based on these findings, the authors concluded that H3K36me2 and H3K36me3 act as a platform guiding DNMT3A-3L in mouse oocytes. Moreover, Shirane *et al.* (2022) explored the link between *de novo* methylation and H3K36me2 and H3K36me3 in mouse male prespermatogonia cells (PSGs) (152). They found that unlike in mouse FGOs, where H3K36me3 established by SETD2 is crucial for directing *de novo* methylation, H3K36me2 catalysed by NSD1 plays a more important role in PSGs. *Nsd1* KO PSGs exhibited a genome-wide decrease in *de novo* DNA methylation, specifically in regions that lost H3K36me2, while H3K36me3 levels were minimally affected. In contrast, *Setd2* deletion led to a significant reduction in H3K36me3, but only mild changes in *de novo* methylation. NSD1 loss also resulted in hypomethylation of pDMRs and the absence of spermatogonia in the adult testicular tubules. Interestingly, in *Nsd1* KO PSGs, regions that lost H3K36me2 gained H3K27me3, leading to gene repression, further suggesting that H3K36me2 might not only shape the *de novo* methylation profile in mouse male gametes but also limit the spread of H3K27me3. Therefore, H3K36me3 and H3K36me2 play vital roles in shaping the sexually dimorphic epigenetic landscapes of mouse gametes, which are important for pre-implantation development.

### 1.1.8. Repressive histone marks

### 1.1.8.1. Histone 3 lysine 9 di- and tri-methylation (H3K9me2, H3K9me3)

This particular modification can be deposited by several histone methyltransferases that are site-specific, such as SETDB1, SUV39H1 and SUV39H2, which can transfer di- and trimethyl groups to lysine 9, while G9A (or EHMT2) and EHMT1 can only add mono- and di-methylation at the same residue (100,135). H3K9me3 is found at constitutive heterochromatin, and it is enriched at repeat-dense regions, such as pericentromeric and telomeric regions composed of satellite repeats (from 5 bp to several hundred bp) (**Figure 1.2A, B**) and ERVs (**Figure 1.2B**) (49). Furthermore, H3K9me3-enriched regions recruit histone deacetylases, followed by the establishment of additional repressive marks such as trimethylation of histone 3 lysines 56 and 64 (H3K56me3, H3K64me3) and trimethylation of histone 4 lysine 20 (H4K20me3) (101). The functions of H3K9me3 and DNA methylation are intertwined, as both of these modifications repress repeat-rich regions (161). SUV39H1/H2 are responsible for establishing H3K9me3 at heterochromatic pericentromeric regions (110) and can also interact with HP1 (107,109,110) and guide DNMT3A or DNMT3B to establish DNA methylation (53,100,162). Additionally, UHRF1 is an important cofactor for DNMT1 and includes several functional domains, such as the SET and RING finger-associated (SRA), Tudor and Plant Homeodomain (TTD-PHD), and the ubiquitin-like (UBL) domain (163–166). The SRA domain is capable of binding to hemimethylated DNA with higher affinity, while the TTD domain can bind to histone 3 lysine 9 dimethylation (H3K9me2) and H3K9me3, enhancing the ubiquitin ligase activity of UHRF1 (164–166). As a result, the RING domain of UHRF1 adds ubiquitin marks to histone 3 at lysine 18 (H3K18ub) and lysine 23 (H3K23ub), which serve as docking sites for DNMT1 (167,168). The UBL domain of UHRF1 interacts with the replication foci targeting sequence (RFTS) domain of DNMT1, which recognises ubiquitinated H3 tails, thereby recruiting DNMT1 to targeted regions. This interaction relieves DNMT1 from its autoinhibited state, allowing for the C-terminus catalytic domain of DNMT1 to methylate hemimethylated CpG sites and thus maintain DNA methylation (169). TFs cannot bind to such hypermethylated heterochromatin regions (**Figure 1.2**) (170), and therefore, such regions remain transcriptionally silenced, which can be found in the “B” compartments at the nuclear periphery (171,172). In both mouse and human PGCs, multiple ERVs are decorated with H3K9me3 (**Figure 1.2B**), and therefore, they remain dormant even when the rest of the genome is in the hypomethylated state (173–175). Moreover, gDMRs or ICRs are enriched with H3K9me3, which protects them from erasure during epigenetic reprogramming in pre-implantation embryos. gDMRs often contain binding motifs for KRAB domain-associated protein 1 (KAP1 or TRIM28) and Krüppel-associated box domain zinc-finger proteins (KRAB-ZFPs), which recruit SETDB1 to establish H3K9me3 at these sites (58,176,177).

### 1.1.8.2. Histone 3 lysine 27 trimethylation (H3K27me3)

This histone PTM is written by PRC2, which can also attach mono- and di-methylation to lysine 27 of histone 3 (178). The three core subunits of PRC2 include either the enhancer of zeste homologue 1 or 2 (EZH1 or EZH2), which catalyses all three forms of H3K27 methylation; embryonic ectoderm development (EED), which acts as a scaffold for other subunits and can bind to H3K27me3 and facilitates the propagation of this mark, and suppressor of zeste 12 (SUZ12), which is required for the regulation of the catalytic EZH1/2 subunit and also aids in chromatin binding (178–182). These three subunits are all essential for methyltransferase activity (178–182). Additional subunits include RBBP4 or RBBP7 (178,183,184). Therefore, based on the combination of subunits incorporated into the complex, two subtypes of PRC2 complexes can be formed, known as PRC2.1 and PRC2.2, which exhibit distinct functions (180,185). Overall, PRC2 can be guided by long non-coding RNAs (lncRNAs) or other accessory proteins to target regions (186). PRC2 has many important functions and is generally considered a repressive mark. It plays a vital role in the repression of developmental genes, such as *HOX* genes (**Figure 1.3B**), after the ZGA or embryonic genome activation (EGA); it is required for establishing heterochromatin, together with H3K4me3 form bivalent domains, has an essential role for XCI and lineage commitment (105,142,144,187). Recently, it has also been shown that it can mediate monoallelic expression in rodent pre-implantation embryos and the placenta, now known as non-canonical imprinting (24,187). In addition, PRC2 binds to CpG-rich regions, and in mammals, it is often found at CpG islands or gene promoters containing CpG islands (188,189) (it can be recruited by BEND3 (190)) that are unmethylated, as H3K27me3 and DNA methylation show an inverse correlation. Different components of PRC2 are frequently mutated in multiple cancers, such as breast or ovarian cancers (178,182). In the mouse inner cell mass (ICM) of the blastocyst, TE-specific genes are repressed by H3K27me3, which prevents ICM differentiation into TE (191). Deletions of PRC2 core component genes in ICM cells lead to their differentiation into PrE (192).

Bivalent domains are considered genomic regions, where H3K4me3 colocalises with H3K27me3 and such regions are located at lineage-specific gene promoters or enhancers during pre-implantation development. In the mouse, such domains are observed at the blastocyst stage (144,193), coinciding with the higher expression of PRC2 genes (194). It is suggested that such domains are required to poise lineage-specific genes from premature expression during earlier stages of mouse development (195) and possibly to protect promoters from hypermethylation (196). Monoallelic bivalent domains have also been found at ICRs when the associated gene is silenced in a tissue-specific manner (197,198).

### 1.1.8.3. Histone 2 A lysine 119 monoubiquitination (H2AK119ub1)

Polycomb repressive complex 1 (PRC1) catalyses this histone modification (199). Like PRC2, the PRC1 complex is built from several subunits, and depending on the incorporated subunits into the final complex, it can be classified into canonical PRC1 and non-canonical or variant PRC1 (199,200). The core components include either the RING1A or RING1B, which have a ubiquitin ligase function and one polycomb group ring-finger protein (PCGF) with several additional subunits. There are six PCGF proteins (PCGF1-6) that are suggested to provide specificity to the PRC1 heterodimer complex and guide the complex to target genomic regions (201,202). The canonical complexes include either PCGF2 or PCGF4 and are called PRC1.2 and PRC1.4, while non-canonical complexes include PCGF1, PCGF3, PCGF5 or PCGF6 and are known as PRC1.1, PRC1.3, PRC1.5 and PRC1.6, respectively (199,200). In the classical model, PRC1 complexes could recognise H3K27me3 through the CBX proteins and establish histone 2 A lysine 119 monoubiquitination (H2AK119ub1) (179,203). After improvements in molecular biology and sequencing technologies that allowed the use of less genetic material to explore histone PTMs, it was found that PRC1 complexes can serve as docking sites for PRC2 complexes (132,204,205). In general, this mark is found at heterochromatin; it represses developmentally important genes, and more recently, it was shown to play an important role during XCI and was also implicated in non-canonical imprinting (131,132,199,206,207).

## 1.2. Epigenetic landscapes of male and female gametes

Before fertilisation, fully differentiated, haploid sperm chromatin fibers are densely packaged into the nucleus. More than tenfold higher condensation in comparison to nucleosome-based chromatin is achieved during the last post-meiotic phase (the spermiogenic phase of spermatogenesis) (5,208). During this time, most nucleosomes are destabilised and eventually displaced by protamines, while DNA becomes hypermethylated so that the resulting chromatin is transcriptionally inactive (5,6,208). Protamines are small proteins (~49 amino acid residues) rich in A residues that are sufficient to neutralise the negatively charged DNA, allowing for tight binding and compaction (209). Such packaging is important for several reasons. Firstly, it makes spermatocytes more hydrodynamic. Therefore, they can move faster through the female

reproductive tract till the oocyte. Secondly, it helps to protect the DNA from physical and chemical damage, especially when sperm is devoid of DNA repair machinery (6).

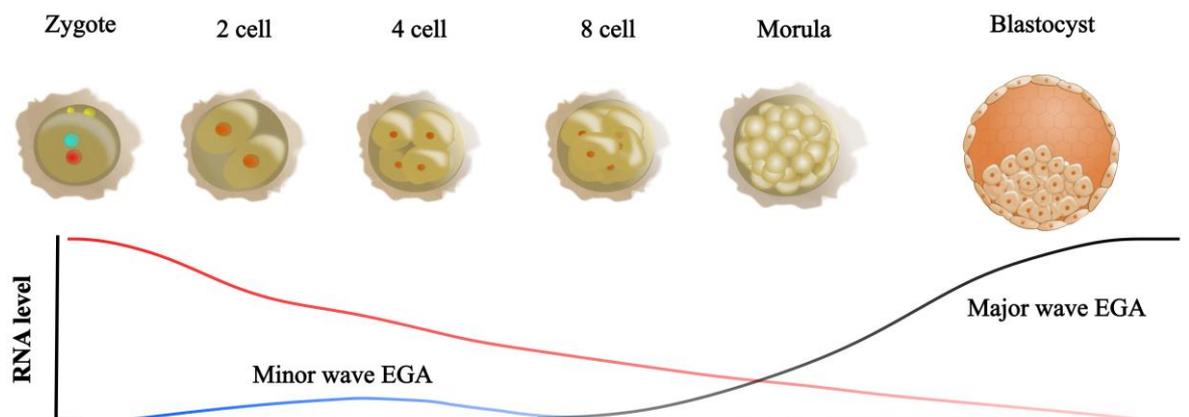
Immediately after fertilisation, sperm DNA starts to decondense and expands, sperm-specific methylation marks are removed, and all protamines are gradually released and replaced by histone variants provided by the oocyte (7). These events lead to a male pronucleus formation, which is overall depleted of epigenetic marks with a few paternal-specific gDMRs remaining. The oocyte and sperm are formed from PGCs that differentiate and mature under different conditions, resulting in asymmetrically distributed epigenetic marks between maternal and paternal genomes that are mostly equalised during pre-implantation development, apart from certain genomic regions. Several studies, in greater detail, examined the methylomes of mouse and human gametes (4,8,9,20,27,88). It was noted that the DNA methylation level of the sperm genome is much higher in comparison to the oocyte (~80% versus ~52%), and the oocyte genome was generally considered to be hypomethylated. In sperm, 5mC densely accumulates in intergenic regions, introns, various repeat and transposable elements, while 5mC is more uniformly distributed in oocytes with a higher level of DNA methylation observed in promoter CpG islands, introns, some repeat elements, and specifically in human oocytes, in tandem repeats within centromeric and pericentromeric regions. Surprisingly, the oocyte contributes a majority of DNA methylation and histone marks to the early embryo, while sperm epigenetic marks are mostly lost (4,8,20,24). For example, 5,438 maternal-specific gDMRs and 48,111 paternal-specific gDMRs were identified in human gametes. After fertilisation, 4,352 maternal-specific gDMRs remained partially methylated in the blastocyst, while only 1% of paternal-specific gDMRs were detected at the same developmental stage (20).

### 1.3. Embryonic genome activation (EGA)

After fertilisation, the paternal and maternal epigenomes must undergo radical changes to ensure the totipotent state required for EGA (17). Uniquely in mice, the first wave or minor wave of transcription appears soon after fertilisation, when the embryo is still at the one-cell stage (known as ZGA) (12,210). In humans, the initial burst of transcripts is detected around the 2-4 cell stage (EGA) (**Figure 1.4**) (11,211). It is suggested that this first wave of gene expression is initiated from the paternal genome, which is temporarily depleted from repressive histone marks and DNA methylation, making it more accessible (212). The pioneer transcripts are lowly expressed and relatively short and may include

retroviral repeats or other factors generally found in heterochromatin. Some of these factors are functionally important for the major wave of transcription, which occurs at later stages (at the 2-cell stage in mice and around the 8-cell stage in humans) (71,213,214). Among such transcripts are *Dux* (in mice) and *DUX4* (the human ortholog), which encode double-homeodomain TFs and are highly conserved across placental mammals. *DUX4* is found within the subtelomeric region of human chromosome 4q35, a region rich in D4Z4 macrosatellite repeats. *DUX4* is expressed for a very short time, with its RNA detected from the oocyte stage until approximately the 4-cell stage, after which it becomes suppressed by H3K9me3 (214). This gene activates the expression of several developmentally important genes, including *ZSCAN4*, *LEUTX*, *KDM4E* and others (71,213). It was also demonstrated that *Dux*-depleted mouse embryos failed to reach the morula and blastocyst stages, presenting impaired ZGA (214).

Nevertheless, it is critical to bear in mind that until the major wave of transcription, the embryo relies entirely on the transcripts (proteins and RNA) that were transcribed in the oocyte (**Figure 1.4**) (10,11). Such oocyte-derived transcripts are gradually degraded, with some maternal products persisting even after the second wave of EGA. Therefore, at around the 8-cell stage in humans, the embryonic genome must be fully activated to ensure normal embryo development and proper transcription throughout life (17,71).



**Figure 1.4. Schematic representation of changes in transcriptome during embryonic genome activation (EGA).**

During oocyte maturation, various maternal transcripts, such as RNAs and proteins, are produced and stored for the embryo (red line). After fertilisation, these maternal transcripts are gradually depleted by the embryo. The minor wave of EGA occurs around the 2 to 4-cell stages, predominantly from the paternal genome (blue line). At this stage, several important transcriptional activators are produced, which then activate other developmentally essential genes at the 8-cell stage, when genes become fully expressed (the major wave of EGA) from the embryonic genome. Light blue circle – paternal pronucleus, red circle – maternal pronucleus and orange circle – embryonic nucleus.

## 1.4. Epigenetic modifications and their role during embryonic development

As discussed previously, early human development is highly dynamic, as the embryo must undergo a few milestones that shape its future development. Firstly, after fertilisation, the embryo has to undergo genome-wide epigenetic reprogramming (4,71,88,136,137,215), during which most DNA methylation and histone PTMs inherited from gametes are removed, leading to increased chromatin accessibility, especially at CpG islands, promoters and enhancers (136,215). These processes are necessary for a successful EGA, which is vital for the embryo's survival (71). After these critical milestones, the embryo has to gradually regain histone PTMs and DNA methylation to specify TE and ICM and later Epi and PrE (71,136,137,216). After 3 to 4 weeks of gestation, the conceptus has to undergo another significant milestone, gastrulation (217), which is poorly understood in humans due to limited access to samples and limited model systems.

Mammalian oocytes, including humans, exhibit genomic regions larger than 10 kb that have low CpG density and display low to intermediate levels of DNA methylation (regions containing a minimum of 20 CpGs within a 10 kb sliding window with an average methylation below 50%). These regions are often observed in intergenic, gene-poor regions or within silent (non-transcribed) gene bodies and are referred to as partially methylated domains (PMDs) (41,136,137).

In several mammalian species, PMDs have been found to overlap with H3K4me3 peaks that are not observed in human oocytes (41,136,137,218,219). In human germinal vesicle (GV) and metaphase II stage (MII) oocytes, H3K4me3 distribution is characterised by sharp and narrow peaks at gene promoters (canonical H3K4me3 distribution (**Figure 1.3A, C, D**)), and such marked genes become highly expressed following EGA (136,137). Similarly, in mouse-developing and mature oocytes, H3K4me3 forms canonical sharp peaks at TSSs of actively expressed genes (137,139). However, most H3K4me3 forms broad, non-canonical domains (covering more than 20% of the mouse MII oocyte genome), which are located distantly from TSSs in intergenic regions and overlap with unmethylated regions and PMDs (136,137,139,140,219). Extensive research from several groups has shown that the methylome of the mouse oocyte displays a bimodal distribution, with relatively few intermediate methylated regions or PMDs, and that the distribution of H3K4me3 and DNA methylation are anti-correlated (139,149,160,220,221). More specifically, H3K4me3 is mainly confined to hypomethylated regions (**Figure 1.3**), such as promoters or bivalent domains, with a preference for higher CpG density, whereas DNA methylation is concentrated at actively transcribed gene bodies marked by

H3K36me3 (**Figure 1.3C, D**) (139,149,159,220). Therefore, H3K4me3 detected at PMDs in mouse oocytes could be a result of somatic cell contamination or differing criteria used to define PMDs. After the 4-cell stage in humans, H3K4me3 peaks become much broader, and some of them can be found in gene-dense or distal regions. A subset of these peaks arises *de novo*, as they are not inherited from gametes (136,137). Interestingly, some of these *de novo* H3K4me3 peaks transiently appear at PMDs just before the onset of the major EGA (136). By the 8-cell stage in human embryos, a proportion of H3K4me3 peaks correlates with high gene expression, especially for genes with CpG-dense promoters, while a large proportion of H3K4me3 is lost from gene promoters, and these genes remain silent. This sudden loss of H3K4me3 was suggested to be associated with the upregulation of the *KDM5B* demethylase and the downregulation of *KMT2C* or *KMT2B* methyltransferases (136). In the same study by Xia and colleagues, it was found that prior to EGA, some distal H3K4me3 domains were associated with *cis*-regulatory elements, and after EGA, some of these domains, in proximity to important lineage-specific genes, remained accessible and gradually obtained H3K27ac (136). In contrast, other regions acquired H3K27me3, formed bivalent domains and were eventually repressed following the loss of H3K4me3. Therefore, after EGA, a large proportion of H3K4me3 was lost and became restricted to active gene promoters, exhibiting the canonical distribution observed in the human blastocyst and somatic tissues.

H3K27me3 distribution in human oocytes is similar to other mammals (136,137,219). Canonical H3K27me3 domains overlap developmental gene promoters (**Figure 1.3B**), while non-canonical H3K27me3 distribution (broad domains) is observed over unmethylated genomic regions and the PMDs (27,136,137,222). Unlike mouse oocytes, such PMDs are not limited to early embryonic development and have been observed in somatic tissues, including the liver, brain and other human tissues (136). In human embryos, H3K27me3, inherited from the oocytes, persists until the 4-cell stage and is largely erased by the 8-cell stage (136,137). In contrast, H3K27me3 domains derived from sperm are almost immediately removed, as in the 2-cell stage embryos, such paternally derived H3K27me3 domains are almost absent (136). Similar observations were reported in other mammalian species, except in the mouse and rat embryos, where maternally derived H3K27me3 is maintained throughout pre-implantation development (24,136,137). This can be explained by the fact that in the mouse oocytes and pre-implantation embryos, the major PRC2 subunit genes, including *Eed*, *Suz12*, and *Ezh2*, are continuously expressed. By contrast, in humans, such components are expressed in female PGCs; *EED* and *SUZ12* become downregulated in oocytes and then become upregulated in the 8-cell stage embryos (136,137). In addition, expression of *KDM6A* and *KDM6B* demethylases are detected in human oocytes and remain expressed throughout pre-implantation development (71,136,137). Therefore, the authors suggested that maternal H3K27me3

peaks are established during human PGC development in females, as all components of PRC2 are available (136). Then, H3K27me3 domains are lost by the 8-cell stage in human embryos because PRC2 cannot protect these domains from erasure. Furthermore, before EGA in humans, H3K27me3 is erased from the promoters of developmental genes, as probably most of these genes become expressed during EGA. After EGA, H3K27me3 becomes restricted to its canonical targets, including developmental genes (**Figure 1.3B**) and genes with CpG-dense regions that can gain bivalency (27,136,137).

Interestingly, in human ICM (D6-D7), H3K27me3 domains were found at Epi- and PrE-specific genes, but they were much less prevalent in TE-specific genes (136). In contrast, in TE cells (D6-D7), Epi- and PrE-specific genes were decorated by H3K27me3, whereas this modification was less abundant on TE-specific genes. This asymmetric distribution of H3K27me3 suggested that cells might be primed to differentiate into TE. A recent study showed that inhibition of PRC2 in naïve hESCs caused their differentiation into either TE or mesoderm, as they continued expressing pluripotency genes with lineage-specific TFs (223). In naïve hESCs, TE and mesoderm genes were decorated by bivalent domains and kept at a transcriptionally poised state.

A recent study by Yuan and colleagues generated human haploid androgenetic (a zygote containing only the paternal pronucleus; AG) and parthenogenetic (a zygote with the maternal pronucleus; PG) embryos to investigate H3K27me3 distribution and non-canonical imprinting in human blastocysts (224). In general, the authors found that genes harbouring H3K27me3 were repressed and uncorrelated with DNA methylation. AG-specific H3K27me3 domains were hypomethylated, but the same regions were hypermethylated in PG embryos, and DNA methylation was inherited from the oocyte. One-fifth of such AG-specific H3K27me3 regions included the DMRs of reported imprinted genes (detected in PG-blastocyst). Interestingly, AG-blastocysts (~77%) harboured many more unique H3K27me3 domains than PG-blastocysts (~23%) (224). The majority of such genes with H3K27me3 enrichment were not expressed in the blastocyst, and only a few genes with AG-specific H3K27me3 domains were expressed in PG-blastocyst, and the opposite was true, with a few genes associated with PG-specific H3K27me3 were expressed in AG-blastocysts. In agreement with previous studies, it was demonstrated that the H3K27me3 profiles of AG- and PG-blastocysts differed from those of gametes, further indicating that H3K27me3 is lost at the 4-cell stage and reestablished after the 8-cell stage in human embryos (27,136,137,224).

So far, H2AK119ub1 has not been explored well in human embryos, but there is a growing interest in this histone PTM in the mouse. In the mouse oocytes, H2AK119ub1 forms broad non-canonical domains found at distal genomic and CpG-rich regions (130,132).

Therefore, this PTM can overlap with H3K27me3 and H3K4me3. In addition, some CpG-rich regions enriched for PRC1 and PRC2 marks can form self-interacting domains known as polycomb-associated domains (PADs) that are inherited by the zygote and maintained during pre-implantation stages (181). H2AK119ub1 is primarily inherited from the oocyte and exhibits a similar distribution; however, it is mainly depleted by the 2-cell stage (130,132). After ZGA, this PTM is reestablished at developmental gene promoters or non-canonical imprints overlapping with H3K27me3 (132).

PRC1 is also important for genes with bivalent domains, as abnormal levels of H2AK119ub1 can disrupt the expression of such genes (130,225). Interestingly, it was reported that after fertilisation, H2AK119ub1 was accumulated in regions that later became bivalent domains during the later stages of mouse development (226), and these regions were inherited from the gametes. Therefore, PRC1 may guide the establishment of such domains, at least in the mouse.

Prior to EGA in humans, H3K27ac forms broad peaks that are also observed in the mouse (123). Most H3K27ac domains were observed at PMDs between 2- to 4-cell stages. Especially such H3K27ac peaks were established over PMDs as H3K27me3 was gradually removed. H3K27ac domains were also found at CpG-rich promoters close to TSSs. These genes showed high expression after initiation of EGA. Interestingly, it was reported that during the 2-4 cell stage developmental window, 75% of H3K4me3 peaks overlapped with H3K27ac domains and were located near CpG-rich regions that became highly upregulated after EGA (123). After the 8-cell stage, both H3K27ac and H3K4me3 domains were reduced to canonical regions, forming narrow domains (**Figure 1.3A, C, D**) (123,136,137). In post-EGA human embryos, H3K27ac was also localised to distal genomic regions. Such regions were hypermethylated at the 8-cell stage but became open (as indicated by ATAC-seq) and hypomethylated in ICM and hESCs (136). Such enhancers were shown to contain the binding sites for several known TFs such as *KLF*, members of the *GATA* family and *TFAP2A/C* that were enriched in 8-cell stage embryos and ICM (136). In contrast, *GSC* and *OTX2* were only detected in the 8-cell stage, while *TEAD* family members were specifically found in the ICM.

Unlike in human oocytes, where H3K9me3 is enriched at gene-dense regions, in embryos, H3K9me3 is found at a much lower level (227). Therefore, it was suggested that H3K9me3 is required for gene repression in human oocytes. Also, in the 4-cell stage human embryos, a proportion of enhancers was decorated by H3K9me3 that was lost between the 4- to 8-cell stage (228). After EGA, such regions became accessible and active. In the mouse, similar regions lost H3K9me3 just before the ZGA or 2-cell stage. Many such active enhancers in humans were present near developmentally important genes (228). The

authors later discovered that such H3K9me3 enriched enhancers contained primate-specific retrotransposons that included LTR12, LTR5\_Hs, LTR7B and HERVH-int elements with important TF binding sites.

## 1.5. Endogenous retroviruses (ERVs)

ERVs are remnants of evolutionary distant exogenous retroviral infections as retroviruses integrated their genetic material in the form of DNA into the host genome and are widely distributed in the mammalian germline (229,230). Long terminal repeats (LTRs) flank the major proviral genes of a retrovirus, encoding structural proteins, enzymes and envelope proteins that all integrate into the host DNA (229,230). ERVs might not be able to carry the infection, but they contain the machinery required for replication and insertion into the host genome by vertical transmission (ERV inheritance through the germline) (231,232). Many LTRs become upregulated in cancers; for example, a primate-specific THE1B is upregulated in Hodgkin's lymphoma (233). Similarly, in male mice, not repressed ERVs cause sterility due to abnormal chromatin conformation and aberrant gene expression (46,234). Therefore, such elements can be harmful to the host genome, and different species evolved several mechanisms to cope with such ERVs. In humans, such elements can be silenced by acquiring *de novo* DNA methylation (**Figure 1.2B**), as such elements are rich in CpG sites (161). Also, KRAB-ZFPs recognise newer ERVs and recruit KAP1, which forms a larger complex incorporating SETDB1, and H3K9me3 becomes deposited at such ERVs. Additionally, H3K9me2/3 can be recognised by UHRF1, which, as a result, brings DNMT1, and such regions remain hypermethylated (161).

ERVs can be useful to the host genome as they can bring innovations important for gene regulation and even drive speciation (235–237). In mammalian cells, similar proviral LTRs can undergo recombination, which results in a loss of virulent genes and the formation of solo-LTRs that retain a promoter function (as they contain TSS for the virulent genes) and can be adapted by the host organism (229,230). Over time, these LTR elements can accumulate mutations that allow them to escape the silencing mediated by KRAB-ZFP-KAP1 but also to form new TF binding sites and drive the expression of novel transcripts (**Figure 1.2B**) (161,229,230). LTRs constitute around 8% and about 10% of the human and mouse genomes (236–238). In general, they are highly active in the mammalian germline and during early embryonic development when the embryonic genome is mostly erased. In the mouse oocytes, around 15% of transcripts are derived from LTR elements and most such transcripts are induced by LTRs of the mouse

transcript (MT) subfamily of MaLR that make a large proportion of all mouse LTRs and are specific to rodent species as they appeared after Hominidae diverged from Muridae (221,235,239). ERV1 family is more prevalent in the human genome than in rodents, and therefore, this family predominantly induces the expression of chimeric transcripts in the human oocytes (235). Such active ERV LTRs in the oocytes become decorated by H3K4me3 (140), which protects them from acquiring *de novo* methylation. In addition, such transcriptionally active regions acquire broad H3K36me3 domains downstream of ERVs as they are established by SETD2, which is present in the RNA polymerase II complex (103,235,236,240). The PWWP domain of the *de novo* methyltransferases recognises H3K36me3 (156), and they hypermethylate these regions (240). Therefore, such regions can contribute to forming gDMRs and novel transcripts in the oocyte. As a result of ERV-induced *de novo* transcription, the major gene promoters can become hypermethylated and such genes show lower expression than compared to ERV-derived transcripts (235,236). In contrast, the integration of ERVs in intragenic regions can result in exon skipping and the evolution of new gene isoforms (**Figure 1.2B**) (235,236). Moreover, ERVs are polymorphic, as diverse ERV families can be shared between species or unique to one species, depending on the timing of the ERV integration into the host genome. For example, MTD, MT2A or MTC elements are shared between rodent species, while THE1B or LTR12C elements are unique to primates (236).

As discussed earlier, ERVs can be utilised as enhancers during human pre-implantation development, and before EGA, some are decorated by H3K9me3 (227,228). During EGA, hominoid-specific ERVs such as LTR12C, LTR5\_Hs, LTR7B, and HERVH-int or SVA retrotransposons (SINE-VNTR-Alu: a fusion of SINE and ERV LTRs (241)) lose H3K9me3 and promote the expression of developmentally important genes. These elements harbour the binding sites for key TFs and EGA-associated genes, including the *DUX* family members and *ZSCAN4*. By employing a dCas9<sup>KRAB</sup> system that can be activated by doxycycline (Dox) in human zygotes to recruit H3K9-methyltransferases to the hominid-specific SVA loci, it was shown that embryos underwent normal cleavage divisions (228). However, these embryos exhibited a developmental delay at the start of EGA. A closer examination of SVA (+Dox) treated samples versus controls revealed differences in H3K9me3 distributions. The treated embryos were more similar to pre-EGA embryos and overall had a lower expression of EGA-associated genes, as they were regulated by SVA-derived enhancers. Thus, the authors concluded that disrupted removal of H3K9me3 from SVA loci could inhibit EGA (228).

In the same study, the authors investigated whether such H3K9me3-enriched ERVs could be important for the first cell lineage commitment (228). By utilising ATAC-seq in combination with ChIP-seq for H3K9me3, the authors found that in TE, hominoid-

specific ERVs such as THE1B-int, MER11C, HERVK9-int, LTR12, and MER11B gained *de novo* H3K9me3 peaks. Curiously, these ERVs contained binding sites for such factors as *POU5F1*, *SOX2*, and *NANOG*, which were expressed in the ICM. However, in the ICM, these ERVs did not harbour H3K9me3. Also, in TE cells, neither ICM-specific TFs nor TE-specific TFs could bind to these ERVs. Therefore, the authors suggested that important ICM-specific TFs might be marked by H3K9me3 and repressed in TE cells, as the binding of such factors could hinder TE differentiation (228).

ERVs can play an essential role during placental development. During gestation, TE can differentiate into syncytiotrophoblasts (STBs) that fuse to form a large multinucleated layer called syncytium. The syncytium acts as a major barrier, preventing maternal blood from directly entering the foetal tissues and performing several other functions during pregnancy that will be discussed later. This cell fusion is caused by the expression of Syncytin derived from ERVs (242).

Recently, one study investigated the function of ERVs in regulating trophoblast-specific genes in hTSCs (243) and primary placental cytotrophoblasts (CTBs) (138). By combining ChIP-seq and CUT&Tag for H3K27ac, H3K4me1, H3K4me3, H3K9me3, and H3K27me3, they identified 18 families of primarily primate-specific ERVs that were mainly enriched for H3K27ac and H3K4me1. These identified elements also included several known ERVs that could regulate gene expression in the placenta. They also identified several ERVs decorated by H3K4me1 (“poised enhancers”) (244), which became active in extravillous trophoblast cells (EVTs) derived from hTSCs. Interestingly, most of these ERVs were not active in hESCs. The authors further investigated which TFs can bind to these active ERVs and found that *ELF5*, *FOXO3*, *GATA3*, *KLF4*, *TEAD4*, *TFAP2C*, *TP63* and multiple other TF binding motifs were preset in LTR10A, LTR23, LTR2B, LTR3A, LTR7C, MER11D, MER21A, MER41C, and MER61E (138). Some of these ERVs were decorated by H3K4me3 and possibly could function as gene promoters. To confirm this, the authors used a CRISPR-Cas9 system to delete several ERVs, including MER41B and LTR10A (138). The deletion resulted in the downregulation of *ADAM9* (associated with pre-eclampsia (PE) (27)), *CSF1R* (important for trophoblast differentiation (245)), and *ENG* (the level of soluble *ENG* correlates with the severity of PE (246), also an important factor in trophoblast differentiation (247)). Finally, it was found that hominid-specific ERVs, such as LTR2B elements, were associated with higher gene expression in the placenta compared to ERVs shared with macaques (138). Hence, ERVs can serve as novel promoters and enhancers that shape the human placental transcriptome and drive its rapid evolution.

## 1.6. Genomic imprinting

Genomic imprinting refers to epigenetic mechanisms that result in the differential marking of parental loci, leading to the monoallelic expression of these loci. It has been reported that genomic imprinting has emerged multiple times independently during evolution, as it can be observed in several arthropod species, flowering plants, and therian mammals, including eutherians and marsupials (248). While imprinting demonstrates some similarities and differences between plants and mammals. I will focus only on mammalian species.

In mammals, genomic imprinting is thought to have emerged around 187 million years ago (249), coinciding with the development of vivipary and the emergence of the placenta, as imprinted genes are absent in monotremes (egg-laying mammals) (250). The first strong evidence for genomic imprinting in mammals came from two seminal studies conducted in the 1980s by the groups of Surani (251), McGrath and Solter (252). Both studies involved experiments with mouse zygotes in which the maternal pronucleus was replaced with a second paternal pronucleus to create diploid AG conceptuses, and vice versa, where the paternal pronucleus was replaced with a second maternal pronucleus to generate diploid gynogenetic conceptuses (a zygote with two maternal pronuclei; GG). These manipulated embryos were then transferred into pseudo-pregnant surrogate females. Interestingly, these embryos failed to survive post-implantation. GG conceptuses showed no development of extra-embryonic tissues, while AG conceptuses exhibited underdeveloped embryonic tissues and overgrowth of extra-embryonic tissues. The phenotype is similar to human androgenetic hydatidiform moles (HM), where an enucleated oocyte, or one with a replication-defective genome, is fertilised by one or two sperms, leading to the overgrowth of trophoblastic tissues (253,254). The experiments by Surani, McGrath and Solter groups demonstrated that parental genomes are epigenetically non-equivalent, and both are essential for a successful pregnancy.

These ground-breaking studies were followed by the discovery of the first imprinted genes, which included maternally expressed Insulin-like growth factor 2 receptor (*Igf2r*) on mouse chromosome 17 (255), paternally expressed Insulin-like growth factor 2 (*Igf2*) on chromosome 7 (256,257), and maternally expressed *H19* (258), located approximately 90 kb downstream from *Igf2* in the mouse genome. Subsequent research revealed that parental chromosomes are differentially methylated at imprinted genes (254,259,260), resulting in the repression of methylated alleles and the expression of unmethylated alleles. Furthermore, comparisons between mice and humans identified that *IGF2* and *H19* are also imprinted in humans (261–263).

## 1.7. The life cycle of imprints

Intergenerational maintenance of imprints is a complicated task. It involves many *cis*-elements and *trans*-acting factors, which all have to work in a highly coordinated fashion to ensure adequate methylation levels in somatic tissues and germ cells and fine-tune the expression of imprinted genes (7). Genomic imprinting involves two rounds of demethylation, followed by two rounds of remethylation (**Figure 1**), which I will describe in more detail further.

### 1.7.1. Pre-implantation epigenetic reprogramming

At the time of zygote formation, the paternal and maternal genomes exhibit asymmetry in their epigenetic landscapes that have to be depleted of epigenetic marks to establish totipotency (4,8,20,88). As noted earlier, after fertilisation, the sperm genome undergoes genome-wide reorganisation, leading to a sudden drop in global methylation level. All these processes occur before pronuclear fusion and the first mitotic division. Sperm DNA loses most 5mC until the 2-cell stage, which is associated with an enzymatic activity of TET3, which is abundant in the oocyte and zygote (**Figure 1.1**) (8,264–266). It was observed that 5mC derivatives, including 5fC and 5caC, showed a gradual decrease in the paternal pronucleus, which suggests that passive demethylation rather than active demethylation is preferentially taking place. The oocyte genome also undergoes genome-wide demethylation, but overall, it demonstrates slower dynamics. Therefore, it is proposed that oocyte-derived 5mC is removed by passive demethylation, especially when oocyte-derived factor DPPA3 was shown to protect the maternal genome from TET3-mediated 5mC demethylation (267,268). DPPA3 (also known as STELLA) interacts with H3K9me2, which is predominantly found in the maternal pronucleus and inhibits TET3 binding. In addition to this, DNMT1 and its co-factor UHRF1 are excluded from the nucleus, restricting DNMT1 activity (163,269) and, in humans, are subject to maternal-transcript decay (71). Only parent-specific methylation marks at ICRs and some repeat elements escape demethylation. A body of evidence indicates that imprints might be protected by a few maternal and zygotic factors, such as DPPA3, ZFP57, ZNF445 and NLRP proteins (17). For example, ZFP57 and ZNF445 are KRAB-zinc finger proteins that each can interact with the TRIM28/KAP1 scaffold protein in a multi-protein complex, including DNMT1, UHRF1, SETDB1 and the histone deacetylation complex NuRD. ZFP57 or ZNF445 recognises specific methylated sequences often found in imprinted DMRs and transposable elements and recruits the co-repressive complex to a targeted sequence

(58,177). In this way, ICRs and transposable elements are possibly protected from demethylation throughout pre-implantation development. This idea is further supported by *Zfp445–Zfp57* double-mutant mice that lost imprinting at 15 ICRs and were embryonic lethal (177). Overall, the lowest methylation level is reached by the blastocyst stage in human and mouse embryos, subsequently leading to remethylating (4,9).

### 1.7.2. Post-implantation maintenance of imprinted regions

After blastocyst implantation, a new wave of remethylation is initiated, leading to a gradual loss of pluripotency (**Figure 1.1**) (7,9,17,216). *DNMT3A* and *DNMT3B* are expressed at high levels and establish *de novo* methylation patterns in the post-implantation embryo. Such *de novo* marks, including imprints, are robustly maintained by DNMT1 after every cell division (270). A global increase in DNA methylation and the expression of lineage-specific marker genes initiate lineage commitment and cellular differentiation, ultimately leading to gastrulation and tissue formation (13,71,216).

It is important to note that unmethylated alleles present at imprinted regions should be protected from *de novo* methylation. A few mechanisms have been proposed that may protect unmethylated alleles from acquiring methylation. Firstly, CTCF is a TF that binds to the hypomethylated maternal chromosome at the well-known *H19* and *Igf2* gDMR, also known as the *H19/Igf2* imprinting domain. In one study employing transgenic RNA interference (RNAi) to reduce CTCF levels in growing oocytes, the maternal chromosome at the *H19/Igf2* gDMR was found to acquire DNA methylation (271). These findings suggested that CTCF plays a protective role against methylation at this imprinting domain during oocyte development. Secondly, CpG islands are usually enriched with H3K4me3. This histone PTM prevents DNMTs, more specifically DNMT3A and DNMT3B, from binding, thereby maintaining an unmethylated state in these regions (272).

### 1.7.3. Genome-wide epigenetic reprogramming in primordial germ cells (PGCs)

PGCs are progenitors of the oocyte and sperm that have to undergo genome-wide reprogramming to erase all parent-specific epigenetic marks to ensure a totipotent state necessary for embryo development (**Figure 1**) (7,17,273). PGCs originate from the Epi of

the pre-gastrulation stage embryo in mice around E6.25 (16). On their way to the genital ridge, these cells must go through global epigenetic reprogramming, during which genome-wide loss of 5mC is detected. Based on observations in mice, it is suggested that the first round of global demethylation appears through passive demethylation, as *de novo* DNMTs and *Uhrf1* are downregulated at this stage. Although the DNMT1 protein is detected at a high level, it is excluded from the nucleus (7,274,275). During this time, the population of PGCs rapidly expands through mitotic division, allowing for replication-coupled dilution of 5mC. This is followed by the second round of demethylation, which involves TET1 and TET2 proteins in mice. These enzymes were shown to be necessary for genomic imprint erasure (16,276,277). Similar findings were made in human PGCs (278–280). By around E13.5 in mice and around weeks 9–11 in humans, methylation marks at imprinted regions are entirely removed, in combination with histone remodelling and changes in the chromatin conformation.

#### 1.7.4. Establishment of sex-specific imprints

Following PGC reprogramming, new sex-specific epigenetic marks are established, giving rise to unique epigenetic profiles of the sperm and oocyte (**Figure 1**). Remethylation in both gametes happens asynchronously at different time points, as gametes are physically separated (173,281). In the male gamete, paternal-specific imprints are obtained before birth (E14 – E17.5), while in the female gamete, maternal-specific imprints are acquired after birth until the GV stage in mice (160,282). *De novo* methylation marks, including maternal-specific imprints, are established by DNMT3A and DNMT3L in the mouse oocyte (282). On the other hand, DNMT3A and DNMT3B are suggested to be more important in mouse sperm (233).

### 1.8. Classical imprinted genes

Imprinted genes are a unique group of genes that are conserved in eutherian and marsupial mammals and demonstrate parent-specific monoallelic expression (284). It is hypothesized that imprinted genes evolved alongside placentation in mammals due to asymmetry in parental investment (285). There are around 197 imprinted genes described in mice and around 165 in humans (<http://igc.otago.ac.nz/>; [www.geneimprint.com](http://www.geneimprint.com)) (18). It is well known that imprinted genes are essential for embryo development, metabolism,

placental formation, and brain development (7,17,286,287). However, recent evidence suggests that these genes may influence behaviour, sleep and the circadian clock (reviewed in (18)). Therefore, *de novo* mutations, abnormal regulation and altered expression levels of imprinted genes are often associated with lifelong imprinting disorders and an increased risk of cancer (17,287).

Imprinted genes are characterised by their unusual allele-specific expression, which is primarily determined by ICRs. Recent studies have identified many maternal and paternal gDMRs, with the majority persisting until the pre-implantation stages (20–22). However, it is important to note that not all such gDMRs function as ICRs. Classical imprinted genes are defined by several features (288). A few genes that demonstrate allele-specific expression (ASE) tend to cluster in imprinted domains, containing at least one ICR and a non-coding RNA (ncRNA) (289–291). The regulation of such genes frequently involves several *cis*-elements and *trans*-acting factors. The most extensively explored and well-described imprinting domain is the *H19/IGF2* locus.

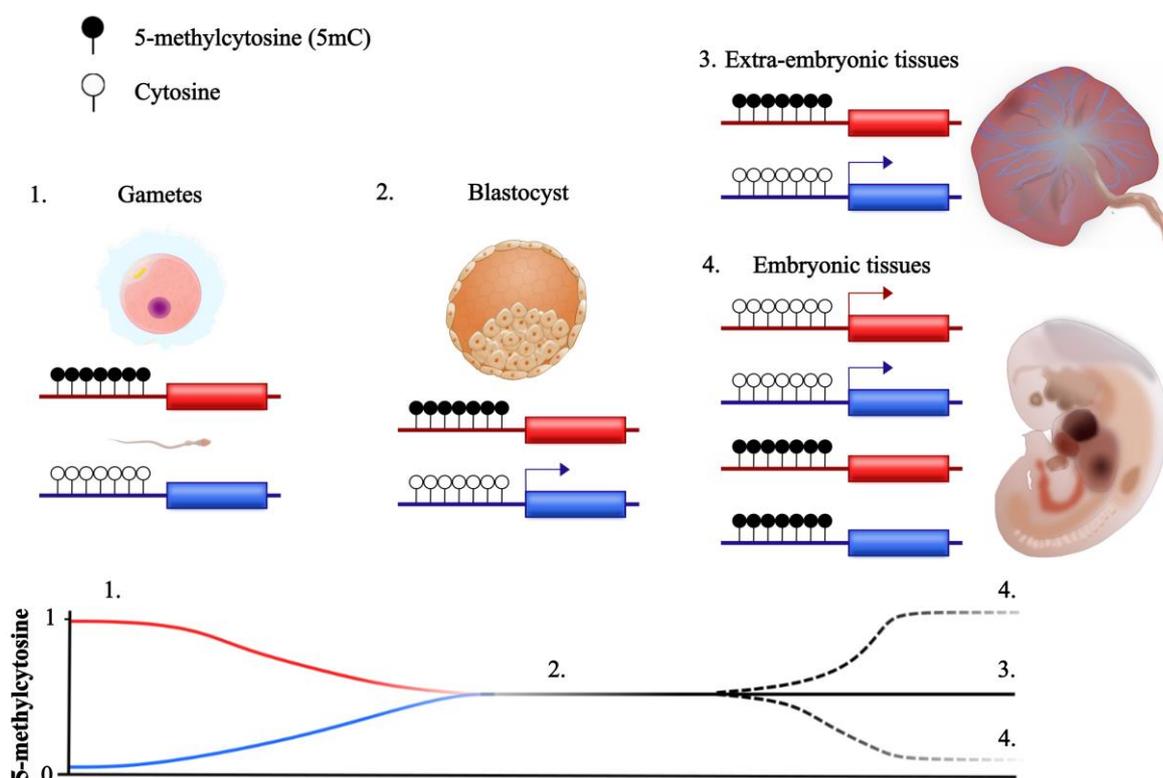
Human chromosome 11p15.5 contains a large imprinted gene cluster, with the telomeric *H19/IGF2* imprinting domain and the centromeric *CDKN1C/KCNQ1OT1* imprinting domain (17,289,291). *H19* is a lncRNA expressed from the maternal chromosome and acts as a growth suppressor. *IGF2* is a paternally expressed gene and is a growth-promoting gene. Imprinting control centre 1 (*IC1*) is a 5 kb intergenetic, paternal gDMR found between *IGF2* and *H19*. *IC1* acts as an insulator, which is rich in tandem repeats recognised by CTCF. It is suggested that CTCF binds to unmethylated *IC1* on the maternal chromosome and protects the maternal *IC1* from gaining methylation (292). On the paternal chromosome, *IC1* is methylated, which inhibits CTCF from binding. On the maternal chromosome, the binding of CTCF induces higher-order chromatin conformational changes, blocking common enhancers from activating the *IGF2* promoter. Instead, these enhancers bind to the *H19* promoter and activate this gene transcription. More recent findings demonstrated that ZFP57 recognises methylated hexameric motifs at *IC1* on the paternal chromosome and protects the paternal allele from losing methylation (293). Consequently, the common enhancers can bind to the *IGF2* promoter on the paternal chromosome and initiate *IGF2* expression. Any (epi)mutations occurring at this imprinted domain can lead to Beckwith-Wiedemann syndrome (BWS) and Silver-Russell syndrome (SRS) (294).

## 1.9.DNA methylation-dependent imprinting

Improving sequencing technologies such as single-cell techniques have allowed us to explore the epigenomes of the oocyte, sperm and pre-implantation embryo with unprecedented depth, providing intriguing new insights. It was observed that oocytes and sperm have many gDMRs, with the majority being erased soon after fertilisation (4,8,20–22,295). Some maternal gDMRs survive the post-fertilisation demethylation process but eventually lose their parental specificity by the implantation stage as they undergo gain or loss of methylation. Therefore, such maternal gDMRs are referred to as transient differentially methylated regions (tDMRs) that are indistinguishable from conventional DMRs. Interestingly, several research groups have demonstrated that some of these maternal tDMRs persist in TE and the placenta (**Figure 1.5**) (8,20–22,295). Although these tDMRs can be polymorphic because the placental samples of different individuals demonstrate varying levels of DNA methylation, with some samples showing low levels of DNA methylation, suggesting a relaxed imprinting mechanism (20,295). In addition, it was experimentally confirmed that a few of these tDMRs might act as ICRs and regulate the expression of proximal genes, including *FAM149A* (20), *DSCAM* (296) and the *GPR1/ZDBF2* locus (23,295,297) in the human pre-implantation embryo and placenta. Probably, the zinc finger, DBF-type containing 2 (*Zdbf2*) locus is perhaps the best-described transiently imprinted gene to date, with similar regulation observed in both mice and humans (23,295). Upstream the TSS of *Zdbf2*, there is an alternative promoter whose activation results in the expression of a *Zdbf2* long isoform named *Liz* (long isoform of *Zdbf2*). *Liz* is a paternally expressed gene controlled by a maternal gDMR (approximately 73 kb upstream of the TSS of *Zdbf2* in the mouse), which disappears from embryonic tissues by implantation as the paternal allele becomes hypermethylated. Therefore, the expression of *Liz* is restricted to the pre-implantation embryo and placenta. Remarkably, the paternal expression of *Liz* disrupts the accumulation of H3K27me3 on the paternal chromosome, allowing for the establishment of antagonizing 5mC (paternal secondary DMR). Secondary or somatic differentially methylated regions (sDMRs) are genomic regions that are methylated either on the paternal or maternal chromosome and originate during pre-implantation development. The appearance of this sDMR contributes to the initiation of *Zdbf2* expression. It was shown that *Liz* KO mice were deficient in *Zdbf2*, which is a growth-promoting gene. Consequently, these KO mice had lower body weight and were smaller than control mice (297). Although the function of *Liz* is not yet fully understood, it is suggested to serve as an alternative mRNA source for the ZDBF2 protein.

Furthermore, cross-species sequence comparisons revealed that there is very little

conservation between human and mouse tDMRs, with most human tDMRs being primate-specific (20,236). Transient imprinting is a relatively recently discovered epigenetic phenomenon, and therefore, the importance of such transiently imprinted genes for embryo development remains unclear. Therefore, it would be interesting to know whether these tDMRs are bioproducts of the oocyte and sperm-specific epigenetic marks and whether they have an important function during early embryo development, with lasting effects throughout an adult's life.



**Figure 1.5. Schematic representation of DNA methylation-dependent transient imprinting.**

(1) In the oocyte, tDMRs are hypermethylated (red line), while in sperm, these regions are hypomethylated (blue line). (2) After fertilisation and until implantation, tDMRs on maternal chromosomes are fully methylated, and the maternal alleles are silenced (red boxes), while the paternal alleles remain unmethylated and are transcriptionally active (blue boxes). Black line – 5mC level in the blastocyst. (3) After implantation, in the TE and future placenta, maternal alleles maintain methylation, whereas paternal alleles are unmethylated and expressed. Black line – 5mC level in extra-embryonic tissues. (4) In embryonic tissues, both maternal and paternal alleles may become hypermethylated (black dashed line), resulting in gene silencing, or hypomethylated (black dashed line), leading to biallelic expression.

## 1.10. DNA methylation-independent imprinting

Genomic regions enriched with repressive histone marks may mediate transient imprinting in mice and potentially in humans. Inoue and colleagues developed a low-input DNase I-sequencing (liDNase-seq) method, which involves digesting DNA from a small number of cells (or even a single cell) with the DNase I enzyme, followed by deep sequencing to identify DNase I footprints or DNase I hypersensitive sites (DHS), representing open chromatin regions (298). They applied this technique in combination with RNA sequencing (RNA-seq) on diploid mouse AG and GG embryos and found that such conceptuses shared thousands of common DHSs, but they also harboured some non-overlapping DHSs that included several imprinted regions (24).

By screening publicly available whole-genome bisulphite sequencing (WGBS) and ChIP-seq datasets from mouse gametes and embryos and by profiling single-nucleotide polymorphisms (SNPs), they discovered that out of 187 AG-specific DHSs (also present on the paternal genome), 105 such regions were hypomethylated and located within H3K27me3 domains in mouse oocytes and ICM (24). Of these, 76 H3K27me3 regions were located near genes, 28 of which were expressed in either AG or GG morulae.

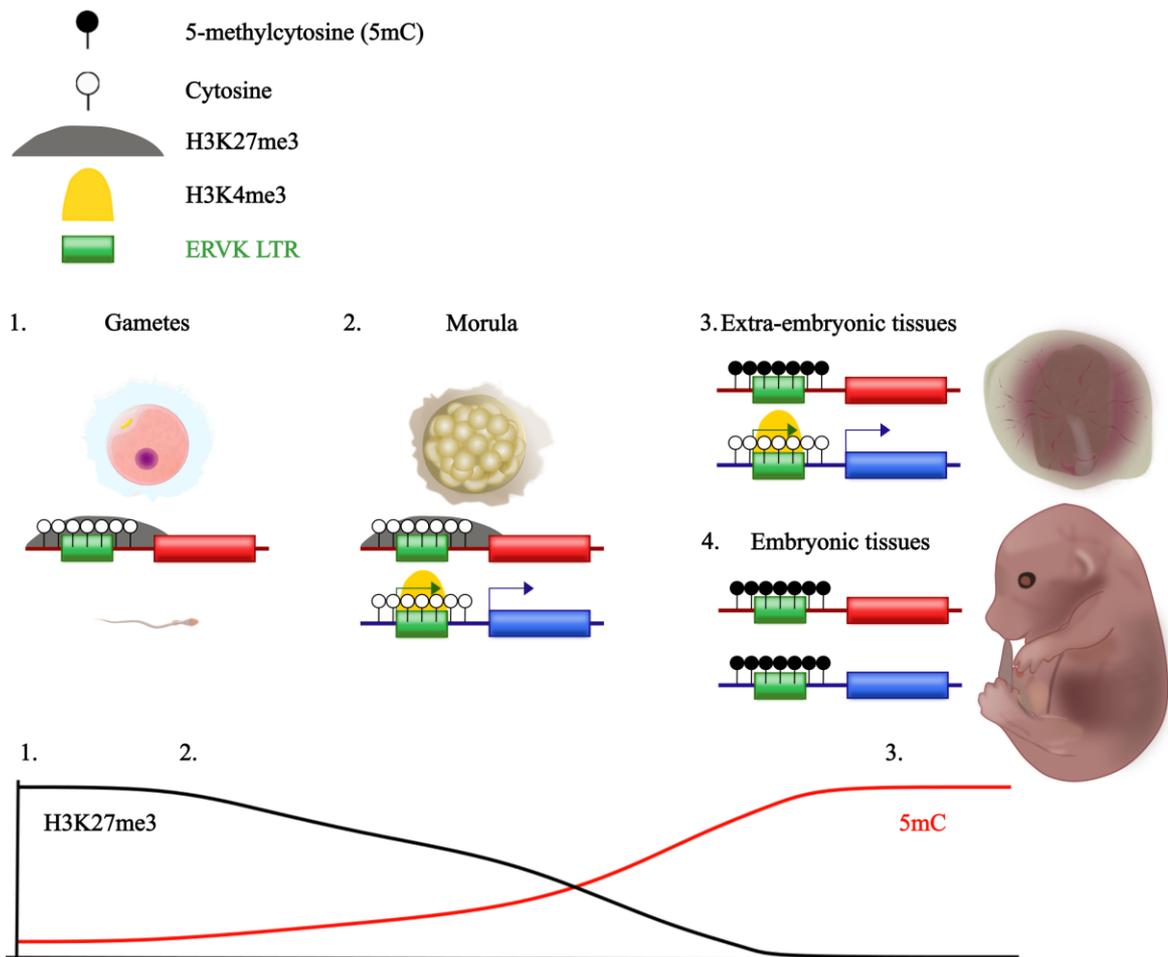
Further analysis in mouse ICM and TE showed that 18 out of 23 genes expressed in TE exhibited paternally biased expression, while 16 out of 24 genes expressed in the ICM also showed paternal-biased expression (**Figure 1.6**) (24). Notably, these 28 genes included *Sfmbt2*, *Gab1*, *Slc38a4*, and *Phf17* (also known as *Jade1*), some of which have been previously shown to be independent of oocyte-derived DNA methylation (299). The researchers further demonstrated that H3K27me3 represses the maternal allele by injecting *Kdm6b* mRNA (H3K27me3-specific demethylase) into one-cell-stage diploid PG embryos (MII oocytes chemically activated), followed by liDNase-seq and RNA-seq at the morula stage (24). This integrated analysis revealed that the loss of H3K27me3 increased the accessibility of the maternal genome, leading to the upregulation of genes with paternal-specific expression, while canonical imprinting remained unaffected. Thus, H3K27me3 was concluded to mediate monoallelic expression independent of DNA methylation, and it was called non-canonical imprinting.

Subsequent studies supported this by generating mouse embryos with maternal knockout (matKO) of *Eed* (an essential component of PRC2), which led to the loss of non-canonical imprints, while canonical imprinted genes remained unaffected in pre-implantation embryos (morula) and extra-embryonic ectoderm (ExE; E6.5) (26,187). Additionally, it was found that non-canonical H3K27me3 domains gradually diminished during mouse

pre-implantation development, disappearing entirely in post-implantation embryos, and some of these regions were overtaken by *de novo* DNA methylation, forming sDMRs that mediate non-canonical imprinting in extra-embryonic tissues (24,26,187). In *Eed* matKO mouse embryos grown to E6.5, the WGBS data of ExE revealed the loss of sDMRs at key genes such as *Gab1*, *Jade1*, *Smoc1*, and *Slc38a4* (maternal gDMR) (26,187). Similarly, mouse embryos with zygotic *Dnmt3a/3b/3B* double knockout (DKO) showed loss of sDMRs in ExE (E6.5; low-input RRBS) for 5 out of 6 non-canonical imprints (with *Sfmbt2* sDMR established by E14.5 in the placenta (300)), while canonical gDMRs retained their methylation (26). Thus, these findings demonstrate that H3K27me<sub>3</sub> domains are transiently retained during pre-implantation development but are replaced by sDMRs in post-implantation ExE.

It was further discovered that many of these non-canonical imprints are associated with ERVK LTRs, which are marked by H3K4me<sub>3</sub> on the paternal allele and the maternal allele at the same region becomes hypermethylated in ExE (sDMR established) (25). These LTRs promote paternal allele-specific expression, as shown by mosaic deletion of RLTR15 via CRISPR-Cas9 in the mouse placenta (E12.5), which resulted in partial loss of imprinting (LOI) at the smaller isoform of *Gab1*. Thus, a deletion of such retroviral elements can disrupt imprinted gene expression.

Following these findings, non-canonical imprinting has been reported in humans, with *FAM101A* showing similar imprinting patterns in human embryos (27). However, this study was limited by its small sample size and low statistical power. Nevertheless, the authors suggested that many such non-canonically imprinted genes may exist in pre-implantation embryos and the placenta.



**Figure 1.6. Model of H3K27me3-dependent imprinting.**

(1) In the oocyte, large regions are enriched with H3K27me3 (black line), while in sperm, such regions are free from H3K27me3. (2) After fertilisation, maternal chromosomes maintain H3K27me3, leading to the silencing of maternal alleles (red boxes), whereas paternal chromosomes lack these marks, allowing ERVK LTRs to activate paternal allele expression (blue boxes). On the paternal chromosome, H3K4me3 overlaps with ERVK LTRs. (3) After implantation in extra-embryonic tissues, H3K27me3 transitions to DNA methylation (red line). As a result, sDMRs are established on maternal chromosomes, silencing maternal alleles. Meanwhile, on paternal chromosomes, ERVK LTRs and paternal alleles remain transcriptionally active. (4) In embryonic tissues, H3K27me3 is replaced by DNA methylation, and both maternal and paternal alleles are silenced.

## 1.11. X chromosome inactivation (XCI)

XCI is a controversial topic in human embryos due to conflicting results reported by different research groups. One explanation for these inconsistent findings is that the mechanism of XCI has been primarily studied in mice (13,301–303).

Until now, it is known that in marsupials, the dosage compensation of X-linked genes is achieved by the inactivation of the paternal X chromosome (Xp) in female embryos (304). In eutherians such as mice, XCI occurs in two waves. During the first wave at the approximately 4-cell stage, the long, *cis*-acting ncRNA called *Xist* is preferentially expressed from Xp, resulting in Xp inactivation. The Xp remains inactive in TE and PrE, but it is reactivated in Epi cells, which leads to the second round of random XCI (this time, either the maternal or Xp can be randomly inactivated). In other eutherian mammals, XCI is a random process, and therefore, it is thought to be the case in humans (13,301–303). Single-cell sequencing technologies allowed exploring this question in much greater detail. One study applied single-cell RNA sequencing (scRNA-seq) to 1,529 cells (88 pre-implantation embryos) and found that X-linked genes were gradually downregulated in female cells from E4 to E7, but *XIST* was biallelically expressed from both X chromosomes during that time window (13). Therefore, they concluded that dosage compensation in female human pre-implantation embryos is achieved by ‘dampening’ both X chromosomes. However, a more recent study (302) using the same data combined with other datasets rejected this hypothesis (a gradual change in the ratio between biallelically and monoallelically expressed genes located on X chromosomes was detected) and, thus, supported a more conventional idea that XCI is random in humans.

## 1.12. Use of uniparental cell lines and embryos to study parental epigenomes

Lately, uniparental embryos and their derived cell lines have been employed to explore the differences between maternal and paternal genomes and their roles in chromatin accessibility, DNA methylation, histone PTMs, and transcriptome profiles during human pre-implantation development (224,305,306). AG embryos are generated by removing the maternal pronucleus using a blastomere biopsy pipette from a newly formed zygote, while PG embryos are produced by removing the paternal pronucleus using an intracytoplasmic sperm injection (ICSI) pipette from the zygote. The use of uniparental cell lines or embryos has become a valuable model for studying processes that occur during early pre-implantation development, such as genomic imprinting, as they provide many advantages over biparental samples (24,224,305). For example, SNPs are extensively used to investigate parent-of-origin-specific effects in mammals, including humans (307–309). In studies using mice as model organisms, two inbred strains can be crossed to generate F1 polymorphic embryos, which are then screened for novel imprinted genes or used to study

the distribution of epigenetic modifications specific to each parental genome (299,307,309). Unfortunately, in humans, only certain genomic regions contain informative SNPs suitable for studying such effects (310). Therefore, uniparental embryos allow for the systematic investigation of contributions from either the maternal or paternal genome to the developing embryo within isolated parental backgrounds.

For instance, Sagi and colleagues used AG, PG, and biparental embryos to derive androgenetic embryonic stem cells (aESCs), parthenogenetic embryonic stem cells (pESCs), and biparental ESCs (305). Transcriptome and methylation profiling of these cell lines revealed similar expression profiles of pluripotency-related genes and most canonical genomic imprints, including genes with sDMRs that demonstrated stable expression and methylation, with some variations at *PEG10* and *H19/IGF2* loci in aESCs and pESCs. By screening for novel genes with parent-of-origin-specific expression, the authors identified *S100A14* - a maternally expressed gene associated with a pDMR not previously linked to known imprinting clusters. Moreover, aESCs demonstrated an intrinsic propensity to differentiate toward extra-embryonic lineages, with placenta-specific genes such as *CGA*, *CGB8*, *ERVFRD-1*, and *ERVW-1* being significantly upregulated in this cell line. Interestingly, teratomas (tumours composed of diverse embryonic tissues) derived from aESCs showed upregulation of liver-specific genes and enhanced proficiency of aESCs to differentiate into hepatocytes. In contrast, teratomas derived from pESCs exhibited increased expression of cerebral cortex-associated genes (305), further supporting the idea that both parental genomes are necessary to generate a viable embryo.

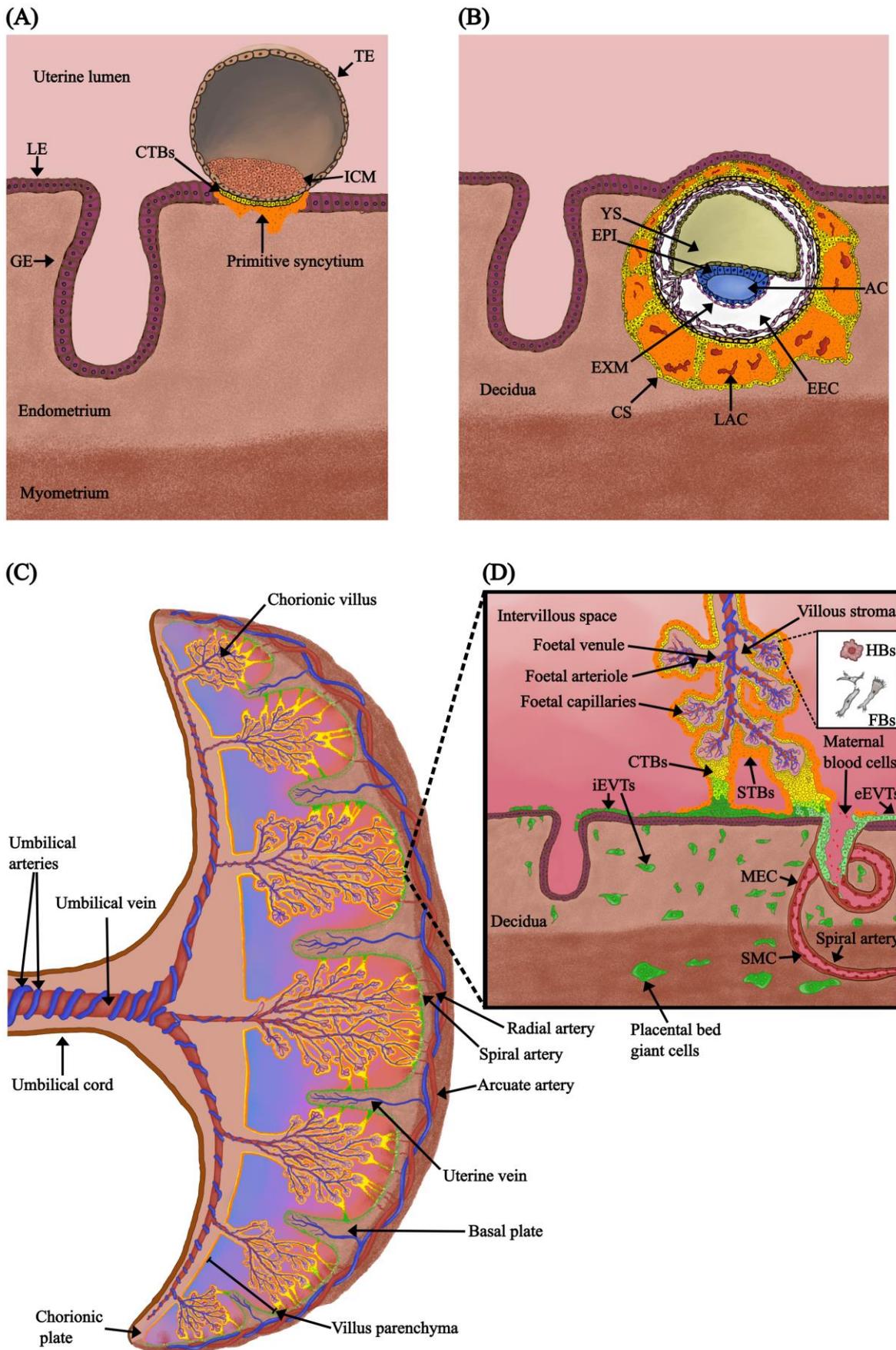
More recently, Yuan and colleagues generated several cleavage stage (CL), morula, and blastocyst-stage AG and PG embryos and found that AG and PG embryos initially resembled biparental embryos (306). However, after the 8-cell stage, PG embryos exhibited a delay in the major EGA. In-depth investigations revealed that *ZNF675* (a primate-specific TF derived from maternally deposited mRNA) and *LSM1* (a component of the deadenylation-dependent mRNA decay complex) were upregulated in 8-cell-stage AG embryos but initially expressed from the paternal genome in biparental embryos. *ZNF675* was required for the upregulation of EGA genes, while *LSM1* was needed to degrade oocyte-derived transcripts and contributed to the delayed EGA observed in PG embryos. These findings suggest that EGA in humans is initiated at the 8-cell stage from the paternal genome (306).

Unfortunately, uniparental embryos are only suitable for studying pre-implantation development, as mouse uniparental embryos die soon after implantation due to imprinting defects that lead to abnormal phenotypes. For instance, PG conceptuses in women can result in ovarian teratomas, while AG pregnancies can lead to complete

hydatidiform moles (CHMs), which form when an enucleated oocyte is fertilised by one or two sperm (253,311–313). Partial hydatidiform moles (PHMs) may arise when a replication-defective oocyte is fertilised by two sperm. Furthermore, Sagi *et al.* (2019) reported that 95 oocytes were required to derive six different aESC lines (7.4% efficiency), 43 oocytes were needed to derive eight pESC lines (19.5% efficiency), and seven oocytes were used to generate two biparental ESC lines (28.6%) (305). Thus, the survival rate of CL AG or PG embryos and the efficiency of ESC line derivation from uniparental embryos remain very low.

## 1.13. Human placenta

The human placenta is the largest transient foetal organ that supports the embryo after the blastocyst implants into the uterine wall until delivery (**Figure 1.7**) (314,315). The placenta performs a plethora of functions, including the secretion of pregnancy-associated hormones, mediation of nutrient, gas and waste exchange between the developing conceptus and the mother, modulation of the immune response, protection against circulating pathogens and modification of the mother's metabolic system (314,315). Therefore, it is not surprising that defects in the placenta can immediately impact the health of both the growing foetus and the mother and may also influence the health of the individual later in life – a concept referred to as the Foetal Origins Hypothesis or Developmental Origin of Health and Diseases (DOHaD) (316,317). In addition, the human placenta is also a unique foetal organ as it shows a distinct epigenome and transcriptome, unlike any somatic tissue, which I will discuss in more detail.



**Figure 1.7. Different stages of human placental development.**

(A) Blastocyst implantation into the uterine wall with trophoblast (TE) differentiating into cytotrophoblasts (CTBs; yellow) and the formation of the primary syncytium (orange); pre-lacunar stage. (B) Formation of primary villi and the emergence and fusion of lacunae within the syncytial

mass (orange). (C) Third-trimester placenta with cotyledons; the pink-to-blue shading in the intervillous space represents oxygen levels in maternal blood. (D) Maternal-foetal interface, showing major placental cell types. LE - luminal epithelium, EXM - extra-embryonic mesoderm (pink), EEC - extra-embryonic coelom, GE - glandular epithelium, YS - yolk sac, LAC - lacunae, ICM - inner cell mass, EPI - epiblast, CS - cytotrophoblastic shell, AC - amniotic cavity, STBs - syncytiotrophoblasts (orange), iEVTs - interstitial extravillous trophoblasts (dark green), eEVTs - endovascular extravillous trophoblasts (light green), MEC - maternal endothelial cells, SMC - smooth muscle cells, FBs - fibroblasts, HBs - Hofbauer cells, TE - trophoblast.

### 1.13.1. Early development of the human placenta

The development of the human placenta starts at the early blastocyst stage at around 5 days post-fertilisation (dpf) when the polar side of the blastocyst's TE that is adjacent to the ICM attaches to the upper layer of the uterine mucosa known as the endometrium (**Figure 1.7A**) (314,315,318). It is suggested that the initial contact between the blastocyst and the epithelial cells of the endometrium is mediated by glycoproteins, such as galectins and selectins found on TE cells, while endometrial epithelium expresses selectin ligands (318–320). In addition, this initial interaction is further strengthened by the heparin-binding EGF-like growth factor (HBEGF) present on the endometrial epithelium that can bind to heparan sulphate proteoglycan and EGF receptors on TE (321,322). This interaction can induce apoptosis of the endometrial epithelium, exposing extracellular matrix components (ECM) of the basement membrane (322). Finally, the blastocyst utilises integrins to bind to diverse ligands on the endometrium, which further aids in implantation (318).

Around 6–8 dpf, different genes become upregulated in TE cells, leading to their proliferation or asymmetric mitosis (71). Some daughter cells remain as progenitor cells, while others exit the G<sub>0</sub> phase of the cell cycle and upregulate *GCM1* (323–325). This leads to the expression of *ERVW-1* and *ERVFRD-1*, as well as *ADAM12*, which triggers cell membrane dissolution, cytosolic fusion and the gradual formation of multinucleated cells known as STBs (323–326). The STBs gradually merge to establish a primary syncytium, a multinucleated monolayer of cells (**Figure 1.7A**). The primary syncytium is highly invasive and quickly penetrates the endometrium, which during the pregnancy is transformed into a specialised tissue called the decidua (314,315). The invasion of the primary syncytium is mediated by hormones secreted by the maternal decidua, while the syncytium itself secretes pregnancy-related hormones such as human chorionic gonadotrophin (hCG) and placental lactogen (327–329). At this point, the ICM of the blastocyst differentiates into Epi and PrE or hypoblast, while the primary yolk sac and the

amniotic cavity (AC) are also formed (**Figure 1.7B**). It is suggested that PrE gives rise to the extra-embryonic mesoderm (ExM) in the growing embryo, which subsequently plays a crucial role in forming placental vasculature (314,315,330). However, Epi cells can also contribute to ExM, as they show the expression of some Epi markers, such as *CREB3L1* (330,331). The growing embryo is completely embedded into the maternal decidua, covered by the surface epithelium, and fully surrounded by the primary syncytium (314,315,332). The rest of the blastocyst TE cells that remain unfused and are able to proliferate are known as placental CTBs, which are placental stem cells (314,328,329,333). CTBs divide and fuse to expand the syncytial mass in which fluid-filled spaces gradually appear, known as lacunae (**Figure 1.7B**). Lacunae further enlarge and merge, partitioning the syncytial mass into a complex lattice of trabeculae (332,334). These trabeculae expand and erode into decidual glands, which provide nourishment for STBs and the growing conceptus. The maternal glands produce glucose oligomers and glycoproteins (histiotrophic nutrition) (335). This form of nutrition persists until 11 weeks of pregnancy; then, the haemochorial system becomes fully established.

At around 12 dpf, CTBs start to rapidly proliferate and push through STB trabeculae by forming primary villi (STBs form the outside layer, and the inner core contains CTBs) (**Figure 1.7B**) (314,315,332,336). The CTBs eventually penetrate the syncytium by forming the cell columns that anchor the developing placenta to the decidua. These CTB-derived cell columns merge laterally to form a CTB shell (precursor of the basal plate) that gradually envelopes the growing conceptus and separates the primary villi from the maternal decidua (~15 dpf). Around this point, the blastocyst forms three germ layers after undergoing gastrulation and the formation of the amnion (217,330,336).

Between 17 to 18 dpf, ExM cells invade the CTB core of the primary villi and secondary villi are formed (314,315,332). The ExM further develops into primitive endothelial tubes, which are the first embryonic vessels, leading to the formation of tertiary villi (**Figure 1.7C, D**) (332,337). Tertiary villi further expand and ramify by cycling through these several stages (the formation of primary villi - trophoblast sprouting, secondary villi - ExM invasion, and tertiary villi - vessel formation) that expand the placental vasculature (332,338). Erythropoietic foci in the villi and chorionic plate produce nucleated red blood cells that migrate to the developing foetus (~22 days dpf). Also, CTBs secrete endothelial growth factor (VEGF) and placental growth factor (PlGF) that promote vessel formation (vasculogenesis, branching and angiogenesis) as their receptors are found on STBs and placental endothelial cells (339). Initially, vessels are formed through vasculogenesis at the tips of growing villi, but later, the vessels mainly grow through angiogenesis (332,340,341). Finally, the complex network of placental vessels is pruned and refined, and from 12 weeks, vessels in the chorionic plate become muscularised.

After 4 weeks of the pregnancy (~32 dpf), the placental vessels connect with the foetal vessels and umbilical circulation becomes established (338). The vascular network is significantly expanded during the second trimester to meet the increasing demands for nutrients and oxygen as the growing foetus develops. During the third trimester, the whole placenta contains between 15 to 28 fully mature villous trees, also known as cotyledons, that anchor the chorionic plate (the foetal side) to the basal plate (the maternal side) (**Figure 1.7C**) (340). Some branches of the cotyledons are attached to the basal side to provide structural support, while other branches, known as terminal villi, contain several foetal capillaries (4-6) that push against the basement membrane of STBs and have a grape-like appearance (**Figure 1.7C, D**) (340). These terminal villi freely float in the intervillous space, which is flooded with maternal blood, where nutrient and gas exchange are most efficient.

### 1.13.2. Spiral artery remodelling and other placental cell types

Around 15 dpf, some CTBs of the cytotrophoblastic shell or from the tips of CTB columns facing the maternal decidua undergo polyploidisation, senescence, and epithelial-to-mesenchymal transition (EMT) to become EVT<sub>s</sub> (**Figure 1.7C, D**) (342). One population of EVT<sub>s</sub>, known as interstitial EVT<sub>s</sub> (iEVT<sub>s</sub>), migrates into the decidual stroma and moves towards the maternal spiral arteries (343). These cells can move deeper into the myometrium, where they lose their migratory properties and differentiate into multinucleated placental bed giant cells (**Figure 1.7D**), which then lose their migratory properties and produce lactogen and *PLAC8* (344,345). The decidua also secretes diverse basement membrane proteins, including fibronectins and laminins, that support the invading cells (346). iEVT<sub>s</sub> are responsible for remodelling spiral arteries, as they express different metalloproteinases and interleukins, such as IL-6 and IL-8, that activate the endothelial cells of spiral arteries (**Figure 1.7D**) (342,347). Following this, the endothelial cells produce cytokines, which attract uterine natural killer (uNK) cells that begin breaking down the ECM of vessel walls with the support of iEVT<sub>s</sub> (348). Gradually, smooth muscles of the vessels undergo dedifferentiation or apoptosis, while some endothelial cells also undergo apoptosis and are replaced by fibrinoids that contain embedded iEVT<sub>s</sub>. At the end of remodelling, these vessels contain some remaining endothelial cells with fibrinoids and low numbers of immune cells in the vicinity. Overall, the vessels lose the ability to contract (vasoconstriction), and the mouths of the spiral arteries open widely, allowing for higher blood flow under lower pressure, which is vital

for efficient placental function and foetal growth (349). However, the maternal blood flow can be restricted by upstream radial arteries, which can undergo vasoconstriction (350).

The second population of EVTs, known as endovascular extravillous trophoblast cells (eEVTs), move along the walls of spiral arteries and plug them to prevent the maternal blood from getting into the intervillous space until the foetal-maternal circulation is established (**Figure 1.7C**) (343,351). This is suggested to prevent reactive oxygen species from getting into the intervillous space that might damage the growing foetus until proper circulation is established. After eEVT plugs are disintegrated, the oxygen levels significantly increase from 2.5% to 8% (352,353). It is suggested that hypoxic conditions promote EVT differentiation during the first trimester until the establishment of foetal-maternal circulation. From this point, the haemochorial interphase becomes established, and then glycolysis-based nourishment is replaced by oxidative phosphorylation. During this time, the placenta produces catalase and superoxide dismutase enzymes that protect the villi and the foetus from damaging reactive oxygen species (354).

Some of EVT markers include *CDH5*, *MMP2*, *MMP3*, *MMP9*, *HLA-G*, *ERBB2*, *HLA-C*, *HLA-G* and *CD56* (314,333).

STBs have microvilli that even further increase the surface area to maximise gas and nutrient exchange between the growing foetus and the mother (355). STBs form a large multinuclear monolayer without cell walls, which is believed to enhance the diffusion of nutrients, gases, and foetal metabolic waste, as well as protect the foetus from external pathogens (**Figure 1.7C, D**). Both sides of STBs are enriched with amino acid and glucose transporters (the apical and basal sides) (356). In addition, STBs secrete hCG, placental lactogen, progesterone and leptin into the maternal circulation to modulate the mother's metabolism (327). STB cells do not express HLA receptors, which help the placenta and the foetus remain invisible to the mother's immune system (314). Additionally, STB cells contain the neonatal Fc receptor (FcRn), which transports maternal IgG to the foetal blood and can activate foetal natural killer (NK) cells (357). Finally, STBs express *ERVW-1* and *ERVFRD-1*, which are essential for syncytial formation (242).

During the maturation of placental villi, the continuous CTB layer beneath the syncytium becomes patchy, and only a single syncytial layer separates placental villi from maternal blood (314,358). CTBs express *GATA2*, *GATA3*, *TFAP2C/A*, *TEAD4*, *KRT7*, *TP63*, and some surface markers such as *EGFR*, *MET* and some members of the WNT family and *NOTCH1* that become restricted to CTB columns (314,328,333).

The villus stroma contains different cell types, including immune cells such as Hofbauer cells (HBs; *CD68* positive), fibroblast (*VIM* positive) and endothelial cells that are *CD34* positive (314,359,360) (**Figure 1.7D**).

### 1.13.3. Placental epigenome

The human placenta possesses a unique epigenome, characterised by distinct distributions of DNA methylation and histone PTMs compared to other embryonic and somatic tissues (361). These differences in epigenetic marks are likely associated with the diverse functions of the placenta and its environmental adaptability or plasticity, which are essential for a healthy pregnancy.

The human placenta exhibits global hypomethylation compared to other somatic tissues. In one of the earliest studies, conducted by Schroeder *et al.* (2013) (362), third-trimester placental samples and somatic tissues, including the cerebral cortex, cerebellum, NK cells, and kidney, were collected for WGBS using MethylC-seq and Illumina Infinium 450K arrays. The study found that placental chorionic villi had lower global DNA methylation levels (5mC - 62.44% and 63.39% in two technical replicates) compared to the cerebral cortex (5mC - 77.38%), cerebellum (5mC - 75.73%), NK cells (5mC - 78.97%), and kidney cells (5mC - 76.8%). Additionally, placental samples from different species, including rhesus monkey, squirrel monkey, mouse, dog, horse, cow, and opossum, confirmed the observation that somatic tissues generally exhibit higher global DNA methylation levels than their respective placentae (363). Interestingly, direct comparisons between species revealed that gene bodies of highly expressed genes tend to become hypermethylated.

More recently, Yuan and colleagues (364) collected first- and third-trimester placental samples and isolated four major cell types from the placental chorionic villi using fluorescence-activated cell sorting (FACS). These cell types included placental endothelial cells, stromal cells, HBs and trophoblasts (primarily CTB cells). Isolated DNA from these cells was used for sodium bisulphite conversion and Illumina Infinium MethylationEPIC arrays to generate methylation reference datasets (**Section 2.8**). Analysis of these datasets revealed that the major contributing factors to variable DNA methylation levels were different placental cell types, followed by gestational age and gender. In the term placenta, trophoblast and HB cells exhibited the most distinct methylation profiles, with both cell types harbouring the largest number of differentially methylated CpGs (placental trophoblasts = 135,553 CpGs; HB cells = 130,733 CpGs) compared to endothelial (75,525 CpGs) and stromal (80,153 CpGs) cells. Most differentially methylated CpGs in

trophoblasts were hypomethylated, while in HB cells, these sites were primarily hypermethylated. Stromal and endothelial cells showed intermediate methylation levels between trophoblasts and HB cells. Notably, similar methylation patterns observed in the term placenta were also present in first-trimester placental samples, though global DNA methylation levels were lower in the first trimester. As gestational age progressed, global DNA methylation increased in placental trophoblasts, HB, and stromal cells, while it decreased in endothelial cells. Overall, DNA methylation increased with advancing pregnancy, potentially linked to the increased proportion of trophoblasts in the term placenta. It was also observed that DNMT expression increased with gestational age, while the expression of TET1 and TET3 enzymes was reduced in term placentae compared to second-trimester samples (365). Although TET2 was highly expressed in term placentae, its cofactors were either downregulated or not expressed. Reduced TET enzyme activity and higher DNMT activity may explain the increase in global DNA methylation as pregnancy progresses.

The placenta shares a similar methylation profile with cancer cells. Several studies have found that the placenta contains large PMDs, which are over 100 kb in length, with methylation levels below 70% (362–364,366). Schroeder and colleagues (367) estimated that PMDs cover 37% of the human genome. These hypomethylated regions were also observed in several cell lines, including IMR90 (foetal lung fibroblasts) (41) and SH-SY5Y (neuroblastoma cells) (368), as well as in cancers such as colon (365,369) and breast (370) cancers. PMDs were maintained throughout gestation, although DNA methylation within PMD regions decreased over time (365). In general, PMDs are gene-poor regions that exhibit low gene expression, with the promoters of these genes being hypermethylated compared to other genomic regions (362). Such genes included tissue-specific genes important for somatic tissues, such as those involved in neuronal development. Interestingly, these hypomethylated regions were flanked by highly methylated domains (HMDs), which contained highly expressed genes with hypermethylated gene bodies (362,363). Notably, PMDs were predominantly found in placental trophoblasts rather than other cell types (364). The function of these regions remains unclear, but it is suggested that they may never have acquired *de novo* methylation following global epigenetic reprogramming in pre-implantation embryos or that they may play a role in regulating placental gene expression (362,366).

The placenta may exhibit a lower global methylation profile as it originates from hypomethylated progenitor cells. One study (216) cultured human blastocysts (day 6) in 8-well plates until day 14 and collected embryos at various time points for scRNA-seq, single-cell bisulphite sequencing (scBS-seq), and single-cell tagged reverse transcription sequencing (STRT-seq). scBS-seq datasets revealed that the fastest increase in DNA

methylation occurred between day 6 (5mC - 23.5%) and day 10 (5mC - 46.3%). After day 12, DNA methylation in TE cells exceeded 50%. In contrast, a more rapid increase in DNA methylation was observed in Epi cells, with levels rising from 26.1% on day 6 to 60% on day 10, though no methylation was reported in Epi cells by day 12. This suggests that TE cells exhibit slower DNA remethylation dynamics.

The placenta harbours more imprinted genes than other somatic tissues (371–373), with placenta-specific imprints mostly maintained in trophoblasts (364). Profiling methylation datasets from four placental cell types revealed that ICRs of canonical imprints showed intermediate methylation levels and were maintained across all cell types. Interestingly, placenta-specific imprinted DMRs were predominantly found in trophoblasts, while these same regions were hypomethylated in HB cells. In stromal and endothelial cells, these placenta-specific DMRs displayed variable methylation. Some placenta-specific DMRs exhibited consistent methylation levels across trophoblasts, stromal, and endothelial cells; an example is the *DNMT1* placenta-specific imprinted gene. However, most regions were primarily methylated in trophoblasts, with stromal and endothelial cells exhibiting lower methylation, as seen in regions such as *DCAF10* and *FGF8*. Only one placenta-specific DMR, *RASGRF1*, was found to be more highly methylated in stromal and endothelial cells than in trophoblasts. Thus, placenta-specific imprinting may be restricted to extra-embryonic cell lineages.

Different histone PTMs also show unique distributions in the human placenta. In addition to large PMDs, smaller hypomethylated regions were found enriched with PRC1 and PRC2 repressive marks, covering less than 1% of the human genome (362). These polycomb-regulated regions could be detected within HMDs, where genes such as *DLX5* and *DLX6* exhibited high expression. Polycomb-regulated regions were also found within PMDs with repressed genes, resulting in even lower DNA methylation levels. Interestingly, these polycomb-enriched regions exhibited higher methylation in the human placenta than compared to other somatic tissues, such as the cerebellum. Furthermore, a recent study (143) profiled histone PTMs in placental cell lines, including hTSCs, hTSCs differentiated into STBs and EVT, CTB cells and hESCs. It was found that bivalent domains were rare in hTSCs, CTBs, STBs, or EVT, unlike hESCs, which contained multiple such regions. A few genes present in bivalent domains in trophoblast cell lines showed low expression, although generally higher expression than genes marked only by H3K27me3 or with no histone PTMs. Notably, trophoblast marker genes such as *KRT7*, *GATA3*, and *MSX2* were located in bivalent domains in hESCs. Thus, the authors concluded that bivalent domains are uncommon in placental trophoblasts.

Further profiling of histone PTMs between these cell lines showed that hTSCs contained

multiple common and 1,661 unique regions marked by H3K4me3 (143), which were mostly lost in STB cells. These hTSC-unique regions included genes such as *TEAD4* and *TP63*, which are associated with epithelial cell proliferation, tissue remodelling, and other functions. Similarly, STB cells gained 646 unique H3K4me3 domains, including genes such as *TBX3* and *GCM1*, which are important for hormone peptide production, metabolism and related functions. Similar observations were made when comparing hTSCs to EVT. hTSCs lost 888 H3K4me3-specific domains upon differentiation into EVTs, while EVTs gained 1,042 H3K4me3 domains, including genes such as *ASCL2* and *MMP2*, associated with lipid storage, immune gene regulation, and placental development. The breadth of H3K4me3 domains correlated with gene expression levels in trophoblast cell lines, and lower H3K4me3 enrichment or its absence led to gene downregulation in hTSC, STB, or EVT cells. Similarly, H3K27ac marked some common and unique enhancer elements in hTSCs, STBs and EVTs, and these enhancers were associated with processes such as morphogenesis, metal ion transport, and other cell-line-specific functions.

Zhang and colleagues (365) recently profiled epigenetic modifications in second- and third-trimester placental samples, finding that PMDs were enriched with H3K9me3, which marked gene-poor, hypomethylated regions. The authors suggested that H3K9me3 might repress genes in CTB cells located in these lowly methylated regions. Additionally, it was observed that H3K9me3, H3K27me3, H3K4me3, and H3K27ac were more abundant in second-trimester placentae and demonstrated reduced levels in term placentae, potentially linked to placental senescence and coming delivery. Interestingly, placental samples affected by PE showed a global increase in H3K27ac even during the third trimester, leading to the expression of genes, such as the pregnancy-specific glycoprotein gene (*PSG*) cluster, that was repressed in normal, term placentae. Another study investigating histone PTM enrichment in placental CTB and STB nuclei found that CTB cells contained higher levels of H3K27me3 and H3K9me3 than STBs in term placentae, despite STB nuclei showing increased nuclear condensation compared to CTBs (374).

#### **1.13.4. Placental transcriptome**

The human placenta has a unique transcriptome, possibly linked to its unique epigenome (361,375). As previously discussed, LTR elements are essential for normal placental development. Some of the best-known genes include *ERVW-1* (encodes Syncytin-1) and *ERVFRD-1* (encodes Syncytin-2) (242), which are required for CTB fusion and the formation of STBs, eventually leading to syncytial formation. It is not surprising that such

genomic regions are often found to be more hypomethylated in the placenta compared to other somatic tissues (364). Many of these elements function as enhancers and can be marked by H3K27ac, H3K4me1, and H3K4me3 that regulate placenta-specific gene expression (138). Thus, they are thought to contribute to the evolution of genomic imprinting (236,240). As noted earlier, genomic imprinting is more prevalent in this tissue, with the placenta harbouring its specific imprints in addition to canonical imprinted genes, which may play a critical role in pregnancy (20–22).

The placenta has one of the least complex transcriptomes in terms of transcribed protein-coding genes compared to 50 other tissues in the GTEx database (375). The complexity of the placental transcriptome was comparable to such tissues as the oesophagus, minor salivary gland, and pituitary gland, while blood demonstrated the least complex transcriptome. This is partially because 71 genes are highly expressed in the placenta, including *CSH1*, *CSH2*, the *PSG* cluster, *CGB3*, *CGB5*, *CGB8*, *ERVW-1*, *ERVFRD-1*, *ERVV-1*, *ERVV-2*, and other important genes during pregnancy. However, the human placenta also shows high expression of small RNAs, such as piRNAs, circular RNAs, and microRNAs (miRNAs). Some of the imprinted miRNA clusters are exclusively expressed in the placenta, such as *C19MC* (376), which encodes 58 mature miRNAs, and *C14MC* (377), which encodes 73 mature miRNAs.

### 1.13.5. Placental pathologies

Abnormal function of the placenta can result in several pregnancy complications that are generally referred to as Obstetric Disorders (378). These include PE, foetal growth restriction (FGR), intrauterine growth restriction (IUGR), miscarriage, stillbirth and others. These pregnancy-related complications can not only have an immediate effect on the growing foetus and the mother but also cause some health-related problems later in life. Here, I will discuss a few more frequent placenta-related pathologies.

#### 1.13.5.1. Pre-eclampsia (PE)

PE is a pregnancy-related condition that is defined by a sudden onset of hypertension (> 20 weeks of gestation) with a systolic blood pressure  $\geq$  140 mmHg or a diastolic blood pressure  $\geq$  90 mmHg, and one or more additional complications, including uteroplacental dysfunction, abnormal renal dysfunction or hepatic dysfunction and/or other maternal organ dysfunction (379). It is suggested that this condition globally affects 4 million

women yearly, resulting in more than 70,000 and 500,000 women and newborn deaths (380). Therefore, it is a serious, life-threatening condition not only for the developing conceptus but also for the mother. Missed diagnosis of PE or not managed adequately, this condition can rapidly progress and can lead to severe headaches, eclampsia (seizures), low platelet count (HELLP) syndrome, renal failure, pulmonary oedema, placental abruption, haemorrhagic stroke or arterial stroke and multiple other symptoms, including death (341,379,381). In general, this condition is more prevalent in developing countries (low- and middle-income), especially in South America and Africa, than in higher-income countries, most likely due to less accessible healthcare services (382). Also, it was shown that certain ethnic groups are more prone to develop PE during pregnancy, with a higher risk observed in black women and women of South Asian descent (383).

Several factors have been identified that are associated with an increased risk of PE, such as a family history of PE, previous pregnancy with PE, current chronic disease or hypertension, diabetes, obesity and use of assisted reproduction technologies (ART) (379,384). The other less predictive factors include advanced maternal age, first pregnancy or previous unsuccessful pregnancies. However, none of these factors are strong predictors for the onset of PE.

Depending on the onset of this pregnancy-associated disease, it is separated into pre-term PE (<37 weeks of gestation), term ( $\geq$ 37 weeks of gestation) and post-partum (381). It is also differentiated by the severity of symptoms, such as mild PE, which includes lower blood pressure (>140/90 mmHg), proteinuria or an increase in either albumin or creatine (379–381). In comparison, severe PE is characterised by extremely high blood pressure (>160/110 mmHg) and one more symptom frequently including HELLP syndrome, haemolysis, and elevated production of liver enzymes that are released in maternal circulation (381,385). Severe PE cases usually show earlier onset and are associated with worse pregnancy outcomes, often resulting in FGR (381).

Depending on the timing of this disease - pre-term or term, it is thought that the underlying aetiology of PE might be different, although both maternal and foetal-placental components might be overlapping (379). Pre-term PE is believed to be caused by reduced proliferation and reduced migration of placental EVT<sub>s</sub>, resulting in incomplete uterine artery remodelling (386). Also, defective decidualisation of the uterine endometrium, abnormal gene expression by decidual cells leading to reduced recruitment of EVT<sub>s</sub>, and resistance to spiral artery remodelling have been suggested as contributing factors (387,388). For example, a reduced population of T<sub>reg</sub> (FoxP3<sup>+</sup>) cells are observed in the maternal decidua of women with early pre-term PE (389). It is suggested that they help modulate the mother's immune response and are required for developing immunological

tolerance to the allogenic foetus (390). They suppress the functions of cytotoxic T cells and uNK cells by releasing cytokines (IL-10 and TGF- $\beta$ ) or direct contact inhibition (391). In contrast, the term PE is suggested to be caused by the earlier senescence of villi STBs or earlier placental senescence (392), possibly due to maternal lifestyle factors.

In general, PE is treated by lowering blood pressure and keeping it at manageable levels with oral antihypertensive drugs, although blood pressure inevitably increases as the pregnancy progresses. In severe cases, pre-term birth can be induced as this can alleviate PE-associated symptoms, but this can increase risks associated with the premature birth of the neonate (380,381).

PE can have an immediate and long-lasting impact on the health of the mother and the newborn. It is not surprising that PE is frequently accompanied by FGR due to placental dysfunction as the foetus develops in hypoxic conditions with reduced maternal nutrient supply (393,394). Such newborns show reduced weight and a smaller placental size, with some damage observed in the placental villi after the delivery (379). Women diagnosed with PE show an increased burden for a range of diseases later in life (379). Thus, PE is a complex disease that is likely caused by abnormal expression of multiple genetic loci in the mother, foetus, and placenta, with these factors being further influenced by various environmental exposures.

#### **1.13.5.2. Mechanism of PE**

By the end of the first trimester, when foetal-maternal circulation is fully established, incomplete remodelling of the spiral arteries can result in elevated blood pressure (343). These arteries may still constrict, resulting in a hypoxic environment within the placenta. The increased velocity of maternal blood flow into intervillous space can create vortices that can damage the placental villi (349,395). In addition, blood might not efficiently escape the spiral arteries (placental reperfusion), which can lead to the accumulation of reactive oxygen species that can induce stress in villi STBs that show decreased efficiency in their diverse functions (349,396). Also, rapid blood flow can damage the syncytial membrane and mitochondria within it, leading to the release of more reactive oxygen species (397). As a result, STBs may undergo apoptosis and shed genomic DNA (gDNA) and other particles into maternal circulation or can secrete inflammatory cytokines, including IL-1 $\beta$  and IL-19 and anti-angiogenic factors such as soluble FLT1 and soluble ENG into the maternal circulation (398). This can affect maternal endothelial cells, leading to systemic inflammation and the development of PE in the mother (379).

As noted earlier, PE is a multifactorial pregnancy-related disease that can affect multiple organs in pregnant women. Thus, abnormal protein levels and other molecules in maternal urine and peripheral blood, in addition to hypertension, are used to diagnose and monitor this condition (380,381). Different placental-secreted factors in maternal blood have been used to predict the onset of PE, with PlGF and soluble anti-angiogenic molecule FLT1 being the most promising and currently implemented in pre-diagnostic tests (399–401). *FLT1* is highly upregulated in women diagnosed with PE, while *PGF* (PlGF gene) is normally downregulated. Thus, the ratio of these two factors can be used to predict and diagnose the onset of PE (402).

Placentae affected by PE show genome-wide aberrations in DNA methylation, especially in pre-term PE cases (403,404). Such cases demonstrate altered methylation at *cis*-regulatory elements controlling diverse genes (403). Some of the identified genes, including *FLT1*, *INHBA* and *WNT2*, showed altered expression and hypomethylation (403,404). Also, *VEGF* and *JUN* (both genes important for vasculature formation) demonstrated higher levels of DNA methylation and H3K9me<sub>3</sub>, resulting in the downregulation of these genes in affected placental samples (405). Furthermore, it was demonstrated that DNA methylation could be used to distinguish between mild PE and severe PE cases, and overall, abnormally methylated regions were associated with genes possibly implicated in seizures, viral infections, immune system diseases, and other PE-associated complications (406). Finally, a recent study reported that placentae affected by PE had a higher level of H3K9me<sub>3</sub>, and especially H3K27ac, which was associated with the overexpression of multiple genes, including pregnancy-specific glycoproteins that in normal placentae were downregulated (365). Also, the H3K27ac profile of most PE-affected placentae was more similar to the second-trimester placentae, suggesting a developmental delay. Several imprinted genes, such as *CDKN1C* (407) or *C19MC* loci (408), were also associated with PE.

### 1.13.5.3. Intrauterine growth restriction (IUGR)

Another frequent pregnancy complication is IUGR (409), which is estimated to affect 3 to 10% of singleton pregnancies (410). This condition is characterised by a significant reduction in foetal growth *in utero* (409,411). IUGR is defined by clinical features of malnutrition and evidence of reduced growth regardless of an infant's birthweight percentile. IUGR is the pathological counterpart of small for gestational age (SGA). IUGR newborns are frequently premature, which is associated with an increased risk of perinatal morbidity and mortality. The clinical definition of IUGR includes signs of malnutrition, such as the absence of buccal fat, decreased skeletal muscle mass, and reduced

subcutaneous fat tissue (411). These foetuses also exhibit *in utero* growth restriction, including reduced height and occasionally smaller head circumference.

IUGR can sometimes be confused with SGA, as both conditions result in reduced size and weight below the 10th percentile (409,411,412). However, SGA foetuses do not display signs of malnutrition. While infants with IUGR often experience catch-up growth after birth, later in life, they may show short stature, poor academic performance (or cognitive impairment), and behavioural issues, such as hyperactivity. Furthermore, individuals with a history of IUGR may have a higher risk of developing metabolic syndromes, including diabetes, insulin resistance, liver and kidney disease, cardiovascular diseases, Alzheimer's disease, and other conditions (409,411,413,414).

Like PE, IUGR is more common in developing countries, with the highest rates observed in Asia, Africa, and Latin America (415).

IUGR is believed to result from various foetal, placental, and maternal factors or a combination of these (409,416–418). Maternal risk factors include advanced maternal age, hypoxia due to high altitudes, ethnicity, certain medications or substance abuse (including smoking), and others. Foetal factors include chromosomal abnormalities, genetic syndromes, metabolic disorders, and multiple gestations. Placental factors can involve low placental weight, avascular villi, decreased redox regulation enzymes, placental infections, and dysfunction.

IUGR is more common in twin pregnancies, particularly in monozygotic twins (10%) that share the same placenta (410), where one twin may develop IUGR. In such cases, the affected twin's placental region often shows advanced villous maturation, infarction, and thrombosis. The affected twin is significantly smaller and exhibits severe malnutrition. One study (419) collected placental samples from 8 pairs of monozygotic monozygotic twins, with one twin affected by IUGR and the healthy twin used as a control. Using the Illumina Infinium HumanMethylation450K BeadChip array, researchers identified differentially methylated regions in these samples. Eight such regions were found to overlap with gene promoters, including *DECR1*, *ZNF300*, *DNAJA4*, *CCL28*, *LEPR*, *HSPA1A/L*, *GSTO1*, and *GNE*, with six of these regions being hypermethylated. The three most differentially methylated regions - *DECR1*, *ZNF300*, and *LEPR* - were validated by pyrosequencing, playing an important role in unsaturated fatty acid (FA) oxidation, lipid metabolism, and transcriptional repression.

IUGR is a common phenotypic feature associated with several imprinting disorders, including transient neonatal diabetes mellitus (TNDM), Temple syndrome (TS14), Prader-

Willi syndrome (PWS), and especially in SRS (420). Previously, Monk group analysed 67 Illumina Infinium Human Methylation450 datasets, including 23 healthy placental samples, 31 affected by PE, and 13 affected by non-syndromic IUGR (295). Profiling these datasets revealed that several IUGR cases were hypomethylated at the *H19* DMR, a finding further supported by pyrosequencing. These samples exhibited upregulation of *H19* and repression of *IGF2*. Additionally, 50 samples were used for microfluidic-based quantitative expression analysis, which showed that *ZDBF2* (a canonical imprint), *GPR1-AS1*, and *ADAM23* (two placenta-specific imprints), located within the same imprinted cluster on chromosome 2, were differentially expressed between IUGR and control placental samples. Thus, several imprinted genes may be associated with IUGR. However, it remains unclear whether the altered expression of these imprinted genes causes IUGR or if IUGR itself induces changes in imprinted gene expression (361).

### 1.13.6. Human trophoblast stem cells (hTSCs)

Lee and colleagues proposed a set of criteria to identify human placental trophoblasts *in vitro* (421). Firstly, cells must express a distinctive combination of trophoblast markers such as KRT7, EGFR, HLA class I molecules, and hCG, along with other markers. As epithelial in origin, placental trophoblasts typically express *KRT7* and *EGFR* (314,328,329,333), which are expressed at very low levels in other placental cell types, according to scRNA-seq data from the Human Protein Atlas (422). However, these surface markers are not exclusive to trophoblasts and are also found in maternal decidual glandular epithelium (GE) (423). Therefore, additional markers should be evaluated in combination with KRT7 or EGFR. Secondly, all human trophoblasts are negative for HLA class II expression (421,424). Moreover, CTBs and STBs do not express HLA class I allotypes (424,425). The exception to this is primary mononuclear EVT, which exhibit high expression of HLA-G, lower expression of HLA-C and HLA-E, and an absence of HLA-A and HLA-B, which are broadly expressed in somatic cells (333,424,425). Multinuclear giant cells also express high levels of HLA-G along with hPL, whereas STBs are characterised by high expression of hCG, placental leucine aminopeptidase, aminopeptidase A and pregnancy-specific glycoproteins (421). Thirdly, the promoter region of *ELF5*, a TF essential for the self-renewal of mouse trophoblast stem cells (mTSCs) (421), is hypomethylated in human trophoblasts but hypermethylated in placental mesenchymal cells (421). Finally, human trophoblasts exhibit high expression of miRNAs from the imprinted *C19MC* cluster, which is usually hypermethylated and silenced in other somatic cell types, with the exception of hESCs (243,376,421,426). Therefore, the unique combination of surface markers, methylation patterns, and

expression profiles at specific loci, along with morphological features, can be used to distinguish human placental trophoblast lineages.

Recently, Okae and colleagues successfully established several human trophoblast stem cell (hTSC) lines derived either from CTBs isolated from first-trimester elective termination placental samples (6–9 weeks gestation; cytotrophoblast (CT) stem cell line 27, 29 and 30) or the outgrowths of cultured human blastocysts (BTS5, BTS11), using a specialised 2D trophoblast culture medium (243). These hTSC lines retained a normal karyotype, demonstrated long-term self-renewal (at least 5 months), and displayed morphological, transcriptional, and epigenetic features similar to those of CTBs. For example, hTSCs expressed genes typically upregulated in CTBs, including *GATA3*, *TEAD4*, *CTNNB1*, *TP63*, *ITGA6*, *FGFR2*, *FZD5*, and *LRP5*. Crucially, these lines demonstrated bipotency: supplementation of the culture media with NRG1, A83-01, and Matrigel induced differentiation into EVT-like cells, while treatment with forskolin led to cell aggregation, fusion and the formation of syncytia. Overall, the hTSC lines fulfilled the molecular and phenotypic criteria for trophoblast identity as defined by Lee *et al.* (2018) (421) and were thus accepted as a good model system for studying human placental trophoblast development (243).

However, subsequent studies employing a range of techniques have uncovered several limitations associated with hTSCs. Advanced transcriptomic analyses using single-cell multi-omic approaches showed that hTSCs were unable to differentiate into eEVTs or trophoblast giant cells (333). Additionally, the other study employing FACS, immunohistochemistry, and related techniques found that most hTSC lines derived from CTBs or blastocysts exhibited detectable expression of HLA-A and HLA-B, even after differentiation into EVT-like cells (425). Interestingly, culturing hTSCs under 3D conditions led to a reduction in HLA class I molecule expression. This reduction was associated with the upregulation of several miRNAs in 3D-cultured hTSCs, which suggested that miRNAs might modulate the expression of HLA class I, and that mechanical forces present in 3D culture are important for culturing these cell lines and may influence cell identity (425). Further transcriptomic profiling revealed that hTSCs exhibit a differentiation bias towards the EVT lineage and are less efficient than recently developed trophoblast organoids at differentiating into STBs (425).

Moreover, although Okae *et al.* (2018) reported that hTSCs retained a methylation profile similar to that of CTBs, these cell lines exhibited significantly lower genome-wide DNA methylation levels (243). Specifically, the average methylation level in CTBs was 52.3%, whereas CTB-derived and blastocyst-derived hTSCs displayed 33.7% and 33.6%, respectively (243). Further investigation revealed that these hTSC lines exhibited global

hypomethylation at placental PMDs, along with altered profiles of repressive histone PTMs in these regions (243,427). Intriguingly, another study showed that ectopic expression of *DNMT3L*, which is not expressed under standard hTSC culture conditions, was capable of restoring intermediate DNA methylation levels at placental PMDs (428). However, persistent overexpression of *DNMT3L* impaired the ERK-CREB signalling pathway required for the induction of the STB transcriptional program and formation of the syncytium. Despite their globally reduced methylation, hTSCs maintained the majority of imprinted genes (243), unlike hTSCs derived from hESCs or human induced pluripotent stem cells (hiPSCs), which typically show complete loss of genomic imprints. For example, among the 33 investigated placenta-specific DMRs, hTSCs derived from CTBs lost methylation at 6 loci, while 8 loci were either hypo- or hypermethylated in blastocyst-derived hTSCs. Affected DMRs included *CYP2J2*, *ZC3H12C*, *GPR1-AS*, and several others, which exhibited methylation levels below 30% or above 70% (243). Although *DNMT3L* overexpression restored normal methylation levels at PMDs and DMRs (428), the allelic analysis indicated that parent-of-origin-specific expression of imprinted genes was irreversibly lost once imprinting was disrupted.

Furthermore, hTSCs established by Okae group have also been successfully employed in genetic screening and gene-editing experiments by using CRISPR technology. Several studies have employed CRISPR-based approaches to identify key TFs, growth-promoting and growth-restricting genes essential for human placental development, as well as to investigate the role of transposable elements in placental gene regulation (138,429,430). These experiments, however, frequently encountered elevated levels of cell death and toxicity under standard culture conditions following lentiviral transfection or antibiotic selection, necessitating optimisation of the trophoblast culture media (138,430). Despite these technical challenges, several intriguing findings have emerged. For instance, TFs essential for mTSCs, such as *CDX2*, *EOMES*, *ESRRB*, and *SOX2*, were found to be dispensable and nearly undetectable in hTSCs (430). Conversely, *DLX3* and *GCM1* were required for hTSC differentiation into EVT and STB (430). *TEAD1* was shown to promote EVT differentiation but hinder STB lineage specification (429). Additionally, imprinted genes such as *CDKN1C* and *GRB10* were identified as growth-suppressive regulators, along with a few other genes, in hTSCs (430). Another interesting finding was the discovery of a transposable element, LTR10A, which functions as an enhancer for *ENG*, whose soluble protein levels are positively associated with the severity of PE (138,246,379,398). In summary, while hTSCs established by Okae group exhibit certain limitations (243), such as restricted differentiation capacity, reduced genome-wide methylation levels and a loss of PMDs or sensitivity to genetic manipulation, they remain a valuable model for investigating human trophoblast development or genomic imprinting, especially in contexts where the risk of maternal cell contamination must be

minimised (431).

## 1.14. PhD aims

The mammalian epigenome is complex and dynamic, comprising numerous *cis*-elements and *trans*-acting factors that regulate DNA methylation, chromatin condensation, and higher-order chromatin structures within the nucleus (7,432). This intricate interplay between the genome and the epigenome is particularly evident during gametogenesis and early embryonic development when parent-specific genomic imprints are erased, and new sex-specific imprints are established. At the onset of zygote formation, both maternal and paternal genomes undergo significant changes, including chromatin remodelling, changes in histone PTMs, and rapid DNA demethylation (71,136,137). These processes are crucial for EGA and normal embryonic development. However, this developmentally sensitive window is vulnerable to environmental influences, such as the quality of maternal diet or chemical exposure *in utero*, which can lead to aberrant epigenetic landscapes (433). Deviations from normal epigenetic modification patterns can result in various diseases. For instance, alterations in DNA methylation have been implicated in cancer (44), Parkinson's disease (434), Alzheimer's disease (435), and imprinting disorders (e.g., loss or gain of methylation at ICRs) (17,420,436). Thus, the correct establishment of epigenetic patterns during early development is critical for an individual's long-term health and might also impact the next generation.

ART includes a series of medical procedures designed to help individuals and couples achieve pregnancy (437,438). According to the World Health Organization (WHO), infertility is defined as a disease of the reproductive system resulting in the inability to achieve a natural pregnancy after 12 months of unprotected intercourse (439). Causes of infertility include advanced parental age, endometriosis, cervical and uterine abnormalities (e.g., polyps, fibroids, tumours), ovarian disorders, hormonal imbalances, poor semen quality, or unexplained infertility (437,438). ART is not exclusive to couples experiencing infertility, as it has recently become more widely used by single women and members of the LGBTQIA+ community (440). ART is an umbrella term encompassing a range of medical procedures, ranging from less invasive techniques such as intrauterine insemination (IUI) to more complex methods, including *c*, ovarian stimulation, and *in vitro* fertilisation (IVF), followed by fresh or frozen embryo transfer (437,438). These procedures often occur during critical periods of epigenetic reprogramming (17,441). ART has been associated with various risks, including ovarian hyperstimulation syndrome

(OHSS), multifoetal pregnancies, gestational diabetes, PE, preterm birth, FGR, SGA, placental abruption, placenta previa and stillbirth. Consequently, the global use of ART treatments is increasing, particularly as infertility rates are predicted to increase in the future (442,443). Therefore, it is crucial to understand the potential risks associated with ART-related procedures.

Several studies have reported an increased incidence of imprinting disorders following ART (444,445), although the findings remain inconsistent due to the limited study sizes and the rarity of these disorders (366). For example, Novakovic *et al.* (2019) used the Illumina Infinium MethylationEPIC BeadChip to assess DNA methylation in ART-conceived versus non-ART-conceived neonates and adults (446). The authors found minimal genome-wide differences in DNA methylation between ART and non-ART neonates, with most of these differences diminished by adulthood. Nevertheless, a slight increase in DNA methylation at the *Kcnq1/KCNQ1* differentially methylated region 1 (*KvDMR1*) region was observed in ART-conceived neonates, which disappeared by adulthood. Conversely, modest changes in DNA methylation at *NAP1L5* were detected in ART-conceived adults but not in neonates. A more recent study by Ye *et al.* (2024) utilised the Swedish national birth registry to examine the prevalence of imprinting disorders in ART-conceived children (445). Among 2,084,127 singleton births between 1997 and 2017, 63,954 (3.1%) were ART-conceived. Of these, 52 children were diagnosed with imprinting disorders such as BWS, PWS or SRS. Overall, ART-conceived children were found to have a higher risk of being diagnosed with an imprinting disorder during childhood. While this risk was somewhat reduced after accounting for parental background factors, it remained elevated among couples with infertility. Notably, ICSI combined with frozen embryo transfer was associated with significantly increased risks of BWS and PWS/SRS compared to children conceived naturally by infertile parents. The authors emphasised the need for further studies to better understand these associations, especially given the rarity of imprinting disorders and the relatively small size of the final study group.

Imprinted genes are essential for normal placental development and its function during pregnancy (371,372). Genome-wide screens of imprinted genes across various human and mouse tissues have revealed that the placenta contains a higher number of imprinted genes compared to other somatic tissues (307–309). In addition, patients with imprinting disorders frequently exhibit abnormalities in placental formation (447,448). In line with this, our group and others have shown that, unlike other somatic tissues, the human placenta retains many gDMRs (20–22). Most of these DMRs originate from oocytes, survive early epigenetic reprogramming, and, in some cases, persist uniquely in the placenta. Further investigation into these tDMRs and placenta-specific mDMRs revealed that some can regulate monoallelic expression, and some of these genes may be involved

in placental pathologies. Interestingly, both tDMRs and placenta-specific mDMRs appear to be specific to humans and, more broadly, to primates but not to evolutionarily more distant mammalian species (20,449). Unfortunately, due to limited cohort sizes and the informativeness of placental samples (e.g., enrichment of polymorphisms) in these studies, it was not possible to determine whether all identified placenta-specific mDMRs induce parent-of-origin-specific expression (20–22). Based on these observations, I hypothesise that the human placenta harbours additional placenta-specific mDMRs that result in monoallelic expression and may be specific to certain placental cell types. In this PhD project, I aimed to use our expanded placental cohort, including samples from both normal and complicated pregnancies, to characterise the methylation and expression profiles of previously identified, but not fully explored, placenta-specific mDMRs, as some may be implicated in placental pathologies and disease (20,295). The specific objectives of this PhD project were:

Objective 1: Review the literature and generate a list of genes with placenta-specific mDMRs that have not been fully characterised.

Objective 2: Use samples from our extended placental cohort alongside a range of molecular biology techniques to explore methylation patterns at placenta-specific mDMRs and the expression of corresponding candidate genes.

Objective 3: Investigate methylation and gene expression in different placental cell types using new placental samples.

Recent studies have shown that rodent oocytes, and subsequently, the maternal genome in pre-implantation embryos, harbour non-canonical H3K27me3 peaks that repress maternal alleles, resulting in paternal-biased expression (24–26,137,139,140,450). Further investigations into H3K27me3-mediated non-canonical imprinting demonstrated that maternal-specific H3K27me3 domains are replaced by sDMRs in the mouse placenta, leading to paternal-specific gene expression. Loss of H3K27me3 due to maternal *Eed* KO (a catalytic subunit of PRC2) results in the loss of non-canonical imprints and causes male-biased lethality in post-implantation stages, as such embryos demonstrated growth retardation and reduced placental size. Moreover, loss of non-canonical imprints has been observed in somatic cell nuclear transfer (SCNT) embryos, which display various placental and body growth abnormalities. More generally, the loss of non-canonical imprints disrupts the post-implantation development of SCNT embryos (451–453). To date, only one study has specifically investigated the conservation of non-canonical imprinting in human embryos, identifying a single candidate gene, *FAM101A*, that is potentially regulated by this mechanism (27). However, the study was significantly limited by a small

embryo cohort. Therefore, I hypothesise that non-canonical imprinting may be conserved in the human placenta, especially since most canonical imprinted genes are known to be conserved between mice and humans (454). In this PhD project, I aimed to use our well-characterised placental cohort and, additionally, an IVF embryo cohort to investigate whether non-canonical imprinting is conserved in the human placenta and embryos using a variety of molecular biology approaches. The specific objectives of this PhD project were:

Objective 1: Conduct a literature search on rodent non-canonical imprints, human genes with LTR-derived promoters, and human genes with placental sDMRs to generate a list of candidate genes potentially regulated by non-canonical imprinting.

Objective 2: Use well-characterised samples from our placental cohort to investigate candidate genes using diverse molecular biology techniques.

Objective 3: Generate and analyse scRNA-seq datasets from human IVF embryos to validate candidate genes of non-canonical imprinting.

Discoveries of H3K27me3-dependent and DNA methylation-dependent transient imprinting mechanisms are still recent subjects of research, and not much is known about their regulation or functional importance during human embryo development (20–26,297). Therefore, such H3K27me3-dependent and DNA-methylation-dependent transient imprinted genes could also be affected by environmental stresses and ART, possibly leading to some pathological phenotypes. Thus, further studies are needed to explore these genes and understand their evolutionary origin, regulation and functional role in placental formation, as well as to determine their approximate number in the mouse and human genomes. During this PhD project, I aimed to address some of these questions.

## **Chapter 2: Methods**

## 2.1. Ethics approval for human samples

The Ethics Committee at the Hospital Sant Joan de Déu (PI35/07) and the Faculty of Medicine and Health Sciences Research Ethics Committee at the University of East Anglia (ETH2122-0856) approved the collection of human samples for this project. Ethical committees granted permission to use placentae, cord and peripheral blood, and saliva samples collected at the Hospital Sant Joan de Déu (Barcelona, Spain) and Norfolk and Norwich University Hospital (NNUH; Norwich, UK). All mothers participating in the study provided written informed consent for themselves and their children in accordance with the declaration of Helsinki. All human samples were obtained after receiving signed informed consent from the study participants.

The use of excessive surplus embryos for research, which were received from the IVI-Valencia (IVF clinic in Valencia, Spain), was approved by the scientific and ethics committee of the Instituto Valenciano de Infertilidad (IVI; 1310-FIV-131-CS), University of East Anglia Faculty of Medicine and Health Sciences Research Ethics Subcommittee (ETH2223-1031), Bellvitge Institute of Biomedical Research, Barcelona (PR292/14), the Centro de Medicina Regenerativa de Barcelona (CMRB CEIC 10/2017), the National Committee for Human Reproduction (CNRHA) and the Regional Health Departments for Valencia and Catalunya (4/2014 & 10/2017).

## 2.2. Samples

### 2.2.1. Placental cohort and parental samples

#### 2.2.1.1. Placental samples

During this project, a large placental cohort was used for genome-wide screening of novel placenta-specific imprinted genes. The majority of placental samples were collected at the Hospital Sant Joan de Déu (Barcelona, Spain) between 2008 and 2012 from normal and complicated pregnancies resulting in live-born infants (**Appendix 1**). A smaller proportion of placentae were obtained from the Norfolk and Norwich University Hospital (NNUH; Norwich, UK) from consecutive births between 2021 and onwards (**Appendix 1**). For each received placenta, several biopsies were taken from the foetal side around the insertion side of the umbilical cord. The collected tissue was thoroughly rinsed in PBS,

snap-frozen in liquid nitrogen and conserved at  $-80^{\circ}\text{C}$  until later use unless it was used for placental cell-type enrichment with Magnetic-Activated Cell Sorting (MACS; Miltenyi Biotec), which can be found in **Section 2.4.1**.

To rule out maternal DNA contamination, all placenta-derived DNA samples obtained from the Hospital Sant Joan de Déu (Barcelona, Spain) and the Norfolk and Norwich University Hospital (NNUH; Norwich, UK) were used for microsatellite repeat analysis, during which the copy-repeat numbers of several highly polymorphic short tandem repeats (microsatellites) are compared between maternal and placenta-derived DNA following PCR amplification and size separation on a gel (455).

#### **2.2.1.2. Blood and saliva samples**

Maternal peripheral and cord blood obtained from the Hospital Sant Joan de Déu (Barcelona, Spain) were collected in EDTA tubes and frozen at  $-20^{\circ}\text{C}$  until further use. Maternal and paternal saliva samples collected at the Norfolk and Norwich University Hospital (NNUH; Norwich, UK) were collected in Oragene Saliva DNA collection tubes (OG-510; DNA Genotek Inc.) and stored at  $4^{\circ}\text{C}$  until later use.

#### **2.2.2. Human embryos**

During this project, two sets of human pre-implantation embryos were used for imprinting analysis. Different stages of human pre-implantation embryos were received from the IVI-Valencia (IVF clinic in Valencia, Spain). The first batch of pre-implantation embryos included 19 different-stage human embryos that were processed into single cells for single-cell methylation and transcriptome sequencing (scM&T-seq), as described in (71). The remainder of SMART-seq2 (Switching mechanism at the end of the 5'-end of the RNA template sequencing 2) full-length cDNA was used during this PhD project. The second batch of embryos was collected in the Eugin Barcelona (Assisted reproductive technology clinic in Barcelona, Spain) and included 15 different-stage human embryos (**Table 4.4**). These embryos were also processed into single cells for scM&T-seq. Only the single-cell transcriptome sequencing data was investigated during this PhD project.

### 2.2.3. Mouse placenta

Mouse placentae were produced by crossing *Mus Musculus Molossinus* (JF1) with C57BL/6 and were collected at E15.5. gDNA and RNA were extracted from 2 placentae that were received as a kind gift from Dr Philippe Arnaud (Institute of Genetics, Reproduction & Development (iGReD), CNRS-Universitié Clermont Auvergne-INSERM, France). Animal care and breeding were carried out following the institutional guidelines of iGReD.

### 2.2.4. Cell lines

All cell lines used during this PhD project were harvested for RNA extraction, these included:

#### 2.2.4.1. Cytotrophoblast stem cells 27 and 30 (CT27 and CT30)

CT27 (female) and CT30 (female) cell lines were established in Prof Hiroaki Okae's laboratory and were kindly given to us. These two cell lines were established from the first trimester placentae and demonstrated similar morphology, global expression and methylation profiles compared to CTB cells (243).

#### 2.2.4.2. Mole 1 and Mole 2

Mole 1 and Mole 2 cell lines were also established in Prof Hiroaki Okae's laboratory and were a kind gift. These cell lines were derived from CHM and may have originated from monospermic fertilisations as this type of molar pregnancy is predominant, and the SNP array adapted to the Japanese population revealed a loss in heterozygosity (311). In general, mole cell lines had a similar morphology, expressed trophoblast markers (*TFAP2C*, *GATA3*, and *KRT7*) and had similar expression and methylation profiles compared to CT cell lines. Interestingly, DNA methylation was lost at most placenta-specific DMRs, resulting in abnormal expression of some imprinted genes. More specifically, the *IC1* of *H19/IGF2* was hypermethylated, while *KvDMR1* was entirely lost, resulting in a very low expression of *H19* and *CDKN1C*, while *GRB10* and *NAA600* had normal expression similar to CT cell lines. Most paternally expressed genes became unregulated, except *PEG3* and *AIM1*, which retained similar expression. Finally, Okae's group reported that loss of *CDKN1C* resulted in the proliferative advantage of these cells

over CT cell lines due to loss of contact inhibition. Thus, Mole 1 and 2 can be used as loss-of-function mutants for placenta-specific imprints.

## 2.3. Molecular Biology techniques

All placental, cord blood, and maternal blood samples collected in Spain were processed into RNA and DNA by previous members of our group, including Dr Marta Sanchez-Delgado, Dr Ana Monteagudo-Sánchez, and others. The placental and parental saliva samples received in Norwich, UK, were processed by previous and current group members, Dr Louise Chappell-Maor, Dr Sarah Russell, Caitlin Bone, Becky Sainty, and Kelly Chen.

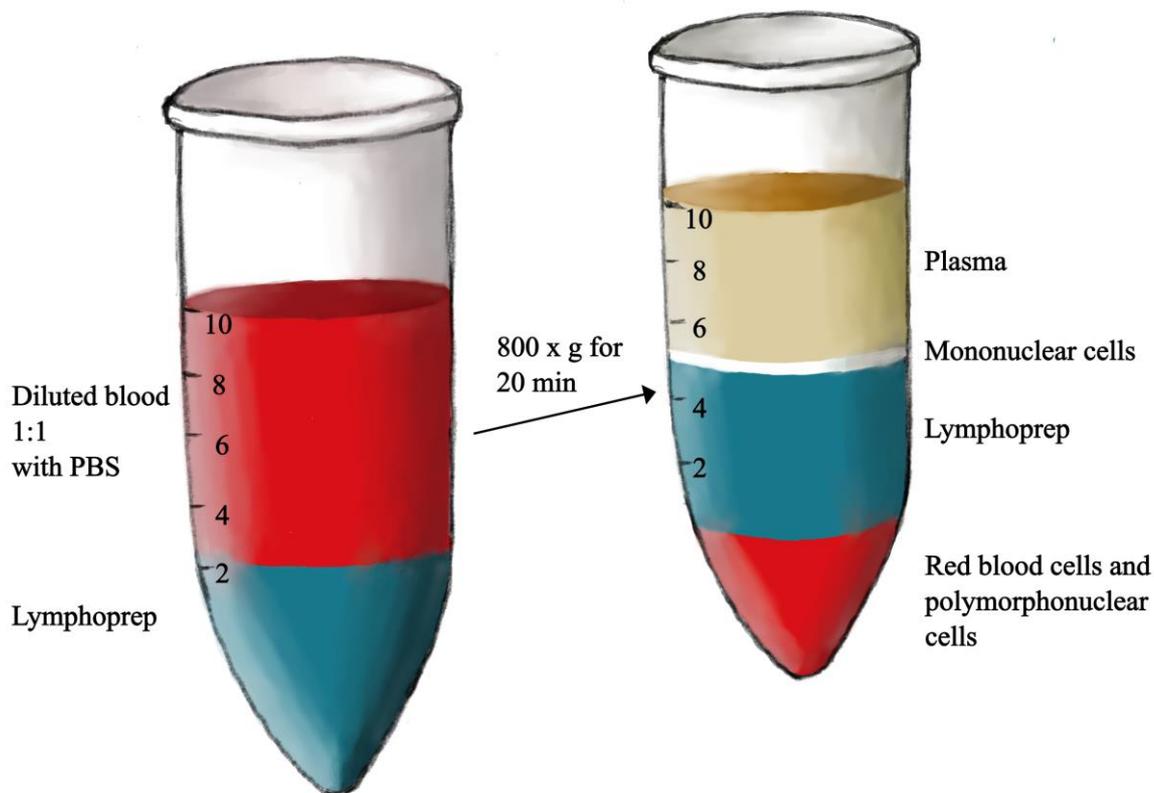
Placental cell enrichment protocol with MACS was developed and optimised by PhD student Becky Sainty. This protocol was carried out by several members of Monk group, including Becky Sainty, Dr Louise Chappell-Maor, Dr Sarah Russell, and Caitlin Bone.

All cell culture work was carried out by PhD student Kelly Chen, Dr Louise Chappell-Maor, Caitlin Bone and Dr Sarah Russell.

### 2.3.1. Mononuclear cell extraction from the cord and maternal blood

Maternal peripheral and cord blood stored in EDTA tubes were used to isolate mononuclear cells by Lymphoprep™ (AXIS-SHIELD) density gradient (**Figure 2.1**). Initially, a blood sample was diluted with an equal volume of PBS and shaken several times by inversion. Well-mixed blood was slowly layered on top of the same volume of the Lymphoprep™ solution (in a new falcon tube). A falcon tube was then centrifuged at 800 x g for 20 minutes with no break (4°C). Mononuclear cells such as monocytes and lymphocytes have a lower density than granulocytes and erythrocytes at the osmotic pressure, and therefore, after centrifugation, mononuclear cells form a single monolayer between blood plasma and the Lymphoprep™ density gradient medium. The upper plasma layer was carefully removed and discarded, and the monolayer with mononuclear cells was carefully removed and transferred to a fresh falcon tube with PBS. Finally, the

sample was centrifuged at 400 x g for 10 minutes to collect the cell pellet. The mononuclear cell pellets were later used for DNA extraction (**Section 2.3.2.1**).



**Figure 2.1. Mononuclear cell isolation from maternal peripheral blood or cord blood with the Lymphoprep™ density gradient.**

## 2.3.2. DNA extraction

### 2.3.2.1. DNA isolation from placental biopsies and blood

For gDNA extraction from a cell pellet, the pellet was washed with PBS and centrifuged at 1,000 x rpm for 5 minutes. Firstly, the cell pellet was resuspended in lysis buffer (10 mM Tris pH = 8.0, 100 mM NaCl, 10 mM EDTA pH = 8.0), while a biopsy of a placenta was homogenised with 1 mL of the same buffer to obtain a suspension. Then, 15 µL of 10% SDS and 15 µL of Proteinase K (20 mg/mL; E00491; Thermo Fisher Scientific Inc.) were added into a 15 mL falcon tube with the homogenised sample or the cell pellet that was further incubated at 56°C in an incubator overnight. The 15 mL MaXtract High Density tube (129065; Qiagen) was centrifuged at 1,500 × g for 2 minutes before loading the sample. 1

mL of the sample was added to the MaXtract High Density tube and mixed with 1 mL of the phenol and chloroform mix (0.5 mL phenol and 0.5 mL chloroform; 15593031; Fisher Scientific), shaken a few times and centrifuged at  $1,500 \times g$  for 2 minutes. This step was repeated 2 more times. Then, 1 mL of chloroform was added to the MaXtract High Density tube with the sample, shaken a few times and centrifuged at  $1,500 \times g$  for 2 minutes. This step was repeated 2 more times. At this stage, DNA had to be separated from cell debris by the MaXtract gel. DNA in the supernatant was transferred to a new 15 mL falcon tube with 2.5 mL of 100% ethanol and gently inverted a few times. The falcon with precipitated DNA was centrifuged at  $1,500 \times g$  for 2 minutes, and the supernatant was gently removed and discarded while the DNA pellet was further washed with 300  $\mu$ L of 70% ethanol and span at  $1,300 \times g$  for 5 minutes. Finally, the supernatant was carefully removed, and the DNA pellet was air-dried at room temperature for 20 minutes and dissolved in 100  $\mu$ L of UltraPure DEPC-Treated Water (UPW; 750023; Thermo Fisher Scientific Inc.). The quality and concentration of cleaned DNA were inspected with the NanoDrop spectrophotometer (ND-1000; Thermo Fisher Scientific Inc.). The precipitated DNA was considered clean if the ratio of absorbance at 260 nm and 280 nm was close to 1.8 (the average absorbance of four nucleotides) and, thus, free of contaminating proteins or other organic compounds. Only the clean DNA was used for genotyping PCR (**Section 2.3.7**) or downstream methylation analysis (**Sections 2.3.4 & 2.3.6**). gDNA was stored at  $-20^{\circ}\text{C}$  until further use.

### 2.3.2.2. DNA extraction from saliva

gDNA from parental saliva samples were extracted following the prepIT•L2P protocol (DNA Genotek Inc.). Briefly, a saliva sample collected in the Oragene Saliva DNA collection tube (OG-510; DNA Genotek Inc.) was inverted and shaken a few times prior to incubation at  $50^{\circ}\text{C}$  in an incubator for a minimum of 2 hours. This step is required for the release of DNA and inactivation of nucleases present in saliva. 500  $\mu$ L of the mixed sample was transferred to a 1.5 mL microcentrifuge tube and mixed with 20  $\mu$ L of the prepIT•L2P buffer (Cat. No. PT-L2P-5; DNA Genotek Inc.) for a few seconds. The microcentrifuge tube was incubated for 10 minutes on ice, followed by centrifugation at  $15,000 \times g$  for 15 minutes. The clear supernatant containing the DNA was transferred to a clean microcentrifuge tube, mixed with 600  $\mu$ L of 100% ethanol and left for 10 minutes at room temperature to precipitate the DNA. The microcentrifuge tube with the precipitated DNA was further centrifuged at  $15,000 \times g$  for 2 minutes. After this, the supernatant was carefully removed, while the DNA pellet was washed with 250  $\mu$ L of 70% ethanol for 1 minute. The supernatant was discarded, and the clean DNA pellet was dissolved in 100  $\mu$ L IDTE (1 x TE Solution; 11-05-01-09; Integrated DNA Technologies IDT). To ensure the DNA was entirely dissolved in TE, the microcentrifuge tube with the DNA was briefly

vortexed and left at room temperature overnight. Finally, the quality and concentration of cleaned DNA were determined with the NanoDrop spectrophotometer (ND-1000; Thermo Fisher Scientific Inc.). A ratio of 1.8 for the 260/280 ratio indicated that DNA was free from contaminating compounds and could be used for genotyping PCR (**Section 2.3.7**). gDNA was stored at -20°C until further use.

### **2.3.3. RNA extraction from the placental samples and cell lines**

RNA was extracted either from placental biopsies or from cell lines, including CT27 & 30 and Mole 1 & 2 (**Sections 2.2.1.1 & 2.2.4**). At the start, 1 mL of the TRI reagent (T9424; Merck Life Science UK Ltd.) was added into a microcentrifuge tube with a cell pellet and mixed, while in case of a placental biopsy, the mixture had to be homogenised. Then, 200 µL of chloroform was added, and the tube was shaken vigorously, followed by incubation at room temperature for 5 minutes. The tube with the sample was microcentrifuged at 12,000 x rpm at 4°C for 15 minutes. At this stage, two layers should be formed: the upper aqueous phase contained the RNA (transparent), while the lower solvent or organic part had DNA and proteins (cloudy). To precipitate RNA, the upper aqueous phase with the RNA was transferred into a new tube with 320 µL of isopropanol (0.8 x) and mixed by inverting the tube a few times. The mix was incubated at room temperature for 10 minutes, followed by microcentrifugation at 12,000 x rpm at 4°C for 60 minutes. The RNA pellet was washed by adding 200-500 µL of 70% ethanol and microcentrifuged at 12,000 x rpm at 4°C for 5 minutes. The supernatant was removed by careful pipetting and discarded. Finally, the RNA pellet was air-dried at room temperature to remove residual ethanol, and dissolved in UPW (750023; Thermo Fisher Scientific Inc.) and stored at -80°C. The quality of RNA was inspected by the NanoDrop spectrophotometer (ND-1000; Thermo Fisher Scientific Inc.). The precipitated RNA sample was considered clean if the ratio of the absorbance at 260 nm and 280 nm was close to 2.0 (the average absorbance of five nucleotides) and, thus, free of contaminating DNA, proteins or other organic compounds. Only the clean RNA was used for cDNA synthesis (**Section 2.3.5**).

### **2.3.4. DNA digestion with methylation-sensitive restriction enzymes**

For methylation analysis, two methylation-sensitive restriction enzymes were utilised:

HpaII (R0171S; NEW ENGLAND Biolabs) and BstUI (R0518S; NEW ENGLAND Biolabs). These enzymes were selected because placental DMR regions frequently contained restriction sites for at least one of these enzymes due to their high CpG content. The restriction site of HpaII is 5'-C/CGG-3' ('/' indicates restriction site), and the restriction site of BstUI is 5'-CG/CG-3'. Both of these enzymes can digest restriction sites if they are unmethylated (456,457). Although it has been reported that HpaII can nick hemimethylated DNA, but it does that extremely slowly, making it negligible (457). For the digestion, 500 ng of placental gDNA was mixed with 1 µL of HpaII or BstUI (10U/µL), 2 µL of the 10 x rCutSmart buffer and 17 µL of UPW (750023; Thermo Fisher Scientific Inc.) to make a total volume of 20 µL. The mix was carefully mixed by pipetting and microcentrifuged for 5 seconds, followed by incubation for 6 hours. Depending on the restriction enzyme used for digestion, the mix was incubated either at 37°C (HpaII) or 60°C (BstUI). To ensure digestion efficiency, 0.5 µL of the same enzyme was added to the mix and incubated for an additional hour. The digested mix was cleaned by ethanol precipitation (**Section 2.3.8.3**) and resuspended in 10 µL of UPW. Either 1 - 2 µL was used for methylation-sensitive genotyping (**Section 2.3.7.1.3**).

### 2.3.5. cDNA synthesis

For cDNA synthesis, only good-quality RNA was used (260 / 280 ≈ 2.0). In total, 1 µg of a placental sample or cell line RNA was dissolved in 8 µL of UPW (750023; Thermo Fisher Scientific Inc.). To avoid possible DNA contamination, the diluted RNA was treated with 1 µL of DNase I (10694233; Fisher Scientific) and with the same amount of the 10 x DNase I Reaction Buffer (10694233; Fisher Scientific) and incubated at room temperature for 15 minutes. The enzyme was deactivated by the addition of 1 µL of 25 mM EDTA (10694233; Fisher Scientific) and incubation in the Veriti™ Thermal Cycler (4375305; Thermo Fisher Scientific Inc.) at 70°C for 10 minutes. After this, the treated RNA sample was immediately placed on ice. In total, 11 µL of the RNA sample was mixed with 9.25 µL of cDNA synthesis Master Mix (**Table 2.1**). The mixed sample was placed into the Veriti™ Thermal Cycler at 37°C for 90 minutes for cDNA synthesis and at 75°C for 10 minutes to heat-inactivate the M-MLV Reverse Transcriptase. A newly synthesised cDNA was stored at -20°C. Prior to the use of a newly synthesised cDNA for allelic RT-PCR (**Section 2.3.7.1.5**) or quantitative real-time reverse transcription PCR (qRT-PCR) (**Section 2.5.3**), it was tested with PCR primers designed for *ACTB*. See the primers in **Appendix 26**. For the PCR conditions, check **Section 2.3.7**.

**Table 2.1. cDNA synthesis Master Mix for one reaction**

<b>Reagents:</b>	<b>Supplier:</b>	<b>Cat. No.</b>	<b>1 x (μL)</b>
M-MLV Reverse Transcriptase	Promega UK Ltd.	M1705	1
M-MLV RT 5 x Buffer	Promega UK Ltd.	M1705	4
dNTP mix (10 mM)	Promega UK Ltd.	U1511	1
Random Primers (random hexadeoxynucleotides)	Promega UK Ltd.	C1181	1
RNasin® Plus Ribonuclease Inhibitor	Promega UK Ltd.	N2611	0.25
UltraPure™ DEPC-Treated Water	Thermo Fisher Scientific Inc.	750023	2
<b>Total:</b>			<b>9.25</b>

### 2.3.6. Sodium Bisulphite DNA conversion

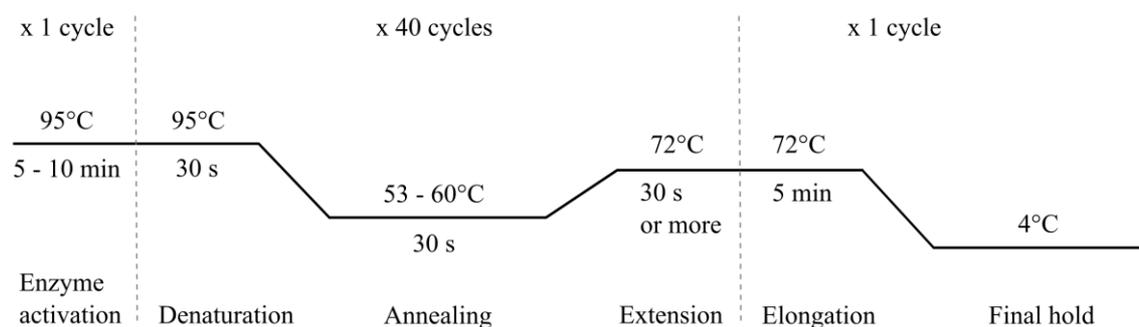
To explore if promoters or placental DMR regions of candidate genes were methylated, placental and blood gDNA samples were treated with sodium bisulphite, which is considered to be a gold-standard method for DNA methylation analyses. During sodium bisulphite conversion, unmethylated cytosines that are present within CG dinucleotides are initially deaminated and then desulphonated, converting cytosines into uracils (U), which after subsequent PCR amplification, are converted into thymines (T) (458). In contrast, methylated cytosines are not converted and remain as cytosines (C). Therefore, this method provides a base-pair resolution view to investigate DNA methylation status at each CpG site.

Sodium bisulphite conversion was carried out with the EZ-96 DNA Methylation-Direct Kit (D5023; Zymo Research Corporation) following the manufacturer's protocol. In short, 1 μg of placental or blood gDNA was dissolved in 20 μL of UPW (750023; Thermo Fisher Scientific Inc.) and mixed with 130 μL of the CT Conversion Reagent, which was prepared in advance and consisted of the CT Conversion Reagent (i.e. sodium metabisulphite), M-Solubilization Buffer, M-Dilution Buffer and M-Reaction Buffer. A Conversion Plate with the samples was placed into the Veriti™ Thermal Cycler (4375305; Thermo Fisher Scientific Inc.) and initially incubated at 98°C for 8 minutes, followed by 64°C for 210 minutes, then finished at 4°C for up to 20 hours. Converted samples were transferred to a fresh Zymo-Spin I-96 Binding Plate, and each sample was mixed with 600 μL of the M-Binding Buffer by pipetting. The Zymo-Spin I-96 Binding Plate placed on a Collection Plate were then centrifuged at 3,100 × g for 5 minutes. This step was followed by three additional washing steps. Firstly, 400 μL of M-Washing Buffer was pipetted into each well, followed by centrifugation. Then, 200 μL of M-Desulphonation Buffer

(desulphonates bisulphite-treated DNA) was loaded into each well, and the plate was incubated at room temperature for 20 minutes, followed by centrifugation. Each sample was then washed by adding 400  $\mu\text{L}$  of the M-Wash Buffer and centrifuged. Finally, the samples were eluted twice. Initially, 15  $\mu\text{L}$  of the M-Elution Buffer was added across the plate and incubated at room temperature for 5 minutes. The Zymo-Spin I-96 Binding Plate placed on the Elution Plate were centrifuged at  $3,100 \times g$  for 3 minutes. This step was repeated, but during the second time, only 7  $\mu\text{L}$  of the M-Elution Buffer was added. The Elution Plate with bisulphite-treated DNA was stored at  $-20^\circ\text{C}$  until later use. The converted bisulphite DNA was either used for bisulphite PCR and cloning (**Sections 2.3.7.1.4 & 2.3.9**) or pyrosequencing (**Section 2.5.4**).

### 2.3.7. Polymerase chain reaction (PCR)

Standard PCR conditions are shown in **Figure 2.2**.



**Figure 2.2. PCR standard amplification conditions with the BIOTAQ or IMMOLASE DNA polymerase.**

The initial stage for the BIOTAQ DNA polymerase is 5 minutes, while 10 minutes are required for the IMMOLASE DNA polymerase. Step 1: enzyme activation; Step 2: PCR amplification, including DNA strand denaturation, primer annealing and extension stages; Step 3: elongation and hold.

Depending on which genomic region was interrogated, different DNA molecules were used as templates for PCR amplification. Usually, 1  $\mu\text{L}$  of placental or parental gDNA ( $\sim 100 \text{ ng}/\mu\text{L}$ ) was used for standard genotyping PCR (**Table 2.2**). For methylation-sensitive genotyping, either 1 or 2  $\mu\text{L}$  of digested placental DNA was applied. For bisulphite PCR, either 2 or 3  $\mu\text{L}$  of bisulphite-converted DNA was used, as DNA during the sodium

bisulphite treatment becomes highly fragmented. Finally, 1 or 2  $\mu\text{L}$  of placental cDNA ( $\sim 50 \text{ ng}/\mu\text{L}$ ) was used for allelic RT-PCR. The Master Mix prepared for bisulphite PCR and allelic RT-PCR usually included the IMMOLASE DNA polymerase as it has higher specificity than the BIOTAQ DNA polymerase (**Table 2.2**). Also, PCR denaturation, annealing and extension steps (Step 2) were usually performed over 45 cycles (to exhaust primers) for bisulphite PCR and over 40 or 45 cycles for allelic RT-PCR (**Figure 2.2**).

**Table 2.2. The reagents used for a standard PCR Master Mix (for one sample)**

<b>Reagents:</b>	<b>Supplier:</b>	<b>Cat. No.</b>	<b>1 x (<math>\mu\text{L}</math>)</b>
10 x NH <sub>4</sub> Reaction Buffer or 10 x ImmoBuffer:	Meridian Bioscience Inc.	BIO-21040, BIO-21047	1.25
MgCl <sub>2</sub> Solution (50 mM)	Meridian Bioscience Inc.	BIO-21040, BIO-21047	0.375
dNTP mix (2 mM)	Promega UK Ltd.	U1511	0.25
Forward primer (0.1 $\mu\text{g}/\mu\text{L}$ )	Merck Life Science UK Ltd.	N/A	0.25
Reverse primer (0.1 $\mu\text{g}/\mu\text{L}$ )	Merck Life Science UK Ltd.	N/A	0.25
Betaine (5 M)	Merck Life Science UK Ltd.	B2629-100G	3.75
BIOTAQ or IMMOLASE DNA Polymerase (5 u/ $\mu\text{L}$ ):	Meridian Bioscience Inc.	BIO-21040, BIO-21047	0.1
UltraPure™ DEPC-Treated Water	Thermo Fisher Scientific Inc.	750023	5.775*
Total:			12

\*The amount of ultrapure water was adjusted based on the amount of template added to the PCR Master Mix.

PCR was performed at various temperatures with all ordered oligonucleotide primer pairs to determine the optimal primer annealing temperature and the minimum DNA template required for a successful PCR experiment. All PCR experiments included a negative control (i.e. a PCR reaction without a template to detect any contamination) and a positive control (i.e. mixed tissue DNA or cDNA, or bisulphite-converted DNA) depending on the PCR experiment performed. All PCR primers can be found in **Appendix 26**, **Appendix 27** and **Appendix 28**.

### 2.3.7.1. Different types of PCR for investigating candidate genes

#### 2.3.7.1.1. Nested PCR

A nested PCR method requires the use of two sets of oligonucleotide primers - outer primers and internal primers that align to the same genomic region. This adapted PCR is a good choice for genes with low expression in different human tissues, isoform-specific expression, and complex genomic regions that are difficult to amplify with standard PCR, as non-specific PCR products can be generated. Such genomic positions can include genes with highly repetitive LTR motifs that are found in multiple locations within a mammalian genome. However, it should be noted that nested PCR often requires many cycles of amplification, which can lead to preferential amplification of one allele and, in some cases, even allelic dropout (459–461). To minimise the risk of PCR-induced bias in the base composition of the sequences, the number of amplification cycles was reduced wherever possible. For variant calling analysis, several primer sets (outer and inner primers) were designed for each region, and only the most efficient primer pairs were used for nested PCR. All informative samples were tested when feasible, and at least two independent PCR runs were performed and used for Sanger sequencing to ensure consistency between results.

The first round of PCRs was performed with the outer primer pair in a total volume of 13  $\mu$ L (11  $\mu$ L of the PCR Master Mix with 2  $\mu$ L of DNA, cDNA or bisulphite-converted DNA) (**Table 2.2**). This PCR was performed for 45 cycles, but after 15 cycles, 1  $\mu$ L of the PCR aliquots was transferred to PCR tubes containing the PCR Master Mix (12  $\mu$ L) with the internal primers, resulting in a total volume of 13  $\mu$ L. This second or nested PCR was run for 30-35 cycles. The first PCR included a non-template negative control and a positive control, depending on the experiment, which was either mixed tissue bisulphite-converted DNA, cDNA or DNA. The nested PCR included a non-template control, a non-template control from the first PCR, and a positive control, which, depending on the experiment, was either mixed tissue bisulphite-converted DNA, cDNA or DNA. All PCR reactions were performed with the IMMOLASE DNA polymerase due to higher specificity. All primers for the first and nested PCR can be seen in **Appendix 26**, **Appendix 27** and **Appendix 28**.

Nested bisulphite PCR products were used for sub-cloning, cleaning and Sanger sequencing (**Sections 2.3.8**, **2.3.9** & **2.3.10**), while nested PCR products generated with cDNA and DNA were cleaned and used for Sanger sequencing (**Sections 2.3.8** & **2.3.10**).

#### 2.3.7.1.2. Genotyping PCR

A standard PCR Master Mix (**Table 2.2**) was used with a placental or parental gDNA. Primers can align to coding and non-coding regions of DNA (introns, 5' and 3' UTR regions) as they flank a polymorphism (**Appendix 26, Appendix 27 and Appendix 28**). The correct size PCR products were cleaned and used for Sanger sequencing. This PCR is used to identify heterozygous placental samples.

#### 2.3.7.1.3. Methylation-sensitive genotyping

A standard PCR Master Mix (**Table 2.2**) was used with a placental gDNA digested with methylation-sensitive restriction enzymes (HpaII or BstUI). Primers were designed to flank a polymorphism specifically within the regions of interest (a DMR or gene promoters) that included multiple restriction sites (**Appendix 26, Appendix 27 & Appendix 28**). Following digestion, only methylated DNA remained intact, acting as a template for PCR (**Figure 2.3A**). Allelic methylation was confirmed when a heterozygous gDNA sample was reduced to homozygosity following digestion with HpaII or BstUI, with the remaining allele representing the methylated chromosome.

Our group has successfully applied this method previously to identify imprinted DMRs (20) because it effectively distinguishes between unmethylated regions (*KLF10* promoter (462)), those showing monoallelic methylation and imprinting (*SNURF: TSS* DMR (420)), biallelically methylated regions (*RASSF1* transcript A promoter (463)), and regions with mosaic/random monoallelic methylation (*DLGAP2* promoter (464)). This technique is summarised in **Figure 2.3**.

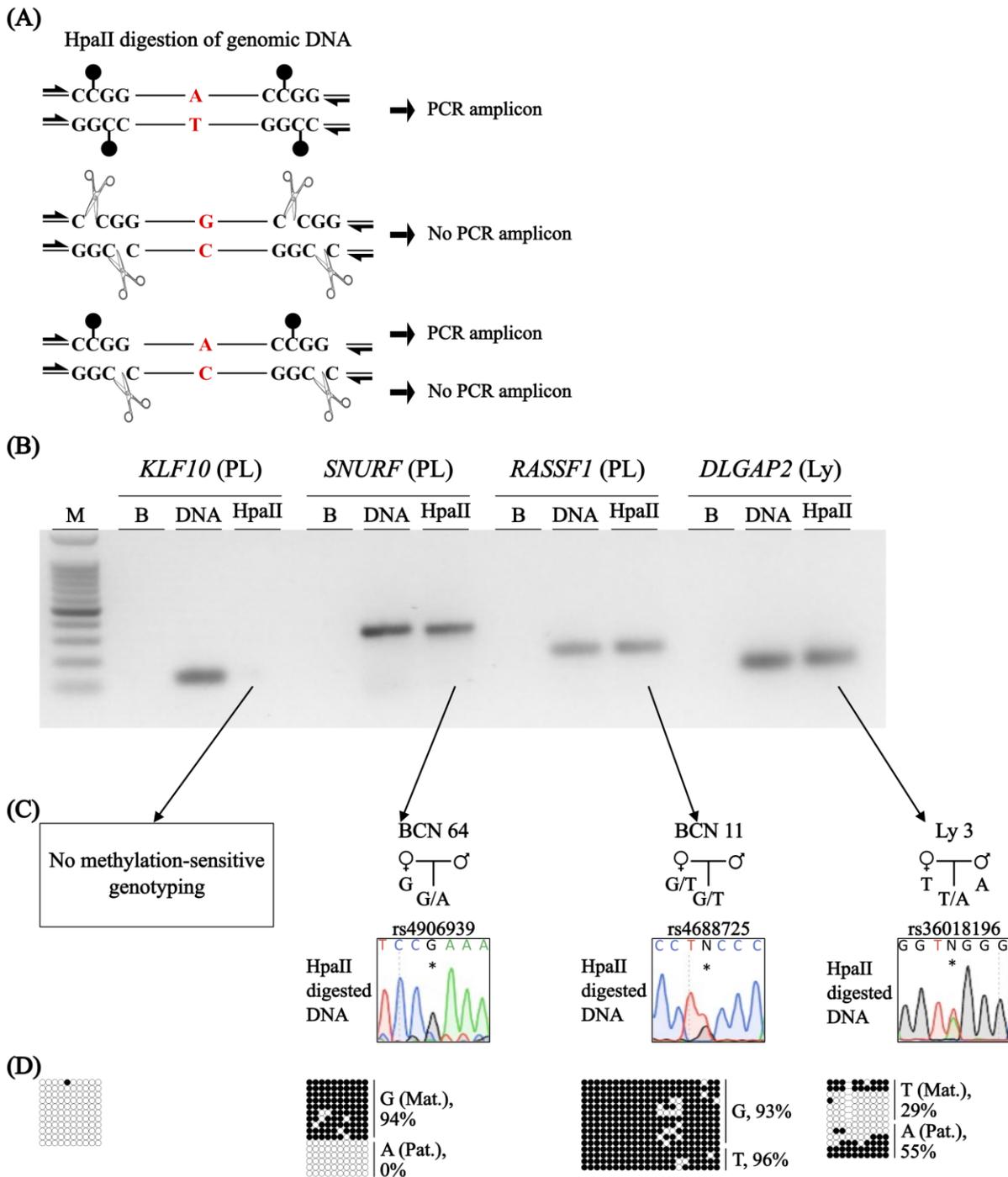
Selected control regions:

- *KLF10* belongs to the family of zinc-finger TFs (462). It is expressed in placental trophoblasts and was more recently found to be important for the transition from CTBs to STBs. According to our placental WGBS dataset, the promoter of this gene is completely unmethylated (**Figure 2.3B, C, D**).

- The *SNURF: TSS* DMR exhibits maternal allele-specific methylation, resulting in paternal-specific expression (420). Gain of methylation on the paternal chromosome at this DMR causes PWS, whereas loss of methylation on the maternal chromosome results in AS (**Figure 2.3B, C, D**).

- Our group previously investigated *RASSF1*, which contains two CpG island promoters (463). The promoter of *RASSF1* transcript A was found to be highly methylated (~80%), while the downstream promoter of *RASSF1* transcripts B and C was hypomethylated (~10%) (**Figure 2.3B, C, D**). The *RASSF1A* promoter showed a significant increase in DNA methylation in placentae from IUGR cases and a significant downregulation in placentae affected by PE compared to controls. Methylation levels were found to be negatively associated with the expression of *RASSF1* transcript A. Interestingly, no significant changes in transcript A expression were observed when comparing placentae from healthy individuals to those from IUGR or PE cases.

- Our group has also shown that *DLGAP2* contains an oocyte-derived DMR that is maintained during pre-implantation stages but transitions to random allele-specific methylation in most foetal tissues after 16 weeks of gestation (**Figure 2.3B, C, D**) (464). An exception to this pattern is observed in the placenta and kidney, where maternal allele-specific methylation is retained. Interestingly, *DLGAP2* is not expressed either in the placenta or kidney, while it demonstrates biallelic expression in other tissues. This gene has clinical significance, as it has been linked to several neurological disorders, including autism spectrum disorder, schizophrenia, obsessive-compulsive disorder, and Alzheimer's disease.



**Figure 2.3. Methylation-sensitive genotyping assay.**

**(A)** Schematic overview of methylation-sensitive genotyping with HpaII. Black circles represent methylated restriction sites (5mC), while unmethylated sites can be digested with HpaII (indicated by scissors). Polymorphic sites are shown in red. Black half-arrows indicate PCR primers. **(B)** Agarose gel showing PCR results for the unmethylated *KLF10* promoter, the imprinted gDMR of *SNURF* and the methylated promoter of *RASSF1* (CpG island with 83 CpGs) in placental samples, as well as random monoallelic methylation at the *DLGAP2* promoter in blood. **(C)** Sequencing chromatograms of heterozygous samples. **(D)** Each region containing the same SNP was confirmed by bisulphite PCR followed by sub-cloning and Sanger sequencing. Methylated cytosines are indicated by (●), and unmethylated cytosines by (○), with each row representing an individual cloned sequence. The parent-of-origin was inferred from SNP genotyping in heterozygous samples.

#### 2.3.7.1.4. Bisulphite PCR or Nested Bisulphite PCR

A PCR Master Mix containing the IMMOLASE DNA polymerase was combined with a placental bisulphite-converted DNA (**Table 2.2**). Primers are targeted to gene promoters or placental DMR regions that are rich in CpG sites and frequently contain polymorphisms (**Appendix 26, Appendix 27 & Appendix 28**). These primers can align to coding and non-coding regions of DNA. The amplified products were used to quantify the level of methylation at the targeted region by pyrosequencing (**Section 2.5.4**) or determine the methylation status (fully methylated region, semi-methylated region or unmethylated region) by sub-cloning with Single-use JM109 Competent Cells (L2005; Promega UK Ltd.) followed by blue and white screening and Sanger sequencing (**Sections 2.3.9 & 2.3.10**).

#### 2.3.7.1.5. Allelic RT-PCR or Nested allelic RT-PCR

A PCR Master Mix containing the IMMOLASE DNA polymerase was applied with placental cDNA or human pre-implantation embryo cDNA (**Table 2.2; Sections 2.2.2 & 2.6**). If possible, primers were designed in different exons, skipping introns and flanking exonic polymorphisms (**Appendix 26, Appendix 27 & Appendix 28**). The generated amplicons were cleaned and used for Sanger sequencing (**Sections 2.3.8 & 2.3.10**). This PCR was used to determine the type of expression: if both alleles of a gene were expressed - biallelic expression, if a single allele of a gene was expressed - monoallelic expression, or if both alleles were expressed, but one allele showed much higher expression signal in a sequencing chromatogram - preferential monoallelic expression.

### 2.3.8. PCR product purification

#### 2.3.8.1. Agarose gel electrophoresis

To verify that PCR worked and that amplified products were specific, PCR amplicons were visualised by agarose gel electrophoresis. Depending on PCR amplicon sizes, either 1% or 2% agarose gels were applied for electrophoresis. For a 1% agarose gel, 1.2 g of agarose (BP160500; Fisher BioReagents) was dissolved in 120 ml of the 0.5 x TAE buffer with 2.5 µL ethidium bromide solution (E1510-10ML; Merck Life Science UK Ltd.). 1 L of the 0.5 x TAE buffer was made by diluting 50 mL of the 10 x TAE stock with 950 mL double-distilled water (ddH<sub>2</sub>O). For 1 L of 10 x TAE, 48.4 g of Tris base was dissolved in 11.42 mL glacial acetic acid and 40 mL EDTA (0.5 M, pH = 8.0). The 0.5 x TAE buffer was used for

gel electrophoresis. Overall, 3  $\mu\text{L}$  of the 100 bp DNA ladder (G2101; Promega) and 3  $\mu\text{L}$  of PCR mixed with 0.5  $\mu\text{L}$  6 x Orange G loading dye (J60562.AC; Thermo Fisher Scientific Inc.) were loaded into an agarose gel. Electrophoresis was carried out in the HU15 Standard Horizontal gel tank unit (Scie-Plas Ltd.) at 120 V by using the PowerPac™ Basic Power Supply (BIO-RAD) for 20-40 minutes. Agarose gels were photographed by the UVP 310 GelDoc-It2 system.

### 2.3.8.2. Gel extraction

A correct-size PCR product was quickly excised from an agarose gel under UV light (the Enprotech TFX-20M UV Transilluminator) to minimise damaging UV illumination. The PCR product was cleaned by the GeneJET Gel Extraction Kit (K0691; Thermo Fisher Scientific Inc.) following the manufacturer's manual. In short, the excised gel slice was dissolved in the Binding Buffer (1:1 ratio of 1% agarose gel weight (g) and the buffer volume (mL)) over 10 minutes at 60°C. Then, 800  $\mu\text{L}$  of the solubilised gel solution was transferred to the GeneJET purification column and microcentrifuged at 12,000 x g for 1 minute. After the spin, the flow-through was discarded, and an additional 100  $\mu\text{L}$  of the Binding Buffer was added into the column, followed by microcentrifugation (12,000 x g, 1 minute). The column was washed by adding 700  $\mu\text{L}$  of the Wash Buffer, followed by the same microcentrifugation. To remove any residual ethanol present within the Wash Buffer from the sample, the GeneJET purification column was microcentrifuged at 12,000 x g for 1 minute. Finally, to increase DNA yield, the clean product was eluted twice in a fresh tube by adding 10  $\mu\text{L}$  of the Elution Buffer to the GeneJET purification column, followed by a 1 minute incubation at room temperature and microcentrifugation at 12,000 x g for 1 minute. The concentration of cleaned PCR product was checked with the NanoDrop spectrophotometer (ND-1000; Thermo Fisher Scientific Inc.). The cleaned product was used either for cloning or Sanger sequencing (**Sections 2.3.9 & 2.3.10**), otherwise, it was stored at -20°C.

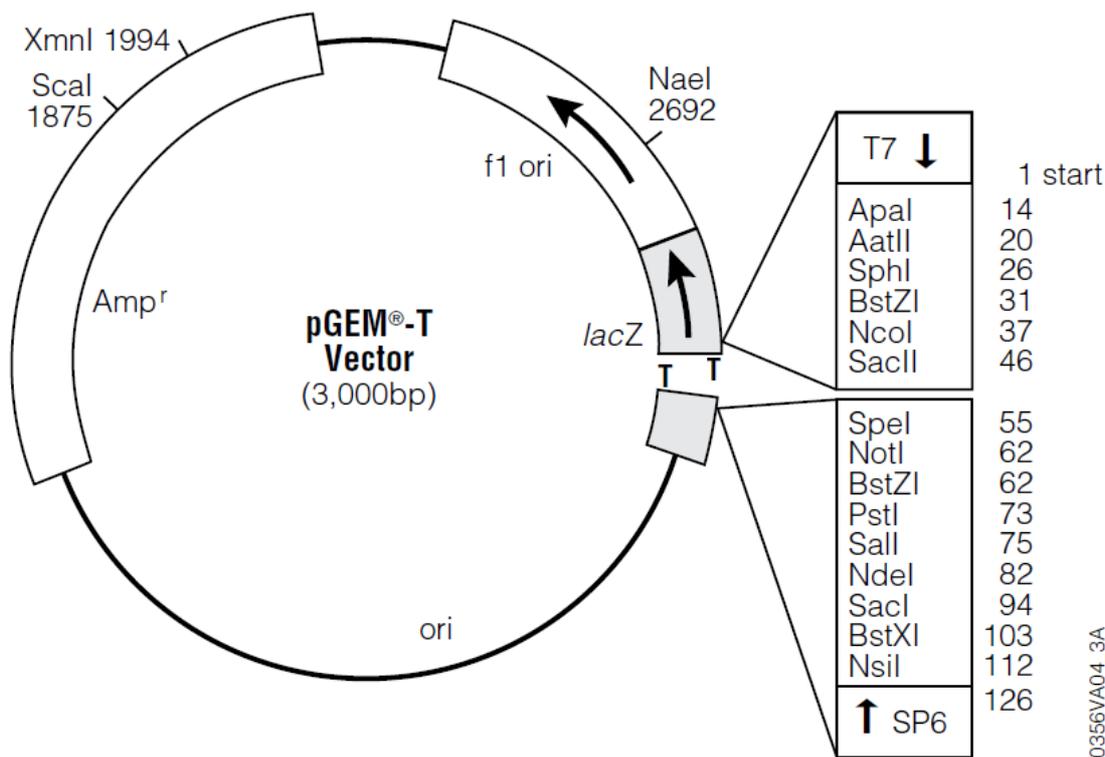
### 2.3.8.3. PCR clean-up by ethanol precipitation

The standard ethanol precipitation method was used to purify PCR products. PCR products were cleaned to remove salts with buffers, unused dNTPs and primers. Briefly, PCR products were loaded into wells of a 96-well PCR plate. Each sample was mixed with 1/10 volume of 3 M sodium acetate (pH = 4.6) by pipetting. Subsequently, 2.5 volumes of 100% ethanol was added into each well and mixed by pipetting. The plate was stored at -20°C for at least 1 hour to improve precipitation. The plate was then centrifuged at 3,700 x rpm for 40 minutes at 4°C to pellet the nucleic acid. Immediately, the plate was quickly inverted or "flicked" over a sink to discard the supernatant. Then, 20  $\mu\text{L}$  of 70% ethanol

was added across the plate, followed by centrifugation at 1,000 x rpm for 10 minutes and at 4°C. To remove residual ethanol, the plate was “flicked” again to discard the supernatant and blotted on a piece of paper towel (upside down), followed by centrifuging at 250 x rpm for 1 minute (4°C). Finally, the cleaned plate was air-dried for 20 minutes. The cleaned PCR pellets were resuspended in 8 µL of UPW (750023; Thermo Fisher Scientific Inc.). To dissolve the cleaned nucleic acid more efficiently, the plate was placed into the Veriti™ Thermal Cycler (4375305; Thermo Fisher Scientific Inc.) at 95°C for 10 seconds, followed by immediate cooling to -20°C in a freezer. The concentration and purity of cleaned PCR samples were checked by the NanoDrop spectrophotometer (ND-1000; Thermo Fisher Scientific Inc.) before Sanger sequencing.

### 2.3.9. Sub-cloning

For sub-cloning, the pGEM®-T Easy Vector System (Promega UK Ltd.) was chosen, as the included vector is small (3,015 bp) and already pre-linearised (**Figure 2.4**). It also contains 3'-T overhangs at the insertion sites that are compatible with the BIOTAQ or IMMOLASE DNA polymerase (Meridian Bioscience Inc.) generated PCR products, which contain 3'-A overhangs that increase the efficiency of cloning. For ligation, a 3:1 ratio of the plasmid and the PCR product was used, as advised by the manufacturer. Thus, 3 µL of PCR was mixed with 1 µL of the pGEM®-T Easy Vector (A137A; Promega UK Ltd.), 5 µL of the Rapid Ligation Buffer (C671A; Promega UK Ltd.) and 1 µL of the T4 DNA Ligase (3 U/µL; M180A, Promega UK Ltd.). The mix was shaken a few times and incubated at 4°C overnight or over the weekend.



**Figure 2.4. pGEM®-T easy vector with the insertion site within *lacZ*.**

Adapted from the Promega manual (2021) (465).

For the bacterial transformation step, 2.5  $\mu\text{L}$  of the ligation product was mixed with 25  $\mu\text{L}$  of Single-use JM109 Competent Cells (L2005; Promega UK Ltd.). The heat shock method was applied to transform the competent cells, which involved incubating the cells on ice for 30 minutes, followed by immediate incubation in a 42°C water bath for 45 seconds, and then incubating the cells on ice for 2 minutes. To increase the efficiency of transformation, bacteria were grown in LB (**Table 2.3**) in a shaking incubator (37°C) for at least 1 hour. Finally, the competent cells were spread on LB-agar plates with ampicillin (**Table 2.4**) and grown at 37°C overnight. The insertion site in the pGEM®-T Easy Vector is present within *lacZ*; therefore, the successful integration of the ligation product interrupts *lacZ* expression, which produces catalytically inactive  $\beta$ -galactosidase that is incapable of catalysing X-gal (then catalysed produces dark blue precipitate). Therefore, transformed bacteria appear white and can be easily selected for subsequent genotyping. White colonies were hand-picked and grown in 50  $\mu\text{L}$  of pre-warmed LB without ampicillin (**Table 2.3**) and further grown for a minimum of 1 hour in an incubator at 37°C. Finally, positive white colonies were subject to PCR genotyping using the PCR primers (**Appendix 26**) designed to flank multiple cloning and insertion sites in the pGEM®-T Easy Vector. 1  $\mu\text{L}$  of LB with bacteria was used as a template for PCR. Appropriate size PCR was either precipitated or gel extracted and prepared for Sanger

sequencing (**Sections 2.3.8 & 2.3.10**). Colonies were either sequenced with M13, SP6 or T7 primers (**Appendix 26**) close to the insertion site inside the pGEM®-T Easy Vector.

**Table 2.3. LB broth**

<b>Reagents:</b>	<b>Supplier:</b>	<b>1 x L</b>
Tryptone	Merck Life Science UK Ltd.	10 g
Sodium chloride	Merck Life Science UK Ltd.	10 g
Yeast extract	Merck Life Science UK Ltd.	5 g
ddH <sub>2</sub> O		1 L

**Table 2.4. LB agar Petri dishes**

<b>Reagents:</b>	<b>Supplier:</b>	<b>1 x L</b>
Tryptone	Merck Life Science UK Ltd.	10 g
Sodium chloride	Merck Life Science UK Ltd.	10 g
Yeast extract	Merck Life Science UK Ltd.	5 g
ddH <sub>2</sub> O		1 L
Agar	Merck Life Science UK Ltd.	15 g
Ampicillin (50 µg/mL)	Merck Life Science UK Ltd.	1 mL
X-gal (20 µg/mL)	ForMedium	2 mL
IPTG (0.1 mM)	Thermo Fisher Scientific Inc.	1 mL

## 2.3.10. Sanger sequencing

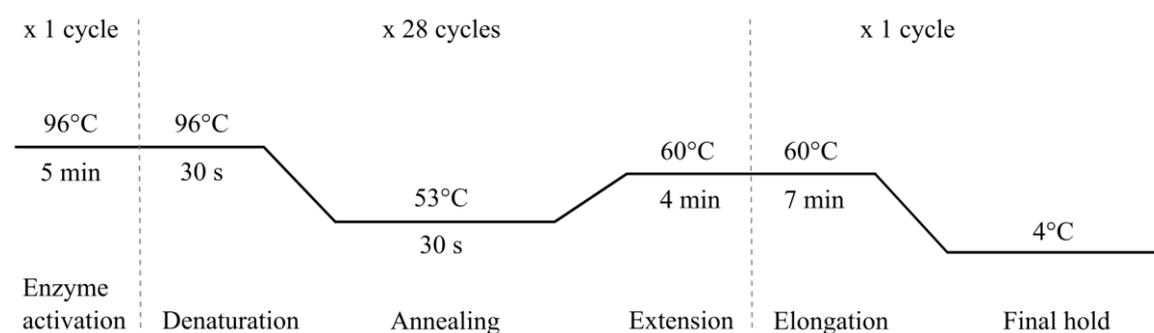
### 2.3.10.1. Sample preparation for sequencing using the BigDye Terminator (BDT)

The cleaned PCR samples were sequenced with the BigDye Terminator v3.1 Cycle Sequencing Kit (BDT; 4337456; Thermo Fisher Scientific Inc.) in 96-well PCR plates. The right amount of the PCR product, determined according to its size (**Table 2.5**), was dissolved in UPW (750023; Thermo Fisher Scientific Inc.) to make a total volume of 5.7 µL. The diluted sequencing template was further mixed with 0.3 µL of a sequencing primer (0.1 µg/µL) and 4 µL aliquot of the BDT mix. For 500 µL of BDT mix, 100 µL of the BDT Ready Reaction mix and 100 µL of the 5 x Sequencing buffer were diluted in 300 µL UPW. The plate with samples was briefly centrifuged and placed into the Veriti™ Thermal Cycler (4375305; Thermo Fisher Scientific Inc.) with the cycling conditions indicated in

**Figure 2.5** and stored at -20°C until the second clean-up.

**Table 2.5. The amount of a PCR product required for sequencing with BDT**

Size (bp):	Amount of template required (ng):
200	6
400	12
600	18
800	25



**Figure 2.5. Cycling conditions for sequencing with BDT.**

### 2.3.10.2. Post-sequencing clean-up and capillary electrophoresis

The sequenced samples might have had unincorporated dye terminators, dNTPs, and salts that could interfere with base calling; therefore, samples had to be purified for the second time. Post-sequencing purification and sequencing files were generated by RevGenUK – Molecular Genetics platform at the John Innes Centre, Norwich, UK. To purify the sequenced samples, Optima DTR™ 96-Well Plates were utilised (Edge BioSystems), while clean samples were run on the 3730xl DNA Analyzer (A41046; Thermo Fisher Scientific Inc.)

### 2.3.10.3. Sanger sequencing data analysis

Electropherograms or sequencing chromatograms were examined using CodonCode Aligner v11.0.2 DEMO (CodonCode Corporation, USA) or SnapGene v7.2.1 (GSL Biotech LLC, USA). The positions of germline variants within the chromatograms were visually inspected by both myself and my primary supervisor, Prof. David Monk. Chromatograms

of poor quality were re-sequenced.

Germline-variant calling with Sanger sequencing: sequenced PCR amplicons contained a SNP and/or short indel that were/was investigated. The sequencing results for **SNPs** were interpreted based on the following criteria:

Heterozygous individuals, biallelic expression or biallelic methylation: two peaks representing two different nucleotides at a single position, both peaks had a similar height (50% : 50% and 25% : 75% ratios).

Homozygous individuals, monoallelic expression or monoallelic methylation: a single peak representing a single nucleotide (92%: 8% signal to background noise ratio).

Preferential expression or methylation: two peaks representing 2 different nucleotides at the same position, but both peaks showing different heights (90% : 10%, 80% : 20% and 74% : 26% ratios).

The sequencing results for **indels** were interpreted based on the following criteria:

Heterozygous individuals, biallelic expression or biallelic methylation: a series of distinct peaks followed by a stretch of overlapping and sometimes distorted peaks indicating a frameshift; overlapping peaks representing two distinct alleles of a variant.

Homozygous, monoallelic expression or monoallelic methylation: distinct peaks showing a single allele of a variant.

All identified informative samples were used for PCR amplification and Sanger sequencing wherever possible. Genomic regions that required a high number of PCR amplification cycles were amplified and sequenced at least twice. Similarly, samples that yielded inconsistent results for a given genomic region were re-amplified and sequenced at least twice to ensure reproducibility. Each sample was determined as homozygous, heterozygous, or exhibiting one preferential allele for the investigated polymorphism based on visual inspection of all sequencing chromatograms. In most cases, a consensus was reached based on the majority of sequencing results. In rare instances, if a single chromatogram clearly showed the presence of both alleles at a given position, the sample was classified as heterozygous - indicating biallelic methylation or expression. For example, one sample showed two chromatograms with preferential expression of one allele and one chromatogram with equal expression of both alleles; it was concluded that the sample exhibited biallelic expression. When a sample was informative for multiple

polymorphisms, its methylation or expression status (monoallelic or biallelic) was determined based on the majority of results. In cases where one polymorphic site showed preferential expression of a single allele while another site exhibited equal expression of both alleles; it was determined that the sample displayed biallelic expression. For some informative samples, methylation at putative DMRs or gene promoters was investigated using bisulphite PCR, sub-cloning and subsequent Sanger sequencing. At the locus level, monoallelic or biallelic methylation and expression were assigned based on the overall pattern observed across the majority of informative samples.

Cloned bisulphite PCR sequences: a CpG site was considered to be methylated if a clear peak indicating C was detected. A CpG site was determined to be unmethylated if a peak for T was observed. If two different peaks for C and T were observed at the same position, the methylation status of the site could not be determined and was usually indicated by a dash (“-”).

## 2.4. Cell isolation and culture techniques

### 2.4.1. Cell isolation from the human placenta by Magnetic-Activated Cell Sorting (MACS)

#### 2.4.1.1. Protocol overview

This cell enrichment protocol relies on a few key stages. Firstly, the dissected placenta is enzymatically digested to obtain a cell suspension. Trophoblasts and other placental cell types are separated from red blood cells by percoll gradient, and finally, the MACS columns are used to positively enrich for placental trophoblast (EGFR positive cells) and stromal cell populations (anti-fibroblasts positive cells). Overall, two enriched cell populations can be used for RNA and DNA isolation as described earlier (**Sections 2.3.2 & 2.3.3**).

#### 2.4.1.2. Percoll gradient

Percoll is a medium containing colloidal silica particles coated with polyvinylpyrrolidone (PVP) and is used for low-viscosity density gradients that are suitable for isolating cells, organelles, or viruses. Accordingly, Kliman *et al.* (1986) (466) first applied this method for

the human placenta to separate placental cell populations such as fibroblasts, CTBs or EVT<sub>s</sub> from cell debris, red blood cells or polymorphonuclear cells. For a percoll gradient, different concentrations of percoll are slowly layered on top of each other in a falcon tube, forming 14 distinct layers that, after centrifugation, contain different cell types (467). Thus, this gradient can be used for the enrichment of positive cells.

#### 2.4.1.3. Magnetic-activated cell sorting (MACS)

MACS is a simple, versatile and fast technique invented by Miltenyi Biotec that is used to enrich different cell populations from a mixture of cells, such as tissues. During this technique, a cell population expressing a unique surface antigen is bound by an antibody that is conjugated to MACS® MicroBeads (**Table 2.6**). The mixture of labelled cells present within a suspension is transferred to the MS column that is surrounded by a strong magnet (OctoMACS™ Separator; **Table 2.6**). Thus, the labelled cells are trapped in the column, while non-labelled cells can freely flow through the column and be collected in a fresh tube (negative cell selection). In addition, the MS column contains a matrix composed of coated ferromagnetic spheres that can increase the magnetic field by 10,000-fold and even further enhance the magnetic field, allowing for the use of a lower amount of antibody to label the cells. In the case of positive selection, the column with captured cells is removed from the magnetic field, allowing the labelled cells to be washed out. The pellets of these cells can be used for conventional RNA and DNA isolation methods, as explained in earlier sections of this thesis (**Sections 2.3.2 & 2.3.3**).

**Table 2.6. Reagents and equipment required for trophoblast and stromal cell enrichment from placental cell suspensions by MACS**

<b>Reagents and equipment:</b>	<b>Quantity:</b>	<b>Supplier:</b>	<b>Cat. No.</b>
MS columns	2	Miltenyi Biotec	130-042-201
OctoMACS™ Separator	1	Miltenyi Biotec	130-042-109
Purified anti-human EGFR Antibody (primary antibody)	20 µl	BioLegend	352902
Anti-Mouse IgG1 MicroBeads (secondary antibody)	80 µl	Miltenyi Biotec	130-047-101
Anti-Fibroblast MicroBeads, human	80 µl	Miltenyi Biotec	130-050-601
Fisherbrand™ Cell Strainers (40 µm)	2	Fisher Scientific	11587522
Fisherbrand™ Cell Strainers (70 µm)	8	Fisher Scientific	11597522
DNA LoBind® Tubes	2	Eppendorf SE	0030108078

#### 2.4.1.4. Isolation of placental cell types

##### 2.4.1.4.1. Placenta dissection

2 cm<sup>2</sup> pieces of a placenta were excised near the insertion site of the umbilical cord at 4°C (**Figure 2.6**). Thin layers from the uterine and foetal sides were removed and discarded, while the placental pieces were further rinsed with PBS, then chopped (0.2 cm<sup>3</sup>) and scraped to remove vessels. Approximately 10 mL of tissue was added into 8 falcon tubes, followed by two consecutive enzymatic digestions and a few washing steps. The first digestion was performed using the Trypsin solution (25 mL per tube; **Table 2.7**) by incubating the samples at 37°C for 30 minutes in a shaking incubator at 100 x rpm. To prepare placental cell suspensions, 2.5 mL of FBS (**Table 2.8**) was added to tubes with digested tissues. The mixes were then transferred to cell strainers (70 µm; **Table 2.6**), and cell suspensions were collected in fresh tubes (kept at 4°C). Undigested pieces of the tissue were further digested with the Collagenase solution (25 mL per tube; **Table 2.7**) at 37°C for 30 minutes in a shaking incubator (100 x rpm). The cell suspensions from the Collagenase digestion were collected in the same way as for the Trypsin digestion (additional 8 tubes). Cell suspensions were centrifuged at 400 x g at 4°C for 10 minutes, the supernatant was discarded, and cell pellets were washed with 10 mL of the Wash buffer (**Table 2.8**). 8 tubes per digestion were combined into 2 falcon tubes. Samples were further centrifuged at 400 x g for 10 minutes (4°C), the supernatant was discarded, and cell pellets were diluted with 5 mL of the Wash buffer (**Table 2.8**). All samples were combined in a falcon tube, making a total of 40 mL.



**Figure 2.6. Dissection of the human placenta from several places.**

**Table 2.7. Enzyme solutions**

<b>Trypsin solution</b>	<b>Stock conc.<sup>1</sup></b>	<b>Amt. per ~25 mL<sup>2</sup></b>	<b>Final conc.<sup>3</sup></b>	<b>Volume for 8 samples</b>	<b>Supplier</b>	<b>Cat. No.<sup>4</sup></b>	<b>Product name</b>
Trypsin solution	0.25%	12 mL	0.12%	96 mL	PAN Biotech UK Ltd.	P10-029500	Trypsin 0.25 %/ 1 mM EDTA
DNase I	5 mg/mL	250 µL	100 µg/mL	2 µL	Merck Life Science UK Ltd.	DN25-100MG	Deoxyribonuclease I from bovine pancreas
MgCl <sub>2</sub>	0.5 M	250 µL	5 mM	2 µL	Fisher Scientific	10418464	MgCl <sub>2</sub> (1 M)
DMEM (with Ca)		6 mL		48 mL	PAN Biotech UK Ltd.	P04-04510	DMEM
1x PBS		6 mL		48 mL			
<b>Collagenase solution</b>	<b>Stock conc.<sup>1</sup></b>	<b>Amt. per ~25 mL<sup>2</sup></b>	<b>Final conc.<sup>3</sup></b>	<b>Volume for 8 samples</b>	<b>Supplier</b>	<b>Cat. No.<sup>4</sup></b>	<b>Product name</b>
Collagenase IV	2% (20 mg/mL)	625 µL	0.5 mg/mL	5 mL	Merck Life Science UK Ltd.	C5138-1G	Collagenase from Clostridium histolyticum
DNaseI – same location as collagenase	5 mg/mL	250 µL	100 µg/mL	2 µL	Merck Life Science UK Ltd.	DN25-100MG	Deoxyribonuclease I from bovine pancreas
MgCl <sub>2</sub>	0.5 M	250 µL	5 mM	2 µL	Fisher Scientific	10418464	MgCl <sub>2</sub> (1 M)
DMEM (with Ca)		12 mL		96 mL	PAN Biotech UK Ltd.	P04-04510	DMEM
1x PBS		12 mL		96 mL			

(1) Stock concentration; (2) Amount per 25 mL; (3) Final concentration; (4) Catalog numbers

**Table 2.8. Buffers**

<b>Wash buffer</b>	<b>Amount to mix</b>	<b>Final conc.<sup>1</sup></b>	<b>Supplier</b>	<b>Cat. No.<sup>2</sup></b>	<b>Product name</b>
FBS Supreme	6 mL	2%	PAN Biotech UK Ltd.	P30-3031HI	FBS Supreme
DMEM (without Ca)	150 mL		PAN Biotech UK Ltd.	P04-04510	DMEM
1x PBS (without Ca)	150 mL		Fisher Scientific	11503387	PBS, pH 7.4

Wash buffer	Amount to mix	Final conc. <sup>1</sup>	Supplier	Cat. No. <sup>2</sup>	Product name
<b>MACS buffer</b>	<b>Amount to mix</b>	<b>Final conc.<sup>1</sup></b>	<b>Supplier</b>	<b>Cat. No.<sup>2</sup></b>	<b>Product name</b>
FBS Supreme	600 µL	2%	PAN Biotech UK Ltd.	P30-3031HI	FBS Supreme
0.5M EDTA	60 µL	1 mM			
DMEM (without Ca)	15 mL		PAN Biotech UK Ltd.	P04-04510	DMEM
1x PBS (without Ca)	15 mL		Fisher Scientific	11503387	PBS, pH 7.4

(1) Final concentration; (2) Catalog numbers

#### 2.4.1.4.2. Percoll gradient

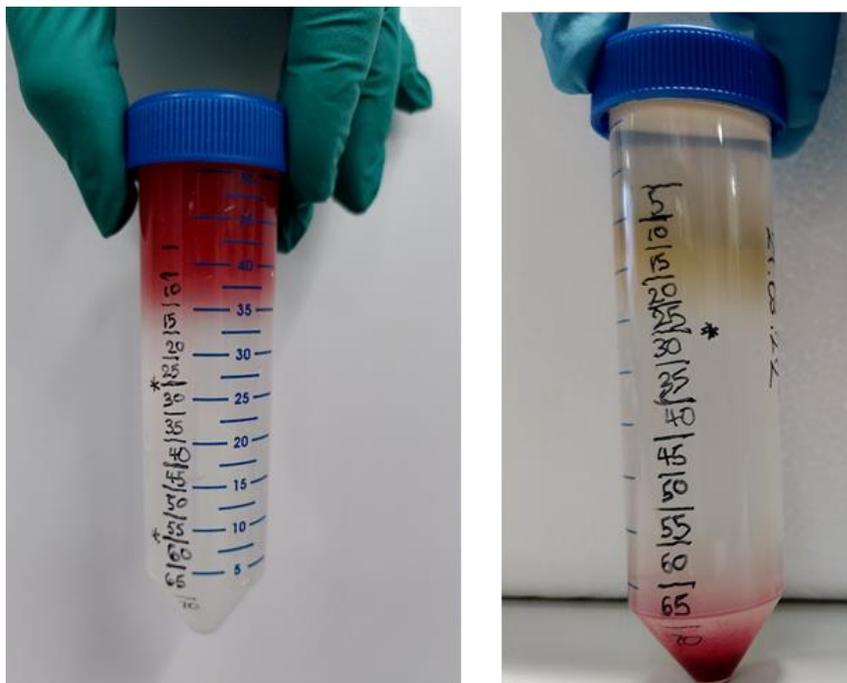
Approximately 10 mL of the cell suspension was carefully pipetted on 4 percoll gradients (**Table 2.9**), as seen in **Figure 2.7** and were centrifuged at 1,600 x g (4 accelerate, 0 brake) for 20 minutes at 21°C. The layers between 30% and 55% of percoll were carefully removed and transferred to fresh falcon tubes, followed by washing with 50 mL of the Wash buffer per tube (**Table 2.8**). The samples were further centrifuged at 400 x g at 4°C for 10 minutes. The supernatant was discarded, while 4 cell pellets were resuspended in 10 mL of the Wash buffer (per sample; **Table 2.8**) and combined, making a total of 40 mL.

**Table 2.9. Preparation of a 90% percoll density gradient for cell separation.**

A percoll stock (90%) was prepared by diluting 117 mL of well-mixed percoll with 13 mL of sterile 10 x PBS without calcium and magnesium.

Percoll concentration (%)	Amount of 90% percoll (mL)	Amount of 1 x PBS (mL)
70	15.6	4.4
65	14.4	5.6
60	13.3	6.7
55	12.2	7.8
50	11.1	8.9
45	10	10
40	8.9	11.1
35	7.8	12.2
30	6.7	13.3
25	5.6	14.4
20	4.4	15.6

Percoll concentration (%)	Amount of 90% percoll (mL)	Amount of 1 x PBS (mL)
15	3.3	16.7
10	2.2	17.8
5	1.1	18.9



**Figure 2.7. Placental cell separation using a percoll gradient.**

(A) A cell suspension is loaded onto the percoll gradient. (B) After centrifugation, the percoll gradient shows separated cells. The layers between 30% and 55% are collected for MACS positive selection.

#### 2.4.1.4.3. MACS

At the start, 40 mL of the cell suspension was split into two and centrifuged at 400 x g for 10 minutes (4°C). The supernatant was discarded from each tube. To do positive enrichment for placental stromal cells, 320 µL of the MACS buffer (**Table 2.8**) was mixed with one cell pellet. Then, 80 µL of Anti-Fibroblast MicroBeads (**Table 2.6**) was added, and the sample was incubated at 4°C for 30 minutes with constant gentle agitation (14 x rpm). The positive selection of placental trophoblasts was done by Purified anti-human EGFR antibody (**Table 2.6**) followed by Anti-Mouse IgG1 MicroBeads (**Table 2.6**). The remaining cell pellet was resuspended in 1.5 mL of the MACS buffer and 20 µL of the Purified anti-human EGFR antibody and incubated at 4°C for 30 minutes with constant

gentle agitation (14 x rpm). This step was followed by the addition of 500 µL of the MACS buffer and centrifugation at 400 x g for 10 minutes (4°C) to wash out the non-bound antibody. The cell pellet, after discarding the supernatant, was resuspended in 320 µL of the MACS buffer, mixed with 80 µL of Anti-Mouse IgG1 MicroBeads and subsequently incubated at 4°C for 15 minutes with constant gentle agitation (14 x rpm). Both tubes with anti-fibroblasts and anti-EGFR bound cells were washed by adding 1 mL of the MACS buffer (per sample), followed by centrifugation at 400 x g for 10 minutes (4°C). The supernatant was discarded, and 1.5 mL of the MACS buffer was mixed with anti-fibroblast labelled cells, while the anti-EGFR labelled cell pellet was rinsed with 1 mL of the MACS buffer. After this, two MS columns (one column per cell type; **Table 2.6**) were placed on the OctoMACS™ Separator (**Table 2.6**) with 15 mL falcon tubes for cell collection and 40 µm cell strainers on top of each MS column to prevent them from clogging. Both cell strainers with the MS columns were pre-wet with 0.5 mL of the MACS buffer, and each cell suspension was separately passed through the strainer while collecting non-labelled cells (negative selection). The columns were washed 3 more times by adding 0.5 mL of the MACS buffer (per sample), each time to remove non-labelled cells that were retained as EGFR and anti-fibroblast negative cell fractions. To collect EGFR and anti-fibroblast positive cell fractions, the MS columns with bound cells were removed from the OctoMACS™ Separator, washed with 1 mL of the MACS buffer (per sample) and a plunger inside each column was used for flushing out captured cells. This washing step was repeated 3 times in total, followed by centrifugation at 400 x g for 10 minutes (4°C) to collect cell pellets that were frozen and stored for DNA and RNA extraction (**Sections 2.3.2 & 2.3.3**).

## 2.4.2. Cell culture conditions for CT and Mole cell lines

Cell lines were grown in specialised cell media (**Appendix 4 & Appendix 6**) at 37°C in a humidified atmosphere of 5% CO<sub>2</sub> and were passaged according to the established protocol shared by Okae's laboratory (243,311). All cell lines were grown in Mycoplasma-free cell culture media.

### 2.4.2.1. Cryogenic storage and cell recovery

#### 2.4.2.1.1. Preparation of cell culture plates for cell lines

6-well cell culture plates for CT27, CT30, Mole 1 and Mole 2 cell lines were prepared in

advance. To each well of a 6-well plate, 2 mL of CT Basal Media (**Appendix 5 & Appendix 4**) and 2  $\mu$ L of iMatrix-511 (**Appendix 4**) were added, and the plates were incubated at 37°C from 10 minutes up to overnight for coating the plates. After incubation, the media is discarded and replaced with pre-warmed CT Working Media (**Appendix 4 & Appendix 6**), followed by the seeding of cells at a density of  $0.5-1 \times 10^6$  cells per well.

All cell lines were cryogenically preserved with CT Working Media containing Dimethyl sulfoxide (DMSO; D2650-100ML; Merck Life Science UK Ltd. ) in liquid nitrogen storage for long-term storage. Before seeding cells, a frozen cell aliquot was fully defrosted at room temperature. CT Working Media (2 mL) was mixed with 1 mL of the defrosted cell aliquot and centrifuged at 1,500 x rpm for 5 minutes. The supernatant was carefully removed, and the cell pellet was resuspended in 2 mL of CT Working Media and spread across a pre-coated 6-well plate.

#### **2.4.2.2. Cell maintenance and harvesting**

Plates with 80% confluency were passaged every few days, depending on the growing speed of these cells. For this, the cell media was removed, and each well of the 6-well plate was rinsed with 1 x sterile PBS to wash out residual FBS present within CT Working Media (Appendix 6), as it can inhibit trypsin. 0.5 mL of the Trypsin solution (P10-029500; PAN Biotech UK Ltd.) with 0.5 mL of 1 x PBS were added to each well, and the plate was incubated at 37°C for 5 minutes. At the end of the incubation, cells were gently scrapped to detach them from the plate. 1 mL of fresh CT Working Media was added to each well and gently mixed by pipetting. The media with detached cells was transferred to 15 mL falcon tubes and centrifuged at 1,500 x rpm for 5 minutes at room temperature. The cell pellet was either washed with 5 mL of PBS (centrifuged at 2,500 x rpm at 4°C for 5 minutes) and used for RNA extraction or resuspended in 2 mL of CT Working Media and seeded onto freshly prepared 6-well plates at a desired concentration (1:2 or 1:4 split ratio).

## **2.5. Quantitative techniques**

### **2.5.1. Placenta-specific imprinted genes with mDMR**

In the first result chapter, quantitative pyrosequencing and qRT-PCR assays were employed to further characterise the possible function of newly discovered genes with mDMRs in the human placenta. More specifically, the techniques were used to quantify

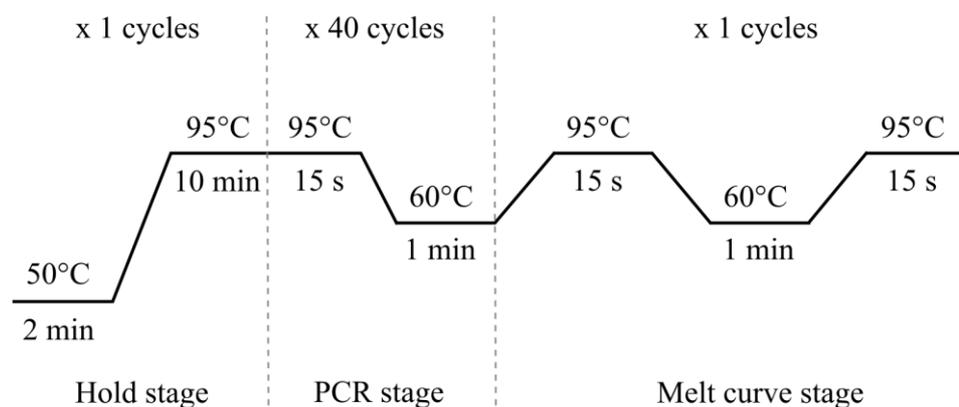
DNA methylation levels and the expression of novel placenta-specific imprinted genes in placentae from normal and complicated pregnancies (**Appendix 2 & Appendix 3**). In addition, qRT-PCR was applied to evaluate MACS-enriched cell populations from the whole placenta and finally to test maternal cell (DNA) contamination in a few placental samples.

### 2.5.2. Non-canonical imprinting in the human placenta and embryos

In the second result chapter of this thesis, qRT-PCR was used to investigate gene dosage for candidate non-canonical imprinted genes in the placenta-derived cell lines (CT and Mole cell lines) and, in the case of *XIST*, in male and female placentae.

### 2.5.3. Quantitative real-time reverse transcription PCR (qRT-PCR)

Several qRT-PCR experiments were carried out during this PhD thesis, and all followed standard conditions unless noted otherwise (**Figure 2.8**).



**Figure 2.8. Standard qRT-PCR cycling conditions for the Power SYBR™ Green PCR Master Mix.**

At the start, all oligonucleotide primers used for quantitative expression assays were vigorously tested before their use with the investigated samples. Firstly, the best primer annealing temperature was determined by running standard PCR (**Section 2.3.7**). Then, the minimal amount of cDNA template (required for a comparative Ct method), the primer specificity and efficiency were determined by a standard curve-based method for qRT-PCR. The primers used for the different qRT-PCR experiments can be found in **Appendix 26**, **Appendix 27** and **Appendix 28**. A range of cDNA samples were used for qRT-PCR experiments with concentrations ranging from ~3.8 ng/ $\mu$ L to 10 ng/ $\mu$ L. Five  $\mu$ L of diluted cDNA template was mixed with 5.5  $\mu$ L of the Power SYBR<sup>™</sup> Green PCR Master Mix (4367659; Thermo Fisher Scientific Inc.), 0.25  $\mu$ L of forward primer (0.1  $\mu$ M/ $\mu$ L) and 0.25  $\mu$ L of reverse primer (0.1  $\mu$ M/ $\mu$ L). All samples were run in triplicates. Also, all plates included a non-template control for each tested gene and a calibrator (e.g., a cDNA mix of human placenta) whenever possible. The experiments were conducted in MicroAmp<sup>™</sup> Optical 384-Well Reaction Plates (4309849; Thermo Fisher Scientific Inc.) in the QuantStudio<sup>™</sup> 5 Real-Time PCR System (A28140; Thermo Fisher Scientific Inc.). The general cycling conditions used for qRT-PCR can be seen in **Figure 2.8**. Only the samples with two technical replicates were included, and the melt or dissociation curves were scrutinised after each experiment to detect any contaminating DNA species or primer dimers. If possible, two endogenous control genes were included for each run, except for one experiment, due to non-specific amplification after inspection of melt curves.

### 2.5.3.1. Comparative $\Delta\Delta$ Ct method

Relative quantification (RQ) of a candidate gene, also known as fold change, was calculated by  $2^{-(\Delta\Delta Ct)}$  (468). During this, the Ct of a target or candidate gene is normalised to the Ct of the endogenous gene(s), which gives  $\Delta$ Ct. *RPL19* and *ACTB* were selected as endogenous control genes as these genes were tested by previous Monk group members and demonstrated relatively stable expression between different placenta samples (similar mean expression, low SD). The tested sample of a target gene is then normalised to a calibrator sample of the same target gene, with the calibrator being a cDNA mix of multiple placentae or a single cell line. This gives  $\Delta\Delta$ Ct. Finally,  $2^{-(\Delta\Delta Ct)}$  for each sample is calculated, assuming that PCR primers are efficient and give 100% amplification efficiency. The calculations for RQ or fold change as follows:

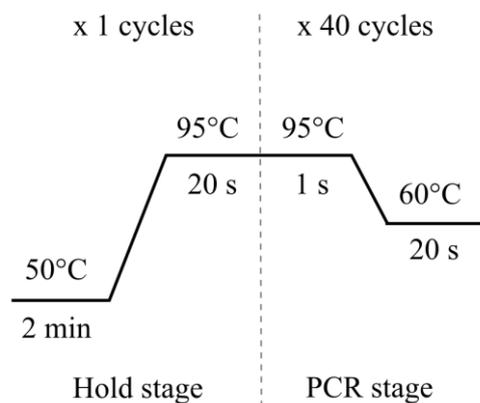
$$\Delta Ct_{(Target\ gene)} = Average\ Ct_{(Target\ gene)} - Average\ Ct_{(Endogenous\ control)} *$$

\*Geometric mean of Ct of multiple endogenous control genes

$$\Delta\Delta Ct_{(Tested\ sample\ of\ a\ target\ gene)} = \Delta Ct_{(Tested\ sample\ of\ a\ target\ gene)} - \Delta Ct_{(Calibrator\ sample\ of\ a\ target\ gene)}$$

$$RQ\ or\ Fold\ change = 2^{-(\Delta\Delta Ct)}$$

For those genes whose expression could not be quantified by the Power SYBR™ Green PCR Master Mix (4367659; Thermo Fisher Scientific Inc.), the TaqMan™ Fast Advanced Master Mix (4444557; Thermo Fisher Scientific Inc.) was employed. In the case of the TaqMan assay, the Hs00377852\_g1 and Hs02338565\_gH TaqMan™ Gene Expression Assay probes for *GoS2* and *RPL19* were purchased from Thermo Fisher Scientific Inc. For this experiment, 5 µL of the TaqMan™ Fast Advanced Master Mix (4444557; Thermo Fisher Scientific Inc.) and 0.5 µL of the TaqMan™ Gene Expression Assay probe were mixed with 4.5 µL of cDNA (9 ng/µL) diluted in UPW (750023, Thermo Fisher Scientific Inc.) with the cycling condition shown in **Figure 2.9**.



**Figure 2.9. Standard TaqMan cycling conditions.**

This assay was performed in the same way as the qRT-PCR experiments carried out with the Power SYBR™ Green PCR Master Mix. Samples were run in triplicates with a non-template control and a calibrator sample. Only samples with 2 acceptable technical replicates were included in the  $2^{-(\Delta\Delta Ct)}$  analysis.

The results from qRT-PCR experiments were generated and inspected using

QuantStudio™ Design & Analysis Software v1.3.1 (Applied Biosystems™ by Thermo Fisher Scientific Inc.). To combine the results from several MicroAmp™ Optical 384-Well Reaction Plates, which included the same genes but different samples (cDNA templates), ExpressionSuite Software v1.3 (Applied Biosystems™ by Thermo Fisher Scientific Inc.) was used. For each gene across multiple plates, the same baseline threshold (the signal or noise level detected during the initial 3 and 15 cycles of qRT-PCR) was set manually, allowing for comparison between different plates. The amplification curves and melting curves across several plates were inspected again by ExpressionSuite Software v1.3. The generated fold change results were further analysed and plotted by Rstudio, an in-house R script (R version 4.3.2).

#### 2.5.4. Pyrosequencing

Pyrosequencing is a sequencing-by-synthesis method, during which the real-time incorporation of a nucleotide into a newly synthesised strand is detected as a light signal (469). This method relies on four enzymes: the Klenow fragment of the DNA polymerase I, ATP sulfurylase, luciferase, and apyrase. At the start of the sequencing reaction, a single nucleotide is injected into wells of a microtiter plate, and it is incorporated at the 3' end of a sequencing primer by the DNA polymerase I, which results in a release of pyrophosphate (PPi). PPi is further used by the ATP sulfurylase to generate ATP from adenosine phosphosulfate (APS). The luciferase uses ATP to oxidise D-luciferin. The resulting product of this reaction is oxyluciferin, which is excited and gradually emits light that is captured by a camera and converted into pyrograms. In the meantime, unincorporated nucleotides are degraded by the apyrase before the injection of a new nucleotide.

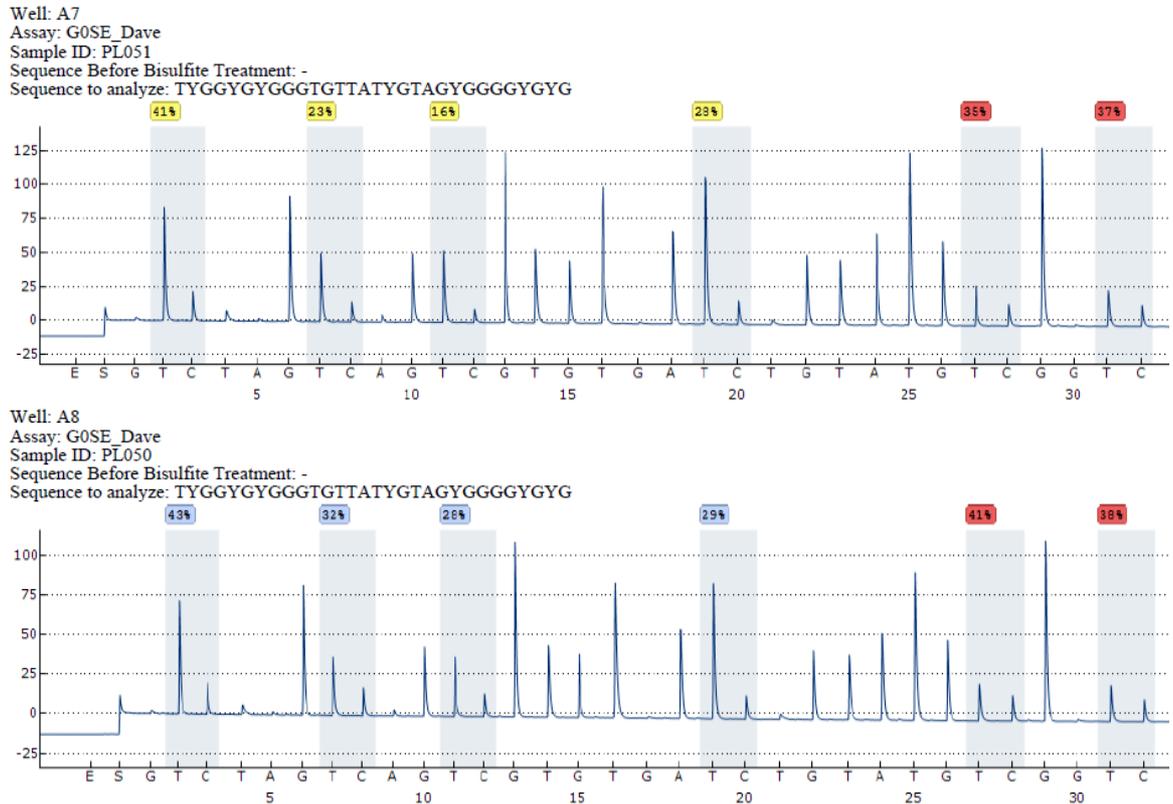
Unlike bisulphite PCR, followed by cloning, which can reveal the DNA methylation status at several CpG sites, pyrosequencing allows quantifying the level of DNA methylation at those CpG sites in percentages. Additionally, this technique is well-suited for complex genomic regions that are rich in highly repetitive elements, such as LINE -1 elements found across the human genome, which could not be sequenced with other NGS technologies such as WGBS or RRBS-seq (71). Unfortunately, only short regions of the genome that are highly enriched with CpG sites can be sequenced by pyrosequencing.

With all this in mind, the pyrosequencing approach was employed to quantify the level of DNA methylation at the promoters with mDMRs of the candidate placenta-specific imprinted genes in the placental cohort. For this assay, 77 placentae were selected with similar distributions in gender. All used samples are listed in **Appendix 2** and **Appendix**

### 3.

For pyrosequencing, the DMR regions densely populated with CpG sites were amplified using bisulphite PCR (**Section 2.3.7.1.4**), with reverse primers being tagged with biotin, while the sequencing primers were designed to anneal to the complementary strand. The primers for these regions can be found in **Appendix 26**. 20  $\mu\text{L}$  of remaining bisulphite PCR products, after running on an agarose gel, was used for pyrosequencing. The downstream procedure was carried out by a former PhD student, Dr Ana Monteagudo-Sánchez, at the Institute Jacques Monod, CNRS & Université Paris-Cité (France). In short, a bisulphite PCR product was diluted in UPW (750023; Thermo Fisher Scientific Inc.) to make a total volume of 40  $\mu\text{L}$ . Each diluted sample was further mixed with 38  $\mu\text{L}$  of the PyroMark Binding Buffer (Qiagen) and 2  $\mu\text{L}$  (10 mg/mL) streptavidin-coated Sepharose® beads (Qiagen) and agitated at 1,600 x rpm for 10 minutes at room temperature. The PCR products, in a 96-well plate, were purified from salts and unused reagents by using the PyroMark Q96 Vacuum Prep Workstation. Single-stranded bisulphite PCR amplicons with the incorporated biotinylated primer were immobilised on streptavidin-coated Sepharose® beads (Qiagen). The bound amplicons were then washed with 70% ethanol, denatured with sodium hydroxide, and resuspended in the PyroMark Buffer (Qiagen). The single-stranded DNA was hybridised to 40 pmol of sequencing primer dissolved in 11  $\mu\text{L}$  of the PyroMark Annealing Buffer (Qiagen) during a 2-minute incubation on a heating block at 80°C. Pyrosequencing was carried out on a PyroMark Q96 instrument.

The pyrosequencing results, such as methylation percentages, were determined from C and T ratios at each CpG site within the sequenced PCR product and were produced with Pyro Q CpG1.0.9 software (Biotage). The pyrograms were inspected, and only the good-quality CpGs were used for further analysis, as shown in **Figure 2.10**. Only the first few CpGs that were marked in blue and yellow, denoted as good quality, were used for the analysis. The methylation percentages recorded in the pyrograms were further analysed and plotted with Rstudio, an in-house R script (R version 4.3.2).



**Figure 2.10. Pyrograms of two placental samples showing the level of DNA methylation (%) at seven CpG sites (grey highlight) in a PCR amplicon.**  
 DNA methylation percentages in blue indicate a good-quality score, in yellow – a score with minor deviations and in red - a poor-quality score.

## 2.6. Single-cell methylation and transcriptome sequencing (scM&T-seq) of human pre-implantation embryos

### 2.6.1. Overview of scM&T-seq

scM&T-seq technique was established by (470). This technique was developed based on a slightly modified version of G&T-seq (471). This method physically separates single-cell gDNA and mRNA with poly(A) tails, allowing them to be processed in parallel according to either the scBS-seq or SMART-seq2 protocols, followed by sequencing on an Illumina platform (472,473). In more detail, cells are sorted into 96-well plates and lysed in the RLT plus lysis buffer (Qiagen) to release gDNA and RNA. Polyadenylated mRNAs are captured by biotinylated oligo-dT primers, which are immobilised on streptavidin-coated

magnetic beads (Dynabeads™ MyOne™ Streptavidin C1), while these beads are captured by a magnet. The supernatant containing gDNA is removed and stored in a new 96-well plate and then frozen at -20°C until further processing. Captured mRNA is thoroughly washed, reverse transcribed, and amplified while still bound to the beads. The subsequent steps are carried out using the Nextera XT DNA Library Preparation Kit (FC-131-1096; Illumina, Inc.) following the manufacturer's instructions. The following key steps include cDNA dilution before tagmentation with the Nextera transposome, additional PCR amplification for incorporation of index adaptors, library pooling and cleaning with Beckman Coulter™ Agencourt AMPure XP beads (10453438; Fisher Scientific) on a magnet followed by paired-end sequencing.

Separated DNA alongside is purified with Beckman Coulter™ Agencourt AMPure XP beads while captured with a magnet. The purified DNA is then used for bisulphite conversion with the EZ DNA Methylation-Direct Kit (D5021; Zymo) while remaining attached to the magnetic beads. The bisulphite-treated DNA is then used for post-bisulphite adaptor tagging (PBAT) with a 3'-random hexamer. Finally, the first synthesised strand is removed by biotin capture, and only the second synthesised strand, after cleaning and PCR amplification, is used for sequencing on an Illumina instrument.

## **2.6.2. Preparation of single-cell RNA sequencing (scRNA-seq) libraries**

In general, this technique is well-suited for samples with low amounts of starting genetic material, such as single cells of human pre-implantation embryos, which require higher sequencing coverage. In addition, this technique was successfully applied by our group previously (71). During this PhD, only the transcriptome part was performed. The scRNA-seq libraries were prepared at the Earlham Institute (Norwich, UK) with the assistance of Dr Louise Chappell-Maor and in collaboration with Dr Iain Macaulay's group.

In brief, the high-quality surplus human IVF embryos donated to research were split into single cells. A Hamilton-Thorne Lykos laser was used to remove one-quarter of the zona pellucida for one embryo with 4 cells and 10 embryos containing between 5 and 12 cells. Blastomeres from these embryos were separated using blastomere biopsy micropipets (Origio, USA) and isolated with a stripper using 120 µm tips. Each blastomere was further washed in 1% PVP and then placed in a sterile PCR tube containing 2.5 µL of PBS and snap-frozen at -80°C until further processing. The ICM and TE of two blastocysts were

separated by micromanipulation using a laser (OCTAX, Herborn, Germany). However, for the two other blastocysts, it was not possible to separately isolate ICM and TE. Separated ICM and TE cells, or a mixture of both, were incubated in an Accutase medium (Chemicon) at room temperature for 10 minutes to isolate single cells by gentle pipetting. These cells were further washed with 1% PVP and then placed in sterile PCR tubes containing 2.5  $\mu$ L of PBS, followed by snap-freezing at  $-80^{\circ}\text{C}$  until processing. Following the scM&T-seq protocol, each isolated cell was transferred to 5  $\mu$ L of the RLT plus lysis buffer (Qiagen) and snap-frozen at  $-80^{\circ}\text{C}$  until downstream processing. In total, 204 cells were collected from different stages of human pre-implantation embryos.

To prepare scRNA-seq libraries, the single-cell lysates were transferred to 96-well plates, with each well containing a single sample. Approximately 5  $\mu$ L of cell lysate was mixed with 10  $\mu$ L of Dynabead mix (**Table 2.10**) containing oligo (dT) and incubated on a ThermoMixer® C (Eppendorf) for 20 minutes at 1,300 x rpm. After this step, the Biomek FX<sup>P</sup> Laboratory Automation Workstation was used to separate mRNA from gDNA, which remained in the supernatant (~40  $\mu$ L) and was stored at  $-20^{\circ}\text{C}$ . While mRNA was washed twice with G&T-seq wash buffer (**Table 2.11**), 5  $\mu$ L of the RT Master Mix (**Table 2.12**) was added into each well for reverse transcription with the cycle conditions present in **Figure 2.11**. Then, 7.5  $\mu$ L of PCR Master Mix was pipetted into each well, and the samples were further amplified following the cycling conditions outlined in **Figure 2.12** and **Table 2.13**.

**Table 2.10. Bead Mix**

Reagent	Supplier	Cat. No. <sup>1</sup>	Reagent	Volume ( $\mu$ L) for 1 reaction
Dynabeads™ MyOne™ Streptavidin C1	Thermo Fisher Scientific Inc.	65001	Dynabeads™ MyOne™ Streptavidin C1	0.5
5 x First-strand buffer	Thermo Fisher Scientific Inc.	18064071	5 x First-strand buffer	1.5
RNase Inhibitor (20 U/ $\mu$ L)	Thermo Fisher Scientific Inc.	N8080119	RNase Inhibitor (20 U/ $\mu$ L)	0.5
Nuclease-Free Water (not DEPC-Treated)	Thermo Fisher Scientific Inc.	AM9938	Nuclease-Free Water (not DEPC-Treated)	7
G&Tseq Oligo dT (1 $\mu$ M)	Integrated DNA Technologies, Inc.	Custom	G&Tseq Oligo dT (1 $\mu$ M)	0.5

(1) Catalog number

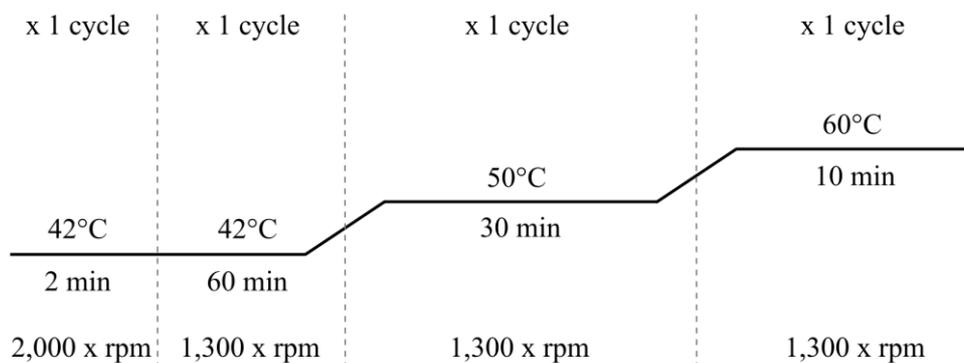
**Table 2.11. G&T-seq wash buffer**

Reagent	Final molarity required	To make up to 50ml
Tris-HCl (pH 8.3)	50 mM	25 ml at 0.1M
KCl	75 mM	1.875 ml at 2M
MgCl <sub>2</sub>	3 mM	300 ul at 0.5M
DTT	10 mM	500 ul at 1M
Tween 20	0.5% (vol/vol)	50 ul at 50% (vol/vol)
Nuclease-free water		21.8 ml

**Table 2.12. RT Master Mix**

Reagent	Supplier	Cat. No. <sup>1</sup>	Volume (μL) for 1 reaction	Final concentration
Nuclease-Free Water (not DEPC-Treated)	Thermo Fisher Scientific Inc.	AM9938	7	
dNTP Mix (10 mM each)	Thermo Fisher Scientific Inc.	R0193	0.5	1 mM
G&T-seq TSO custom LNA oligonucleotide (100 μM)	Integrated DNA Technologies, Inc.	Custom	0.05	1 μM
MgCl <sub>2</sub> (1 M)	Thermo Fisher Scientific Inc.	AM9530G	0.03	6 mM
Betaine solution (5M)	Merck Life Science UK Ltd.	B0300-1VL	1	1 M
5 x First-strand buffer	Thermo Fisher Scientific Inc.	18064071	1	1 x
DTT (100 mM)	Thermo Fisher Scientific Inc.	18064071	0.25	5 mM
SuperScript II Reverse Transcriptase (200 U/μL)	Thermo Fisher Scientific Inc.	18064071	0.25	10 U/μL
RNase Inhibitor (20 U/μL)	Thermo Fisher Scientific Inc.	N8080119	0.125	0.5 U/μL

(1) Catalog number

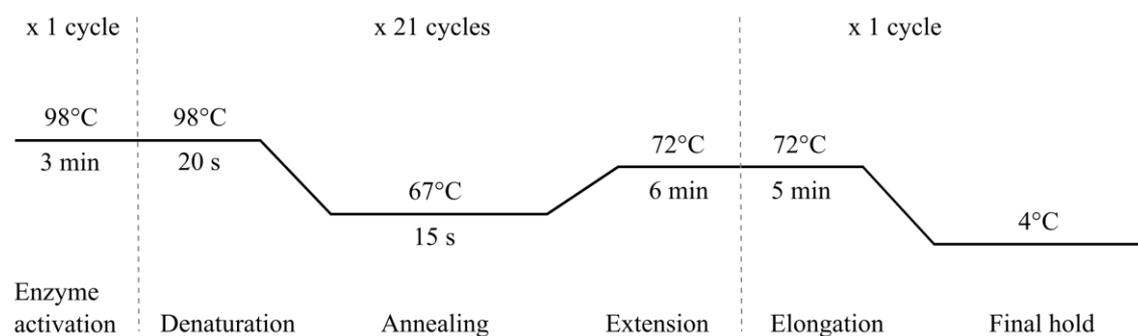


**Figure 2.11. Reverse transcription conditions with RT Master Mix.**

**Table 2.13. PCR Master Mix**

Reagent	Supplier	Cat. No. <sup>1</sup>	Volume (μL) for 1 reaction	Final concentration
Phusion Hot Start II High-Fidelity PCR Master Mix (2 x)	Thermo Fisher Scientific Inc.	F565L	6.25	1 x
IS PCR primer (10 μM)	Integrated DNA Technologies, Inc.	Custom	0.125	0.1 μM
Nuclease-free water	Thermo Fisher Scientific Inc.	F565L	1.125	

(1) Catalog number

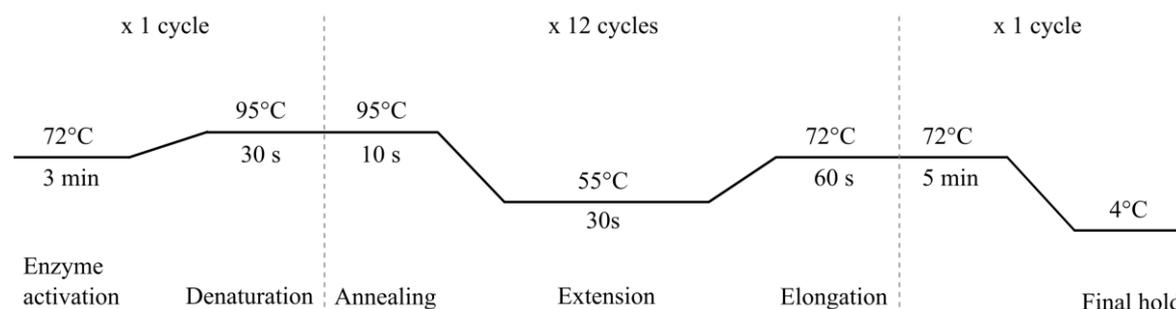


**Figure 2.12. PCR cycling conditions for KAPA HiFi HotStart PCR Master Mix.**

Amplified cDNA was subject to post-PCR amplification clean-up, which was performed with the Biomek NXP Automated Workstation. cDNA in a volume of 25 μL was cleaned with Beckman Coulter™ Agencourt AMPure XP beads (10453438; Fisher Scientific) and with two consecutive 80% ethanol washes followed by elution in 20 μL of Nuclease-Free

Water (AM9938; Thermo Fisher Scientific Inc.). The concentration and quality of cDNA were evaluated by a Qubit Fluorometer (Thermo Fisher Scientific Inc.) and a 2100 Bioanalyzer Instrument (Agilent Technologies Inc.).

ScRNA-seq was performed as described in (470,471) with some adjustments. The libraries were made with the Nextera XT DNA Library Preparation Kit (FC-131-1096; Illumina, Inc.) according to the manufacturer's instructions, except that 1/12.5 of the original kit volume was used. Before the start, amplified cDNA was diluted to 0.2 ng/ $\mu$ L in 15  $\mu$ L. The subsequent steps were automated and carried out by the mosquito LV genomics (SPT Labtech Ltd.). 0.4  $\mu$ L of diluted cDNA was mixed with 1.2  $\mu$ L of ATM and TD mix in each well (for one 96-well plate: 55.038  $\mu$ L ATM mixed with 102.213  $\mu$ L TD). The plate was centrifuged and incubated in a PCR thermocycler at 55°C for 10 minutes for tagmentation. To stop the tagmentation reaction, 0.4  $\mu$ L of 0.2% SDS was added to each well and mixed well, followed by a 5-minute incubation at room temperature. For the indexing part, 1.2  $\mu$ L of NPM mix and 0.8  $\mu$ L of each index were added into each well and mixed. The plate was placed into the PCR thermocycler with the following cycling conditions (**Figure 2.13**).



**Figure 2.13. PCR cycling conditions for Nextera PCR Master (NPM) Mix to incorporate indexes**

Ninety-six indexed samples (1  $\mu$ L) were pooled in a total volume of 96  $\mu$ L; the same step was repeated for the second plate. The pooled libraries were cleaned again with 0.8 x AMPure XP beads and 2 consecutive washes with 80% ethanol using the Biomek NX<sup>P</sup> Automated Workstation. The cleaned libraries resuspended in ~20  $\mu$ L of Nuclease-Free Water were tested with the Qubit Fluorometer and the 2100 Bioanalyzer Instrument. Overall, pooled libraries were diluted to 10 nM and combined in equal amounts for sequencing on an Illumina NovaSeq X Plus 10B flow cell, generating 150 bp paired-end reads.

## 2.7. Publicly available methyl-seq datasets

During this thesis, various human cell-type and tissue methylation datasets were implemented as methylation tracks in the UCSC genome browser (474) (<http://genome.ucsc.edu>) to screen for placental DMRs or germline-derived DMRs. The analysis of these tracks was carried out by our previous group members and explained in more detail in (20). In summary, these tracks included 25 methylome datasets available at Gene Expression Omnibus (GEO) or NBDC repositories. The analysed tracks included methylomes for human oocytes (JGAS000000000006), 5 for human sperm (JGAS000000000006 and GSE30340), 2 for the brain (GSM913595, GSM916050), 3 for CD4+ lymphocytes (GSE31263), and single datasets for blastocysts (JGAS000000000006), muscle (GSM1010986), CD34+ cells (GSM916052), sigmoid colon (GSM983645), lung (GSM983647), aorta (GSM983648), oesophagus (GSM983649), small intestine (GSM983646), pancreas (GSM983651), spleen (GSM983652), liver (GSM916049), adrenal (GSM1120325) and adipose tissue (GSM1010983). These methylation tracks were mapped to the GRCh37 genome. Only CpG sites covered by at least 5 reads were analysed, and the average methylation estimates were obtained for samples with several technical replicates, except for human oocytes with poor coverage. For the oocytes, the methylated and unmethylated calls from the two experiments were summed to estimate the methylation ratio.

Our group previously performed WGBS for the brain, liver and third-trimester human placenta, and these datasets are described in (449) and can be found in the GEO repository (GSM1134680, GSM1134681 and GSM1134682, respectively). In short, the WGBS datasets were aligned to the GRCh37 reference genome, and the percentage of DNA methylation at a single CpG site was calculated by dividing all methylated reads by the total number of reads (methylated and unmethylated). Only CpG sites covered by a minimum of 5 reads were included in this estimate.

For comparative analysis, methylation datasets for mouse ICM, oocyte, sperm, 2-cell stage and 4-cell stage embryo, E6.5 embryo, mESC, placenta and cerebellum were also included to look at the gene promoters from (475–477) (GSE56697, GSE30206, GSE42836). The methyl-seq datasets were aligned to the GRCm38 mouse reference genome, and CpG methylation was estimated in the same way as for human methyl-seq datasets.

## 2.8. Publicly available Illumina methylation array datasets

During this PhD, 22 placental methylation datasets generated with the Illumina Infinium Human Methylation450 BeadChip array were used to aid in the screening of placenta-specific mDMRs and placental sDMRs. These datasets were downloaded from the GEO repository under accession number GSE120981. The placental samples used in these datasets are described in greater detail in Monteagudo-Sánchez *et al.* (2019) (295) and include 9 normal and 13 placental samples affected by IUGR.

In brief, control probes were used to remove background signals and reduce inter-plate variation in BeadStudio (version 2011.1\_Infinium HD). Methylation probes were excluded from downstream analysis if they had a detection p-value  $> 0.01$ , contained SNPs in the interrogation or extension bases, exhibited cross-reactivity due to multiple homologous sequences, or showed no signal in one or more placental samples. Probes located in candidate genomic regions, such as placenta-specific mDMRs or placental sDMRs, were used to inspect beta values, with promising candidates displaying approximately 50% methylation.

To investigate placental cell-type-specific methylation in candidate regions, 95 placental cell-type methylation datasets, generated using the Infinium MethylationEPIC BeadChip array, were downloaded from the GEO repository under accession number GSE159526. The placental samples and cell-type-specific isolation processes were extensively described by Yuan *et al.* (2021) (364). Briefly, four placental cell types, including HBs, trophoblasts, stromal, and endothelial cells, along with matched whole chorionic villi, were isolated from 19 third-trimester placental samples using FACS with cell-type-specific antibodies such as 7-AAD, CD235a FITC, CD45 APC-eFluor780, CD14 PE, CD34 APC, and EGFR PeCy7. DNA from the isolated cell types was bisulphite-converted using the EZ DNA Methylation Kit prior to hybridisation on the Infinium MethylationEPIC BeadChip array.

Methylation probes were excluded from further analysis if they had a detection p-value  $> 0.01$ , a bead count  $< 3$ , exhibited cross-reactivity due to multiple homologous sequences, contained SNPs within 5 bp of the interrogation site, or were located on sex chromosomes. Samples were excluded if mismatched genotypes were observed between different cell types and their corresponding whole chorionic villus samples or if maternal contamination was detected. Good quality probes were normalised across samples to

remove technical variance. The beta values from the 95 placental cell-type-specific methylation datasets were extracted for the cell-type-specific comparisons of candidate genomic regions.

## 2.9. Bioinformatics analysis

### 2.9.1. Gene synteny analysis

Several human imprinting clusters have been identified through sequence homology or synteny by comparing mouse imprinted genes with their human homologues (18,261,371,436). Also, novel imprinted genes have often been identified based on their proximity to known imprinting clusters (18,261,371,436). In classical genetics, synteny refers to homologous genes located on corresponding chromosomes across different species (478,479). In modern genetics, it refers to genes or chromosomal segments with a conserved order (collinearity) across the chromosome sets of two or more species (480). During this PhD, synteny analysis was conducted to determine whether chromosomal regions containing non-canonical imprints in mice and rats are highly conserved in humans and whether non-canonical imprints might also form clusters in the mouse, rat and human genomes - a feature characteristic of canonical imprints.

Macro-synteny plots between the mouse and human and between the rat and human were generated using a ShinySyn application (481) developed with the Shiny package (R package). It allows an interactive visualisation of synteny analysis results generated by a multiple collinearity scan (MCscan) algorithm (482–484). This algorithm identifies putative homologous chromosomal segments across multiple genomes by using gene pairs with high pairwise collinearity as anchors (reference points) and performs multiple alignments for those homologous chromosomal regions. Initially, MCscan uses BLASTP to compare several genomes and retrieves the most highly scoring hits, which are then sorted by gene positions. The hits are later applied in dynamic programming to identify collinearity blocks between different chromosomes and, in other words, to find regions of synteny and collinearity.

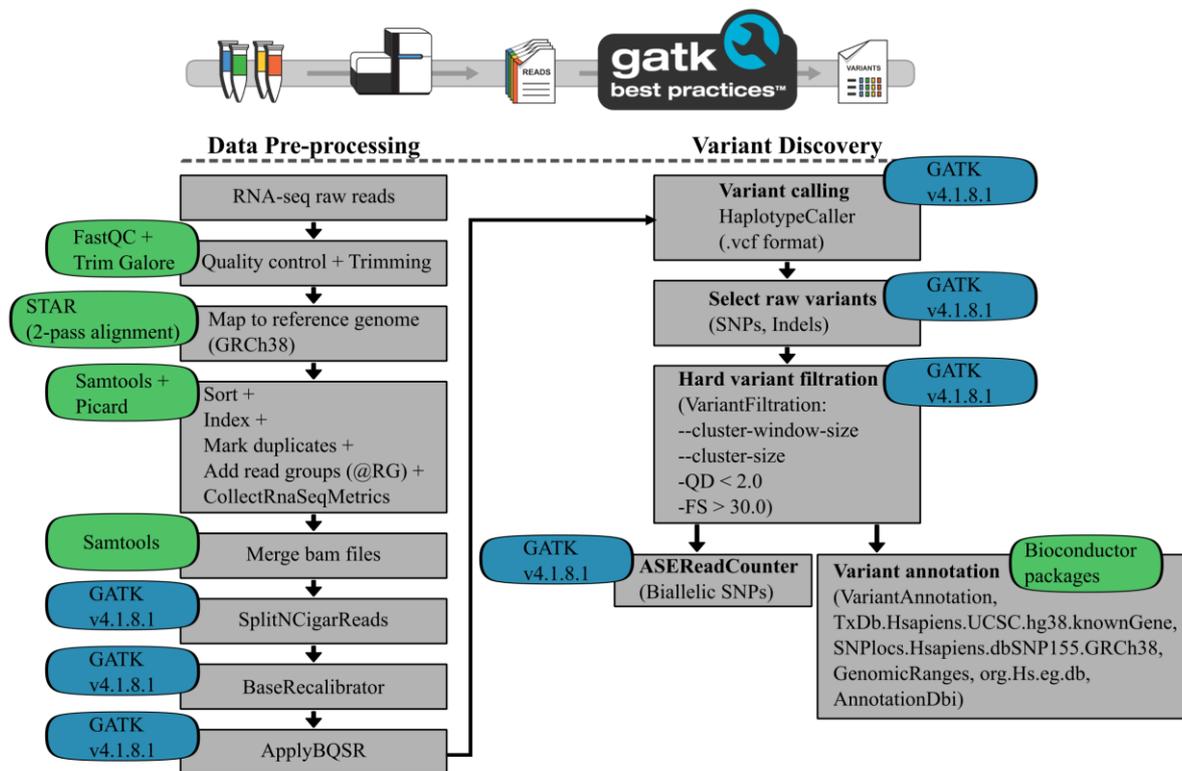
For the ShinySyn application, the FASTA files of the most recent versions of the human (GRCh38) and mouse (GRCm39) genomes, as well as GTF files for transcript and gene annotations, were downloaded from the GENCODE database, while the rat genome (mRatBN7.2) was downloaded from the Ensembl genome database. These files were

applied to the MCscan pipeline, which generated BED and anchor gene files, used as inputs to create macro-synteny, micro-synteny and dot plots. The list of mostly mouse and rat of non-canonical imprints was checked in the output files to create macro-synteny plots.

## 2.9.2. ScRNA-seq processing

### 2.9.2.1. Overview of scRNA-seq analysis

The scRNA-seq libraries were sequenced at the Earlham Institute (Norwich, UK). The pipeline used to process scRNA-seq FASTQ files consists of several stages, as illustrated in **Figure 2.14**. The first stage involves checking the quality of raw reads, trimming adapters, and excluding poor quality and short reads prior to alignment with the human reference genome. The second stage involves aligning reads to the reference genome, followed by sorting, indexing and checking the alignment quality. The third stage involves marking duplicated reads in aligned BAM files, adding Read Group (RG) tags, and inspecting the duplication rate and alignment rate of nucleotides across genomic regions such as untranslated regions (UTRs, introns, intergenic regions) and peptide coding regions (exons). During the fourth stage, the single-cell aligned BAM files are merged into pseudo bulks to make whole embryos; this is done to increase coverage for polymorphisms. The fifth stage consists of several GATK tools (485) that are used to adjust and recalibrate the BAM files for variant calling (HaplotypeCaller). In the final stage, different GATK tools are used to select and filter polymorphisms and perform ASE analysis, while a range of Bioconductor/R packages are applied to annotate polymorphisms identified at the earlier stages.



**Figure 2.14. Summary of variant calling analysis in pseudo bulk scRNA-seq datasets.**

Green and blue boxes represent the bioinformatic tools and their respective versions used at each analysis step, which are indicated in grey boxes.

### 2.9.2.2. ScRNA-seq analysis

At the start, 374 raw FASTQ files (187 cells and 5 non-template controls) were inspected with FastQC v0.11.9, followed by trimming the Nextera adapters, excluding short (15 bp) and poor quality ( $-q\ 20$ ) reads by Trim Galore! v0.6.5 ( $--phred33, --paired$ ). The remaining reads were aligned to the GRCh38.p14 reference genome from the GENCODE database by STAR v2.7.10a (486) ( $2\text{-pass}\ mode$  - more sensitive to splice junctions). The aligned reads were sorted and indexed by SAMtools v1.16.1 (487). Picard v3.1.1 was used to mark duplicated reads with  $--TAGGING\_POLICY\ All\ --OPTICAL\_DUPLICATE\_PIXEL\_DISTANCE\ 2500\ --REMOVE\_DUPLICATES\ false$ , add RG tags with `AddOrReplaceReadGroups` and inspect the aligned base distribution within different genomic regions with  $-REF\_FLAT\ refFlat.txt, -STRAND\ NONE$ . Pseudo bulk files were generated by SAMtools v1.16.1 merge by merging single-cell aligned files. SplitNCigarReads from GATK v4.9.1 (485,488–491) was used to adjust the format of pseudo bulk files, while BaseRecalibrator ( $--known-sites\ Homo\_sapiens\_assembly38.dbsnp138.vcf\ --known-sites\ Homo\_sapiens\_assembly38.known\_indels.vcf$ ) and ApplyBQSR to adjust the

quality scores in pseudo bulk files. `HaplotypeCaller` generated raw VCF files containing all polymorphisms. `SelectVariants` option generated separate VCF files for SNPs and indels for `VariantFiltration` with `--cluster-window-size 35 --cluster-size 3 --filter-name "QD_filter" -filter "QD < 2.0" --filter-name "FS_filter" -filter "FS > 30.0"`. Filtered VCF files with SNPs were applied to `ASEReadCounter` to find biallelic SNPs, and filtered VCF files with SNPs were also annotated by Bioconductor packages (R v4.3.1, `VariantAnnotation`, `TxDb.Hsapiens.UCSC.hg38.knownGene`, `SNPlocs.Hsapiens.dbSNP155.GRCh38`, `GenomicRanges`, `org.Hs.eg.db`, `AnnotationDbi`, `BiocManager`). Summary reports were generated by `MultiQC v1.17` (492).

### 2.9.2.3. Bioinformatic programs and tools

More detailed descriptions of the programs and tools used in this pipeline are provided below.

#### 2.9.2.3.1. STAR aligner

scRNA-seq datasets were aligned using the `STAR v2.7.10a` aligner (486), which stands for Spliced Transcripts Alignment to a Reference (STAR) software. It is a widely used aligner that is fast, sensitive and splice-aware, which can be used for short-read sequencing produced by Illumina instruments but is also suitable for longer reads generated by third-generation sequencing technologies (493,494). This aligner was designed for bulk RNA-seq, but it is also a popular choice for scRNA-seq (495).

The developed algorithm for this aligner includes two major steps (486). The first step involves a seed search, and the second step involves clustering, stitching and scoring of the discovered seeds. Initially, the algorithm searches for the longest complementary stretches of a genome that map to a sequencing read (known as seeds), starting from the first base of that read until a splice site is reached. The first part of the read or the seed is mapped to a donor splice site, while the second part of the read is mapped to an acceptor splice site. At this stage, genetic variants and sequencing adapters are detected, while parts of the reads with a poor alignment rate to the genome undergo soft-end clipping. During the second stage, all mapped seeds belonging to the same sequencing read are clustered together based on their proximity. The seeds are then stitched and scored according to their local alignment, in which indels, mismatches, and splice junctions are penalized. The highest-scoring stitched combination is determined as the best alignment of that read. For reads that map to multiple regions of the genome, all alignments (stitched combinations)

are reported that are above a predefined threshold determined by a user.

#### 2.9.2.3.2. **GATK**

The Genome Analysis Toolkit (GATK) is a pipeline developed by the Broad Institute of MIT and Harvard, which includes a series of tools to discover and further process germline and somatic variants found in different sequencing datasets (485,488–491). It became widely used for large cohorts with hundreds of samples, as this workflow offers joint genotyping analysis, which accurately infers SNPs and small variant copy changes (indels, deletions) across multiple samples simultaneously, thereby dramatically reducing computational resources.

#### 2.9.2.3.3. **GATK pipeline for RNA-seq:**

##### 2.9.2.3.3.1. **SplitNCigarReads**

The first step after aligning reads is to use the SplitNCigarReads tool, which is required for variant calling with HaplotypeCaller to reduce the number of false positives. This tool splits aligned reads that contain Ns' in their cigar strings, with Ns' indicating splicing events. In the process, multiple supplementary alignments are produced with mismatching overhangs being trimmed. Also, the mapping quality score is reassigned to match DNA conventions.

##### 2.9.2.3.3.2. **BaseRecalibration & ApplyBQSR**

BaseRecalibration creates an empirical error model that is applied to adjust the base quality scores provided by an aligner, which may be biased by systemic technical errors made by the sequencing machine, resulting in under- or over-estimation of base quality score results. In more detail, the algorithm detects bases that do not align with the reference genome (reference mismatches). It groups those mismatches based on four major sources of systemic errors, which are a machine cycle, dinucleotide context, read group and the base reported quality score and calculates covariates. These covariates are used to derive error estimates, which are applied to recalibrate the base quality scores with ApplyBQSR in the input/alignment files. The known variants that are known to vary in the human population are not corrected by this empirical error model.

##### 2.9.2.3.3.3. **HaplotypeCaller**

Variant calling was performed with the HaplotypeCaller tool, which looks for biallelic and

multiallelic SNPs and indels by local *de novo* assembly of haplotypes. Then, a variation in a sequencing read is detected, and HaplotypeCaller resembles the alignment at that position, which is useful for overlapping variants present in proximity to each other (where other variant callers can struggle). Also, it can process splice junctions present in RNA-seq data that can produce false negative calls.

This tool implements several algorithms in a few-step process. Firstly, it identifies regions with possible signs of variation (active regions). For each such active region, a De Bruijn-like graph is produced that shows all possible haplotypes for each active region (haplotypes are identified during de-novo reassembly of the active region). Then, variant sites are detected by the Smith-Waterman algorithm that realigns each possible haplotype to the reference haplotype. The PairHMM algorithm is further applied for a pairwise alignment, aligning all reads to all possible haplotypes, resulting in a matrix with likelihoods. These generated likelihoods are further marginalised to obtain likelihoods for variants. Finally, Bayes' theorem rules are used to implement previously calculated likelihoods for variants to determine the most likely genotype of a sample. The given output of this tool is a list of raw unfiltered SNP and indel calls (genotype calls).

#### 2.9.2.3.3.4. Hard filtering for RNA-seq:

Hard filters are selected thresholds by a user that are applied to variant annotations (statistical estimates). Variants with annotations below or above thresholds are excluded from the final list of variants.

**Cluster-window-size** - looks at 35 bp windows with at least 2 SNPs making a cluster.

**QualByDepth (QD)** or Quality score by depth is generated after a variant quality is normalised by its coverage. The quality of a variant might be inflated due to deep sequencing.

**FisherStrand (FS)** - the phred-scaled probability indicating a strand bias at a variant site. It is the probability that an alternative allele was detected at a higher frequency on each sequencing read strand than a reference allele.

#### 2.9.2.3.3.5. ASEReadCounter

ASEReadCounter tool calculates read counts at the heterozygous SNP positions that are biallelic after applying filters on mapping quality, base quality, coverage depth,

overlapping paired-end reads and deletions for ASE analysis. Provides a text table with allele counts at each heterozygous SNP.

**Chapter 3: *PIK3R1* and *GOS2* are human  
placenta-specific imprinted genes  
associated with germline-inherited  
maternal DNA methylation**

### 3.1.Introduction

Genomic imprinting is vital for normal placental development as abnormal expression of imprinted genes is observed in various placental pathologies (371–373). BWS can be caused by several genetic and epigenetic aberrations, including hypomethylation at *KvDMR1*, hypermethylation at the *IC1* of *H19/IGF2*, mutations in *CDKN1C*, paternal uniparental disomy 11p15 or paternal 11p15 duplication (420,496,497). Patients with BWS can demonstrate placentomegaly, placental mesenchymal dysplasia, chorangioma/chorangiomatosis and extravillous trophoblastic cytomegaly (447,448). Interestingly, a subset of BWS fetuses carrying a mutation in the *CDKN1C* gene frequently caused PE in the mothers during pregnancy (407). *CDKN1C* is a maternally expressed gene that encodes a Cyclin-dependent kinase inhibitor 1C and is one of the multiple genes present in a known imprinting cluster controlled by the *KvDMR1* (291,498). More recently, it was shown that *CDKN1C* was downregulated in trophoblast cells derived from CHM, which are characterised by excessive growth of trophoblast tissue (311,499). This gene was shown to be responsible for cell cycle arrest under direct contact inhibition in high-density cell culture conditions as CT cells derived from normal first-trimester placenta stopped dividing, while CHM-derived cells continued to proliferate. Therefore, it is crucial to study the function of imprinted genes as they might be directly involved in placental-associated diseases.

Significant improvements in high-throughput sequencing technologies have enabled researchers to explore DNA methylation profiles of human gametes and pre-implantation embryos at previously unrepresented levels and divulge some exciting findings (71,88,500). The fully mature human oocyte (MII stage) had a much lower level of DNA methylation than compared to sperm, but after fertilisation, both parental genomes underwent global erasure of DNA methylation that occurred at different rates in parental pronuclei, resulting in thousands of genomic regions with differential methylation (4,88,501). Surprisingly, many of the DMRs that were inherited from the oocytes (the maternal allele methylated) survived in pre-implantation stage embryos and could transiently induce monoallelic expression (20–22). Our group and others later showed that many of these oocyte-derived DMRs were maintained beyond pre-implantation stages but only in extra-embryonic tissues (20–22,449). More detailed characterisation of these placental mDMRs revealed that most were found near gene promoters, but only some could mediate monoallelic expression. More interestingly, these placental mDMRs were absent in a small fraction of placentae, revealing their polymorphic nature between individuals (295). Also, these placental mDMRs are poorly conserved in non-human mammals, with none observed in mice (20,449). Hamada and colleagues observed that

some of these placental mDMRs might have important biological functions during placental development (21). For example, *CYP2J2* (502), which encodes an arachidonic acid lipoxygenase, was shown to be upregulated in placentae affected by PE. At the same time, *CUL7* (an E3 ubiquitin ligase scaffold protein) was reported to be abnormally expressed in IUGR placentae (503), and *Cul7*-deficient mice exhibited IUGR symptoms with affected trophoblast differentiation, abnormal vasculature, and, in general, KO mice were embryonic lethal (504). Nevertheless, it still remains unclear what the role of these placental mDMRs is and if any of the genes regulated by these mDMRs could be associated with placental pathologies.

Our group previously performed a genome-wide screen of placental mDMR in WGBS datasets from human oocytes, sperm, blastocysts, placenta, and other somatic tissues, identifying 551 candidate regions (20). Unfortunately, due to low heterozygosity rates in our placental cohort, multiple candidate regions were not confirmed in the initial publication. Hence, in this PhD chapter, I revisited these previously discovered regions and compared them with candidate placental DMRs identified by two other research groups (21,22) and assessed their allelic regulation in an expanded placenta cohort. Through this screen, I identified two candidate genes with placenta-specific mDMRs, Phosphoinositide-3-Kinase Regulatory Subunit 1 (*PIK3R1*) and Go/G1 Switch Regulatory Protein 2 (*GoS2*), for which I confirmed polymorphic, maternal allele-specific methylation. In addition, I profiled their transcript-specific allelic expression and investigated whether the aberrant regulation of these genes might be associated with pregnancy complications.

## 3.2. The datasets used for identifying candidate genes with placenta-specific maternal gDMRs

All gDMR regions were identified in the human oocyte and sperm methylation tracks using a sliding window analysis described in (20) (**Section 2.7**). The regions showing opposite methylation profiles for genomic regions containing more than 25 overlapping CpG sites were determined as gDMR regions.

Placenta-specific maternal gDMRs were identified by screening the WGBS datasets of blastocysts, placenta and other somatic tissues after identifying gDMRs (in gametes) (**Section 2.7**). For this, genomic regions containing 25 CpG sites, whose average methylation  $-/+ 1.5$  SD of 25 CpGs was greater than 20% but less than 80%, were

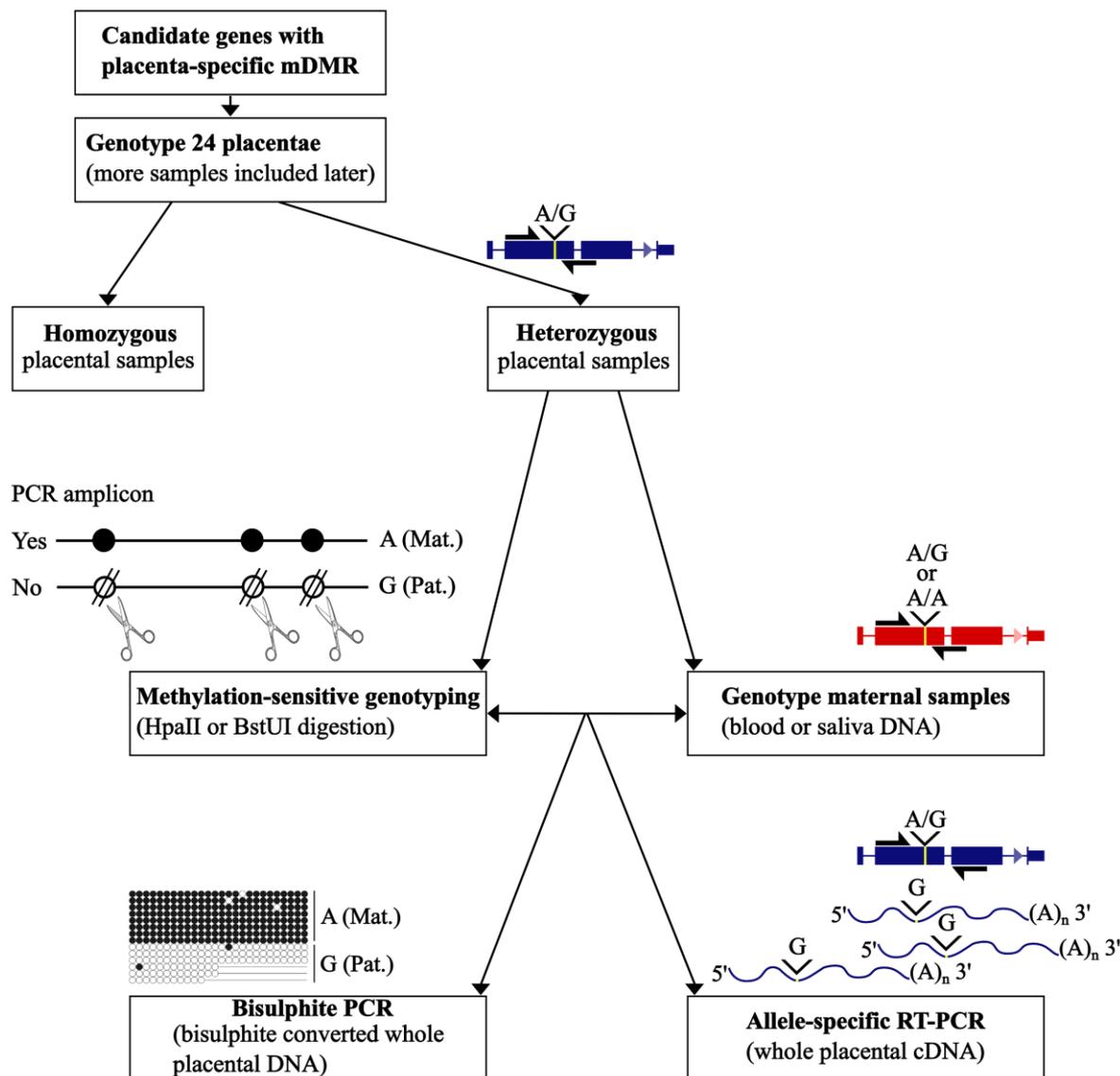
identified. The highly repetitive regions, whose coordinates were retrieved from the UCSC genome browser, were excluded from this sliding window analysis. Placenta-specific mDMRs were defined as at least 500 bp length regions that were methylated in oocytes; such regions retained partial methylation (~50% methylation) in the blastocysts and were partially methylated in the placenta but mostly not preserved in other somatic tissues. A large number of such identified placental DMRs were analysed by our group previously (20,295). The remaining placenta-specific mDMRs were compared to findings from other groups (21,22), and only uncharacterised regions were further investigated during this PhD thesis in the hope of finding novel placenta-specific imprinted genes.

### 3.3. Gene selection criteria and analysis

Placenta-specific mDMRs were further screened if they were close to gene bodies (less than 10 or 5 kb apart), if the genes were expressed in the human placenta according to the Human Protein Atlas (422), and if they had polymorphisms with a minor allele frequency (MAF) no less than 0.1.

The initial screen began with the 24 most polymorphic placental samples; however, due to a low heterozygosity rate, this was increased to 92 samples (**Figure 3.1**). To identify heterozygous samples for polymorphisms, these samples were genotyped. Heterozygous samples were then subjected to methylation-sensitive genotyping in combination with parental genotyping to determine the methylated allele. Bisulphite PCR, sub-cloning, and sequencing were carried out to determine the methylation status at the placenta-specific mDMRs. Finally, the expression of these genes was investigated by allelic RT-PCR in the most informative samples.

Identified informative genes were further investigated by using the bisulphite-converted DNA of MACS-positive cell fractions (EGFR-positive cells and stromal cells).



**Figure 3.1. Applied strategy to characterise the methylation and expression patterns of candidate genes with placenta-specific mDMRs.**

Genes or mRNAs are shown in dark blue for placental samples, while genes in red represent maternal samples. Thicker bars indicate exons, and yellow stripes within genes or mRNAs highlight polymorphic sites, with corresponding genotypes shown above. PCR primers are represented as black half arrows. Black circles represent methylated CpG sites within PCR amplicons, and white circles indicate unmethylated CpG sites.

### 3.4. Discovery of novel placenta-specific maternal DMRs (mDMRs) in the human placenta

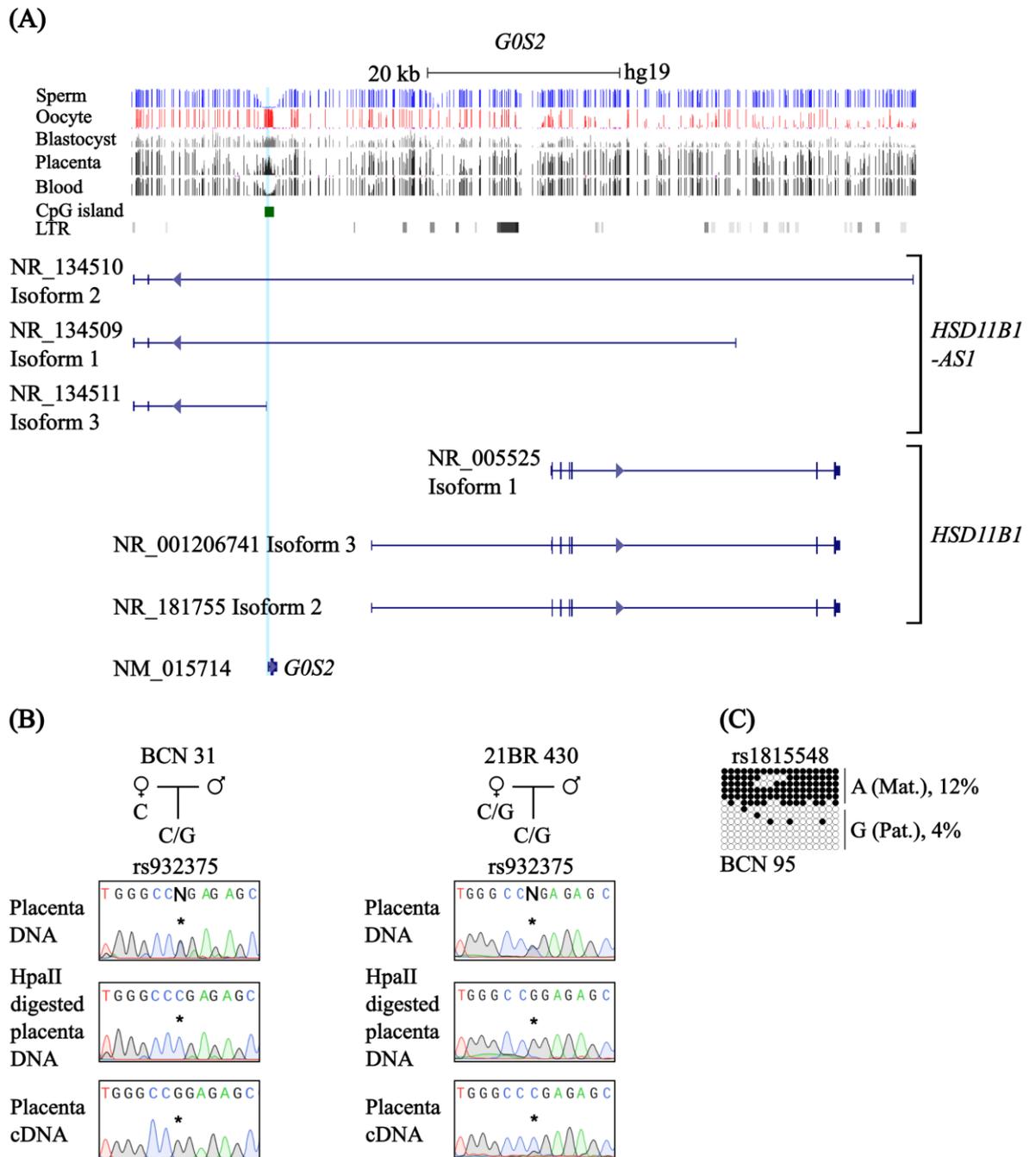
To identify candidate genes for further study, the list of placenta-specific mDMRs identified by Monk group (20) was compared to candidate genomic regions identified by others (21,22). In total, our group previously identified 551 genomic regions with oocyte-

derived mDMRs, from which 170 were found in gene promoters (**Appendix 7**). Of these, 60 mDMRs were previously characterised by our group, while 110 other mDMRs (78 mDMRs contained CpG islands with a minimum of 26 CpGs) found in promoters remained not investigated. After further inspection of these genomic intervals, 21 genes showed higher expression in the placenta ( $TPM \geq 30$ ) and could be explored by allelic RT-PCR. Among these, 9 genes contained genetic variants ( $MAF > 0.1$ ) located within placenta-specific mDMRs that could be used to assess the methylation and expression patterns of candidate genes. However, 7 of the 9 genes demonstrated little to no expression in placental trophoblasts or showed high expression in diverse immune cell types, based on scRNA-seq data from the Human Protein Atlas (422). The remaining 2 candidate genes - *GOS2* and *PIK3R1* - were expressed in placental trophoblasts (the major cell types in the placenta) (**Appendix 7**) as well as other placental cell types and were selected for more in-depth characterisation during this PhD project.

### 3.5. Allelic methylation at *GOS2* mDMR in the human placenta

The genome-wide sliding window analysis of the placental WGBS data identified a placenta-specific mDMR located on chromosome 1 between 209,847,680 and 209,849,302 bp, which overlapped with *GOS2* (chr1:209848757-209849735). According to different methyl-seq datasets, this placenta-specific mDMR was ~40% methylated in the human placenta but not in other somatic tissues such as blood (**Figure 3.2A**). It contains a large CpG island (chr1:209848444-209849428) with 76 CpG dinucleotides, of which 48 CpGs were present within the promoter and gene body of *GOS2* (“hg19 CpG Island Info”). *GOS2* encodes a small protein made of 103 amino acids and has a molecular weight of 11,321 Da (UniProt database). Intriguingly, this gene was flanked by lncRNAs that were called *HSD11B1* antisense RNA 1 (*HSD11B1-AS1*) (in the RefSeq database). Under closer inspection, *GOS2* was present in an intron shared by two larger lncRNA transcripts (NR\_134510.1 and NR\_134509.1) that possibly originate from ERVs. The promoter of NR\_134510.1 was near an LTR element (MER34) with a high SW score (the Smith-Waterman algorithm identifies local alignments between sequences). Interestingly, the *GOS2* CpG island serves as a bidirectional promoter, encompassing the TSS for the smallest *HSD11B1-AS1* transcript (NR\_134511.1). Unfortunately, a lack of informative genetic variants made it impossible to determine whether these lncRNAs are imprinted in the human placenta.

To determine whether the *GoS2* gene is imprinted in the human placenta, I screened for highly polymorphic SNPs so that alleles could be distinguished. Two variants were identified using the UCSC genome browser: rs1815548 (MAF = 0.175), located in the 5' UTR of *GoS2*, and rs932375 (MAF = 0.121), mapped within the first exon (**Figure 3.2A**). Initially, 24 most polymorphic placentae were used for the initial screening of heterozygous samples. However, due to a low rate of heterozygosity in our placental cohort, this number was increased to 92. After genotyping placenta-derived DNA, I found 13 heterozygous placentae for at least one of the SNPs. More specifically, 7 heterozygous samples for rs1815548 and 8 for rs932375, which are summarised in **Table 3.1** and **Appendix 8**. To determine if this locus exhibits allelic-methylation, I genotyped corresponding parental DNA samples and carried out methylation-sensitive genotyping, during which only the methylated allele is amplified and can generate a sequencing trace (**Sections 2.3.4 & 2.3.7.1.3**). Genotype primers were carefully designed to incorporate both the SNP and multiple HpaII methylation-sensitive restriction sites (5' - C/CGG - 3') (**Appendix 26**). Following amplification of digested DNA samples, 10 (77%) showed monoallelic methylation, of which 5 (39%) exhibited maternal allele-specific methylation for at least one of the SNPs (**Table 3.1**). In more detail, 5 showed monoallelic methylation for rs1815548, with one sample confirming maternal methylation (**Appendix 8**). Similarly, 6 samples exhibited monoallelic methylation for rs932375, of which 4 placentae had maternal methylation (**Appendix 8**). Therefore, the majority of samples demonstrated monoallelic methylation, and when informative, DNA methylation was restricted to the maternally inherited allele (**Figure 3.2B, Table 3.1**). To confirm that allelic methylation was not restricted to a few CpGs within HpaII restriction sites, I carried out bisulphite PCR and sub-cloning (**Figure 3.2C**). Characterisation of a single placental sample (BCN 95) confirmed that the maternal methylation extended for at least 21 CpGs (271 bp), with the paternal allele being largely devoid of methylation. Finally, to determine if this placenta-specific mDMR regulates monoallelic expression, I carried out allelic RT-PCR by amplifying the region with the exonic SNP (rs932375) (**Table 3.1**). 2 (25%) placentae out of 8 demonstrated monoallelic paternal expression, while the other 4 (50%) samples revealed monoallelic maternal expression (**Figure 3.2B, Table 3.1, Appendix 8**). Surprisingly, a similar finding was reported by Hamada and colleagues (21). The authors reported that *GoS2* showed maternal expression and suggested that this was likely due to residual maternal contamination in placental RNA samples, as this gene was ~400-fold more abundant in maternal peripheral blood cells.



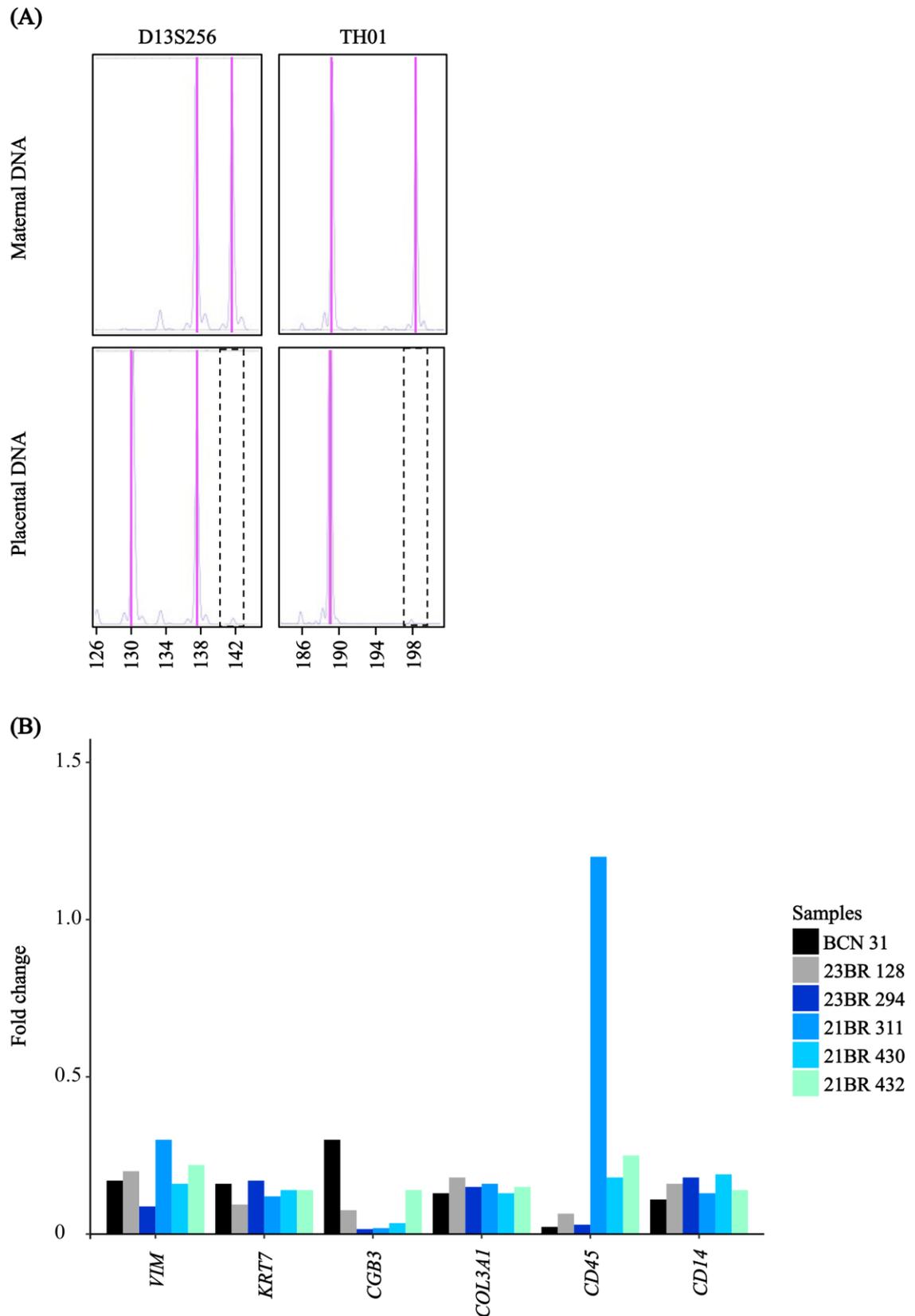
**Figure 3.2. DNA methylation and allelic expression patterns of *G0S2*.**

(A) The genomic map of the *G0S2* locus displays DNA methylation profiles from methyl-seq datasets of human sperm, oocyte, blastocyst, placenta and blood. The map also includes the antisense ncRNA *HSD11B1-ASI* (NR\_134511), with its TSS located within the *G0S2* CpG island. Gene transcripts are shown in dark blue, with thicker bars representing exons, CpG islands in dark green, and ERV LTRs in grey. Vertical lines in the methyl-seq tracks represent the mean methylation levels of individual CpG dinucleotides. (B) Two placental samples show maternal and monoallelic methylation at the *G0S2* placenta-specific mDMR, determined by methylation-sensitive genotyping. Paternal-specific expression and preferential monoallelic expression were confirmed by including the exonic SNP (rs932375) in RT-PCR products. (C) Maternal-specific methylation at the *G0S2* placenta-specific mDMR was confirmed via bisulphite PCR and sub-cloning of placenta-derived DNA. Methylated cytosines are represented by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with parent-of-origin inferred from SNP genotyping.

**Table 3.1. Result summary for the *GOS2* placenta-specific mDMR**

Total no. of heterozygous samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
13	rs1815548, rs932375	Biallelic	0	0%	Biallelic	0	0%
		Pref. monoallelic	3	23%	Pref. monoallelic	2	15%
		Monoallelic	5	38%	Monoallelic	2	15%
		Maternal	5	38%	Maternal	4	31%
		Uninformative	0	0%	Uninformative	5	38%

To confirm the presence of maternal contaminating cells in our placental samples, a Short Tandem Repeat (STR) analysis was performed. No evidence of maternal contamination was found in any of the placental samples, although it should be noted that this method would only detect contamination down to ~5% (**Figure 3.3A**). Subsequently, I performed qRT-PCR of known placental cell-type-specific marker genes, as this method has higher sensitivity. Six biomarkers were selected, including *VIM* (stromal mesenchymal marker), *KRT7* (trophoblast marker), *CGB3* (STB), *COL3A1* (fibroblast and smooth muscle marker), *CD45* (maternal haematopoietic marker) and *CD14* (HBs) (314,360,364,505). I tested the expression of these genes in 6 placental samples informative for rs932375. Two placental samples (BCN 31 and 23BR 128) with the paternal-specific expression of *GOS2* demonstrated the lowest expression of *CD45*, which is a marker of immune cells and is especially highly expressed in T-cells and monocytes (**Figure 3.3B**). In comparison, three placental samples that exhibited the maternal-specific expression of *GOS2* (23BR 294, 21BR 311 and 21BR 432) showed the highest levels of *CD45* and the lowest levels of the STB marker *CGB3*, which was especially notable in 21BR 311. These results suggested maternal contamination could account for the observed maternal expression, as all mothers were homozygous. Even a few invading maternal immune cells could contaminate placental RNA and influence *GOS2* expression.



**Figure 3.3. Investigation of residual maternal contamination in placental samples.**

**(A)** Results of Short Tandem Repeat (STR) analysis showing no maternal contamination in placental DNA. Fragment sizes are displayed beneath the plots. Pink lines indicate different fragment sizes detected in maternal and placental DNA. Dashed squares highlight the absence of fragments indicative of maternal contamination in placental DNA. **(B)** Quantitative expression profiles of various cell marker genes in placental samples informative for rs932375 located at the *GOS2* placenta-specific mDMR. Samples 23BR 294, 21BR 311, and 21BR 432 exhibited high

expression of *CD45* (a maternal haematopoietic marker), indicating residual contamination from maternal immune cells in the placental cDNA.

### 3.6. Allelic methylation at placenta-specific mDMR of *PIK3R1* isoform 3 in the human placenta

The promoter of *PIK3R1* isoform 3 contains a placenta-specific mDMR, which is 2 kb upstream of the TSS of *PIK3R1* isoform 2 (**Figure 3.4A**). The mDMR spans a ~1.3 kb interval (67,583,849 to 67,585,202 bp) that extends beyond the CpG island containing 26 CpG sites (chr5:67,584,214-67,584,451). Based on methyl-seq datasets for different human tissues, this region was exclusively methylated in the placenta (~50% methylation) and unmethylated in other somatic tissues such as blood, liver, brain, pancreas and others (data not shown). An inspection of transcripts revealed two polymorphisms, rs138814985 (indel, MAF = 0.264) and rs2888323 (SNP, MAF = 0.276), both of which map within the first exon of isoform 3 and could be used for allelic RT-PCR and methylation analysis.

In total, 34 placentae were genotyped, and 19 samples were found to be heterozygous for at least one of these polymorphisms within the placenta-specific mDMR (13 heterozygous for rs138814985 and 9 heterozygous for rs2888323) (**Table 3.2, Appendix 9**). Allelic methylation was assessed using methylation-sensitive genotyping. In total, 18 samples were informative for at least one of the variants (12 informative for rs138814985 and 8 informative for rs2888323) (**Table 3.2, Appendix 9**). Among these, 14 (78%) placentae demonstrated monoallelic methylation, of which 4 (22%) exhibited maternal allele-specific methylation (**Figure 3.4B, Table 3.2**). Specifically, 12 samples exhibited monoallelic methylation for the indel rs138814985, with 3 showing methylation of the maternal allele (**Appendix 9**). For the SNP rs2888323, sequencing results were more variable, with only 4 out of 8 informative samples showing monoallelic methylation. One of these samples was highly informative and demonstrated maternal allele-specific methylated (**Figure 3.4B, Appendix 9**). These findings were further supported by bisulphite PCR and sub-cloning, which confirmed that 26 CpGs on the maternal allele were methylated within the *PIK3R1* placenta-specific mDMR region (**Figure 3.4C**).

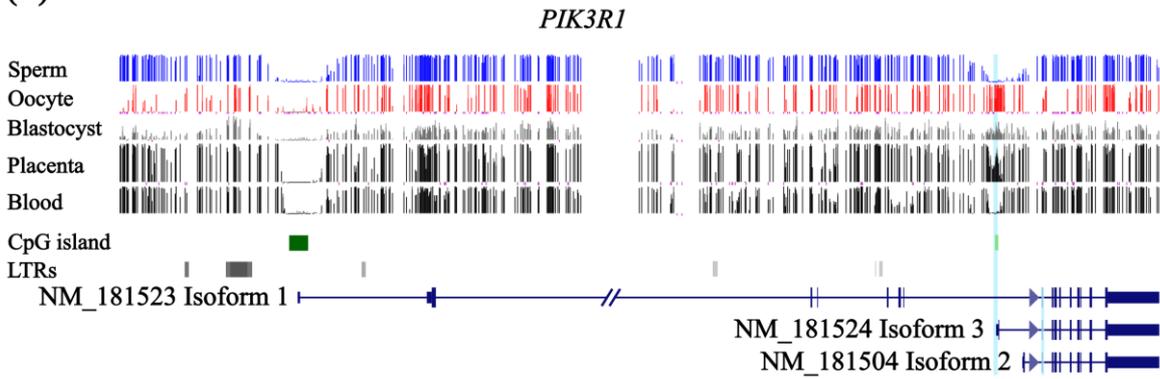
Allelic expression of isoform 3 was further investigated by RT-PCR incorporating both polymorphisms into the amplicon (**Figure 3.4D, Table 3.2, Appendix 9**). Fourteen out of 19 heterozygous samples were found to be informative for at least one polymorphism (9 informative for rs138814985 and 8 informative for rs2888323), with 6 (43%) samples

demonstrating monoallelic expression, while the remaining 8 (57%) showed expression from both parental alleles (**Table 3.2, Appendix 9**). More specifically, 4 placentae exhibited monoallelic expression for the indel rs138814985, while 5 others showed biallelic expression (**Appendix 9**). A similar observation was made for the SNP rs2888323, with 3 placentae demonstrating monoallelic expression and 5 exhibiting biallelic expression. Overall, 13 samples were informative for both allelic methylation and expression analyses, of which 6 (46%) exhibited monoallelic expression accompanied by preferential methylation of one allele (**Appendix 9**).

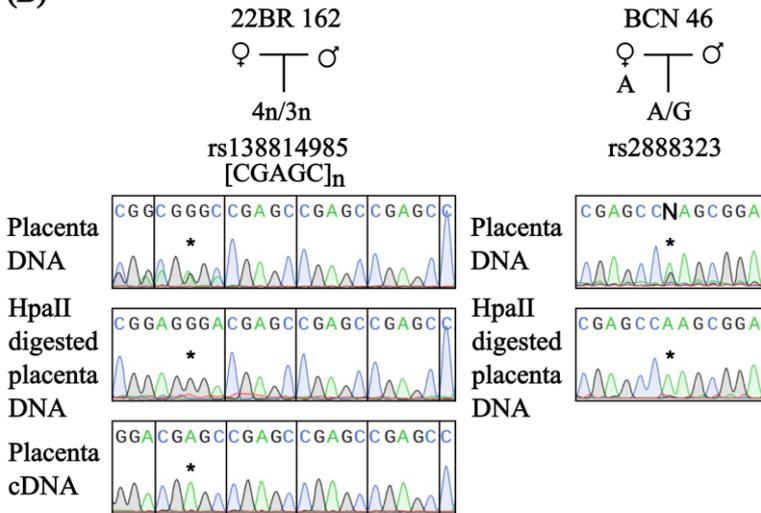
Our group and others (20–22,247) previously reported that placenta-specific imprinted genes are polymorphic (**Appendix 10**), with some individuals showing complete loss of DNA methylation at the associated placenta-specific mDMRs, which could lead to biallelic expression. To confirm this, I carried out bisulphite PCR and sub-cloning for two samples with biallelic expression for rs2888323 (BCN 6) and rs3730089 (BCN 44), both of these samples showed complete loss of methylation (**Figure 3.4D, Appendix 9**).

The *PIK3R1* gene encodes several different isoforms. *PIK3R1* isoform 1 is the major isoform and is ubiquitously expressed across different tissues. Its promoter is located ~75 kb upstream of the placenta-specific mDMR and falls within a large CpG island containing 144 CpGs (**Figure 3.4A**). Unlike isoform 3, it is unmethylated in all somatic tissues, including the placenta. The promoter of isoform 2 is ~2 kb downstream of the placenta-specific mDMR and lies within the oocyte-specific methylated region, which extends beyond the placenta-specific mDMR in gametes and becomes fully methylated after implantation. To determine if these additional isoforms are imprinted, RT-PCR was performed across SNP rs3730089, an exonic variant within the first shared exon of all three isoforms. Unfortunately, I could not determine if isoform 2 of *PIK3R1* was imprinted, as its expression was extremely low in the term placenta. Consistent with isoform 1 being transcribed from an unmethylated promoter, biallelic expression was observed in 10 informative heterozygous placentae, whereas the expression of isoform 3 was highly variable across the same placental samples (**Figure 3.4E**). Of the 10 informative heterozygous samples, only 2 (20%) showed monoallelic expression, and the other 4 (40%) demonstrated preferential expression of one allele (**Appendix 9**). Four samples informative for this SNP (rs3730089) were also informative for either the indel rs138814985 or the SNP rs2888323, and only one sample (22BR 162) demonstrated monoallelic expression of isoform 3 across both variants (**Appendix 9**). Taken together, the placenta-specific mDMR of *PIK3R1* exhibits polymorphic imprinting, which in some placental samples leads to monoallelic expression of isoform 3.

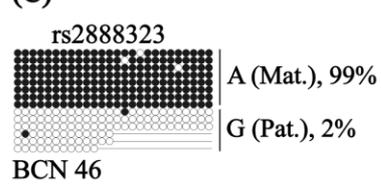
(A)



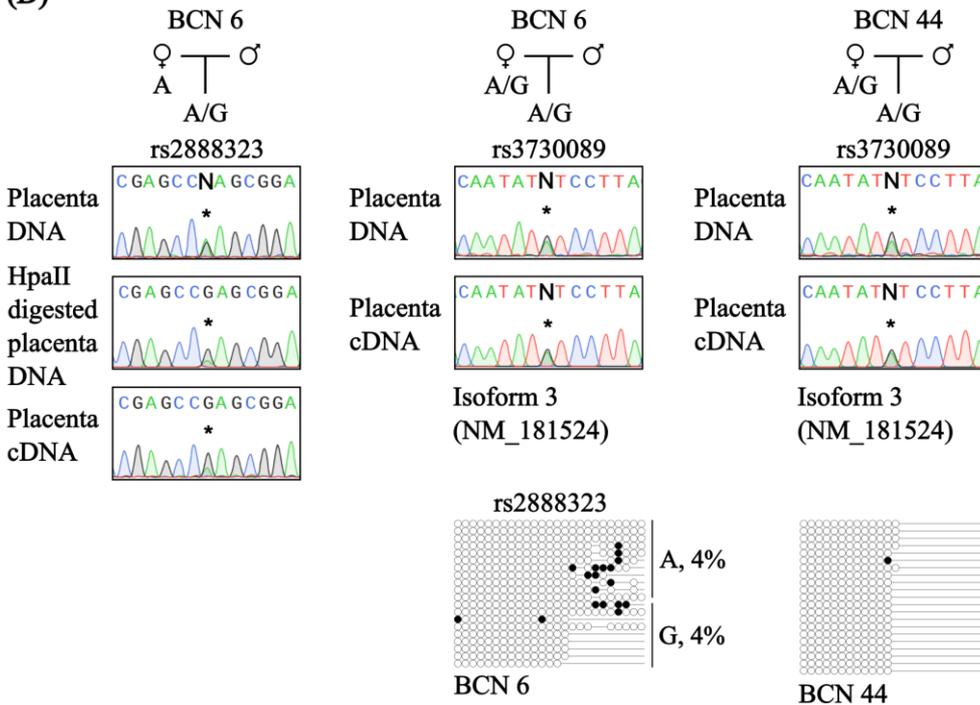
(B)

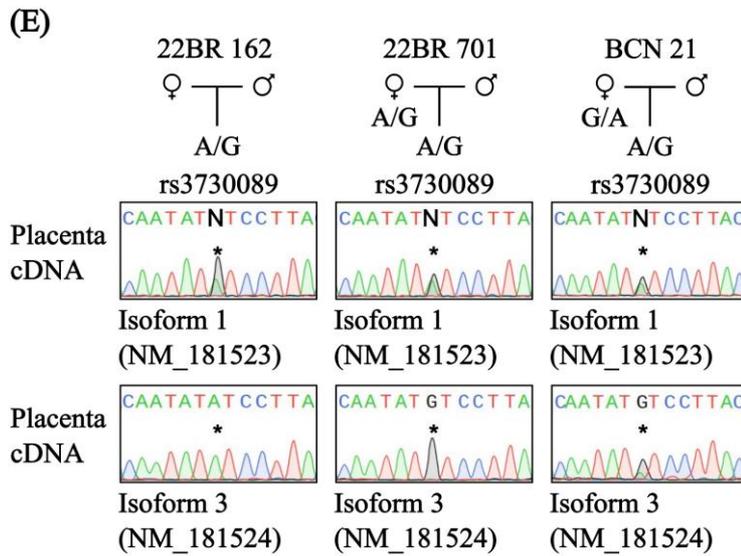


(C)



(D)





**Figure 3.4. Characterisation of DNA methylation and allelic expression for *PIK3R1* isoform 3.**

(A) The genomic map displays three *PIK3R1* isoforms with distinct TSSs. The placenta-specific mDMR of *PIK3R1* isoform 3, highlighted in light blue, exhibits hypermethylation in oocytes, no methylation in sperm, intermediate methylation in blastocysts and placental tissues, and no methylation in somatic tissue methyl-seq datasets. Gene transcripts are shown in dark blue, with thicker bars representing exons, while CpG islands are depicted as dark green bars and ERV LTRs as grey bars. Vertical lines in the methyl-seq tracks represent the mean methylation levels for individual CpG dinucleotides. (B) Two placental samples show monoallelic and maternal allele-specific methylation for the indel and SNP, determined by methylation-sensitive genotyping. Monoallelic expression was confirmed for 22BR 161 by including the indel rs138814985 in RT-PCR products. (C) Maternal-specific methylation at the placenta-specific mDMR of *PIK3R1* isoform 3 was validated using bisulphite PCR and sub-cloning of placenta-derived DNA. (D) Two placental samples (BCN 6 and BCN 44) demonstrated biallelic expression of *PIK3R1* isoform 3, possibly due to the absence of methylation at the placenta-specific mDMR, which was confirmed by bisulphite PCR and sub-cloning of placenta-derived DNA. Methylated cytosines are indicated by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with the parent-of-origin inferred from SNP genotyping if the placental sample was heterozygous. (E) Two placental samples (22BR 162 and 22BR 701) demonstrated biallelic expression of *PIK3R1* isoform 1 but monoallelic expression of *PIK3R1* isoform 3, while BCN 21 demonstrated preferential monoallelic expression of *PIK3R1* isoform 3.

**Table 3.2. Result summary of *PIK3R1* isoform 3 placenta-specific mDMR**

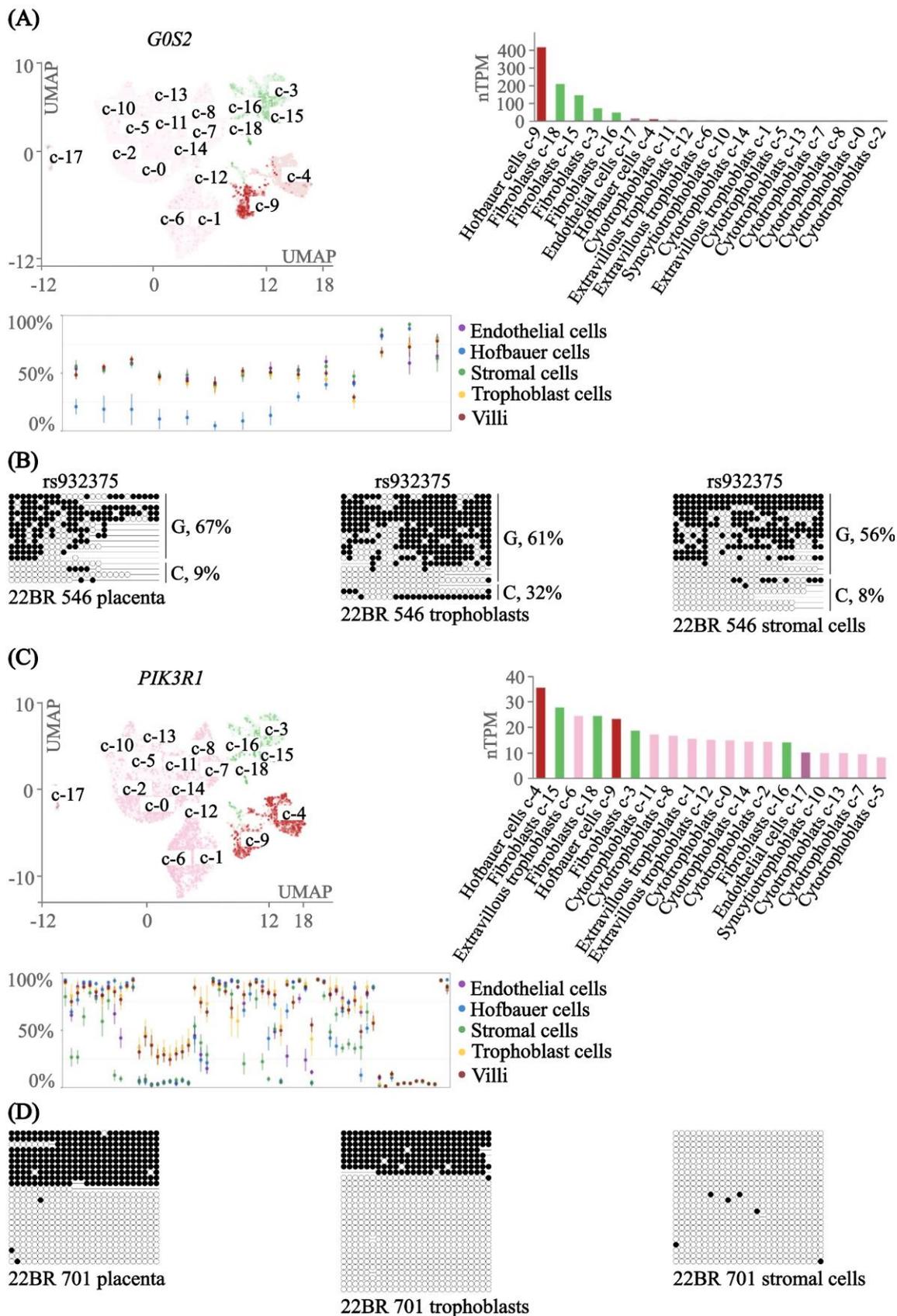
Total no. of heterozygous samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
24	rs138814985, rs2888323, rs3730089	Biallelic	2	8%	Biallelic	11	46%
		Pref. monoallelic	2	8%	Pref. monoallelic	2	8%
		Monoallelic	10	42%	Monoallelic	6	25%
		Maternal	4	17%	Maternal	1	4%
		Uninformative	6	25%	Uninformative	4	17%

### 3.7. Cell-type specific methylation of *GOS2* and *PIK3R1*

By interrogating publicly accessible scRNA-seq datasets, we observed that *GOS2* is expressed in a few placental cell populations, including fibroblasts, and, in general, was lowly expressed in placental trophoblast lineages (**Figure 3.5A**). To determine cell-type-specific methylation profiles, I interrogated the recently published Illumina Infinium MethylationEPIC array datasets for 4 placental cell types generated by Yuan and colleagues (364). This revealed that the *GOS2* placenta-specific mDMR is ~ 50% methylated in placental villi and three other cell types, including trophoblasts, stromal and endothelial cells, but hypomethylated in HBs (**Figure 3.5A, Appendix 10**). Therefore, we speculated that *GOS2* should be imprinted in placental stromal, endothelial and trophoblast cells but not in HBs. To confirm this, our group established the placental cell type enrichment protocol, which is described in detail in **Section 2.4.1**. Overall, this protocol utilises MACS for the enrichment of EGFR-positive trophoblasts and anti-fibroblast-positive stromal cells after depleting blood cells by a continuous percoll gradient. Enrichment was confirmed using qRT-PCR targeting cell-type specific biomarkers (**Appendix 11**). Extracted gDNA from the EGFR-positive and fibroblast-positive fractions of a term placental sample (22BR 546) was subject to bisulphite PCR and sub-cloning. Following genotyping, the sample was found to be heterozygous for the SNP rs932375. After mapping the sequencing results, it was observed that the G allele was preferentially methylated in the whole placental sample and trophoblast cell fraction, but this finding was most apparent in stromal cell fraction, which further supported our hypothesis (**Figure 3.5B**). Unfortunately, I could not determine the allelic expression of *GOS2* in this sample, as no placental RNA remained.

A similar observation was noted for *PIK3R1*. According to single-cell data analysis, *PIK3R1* is abundant in all placental cell types (**Figure 3.5C**). However, it was not possible to determine isoform-specific expression in different placental cell types because all scRNA-seq datasets were generated using 10 x Genomics sequencing technology, which utilises 3'-end short-read sequencing. After interrogating the Illumina Infinium MethylationEPIC array data for 4 placental cell types, I found that the placenta-specific mDMR of *PIK3R1* was ~40% methylated in placental villi and trophoblast cells but was unmethylated in other placental cell types (**Figure 3.5C, Appendix 10**). Thus, it is anticipated that *PIK3R1* isoform 3 should be imprinted in placental trophoblast but not in other placental cell types. As for *GOS2*, we utilised EGFR-positive trophoblasts and anti-fibroblasts-positive stromal cells from a single placenta (22BR 701) for bisulphite PCR and sub-cloning. Unfortunately, this sample was homozygous for both polymorphisms within

the *PIK3R1* placenta-specific mDMR, but ~40% of clones were methylated in the whole placenta, 30% of clones were methylated in placental trophoblasts, but no methylation was detected in placental stromal cells (**Figure 3.5D**). In accordance with the findings, we believe that *PIK3R1* should exhibit isoform-specific imprinting in placental trophoblasts.



**Figure 3.5. Characterisation of allelic methylation of *GOS2* and *PIK3R1* placenta-specific mDMRs in different placental cell types.**

(A) *GOS2* shows high expression in Hofbauer cells (HBs) and placental fibroblasts, according to the Human Protein Atlas (10 x Genomics scRNA-seq datasets) (422). The placenta-specific mDMR of *GOS2* exhibited intermediate methylation in three placental cell types and was

hypomethylated in HB cells, based on the Infinium MethylationEPIC array datasets (364). **(B)** DNA methylation patterns at the *GOS2* placenta-specific mDMR were observed in bulk placental samples, as well as in placental trophoblast and stromal cell fractions isolated by MACS, and confirmed by bisulphite PCR followed by sub-cloning. **(C)** *PIK3R1* expression is detected in all placental cell types, according to the Human Protein Atlas (10 x Genomics scRNA-seq datasets) (422). The placenta-specific mDMR of *PIK3R1* showed partial methylation in placental trophoblast cells, as indicated by the Infinium MethylationEPIC array datasets (364). **(D)** Cell type-specific methylation of the *PIK3R1* placenta-specific mDMR was confirmed by bisulphite PCR and sub-cloning. This mDMR showed intermediate methylation in the whole placenta and placental trophoblasts but no methylation in stromal cells. Methylated cytosines are represented by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with the parent-of-origin inferred from SNP genotyping if the placental sample was heterozygous.

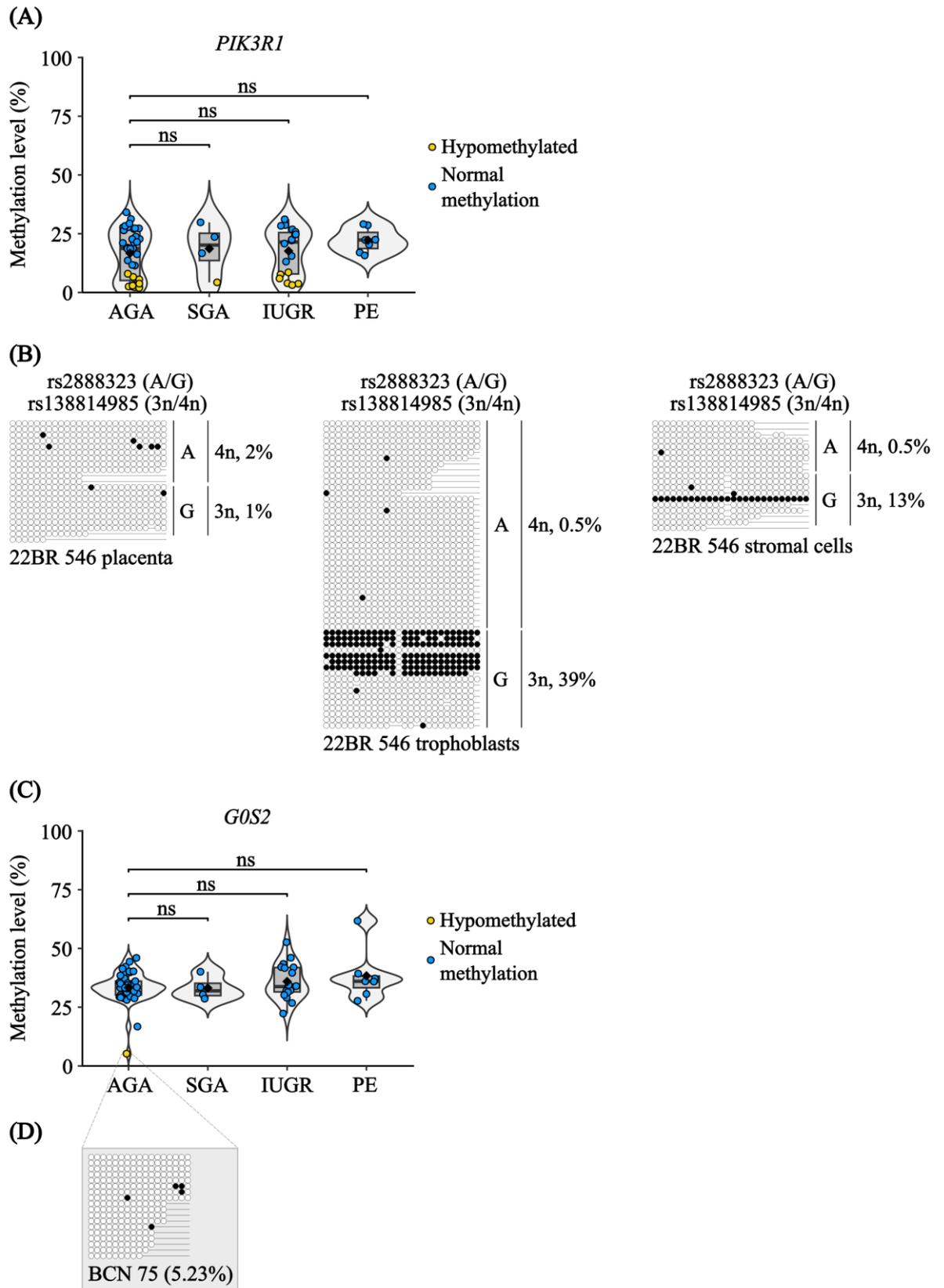
### 3.8. Polymorphic imprinting events in the placental cohort

Gong and colleagues (375) previously conducted long and short placental RNA-seq and differential gene expression analysis for 155 normal placentae, 82 placentae affected by PE and 40 samples affected by FGR that were all collected for the Pregnancy Outcome Prediction (POP) study. The authors created a Shiny app for the research community to explore this data interactively (<https://www.obgyn.cam.ac.uk/placentome/>). After investigating the results of differential gene expression analysis for this placental cohort, it was observed that *PIK3R1* and *GOS2* were weakly associated with PE. More specifically, *PIK3R1* was downregulated in PE cases (p-value = 0.016,  $\log_2(\text{Fold change}) = -0.11$ ), while *GOS2* was upregulated (p-value = 0.007,  $\log_2(\text{Fold change}) = 0.21$ ). However, after the Benjamini-Hochberg correction for multiple comparisons, both significant associations were lost. No differential expression was reported for other pregnancy complications. Thus, with this in mind, I investigated the role of *PIK3R1* and *GOS2* in pregnancy complications.

The level of DNA methylation at the *PIK3R1* and *GOS2* placenta-specific mDMRs and associated gene expression levels were determined in our placental cohort, which included samples from normal pregnancies (a baby appropriate for gestational age or AGA), PE, SGA (a baby that is small for gestational age) and IUGR cases (**Appendix 2, Appendix 3**). DNA methylation was quantified by pyrosequencing. In total, I screened 69 placentae for the placenta-specific mDMR associated with *PIK3R1*. The average DNA methylation of 5 CpG sites present within this mDMR was 16.68% (SD = 10.44%, n = 40) for AGA, 18.62% (SD = 10.96%, n = 4) for SGA, 22.24% (SD = 5.18%, n = 7) for PE and 17.62% (SD = 9.88, n = 18) for IUGR groups (**Figure 3.6A**). This was lower than anticipated but is

explainable by the high frequency of polymorphic unmethylated samples. Unfortunately, no significant changes in DNA methylation levels were detected between AGA and SGA cases, AGA and IUGR cases, and AGA and PE cases (the Wilcoxon Mann-Whitney rank-sum test, two-sided). Interestingly, I observed that 20 placental samples were hypomethylated at this genomic region, as they had less than 10% DNA methylation. This included 13 AGA samples, 1 SGA and 6 IUGR cases. The pyrosequencing results for 2 placental samples (BCN 6 (AGA) - 2.270%, BCN 44 (IUGR) - 3.926 %) were concordant with the biallelic lack of methylation as determined by bisulphite PCR and sub-cloning as well as demonstrating biallelic expression of *PIK3R1* isoform 3 (**Figure 3.4D**, **Appendix 12**). In addition, while investigating cell-type specific methylation of *PIK3R1* mDMR, I discovered one sample with a complete loss of DNA methylation in the whole placenta (**Figure 3.6B**). Surprisingly, in the trophoblast cell fraction of this sample, one allele retained some allelic methylation, while the same region was almost unmethylated in the stromal cell fraction.

Similar results were noted for the placenta-specific mDMR of *GOS2*, for which I screened 70 placentae. Overall, I found that this region had a higher DNA methylation level when compared to *PIK3R1* mDMR, as the average DNA methylation of 4 CpG sites within this mDMR was 33.34% (SD = 6.89%, n = 41) for AGA, 33.16% (SD = 5.05%, n = 4) for SGA, 38.36% (SD = 11.06%, n = 7) for PE and 35.92 (SD = 7.59%, n = 18) for IUGR groups (**Figure 3.6C**). After comparing DNA methylation levels between different groups, I found no significant changes (the Wilcoxon Mann-Whitney rank-sum test, two-sided). Only one placental sample in the AGA group showed methylation less than 10% (BCN 75 - 5.23%), indicating polymorphic imprinting was less frequent at this locus (**Figure 3.6D**, **Appendix 12**). This profile was confirmed using bisulphite PCR and sub-cloning. These results agree with previous studies (20–22,295) suggesting that genes with placenta-specific mDMRs are polymorphic in the human population.

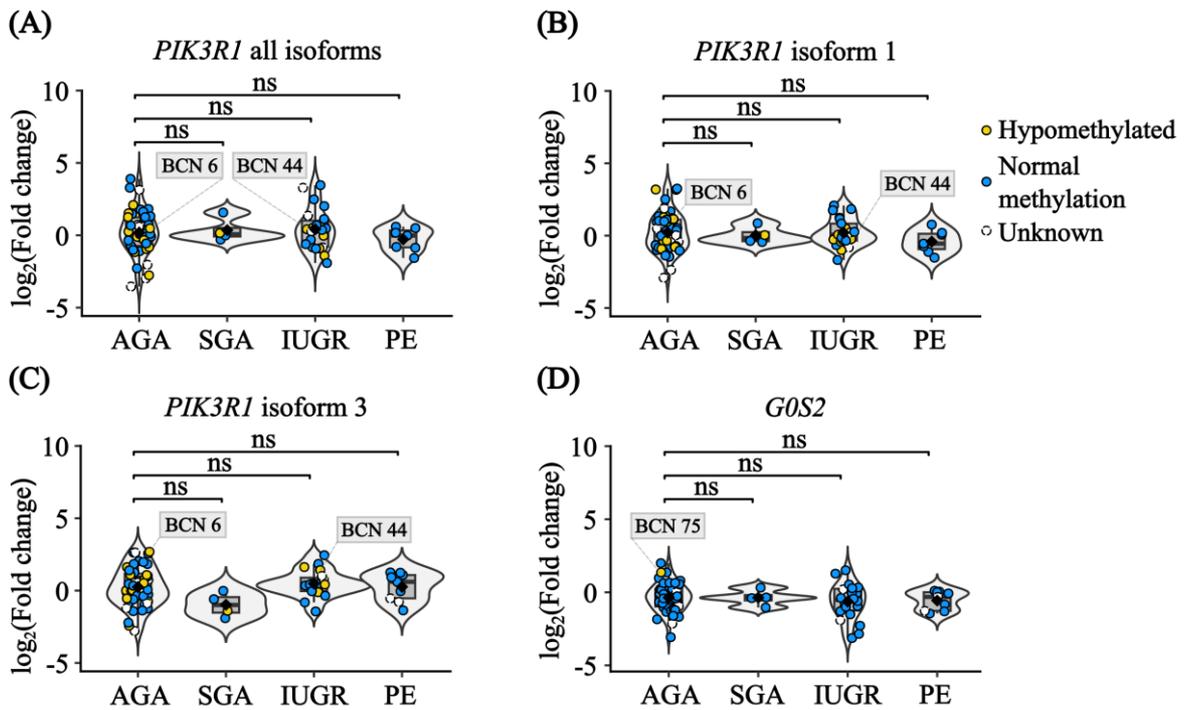


**Figure 3.6. Polymorphic placenta-specific mDMRs of *PIK3R1* isoform 3 and *G0S2*.**

(A) Quantified DNA methylation levels at the placenta-specific mDMR of *PIK3R1* across four placental groups, measured by pyrosequencing. No significant changes in DNA methylation were observed between the groups, although many placental samples showed a loss of methylation at this mDMR. (B) A placental sample showed loss of methylation in the whole placental tissue but retained some residual methylation in the placental trophoblast cell fraction, isolated using MACS. Bisulphite-converted placental or cell-type DNA was used for bisulphite PCR and sub-cloning. (C)

Quantified DNA methylation levels at the placenta-specific mDMR of *GOS2* across four placental groups measured by pyrosequencing. No significant changes in DNA methylation were observed between groups, with only one sample showing loss of methylation at this mDMR. **(D)** Low methylation in BCN 75 was confirmed by bisulphite PCR and sub-cloning of placenta-derived DNA. Methylated cytosines are represented by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with the parent-of-origin inferred from SNP genotyping if the placental sample was heterozygous. Results are shown as violin plots containing box plots, which extend from the first to third quartiles (25th to the 75th percentiles), with whiskers indicating 1.5 times the interquartile range below the first or above the third quartiles. Samples outside this range are considered outliers. Black diamonds indicate the mean, and the black line within the box represents the median (50th percentile). Each circle represents an individual placental sample, with hypomethylated samples (less than 10% DNA methylation) indicated by yellow circles. The Wilcoxon Mann-Whitney rank-sum test (two-sided) was used to compare mean methylation between groups (ns - not significant). Groups include AGA (appropriate for gestational age), SGA (small for gestational age), IUGR (intrauterine growth restriction), and PE (pre-eclampsia).

As noted earlier, I also quantified the expression of *PIK3R1* and *GOS2* in our placenta cohort by applying qRT-PCR (**Appendix 2, Appendix 3**). For *PIK3R1*, I designed primer pairs that were able to amplify regions common to all isoforms (shared by several isoforms) or were isoform-specific (**Appendix 26**). Unfortunately, I found no significant changes in mRNA levels of different *PIK3R1* isoforms when comparing the control placentae of normal pregnancies to those affected by PE, SGA or IUGR (**Figure 3.7 A, B, C**). I also observed similar expression levels of *GOS2* between groups, which resulted in no significant differences (**Figure 3.7D**).

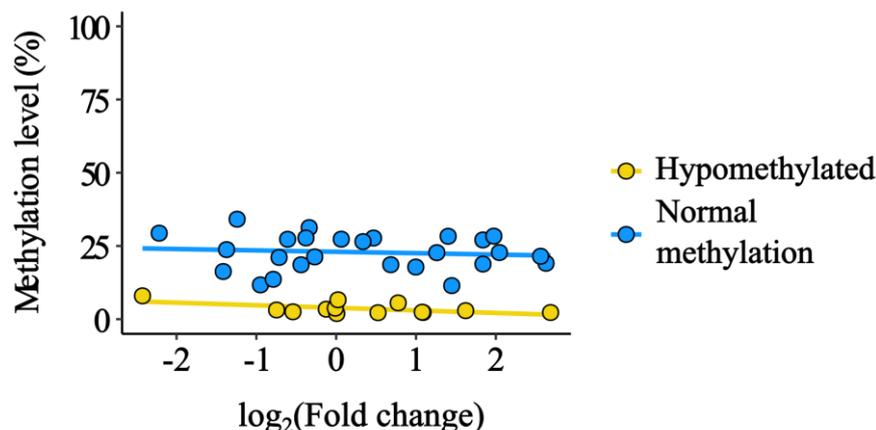


**Figure 3.7. Quantified expression of *G0S2* and two *PIK3R1* isoforms in normal and disease-affected placental samples.**

Expression of (A) all *PIK3R1* isoforms, (B) *PIK3R1* isoform 1, and (C) *PIK3R1* isoform 3, which contains the placenta-specific mDMR within its promoter. (A-C) Expression was quantified by qRT-PCR using Power SYBR™ Green Master Mix with slightly modified qRT-PCR cycling conditions (Appendix 13). (D) Expression of *G0S2* was quantified by TaqMan qRT-PCR. No significant changes in *G0S2* or *PIK3R1* isoform-specific expression were observed across different placental groups. Samples within grey boxes represent hypomethylated samples at the placenta-specific mDMRs, identified by pyrosequencing and further confirmed by bisulphite PCR and sub-cloning. Expression levels were normalized to the *RPL19* endogenous control. Results are presented as violin plots containing box plots, which extend from the first to third quartiles (25th to 75th percentiles), with whiskers indicating 1.5 times the interquartile range below the first or above the third quartiles. Samples outside this range are considered outliers. Black diamonds indicate the mean, and the black line within the box represents the median (50th percentile). Each circle represents an individual placental sample, with hypomethylated samples (less than 10% DNA methylation) indicated by yellow circles and white circles indicating placental samples with no corresponding bisulphite converted DNA. The Wilcoxon Mann-Whitney rank-sum test (two-sided) was used to compare mean methylation between groups (ns - not significant). The groups include AGA (appropriate for gestational age), SGA (small for gestational age), IUGR (intrauterine growth restriction), and PE (pre-eclampsia).

Our group previously showed that gene expression is not always associated with the level of DNA methylation at mDMRs. By exploring pyrosequencing and qRT-PCR datasets for *PIK3R1*, a weak correlation between the level of DNA methylation at the placenta-specific mDMR and the expression of *PIK3R1* isoform 3 in the AGA group was observed (Figure 3.8). In total, there were 39 placentae in the AGA group, 26 samples showed normal methylation (above 10%), and 13 placentae were hypomethylated (less than 10%

methylation at the mDMR). The Spearman's rank correlation coefficient between DNA methylation and expression of isoform 3 in the normal samples was = -0.05, and in the hypomethylated samples, it was -0.45.



**Figure 3.8. Correlation between *PIK3R1* isoform 3 expression and DNA methylation levels at the placenta-specific mDMR in normal placental samples.**

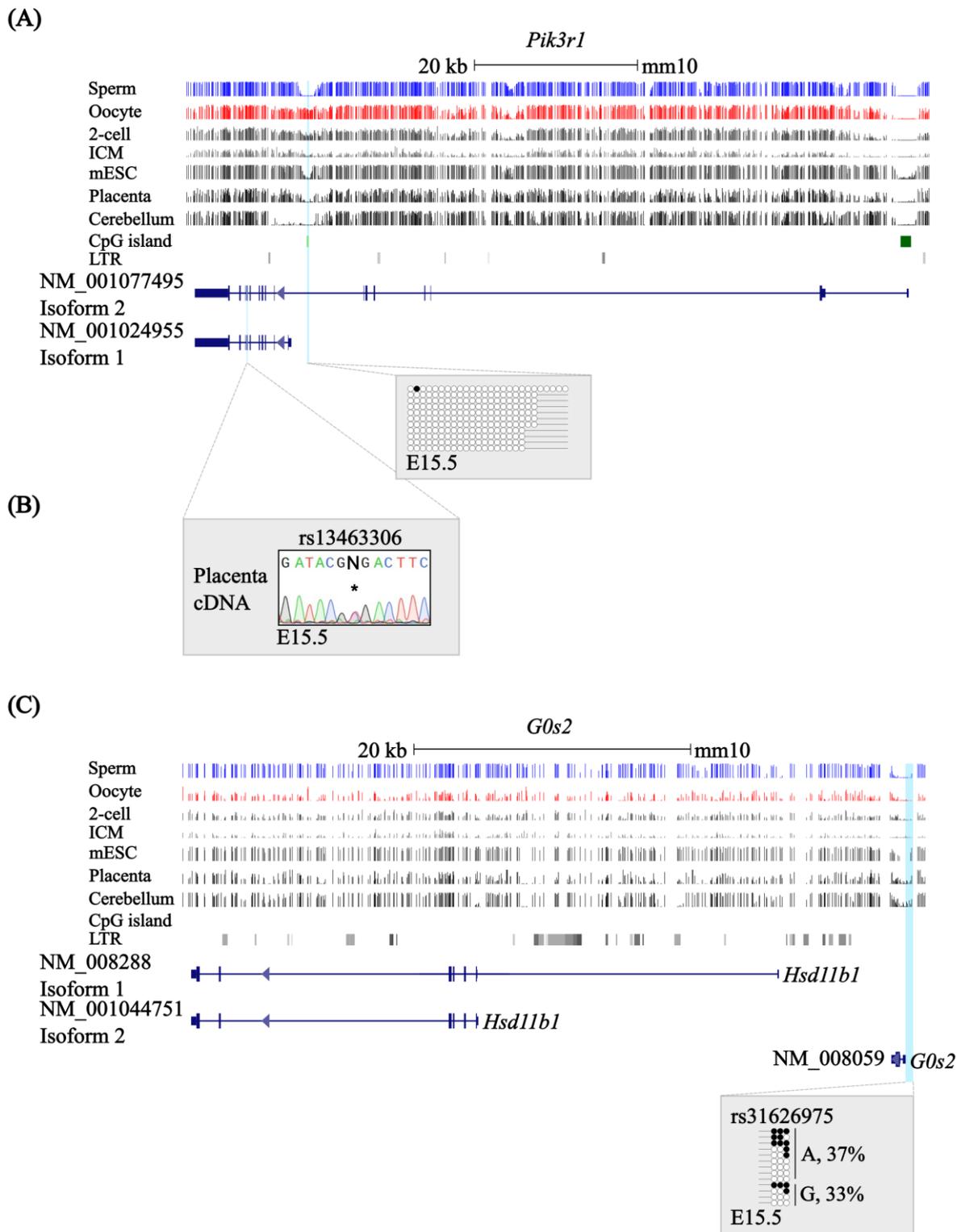
Only placental samples from the AGA (appropriate for gestational age) group with both expression and methylation data were included. Each circle represents an individual placental sample, with hypomethylated samples (less than 10% DNA methylation) shown as yellow circles. No correlation between DNA methylation levels and *PIK3R1* isoform 3 expression was observed in normal placental samples, although a weak negative correlation was noted in hypomethylated samples.

### 3.9. Placenta-specific mDMRs are not conserved in the mouse placenta

Previously, our group showed that placenta-specific imprinted genes (or mDMRs) are only confined to primates as they are not conserved in other mammalian species (20,449). To confirm that the placenta-specific mDMRs associated with *PIK3R1* and *GOS2* are not conserved in mice, I used DNA from C57BL(6) vs JF1 hybrid mouse placentae for sodium bisulphite treatment, followed by bisulphite PCR and sub-cloning (**Appendix 27**). *Pik3r1* has two isoforms in the mouse genome, and I targeted a CpG island region (30 CpGs) upstream of the promoter of *Pik3r1* isoform 1 (the equivalent location to the imprinted human orthologue), which was methylated in mouse oocytes and different stages of mouse embryos, but mostly hypomethylated in the placenta (**Figure 3.9A**). I found that this region was unmethylated in the E15.5 mouse placenta. In addition, I identified an exonic SNP, which revealed biallelic expression (**Figure 3.9B**). However, *Pik3r1* isoforms 1 and

2 were not tested separately.

In addition, I investigated the methylation at the *Gos2* promoter, which contained only a few CpG sites and was hypomethylated in mouse gametes, embryos and the placenta, according to publicly available methyl-seq datasets (**Figure 3.9C**). This genomic region demonstrated mosaic methylation across both alleles. Also, I found that this gene is not expressed in the mouse placenta, as no PCR amplicons were generated by RT-PCR. Thus, both genes are unlikely to be imprinted in the mouse placenta.



**Figure 3.9. The mouse orthologs of human *PIK3R1* and *G0S2* are not conserved in the mouse placenta.**

(A) A genomic map of the mouse *Pik3r1* gene showing DNA methylation profiles from various methyl-seq datasets. The CpG island near the smaller *Pik3r1* isoform is highlighted in light blue and is methylated in mouse oocytes, unmethylated in sperm, and shows low methylation in the inner cell mass (ICM) and placenta. The absence of DNA methylation in this region was confirmed by bisulphite PCR and sub-cloning using placental DNA from C57BL6 and JF1 hybrids. (B) Biallelic expression of *Pik3r1* was demonstrated by including an exonic SNP in RT-PCR products, with the corresponding SNP ID shown above the sequencing chromatograms. (C) A genomic map of *G0s2* showing DNA methylation profiles across different methyl-seq datasets. The promoter

region of *G0s2*, highlighted in light blue, is hypomethylated across these datasets. Biallelic methylation in this region was confirmed by bisulphite PCR and sub-cloning of placental DNA from C57BL6 and JF1 hybrids. Gene transcripts are shown in dark blue, with thicker bars indicating exons, CpG islands depicted as dark green bars, and ERV LTRs as grey bars. Vertical lines in the methyl-seq tracks represent the average methylation levels for individual CpG dinucleotides. Methylated cytosines are represented by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with the parent-of-origin inferred from SNP genotyping if the placental sample was heterozygous.

**Chapter 4: Non-canonical imprinting,  
manifesting as post-fertilisation placenta-  
specific parent-of-origin methylation, is  
not conserved in humans**

## 4.1. Introduction

Over the years, extensive research has elucidated the molecular mechanisms and functions of DNA methylation-dependent imprinting, now referred to as canonical imprinting (436). This type of regulation is associated with imprinting clusters primarily controlled by ICRs, which are decorated by DNA methylation in a parent-of-origin-specific manner (254,259,260,270). These ICRs, as noted earlier, are erased in PGCs and established during oogenesis and spermatogenesis at different developmental stages in mice (160,173,281,282). To prove that ICRs were able to induce monoallelic expression throughout clusters, ICRs were deleted (506), which often resulted in a LOI and, in some cases, embryonic lethality. Similar observations were noted in mouse mutants lacking *de novo* or maintenance methyltransferases, exhibiting not only a global reduction of DNA methylation but also a loss of most genomic imprints (507–510).

More recently, advances in next-generation sequencing technologies and the development of sophisticated molecular biology techniques have allowed us to process more samples with fewer cells at a lower cost and offer higher sequencing resolution, allowing us to identify many more imprinted genes (4,300,501,511,512). These technical advances enabled the profiling of multiple mouse and human tissues, building imprinting atlases, or “imprintomes,” which not only facilitated the discovery of many more imprinted genes but also revealed some genes that were transiently imprinted in pre-implantation embryos or demonstrated tissue-specific imprinting with a strong bias towards extra-embryonic tissues (307–309). This was later demonstrated by several groups, including ours, showing that the human placenta is enriched with imprinted genes that are polymorphic (20–22,295).

Surprisingly, these discoveries uncovered a handful of genes, such as *Gab1*, *Sfmbt2* and *Slc38a4*, that were imprinted in the mouse placenta by novel mechanisms not associated with gDMRs (299,513), and their imprinting status was maintained in mouse conceptuses without oocyte-derived DNA methylation (i.e. *DNMT3* KOs), which puzzled the research community. Nevertheless, some earlier work conducted by several groups suggested that repressive histone PTMs were associated with the *Kcnq1/Kcnq1ot1* and *Igf2r/Airn* imprinting clusters in the mouse placenta (514–516). Also, it was known that repressive histone marks could repress genes with CpG island-promoters (monoallelic bivalent chromatin domains) (188,189,197,198).

Inoue and colleagues (298) explored the regulatory landscape (promoters and enhancers) of mouse pre-implantation embryos (from 1-cell stage to morula) by applying liDNase-seq,

which revealed that the paternal chromatin was highly dynamic and by a pronuclear 3 stage (PN3; 7h after fertilisation) had similar chromatin accessibility comparable to the maternal chromatin, with an exception at imprinted loci. The same group followed up this study by utilising liDNase-seq with RNA-seq to screen mouse gametes, pre- and post-implantation embryos, and AG and GG embryos (24). They also combined these datasets with the publicly accessible data of WGBS (mouse gametes) and CHIP-seq (mouse gametes and pre-implantation embryos). They found that the mouse oocyte and, later, the maternal genome in mouse embryos harboured numerous paternal-specific DHS sites that were hypomethylated and decorated by H3K27me3, which inhibited the expression from the maternal allele. Many of the identified genes demonstrated paternal-biased expression in pre-implantation embryos and blastocysts. However, only a few genes retained paternal-specific expression in the mouse placenta (E9.5), including *Gab1*, *Phf17*, *Sfmbt2*, *Slc38a4* and *Smoc1*, some of which were previously shown to be independent of oocyte DNA methylation (299,513).

Several subsequent studies (25,26,105,187) provided further evidence that maternal H3K27me3 peaks were transient and eventually became hypermethylated, forming sDMRs in post-implantation embryos that orchestrated monoallelic expression of these loci, and most of them contained alternative promoters derived from murine-specific ERVs that drove the expression of LTR-derived transcripts. Recently, non-canonical imprinting was reported in rats, revealing some similarities and differences between the two rodent species (450). Furthermore, several studies explored the function of non-canonical imprinting during mouse development. For example, *Slc38a4* LOI resulted in placental overgrowth in mouse embryos derived from SCNT (451), while mouse pups with paternal KO of *Slc38a4* had intrauterine growth restriction and reduced placental sizes (452). In addition, non-canonical imprints were shown to be lost in SCNT embryos (451–453). Taken together, non-canonical imprints are important for proper mouse embryo and placental development.

To date, there is limited information on the conservation of non-canonical imprinting in humans, as only a few studies have attempted to address this question in human embryos and the placenta (27,136,224,517), resulting in inconclusive findings. Therefore, in this chapter, I investigated whether H3K27me3-mediated imprinting is conserved in humans. As noted earlier, in the rodent post-implantation ExE, oocyte-derived H3K27me3 is replaced by sDMRs exhibiting maternal allele-specific methylation, which in most cases overlap with rodent-specific ERVK LTR elements, leading to paternal allele-specific expression. Accordingly, I aimed to discover candidate genes harbouring placenta-specific sDMRs within their promoters (exhibiting maternal allele-specific methylation) that could potentially drive paternal-biased expression. Nevertheless, as discussed earlier, human

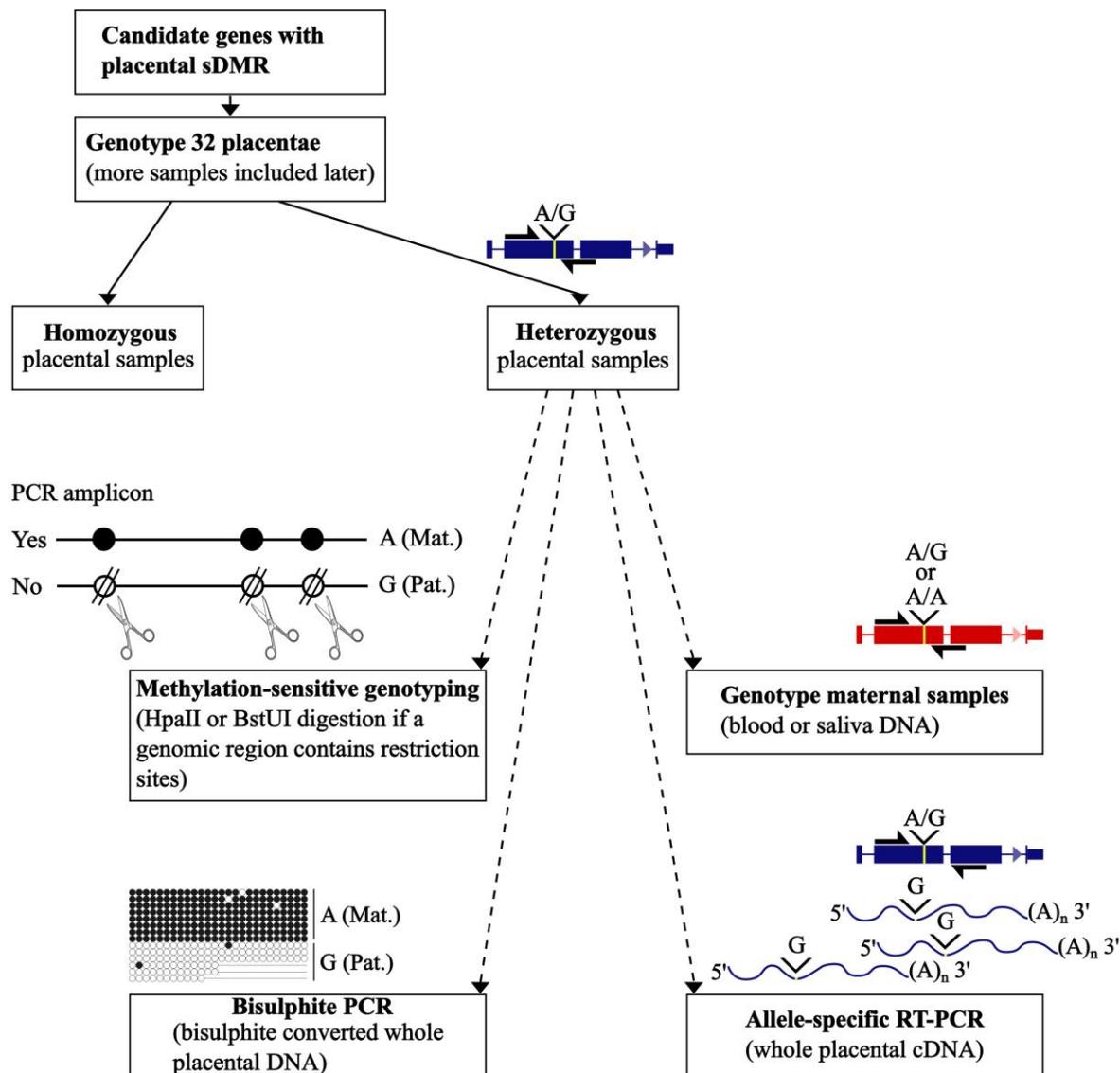
placenta-specific imprints demonstrate relaxed and polymorphic imprinting, with some individuals showing a complete lack of methylation at such genomic regions and often exhibiting biallelic expression (20–22,295). As a result, identifying and robustly validating such candidate placental imprinted genes is a challenging task. However, based on the results of the methylation-sensitive genotyping analysis from our previous study (20), genes with fewer than 48% of samples showing biallelic methylation were more likely to be imprinted in the human placenta (**Appendix 14**). Methylation-sensitive genotyping enables the efficient screening of multiple candidate genes across many placental samples, requiring only minimal gDNA and avoiding more labour-intensive methods such as bisulphite PCR followed by sub-cloning and sequencing. Nonetheless, allelic methylation and expression of promising candidate genes should be further validated using alternative and preferably more quantitative approaches to distinguish imprinted genes from genomic regions exhibiting random or allele-specific methylation due to *cis*-acting DNA variants (464,518). Therefore, in this project, I employed methylation-sensitive genotyping in combination with other techniques to interrogate candidate non-canonical imprinted genes in human embryos and the placenta. Firstly, I examined orthologs of mouse and rat non-canonical imprinted genes in the human placenta and embryos. I then explored human- and primate-specific LTR-driven transcripts associated with putative placental sDMRs, as well as genes previously proposed to be imprinted in human morulae. Additionally, I analysed the methylation and expression of *XIST*, a crucial component of the XCI process, and profiled candidate non-canonical imprinted genes with putative placenta-specific sDMRs. During this screening, I also identified several novel genes containing germline-derived mDMRs specific to the human placenta (20,24,519,25–27,187,235,373,450,517). Overall, I found no evidence of non-canonical sDMRs residing in the human term placenta, suggesting that imprinting in this human embryonic organ is regulated exclusively by gDMRs.

## 4.2. Selection of human candidate genes for non-canonical imprinting

For the screening of candidate genes, the same methyl-seq datasets were used as described in earlier **Section 2.7**. Candidate genes with placenta-specific sDMRs were identified by performing sliding window analysis in the placental WBS dataset as described by Court *et al.* (2014) (449). Genomic regions were considered as placental sDMR regions if the average methylation was  $25\% < \text{mean of } 25 \text{ CpGs } \pm 1.5 \text{ SD} < 75\%$  and these regions were unmethylated in the oocyte, sperm and blastocyst methyl-seq

datasets. Only those placenta sDMR regions were further analysed if they were close to gene bodies (less than 10 or 5 kb apart). These genes were expressed in the human placenta, according to the Human Protein Atlas (422), and finally had polymorphisms with a MAF of at least 0.1. During this analysis, genes showing signs of a maternal gDMR were also considered for screening. The promoters of such genes demonstrated methylation in the oocyte, no methylation in sperm, partial methylation in the blastocyst and mostly no methylation in somatic tissues.

To determine if candidate-selected genes were imprinted, a similar strategy was employed, as described in **Section 3.3**. Firstly, I genotyped the 32 most polymorphic placentae to identify heterozygous samples by genotyping PCR (**Figure 4.1, Appendix 28**). Then, methylation-sensitive genotyping was applied to placental sDMRs and genes with LTR promoters or genes with placenta-specific mDMRs to identify the methylated allele of a previously identified polymorphism. The human orthologs of mouse and rat non-canonical imprints were screened by methylation-sensitive genotyping to determine if their promoters contained restriction sites. For very informative placental samples, corresponding parental DNA was also genotyped if available. Bisulphite PCR, cloning and sequencing followed to determine the methylation status at placental DMR regions or gene promoters with CpG-rich regions. Finally, the most informative heterozygous samples for genes with placental sDMRs, genes with LTR promoters or genes with placenta-specific mDMRs were tested by allelic RT-PCR, while all human orthologs of mouse and rat non-canonical imprints were tested by this method.



**Figure 4.1. Applied strategy to investigate the methylation and expression patterns of candidate genes with placental sDMRs.**

Genes or mRNAs are shown in dark blue for placental samples, while genes in red represent maternal samples. Thicker bars indicate exons, and yellow stripes within genes or mRNAs highlight polymorphic sites, with corresponding genotypes shown above. PCR primers are represented as black half arrows. Black circles represent methylated CpG sites within PCR amplicons, and white circles indicate unmethylated CpG sites. Dashed arrows indicate optional techniques used to explore candidate genes.

### 4.3. Human orthologs of mouse non-canonical imprints are not conserved in the human placentae

Inoue and colleagues (24) provided strong evidence that H3K27me3 covers large genomic

regions in the mouse oocyte, which are passed on to the zygote and influence monoallelic expression until blastocyst implantation. They also showed that the depletion of these H3K27me3 domains from the maternal chromosome resulted in the biallelic expression of several loci, whereas under normal conditions, these genes were paternally expressed. Interestingly, H3K27me3-dependent imprinting was maintained at a few loci in extra-embryonic tissues, including *Gab1*, *Phf17*, *Sfmbts* and *Slc38a4*. Further work demonstrated that the oocyte-derived H3K27me3 is gradually replaced by DNA methylation (by E6.5) in extra-embryonic tissues, where such hypermethylated regions serve as maternal sDMRs in post-implantation embryos (26). In contrast, the paternal alleles remain active because they acquire H3K4me3, which repels DNA methylation. Another study revealed that these H3K27me3-dependent imprinted regions are proximal to active ERVK LTRs that gain H3K4me3 and function as alternative promoters or enhancers for paternal alleles in extra-embryonic tissues (25). Deletions of such retroviral elements have been shown to disrupt the imprinted gene expression.

To investigate whether non-canonical imprinting is conserved in the human placenta, I evaluated DNA methylation levels at the promoters of human orthologs of mouse non-canonical imprints and investigated their expression. These genes included *Sfmbt2*, *Jade1* (or *Phf17*), *Smoc1*, *Gab1*, *Slc38a4* and *Sall1* (**Table 4.1, Appendix 15**). Although *Platr20* and *Gm32885* were also reported to be imprinted in the mouse placenta, these genes do not have human orthologs. Firstly, I examined the methylation profiles of human orthologs using bisulphite PCR and sub-cloning, followed by allelic RT-PCR if the gene had highly informative exonic SNPs (**Figure 4.1, Appendix 28**).

**Table 4.1. The list of mouse non-canonical imprinted genes and their status in humans**

Gene	Mouse		Rat		Human	
	Allelic expression	Allelic methylation	Allelic expression	Allelic methylation	Allelic expression	Allelic methylation
<i>Jadel</i> (24,25,187)	Paternal	mat sDMR	nd*	nd*	Biallelic	Unmethylated
<i>Gab1</i> (24,25,299)	Paternal	mat sDMR	Paternal	mat sDMR	Biallelic	Long isoform - unmethylated; small isoform - biallelic methylation
<i>Sfmbt2</i> (24,187,520)	Paternal	mat sDMR	Paternal	mat sDMR	Biallelic	Unmethylated
<i>Slc38a4</i> (24,25,187)	Paternal	mat gDMR	Paternal	mat sDMR	Biallelic	Long isoform - biallelic methylation; small isoform - unmethylated
<i>Smoc1</i> (24,187)	Paternal	mat sDMR	nd*	nd*	Biallelic	Unmethylated
<i>Sall1</i> (24,187,450)	Paternal	mat sDMR	Paternal	mat sDMR	Biallelic	Biallelic / mosaic

\*(nd) – not determined

### 4.3.1. *SFMBT2*

In the mouse genome, *Sfmbt2* encodes several isoforms, and the promoter region contains three proximal CpG islands (48, 18 and 39 CpGs), which were unmethylated in gametes, pre-implantation embryos and the placenta (**Appendix 15A**). Interestingly, an LTR element, ~5 kb upstream of this promoter, mediated paternal allele-specific expression (25). This LTR was hypomethylated in pre-implantation embryos but gained methylation in E6.5 embryos and other somatic tissues except for the placenta, where it showed around 35% methylation, illustrating the presence of sDMR. As it was reported previously, the sDMR for this locus was established relatively late in extra-embryonic tissues (300).

*SFMBT2* contains a large CpG island (418 CpGs) within its promoter, which contains the TSS for multiple isoforms. Based on methyl-seq datasets, this region was hypomethylated in gametes, blastocysts and other somatic tissues, including the human placenta, which was confirmed in a term placental sample (BCN 8) (**Figure 4.2A**). A single EST (a human-expressed sequence tag in GenBank) revealed a rarely used alternative promoter

that originated from an LTR repeat (~5 kb upstream THE1C) in a similar location as the mouse imprinted TSS. This region was also investigated using bisulphite PCR, which revealed that both alleles were methylated (**Figure 4.2A**). I also identified a SNP at the 3' UTR that was shared by several isoforms. Four informative placentae demonstrated biallelic expression (**Figure 4.2A, Table 4.2, Appendix 16**).

### 4.3.2. *JADE1*

In the mouse genome, *Jade1*, also known as *Phf17*, encodes a few isoforms that are expressed from a hypomethylated CpG-rich promoter (comprising two CpG islands: 52 and 153 CpGs) (**Appendix 15B**). It was reported that this gene had a rodent-specific LTR upstream of the major promoter (521), but it showed opposite expression relative to *Jade1*. This LTR was slightly methylated in mouse oocytes and unmethylated in pre-implantation embryos. It showed 40% methylation in the placenta, where the sDMR became established (300), leading to paternal-biased expression in mouse extra-embryonic tissues (24–26).

In humans, *JADE1* encodes several isoforms that are mostly different at their 3' UTRs (**Figure 4.2B**). All these isoforms share the same promoter region that contains 2 CpG islands: the smaller one contains 34 CpGs, and the bigger one includes 174 CpG sites. This promoter was unmethylated in human gametes, blastocysts and other tissues, including the term placenta, which I confirmed by bisulphite PCR and sub-cloning. Two SNPs were identified in the alternative 3' UTR associated with the longer or shorter isoforms. For each SNP, 4 heterozygous samples revealed biallelic expression after inspecting sequencing traces of allelic RT-PCR (**Figure 4.2B, Table 4.2, Appendix 16**).

### 4.3.3. *SMOC1*

In mice, the promoter of *Smoc1* is located within a CpG island (52 CpGs), which, according to methyl-seq datasets, was unmethylated in gametes, pre-implantation embryos, and the placenta (**Appendix 15C**). Upstream this promoter, there was an LTR element, which was slightly methylated in the mouse oocyte and the placenta but hypomethylated in sperm (25,521). In post-implantation embryos and somatic tissues, this region became hypermethylated. Additionally, this LTR was shown to mediate paternal-specific

expression (521).

*SMOC1* is transcribed from a CpG island promoter, which includes 63 CpG sites, and it was unmethylated in all methyl-seq datasets I inspected (**Figure 4.2C**). This promoter was hypomethylated in the term placenta (BCN 26) after sequencing cloned bisulphite PCR products. To test the expression of this gene, I performed allelic RT-PCR for 4 heterozygous placentae for an exonic variant located in the 12th exon of this gene, and all samples demonstrated biallelic expression (**Figure 4.2C, Table 4.2, Appendix 16**).

#### 4.3.4. *GAB1*

*Gab1* encodes several isoforms in the mouse genome that are transcribed from the CpG island promoter (110 CpGs), which was unmethylated in all methyl-seq datasets (**Appendix 15D**). It was reported that an LTR element located within the first intron of this gene could drive paternal-specific expression of a chimeric transcript (25,521). The promoter of this transcript was lowly methylated in the oocyte, unmethylated in sperm and pre-implantation embryos, but showed around 40% methylation in the placenta, with mixed methylation in other somatic tissues (**Appendix 15D**).

*GAB1* has a few isoforms in the human genome, and like in the mouse, the promoter of the shorter isoform was present within the first intron of the largest isoform (**Figure 4.2D**). Similar to the mouse genome, the promoter of the major isoform contained a large CpG island promoter with 161 CpGs and was unmethylated in human gametes, blastocysts, other somatic tissues and the placenta, which I confirmed by cloned bisulphite PCR (BCN 6) (**Figure 4.2D**). On the other hand, the promoter of the smaller isoform was mostly methylated in human gametes and some somatic tissues and slightly methylated in blastocysts and the placenta. I found that placenta 21BR 309 was informative for rs62337524, with both parental alleles being mostly methylated and associated with biallelic expression (**Figure 4.2D, Table 4.2, Appendix 16**). This is not surprising as, unlike in the mouse genome, there is no functional LTR element reported near this promoter that could drive the expression of a smaller AK295684 isoform. Finally, I found a SNP that mapped within the shared 3' UTR, revealing biallelic expression for 4 samples (**Figure 4.2D, Table 4.2, Appendix 16**).

### 4.3.5. *SLC38A4*

*Slc38a4* has several isoforms that are expressed from different TSSs (**Appendix 15E**). The longest isoform of this gene is expressed from gDMR, which contains a CpG island (36 CpGs) that was methylated in the oocyte and somatic tissues but unmethylated in sperm, with lower methylation observed in the ICM and placenta. Bogutz *et al.* (2019) (236) reported an LTR (MT2A) element located ~15 kb upstream of this gDMR, which induced the expression of a chimeric transcript in the mouse oocytes. After deletion of this LTR with CRISPR-Cas9, imprinted expression was lost. This LTR element was hypomethylated in the oocyte and pre-implantation embryos but methylated in sperm and other somatic tissues, including the placenta. Hanna *et al.* (2019) (25) reported another LTR element (MLTR31F\_Mm), located ~100 kb upstream of the gDMR, which promoted the expression of a ncRNA with paternal-specific expression.

According to the ENSEMBL and GenBank databases, *SLC38A4* encodes several isoforms, with one less annotated isoform being expressed from the CpG island (64 CpGs), which demonstrated around 60% methylation exclusively in the term placenta and suggested the presence of the sDMR (**Figure 4.2E**). To determine if this methylation was allelic, I performed methylation-sensitive genotyping as this interval contains 4 HpaII restriction sites and 2 informative SNPs (rs4994910 and rs74851348). In total, I found 13 heterozygous samples across both SNPs, with the majority showing methylation on both alleles (**Figure 4.2E, Table 4.2, Appendix 16**). I further confirmed this by cloning the bisulphite PCR product and sequencing it. In term placental sample 21BR 19, both alleles were hypermethylated at this sDMR (**Figure 4.2E**). The TSS for most isoforms was ~1 kb downstream of this methylated interval, and this region was hypomethylated in most methyl-seq datasets, which I confirmed in a placental sample (BCN 8). I also found an exonic SNP (rs2429467) that is common to most isoforms, which was informative in 3 samples that all exhibited biallelic expression (**Figure 4.2E, Table 4.2, Appendix 16**).

### 4.3.6. *SALL1*

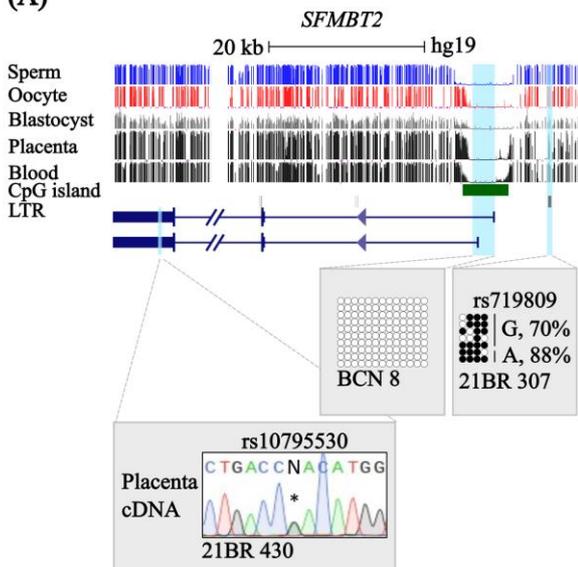
In the mouse, *Sall1* has a few isoforms that all share the same CpG island promoter (2 CpG islands in proximity: 261 and 32 CpGs), which acts as a bidirectional promoter as *Gm3134* is expressed from the opposite strand (not shown, **Appendix 15F**). This region was unmethylated in gametes, embryos, and somatic tissues. It has been reported that an LTR upstream of the *Sall1* promoter was responsible for the paternal-specific expression of

ncRNA in extra-embryonic tissues (25,521).

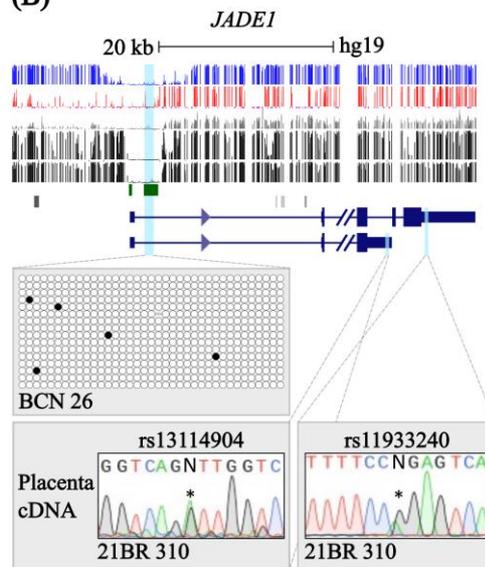
According to the NCBI RefSeq database, *SALL1* encoded 2 isoforms in the human genome (**Figure 4.2F**). Both isoforms utilise unique TSS but originate from the same genomic region densely populated with CpG dinucleotides (a larger CpG island contained 365 CpGs and a smaller included 47 CpG, respectively). This region was mostly unmethylated in all methyl-seq datasets I investigated, with slightly higher methylation levels in the placenta, which I investigated further by bisulphite PCR and sub-cloning. Two placenta samples revealed mosaic methylation, with one sample being informative for both parental alleles (**Figure 4.2F**). Therefore, I have not investigated this region further. I also found an exonic SNP (rs11645288) that showed biallelic expression in 4 placentae (**Table 4.2, Appendix 16**).

Taken together, I concluded that mouse non-canonical imprints were not conserved in the human placenta. However, as shown in the mouse studies (24,511), the majority of non-canonical imprints exist at pre-implantation stages, and only a few are retained in the mouse placenta. Therefore, it is possible that these genes could be imprinted in human pre-implantation embryos.

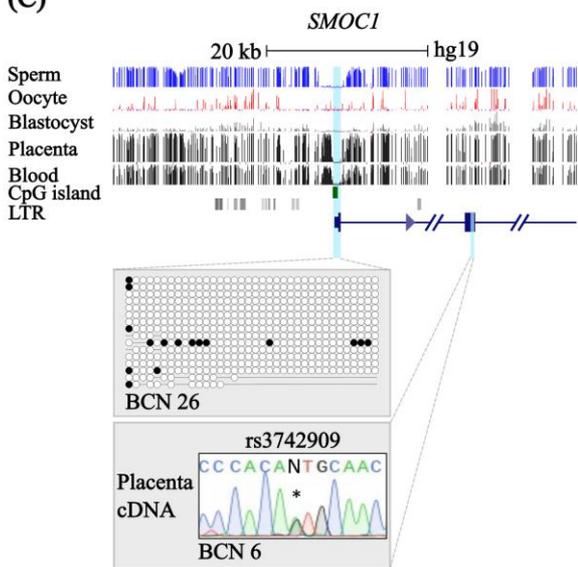
(A)



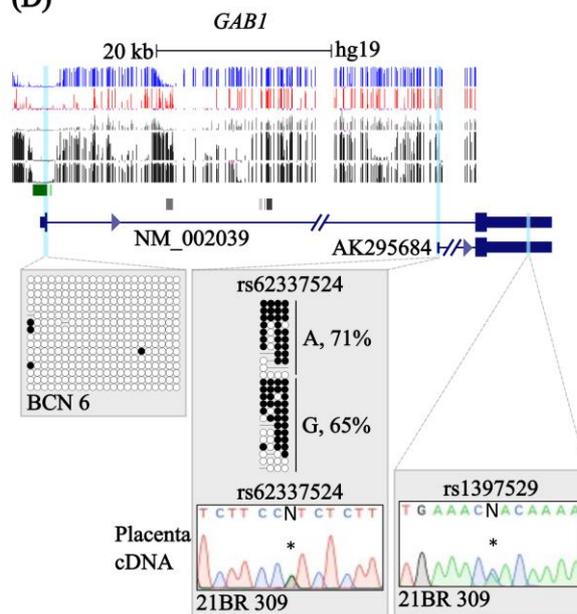
(B)

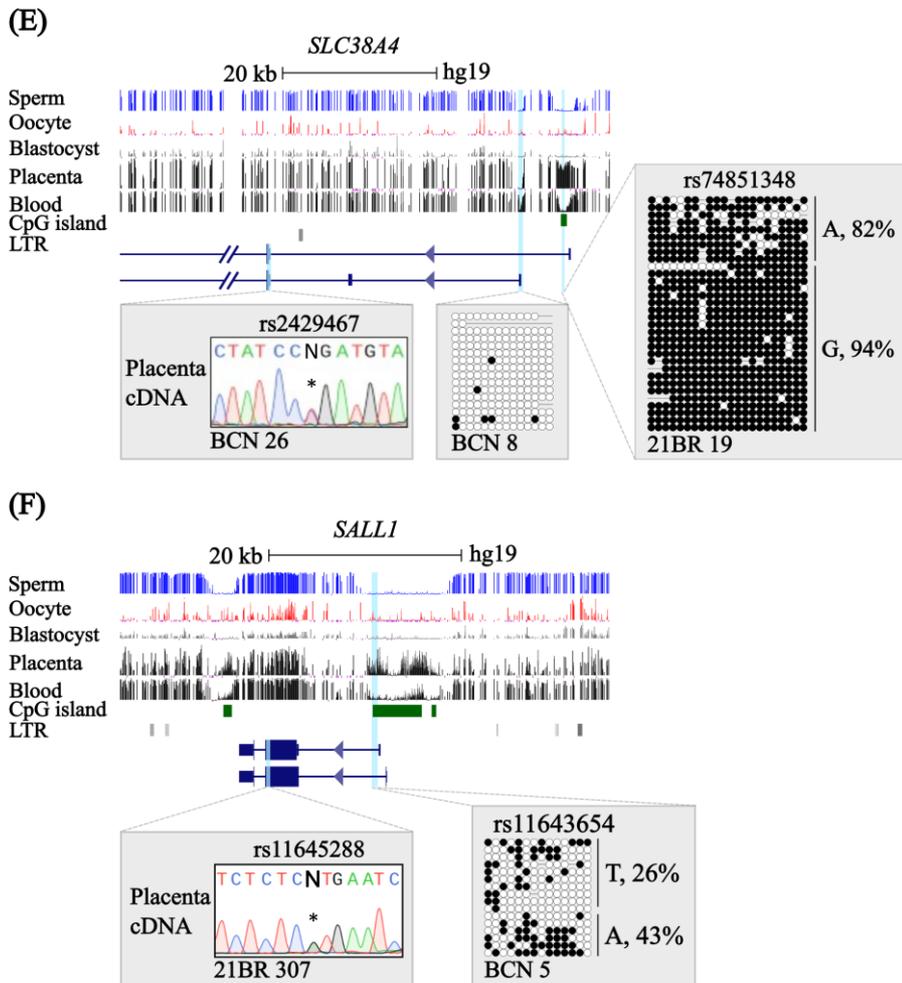


(C)



(D)





**Figure 4.2. Characterisation of DNA methylation and allelic expression for the human orthologs of mouse non-canonical imprinted genes.**

Genomic maps display the human orthologs of mouse placental sDMRs (highlighted in light blue) for the following genes: (A) *SFMBT2*, (B) *JADE1*, (C) *SMOC1*, (D) *GABI*, (E) *SLC38A4* and (F) *SALL1*. Gene transcripts are shown in dark blue, with thicker bars representing exons, while CpG islands are depicted as dark green bars and ERV LTRs as grey bars. For each gene, DNA methylation profiles from methyl-seq datasets of human sperm, oocyte, blastocyst, placenta, and blood are presented. Vertical lines in the methyl-seq tracks represent the mean methylation levels for individual CpG dinucleotides. Promoter methylation was confirmed using bisulphite PCR and sub-cloning of placental DNA. Methylated cytosines are indicated by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with the parent-of-origin inferred from SNP genotyping if the placental sample was heterozygous. Allelic expression was assessed by including SNPs (highlighted in light blue) within RT-PCR products, with SNP IDs shown above the corresponding sequencing chromatograms.

**Table 4.2. Result summary for the human orthologues of mouse and rat non-canonical imprinted genes**

Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
4	<i>JADE1/PHF17</i>	NM_001287441, NM_024900	4	rs13114904	Biallelic	-	-	Biallelic	4	100%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%
					Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%
		NM_001287437	4	rs11933240	Biallelic	-	-	Biallelic	4	100%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%
					Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%
4	<i>GAB1</i>	All isoforms	4	rs1397529	Biallelic	-	-	Biallelic	4	100%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%
					Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%
		AK295684	4	rs62337524	Biallelic	-	-	Biallelic	1	25%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%
					Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	3	75%
10	<i>SFMBT2</i>	-	5	rs10795530	Biallelic	-	-	Biallelic	3	60%
					Pref. monoallelic	-	-	Pref. monoallelic	1	20%
					Monoallelic	-	-	Monoallelic	0	0%

Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)		Allelic expression			
12	<i>SLC38A1</i>	-	6	rs1045278	Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	1	20%
					Biallelic	-	-	Biallelic	6	100%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%
					Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%
12	<i>SLC38A4</i>	-	14	rs4994910, rs74851348, rs2429467	Biallelic	5	36%	Biallelic	3	21%
					Pref. monoallelic	6	43%	Pref. monoallelic	0	0%
					Monoallelic	1	7%	Monoallelic	0	0%
					Maternal	0	0%	Maternal	0	0%
					Uninformative	2	14%	Uninformative	11	79%
14	<i>SMOC1</i>	-	4	rs3742909	Biallelic	-	-	Biallelic	4	100%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%
					Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%
16	<i>SALL1</i>	-	4	rs11645288	Biallelic	-	-	Biallelic	4	100%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%
					Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%
18	<i>ZNF516</i>	-	1	rs690353	Biallelic	-	-	Biallelic	1	100%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%

Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)		Allelic expression			
					Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%
					Biallelic	-	-	Biallelic	2	100%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%
20	<i>ZFP64</i>	-	2	rs3746413	Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%
					Biallelic	-	-	Biallelic	2	100%
					Pref. monoallelic	-	-	Pref. monoallelic	0	0%
X	<i>XIST</i>	-	2	rs1894271	Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%

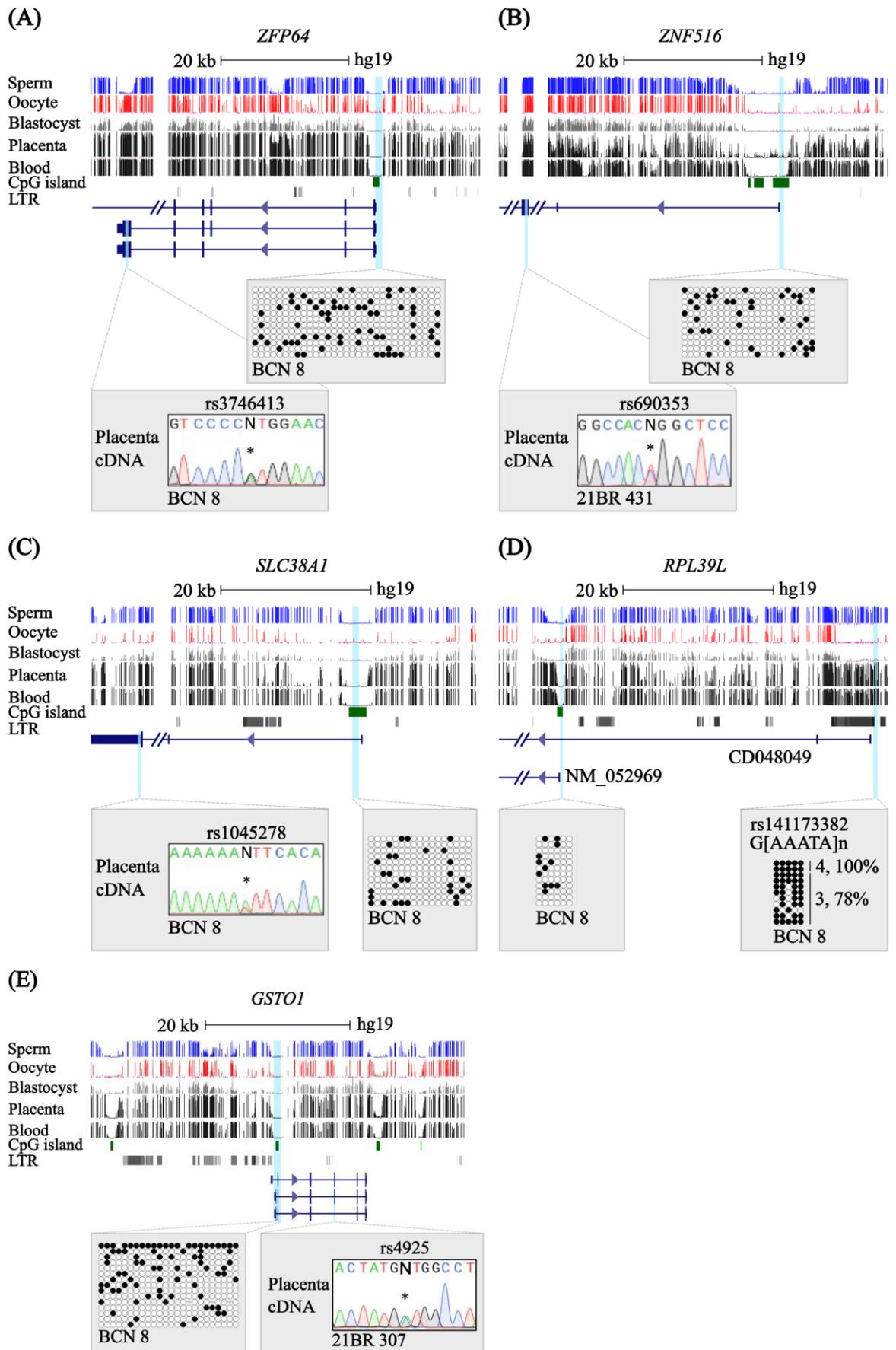
## 4.4. Human orthologs of rat non-canonical imprints are not conserved in the human placenta

The rat is a widely used model organism for biomedical studies, as it has a more similar physiology to humans than the mouse (522). Therefore, it is frequently used to model different human diseases. However, in imprinting studies, the rat is not widely used. Before the discovery of non-canonical imprinting, only 12 genes were known to be imprinted in the rat genome, and these genes were identified by comparing their homology to mouse and human canonical imprints (520,523). None of the known genes were unique to this rodent species. Therefore, Albert *et al.* (2023) (450) conducted an unbiased screening to discover novel canonical and non-canonical imprinted genes in the rat genome. For this study, the researchers used 3 genetically distinct rat strains, including BN/NCrIcrlj, WKY/NCrIcrlj and F344/NSlc, and two mouse strains, including C57BL/6N and JF1/Ms, to generate F1 progeny from different rat and mouse crosses. The hybrid embryos of F1 progeny were used to collect Epi and ectoplacental cone (EPC); for rat embryos, this was performed at E8.5, while for mice, it was performed at E7.25. The collected Epi and EPC were used for WGBS along with strand-specific RNA-seq. Throughout the study, various publicly accessible datasets, including CUT&RUN, ChIP-seq, WGBS and RNA-seq from either rat gametes, pre-implantation embryos or somatic tissues, were utilised. Similar datasets of different mouse samples were also investigated for thorough comparisons between the rat and mouse genomes. By applying a multiomics approach, the authors identified 45 genes demonstrating monoallelic expression (paternal and maternal), 18 of these were canonical and present in known imprinted clusters such as *H19/Igf2*, *Trpm5/Tssc4/Alsc2/Cd81*, and *Peg10/Sgce*. The other genes demonstrated monoallelic paternal expression exclusively in rat EPCs and included *Sfmbt2*, *Gab1* and *Sall1* and 8 novel rat-specific imprints with paternal allele-specific expression, such as *Zfp516*, *Slc38a1*, *Zfp64*, *Gsto1*, *Rpl39l*, *Syt16AS*, *Gadl1-3'UTR*, and *LOC108350526*. They also found an additional 33 maternally expressed genes in the EPCs of the rat, but these genes were highly expressed in adult blood, suggesting maternal contamination (similar to human *GOS2*); thus, these genes were not investigated further. By screening methylation datasets, they identified 45 gDMRs derived from rat gametes, and some of these DMRs overlapped known imprinting clusters with CpG island promoters or intragenic CpG islands, while others were near rat-specific imprinted genes, including *Zfp516*, *Zfp64*, *Syt16-AS*.

During this study (450), they uncovered multiple interesting differences and similarities between the mouse and rat genomes. They identified 18 genes following canonical or DNA methylation-dependent imprinting and 11 non-canonical imprinted genes (*Smoc1* and

*Jade1* did not contain informative SNPs) established by maternal H3K27me3, which in EPCs became replaced by sDMRs. Eight non-canonical imprints were exclusive to rat EPCs (at least 3 non-canonical imprints conserved between mouse and rat and 8 being exclusive to the rat genome). Overall, this study demonstrated that although the mouse and rat diverged ~13 million years ago (524), canonical imprinted genes are highly evolutionarily conserved between rodent species, as most of these genes are also conserved in humans. However, less conservation is observed in extra-embryonic tissues, where species-specific imprinted genes are much more prevalent (20,372,523).

To screen rat non-canonical imprints, I used the same approach as for the mouse non-canonical imprints. Initially, I targeted the *Zfp64*, *Zfp516*, *Slc38a1*, *Rpl39l* and *Gsto1* genes, which had orthologous regions in the human genome (**Table 4.3**). To characterise the methylation status at promoters of these genes, I performed bisulphite PCR and sub-cloning (**Figure 4.1**). I found that the promoters of these 5 genes had mosaic methylation (**Figure 4.3**). Also, after closer inspection of *RPL39L* in the UCSC genome browser, I found that it had 3 isoforms (GENCODE), and the promoter of the largest isoform originated from within a large LTR cluster, including 5 LTRs, all from the ERV1 family (**Figure 4.3D**). Interestingly, this isoform was not recorded in the RefSeq database. Thus, this suggested the expression of a chimeric LTR-derived transcript. According to the methyl-seq datasets, the promoter containing the LTRs was only unmethylated in the human oocytes and blastocysts but hypermethylated in other samples. I confirmed this by amplifying and cloning this region in a term placenta sample (BCN 8), which revealed biallelic methylation as it was heterozygous for the copy number variant (indel) (**Figure 4.3D**). Thus, it is unlikely to be a sDMR in the human term placenta. Finally, I performed allelic RT-PCR followed by Sanger sequencing for *ZFP64*, *ZNF516*, *SLC38A1* and *GSTO1*, demonstrating biallelic expression (**Figure 4.3A-C, E; Table 4.2, Appendix 16**). Unfortunately, no heterozygous placental samples could be found for *RPL39L* in our placental cohort. Based on these observations, it is unlikely that rat non-canonical imprints are imprinted in the human term placenta, which is unsurprising as these genes were not conserved in mouse EPCs (450). It remains to be determined if they are transiently imprinted in human pre-implantation embryos.



**Figure 4.3. Characterisation of DNA methylation and allelic expression for the human orthologs of rat non-canonical imprinted genes.**

Genomic maps display the human orthologs of rat placental sDMRs (highlighted in light blue) for the following genes: (A) *ZFP64*, (B) *ZNF516*, (C) *SLC38A1*, and (E) *GSTO1*. Gene transcripts are

shown in dark blue, with thicker bars representing exons. CpG islands are represented as dark green bars, and ERV LTRs are represented as grey bars. For each gene, DNA methylation profiles from methyl-seq datasets of human sperm, oocyte, blastocyst, placenta, and blood are shown. Vertical lines in the methyl-seq tracks represent the mean methylation levels of individual CpG dinucleotides. Promoter methylation was confirmed using bisulphite PCR and sub-cloning of placental DNA. Methylated cytosines are indicated by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with the parent-of-origin inferred through SNP genotyping when the placental sample was heterozygous. Allelic expression was determined by including SNPs (highlighted in light blue) in RT-PCR products, with SNP IDs displayed above the corresponding sequencing chromatograms. **(D)** A genomic map shows the DNA methylation profile from methyl-seq datasets for the main TSS and the LTR-derived promoter of the human *RPL19L* loci. DNA methylation at both promoters was tested in the same placental sample using bisulphite PCR and sub-cloning.

**Table 4.3. The list of rat non-canonical imprinted genes and their status in humans**

Gene	Rat		Human	
	Allelic expression	Allelic methylation	Allelic expression	Allelic methylation
<i>Rpl39l</i> (450)	Paternal	mat sDMR	-	LTR promoter methylated; major promoter unmethylated
<i>Slc38a1</i> (450)	Paternal	mat sDMR	Biallelic	Unmethylated
<i>Zfp516</i> (450)	Paternal	mat sDMR	Biallelic	Unmethylated
<i>Zfp64</i> (450)	Paternal	mat gDMR	Biallelic	Unmethylated
<i>Gstol</i> (450)	Paternal	nd*	Biallelic	Unmethylated

\*(nd) – not determined

## 4.5. Mouse and rat non-canonical imprinted genes demonstrate biallelic expression in human pre-implantation embryos

To determine whether rat and mouse non-canonical imprints could orchestrate allelic expression at the earlier stages of human development (24,26,450,525). I took advantage of 15 human IVF embryos that were processed into individual cells, summarised in **Table 4.4**. In total, 187 single cells were used for scM&T-seq, but only the transcriptome datasets were used during this PhD thesis (**Sections 2.2.2, 2.6 & 2.9.2**). Raw reads of

single cells were trimmed to remove sequencing adapters and short reads, and reads of poor quality were excluded from further analysis. Processed reads were aligned with STAR v2.7.10a to the human GRCh38.p14 reference genome.

**Table 4.4. Human pre-implantation embryos collected for scM&T-seq and used for single-cell transcriptome sequencing**

Couple	Number of embryos that survived	Day	Reported morphology	Survival rate	Cell types	Stage	Number of cells reported per embryo	Number of cells sequenced (scRNA-seq)	Embryo name	Duplicated reads (%)	GC (%)	Total sequences (millions)	Unique reads (millions)	Unique reads per embryo (millions)
1	1	5	Compacted	100%	Blastomere	Day 3	12c	13	1_12a	70.90%	49%	248.7	72.4	72.4
2	1	3	10c 5%	100%	Blastomere	Day 3	10c	10	2_10a	76.50%	48%	212.4	49.9	49.9
			7c 20%	100%	Blastomere	Day 3	7c	7	3_7a	74.40%	50%	174.3	44.5	44.5
			5c 20%	100%	Blastomere	Day 3	5c	5	3_5a	77.00%	49%	109.5	25.2	25.2
3	4	3	10c 10%: blastocyst 4CB	100%	TE	Blastocyst	7c BL(ICM) + 41c TE	41	3_TeA	78.50%	50%	988.2	212.0	248.7
			10c 10%: blastocyst 4CB	100%	ICM		7c BL(ICM) + 41c TE	7	3_Bla	46.40%	50%	159.2	85.3	
			9c 10%: blastocyst 4CD	100%	ICM	Blastocyst	8c BL(ICM)	8	3_Blb	55.10%	50%	168.3	75.5	75.5
4	2	6	4BB	90%	ICM	Blastocyst	11c of 25	11	4_BLa	62.80%	49%	104.8	39.0	100.8
			4BB	90%	TE		15c of 25	15	4_TeA	72.10%	50%	264.6	73.9	

Couple	Number of embryos that survived	Day	Reported morphology	Survival rate	Cell types	Stage	Number of cells reported per embryo	Number of cells sequenced (scRNA-seq)	Embryo name	Duplicated reads (%)	GC (%)	Total sequences (millions)	Unique reads (millions)	Unique reads per embryo (millions)
			4CB	70% collapsed blastocyst	ICM + TE	Blastocyst	43c total	26 cells of 43	4_TEb	64.00%	50%	282.2	101.5	101.5
6	1	3	10c 10%	90%	Blastomere	Day 3	9c	9	6_9a	80.20%	48%	250.9	49.6	49.6
7	2	3	4c 20%	50%	Blastomere	Day 2	2c	2	7_2a	54.10%	52%	68.8	31.6	31.6
			8c 20%	100%	Blastomere	Day 3	8c	8	7_8a	73.20%	48%	145.3	39.0	39.0
8	2	3	7c 10%	100%	Blastomere	Day 3	7c	7	8_9a	69.90%	48%	119.5	35.9	35.9
			8c 15%	100%	Blastomere	Day 3	8c	8	8_7a	71.80%	49%	179.9	50.8	50.8
9	2	3	5c 20%	100%	Blastomere	Day 3	5c	5	9_5a	70.30%	49%	103	30.6	30.6
			5c 20%	100%	Blastomere	Day 3	5c	5	9_5b	72.10%	50%	98.1	27.4	27.4
								Total cells sequenced: 187	Total unique embryos: 15	Average: 216.3			Average: 65.6	
								Total cells: 204						

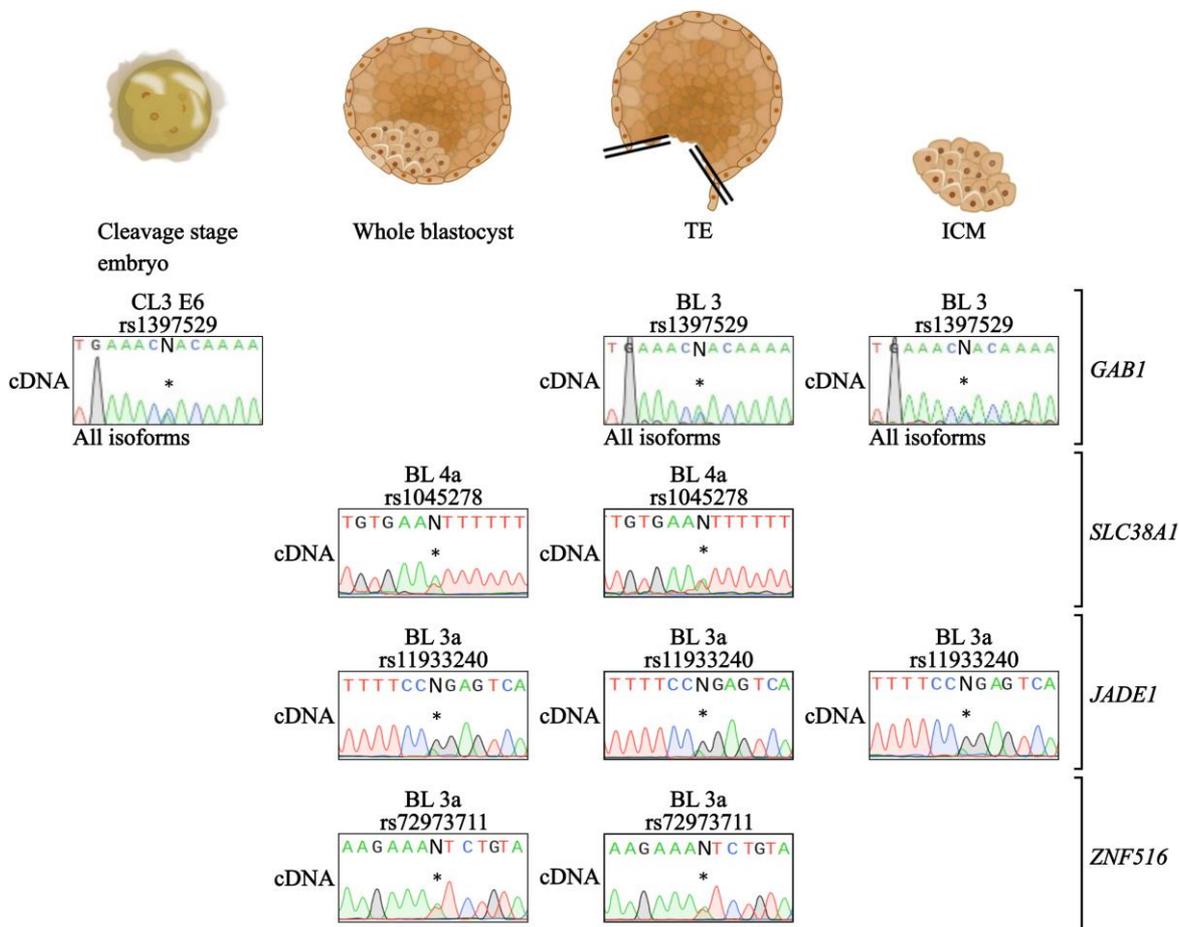
The read mapping rate was highly variable across single cells, with an average of 5.613 million unique mapped reads (the average of total mapped reads was 7.039 million) (**Appendix 17**). Deeper inspection of these datasets revealed a very high duplication rate, which is common with RNA-seq data (526) (**Appendix 17**). I further decided to generate pseudo bulk RNA-seq embryo samples by combining individual cells derived from the same embryos to increase the total coverage across the transcriptome. In total, I generated 19 pseudo bulk samples for whole embryos and separately for ICM and TE of the blastocysts where possible (**Table 4.4; Section 2.9.2**).

The average mapping rate across the pseudo bulk samples was around 65 million unique mapped reads (~ 216.3 million total mapped reads), ranging from 25 to 248 million, which was highly correlated with the number of single cells used to generate a pseudo bulk sample (**Section 2.9.2**). The pseudo bulk samples were subsequently used for germline variant calling, for which I employed the GATK tool kit (**Section 2.9.2**). This method is widely used for variant calling in whole-genome sequencing (WGS) and whole-exome sequencing (WES) datasets and has more recently been applied to bulk RNA-seq and scRNA-seq datasets.

On average, I identified 148,081 SNPs across the whole embryo pseudo bulk samples, of which 92,693 (average) were unique annotated SNPs with rsID from dbSNP and 46,630 (average) were unique novel SNPs without associated rsID (**Appendix 19**). Surprisingly, after further inspection of SNPs across different genomic features, I found that many SNPs mapped to intronic regions (**Appendix 18, Appendix 19**). In addition, many reads of pseudo bulk samples mapped to intergenic and intronic regions. This could be due to the enrichment of pre-mRNA (527) that contains introns as well as multiple spurious transcripts (528) that are known to be expressed during early pre-implantation stages.

In the absence of gDNA to generate individual embryo genotypes, I used the GATK toolkit to screen for biallelic expression directly. This revealed 35,295 biallelically expressed variants across 15 whole embryo pseudo bulk samples (**Appendix 19, Table 4.4**). This was further refined to include only those SNPs that were located within coding regions of the genome (UTRs and exons), that were detected in at least 2 single cells, that were associated rsID and were covered by 10 reads with a similar distribution for the reference and alternative allele genotypes. Finally, I focused on those SNPs that were present in *JADE1*, *SLC38A1*, *SLC38A4*, *SMOC1*, *GAB1*, *SFMBT2*, *ZNF516*, *ZNF64*, *GSTO1*, and *SALL1*. In total, 27 candidate SNPs were identified, of which 12 were selected for allelic RT-PCR on left-over amplified embryonic cDNAs produced during SMART-seq2 preparation. The amplification and Sanger sequencing of individual embryo pseudo bulk cDNA samples (whole embryo, ICM, or TE) confirmed 10 of the 12 biallelic SNPs (**Figure**

**4.4, Table 4.5).** Biallelic expression was observed for *GAB1*, *SLC238A1*, *JADE1* and *ZNF516*. In summary, our results showed that the human orthologs of mouse and rat non-canonical imprinted genes are not conserved in human pre-implantation embryos.



**Figure 4.4. Allelic expression of human orthologs of mouse and rat non-canonical imprinted genes in human pre-implantation embryos.**

Human embryos include day 3 CL embryos (8-cell stage), whole blastocysts, or blastocysts surgically separated into the inner cell mass (ICM) and trophectoderm (TE). The generated RT-PCR products contain SNPs, with the corresponding SNP IDs shown above the sequencing chromatograms.

**Table 4.5. Summary of allelic expression of the orthologs of mouse and rat non-canonical imprinted genes in human pre-implantation embryos**

Chr.	Gene	SNP	Total no. of informative samples	Embryo stage	Allelic expression
4	<i>JADE1</i>	rs11933240	2	Blastocyst	1 (TE) - biallelic, 1 (ICM) - biallelic, 1 (whole embryo) - biallelic
4	<i>GABI</i>	rs1397529	7	CL 2	1 - biallelic
				CL 3	2 - biallelic
				Blastocyst	1 (TE) - biallelic, 2 (ICM) - biallelic
				CL 2	1 - biallelic
	<i>GABI</i>	rs1360288278		CL 2	1 - biallelic
	<i>GABI</i>	rs28924077		Morula	1 - biallelic
12	<i>SLC38A1</i>	rs1045278	3	Blastocyst	1 (TE) - biallelic, 2 (whole blastocysts) - biallelic, 1 (whole blastocyst) - pref. monoallelic
				Blastocyst	1 (whole blastocyst) - biallelic
				Blastocyst	1 (whole blastocyst) - biallelic
				Blastocyst	1 (whole blastocyst) - pref. monoallelic
18	<i>ZNF516</i>	rs72973711	1	Blastocyst	1 (TE) - biallelic, 1 (whole embryo) - biallelic
		rs2074488845		Blastocyst	1 (TE) - pref. monoallelic, 1 (whole embryo) - pref. monoallelic

## 4.6. X chromosome demonstrates random inactivation in the human placenta and embryos

Different mammalian species employed diverse mechanisms to achieve a balanced dosage of X-coupled genes between males and females (303,304). In mice, XCI is imprinted as the Xp becomes preferentially silenced in extra-embryonic lineages in female mice. This is initiated by *Xist* transcription from the Xp. In mouse Epi, the inactive Xp becomes reactivated, and both X chromosomes co-exist in an active state for a brief moment while one X is randomly inactivated.

After discovering non-canonical imprinting, Inoue and colleagues hypothesised that the imprinting of *Xist* could be controlled by PRC2 established repressive mark (105). To explore this possibility, the authors examined publicly available HiDNase-seq, WBS and

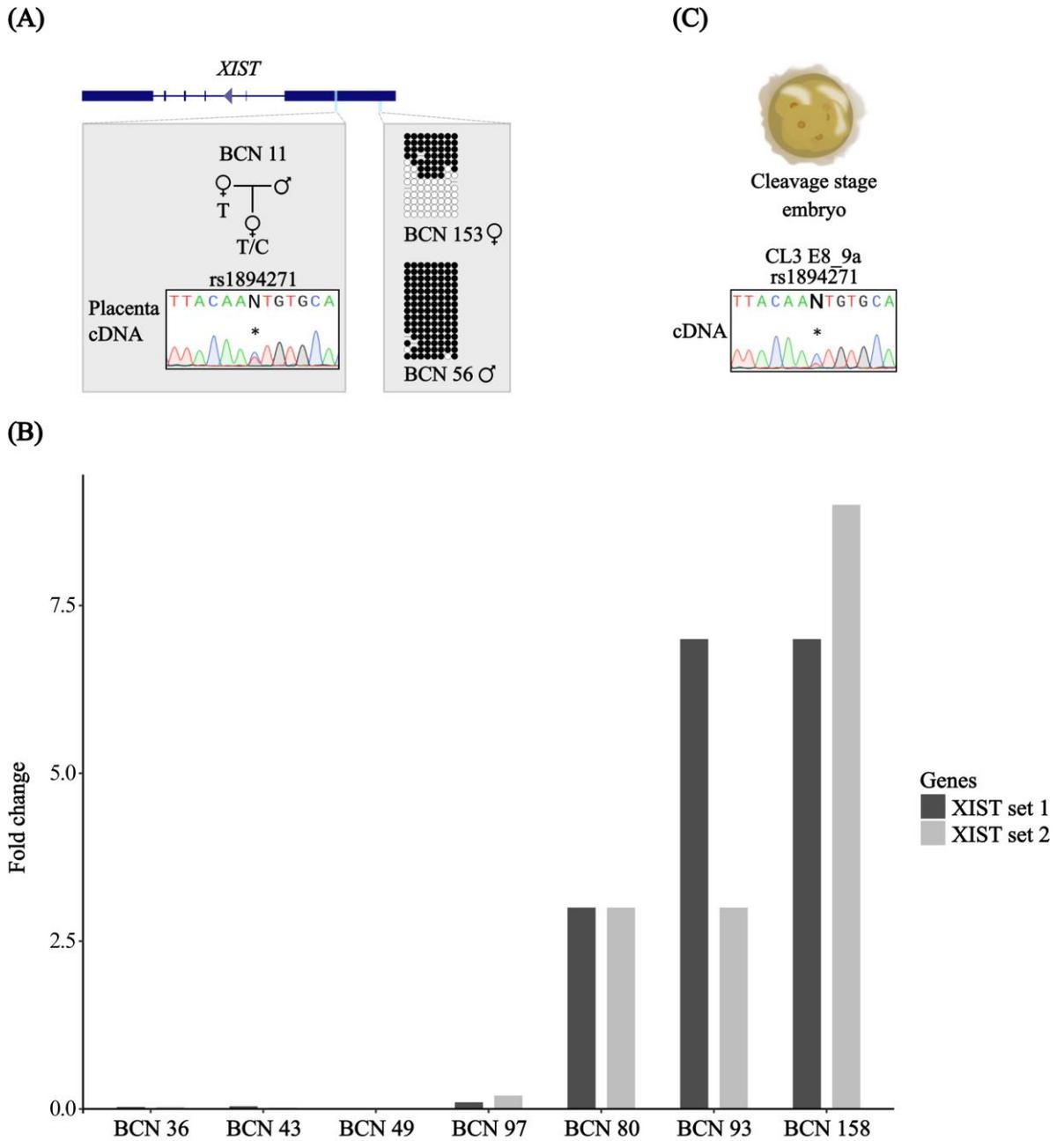
H3K27me3 ChIP-seq datasets for mouse oocytes and H3K27me3 ChIP-seq datasets for mouse post-implantation embryos. They found that *Xist* is covered by a broad H3K27me3 peak (~450 kb) in mature oocytes, and this region was hypomethylated and exhibited low chromatin accessibility. This H3K27me3 enrichment over *Xist* was maintained in pre-implantation embryos (1-cell stage to blastocyst) but was lost in the Epi (E6). To further confirm that H3K27me3 could be responsible for maternal *Xist* silencing, they injected *Kdm6b* mRNA (to remove H3K27me3) into zygotes that were grown to morulae and used for H3K27me3 ultralow input native ChIP-seq (ULI-NChIP) or blastocysts to perform RNA-seq. *Kdm6b* injected morulae not only showed a reduction in global H3K27me3, but this mark was also lost at the maternal *Xist* locus. At the same time, analysis of RNA-seq revealed that genes located on the maternal X were downregulated, which suggested maternal XCI. In the following study, the authors generated *Eed* matKO mouse morulae (derived from *Eed* KO oocytes) and performed an H3K27me3 CUT&RUN assay that further supported their previous findings as H3K27me3 was lost at the maternal *Xist* locus (187). Finally, RNA fluorescent in situ hybridisation (FISH) analysis revealed the expression of *Xist* from both parental chromosomes in female *Eed* matKO morulae and maternal *Xist* expression in male *Eed* matKO morulae. Therefore, the authors concluded that non-canonical imprinting is required for silencing the maternal *Xist*, which in turn prevents maternal XCI. In addition, Alber *et al.* (2023) (450) found that the *Xist* locus was covered by H3K27me3 (CUT&RUN dataset), and the genes present on the Xp were downregulated in rat EPCs, suggesting a conserved mechanism between mice and rats.

In humans, XCI remains a controversial topic mainly due to limited access to early-stage and good-quality human embryos (303,304). However, it is generally accepted that a random X chromosome becomes inactivated in female post-implantation embryos (13,302). In the past, several studies reported that the Xp was preferentially inactivated in human trophoblasts (529,530). However, more recent studies have shown that the placenta is composed of large patches of clonal cells with one of the parental X chromosomes inactivated (531,532). Thus, it was concluded that biopsies from several placental sites should be investigated. Interestingly, more recently, Hamada *et al.* (2016) (21) reported that at least in CTB cells from the first trimester placentae, the Xp was preferentially inactivated in their tested samples according to RNA-seq.

Thus, I decided to explore XCI in our placental cohort. Firstly, I investigated the methylation status at the P2 promoter that is located in the first exon of the *XIST* locus. For this, I selected 2 placentae derived from male and female offspring that I used for bisulphite PCR and sub-cloning (**Figure 4.5A**). In the male placenta, this promoter was fully methylated, while half of the cloned products were methylated in the female placenta, consistent with one active and one inactive *XIST* allele. Unfortunately, this region did not

contain any informative SNPs. To determine the allelic origin of *XIST* expression in the third-trimester placentae, I identified a common exonic SNP in the first exon of *XIST*, which I utilised for allelic RT-PCR (**Figure 4.5A**). Two informative female samples were identified that both demonstrated biallelic expression, supporting random XCI (Error! Reference source not found.). I further evaluated *XIST* expression in male and female placentae via qRT-PCR, for which I used whole placental cDNA from 4 male and 3 female placentae (**Figure 4.5B**). The three female samples showed abundant expression of *XIST*, suggesting continuous *XIST* expression to maintain inactivated X chromosomes across placental cells, while only residual expression was observed in the male samples.

The allelic expression of *XIST* was also investigated in female human pre-implantation embryos (**Figure 4.5C**). All female samples expressed *XIST*, but due to low rates of heterozygosity in our embryo cohort, only one embryo demonstrated biallelic expression. This further corroborates previous reports that *XIST* is biallelically expressed at human pre-implantation stages as both X chromosomes are active at this developmental window (529,530). Taken together, this data supports the notion that the X chromosomes undergo random XCI in the female placenta, as the expression of both *XIST* alleles is readily detected.



**Figure 4.5. Expression and methylation patterns of *XIST* in human pre-implantation embryos and term placental samples.**

(A) A schematic of the *XIST* locus showing an exonic SNP and the P2 promoter, both highlighted in light blue. The SNP rs1894271 showed biallelic expression in (A) female placental samples and (C) the day 3 CL female embryo. (A) DNA methylation at the P2 promoter was confirmed by bisulphite PCR and sub-cloning of placenta-derived DNA, revealing distinct methylation patterns in female and male placental samples. Methylated cytosines are indicated by (●), and unmethylated cytosines by (○), with each row corresponding to an individual cloned sequence. (B) *XIST* expression was quantified in male (BCN 36, BCN 43, BCN 49, BCN 97) and female (BCN 80, BCN 93, BCN 158) placental samples using different qRT-PCR primer pairs. The relative expression of *XIST* was normalized to *RPL19* and *ACTB* endogenous controls.

## 4.7. LTR-driven human transcripts are not imprinted in the placenta

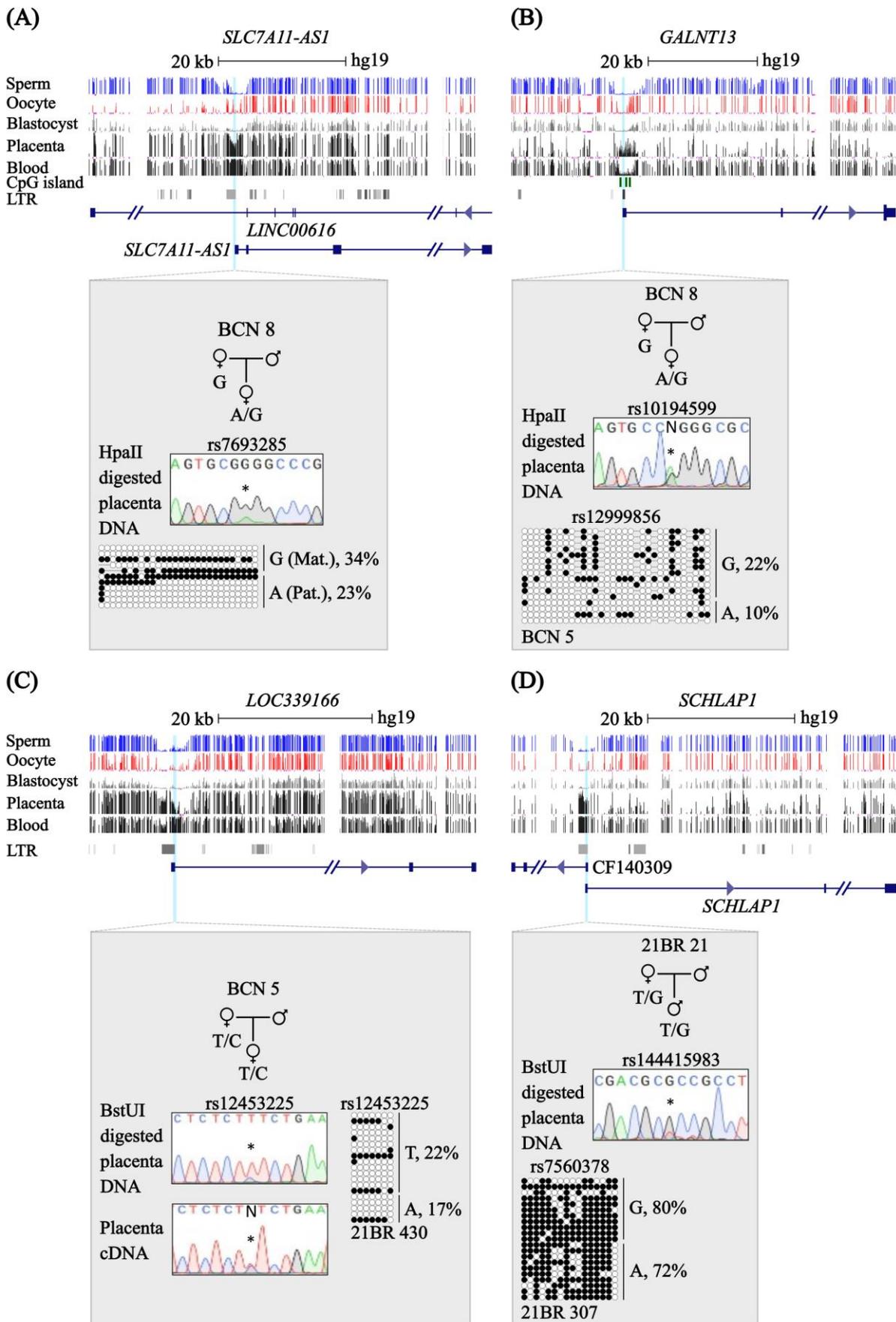
One study exploring the maintenance of non-canonical imprints in the mouse post-implantation ExE and later in the placenta made an interesting observation that some of these genes contained solo-LTRs near their promoters, which act as alternative promoters or enhancers. Hanna *et al.* (2019) (25) generated mouse embryos (C57BL6/BabR and CAST/EiJ) with maternal deletions of *Dnmt3a* and *Dnmt3b* (maternal double knockouts (matDKO)) and performed RNA-seq, PBAT (or WGBS) and low-input ChIP-seq to profile H3K4me<sub>3</sub>, H3K36me<sub>3</sub> and H3K27me<sub>3</sub> in Epi (E6.5) and ExE (E6.5). They discovered some non-canonical H3K4me<sub>3</sub> peaks that were not associated with genic CpG islands established on paternal alleles that were located near the promoters of several non-canonical imprints. These H3K4me<sub>3</sub> peaks contained the solo-LTRs of ERVs, in particular, endogenous retroviruses-K (ERVks). It was found that these ERVK LTRs were rich in CpG sites, were around 450 bp in length, shared the same orientation with the genes, overlapped the oocyte-derived H3K27me<sub>3</sub> peaks and acquired *de novo* DNA methylation in post-implantation ExE by forming sDMRs. Strikingly, the mosaic deletion of such LTR (RLTR15) found within the first intron of *Gab1* led to the upregulation of the smaller imprinted *Gab1* isoform (partial loss of paternal allele-specific expression).

Finally, a few studies have revealed that LTRs are important in mammalian oocytes and the placenta, as they drive the expression of novel transcripts and, at the same time, can form novel DMR regions that result in the monoallelic expression of certain loci (235,236,240).

Rodent ERVK elements responsible for non-canonical imprinting in the mouse and rat extra-embryonic tissues are not conserved in the human genome (25,450). Thus, I hypothesised that primate-specific ERVs could possibly drive the expression of human-specific non-canonical imprints in the term placenta. For this, I utilised a list of identified LTRs that initiate the expression of chimeric transcripts in human oocytes, which was discovered by Brind'Amour *et al.* (2019) (235). To identify active LTRs, the authors utilised PBAT libraries of human oocytes and discovered that 1/3 of hypermethylated regions were located in intergenic regions. They hypothesised that this could be due to active ERV LTRs, as many of such elements triggered expression in the mouse oocytes and were vital for mouse pre-implantation development. To test this hypothesis, they identified all annotated LTR elements present within intergenic regions and analysed RNA-seq of GVO and MII human oocytes using the LIONS pipeline, which allowed them to perform *de novo* transcriptome assembly. They identified 1056 expressed transcripts

overlapping LTRs in GVO and MII human oocytes. Thus, I further screened these transcripts in combination with methyl-seq datasets to identify putative candidate LTRs that could drive the expression of non-canonical imprints in the human placenta.

I identified four LTR elements that were hypomethylated in human gametes and blastocysts and showed around ~40% methylation in the term placenta, suggesting the presence of sDMR (**Figure 4.6**). Also, these LTRs were mostly hypermethylated in somatic tissues, except for one LTR present within the promoter of *GALNT13* (**Figure 4.6B**). All four LTR elements belonged to the ERV1 family, which is unsurprising, as it is one of the most abundant LTRs in primates (235). 3 LTR elements contained TSSs of ncRNAs, including *SLC7A11-AS1*, *LOC339166* and *SCHLAP1*, while one LTR was found upstream of the *GALNT13* promoter that is a protein-coding gene, encoding multiple isoforms with different TSSs in the human genome (**Figure 4.6A, C, D, B**). After inspecting these LTR regions, I found informative SNPs that I utilised for methylation-sensitive genotyping with either HpaII or BstUI to determine their methylation status in our placental cohort. Generated sequencing traces revealed the presence of methylation at both parental alleles, as the results were highly variable in the selected panel of placentae (**Figure 4.6, Table 4.6, Appendix 20**). To corroborate this further, I performed bisulphite PCR and sub-cloning, which confirmed the previous findings exhibiting different degrees of methylation on both alleles. Finally, I wanted to determine the allelic expression of ncRNAs (that were *LOC339166*, *SCHLAP1* and *SLC7A11-AS1*) by carefully designing several primers around exonic SNPs (**Appendix 28**). Unfortunately, I could only generate sequencing traces for *LOC339166* ncRNA that revealed biallelic expression (**Figure 4.6C, Table 4.6, Appendix 20**). Overall, this data suggests that these 4 transcripts located near ERV1 LTRs are not imprinted in the human placenta. In addition, it was shown that these ERV1 LTRs can drive the expression of chimeric transcripts in the human oocytes, but it is unclear if they are expressed in the third-trimester placenta.



**Figure 4.6. Characterisation of allelic methylation at gene promoters containing primate-specific LTRs.**

Genomic maps display the partially methylated LTR-derived promoters of human genes: (A) *SLC7A11-ASI*, (B) *GALNT13*, (C) *LOC339166*, and (D) *SCHLAP1* loci in the placental methyl-seq

dataset. Vertical lines in the methyl-seq tracks represent the mean methylation levels of individual CpG dinucleotides. Gene transcripts are shown in dark blue, with thicker bars representing exons, while CpG islands are shown as dark green bars. The locations of LTRs (grey bars) were retrieved from the UCSC RepeatMasker track. For each gene, promoter methylation was confirmed using methylation-sensitive genotyping and bisulphite PCR, followed by sub-cloning of placenta-derived DNA. Methylated cytosines are indicated by (●), and unmethylated cytosines by (○), with each row representing an individual cloned sequence. The parent-of-origin was inferred from SNP genotyping if the placental sample was heterozygous. (C) Allelic expression of *LOC339166* was assessed by including SNP rs12453225 (highlighted in light blue) within RT-PCR products.

**Table 4.6. Result summary for human genes with LTR-associated promoters**

Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
2	<i>GALNT13</i>	-	10	rs62174125, rs12999856, rs10194599	Biallelic	6	60%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	2	20%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	2	20%	Uninformative	-	-
2	<i>SCHLAPI</i>	-	5	rs144415983, rs148398319, rs7560378	Biallelic	2	40%	Biallelic	-	-
					Pref. monoallelic	2	40%	Pref. monoallelic	-	-
					Monoallelic	1	20%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
4	<i>SLC7A11 -AS1</i>	-	4	rs7693285, rs7699108	Biallelic	2	50%	Biallelic	-	-
					Pref. monoallelic	2	50%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
17	<i>LOC339166</i>	-	10	rs12450161, rs12450165, rs12453225	Biallelic	2	20%	Biallelic	4	40%
					Pref. monoallelic	0	0%	Pref. monoallelic	1	10%
					Monoallelic	5	50%	Monoallelic	3	30%
					Maternal	2	20%	Maternal	1	10%
					Uninformative	1	10%	Uninformative	1	10%

## 4.8. Previously identified human candidates of non-canonical imprinting demonstrate biallelic expression in the human placenta

Since the discovery that H3K27me3 can mediate imprinting independent of DNA methylation in rodents, several teams have attempted to investigate whether this type of imprinting could be conserved in humans. One of the first studies conducted by Zhang *et al.* (2019) (27) performed a CUT&RUN assay to investigate global H3K27me3 distribution in good-quality human morulae remaining from IVF patients. To perform this technique, the authors combined 7 and 8 morulae from two different couples. In addition, they performed RNA-seq for 15 human morulae from 5 couples and either WGS or WES for cumulus cells collected from 5 women. By profiling SNPs in embryo RNA-seq in combination with maternal WGS or WES datasets, authors identified 44 paternally expressed genes (paternal-biased expression detected in at least two morulae; SNPs were considered if covered by at least 10 reads). By utilising publicly available WGBS of human gametes and morulae, they identified that 17 out of 44 genes were most likely controlled by canonical imprinting, as these genes were located near oocyte hypermethylated regions or mDMRs. This left 27 paternally expressed genes, of which the promoters of 5 genes overlapped H3K27me3 domains. Due to limited coverage and lack of informative exonic SNPs, they determined that *FAM101A* was a paternally expressed gene with maternal H3K27me3.

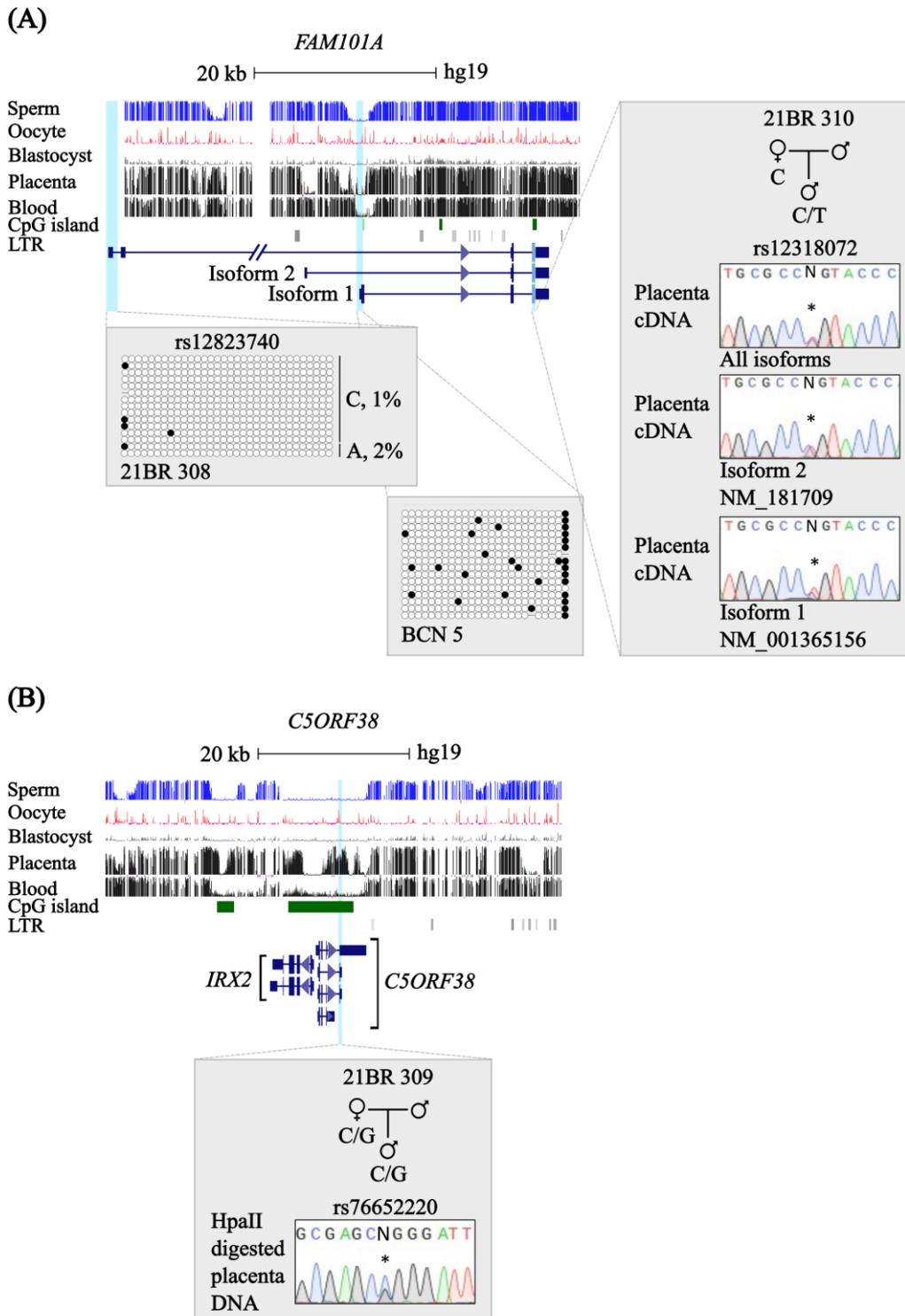
Since *FAM101A* has been reported to be the only non-canonical imprinted gene in human pre-implantation embryos, and its imprinting status is unreported in placenta, I characterised the allelic expression and methylation of this gene in human third-trimester placental samples. *FAM101A*, also known as *RFLNA*, has several isoforms that all have unique TSS (**Figure 4.7A**). I mainly focused on isoform 1, as its promoter contains a small CpG island (18 CpGs) that is unmethylated in human gametes and blastocyst and maintains slight methylation in the placenta. The unmethylated status was confirmed by cloning and sequencing of bisulphite PCR products. Interestingly, this gene can also form a long fusion transcript with the upstream *ZNF664* gene called *ZNF664-RFLNA*, which originates from a CpG island promoter (194 CpGs). This interval is unmethylated in all our methyl-seq datasets and is devoid of methylation in the term placenta (**Figure 4.7A**).

To investigate if *FAM101A* demonstrates paternal allele-specific expression in the placenta, I performed nested allelic RT-PCR as rs12318072 - an exonic SNP was shared between two *FAM101A* isoforms and the fusion transcript (**Appendix 28**). After inspecting the sequencing chromatograms, I found that both isoforms were biallelically

expressed in term placenta, and all transcripts together were biallelically expressed in seven samples (**Figure 4.7A, Table 4.7, Appendix 21**). Therefore, I conclude that this gene is not imprinted in the third-trimester placenta. Additionally, none of the pre-implantation embryos was informative for *FAM101A*, and therefore, I was unable to investigate the allelic expression of this gene at earlier stages of human development.

In their review article, Kelsey and Hanna (2021) (517) screened for non-canonical imprinting candidates. Their approach involved interrogating publicly available WGBS (PBAT libraries) and H3K27me3 CUT&RUN datasets from human oocytes for regions that were hypomethylated (less than 25% methylation) and marked by H3K27me3. This analysis revealed 65 putative placental sDMRs. By analysing SNP information in WGBS datasets from first-trimester placental trophoblasts, they narrowed the 65 candidate sDMRs down to 26 placental DMRs. Using the SNPsplit mapping program, they further identified 16 placental DMR regions that showed less than a 10% difference in allelic DNA methylation, suggesting that these genes could be controlled by oocyte-derived H3K27me3 and might be imprinted in the human placenta.

Using our methyl-seq datasets for different human samples (cells and tissues), I assessed the 65 putative placental sDMRs in greater detail. By applying our previous criteria, which require that regions must be unmethylated in gametes and the blastocyst and maintain around ~50% methylation in the placenta, only one candidate region was identified. This region was associated with a large CpG island with 648 CpG dinucleotides within the gene body of *C5ORF38* (**Figure 4.7B, Table 4.7, Appendix 21**). This gene encodes a ncRNA and has two highly similar isoforms, sharing 2 informative exonic SNPs. In addition, this region contained 8 HpaII restriction sites. To investigate the allelic methylation profile within this interval, I employed our methylation-sensitive genotyping assay. This revealed that 2 placentae showed biallelic methylation for rs62333235, and 3 samples demonstrated biallelic methylation for rs76652220 (**Table 4.7, Appendix 21**). Therefore, it is unlikely that this region represents a non-canonical sDMR in the human placenta.



**Figure 4.7. Investigation of previously reported human candidate non-canonical imprinted genes.**

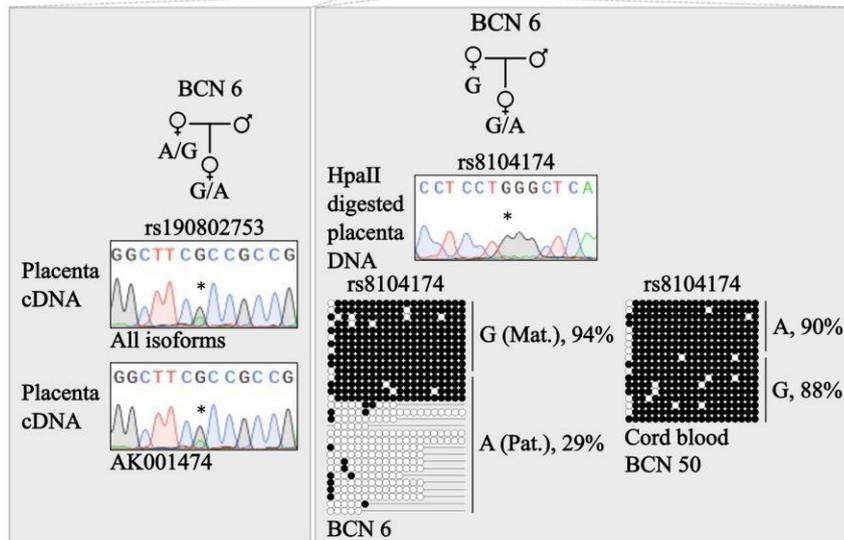
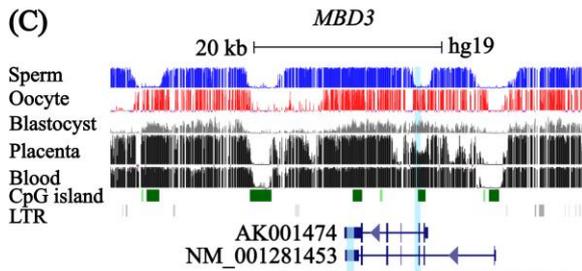
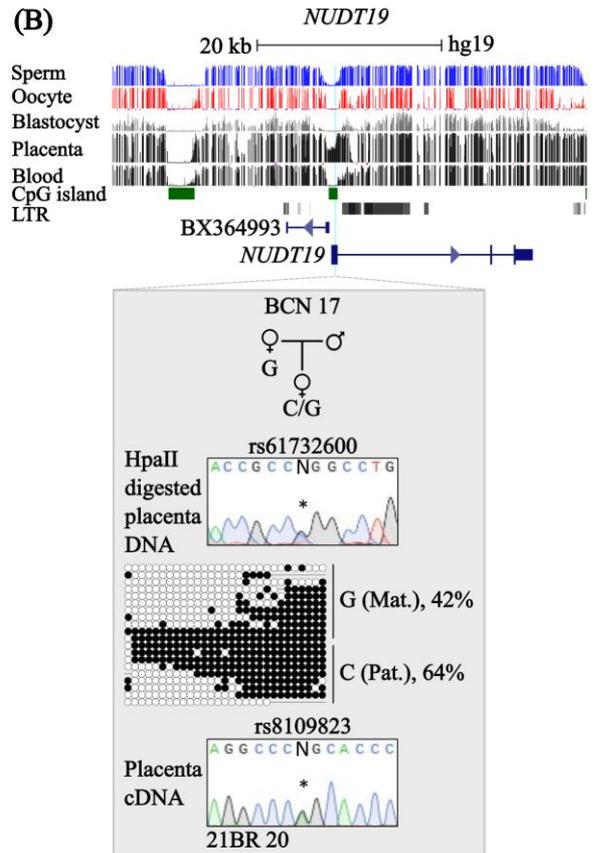
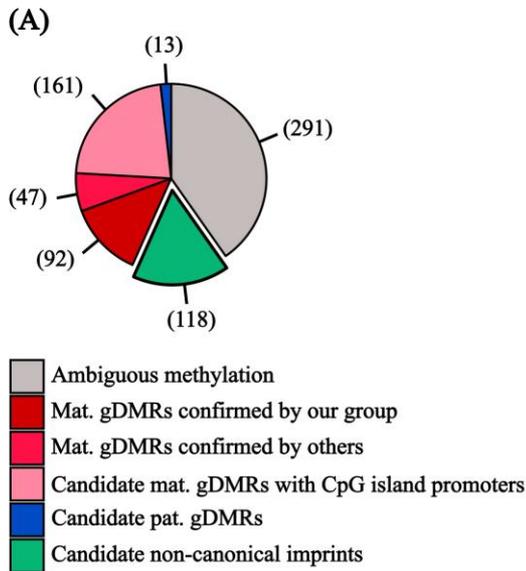
Genomic maps of (A) *FAM101A* and (B) *C5ORF38* display DNA methylation profiles from methyl-seq datasets of human sperm, oocyte, blastocyst, placenta, and blood. Vertical lines in the methyl-seq tracks represent the mean methylation levels of individual CpG dinucleotides. Gene transcripts are shown in dark blue, with thicker bars representing exons, while CpG islands are depicted as dark green bars and ERV LTRs as grey bars. (A) For *FAM101A*, DNA methylation at isoform-specific promoters was confirmed using bisulphite PCR and sub-cloning of placenta-derived DNA. Methylated cytosines are indicated by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with the parent-of-origin inferred through SNP genotyping if the placental sample was heterozygous. Allelic isoform-specific expression was

determined by including SNPs (highlighted in light blue) in RT-PCR products generated using isoform-specific PCR primers, with SNP IDs displayed above the corresponding sequencing chromatograms. **(B)** For *C5ORF38*, allelic methylation was examined through methylation-sensitive genotyping at the 3' UTR, which contains the placental sDMR.

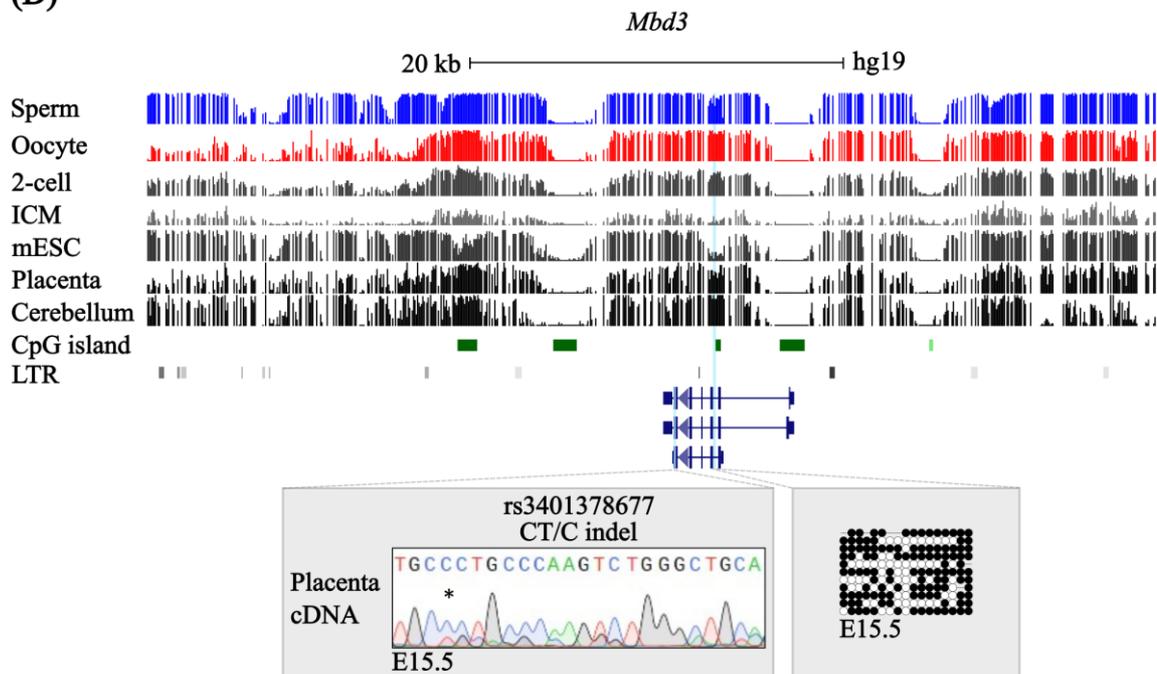
## 4.9. Human genes with placenta sDMR regions are not imprinted in the human placenta

While cross-species comparisons between the mouse and human and between the rat and human failed to identify any non-canonical imprints, it is possible that humans harbour a unique set of genes that could be controlled independently of germline imprinted DMRs in the placenta.

To identify such candidates, I explored our previously identified 722 partially methylated regions specific to the human placentae (449) (**Figure 4.8A**). I further selected those regions that were unmethylated in both gametes and blastocysts, as determined by methyl-seq datasets, which revealed 118 such genomic regions, of which 94 were located close to genes. I further screened these genes for common polymorphisms ( $MAF \geq 0.1$ ) and selected 14 promising candidate genes (**Table 4.7, Appendix 21**). After genotyping our placental samples for these loci and comparing their genotypes with the sequence traces generated by methylation-sensitive genotyping, I found that all placental sDMRs were randomly methylated, demonstrating the presence of methylation on both parental alleles (**Table 4.7, Appendix 21, Appendix 22**). I decided to explore one of these loci in more detail. For this, I selected *NUDT19*, which had a CpG island promoter (113 CpGs) shared with BX364993 anti-sense transcript (**Figure 4.8B**). I performed bisulphite PCR, followed by sub-cloning and sequencing, to ensure that the biallelic methylation observed following methylation-sensitive genotyping represented the entire promoter region. Analysis of cloned bisulphite PCR products confirmed the previous findings, showing the presence of methylation on both parental alleles. I also tested the allelic expression of this gene by selecting 3 most informative placental samples for allelic RT-PCR, which revealed biallelic expression (**Figure 4.8B, Table 4.7, Appendix 21**). Hence, our identified 14 genes with placental sDMRs were not imprinted in the human placenta, and I believe that non-canonical imprinting might be specific to the rodent lineage.



(D)



**Figure 4.8. Systematic screening of human loci with placental sDMRs.**

(A) Pie chart showing the distribution of 722 partially methylated placental regions identified in methyl-seq datasets. (B) Genomic map of the human *NUDT19* locus displaying DNA methylation profiles in human sperm, oocyte, blastocyst, placenta, and blood. Methylation at the *NUDT19* promoter (highlighted in light blue) was assessed using methylation-sensitive genotyping and bisulphite PCR, followed by sub-cloning, revealing biallelic methylation. Biallelic expression of *NUDT19* was determined by including SNP rs8109823 (highlighted in light blue) in RT-PCR products. (C) Genomic map of human *MBD3* isoforms, showing DNA methylation profiles in several human methyl-seq datasets. The smaller AK001474 isoform contains a placenta-specific mDMR within its alternative promoter (highlighted in light blue). Partial methylation at this region was confirmed via methylation-sensitive genotyping and bisulphite PCR, followed by sub-cloning of placental DNA, while hypermethylation was observed in cord blood. Allelic and isoform-specific expression was investigated using an exonic SNP (highlighted in light blue) in RT-PCR products. (D) Genomic map of mouse *Mbd3*, displaying DNA methylation profiles from various mouse methyl-seq datasets. Biallelic methylation at the promoter of the smaller *Mbd3* isoform (highlighted in light blue) was confirmed through bisulphite PCR and sub-cloning of C57BL6 and JF1 hybrid placental DNA. Biallelic expression was determined using the indel rs3401378677 (highlighted in blue; reported in mouse GRCm39) in RT-PCR products. Vertical lines in the methyl-seq tracks represent the mean methylation levels of individual CpG dinucleotides. Gene transcripts are shown in dark blue, with thicker bars representing exons. CpG islands are shown in dark green, and ERV LTRs are depicted as grey bars. Methylated cytosines are indicated by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with the parent-of-origin inferred through SNP genotyping if the placental sample was heterozygous.

**Table 4.7. Result summary for human genes with placental sDMRs**

Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
1	<i>DNAJC6</i>	-	6	rs577841	Biallelic	2	33%	Biallelic	-	-
					Pref. monoallelic	3	50%	Pref. monoallelic	-	-
					Monoallelic	1	17%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
2	<i>C2ORF40 &amp; ECRG4</i>	-	12	rs4271786, rs4266035, rs73949223, rs4477942	Biallelic	8	67%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	3	25%	Monoallelic	-	-
					Maternal	1	8%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
4	<i>CRMP1</i>	-	2	rs139357095	Biallelic	2	100%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
4	<i>CWH43</i>	-	14	rs3747690	Biallelic	8	57%	Biallelic	-	-
					Pref. monoallelic	3	21%	Pref. monoallelic	-	-
					Monoallelic	1	7%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	2	14%	Uninformative	-	-
5	<i>C5ORF38</i>	-	3	rs62333235, rs76652220	Biallelic	3	100%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-

Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
5	<i>ANKDD1B</i>	-	6	rs72633976, rs1489, rs61516153	Uninformative	0	0%	Uninformative	-	-
					Biallelic	0	0%	Biallelic	-	-
					Pref. monoallelic	1	17%	Pref. monoallelic	-	-
					Monoallelic	5	83%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
6	<i>TFAP2B</i>	-	3	rs4628086, rs62405419	Biallelic	3	100%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
8	<i>SULF1</i>	-	4	rs2704035	Biallelic	4	100%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
8	<i>RGS22</i>	-	12	rs2453627	Biallelic	11	92%	Biallelic	-	-
					Pref. monoallelic	1	8%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
12	<i>KRT86</i>	-	2	rs2078294	Biallelic	2	100%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-

Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
12	<i>FAM101A</i>	All isoforms	7	rs12318072	Uninformative	0	0%	Uninformative	-	-
					Biallelic	-	-	Biallelic	5	71%
					Pref. monoallelic	-	-	Pref. monoallelic	2	29%
					Monoallelic	-	-	Monoallelic	0	0%
					Maternal	-	-	Maternal	0	0%
					Uninformative	-	-	Uninformative	0	0%
					Biallelic	-	-	Biallelic	2	29%
		Pref. monoallelic			-	-	Pref. monoallelic	0	0%	
		Monoallelic			-	-	Monoallelic	0	0%	
		Maternal			-	-	Maternal	3	43%	
		Uninformative			-	-	Uninformative	2	29%	
		Biallelic			-	-	Biallelic	7	100%	
		Pref. monoallelic			-	-	Pref. monoallelic	0	0%	
		Monoallelic			-	-	Monoallelic	0	0%	
Maternal	-	-	Maternal	0	0%					
Uninformative	-	-	Uninformative	0	0%					
15	<i>LTK</i>	-	4	rs1077809	Biallelic	4	100%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
17	<i>PLXDC1</i>	-	4	rs188501857	Biallelic	2	50%	Biallelic	-	-
					Pref. monoallelic	1	25%	Pref. monoallelic	-	-
					Monoallelic	1	25%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-

Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
19	<i>NUDT19</i>	-	15	rs8108621, rs8109823, rs61732600	Uninformative	0	0%	Uninformative	-	-
					Biallelic	6	40%	Biallelic	3	20%
					Pref. monoallelic	4	27%	Pref. monoallelic	1	7%
					Monoallelic	4	27%	Monoallelic	0	0%
					Maternal	1	7%	Maternal	0	0%
					Uninformative	0	0%	Uninformative	11	73%
19	<i>FFAR1</i>	-	8	rs2301151	Biallelic	6	75%	Biallelic	-	-
					Pref. monoallelic	2	25%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
20	<i>TSPY26P &amp; PLAGL2</i>	-	12	rs11907716, rs11907235	Biallelic	0	0%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	10	83%	Monoallelic	-	-
					Maternal	2	17%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-

## 4.10. Novel candidate imprinted genes with placenta-specific mDMRs

While searching for non-canonical candidate genes with sDMRs in the list of 722 partially methylated regions in the human placenta, I observed multiple regions harbouring maternal gDMRs or mDMRs derived from the oocyte. I identified 300 regions containing mDMRs, of which 139 had been characterised previously by several groups, including ours (20–22,295) (**Figure 4.8A**). I further screened the remaining 161 mDMRs for CpG island promoters that contained common polymorphisms (MAF > 0.1). Only 8 loci fulfilled these criteria, including *DYRK1B*, *LRRC8D*, *WNT7B*, *EID3*, *CLDN23*, *PRKAG2*, *STARD13* and *MBD3* (**Table 4.8, Appendix 23, Appendix 24**). Interestingly, all 8 genes were unmethylated in sperm and other somatic tissues, except the mDMR at *MBD3*, which was fully methylated in all somatic tissues (**Figure 4.8C**). I confirmed this observation by applying bisulphite PCR and sub-cloning to cord blood samples, where this region was methylated on both parental alleles. In contrast to somatic tissues, the maternal allele was methylated at this region in the third-trimester placenta samples. This *MBD3*-associated mDMR overlapped a CpG island (89 CpGs) that is located between the second and third exons of the full-length *MBD3* transcripts, which suggested that it might behave as an alternative promoter. After inspecting GENCODE and GenBank, the NIH genetic sequence database, I found that the mDMR encompasses the TSS of an alternative *MBD3* isoform (GenBank accession number AK001474). To determine the allelic expression of AK001474, I designed RT-PCR primers around an exonic SNP shared by *MBD3* isoforms. Unfortunately, due to low heterozygosity in our placental cohort, only 2 heterozygous placental samples were identified (**Figure 4.8C, Table 4.8, Appendix 23**). One sample exhibited preferential monoallelic expression, while the other sample demonstrated biallelic expression.

Based on previous reports and my previous data for *PIK3R1* and *GoS2*, human placenta-specific mDMRs are not conserved in non-primate mammals such as mice (20,449). The imprinting status of the *MBD3* ortholog in the mouse was determined using placental DNA from a hybrid mouse (E15.5). Similar to humans, the mouse has several *Mbd3* isoforms, all originating from different CpG island promoters (**Figure 4.8D**). The smaller *Mbd3* isoform (CV675626) mapped to a similar location as human AK001474. Upon characterisation, unlike in the human genome, this CpG island promoter was only hypomethylated in mouse ICM and mESCs and hypermethylated in other investigated methyl-seq datasets (**Figure 4.8D**). Following the cloning of bisulphite PCR products, I observed mosaic methylation overlapping the promoter of the smaller *Mbd3* isoform. Allelic RT-PCR revealed biallelic expression of *Mbd3* for the rs3401378677 indel, which is

included in the GRCm39 mouse reference genome. Overall, I identified an additional 8 germline-derived placenta-specific mDMRs in the human placenta, reassuring us that should allelic methylation be present, our molecular approaches would readily detect it.

**Table 4.8. Result summary for human genes with placenta-specific mDMRs**

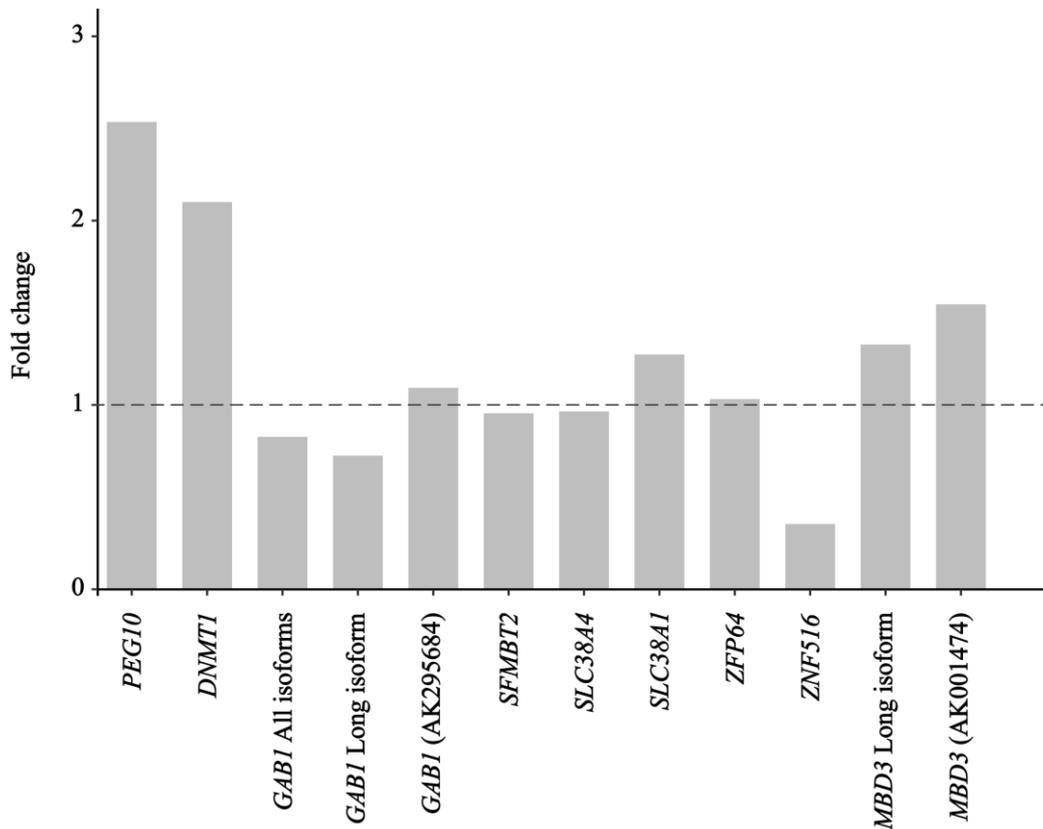
Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
1	<i>LRR8D</i>	-	4	rs114770365, rs115363384, rs114208181, rs113834473	Biallelic	1	25%	Biallelic	0	0%
					Pref. monoallelic	0	0%	Pref. monoallelic	0	0%
					Monoallelic	1	25%	Monoallelic	0	0%
					Maternal	2	50%	Maternal	0	0%
					Uninformative	0	0%	Uninformative	0	0%
7	<i>PRKAG2</i>	NM_024429	11	rs6964957, rs8961	Biallelic	1	9%	Biallelic	2	18%
					Pref. monoallelic	7	64%	Pref. monoallelic	5	45%
					Monoallelic	0	0%	Monoallelic	2	18%
					Maternal	0	0%	Maternal	1	9%
					Uninformative	3	27%	Uninformative	1	9%
8	<i>CLDN23</i>	-	9	rs9644774, rs11995449	Biallelic	0	0%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	2	22%	Monoallelic	-	-
					Maternal	3	33%	Maternal	-	-
					Uninformative	4	44%	Uninformative	-	-
12	<i>EID3</i>	-	5	rs7488680, rs58078551	Biallelic	4	80%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	0	0%	Monoallelic	-	-
					Maternal	0	0%	Maternal	-	-
					Uninformative	1	20%	Uninformative	-	-
13	<i>STARD13</i>	AK308453	13	rs5011113, rs495680	Biallelic	0	0%	Biallelic	7	54%
					Pref. monoallelic	0	0%	Pref. monoallelic	1	8%
					Monoallelic	3	23%	Monoallelic	1	8%
					Maternal	2	15%	Maternal	0	0%

Chr.	Gene	Isoform	Total no. of informative samples	Variants	Methylation-sensitive genotyping (HpaII)			Allelic expression		
19	<i>MBD3</i>	All isoforms	12	rs8104174, rs190802753	Uninformative	8	62%	Uninformative	4	31%
					Biallelic	3	25%	Biallelic	1	8%
					Pref. monoallelic	1	8%	Pref. monoallelic	1	8%
					Monoallelic	4	33%	Monoallelic	0	0%
					Maternal	4	33%	Maternal	0	0%
					Uninformative	0	0%	Uninformative	10	83%
		AK001474	Biallelic	3	25%	Biallelic	1	8%		
			Pref. monoallelic	1	8%	Pref. monoallelic	1	8%		
			Monoallelic	4	33%	Monoallelic	0	0%		
			Maternal	4	33%	Maternal	0	0%		
			Uninformative	0	0%	Uninformative	10	83%		
19	<i>DYRK1B</i>	-	6	rs2354800	Biallelic	2	33%	Biallelic	-	-
					Pref. monoallelic	1	17%	Pref. monoallelic	-	-
					Monoallelic	1	17%	Monoallelic	-	-
					Maternal	2	33%	Maternal	-	-
					Uninformative	0	0%	Uninformative	-	-
22	<i>WNT7B</i>	-	10	rs62226057	Biallelic	0	0%	Biallelic	-	-
					Pref. monoallelic	0	0%	Pref. monoallelic	-	-
					Monoallelic	1	10%	Monoallelic	-	-
					Maternal	4	40%	Maternal	-	-
					Uninformative	5	50%	Uninformative	-	-

## 4.11. Candidate gene expression in different placental cell lines

I took advantage of recently derived human CTs (CT30 cell line) and cells derived from molar pregnancies (Mole 1), both established by Okae group at the University of Tohoku (243,311). The CT30 cell line is derived from a first-trimester placenta and maintains the most ubiquitous and placenta-specific imprints. In contrast, allelic methylation was lost in the mole-derived cell line, consistent with its AG nature. These cell line models allow for the investigation of imprinting status without relying on genetic variants.

Imprint expression was investigated by comparing the expression levels of several orthologs of mouse and rat non-canonical imprinted genes in cDNA derived from CT30 and Mole 1. To ensure this approach would identify imprinting profiles, the expression of two known paternally expressed genes was used as a control. Both *PEG10* and *DNMT1* showed a 2-fold increase in Mole 1 compared to CT30, consistent with two active paternally-derived chromosomes. qRT-PCR primer sets were designed to distinguish the expression of different *GAB1* isoforms (**Appendix 28**). The expression levels of the short and long *GAB1* isoforms, as well as total expression, were similar in both Mole 1 and CT30 cells (**Figure 4.9**). Similar results were observed for four other orthologs of mouse and rat non-canonical imprinted genes (*SFMBT2*, *SLC38A4*, *SLC38A1* and *ZFP64*), except for *ZNF516*, whose transcription was reduced greatly in the Mole 1 cell line. In addition, expression of *MBD3* isoforms, including AK001474 (originating from the placenta-specific mDMR) and the longer NM\_001281453 transcript, was assessed and found to be upregulated in Mole 1 cells. This was especially pronounced for AK001474, which showed ~1.7 times higher expression in Mole 1 cells compared to CT30 cells. In summary, I showed that known canonical imprints and our candidate genes with placenta-specific mDMRs became upregulated in the Mole 1 cell line than compared to CT30 cells, while non-canonical imprinted gene orthologs displayed comparable expression levels, indicating biallelic expression.



**Figure 4.9. Expression profiling of human canonical imprinted genes, orthologs of mouse and rat non-canonical imprinted genes, and human genes with placenta-specific mDMRs in Mole 1 cells.**

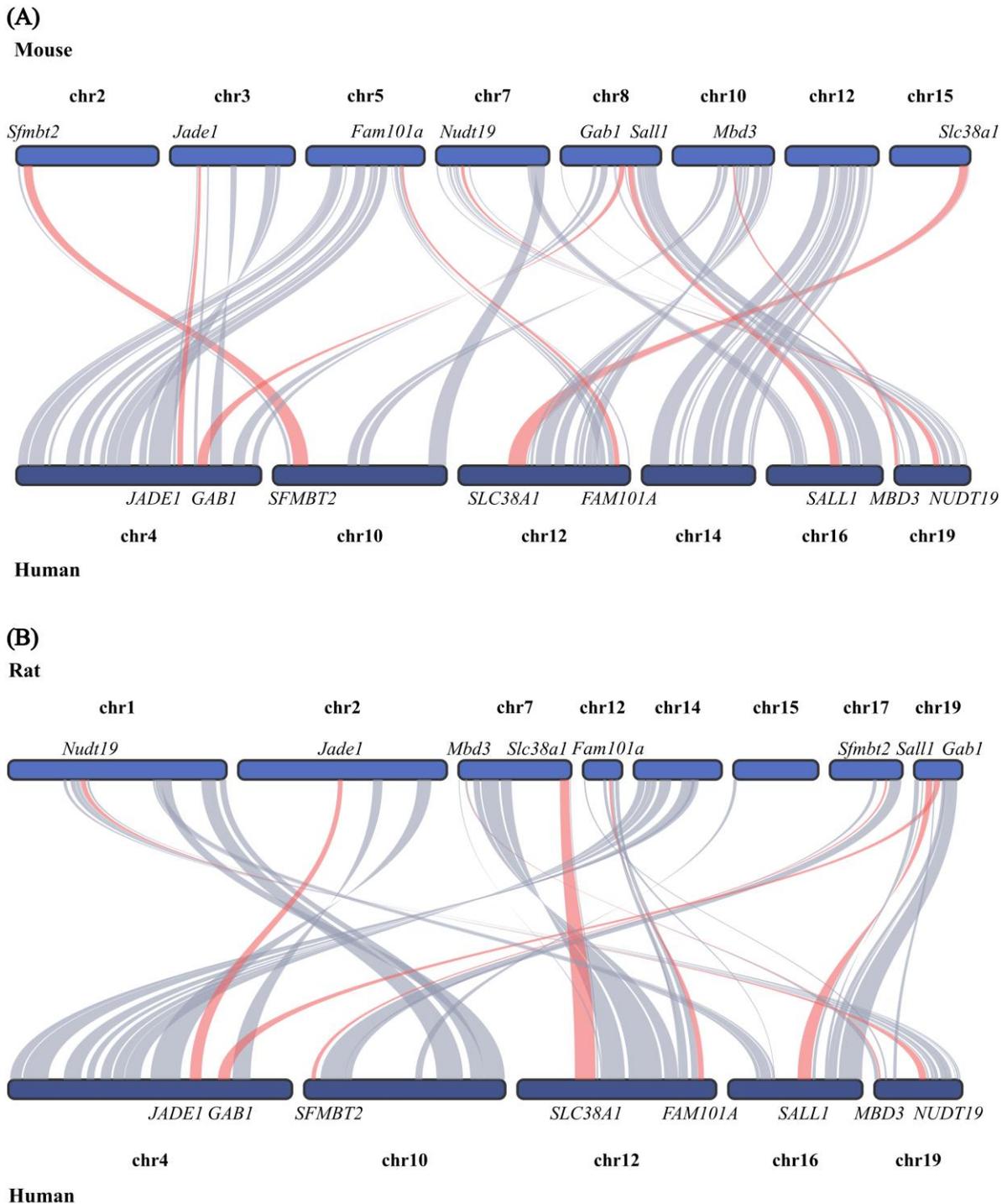
The expression of all genes in Mole 1 cells was normalized to the expression levels detected in CT 30 cells, indicated by a dashed line. *PEG10* and *DNMT1*, known paternally expressed genes, were used as controls. *GABI*, *SFMBT2*, *SLC38A4*, *SLC38A1*, *ZFP64*, and *ZNF516* are orthologs of mouse and rat non-canonical imprinted genes, while *MBD3* (AK001474) contains a placenta-specific mDMR.

## 4.12. Gene synteny

Synten refers to highly conserved regions in a genome that are shared between species as they arose from a common ancestor (480). Such syntenic regions frequently exhibit the same order of genes (collinearity) between species (483,533). Imprinting clusters controlled by DNA methylation-dependent imprinting are highly conserved among placental mammals, as many of them are shared between humans and mice (18,534). Genes present within these clusters overall show high collinearity with a few more recent evolutionary events as a few genes were inserted into imprinting clusters after a human and mouse divergence, such as *U2AF1-RS1* (maternal allele imprinted) or *Zim1* (paternal allele imprinted) mouse imprinted genes that do not have orthologs in the human genome (454). Surprisingly, many genes from imprinting clusters are highly conserved in the

chicken genome and, to some degree, in lower vertebrates, including zebrafish or spotted green pufferfish.

Thus, I explored the conservation of mouse and rat non-canonical imprinted gene orthologues in the human genome. For this, I used the *in silico* tool ShinySyn (481), which utilised the MCscan algorithm (482–484) (**Section 2.9.1**). This algorithm employs highly conserved homologous genes as anchors to identify conserved syntenic regions between comparable genomes. Interestingly, out of 10 mouse and rat non-canonical genes, I identified 5 homologous genes in humans (**Figure 4.10**). This suggested that human orthologs of *Smoc1*, *Slc38a4*, *Znf516*, *Zfp64* and *Gst01* were less conserved in the human genome and present within more rearranged chromosomal regions. Alternatively, these genes may have been missed due to poorer annotation between different genomes, as different databases were used to retrieve the reference genomes. The other five genes are highly conserved across species, as they were found in syntenic blocks. Unlike mouse canonical imprints, all identified non-canonical imprints were present on different chromosomes in both rodent species (**Figure 4.10A, B**). Similarly, non-canonical imprinted gene orthologs were scattered across the human genome. I also looked at the conservation of several human candidate genes discussed in this thesis, including *MBD3*, *NUDT19* and *FAM101A*, which are located on human chromosomes 19 and 12. These genes were also detected in different syntenic regions across different mouse and rat chromosomes. In summary, mouse and rat non-canonical imprints form no clusters, unlike canonical imprints (535), as they are single isolated genes distributed across the genomes, although they are present in macro-syteny regions, indicating higher conservation between humans and rodents.



**Figure 4.10. Macro-synteny maps showing highly conserved genomic regions between mouse, rat, and human chromosomes.**

Only chromosomes containing mouse and rat non-canonical imprinted genes are shown. Blue and red ribbons indicate syntenic regions with highly conserved orthologous genes (anchors) exhibiting high collinearity. Red ribbons also highlight mouse and rat non-canonical imprinted genes and, additionally, three human candidate genes: *NUDT19* (placental sDMR), *MBD3* (placenta-specific mDMR), and *FAM101A* (a previously reported human candidate gene for non-canonical imprinting). (A) Syntenic regions were identified between mouse and human chromosomes using ShinySyn. (B) Syntenic regions were identified between rat and human chromosomes using ShinySyn.

## **Chapter 5: Discussion**

## 5.1. Polymorphic imprinting of *GOS2* and *PIK3R1* in the human placenta

Abnormal expression of canonical imprinted genes regulated by ICR regions results in rare, multifactorial, and often severe developmental disorders, including SRS, BWS, and Kagami-Ogata syndrome (KOS) (420,496,536). Some of these conditions are often associated with diverse placental pathologies, including placental mesenchymal dysplasia, placentomegaly, or placental hypoplasia with hypoplastic chorionic villi. LOI is frequently implicated in the early development and progression of cancer (537). For example, a higher level of *IGF2* due to LOI has been associated with childhood Wilms tumours (538) as well as colorectal cancer (539,540). Thus, imprinted genes are known for their importance during development, but they can also play a significant role in early development and fertility.

Recent studies applying high-throughput sequencing techniques to investigate DNA methylation profiles have identified multiple differentially methylated regions (DMRs) between human gametes, with the majority being oocyte-specific (4,21,88,501). Monk group and others have shown that many of these regions persist in the human placenta but not in other embryonic tissues (20–22,449). In the placenta, the maternal allele retains methylation, forming placenta-specific mDMRs. Our group previously identified 551 such placental mDMRs (20), while Hanna *et al.* (2016) (22) reported 882, and Hamada *et al.* (2016) (21) identified 3,676 candidate mDMRs in the human placenta. These three studies highlight the high frequency of placental mDMRs, with the variable numbers likely due to differences in screening techniques and bioinformatic criteria. Thus, these regions are highly prevalent, but they are restricted to this foetal organ. Interestingly, all groups concluded that placenta-specific mDMRs are highly polymorphic, unlike canonical imprints orchestrated by ICRs, making them challenging to study (309). It has been suggested that placenta-specific mDMRs may form due to incomplete maternal DNA methylation erasure during pre-implantation development (7,20,21). Alternatively, these regions could form during post-implantation stages due to incomplete *de novo* methylation or inefficient maintenance of DNA methylation at such genomic regions. To fully determine if these mDMRs are polymorphically established or show variation in their maintenance during the pre-implantation stage, single-cell methylomes from many oocytes would be required. Furthermore, the function of these placenta-specific mDMRs remains unknown, but a few placenta-specific imprinted genes reported by Hamada were associated with pregnancy-associated diseases (21).

A large proportion of the placenta-specific mDMRs identified by our group remained uncharacterised due to low heterozygosity in the previous placental cohort (20). Therefore, in this PhD project, we revisited these regions and compared them with placenta-specific mDMRs identified by other groups (21,22). In the process, we identified two promising candidate genes near placenta-specific mDMRs: *PIK3R1* and *GoS2*. We characterised the methylation and expression patterns of these genes in our extended placental cohort, which included both normal samples and those affected by pregnancy complications.

We investigated *PIK3R1* and *GoS2* using our previously developed strategy, which allowed us to identify nine novel paternally expressed genes with placenta-specific mDMRs (20). This approach employed several techniques, including methylation-sensitive genotyping, bisulphite PCR and allelic RT-PCR, which we used in combination with methyl-seq datasets from human gametes, blastocysts, and other somatic tissues to investigate the methylation and expression patterns of candidate genes in the placental cohort. The combined results revealed that the promoter of *GoS2* contains a placenta-specific mDMR with maternal allele-specific methylation. The CpG island overlapping with the *GoS2* promoter is shared with the ncRNA *HSD11B1-AS1* (NR\_134511.1), which unfortunately could not be investigated further due to a lack of informative SNPs. Bidirectional TSSs have previously been observed at other imprinted DMRs, including *ZNF597* and *NAA60*, as well as the CpG island shared by *PEG10* and *SGCE* (449). Unexpectedly, we observed that 4 out of 8 placental samples exhibited maternal-specific expression for *GoS2*. Detailed investigations revealed residual maternal contamination in the bulk placental RNA. Since *GoS2* is highly expressed in maternal blood cells, magnitudes higher than in placental stromal cells, we believe this accounts for the observed maternal expression. Similar results were reported by Hamada and colleagues (21). Additionally, our placentae were investigated using STR analysis, which revealed no maternal contamination at the DNA level. However, the sensitivity of this analysis is limited to a 5% threshold. Maternal contamination hampers the accurate reporting of imprinted expression. Previous cases of maternal expression have been associated with maternal decidual contamination in mouse placentae (e.g., *Dcn* and *Gatm*) (541). While some maternally expressed genes, such as *Tfpi2* (542,543), within large co-regulated clusters, show expression in maternal decidua and placenta imprinting, many cases simply result from maternal contamination. Proudhon and Bourc'his proposed a genetic strategy to distinguish true maternal expression from maternal contamination based on the dam's genotype using inbred mouse strains (544). If heterozygous mothers are crossed with homozygous fathers, maternal contamination will always manifest as biallelic expression. Unfortunately, all mothers in our study were homozygous for the *GoS2* SNPs used.

Exploring publicly available datasets, we found that *GOS2* was predominantly expressed in placental HB cells and fibroblasts but only minimally in trophoblasts. The *GOS2* placenta-specific mDMR exhibited approximately 50% methylation in placental endothelial, stromal, and trophoblast cells (364). We confirmed these findings in placental trophoblast and stromal cell fractions obtained from the same placental sample using the MACS cell enrichment method. Furthermore, we investigated the methylation and expression of *GOS2* using pyrosequencing and qRT-PCR in our extended placental cohort, which included normal placentae from healthy pregnancies as well as samples affected by pregnancy complications such as PE, IUGR and SGA. We failed to observe significant changes in either DNA methylation or gene expression levels between the different groups. Consistent with previous findings (20–22,295), we identified a single sample with a complete absence of the mDMR, suggesting that the *GOS2* placenta-specific mDMR is polymorphic in the human population, but it is a rare event. Finally, we demonstrated that the *GOS2* placenta-specific mDMR is not conserved in the mouse placenta, and the gene showed no expression at E15.5 in the mouse placenta. Thus, we have characterised *GOS2* as a new polymorphic, placenta-specific imprinted gene exhibiting cell-type-specific imprinting in the human placenta.

We applied the same strategy to investigate allelic usage for *PIK3R1*, which encodes several isoforms originating from different promoters. The promoter of *PIK3R1* isoform 3 contained a placenta-specific mDMR, which is only 2 kb upstream of the isoform 2 TSS. Unfortunately, we were unable to investigate *PIK3R1* isoform 2 expression in our placental cohort due to its extremely low and tissue-restricted expression (545–547). Using the same techniques employed for *GOS2*, we found that *PIK3R1* isoform 1, the predominant isoform, was not imprinted in the human placenta. In contrast, isoform 3, which contains the placenta-specific mDMR within its promoter, exhibited maternal allele methylation. This type of isoform-specific imprinting, where an upstream promoter is biallelically expressed and intergenic transcripts originate from mDMRs, is widely observed at imprinted loci, including *MEST*, *GRB10*, *ERLIN2*, *RB1*, *ZNF331*, *WRB* and *SNU13* (449,548). It is likely that the transcription across the intergenic mDMR in oocytes is responsible for the establishment of the ICRs (549).

Further investigation into isoform-specific expression revealed preferential monoallelic expression of *PIK3R1* isoform 3, as most corresponding maternal samples were either heterozygous or not informative due to high heterozygosity. According to publicly available datasets, *PIK3R1* was expressed in all placental cell types; however, isoform-specific expression could not be determined from these datasets because they were generated using short-read sequencing (422). Additionally, the placenta-specific mDMR of *PIK3R1* isoform 3 was maintained in whole placental villi and trophoblast cells but not

in stromal, endothelial or HB cells, a finding we confirmed in trophoblast and stromal cell fractions from the same placental sample using the MACS cell enrichment method.

Moreover, we examined *PIK3R1* isoform 3-specific methylation and expression in our extended placental cohort, which included samples from both normal and complicated pregnancies. We observed no significant differences in DNA methylation or expression between the groups. We also found that the placenta-specific mDMR of *PIK3R1* was highly polymorphic within our placental cohort, with unmethylated individuals present in both controls and complicated pregnancies at similar frequencies. Finally, we demonstrated that the orthologous *Pik3r1* isoform 3 promoter lacked allelic methylation and was biallelically expressed. Collectively, our results indicate that *PIK3R1* exhibits highly polymorphic isoform-specific imprinting due to the presence of the oocyte-derived mDMR that is exclusive to placental trophoblasts.

Our group previously analysed placental WGBS data alongside methyl-seq data from human gametes, blastocysts, and other somatic tissues to identify placenta-specific mDMRs (20). To detect these regions, we performed a sliding window analysis, focusing on regions containing 25 CpGs with an average methylation between 20% and 80%. These regions had to be hypermethylated in oocytes and show intermediate methylation in blastocysts. This analysis uncovered 551 loci, including the placenta-specific DMRs of *GoS2* and *PIK3R1* isoform 3.

Two other research groups also investigated placenta-specific imprinting using different approaches to identify candidate regions. Hanna *et al.* (2016) (22) utilised Illumina HumanMethylation450K array data from triploid placental samples (diandric and digynic triploid pregnancies) that allowed to identify 882 DMRs. These DMRs were further compared with low-input RRBS datasets from human gametes and blastocysts, leading to the discovery of 101 novel mDMRs that met specific selection criteria. The analysed regions had to overlap CpG islands, be differentially methylated between gametes (>50%), and show intermediate methylation in blastocysts (15%–60%). Of these, 72 were determined to be placenta-specific mDMRs, exhibiting intermediate methylation in the placenta (25%–75%) and mainly being hypomethylated (<25%) in other somatic tissues. The placenta-specific mDMR of *GoS2* was included in this list, but the placenta-specific DMR of *PIK3R1* was not, likely due to differing stringency criteria or its polymorphic nature. Additionally, the Illumina HumanMethylation450K array covers only a small portion of the human genome, primarily targeting CpG-rich regions such as promoters, with limited probe coverage in distal genomic regions (550,551). Furthermore, in this study, the allelic expression of the genes associated with placenta-specific DMRs was not investigated further.

In another study, Hamada and colleagues used an immunomagnetic isolation method to obtain CTBs, primarily from first-trimester placentae, which were then subjected to WGBS and RNA-seq (21). These samples were analysed alongside methylation datasets from human gametes, blastocysts, and other embryonic and somatic tissues. A sliding window analysis was conducted to identify placenta-specific mDMRs, focusing on regions containing 20 CpGs that showed more than a 30% difference in DNA methylation between maternal and paternal alleles. These regions had to be hypermethylated (>80%) in oocytes, hypomethylated (<20%) in sperm, and show intermediate methylation in blastocysts, resulting in the discovery of 3,676 candidate mDMRs, including those of *GoS2* and *PIK3R1*. Although some regions were further validated by targeted bisulphite sequencing and RNA-seq, neither *GoS2* nor *PIK3R1* met the final criteria for paternal allele-specific expression (paternal allele >65% and maternal allele <35%). However, it is worth noting that *GoS2* was excluded from further analysis due to maternal contamination. Also, the authors reported that additional mDMRs could regulate allelic expression that simply did not pass the final stringent selection criteria.

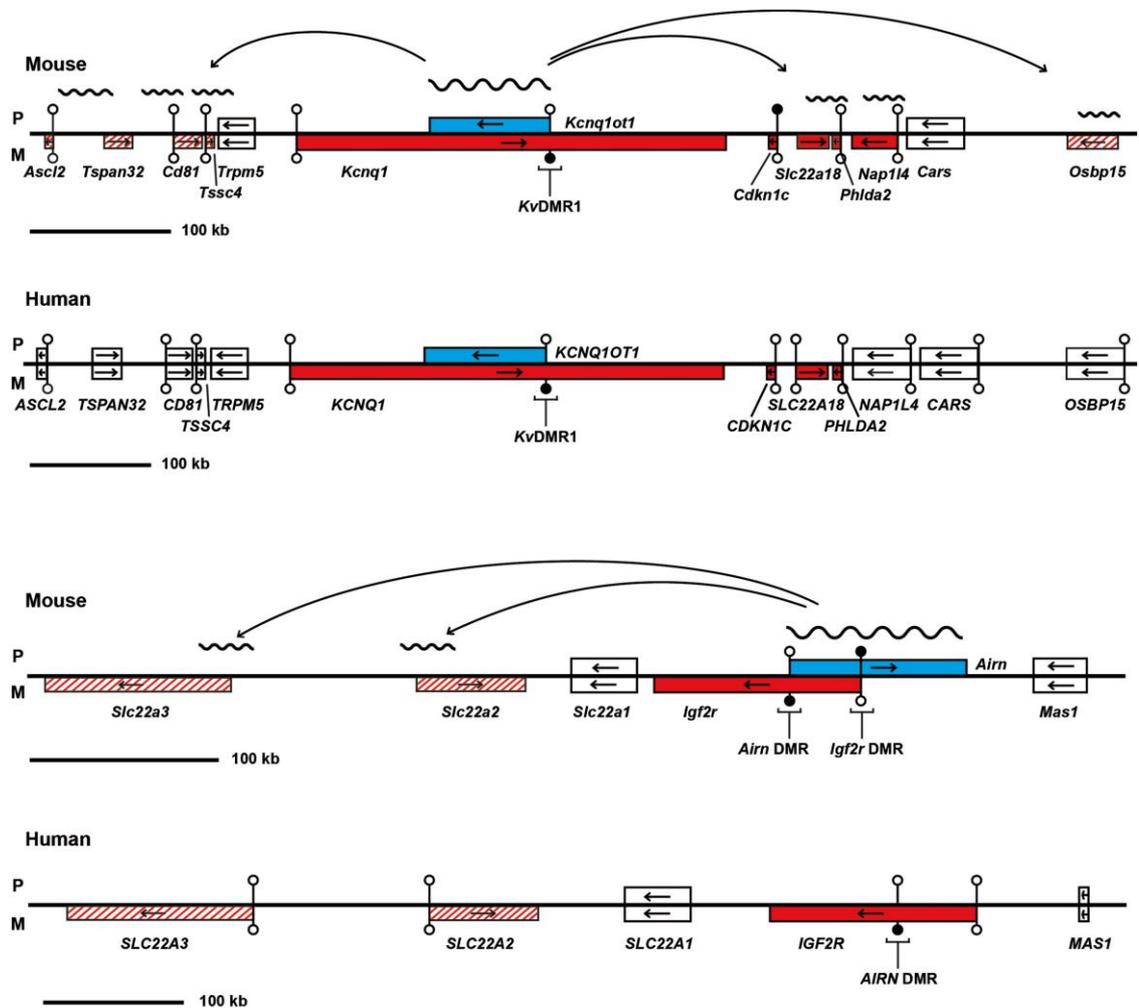
By investigating pyrosequencing and qRT-PCR datasets of *PIK3R1* isoform 3, I found a weak negative correlation (not significant) between the level of DNA methylation at the placenta-specific mDMR of *PIK3R1* and the expression of its isoform 3 in the AGA group (hypomethylated samples versus normal samples). As reported in the literature, DNA methylation at gene promoters is associated with gene repression (49), as often observed in cancer, where promoters of tumour suppressor genes become hypermethylated (369,370). Consequently, placental samples exhibiting hypomethylation at the placenta-specific mDMR overlapping the promoter of *PIK3R1* isoform 3 would be expected to show an upregulation of this isoform compared to samples maintaining normal methylation at this region. There can be several reasons why the degree of DNA methylation at the placenta-specific mDMR does not correlate with imprinted gene expression.

Several histone PTMs can influence the expression of imprinted genes in combination with DNA methylation. As noted earlier throughout this thesis, several groups, including ours, have found that a few individuals exhibit low levels of methylation at placenta-specific mDMRs, and thus, these regions are polymorphic (20–22,295). Additionally, they are restricted to human pre-implantation stages and the placenta (20,449). Consequently, Hanna and colleagues (22) speculated whether these regions could be remnants of embryonic methylation patterns or selectively protected from demethylation in the human placenta by ZFP57 and H3K9me2/3 (58,176,267). Further investigation suggested that placenta-specific mDMRs were slightly enriched for ZFP57 binding motifs and H3K9me3 compared to other somatic tissues and could, therefore, potentially contribute to the maintenance of these regions in the human placenta. Moreover, our group investigated

whether placenta-specific mDMRs could be associated with other histone PTMs that might account for the lack of methylation in these regions in some individuals (295). After extensive methylation profiling in our placental cohort, two polymorphic imprinted genes, including *LIN28B* (~12% of samples were hypomethylated) and *R3HCC1* (53% of samples were hypomethylated), had informative samples and were extensively analysed using ChIP and qRT-PCR for several permissive marks such as Histone 3 lysine 4 dimethylation (H3K4me2) and H3K4me3, and repressive marks such as H3K9me2 and H3K9me3. In samples that maintained the placenta-specific mDMR of *LIN28B* with paternal allele-specific expression, the paternal chromosome was decorated with permissive histone marks, and the maternal chromosome was enriched with repressive histone PTMs. Surprisingly, both repressive and permissive histone PTMs were observed on both parental chromosomes in samples that showed a loss of methylation at this DMR (295). An even more interesting case was observed for *R3HCC1*, where some samples maintained the placenta-specific mDMR, resulting in monoallelic expression, some maintained the mDMR but showed biallelic expression, and others showed biallelic expression along with loss of the mDMR. As in the previous example, in the sample that maintained the placenta-specific mDMR and showed paternal allele-specific expression, repressive and permissive histone marks were detected on the opposite chromosomes. In contrast, in the samples that demonstrated biallelic expression of *R3HCC1*, both repressive and permissive histone PTMs were detected on both parental chromosomes, with a higher enrichment of permissive histone PTMs (295). Therefore, an increase in permissive histone marks, irrespective of DNA methylation levels at the placenta-specific mDMR, could contribute to a more open chromatin configuration and biallelic expression.

Choux *et al.* (2020) reported that ubiquitous gDMRs, including *H19/IGF2*, *KCNQ1OT1* and *SNURF*, were hypomethylated in the placentae of children conceived through IVF or ICSI compared to those of naturally conceived children (552). An in-depth analysis of these regions revealed that changes in DNA methylation at these DMRs did not result in altered gene expression between the test and control groups. Profiling of histone PTMs showed a significant decrease in H3K9me2/3 at the *H19/IGF2 IC1* and *SNURF:TSS* DMR in the IVF/ICSI group relative to the natural conception group, along with a significant increase in H3K4me2 at the *H19/IGF2 IC1* and *KvDMR1*. Informative heterozygous samples showed enrichment of H3K4me2 and histone 3 lysine 9 acetylation (H3K9ac) on the methylated and repressed allele, but only at the *H19/IGF2* imprinting cluster. A similar study investigated *SNRPN* (553,554), *PEG10* (555) and *MEST* (372,556) - known paternally expressed imprinted genes with growth-promoting functions during embryonic and placental development and their associations with pregnancy complications (553). It was found that the expression of all three genes significantly decreased from the first trimester to the term placenta in normal pregnancies. These genes were also

downregulated in placentae affected by PE, especially *SNRPN*, while showing significant upregulation in samples from molar pregnancies. Profiling of epigenetic modifications in healthy samples revealed increased DNA methylation and H3K27me3 at all three DMR regions, with H3K9me3 specifically increased at the *PEG10* and *MEST* DMRs as pregnancy advanced. PE samples showed no significant change in DNA methylation but exhibited increased H3K27me3 at all DMRs, along with a notable increase in H3K9me3 at the *PEG10* DMR. In contrast, molar pregnancies were characterised by reduced levels of H3K27me3. These findings suggest that dysregulation of repressive histone PTMs and DNA methylation at the *SNRPN*, *PEG10*, and *MEST* DMRs could contribute to pregnancy complications such as PE or molar pregnancies. Furthermore, a large imprinting cluster located on distal chromosome 7 in the mouse genome is regulated by the maternally methylated *KvDMR1* (**Figure 5.1**) (291,371,372,514,515). This ICR lies within the maternally expressed *Kcnq1* transcript and functions as a promoter for the paternally expressed antisense lncRNA *Kcnq1ot1*. This lncRNA recruits G9A and interacts with PRC2 via HNRNPK, an RNA-binding protein, to silence several genes, such as *Ascl2*, *Cd81* or *Tssc4*, that flank the *Kcnq1* cluster in the mouse placenta (514,515,557,558). As a result, these genes exhibit maternal-biased expression uniquely in this embryonic organ. Taken together, as illustrated by these examples, histone PTMs play an important role in regulating imprinted gene expression in both the human and mouse placenta. Moreover, the loss of DNA methylation at gDMR regions may be compensated by the redistribution of histone PTMs such as H3K9me2/3 or H3K27me3, ultimately resulting in either preserved or altered imprinted gene expression in the placenta(295,552).



**Figure 5.1. Representation of the *Kcnq1ot1* imprinting cluster controlled by *KvDMR* in the mouse (chr 7qF5) and human (chr 11p15) genomes.**

The paternally expressed *Kcnq1ot1* transcript (blue) silences flanking genes within this imprinting cluster, resulting in maternal allele-specific expression (red). Red hatching indicates placenta-specific imprinted genes in the mouse but not in humans. Black arrows denote transcriptional orientation. Black lollipops represent methylated CpG islands, while white lollipops indicate unmethylated CpG islands. White boxes represent biallelically expressed transcripts. Adapted from Monk, 2015 (371).

To explore the expression and methylation profiles of the *PIK3R1* locus, I used RNA and DNA extracted from bulk placental samples. As shown by Vento-Tormo *et al.* (2018) and several other studies (328,329,364,365,553), the human placenta contains multiple cell types with distinct transcriptomes and epigenetic landscapes that undergo changes throughout gestation. A few reported placenta-specific imprinted genes, such as *THAP3* or *LIN28B*, exhibit polymorphic imprinting in humans (22,295). Interrogation of placenta-specific mDMRs using Infinium MethylationEPIC BeadChip array datasets from different placental cell types, generated by Yuan and colleagues (364), suggests that these mDMRs are primarily maintained in placental trophoblasts, in some cases, in endothelial or

stromal cells depending on the individual (**Appendix 25**). Therefore, these genes may be imprinted in specific placental cell types, similar to what we observed for *PIK3R1* isoform 3. Additionally, several studies investigating *PIK3R1* isoform expression have found that isoform 3 shows tissue-specific expression in both humans and rodents, while isoform 1 is the predominant transcript (545–547). Consequently, cell type-specific signatures may be lost in bulk samples containing RNA and DNA from all placental cell types (528,559). It is, therefore, possible that *PIK3R1* isoform 3 is upregulated in hypomethylated placentae, but this would not be detected in bulk RNA samples, particularly if this isoform is lowly expressed in placental trophoblasts.

The placenta efficiently mediates communication between the growing foetus and the maternal decidua and, therefore, exhibits high plasticity in its epigenetic and transcriptomic profiles (314,361,364,371,553,560). Changes in its epigenome and transcriptome can alter multiple placental functions, as noted throughout this thesis, allowing better adaptation to the increasing demands of the foetus during gestation. The placenta contains many imprinted genes with either growth-promoting or growth-restricting functions, some of which are part of a genetic growth network regulating embryonic development (371–373,560). This network is controlled by *PLAGL1* (also known as *Zac1*) (561), a paternally expressed gene that encodes a C2H2 ZNF and is regulated by the *PLAGL1/HYMAI* ICR on chromosome 6q24 (562). Overexpression of *PLAGL1* is associated with TNDM (420,563,564), a condition characterised by severe IUGR, hyperglycemia and other symptoms. Similarly, heterozygous and homozygous *Plagl1*-deleted mouse pups exhibit embryonic growth restriction (565). It has been suggested that in response to adverse intrauterine environmental exposures - such as smoking, chemical exposure or others - the placenta may alter the expression of genes belonging to this growth network to improve conditions for foetal development (317,361,371,560,561). Temporal alterations in the expression of co-regulated genes may include changes in nutrient transporter density, placental cell number, trophoblast migration and invasiveness, vascularization, and placental weight - all of which can modulate nutrient, gas and waste exchange (560). Such examples of placental adaptation have been observed in ART-conceived children and animal studies following ART procedures (560). For instance, one study investigated the effects of superovulation followed by embryo transfer at E3.5 on canonical genomic imprints in mouse embryos and placentae (566,567). *H19* was found to be biallelically expressed in a large proportion of E9.5 mouse embryos and placentae, while *Igf2* expression was significantly elevated in E9.5 placentae but remained monoallelic (566). However, by E14.5 and E18.5, normal monoallelic expression of *H19* and normal *Igf2* expression levels were observed, with no significant differences in placental or embryonic weight and size between ART and control groups (567). Similarly, another study reported that *IGF2* expression positively correlated

with birth weight and crown-rump length during the first trimester in humans, although this association was not observed at term (568). In addition, heterozygous or homozygous deletion of the Phosphatidylinositol 3-kinase (PI3K) p110 $\alpha$  catalytic subunit (*Pik3ca*) resulted in reduced surface area and foetal capillaries in E19 mouse placentae, but an increase in glucose transport, likely as compensation for impaired placental function to support foetal growth (569). Therefore, it is possible that hypomethylated samples at the *PIK3R1* placenta-specific mDMR may exhibit significant changes in isoform 3 expression during earlier stages of pregnancy but not near term, especially when this gene encodes a regulatory subunit of PI3K involved in growth-regulating and metabolic pathways such as *IGF2* and mTORC signalling (570,571).

Our group previously investigated changes in imprinted gene expression within the genetic growth network in placentae from IUGR cases or children conceived using ART. The non-coding transcript *HYMAI*, part of the *PLAGL1* imprinting cluster, was found to be upregulated in IUGR cases compared to controls (561). Also, *PLAGL1* was significantly downregulated in the placental samples from babies conceived with the help of ART. Interestingly, quantification of DNA methylation revealed no changes at the *PLAGL1/HYMAI* ICR. Therefore, it is possible that other *trans*-acting factors, such as TF, may modulate the expression of these genes independently of DNA methylation.

As shown by several studies, the placenta contains many DMRs, some of which are inherited from the germline, while others are established *de novo* uniquely in this transient embryonic organ (20,21,295,449). However, many of these DMRs do not induce monoallelic expression of nearby genes, as demonstrated in our previous work (20) and throughout this PhD thesis. Consequently, hypomethylation at such placenta-specific DMRs may not result in any changes in gene expression. It is, therefore, possible that these DMRs are not efficiently erased during early embryonic development and instead represent remnants that persist in the term placenta.

### 5.1.1. Functional role of *GOS2* during gestation and the placenta

*Gos2*, also known as Go/G1 switch 2, is a highly conserved gene in vertebrates (572,573). For instance, the human and mouse orthologs share 78% homology. This gene encodes a small protein of 103 amino acids that folds into two alpha helices separated by a beta-sheet. *Gos2* expression has been detected in various mouse and human cell types but is particularly abundant in metabolically active tissues such as adipose tissue, liver, heart,

and skeletal muscle. It has been implicated in a range of cellular processes, including proliferation, apoptosis, inflammation, and oxidative phosphorylation.

Early studies on *GoS2* suggested it was required for mononuclear hematopoietic cells to enter the G1 phase of the cell cycle (574,575). However, more recent research has shown that *GoS2* is required to maintain the quiescent state of hematopoietic stem cells (576). Despite this, *GoS2* is best known for its role in inhibiting adipose triglyceride lipase (ATGL), which is essential for the initial step of adipose lipolysis, converting triacylglycerols (TGs) into FAs and glycerol (572,573). *GoS2* inhibits ATGL by directly binding to its hydrolytic domain (HD) at the patatin-like region of ATGL. Notably, *GoS2* is highly specific to ATGL and does not affect the activity of other lipases, such as hormone-sensitive lipase (HSL), lipoprotein lipase (LPL), monoacylglycerol lipase (MGL), or lysophospholipases (577). Interestingly, the *GoS2* protein has a very short half-life of only 15 minutes, and it was suggested that ATGL may be required to stabilise it (578). Additionally, it has been reported that *GoS2* can bind to Bcl-2 at the mitochondria and disrupt the formation of the anti-apoptotic Bcl-2/Bax heterodimeric complex (579). Thus, while *GoS2* has diverse functions, some of which are not yet fully understood, it is critical for maintaining metabolic homeostasis.

The function of *GoS2* during pregnancy is not well understood, but a recent study has highlighted its role in recurrent spontaneous abortions (RSA). It was reported that in decidualised primary human endometrial stromal cells (HESCs) and decidualised, immortalised HESCs with *JAZF1* KO, increased apoptosis was observed due to the upregulation of *GoS2* (580). It was found that *JAZF1* could inhibit Pur $\beta$ , which, in the absence of *JAZF1*, was able to bind to the *GoS2* promoter and upregulate its transcription, leading to an increased rate of apoptosis. *GoS2* was also found to be upregulated in the decidua of women with RSA. In addition, the loss of *JAZF1* was also shown to impair the invasion of HTR-8/SVneo cells (a trophoblast cell line similar to EVT<sub>s</sub>). Overall, this suggested that *GoS2* plays a significant role in endometrial stromal decidualization and is implicated in miscarriage. Moreover, another study reported that *GoS2* was downregulated in circulating monocytes of pregnant women during the first trimester (581). The lower expression of *GoS2* was suggested to be important for the immune suppression required during pregnancy.

*GoS2* was investigated alongside ATGL and other placental lipases in pregnancies complicated by gestational diabetes mellitus (GDM) (582). Neonates born to women with GDM are often larger and have more body fat, potentially due to increased maternal lipid supply to the placenta and foetus. The study found that the mRNA level of *GoS2* remained unchanged in placentae affected by GDM compared to control samples. Interestingly, a more recent study explored the relationship between placental polar lipid composition

and birth outcomes (583). The authors collected 99 term placental samples for mass spectrometry and used placental RNA to analyse a panel of 30 genes alongside various maternal, neonatal, and placental measurements. They identified 75 placental polar lipids and performed PCA analysis. Only PCA2 was significantly associated with birth outcomes, including higher placental weight, birth weight, and neonatal lean mass. PCA2 explained 12.7% of the variation in placental lipid composition and was linked to acyl-alkyl-glycerophosphatidylcholines and lipid species containing DHA. PCA2 was also associated with the higher expression of several genes, including *GoS2*. These findings suggest that *GoS2* may influence placental polar lipid composition, possibly through interactions with ATGL, and influence placental and neonatal weight.

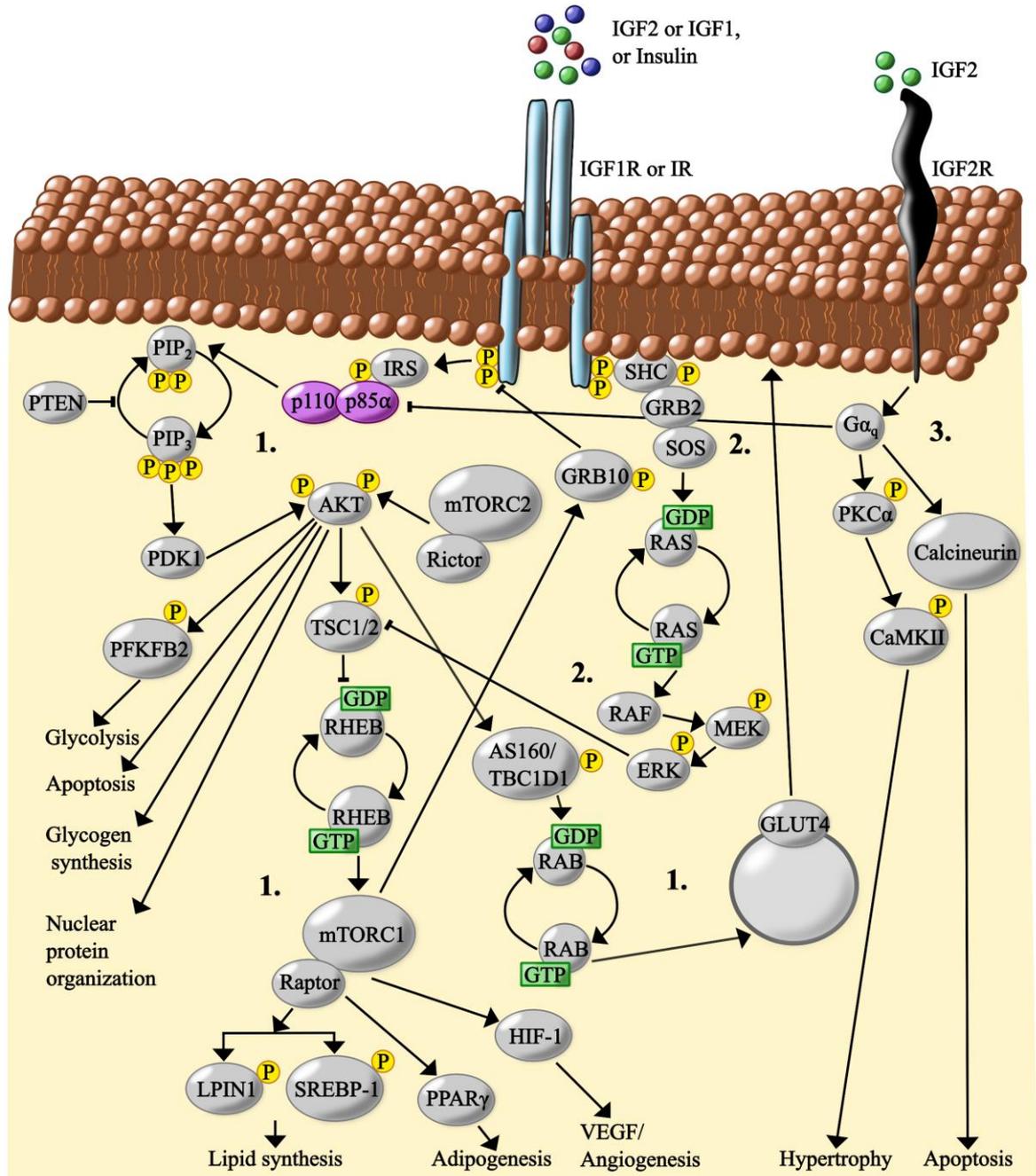
Finally, Bellazi *et al.* (2011) (584) investigated the role of placental alkaline phosphatase (*ALPP*) in the human placenta, as it is expressed throughout pregnancy, and some of it is secreted into the maternal circulation. Interestingly, they found that several genes, including *GoS2*, were upregulated following the overexpression of *ALPP* in HTR-8/SVneo cells. These cells exhibited increased proliferation and DNA synthesis. Further analysis of differentially expressed genes revealed that the upregulated genes were associated with cell signalling, proliferation, and growth pathways. In general, the authors suggested that the secretion of placental alkaline phosphatase may stimulate CTB proliferation, and it is possible that *GoS2* could also be associated with this process.

### 5.1.2. Functional role of *PIK3R1* during gestation and the placenta

The second candidate gene we characterised during this PhD project was *PIK3R1*. This gene encodes a regulatory subunit of class IA PI3K (570,585). *PIK3R1* produces several isoforms, as discussed earlier. The major isoform, isoform 1, is translated into p85 $\alpha$ , while isoforms 2 and 3 are splice variants translated into p55 $\alpha$  and p50 $\alpha$  subunits, respectively. Isoform 1 is widely expressed across various tissues and has been extensively studied in cancer and metabolic research. In contrast, isoforms 2 and 3 are restricted to a few tissues in humans and other mammals, such as muscle, brain, heart, and adipose tissue, and are generally understudied (545–547).

The p85 $\alpha$  regulatory subunit forms a heterodimer with the catalytic subunit, such as p110 $\alpha$  (though other catalytic isoforms like p110 $\beta$ , p110 $\gamma$ , and p110 $\delta$  also exist) (570,585). The regulatory subunit stabilizes PI3K and regulates the activity of the catalytic subunit.

p85 $\alpha$  contains two SH2 domains separated by an inter-SH2 domain at the C-terminus, as well as SH3 and Bcr homology (BH) domains at the N-terminus. These two N-terminal domains are absent in the p55 $\alpha$  and p50 $\alpha$  subunits encoded by isoforms 2 and 3 of *PIK3R1* (570,585). The inter-SH2 and SH2 domains are thought to bind to the p110 $\alpha$  subunit and inhibit its function (586). When an active tyrosine kinase receptor or adaptor protein like Insulin receptor substrate 1 (IRS-1) binds to the SH2 domains of p85 $\alpha$ , its inhibitory effect is relieved, allowing the p110 $\alpha$  subunit to move closer to the plasma membrane and convert phosphatidylinositol (4,5)-bisphosphate (PIP2) to Phosphatidylinositol (3,4,5)-trisphosphate (PIP3), a secondary messenger molecule (**Figure 5.2**) (570,585). The SH3 domain facilitates binding with cytoskeletal components, while the BH domain interacts with small GTPases (570,585).



**Figure 5.2. PI3K functions in different signalling pathways.**

(1) Activated insulin receptor (IR) or insulin-like growth factor 1 receptor (IGF1R) phosphorylates IRS(570,587), which then binds to PI3K (violet) (570,587). The release of the p110 subunit from the inhibitory effect of p85 $\alpha$  enables the conversion of PIP<sub>2</sub> into PIP<sub>3</sub>. This can initiate the AKT–mTORC1 signalling cascade (571), which drives various cellular processes, or promote glucose uptake via activation of the AKT–GLUT4 signalling pathway (588). mTORC1 can also phosphorylate GRB10 (589), which in turn inhibits IGF1R or IR. (2) Activated IR or IGF1R can also recruit and phosphorylate the SHC adaptor protein, which, with the help of GRB2 and SOS, activates the RAS–ERK pathway. This pathway can negatively regulate the AKT–mTORC1 signalling cascade. (3) IGF2 can additionally bind to IGF2R, which activates G $\alpha_q$  (590,591). This activation can inhibit PI3K or stimulate several other signalling pathways (592). Phosphorylation is indicated by yellow circles containing “P”. Green squares represent GTP-to-GDP conversion. PI3K is depicted as two subunits: the catalytic subunit p110 and the regulatory subunit p85 $\alpha$ , encoded by *PIK3R1* isoform 1. *IGF2* and *GRB10* are canonical imprinted genes. IR – Insulin receptor; IGF1R – Insulin-like growth factor 1 receptor; IGF2R – Insulin-like growth factor 2 receptor; IRS –

Insulin receptor substrate; PI3K – Phosphatidylinositol 3-kinase; PIP2 – Phosphatidylinositol (4,5)-bisphosphate; PIP3 – Phosphatidylinositol (3,4,5)-trisphosphate; AKT – Protein kinase B; mTORC1 – Mammalian target of rapamycin complex 1; mTORC2 – Mammalian target of rapamycin complex 2; GLUT4 – Glucose transporter type 4; Gαq – G protein alpha q subunit; GTP – Guanosine triphosphate; GDP – Guanosine diphosphate.

PI3K is a key intermediate signalling molecule in insulin and IGF2 signalling pathways (**Figure 5.2**) (570,585). Notably, *PIK3R1* isoform 3 may be an additional imprinted gene within the IGF2 signalling pathway alongside IGF2 (257,261,593), IGF2R (255,542), and GRB10 (**Figure 5.2**) (197,594). Consequently, it is not surprising that *PIK3R1* could be associated with growth-related conditions. Several *de novo* mutations in *PIK3R1* are associated with SHORT syndrome (Short stature, Hyperextensibility of joints/hernia, Ocular depression, Rieger anomaly and Teething delay), a rare genetic disorder characterised by short stature, joint hyperextensibility, ocular depression, Rieger anomaly, and teething delays (595–597). Mutations in *PIK3R1* have been suggested to cause insulin resistance and/or lipodystrophy. This syndrome shares many phenotypic similarities with SRS – a known imprinting disorder (420,597). One study investigated the mechanisms associated with advanced maternal age and accelerated placental ageing in PE (598). The expression of 307 genes linked to ageing was analysed using microarray datasets from 80 placental samples affected by PE and 77 normal samples. Of these, 58 genes were found to be differentially expressed between PE and normal samples. The top five differentially expressed genes included known PE-associated genes, such as *FLT1* and *LEP*, as well as *PIK3R1*. These five genes were incorporated into a diagnostic model for PE, which demonstrated good predictive ability. Thus, *PIK3R1* may be associated with accelerated ageing in placentae affected by PE.

Rosario *et al.* (2021) (599) explored the role of mammalian target of rapamycin complex 2 (mTORC2) in primary human trophoblast cells from normal placentae. After inhibiting mTORC2 with RICTOR siRNA (a key component of mTORC2), 307 genes were found to be upregulated and 102 downregulated, including *PIK3R1*. The upregulated genes were primarily involved in pro-inflammatory signalling pathways, such as VEGF-A, IL-6, leptin, and SAPK/JNK pathways, while downregulated genes were associated with multivitamin transport (*SLC5A6*) and angiogenesis (osteopontin). Additionally, osteopontin and *SLC5A6* were shown to be downregulated in IUGR cases following reduced mTORC2 activity. The authors suggested that mTORC2 inhibition may suppress the activity of osteopontin and PI3K, potentially contributing to placental insufficiency and reduced foetal growth in IUGR cases.

In animal models, a homozygous deletion of *Pik3ca* results in embryonic lethality (569). The heterozygous deletion of *Pik3ca* in the mouse placenta led to a reduction in maternal blood space and a decrease in foetal capillary and surface area. Increased apoptosis was also observed in the junctional zone of the affected placenta. Additionally, placentae with this deletion showed increased transfer of non-metabolizable glucose and neutral amino acids, although foetal growth was slightly restricted at later stages of development. Gene expression analysis revealed that downregulated genes were associated with processes such as cell death, proteolysis, immune regulation, cytolysis, and oxygen transport, while upregulated genes were associated with hormone metabolism. Interestingly, the deletion of *Pik3ca* in mouse trophoblast cell lines resulted in the reduced expression of beta-2 microglobulin (*B2m*), a component of the MHC I complex (600), while no other genes were significantly affected by this mutation. Hence, PI3K plays an important role in regulating placental nutrient supply, foetal growth, and overall placental development.

Several studies exploring the function of *PIK3R1* during gestation suggest that this gene is important for placental and embryonic development, as it can influence maternal nutrient supply to the foetus (569,598,599). However, the role of the imprinted isoform 3 in the human placenta remains unclear, as no studies have investigated the specific function of this isoform or its encoded protein during embryonic development or placentation.

### 5.1.3. Study limitations

To characterise the methylation and expression profiles of the candidate *GOS2* and *PIK3R1* genes, I used RNA and gDNA extracted from bulk term placental samples. As discussed earlier, investigating imprinted genes in bulk samples can be challenging, especially in the placenta, where genes might be imprinted in specific cell types. For example, *Ube3a*, a canonical imprinted gene conserved in both mice and humans, is maternally expressed in mouse neurons but shows biallelic expression in glial cells of the embryonic and postnatal mouse brain (601,602). Loss or abnormal expression of the maternal *UBE3A* allele causes AS (420). Another example is *GRB10*, a canonical imprint that demonstrates isoform- and tissue-specific expression (197). In the human embryonic brain, this gene shows paternal-specific expression, while in the placental villous trophoblasts, it exhibits maternal allele-specific expression, and biallelic expression is observed in other embryonic tissues. Genes uniquely imprinted in specific cell types may be undetectable in bulk samples or produce inconclusive or contradictory results, particularly if imprinting occurs in a rare cell population within the tested tissue or organ. As a result, such genes may be incorrectly dismissed as non-imprinted. This appears to be

the case for *PIK3R1*, which maintains its placenta-specific mDMR only in placental trophoblasts and is likely to exhibit paternal-specific expression exclusively in these cells. Therefore, it is not surprising that several samples showed biallelic expression of *PIK3R1* isoform 3 in bulk placental cDNA, especially considering that this isoform is likely expressed at low levels in the placenta. Consequently, genes that might be imprinted in specific cell types should be investigated using isolated cell populations or single-cell omics techniques.

As discussed earlier, several placental samples in our cohort demonstrated maternal-biased expression of *GOS2*, likely due to maternal contamination, which we confirmed in several samples via qRT-PCR. This is not surprising, as primates and rodents have hemochorial placentae, which are highly invasive compared to those of other mammalian groups (603). This invasiveness allows maternal cells to intermingle with placental cells, particularly in the intervillous space, where maternal blood flows from remodelled spiral arteries. As a result, several mouse genes have previously been falsely identified as imprinted in the placenta (299). Unfortunately, for our study, *GOS2* is also highly expressed in various blood cell types. Therefore, even a small number of maternal blood cells in the placental biopsy could lead to maternal contamination, as it was reported by Hamada (21).

To investigate the presence of placenta-specific mDMRs in different placental cell types, we applied a continuous percoll gradient to deplete contaminating blood cells and used the MACS cell enrichment method to obtain placental trophoblast and stromal cell fractions. Unfortunately, these methods have several limitations. Firstly, the percoll gradient does not entirely remove all blood cells from a cell suspension, requiring its combination with other techniques. Secondly, MACS columns rely on antibodies that must be highly specific to the targeted cells. If these antibodies lack sufficient specificity, they may bind to other cell types, leading to contamination in the positive cell fraction, which we observed in our samples. Quality control for cell-type enrichment suggested that the anti-EGFR antibody selected pan-trophoblasts, while the anti-fibroblast antibody enriched cells of mesenchymal origin (and strongly depleted trophoblasts). Since the procedure requires fresh samples and is extremely laborious (~10 hours from the time of delivery), only a subset of our placenta samples was processed. As a result, we were unable to confirm the monoallelic expression of *GOS2* and *PIK3R1* isoform 3 in different placental cell types due to the lack of informative samples.

To determine whether *GOS2* and *PIK3R1* isoform 3 might be imprinted in the human placenta, I applied a range of PCR-based methods followed by Sanger sequencing. Sanger sequencing is still widely used in forensic investigations and in clinical genetics

laboratories for STR or mitochondrial DNA analyses, diagnosing genetic diseases, and verifying SNPs and copy number variants (CNVs) identified through NGS techniques (604,605). It remains the gold standard for nucleic acid analysis, as it provides several advantages over other methods. Firstly, Sanger biochemistry enables the generation of long sequences, up to 1000 bp, with high accuracy (up to 99%) when good-quality samples and optimised conditions are used (606). It is cost-effective for the small number of samples, works well with low-quality DNA, follows a well-established protocol, and allows relatively straightforward data analysis using a wide range of available software tools.

However, like most laboratory techniques, Sanger sequencing has several limitations that have to be considered. It requires the selection of candidate regions with optimised sequencing primers, which prevents its use for genome-wide screening of novel imprinted genes. Most importantly, poor-quality samples, suboptimal PCR amplification or sequencing conditions, or the presence of inhibitors in the sample can lead to low-quality chromatograms, characterised by a high background signal, distorted or stutter peaks, or other sequencing artefacts that complicate interpretation (604,605). PCR amplification can also introduce base-composition bias by depleting loci with high or low GC content (>65% GC or <12% GC), through polymerase slippage and misincorporation of nucleotides in sequences with extreme base compositions or highly repetitive regions, and, in general, can result in uneven amplification in sequencing libraries (459,460). Additionally, PCR overamplification at polymorphic sites such as SNPs or CNV can result in preferential amplification of one allele or complete allelic dropout, potentially leading to the false identification of imprinted loci (461). A recent study used Sanger sequencing to confirm 866 high-quality SNPs identified in 825 clinical exomes generated by Illumina sequencing (607). Three SNPs identified in *NOTCH3*, *TPRN* and *C1QTNF5* were missed and appeared as homozygous in the Sanger chromatograms due to preferential amplification of one allele. Two of these SNPs, located in *NOTCH3* and *TPRN*, were later confirmed by redesigning the primers, while the third SNP in *C1QTNF5* could not be verified due to the complex nature of its transcript. Furthermore, 7 heterozygous SNPs were initially missed in 170 samples during the first round of Sanger sequencing due to issues with primers or PCR conditions. As a result, the authors decided to discontinue the use of Sanger sequencing for SNP validation to reduce associated costs and turnaround time, although they encouraged its continued use for CNV validation. Moreover, in a high-quality chromatogram, the height and width of a peak are generally proportional to the amount of DNA in the sample (604,605). At polymorphic sites such as a SNP, two peaks represent two alleles and can be used to roughly estimate allelic ratios. However, Sanger sequencing is less sensitive than NGS, which offers more accurate quantification of allele-specific expression. Additionally, peaks in chromatograms often vary in size, which can be caused by multiple factors, such as polymerase slippage at homopolymers (stretches of DNA or

RNA composed of one type of nucleotide), incorrect template concentrations, inaccurate reagent volumes, or suboptimal amplification conditions. In such cases, peak sizes do not represent the amount of molecules or alleles in the sample. Accurate allelic ratio estimates are important for studies of imprinting and XCI. A recent study investigated gene expression on active and inactive X chromosomes across 29 human tissues and 940 single-cell transcriptomes from GTEx, revealing that 23% of genes, or “escapes”, including *CHM*, *ZMAT1*, *NAA10* or *PIN4*, escape XCI in humans (608). Interestingly, escape genes in the non-pseudoautosomal (nonPAR) region of the X chromosomes showed higher expression in females, while escapes located in the pseudoautosomal region PAR1, upstream of the nonPAR region, exhibited male-biased expression. On average, escapes on the inactive X chromosome were expressed at only 33% of the level observed from the active X chromosome. Thus, although Sanger sequencing remains a reliable and accurate method, it is not a sufficiently quantitative or sensitive technique, particularly for ASE analyses of imprinted genes or XCI, where precise estimates of allelic dosage are critical. In such cases, NGS or single-cell sequencing techniques are more appropriate, especially for analyses focused on specific cell types or tissues.

Our results showed no significant differences in DNA methylation levels at the *GOS2* and *PIK3R1* placenta-specific mDMRs, nor in their mRNA levels, across our extended placental cohort, which included AGA, IUGR, PE, and SGA cases. Firstly, our test groups were not well balanced, with the majority of samples belonging to the AGA group, while only a few samples were in the PE group (**Appendix 3**). Secondly, the differences in DNA methylation or gene expression may be quite subtle. In addition, studying placenta-specific imprinted genes is challenging due to their highly polymorphic nature (20–22,295), as seen with the placenta-specific mDMR of the *PIK3R1* isoform 3. They are also largely not conserved in other species, with some exceptions among primates (20,449). This lack of conservation limits experiments in model systems such as mice, thereby restricting the options for investigating genes with placenta-specific imprinting. Therefore, a larger and more balanced placental cohort is necessary to detect any potential differences in DNA methylation at the placenta-specific mDMRs or in their regulated gene expression, if such differences even exist. Members of the Monk group are currently quantifying these genes in a second, larger cohort of AGA versus IUGR samples obtained from the Baby Bio Bank repository (609). Additional experiments are planned to expand this to severe PE and GDM.

#### 5.1.4. Future research

In future studies, the placental cell-type-specific imprinting, including *GoS2* and *PIK3R1* isoform 3, should be explored using more advanced techniques. To minimize maternal cell contamination and assess cell-type-specific imprinting, placental biopsies from the second and third trimesters could be subjected to FACS (364). This technique allows for the isolation of distinct cell populations based on antibodies, as well as cell size and granularity. Isolated populations, such as trophoblasts, fibroblasts, endothelial cells, and HB cells, could then be used for DNA and RNA isolation.

These samples could be analysed using techniques employed throughout this PhD project, including bisulphite PCR and sub-cloning, as well as allelic RT-PCR followed by Sanger sequencing. Similarly, RNA from different placental cell types could be subjected to long-read sequencing methods, such as PacBio Iso-Seq or Nanopore direct RNA sequencing (DRS) (610,611). These techniques would provide insights into isoform-specific expression. This approach would enable a detailed investigation of other placental-specific imprinted genes and their isoforms for cell-type-specific imprinting, as in the case of *DNMT1* (612), whose placenta-specific mDMR is maintained across all placental cell types but not in HB cells (364).

The gene expression and functions of the placenta-specific mDMRs of *GoS2* and *PIK3R1* isoform 3 could be further explored using hTSCs. Parent-of-origin-specific expression of *GoS2* and *PIK3R1* isoform 3 could be investigated in CT and Mole cell lines. As noted in the introduction chapter, CT cell lines (243) are a good model for studying major placental cell types and placenta-specific imprinting, as most of these imprints are retained. In contrast, Mole cell lines (311), which are established from CHM pregnancies, lack genomic imprints (based on in-house EPIC data). Therefore, RNA from these two cell lines could be analysed using qRT-PCR. The expression of *GoS2* and *PIK3R1* isoform 3 should be upregulated in Mole cell lines. Consistent with this, in-house Illumina Infinium MethylationEPIC array data revealed that *PIK3R1* and *GoS2* mDMRs are present in CT27 cells and show relative hypomethylation in the Mole 1 line. This would provide additional confirmation that they are imprinted in placental trophoblasts and would warrant deeper epigenetic profiling, including histone modifications.

To explore the function of the placenta-specific mDMRs of *GoS2* and *PIK3R1* isoform 3, various CRISPR-Cas9 systems could be employed to modify methylation patterns in CT cell lines (613). For instance, a recent study utilised a dCas9-GCN4 fusion protein with gRNA to transiently demethylate the IG-DMR (hypermethylated paternal allele) of the *Dlk1-Dio3* region (614). This transient depletion of methylation at the ZFP57 binding site within the IG-DMR led to the upregulation of *Meg3* and *Mirg* in mESCs. Therefore, if the placenta-specific mDMRs of *GoS2* and *PIK3R1* isoform 3 are functionally significant, their

transcripts should be upregulated in CT cells following hypomethylation of these mDMRs. Reciprocal experiments could use the recently described iCRUSH system to deposit repressive histone modifications and silence the active alleles of *PIK3R1* and *GOS2* (615). In addition, CT cells might exhibit observable phenotypic changes as a result of these modifications.

Additionally, to further investigate the functions of *GOS2* and *PIK3R1* isoform 3, CRISPR-Cas9 could be used to delete essential exons/transcripts in expressing cell lines. For *GOS2*, targeting the 3' end of the gene could be a strategy to avoid disrupting the *HSD11B1-AS1* ncRNA (NR\_134511.1) or other adjacent transcripts. For *PIK3R1* isoform 3, the promoter region could be deleted, as it resides within the intron of *PIK3R1* isoform 1. If these transcripts are functionally important, CT cells will exhibit phenotypic changes and potential alterations in other gene expression.

### 5.1.5. Conclusions

In this project, I revisited candidate placenta-specific mDMRs previously discovered by our group and identified two promising regions that overlap the promoters of *GOS2* and *PIK3R1* isoform 3. To investigate these placenta-specific mDMRs in our placental cohort, I applied several techniques, including methylation-sensitive genotyping, bisulphite PCR and sub-cloning, and allelic RT-PCR. The combined results of these methods revealed that the placenta-specific mDMRs of *GOS2* and *PIK3R1* isoform 3 were polymorphic and exhibited maternal allele-specific methylation, with both genes showing preferential monoallelic expression.

For *GOS2*, many samples demonstrated preferential maternal expression, most likely due to residual maternal contamination, as this gene is highly transcribed in maternal blood cells. By exploring the Human Protein Atlas (422), I found that *GOS2* is predominantly expressed in immune cells, fibroblasts, and, to a lesser extent, trophoblasts, while *PIK3R1* is expressed across all placental cell types. Further analysis of the Illumina Infinium MethylationEPIC array datasets generated by the Robinson group (364) revealed that the placenta-specific mDMR of *GOS2* was unique to endothelial, stromal, and trophoblast cells in the third-trimester placenta, whereas the mDMR of *PIK3R1* isoform 3 was exclusive to placental trophoblasts. We confirmed these observations in placental stromal and trophoblast cell fractions obtained through continuous percoll gradient separation and MACS cell enrichment techniques.

I also investigated the methylation and expression of *GoS2* and *PIK3R1* isoform 3 in our extended placental cohort, including placentae from normal pregnancies and those affected by IUGR, PE, and SGA. Unfortunately, no significant differences in methylation or expression were found between the control and affected groups. Additionally, I observed a few hypomethylated samples within the control group, further highlighting the polymorphic nature of these placenta-specific mDMRs in humans. Finally, I showed that these placenta-specific mDMRs are not conserved in the mouse placenta.

Collectively, I identified two genes with placenta-specific mDMRs that exhibit cell type-specific imprinting and, in the case of *PIK3R1*, isoform-specific imprinting. While the roles of these genes in placental function remain unclear, *GoS2* may play a role in placental lipid composition and cell proliferation, whereas *PIK3R1* could regulate maternal nutrient supply and may be associated with accelerated placental senescence in PE cases. Notably, *PIK3R1* isoform 3 has not yet been thoroughly studied in the context of development. Therefore, further research is needed to explore the functions of *GoS2* and *PIK3R1* isoform 3 in human placental development and pathology.

## 5.2. Non-canonical imprinting might not be conserved in humans

Many studies have shown that canonical imprinted genes, controlled by ICRs, are highly conserved across most mammalian species (18,616). For example, the *H19/Igf2* cluster is also imprinted in marsupials, which diverged from eutherian mammals approximately 160 million years ago (250,617). Genome-wide screenings for imprinted genes across multiple tissues have supported previous findings that canonical imprinted genes are highly conserved between mice and humans (307–309). Additionally, these ICRs are robust and consistently found across various tissues. Surprisingly, it was recently discovered that extra-embryonic tissues have a higher number of genes with monoallelic expression (307,309). Following these findings, several groups, including ours, have demonstrated that many imprinted genes in humans are controlled by maternal gDMRs, which tend to be polymorphic (20–22,295). It was also shown that these placenta-specific imprinted genes are not conserved in mice or other mammalian species (20), except for primates. This suggests that species-specific differences exist, particularly in extra-embryonic tissues, and that imprinting may evolve more rapidly in these tissues.

More excitingly, several groups have described H3K27me<sub>3</sub>-mediated imprinting, also referred to as non-canonical imprinting, in mice and rats (24–26,187,450,511). In this form of imprinting, several genes, often located near rodent-specific LTRs, become decorated with broad H3K27me<sub>3</sub> domains derived from the oocyte. In post-implantation embryos, these domains are replaced by sDMRs, leading to paternal-specific expression in the rodent placenta. Several studies have explored the function of non-canonically imprinted genes during mouse development, revealing that aberrant expression of these genes can impact both embryonic and placental development. For instance, paternal deletion of *Slc38a4*, an amino acid transporter, results in reduced spongiotrophoblast and labyrinth layers in the placenta, decreased mouse pup weight, and frequent death shortly after birth (452). Additionally, the loss of non-canonical imprinting at a miRNA cluster within *Sfmbt2* leads to placental enlargement in mouse embryos derived through SCNT (618,619). These findings underscore the importance of such genes for normal rodent development.

To date, several studies have attempted to identify non-canonically imprinted genes in human embryos, but the results have been largely conflicting (27,136,224,517). Therefore, in this study, we aimed to discover non-canonically imprinted genes in human pre-implantation embryos and the placenta.

In this project, we analysed several methylation datasets from human gametes, blastocysts, and various somatic tissues, including the placenta, to identify placental sDMRs (regions hypomethylated in gametes and blastocysts but partially methylated in the placenta). First, we characterised 11 orthologs of mouse and rat non-canonical imprinted genes, such as *Gab1*, *Slc38a1*, *Slc38a4*, *Sfmbt2* and several others, using different molecular biology techniques to assess allelic expression and DNA methylation in our placental cohort. We found that the promoters of these genes were mainly unmethylated in the placenta and exhibited clear biallelic expression across multiple samples. This expression pattern was also observed in our pre-implantation embryo cohort.

*Xist* in mice is classified as a non-canonical imprinted gene, as the Xp is exclusively inactivated in the early embryo, TE, PrE and placenta (303,304,620). We investigated *XIST* methylation and expression patterns, with a primary focus on human placental samples. Our results showed that the P2 promoter of *XIST* was fully methylated in males, while in females, it exhibited allelic methylation. This sex-specific regulation was confirmed through qRT-PCR, which revealed almost no *XIST* expression in male placental samples but high levels of *XIST* expression in female samples. Allelic RT-PCR also revealed that *XIST* was biallelically expressed in female placental samples, and this

pattern was also observed in one female embryo (CL 3), consistent with random XCI in humans.

Given that non-canonical imprints in mice and rats are often located near murine-specific LTR elements (25), we screened a recent list of primate-specific LTRs, which are reported to be highly active in human oocytes (235). We identified four candidate genes with LTR-derived promoters, all of which exhibited a variable degree of methylation on both parental alleles in placental samples. Using RT-PCR across informative SNPs, we were able to show that one LTR-driven chimeric transcript was biallelically expressed. Using the same approach, we also examined previously suggested candidate genes, including *FAM101A*, discovered in human morulae (27), and *C5ORF38*, proposed as a candidate for non-canonical imprinting in term placentae (517). Both genes demonstrated biallelic methylation, while different isoforms of *FAM101A* showed biallelic expression in our placental cohort.

We further explored our previously identified 722 partially methylated regions (~50% methylation) in the placental WGBS dataset through a sliding window analysis (449). Interestingly, many of these regions contained potential placental sDMRs that were not derived from gametes. After a detailed investigation, we identified 14 promising candidate genes, all of which exhibited varying degrees of DNA methylation on both parental alleles. Among these, we followed up on one candidate gene, which revealed biallelic expression. The inability to identify sDMRs was not due to our molecular approach, as methylation-sensitive genotyping and allelic bisulphite PCR identified eight *bona fide* genes with placenta-specific mDMRs, three of which had been previously reported by Hamada (21). Finally, we took advantage of newly established hTSCs (CT27, CT30) derived from first-trimester placental samples, as well as trophoblast cells derived from CHM pregnancies (Mole 1, Mole 2). Initial investigation of the methylation profiles using the Illumina Infinium MethylationEPIC array in these cells revealed largely stable imprinting in biparental CT cell lines, while Mole cell lines lacked DMRs due to their AG nature. Using these cell lines, we demonstrated that most orthologs of mouse and rat non-canonical imprinted genes were not differentially expressed between the two cell lines. However, the expression of the *PEG10*, a canonical imprinted transcript, and *DNMT1*, a placenta-specific imprinted gene, was upregulated approximately 2-fold in the Mole cells. In conclusion, our findings suggest that non-canonical imprinting is unlikely to be conserved in human extra-embryonic tissues.

Epigenetic reprogramming is highly conserved among placental mammals; however, some species-specific differences have been noted, which may explain why non-canonical imprinting is not conserved in humans - a topic I will discuss further below.

In rodent oocytes and early pre-implantation embryos, non-canonical distributions of H3K27me<sub>3</sub> and H3K4me<sub>3</sub> are observed. Multiple studies have reported that mouse and rat oocytes harbour broad, distal, non-canonical domains of H3K27me<sub>3</sub> and H3K4me<sub>3</sub>, often overlapping with large PMDs (131,136,137,139,219). These domains are inherited by the zygote and retained throughout the pre-implantation stages that later are replaced by canonical distributions of these marks. As rodent embryos progress past implantation, the H3K4me<sub>3</sub> domains shrink significantly, becoming restricted mainly to promoter regions or bivalent domains - regions marked by both H3K4me<sub>3</sub> and H3K27me<sub>3</sub> (136,144,193,194). Similarly, H3K27me<sub>3</sub> domains become smaller and are primarily localized at developmental gene promoters, enhancers, or bivalent domains. Non-canonical imprinted genes are typically located within these non-canonical H3K27me<sub>3</sub> domains, which, in post-implantation embryos, are replaced by DNA methylation, forming sDMRs (24–26). Concurrently, the paternal chromosome often becomes enriched with H3K4me<sub>3</sub>. Interestingly, even between closely related rodent species, some differences in H3K27me<sub>3</sub> and H3K4me<sub>3</sub> patterns have been noted during the pre-implantation stages (450). For instance, rat oocytes exhibit larger H3K27me<sub>3</sub> domains than mouse oocytes, while broad H3K4me<sub>3</sub> domains are larger in mice.

These distinctive patterns in rodents contrast significantly with the distributions of H3K27me<sub>3</sub> and H3K4me<sub>3</sub> observed in human pre-implantation embryos, which I will discuss in more detail later. Understanding these differences may provide insights into why non-canonical imprinting, prevalent in rodents, is not conserved in human extra-embryonic tissues.

As previously discussed, several groups have attempted to identify non-canonical imprinted genes in humans, leading to intriguing but sometimes contradictory findings. In the first study by Zhang and colleagues, RNA-seq and CUT&RUN were used to profile the distribution of H3K27me<sub>3</sub> in human morulae from five couples (27). By identifying SNPs from WES or whole-genome sequencing datasets of maternal cumulus cells, they were able to identify 44 paternally expressed genes, which included well-known imprinted genes such as *PEG10* and *SNRPN*. When profiling H3K27me<sub>3</sub> distributions, they observed that the promoters of five paternally expressed genes - *DUSP4*, *EDNRB*, *ERO1LB*, *FAM101A*, and *MAGEB2* - were colocalised with H3K27me<sub>3</sub> domains. Unfortunately, due to low heterozygosity in the collected samples, they only identified *FAM101A* as having a maternally derived H3K27me<sub>3</sub> domain, suggesting that this gene might be regulated by H3K27me<sub>3</sub>-mediated imprinting. The authors concluded that additional paternally expressed genes could be controlled by maternally derived H3K27me<sub>3</sub> domains, warranting further investigation into the human genome.

In a subsequent study (136), key histone modifications were profiled across human GV oocytes, 2-cell, 4-cell, and 8-cell stage embryos, as well as in the ICM and sperm, using CUT&RUN. The researchers observed a gradual decrease in H3K27me3 levels after fertilisation, with the lowest levels in 8-cell stage embryos. In contrast, in the ICM, H3K27me3 was predominantly located at canonical PRC2 targets. When examining whether mouse non-canonical imprinted genes were conserved in humans, they found that 11 orthologous genes had H3K27me3 domains in human GV oocytes and the ICM but not in 8-cell stage embryos. Of these, only *OTX2* was expressed in early human embryos. They also revisited the 44 paternally expressed genes identified in the earlier study and found that just four of these genes were enriched with H3K27me3 in GV oocytes or ICM, but no such domains were present in 8-cell stage embryos. They suggested that alternative molecular mechanisms, aside from H3K27me3-mediated imprinting, might regulate paternal expression in humans, prompting further investigation.

In the most recent study (224), haploid AG and PG human embryos were examined at the blastocyst stage to profile DNA methylation, gene expression, open chromatin, and H3K27me3 distribution. Surprisingly, upon examining the distribution of H3K27me3, the researchers discovered that AG-blastocysts exhibited a significantly higher number of H3K27me3 domains compared to PG-blastocysts. Despite this, the majority of genes associated with AG-specific H3K27me3 domains at their promoters were not expressed at this stage. The authors further found that 7 genes with AG-specific H3K27me3 domains at their promoters demonstrated PG-specific expression and vice versa. Importantly, they observed that human genes harbouring AG- or PG-specific H3K27me3 domains were not conserved in mice. Moreover, none of the 76 previously identified mouse candidate non-canonical imprints showed PG-specific enrichment of H3K27me3 in humans, leading the authors to conclude that non-canonical imprinting is unlikely to be conserved in humans.

In contrast to rodents, in human oocytes and pre-implantation embryos, H3K4me3 and H3K27me3 form much smaller canonical domains. In human pre-implantation embryos, H3K4me3 becomes reduced just before EGA, while H3K27me3 is mostly depleted between the 4-cell and 8-cell stages and re-established between the morula and blastocyst stages (136,137). This creates a pattern that is markedly different from the H3K27me3 distribution seen in mouse oocytes. Moreover, the core proteins of the PRC2 complex, which are responsible for depositing H3K27me3, are expressed immediately after EGA in humans, whereas in rodents, these genes are expressed throughout pre-implantation development.

As discussed in the introduction, several histone modifications are implicated in non-canonical imprinting. Two studies utilised CUT&RUN and ChIP-seq, combined with RNA-

seq, to investigate the profile of H2AK119ub1 established by PRC1 in mouse oocytes and different-stage embryos (130,132). Both studies demonstrated that H2AK119ub1 forms broad domains that colocalise with H3K27me3, with most H2AK119ub1 regions originating from oocytes. It was also observed that H2AK119ub1 is erased at the 2-cell stage and re-established by the morula stage, exhibiting a distribution different from that in the zygote.

Mei and colleagues deleted the PCGF1 and PCGF6 components of the non-canonical PRC1 complex in mouse FGOs, resulting in a significant reduction of H2AK119ub1 in these FGOs. They found that a subset of genes lost H3K27me3 domains in *Pcgf1/6* KO FGOs, and this loss was inherited by matKO embryos (132). These regions included 16 non-canonical imprinted genes. By the morula stage, 9 of these non-canonical imprints became biallelically expressed, whereas DNA methylation-dependent imprints maintained their allele-specific expression. At E6.5, six non-canonical imprinted genes showed either a complete LOI or a milder allelic bias in the extra-embryonic ectoderm. Additionally, matKO morulae exhibited biallelic and monoallelic expression of *Xist* in female and male embryos, respectively. RNA-seq data indicated repression of genes on the maternal X chromosome, suggesting an aberrant maternal XCI, though normal XCI was established by E6.5.

A different approach was used in another study to remove H2AK119ub1 from developing mouse embryos (130). Mutant mRNAs of *Bap1* and *Asxl1*, a part of the Polycomb repressive deubiquitinase (PR-DUB) complex, were co-injected into mouse zygotes. In 4-cell stage embryos, the level of H3K27me3 remained mostly unaffected, and only some distal genomic regions lost H3K27me3 following H2AK119ub1 depletion. Importantly, non-canonical imprints retained their H3K27me3 domains and showed no changes in expression. However, these embryos experienced embryonic arrest at the 4-cell stage, likely due to the premature expression of early developmental genes resulting from the depletion of H2AK119ub1.

Based on these studies, PRC1 appears to be crucial for the establishment of non-canonical imprinted genes in developing oocytes by potentially recruiting PRC2 to these genomic regions. However, PRC1 may not be required in growing embryos where non-canonical H3K27me3 domains are already established. Thus, the necessity of PRC1 for non-canonical imprinting remains unclear and warrants further investigation. Unfortunately, datasets for human gametes or pre-implantation embryos are not yet available for direct comparison, but it is possible that H2AK119ub1 follows a similar pattern in human embryos.

Additionally, G9A, also known as EHMT2, is an H3K9 methyltransferase responsible for catalysing H3K9me2, which is essential for repressing ERVs in mouse embryos (515). H3K9me2 is enriched at certain mouse ICR elements and its depletion results in the upregulation of several imprinted genes within the *Kcnq1* domain in the mouse placenta (621). EHMT2 can recruit *de novo* methyltransferases to hypermethylate H3K9me2-decorated genomic regions. A recent study involving homozygous and maternal heterozygous mouse mutants with a deleted SET catalytic domain of EHMT2 found that all non-canonical imprinted genes were upregulated in the EPCs of homozygous mutants, while only some imprints were lost in maternal mutants (521). This suggests that embryonic EHMT2 may be a critical factor for repressing ERVs associated with non-canonical imprinted genes in mouse extra-embryonic tissues. Further investigation into the role of EHMT2 is needed.

Most mouse and rat non-canonical imprinted genes are associated with Muridae-specific LTR elements that are important for imprinting. These genes are often near solo-LTR elements, specifically ERVKs, which share the same orientation with non-canonical imprinted genes (25,450). For example, in the case of *Gab1*, the RLTR15 element within the first intron of this gene acts as an alternative promoter. RLTR15 is marked by H3K4me3 on the paternal chromosome, protecting it from *de novo* DNA methylation, while RLTR15 becomes hypermethylated on the maternal chromosome and forms the sDMR in the mouse ExE. Interestingly, the alternative *Gab1* isoform driven by RLTR15 is more highly expressed than the major isoform transcribed from the primary promoter. Mosaic paternal deletion of RLTR15 results in partial *Gab1* alternative isoform LOI in the mouse placenta, indicating the importance of this LTR element for *Gab1* imprinting (25). This LTR element is also present in the rat, where *Gab1* shows monoallelic expression in extra-embryonic tissues (450). During this PhD project, we investigated *GAB1* in human embryos and placentae and found that the RLTR15 element present in rodents is not conserved in humans. We also observed that the smaller isoform of *GAB1* is not imprinted in human pre-implantation embryos or placentae, suggesting that this ERVK LTR element is seemingly required for *Gab1* imprinting in rodents.

Bogutz *et al.* (2019) (236) explored the role of species-specific LTR elements in imprinting across humans, primates and rodents. They found that several human imprinted genes, such as *RHOBTB3*, *GLIS3*, *MCCC1*, *ST8SIA1* and others, are located near LTR elements specific to the Catarrhini or Hominoidea lineages. These genes exhibit imprinting in the human placenta but are not imprinted in mice, as such LTRs are absent from the mouse genome. During this PhD, we also examined the potential association between primate-specific LTRs and non-canonical imprinting. We investigated 1,031 recently identified primate-specific LTR elements located on autosomes that are actively transcribed in

human GV and MII oocytes, forming chimeric LTR fusion transcripts with nearby genes (235). We identified four candidate genes, primarily encoding ncRNAs, with promoters colocalised with these LTRs. Surprisingly, our results showed that both parental alleles were methylated to some extent, and only one ncRNA exhibited biallelic expression in the human placenta. Notably, all the LTRs associated with imprinted gene expression in the placenta, as reported by Bogutz and colleagues (236), were found upstream of gene promoters or within intronic regions, likely functioning as alternative promoters. Transcription initiated from such LTR elements would result in the deposition of transcription-coupled H3K36me3 (159) and subsequent high levels of DNA methylation by DNMT3A-DNMT3L in developing oocytes (282), potentially contributing to the formation of new maternal gDMRs at CpG islands overlapping these transcribed regions (235,236,240,549). This raises questions about the differences between our tested LTR elements and those previously reported and whether primate-specific LTRs mediate non-canonical imprinting in humans.

Several elegant studies have demonstrated that mammalian placentae contain numerous distal *cis*-regulatory elements that are frequently enriched with species-specific transposable elements (138,622,623). As noted in the introduction, the placenta is unique in many ways and is likely one of the most diverse organs across eutherian mammals (603). To understand the potential causes of this diversity observed in this transient embryonic organ, Chuong and colleagues investigated the regulatory landscape of mouse and rat trophoblast stem cells (rTSCs) (622). They found that enhancers were often species-specific, marked by permissive histone PTMs such as H3K27ac and/or H3K4me1, and located near genes functionally important for placental development. In contrast, the promoter regions of placental genes were generally highly conserved between closely related species. Notably, many species-specific enhancers contained species-specific ERVs, with the RLTR13D5 element frequently detected in mTSCs but not in rTSCs. This element includes binding motifs for key TFs in mTSCs, including *Cdx2*, *Eomes* and *Elf5*. Transfection of RLTR13D5 along with the mouse-specific *Apoceb3* gene into Rcho-1 rTSCs led to significant upregulation of this gene, as demonstrated by luciferase assay. Interestingly, these species-specific ERV enhancers were unique to the placenta and were not detected in other somatic tissues. Other tissues generally harboured enhancer elements composed of more ancient transposable elements shared among rodents and humans. Only other hypomethylated tissues, such as testes, also contained species-specific ERV enhancers. Similar findings were reported by Sun *et al.* (2021), who compared late-gestation placentae from humans, macaques and mice (623). Their study identified many placenta-specific genes located near enhancer elements unique to the human placenta, which were significantly enriched for various classes of transposable elements, including SINEs, LINEs, DNA transposons, and most notably, ERV LTRs. Similarly, the chorion was

found to contain many species-specific enhancers with human-specific ERV LTRs. A more recent study by Frost and colleagues confirmed these observations by examining enhancer elements containing human-specific ERVs in CTBs and hTSCs (138). They found that, unlike the majority of enhancers decorated by H3K27ac and/or H3K4me1, a small subset of ERV-containing enhancers was marked by H3K4me3. Although some of these enhancer/ promoter elements were located near key placental genes such as *CYP19A1*, *PTN* or *PRL*, most were found adjacent to transcripts with low expression in CTBs. Furthermore, it was observed that a shorter distance between a gene and an ERV-containing enhancer correlated with stronger regulatory activity of the element. Based on these findings (138,143,375,622,623), enhancers containing human-specific ERVs appear to play an important role in placental gene expression. However, they are unlikely to contribute to the formation of non-canonical imprints in humans. As shown previously (25,521), LTRs associated with non-canonical imprints are enriched for H3K4me3 on the paternal chromosomes, share the same orientation as the non-canonical imprinted genes, and function as alternative promoters. As demonstrated by Frost *et al.* (2023), only a few human-specific ERV enhancers are decorated by H3K4me3, and these were mostly associated with lowly expressed genes in human CTBs (138). Additionally, several canonical imprints, including *Mcts2*, *Nap1l5* and *Inpp5f\_v2*, or their human orthologs, are suggested to have originated from retrotransposon insertions within host genes (624,625). *Peg10* and *Rtl1*, for instance, both encode GAG and POL proteins with strong homology to suchi-ichi LTR retrotransposons (626), while the *Zdbf2/Liz* imprinting cluster (*ZDBF2/GPR1-AS* in humans) is proposed to have gained imprinting via integration of a MER21C LTR, which acts as an alternative promoter for *Liz* in mice or *GPR1-AS* in humans (627). Nonetheless, it remains possible that primate-specific LTR-derived enhancers may contribute to the regulation of human-specific imprinted genes during early pre-implantation stages and in the placenta. To explore this further, techniques such as Nanopore sequencing (628) in combination with CUT&Tag could be used to screen for novel monoallelically expressed genes located near active primate-specific LTR-derived enhancers in human CTBs, with subsequent experimental validation in hTSCs (138,365).

Finally, it should be noted that the rat has more non-canonical imprinted genes in extra-embryonic tissues than the mouse (450), despite the two species diverging only 13 million years ago (524). While mouse and rat share some non-canonical imprinted genes such as *Gab1*, *Jade1*, *Sfmbt2*, and *Sall1*, the rat has additional genes controlled by H3K27me3-mediated imprinting, including *LOC108350526*, *Zfp516*, *Slc38a1*, *Gadl1-3'UTR* and others. This suggests potential species-specific differences. For example, *Slc38a1*, surrounded by a broad H3K27me3 domain derived from the oocyte in both species, shows paternally biased expression only in the rat. Further investigation revealed that *Slc38a1*

contains two LTR elements: an MTD retroelement overlapping its first intron, which is also found in mice, and a second LTR element (RLTR51) within its promoter, which is absent in mice. The MTD retroelement in rats forms a sDMR and includes a ZFP57 binding motif absent in mice. Therefore, it remains unclear which of these factors is critical for imprinting of *Slc38a1* in the rat. Interestingly, *Smoc1* demonstrates paternal-biased expression in the mouse due to H3K27me3-mediated imprinting, while in reciprocal dwarf hamster hybrids (*Phodopus sungorus* and *Phodopus campbelli*), *Smoc1* shows maternal-biased expression (629). This suggests that non-canonical imprinting may evolve rapidly, even among closely related species.

Additionally, a large miRNA cluster within the tenth intron of *Sfnbt2* has been proposed to contribute to imprinting at this locus in mice and rats. Recent findings indicate that this miRNA cluster overlaps ERVK repeats, which might be responsible for imprinting (520). Thus, multiple factors, in addition to H3K27me3, may be required for non-canonical imprinting.

### 5.2.1. Study limitations

It is important to note that this study focused on identifying non-canonical imprints in placental tissues rather than in pre-implantation embryos. As observed in mice and rats, non-canonical imprinting predominantly occurs during pre-implantation stages, with only a few genes establishing sDMRs that maintain imprinted expression in extra-embryonic tissues after implantation. Most non-canonical imprints are lost after implantation, as H3K27me3 domains are erased. Therefore, there may be human-specific non-canonical imprints present during pre-implantation stages. Additionally, our analysis utilised various methylation datasets from human gametes, embryos, and somatic tissues, but we did not explore other critical epigenetic modifications, such as H3K27me3, H3K4me3, and H3K9me2/ H3K9me3, which are involved in non-canonical imprinting. We also did not assess ZFP57 and ZNF445 binding sites in our candidate genes. These ZNFs are important for maintaining canonical imprints (176,177,630), and, as demonstrated in rats for *Slc38a1*, they might also be involved in non-canonical imprinting (450). The lack of resolution regarding ZFP57 and ZNF445 binding sites in human embryo and stem cell models is partly due to the absence of suitable antibodies for ChIP, with current binding sites defined using epitope-tagged proteins in overexpressing cell models (176,177,630).

### 5.2.2. Future research

In the future, different stages of human pre-implantation embryos should be collected along with corresponding parental samples. Embryos at the 4 to 8-cell stage, morula, and blastocyst stages should be considered, as these stages exhibit significant epigenetic changes crucial for understanding non-canonical imprinting (71,136). However, this is extremely challenging due to ethical reasons. Furthermore, each embryo possesses a unique genotype, unlike when using inter-specific mouse crosses, meaning multiple embryos from different donor couples are required to draw solid conclusions. Additionally, these stages would not contain maternal transcripts derived from oocytes, which could complicate downstream analysis, so only transcripts subject to EGA can be assessed (71). Such embryos should be analysed using scRNA-seq, Nanopore, or PacBio HiFi technologies to profile all expressed genes (528). The data should be combined with parental WGS or WES datasets to determine allelic expression. gDNA from these embryos should also be sequenced to confirm embryo genotypes and address any errors observed in RNA-seq datasets. This approach will help identify genes that are monoallelically expressed at these embryonic stages. Candidate genes identified from these embryos could be validated in placental samples using the same techniques employed in this PhD thesis. Additionally, more embryos at the same developmental stage could be collected for low-input CUT&Tag or multi-CUT&Tag to profile H3K27me3 and H3K4me3 modifications (631,632). Combining these multi-omic datasets would facilitate an unbiased screening for non-canonical imprinted genes in human pre-implantation embryos if such genes exist in humans. These genes may be important for pre-implantation development and could be associated with miscarriage or other pregnancy-related complications, as aberrant expression of non-canonical imprinted genes has been linked to placental and embryonic development issues in mice (451,452).

### 5.2.3. Conclusions

During this project, I investigated non-canonical imprinted genes in the human placenta. Using various methylation datasets and molecular approaches, I failed to find robust evidence for the non-canonical imprinting of novel mouse and rat orthologous genes. Most genes showed similar methylation profiles on both parental alleles and were associated with biallelic expression. In addition, some mouse and rat non-canonical imprints investigated in human pre-implantation embryos were also biallelically expressed, indicating they were not imprinted at early time points. However, we must be

cautious, as additional genes may be subject to temporal imprinting during pre-implantation stages that were not screened during this thesis. In summary, we believe that non-canonical imprinting is not conserved in the human placenta. Nevertheless, future studies should further explore the early stages of human post-EGA embryos using continuously improving multi-omic techniques, as these genes may play an important role in determining embryo viability and developmental potential.

# References

1. Putiri EL, Robertson KD. Epigenetic mechanisms and genome stability. *Clin Epigenetics* [Internet]. 2011 Aug 29;2(2):299–314. Available from: <http://www.clinicalepigeneticsjournal.com/content/2/2/>
2. Anifandis G, Messini CI, Dafopoulos K, Messinis IE. Genes and Conditions Controlling Mammalian Pre- and Post-implantation Embryo Development. *Curr Genomics* [Internet]. 2015 Jan 28;16(1):32–46. Available from: <http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1389-2029&volume=16&issue=1&spage=32>
3. Weaver JR, Susiarjo M, Bartolomei MS. Imprinting and epigenetic changes in the early embryo. *Mamm Genome* [Internet]. 2009 Oct 16;20(9–10):532–43. Available from: <http://link.springer.com/10.1007/s00335-009-9225-2>
4. Zhu P, Guo H, Ren Y, Hou Y, Dong J, Li R, et al. Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat Genet* [Internet]. 2018 Jan 18 [cited 2021 Feb 20];50(1):12–9. Available from: <https://www.nature.com/articles/s41588-017-0007-6>
5. Miller D, Brinkworth M, Iles D. Paternal DNA packaging in spermatozoa: more than the sum of its parts? DNA, histones, protamines and epigenetics. *REPRODUCTION* [Internet]. 2010 Feb;139(2):287–301. Available from: <https://rep.bioscientifica.com/view/journals/rep/139/2/287.xml>
6. Braun RE. Packaging paternal chromosomes with protamine. *Nat Genet* [Internet]. 2001 May;28(1):10–2. Available from: [http://www.nature.com/articles/ng0501\\_10](http://www.nature.com/articles/ng0501_10)
7. Monk D. Germline-derived DNA methylation and early embryo epigenetic reprogramming: The selected survival of imprints. *Int J Biochem Cell Biol* [Internet]. 2015;67:128–38. Available from: <http://dx.doi.org/10.1016/j.biocel.2015.04.014>
8. Okae H, Chiba H, Hiura H, Hamada H, Sato A, Utsunomiya T, et al. Genome-Wide Analysis of DNA Methylation Dynamics during Early Human Development. *PLoS Genet* [Internet]. 2014;10(12):1004868. Available from: [www.plosgenetics.org](http://www.plosgenetics.org)
9. Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, et al. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* [Internet]. 2012 Mar 28;484(7394):339–44. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22456710>
10. Walser CB, Lipshitz HD. Transcript clearance during the maternal-to-zygotic transition. *Curr Opin Genet Dev* [Internet]. 2011 Aug;21(4):431–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21497081>
11. Vassena R, Boué S, González-Roca E, Aran B, Auer H, Veiga A, et al. Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. *Development* [Internet]. 2011 Sep 1;138(17):3699–709. Available from: <https://journals.biologists.com/dev/article/138/17/3699/44606/Waves-of-early-transcriptional-activation-and>
12. Schulz KN, Harrison MM. Mechanisms regulating zygotic genome activation. *Nat Rev Genet* [Internet]. 2019 Apr 20;20(4):221–34. Available from: <http://www.nature.com/articles/s41576-018-0087-x>
13. Petropoulos S, Edsgård D, Reinius B, Deng Q, Panula SP, Codeluppi S, et al. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell*. 2016;165(4):1012–26.
14. Cockburn K, Rossant J. Making the blastocyst: lessons from the mouse. *J Clin Invest* [Internet]. 2010 Apr 1;120(4):995–1003. Available from: <http://www.jci.org/articles/view/41229>
15. Johnson AA, Akman K, Calimport SRG, Wuttke D, Stolzing A, de Magalhães JP. The Role of DNA Methylation in Aging, Rejuvenation, and Age-Related Disease. *Rejuvenation Res* [Internet]. 2012 Oct;15(5):483–94. Available from: <https://www.liebertpub.com/doi/10.1089/rej.2012.1324>
16. Hill PWS, Leitch HG, Requena CE, Sun Z, Amouroux R, Roman-Trufero M, et al. Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. *Nature* [Internet]. 2018 Mar 7;555(7696):392–6. Available from: <http://www.nature.com/articles/nature25964>
17. Monk D, Mackay DJG, Eggermann T, Maher ER, Riccio A. Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. *Nat Rev Genet* [Internet]. 2019 Apr 15;20(4):235–48. Available from: <http://www.nature.com/articles/s41576-018-0092-0>
18. Tucci V, Isles AR, Kelsey G, Ferguson-Smith AC, Bartolomei MS, Benvenisty N, et al. Genomic Imprinting and Physiological Processes in Mammals. *Cell*. 2019;176(5):952–65.

19. Perez MF, Lehner B. Intergenerational and transgenerational epigenetic inheritance in animals. *Nat Cell Biol* [Internet]. 2019 Feb 2;21(2):143–51. Available from: <http://www.nature.com/articles/s41556-018-0242-9>
20. Sanchez-Delgado M, Court F, Vidal E, Medrano J, Monteagudo-Sánchez A, Martín-Trujillo A, et al. Human Oocyte-Derived Methylation Differences Persist in the Placenta Revealing Widespread Transient Imprinting. Bartolomei MS, editor. *PLOS Genet* [Internet]. 2016 Nov 11 [cited 2021 Feb 6];12(11):e1006427. Available from: <https://dx.plos.org/10.1371/journal.pgen.1006427>
21. Hamada H, Okae H, Toh H, Chiba H, Hiura H, Shirane K, et al. Allele-Specific Methylome and Transcriptome Analysis Reveals Widespread Imprinting in the Human Placenta. *Am J Hum Genet* [Internet]. 2016 Nov;99(5):1045–58. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0002929716303780>
22. Hanna CW, Peñaherrera MS, Saadeh H, Andrews S, McFadden DE, Kelsey G, et al. Pervasive polymorphic imprinted methylation in the human placenta. *Genome Res* [Internet]. 2016 Jun;26(6):756–67. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26769960>
23. Duffié R, Ajjan S, Greenberg M V., Zamudio N, del Arenal Escamilla M, Iranzo J, et al. The *Gpr1/Zdbf2* locus provides new paradigms for transient and dynamic genomic imprinting in mammals. *Genes Dev.* 2014;28(5):463–78.
24. Inoue A, Jiang L, Lu F, Suzuki T, Zhang Y. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature.* 2017;547(7664):419–24.
25. Hanna CW, Pérez-Palacios R, Gahurova L, Schubert M, Krueger F, Biggins L, et al. Endogenous retroviral insertions drive non-canonical imprinting in extra-embryonic tissues. *Genome Biol* [Internet]. 2019 Dec 29;20(1):225. Available from: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1833-x>
26. Chen Z, Yin Q, Inoue A, Zhang C, Zhang Y. Allelic H3K27me3 to allelic DNA methylation switch maintains noncanonical imprinting in extraembryonic cells. *Sci Adv* [Internet]. 2019 Dec 20;5(12):eaay7246. Available from: <https://advances.sciencemag.org/lookup/doi/10.1126/sciadv.aay7246>
27. Zhang W, Chen Z, Yin Q, Zhang D, Racowsky C, Zhang Y. Maternal-biased H3K27me3 correlates with paternal-specific gene expression in the human morula. *Genes Dev.* 2019;33(7–8):382–7.
28. Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. *Neuropsychopharmacology* [Internet]. 2013 Jan 11;38(1):23–38. Available from: <http://www.nature.com/articles/npp2012112>
29. Kim M, Costello J. DNA methylation: an epigenetic mark of cellular memory. *Exp Mol Med* [Internet]. 2017 Apr 28;49(4):e322–e322. Available from: <http://www.nature.com/articles/emm201710>
30. Taudt A, Colomé-Tatché M, Johannes F. Genetic sources of population epigenomic variation. *Nat Rev Genet* [Internet]. 2016 Jun 9;17(6):319–32. Available from: <http://www.nature.com/articles/nrg.2016.45>
31. Detich N, Hamm S, Just G, Knox JD, Szyf M. The Methyl Donor S-Adenosylmethionine Inhibits Active Demethylation of DNA. *J Biol Chem* [Internet]. 2003 Jun;278(23):20812–20. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S002192582073382X>
32. Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL, et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet* [Internet]. 2010 May 13;6(5):e1000952. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20485568>
33. McClay JL, Shabalin AA, Dozmorov MG, Adkins DE, Kumar G, Nerella S, et al. High density methylation QTL analysis in human blood via next-generation sequencing of the methylated genomic DNA fraction. *Genome Biol* [Internet]. 2015 Dec 23;16(1):291. Available from: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0842-7>
34. Rowlatt A, Hernández-Suárez G, Sanabria-Salas MC, Serrano-López M, Rawlik K, Hernandez-Illan E, et al. The heritability and patterns of DNA methylation in normal human colorectum. *Hum Mol Genet* [Internet]. 2016 Mar 2;25(12):2600–11. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddw072>
35. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* [Internet]. 2002 Jan 1;16(1):6–21. Available from: <http://genesdev.cshlp.org/lookup/doi/10.1101/gad.947102>
36. Youk J, An Y, Park S, Lee JK, Ju YS. The genome-wide landscape of C:G & T:A polymorphism at the CpG contexts in the human population. *BMC Genomics* [Internet]. 2020 Dec 30;21(1):270. Available from: <https://bmcbgenomics.biomedcentral.com/articles/10.1186/s12864-020-6674-1>
37. Russell GJ, Walker PMB, Elton RA, Subak-Sharpe JH. Doublet frequency analysis of fractionated vertebrate nuclear DNA. *J Mol Biol* [Internet]. 1976 Nov;108(1):1–20. Available

- from: <https://linkinghub.elsevier.com/retrieve/pii/S0022283676800903>
38. Bird AP. CpG-rich islands and the function of DNA methylation. *Nature* [Internet]. 1986 May;321(6067):209–13. Available from: <https://www.nature.com/articles/321209a0>
  39. Cooper DN, Youssoufian H. The CpG dinucleotide and human genetic disease. *Hum Genet* [Internet]. 1988 Feb;78(2):151–5. Available from: <http://link.springer.com/10.1007/BF00278187>
  40. Zhou Y, He F, Pu W, Gu X, Wang J, Su Z. The Impact of DNA Methylation Dynamics on the Mutation Rate During Human Germline Development. *G3 Genes|Genomes|Genetics* [Internet]. 2020 Sep 1;10(9):3337–46. Available from: <https://academic.oup.com/g3journal/article/10/9/3337/6060130>
  41. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* [Internet]. 2009 Nov 19;462(7271):315–22. Available from: <https://www.nature.com/articles/nature08514>
  42. Loyfer N, Magenheimer J, Peretz A, Cann G, Bredno J, Klochendler A, et al. A DNA methylation atlas of normal human cell types. *Nature*. 2023;613(7943):355–64.
  43. Scelfo, Fachinetti. Keeping the Centromere under Control: A Promising Role for DNA Methylation. *Cells* [Internet]. 2019 Aug 16;8(8):912. Available from: <https://www.mdpi.com/2073-4409/8/8/912>
  44. Vera E, Canela A, Fraga MF, Esteller M, Blasco MA. Epigenetic regulation of telomeres in human cancer. *Oncogene* [Internet]. 2008 Nov 1;27(54):6817–33. Available from: <http://www.nature.com/articles/onc2008289>
  45. Choi SH, Worswick S, Byun HM, Shear T, Soussa JC, Wolff EM, et al. Changes in DNA methylation of tandem DNA repeats are different from interspersed repeats in cancer. *Int J Cancer* [Internet]. 2009 Aug 1;125(3):723–9. Available from: <http://doi.wiley.com/10.1002/ijc.24384>
  46. Bourc'his D, Bestor TH. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* [Internet]. 2004 Sep 18;431(7004):96–9. Available from: <https://www.nature.com/articles/nature02886>
  47. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation. *Science* (80- ) [Internet]. 2010 May 14;328(5980):916–9. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1186366>
  48. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* [Internet]. 2007 Apr 4;39(4):457–66. Available from: <https://www.nature.com/articles/ng1990>
  49. Saksouk N, Simboeck E, Déjardin J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* [Internet]. 2015 Dec 15;8(1):3. Available from: <https://epigeneticsandchromatin.biomedcentral.com/articles/10.1186/1756-8935-8-3>
  50. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* [Internet]. 2017 May 5;356(6337):eaaj22239. Available from: <https://www.science.org/doi/10.1126/science.aaj2239>
  51. Chen X, Li Y, Rubio K, Deng B, Li Y, Tang Q, et al. Lymphoid-specific helicase in epigenetics, DNA repair and cancer. *Br J Cancer* [Internet]. 2022 Feb 1;126(2):165–73. Available from: <https://www.nature.com/articles/s41416-021-01543-2>
  52. Geiman TM, Sankpal UT, Robertson AK, Zhao Y, Zhao Y, Robertson KD. DNMT3B interacts with hSNF2H chromatin remodeling enzyme, HDACs 1 and 2, and components of the histone methylation system. *Biochem Biophys Res Commun* [Internet]. 2004 May;318(2):544–55. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0006291X04007764>
  53. Fuks F. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* [Internet]. 2003 May 1;31(9):2305–12. Available from: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkg332>
  54. Fuks F. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J* [Internet]. 2001 May 15;20(10):2536–44. Available from: <http://emboj.embopress.org/cgi/doi/10.1093/emboj/20.10.2536>
  55. Du Q, Luu PL, Stirzaker C, Clark SJ. Methyl-CpG-Binding Domain Proteins: Readers of the Epigenome. *Epigenomics* [Internet]. 2015 Sep 30;7(6):1051–73. Available from: <https://www.tandfonline.com/doi/full/10.2217/epi.15.39>
  56. Jones PL, Veenstra GJC, Wade PA, Vermaak D, Kass SU, Landsberger N, et al. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* [Internet]. 1998 Jun;19(2):187–91. Available from: [https://www.nature.com/articles/ng0698\\_187](https://www.nature.com/articles/ng0698_187)

57. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* [Internet]. 1998 May;393(6683):386–9. Available from: <https://www.nature.com/articles/30764>
58. Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, et al. In Embryonic Stem Cells, ZFP57/KAP1 Recognize a Methylated Hexanucleotide to Affect Chromatin and DNA Methylation of Imprinting Control Regions. *Mol Cell* [Internet]. 2011 Nov;44(3):361–72. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276511007659>
59. Denslow SA, Wade PA. The human Mi-2/NuRD complex and gene regulation. *Oncogene* [Internet]. 2007 Aug 13;26(37):5433–8. Available from: <https://www.nature.com/articles/1210611>
60. Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, et al. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet* [Internet]. 1999 Sep;23(1):58–61. Available from: [https://www.nature.com/articles/ng0999\\_58](https://www.nature.com/articles/ng0999_58)
61. Schübeler D. Function and information content of DNA methylation. *Nature* [Internet]. 2015 Jan 15;517(7534):321–6. Available from: <https://www.nature.com/articles/nature14192>
62. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet* [Internet]. 2013 Mar 12;14(3):204–20. Available from: <https://www.nature.com/articles/nrg3354>
63. Chédin F. The DNMT3 Family of Mammalian De Novo DNA Methyltransferases. In 2011. p. 255–85. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B978012387685000007X>
64. Okano M, Bell DW, Haber DA, Li E. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell* [Internet]. 1999 Oct;99(3):247–57. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867400816566>
65. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the Establishment of Maternal Genomic Imprints. *Science* (80- ) [Internet]. 2001 Dec 21;294(5551):2536–9. Available from: <https://www.science.org/doi/10.1126/science.1065848>
66. Petrusa L, Van de Velde H, De Rycke M. Dynamic regulation of DNA methyltransferases in human oocytes and preimplantation embryos after assisted reproductive technologies. *MHR Basic Sci Reprod Med* [Internet]. 2014 Sep;20(9):861–74. Available from: <https://academic.oup.com/molehr/article-lookup/doi/10.1093/molehr/gau049>
67. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* [Internet]. 2013 Sep 11;20(9):1131–9. Available from: <https://www.nature.com/articles/nsmb.2660>
68. Theunissen TW, Friedli M, He Y, Planet E, O'Neil RC, Markoulaki S, et al. Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell* [Internet]. 2016 Oct;19(4):502–15. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1934590916301618>
69. Kurihara Y, Kawamura Y, Uchijima Y, Amamo T, Kobayashi H, Asano T, et al. Maintenance of genomic methylation patterns during preimplantation development requires the somatic form of DNA methyltransferase 1. *Dev Biol* [Internet]. 2008 Jan;313(1):335–46. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0012160607014844>
70. Branco MR, Oda M, Reik W. Safeguarding parental identity: Dnmt1 maintains imprints during epigenetic reprogramming in early embryogenesis. *Genes Dev* [Internet]. 2008 Jun 15;22(12):1567–71. Available from: <http://www.genesdev.org/cgi/doi/10.1101/gad.1690508>
71. Hernandez Mora JR, Buhigas C, Clark S, Del Gallego Bonilla R, Daskeviciute D, Monteaudo-Sánchez A, et al. Single-cell multi-omic analysis profiles defective genome activation and epigenetic reprogramming associated with human pre-implantation embryo arrest. *Cell Rep*. 2023;42(2).
72. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* [Internet]. 1992 Jun;69(6):915–26. Available from: <https://linkinghub.elsevier.com/retrieve/pii/009286749290611F>
73. Hackett JA, Sengupta R, Zyllicz JJ, Murakami K, Lee C, Down TA, et al. Germline DNA Demethylation Dynamics and Imprint Erasure Through 5-Hydroxymethylcytosine. *Science* (80- ) [Internet]. 2013 Jan 25;339(6118):448–52. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1229277>
74. Song CX, He C. Potential functional roles of DNA demethylation intermediates. *Trends Biochem Sci* [Internet]. 2013 Oct;38(10):480–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23932479>

75. He B, Zhang C, Zhang X, Fan Y, Zeng H, Liu J, et al. Tissue-specific 5-hydroxymethylcytosine landscape of the human genome. *Nat Commun* [Internet]. 2021 Jul 12;12(1):4249. Available from: <https://www.nature.com/articles/s41467-021-24425-w>
76. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science* (80- ) [Internet]. 2011 Sep 2;333(6047):1300–3. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1210597>
77. Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, et al. Base-Resolution Analysis of 5-Hydroxymethylcytosine in the Mammalian Genome. *Cell* [Internet]. 2012 Jun;149(6):1368–80. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S009286741200534X>
78. Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System. *Cell* [Internet]. 2012 Dec;151(7):1417–30. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867412014079>
79. Song CX, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, et al. Genome-wide Profiling of 5-Formylcytosine Reveals Its Roles in Epigenetic Priming. *Cell* [Internet]. 2013 Apr;153(3):678–91. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867413004005>
80. Shen L, Wu H, Diep D, Yamaguchi S, D'Alessio AC, Fung HL, et al. Genome-wide Analysis Reveals TET- and TDG-Dependent 5-Methylcytosine Oxidation Dynamics. *Cell* [Internet]. 2013 Apr;153(3):692–706. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867413004017>
81. Heyn H, Esteller M. An Adenine Code for DNA: A Second Life for N6-Methyladenine. *Cell* [Internet]. 2015 May;161(4):710–3. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867415004389>
82. Xiao CL, Zhu S, He M, Chen D, Zhang Q, Chen Y, et al. N6-Methyladenine DNA Modification in the Human Genome. *Mol Cell* [Internet]. 2018 Jul;71(2):306–318.e7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S109727651830460X>
83. Clancy MJ. Induction of sporulation in *Saccharomyces cerevisiae* leads to the formation of N6-methyladenosine in mRNA: a potential mechanism for the activity of the IME4 gene. *Nucleic Acids Res* [Internet]. 2002 Oct 15;30(20):4509–18. Available from: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkf573>
84. Bodi Z, Zhong S, Mehra S, Song J, Graham N, Li H, et al. Adenosine Methylation in Arabidopsis mRNA is Associated with the 3' End and Reduced Levels Cause Developmental Defects. *Front Plant Sci* [Internet]. 2012;3. Available from: <http://journal.frontiersin.org/article/10.3389/fpls.2012.00048/abstract>
85. Wu TP, Wang T, Seetin MG, Lai Y, Zhu S, Lin K, et al. DNA methylation on N6-adenine in mammalian embryonic stem cells. *Nature* [Internet]. 2016 Apr 21;532(7599):329–33. Available from: <http://www.nature.com/articles/nature17640>
86. Broche J, Köhler AR, Kühnel F, Osteresch B, Chandrasekaran TT, Adam S, et al. Genome-wide deposition of 6-methyladenine in human DNA reduces the viability of HEK293 cells and directly influences gene expression. *Commun Biol*. 2023;6(1):1–12.
87. Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, et al. Global Epigenomic Reconfiguration During Mammalian Brain Development. *Science* (80- ) [Internet]. 2013 Aug 9;341(6146):1237905. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1237905>
88. Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, et al. The DNA methylation landscape of human early embryos. *Nature*. 2014;511(7511):606–10.
89. Tomizawa S ichi, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G, et al. Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development* [Internet]. 2011 Mar 1;138(5):811–20. Available from: <https://journals.biologists.com/dev/article/138/5/811/44864/Dynamic-stage-specific-changes-in-imprinted>
90. Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, et al. Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nat Neurosci* [Internet]. 2014 Feb 22;17(2):215–22. Available from: <https://www.nature.com/articles/nn.3607>
91. Ziller MJ, Müller F, Liao J, Zhang Y, Gu H, Bock C, et al. Genomic Distribution and Inter-Sample Variation of Non-CpG Methylation across Human Cell Types. Schübeler D, editor. *PLoS Genet* [Internet]. 2011 Dec 8;7(12):e1002389. Available from: <https://dx.plos.org/10.1371/journal.pgen.1002389>
92. Arand J, Spieler D, Karius T, Branco MR, Meilinger D, Meissner A, et al. In Vivo Control of CpG and Non-CpG DNA Methylation by DNA Methyltransferases. Schübeler D, editor. *PLoS*

- Genet [Internet]. 2012 Jun 28;8(6):e1002750. Available from: <https://dx.plos.org/10.1371/journal.pgen.1002750>
93. Liao J, Karnik R, Gu H, Ziller MJ, Clement K, Tsankov AM, et al. Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat Genet* [Internet]. 2015 May 30;47(5):469–78. Available from: <https://www.nature.com/articles/ng.3258>
  94. Barrès R, Osler ME, Yan J, Rune A, Fritz T, Caidahl K, et al. Non-CpG methylation of the PGC-1 $\alpha$  promoter through DNMT3B controls mitochondrial density. *Cell Metab* [Internet]. 2009 Sep;10(3):189–98. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19723495>
  95. Inoue S, Oishi M. Effects of methylation of non-CpG sequence in the promoter region on the expression of human synaptotagmin XI (syt11). *Gene* [Internet]. 2005 Mar;348:123–34. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S037811904007954>
  96. Barrès R, Osler ME, Yan J, Rune A, Fritz T, Caidahl K, et al. Non-CpG Methylation of the PGC-1 $\alpha$  Promoter through DNMT3B Controls Mitochondrial Density. *Cell Metab* [Internet]. 2009 Sep;10(3):189–98. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1550413109002290>
  97. Lai WKM, Pugh BF. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat Rev Mol Cell Biol* [Internet]. 2017 Sep 24;18(9):548–62. Available from: <https://www.nature.com/articles/nrm.2017.47>
  98. Luger K, Dechassa ML, Tremethick DJ. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat Rev Mol Cell Biol* [Internet]. 2012 Jul 22;13(7):436–47. Available from: <https://www.nature.com/articles/nrm3382>
  99. Gillette TG, Hill JA. Readers, Writers, and Erasers. *Circ Res* [Internet]. 2015 Mar 27;116(7):1245–53. Available from: <https://www.ahajournals.org/doi/10.1161/CIRCRESAHA.116.303630>
  100. Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. *Exp Mol Med* [Internet]. 2017 Apr 28;49(4):e324–e324. Available from: <https://www.nature.com/articles/emm201711>
  101. Millán-Zambrano G, Burton A, Bannister AJ, Schneider R. Histone post-translational modifications – cause and consequence of genome function. *Nat Rev Genet* [Internet]. 2022 Sep 25;23(9):563–80. Available from: <https://www.nature.com/articles/s41576-022-00468-7>
  102. Musselman CA, Khorasanizadeh S, Kutateladze TG. Towards understanding methyllysine readout. *Biochim Biophys Acta - Gene Regul Mech* [Internet]. 2014 Aug;1839(8):686–93. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1874939914000789>
  103. Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD. A Novel Domain in Set2 Mediates RNA Polymerase II Interaction and Couples Histone H3 K36 Methylation with Transcript Elongation. *Mol Cell Biol* [Internet]. 2005 Apr 1;25(8):3305–16. Available from: <https://www.tandfonline.com/doi/full/10.1128/MCB.25.8.3305-3316.2005>
  104. Marsano RM, Dimitri P. Constitutive Heterochromatin in Eukaryotic Genomes: A Mine of Transposable Elements. *Cells* [Internet]. 2022 Feb 22;11(5):761. Available from: <https://www.mdpi.com/2073-4409/11/5/761>
  105. Inoue A, Jiang L, Lu F, Zhang Y. Genomic imprinting of Xist by maternal H3K27me3. *Genes Dev* [Internet]. 2017 Oct 1;31(19):1927–32. Available from: <http://genesdev.cshlp.org/lookup/doi/10.1101/gad.304113.117>
  106. Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, et al. A model for transmission of the H3K27me3 epigenetic mark. *Nat Cell Biol* [Internet]. 2008 Nov 19;10(11):1291–300. Available from: <https://www.nature.com/articles/ncb1787>
  107. Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, et al. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* [Internet]. 2001 Mar;410(6824):120–4. Available from: <https://www.nature.com/articles/35065138>
  108. Wang T, Xu C, Liu Y, Fan K, Li Z, Sun X, et al. Crystal Structure of the Human SUV39H1 Chromodomain and Its Recognition of Histone H3K9me2/3. Ballestar E, editor. *PLoS One* [Internet]. 2012 Dec 28;7(12):e52977. Available from: <https://dx.plos.org/10.1371/journal.pone.0052977>
  109. Lachner M, O’Carroll D, Rea S, Mechtler K, Jenuwein T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* [Internet]. 2001 Mar;410(6824):116–20. Available from: <https://www.nature.com/articles/35065132>
  110. Padeken J, Methot SP, Gasser SM. Establishment of H3K9-methylated heterochromatin and its functions in tissue differentiation and maintenance. *Nat Rev Mol Cell Biol* [Internet]. 2022 Sep 13;23(9):623–40. Available from: <https://www.nature.com/articles/s41580-022-00483-w>
  111. Janssen SM, Lorincz MC. Interplay between chromatin marks in development and disease.

- Nat Rev Genet. 2022;23(3):137–53.
112. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* [Internet]. 2012 May 11;485(7398):376–80. Available from: <https://www.nature.com/articles/nature11082>
  113. de Wit E, Vos ESM, Holwerda SJB, Valdes-Quezada C, Verstegen MJAM, Teunissen H, et al. CTCF Binding Polarity Determines Chromatin Looping. *Mol Cell* [Internet]. 2015 Nov;60(4):676–84. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276515007625>
  114. Kim Y, Shi Z, Zhang H, Finkelstein IJ, Yu H. Human cohesin compacts DNA by loop extrusion. *Science* (80- ) [Internet]. 2019 Dec 13;366(6471):1345–9. Available from: <https://www.science.org/doi/10.1126/science.aaz4475>
  115. Kubo N, Ishii H, Xiong X, Bianco S, Meitinger F, Hu R, et al. Promoter-proximal CTCF binding promotes distal enhancer-dependent gene activation. *Nat Struct Mol Biol* [Internet]. 2021 Feb 4;28(2):152–61. Available from: <https://www.nature.com/articles/s41594-020-00539-5>
  116. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragozy T, Telling A, et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* (80- ) [Internet]. 2009 Oct 9;326(5950):289–93. Available from: <https://www.science.org/doi/10.1126/science.1181369>
  117. Briand N, Collas P. Lamina-associated domains: peripheral matters and internal affairs. *Genome Biol* [Internet]. 2020 Dec 2;21(1):85. Available from: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02003-5>
  118. Cremer T, Cremer M. Chromosome Territories. *Cold Spring Harb Perspect Biol* [Internet]. 2010 Mar 1;2(3):a003889–a003889. Available from: <http://cshperspectives.cshlp.org/lookup/doi/10.1101/cshperspect.a003889>
  119. Altemose N, Logsdon GA, Bzikadze A V., Sidhwani P, Langley SA, Caldas G V., et al. Complete genomic and epigenetic maps of human centromeres. *Science* (80- ) [Internet]. 2022 Apr;376(6588). Available from: <https://www.science.org/doi/10.1126/science.abl4178>
  120. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al. Integrative analysis of 111 reference human epigenomes. *Nature* [Internet]. 2015 Feb 19;518(7539):317–30. Available from: <https://www.nature.com/articles/nature14248>
  121. Cosgrove MS, Boeke JD, Wolberger C. Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol* [Internet]. 2004 Nov 2;11(11):1037–43. Available from: <https://www.nature.com/articles/nsmb851>
  122. Wang H, Fan Z, Shliha P V., Miele M, Hendrickson RC, Jiang X, et al. H3K4me3 regulates RNA polymerase II promoter-proximal pause-release. *Nature* [Internet]. 2023 Mar 9;615(7951):339–48. Available from: <https://www.nature.com/articles/s41586-023-05780-8>
  123. Wu K, Fan D, Zhao H, Liu Z, Hou Z, Tao W, et al. Dynamics of histone acetylation during human early embryogenesis. *Cell Discov*. 2023;9(1).
  124. Zhang W. Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. *EMBO J* [Internet]. 1998 Jun 1;17(11):3155–67. Available from: <http://emboj.embopress.org/cgi/doi/10.1093/emboj/17.11.3155>
  125. Di Lorenzo A, Bedford MT. Histone arginine methylation. *FEBS Lett* [Internet]. 2011 Jul 7;585(13):2024–31. Available from: <https://febs.onlinelibrary.wiley.com/doi/10.1016/j.febslet.2010.11.010>
  126. Sabari BR, Tang Z, Huang H, Yong-Gonzalez V, Molina H, Kong HE, et al. Intracellular Crotonyl-CoA Stimulates Transcription through p300-Catalyzed Histone Crotonylation. *Mol Cell* [Internet]. 2015 Apr;58(2):203–15. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276515001434>
  127. Dawson MA, Foster SD, Bannister AJ, Robson SC, Hannah R, Wang X, et al. Three Distinct Patterns of Histone H3Y41 Phosphorylation Mark Active Genes. *Cell Rep* [Internet]. 2012 Sep;2(3):470–7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2211124712002410>
  128. Zhang D, Tang Z, Huang H, Zhou G, Cui C, Weng Y, et al. Metabolic regulation of gene expression by histone lactylation. *Nature* [Internet]. 2019 Oct 24;574(7779):575–80. Available from: <https://www.nature.com/articles/s41586-019-1678-1>
  129. Bao X, Liu Z, Zhang W, Gladysz K, Fung YME, Tian G, et al. Glutarylation of Histone H4 Lysine 91 Regulates Chromatin Dynamics. *Mol Cell* [Internet]. 2019 Nov;76(4):660–675.e9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276519306574>
  130. Chen Z, Djekidel MN, Zhang Y. Distinct dynamics and functions of H2AK119ub1 and H3K27me3 in mouse preimplantation embryos. *Nat Genet* [Internet]. 2021 Apr 5;53(4):551–63. Available from: <http://dx.doi.org/10.1038/s41588-021-00821-2>

131. Condemi L, Mocavini I, Aranda S, Di Croce L. Polycomb function in early mouse development. *Cell Death Differ* [Internet]. 2024 Jul 12;(June):1–10. Available from: <http://dx.doi.org/10.1038/s41418-024-01340-3>
132. Mei H, Kozuka C, Hayashi R, Kumon M, Koseki H, Inoue A. H2AK119ub1 guides maternal inheritance and zygotic deposition of H3K27me3 in mouse embryos. *Nat Genet* [Internet]. 2021 Apr 5;53(4):539–50. Available from: <http://dx.doi.org/10.1038/s41588-021-00820-3>
133. Diamant I, Clarke DJB, Evangelista JE, Lingam N, Ma'ayan A. Harmonizome 3.0: integrated knowledge about genes and proteins from diverse multi-omics resources. *Nucleic Acids Res* [Internet]. 2025 Jan 6;53(D1):D1016–28. Available from: <https://academic.oup.com/nar/article/53/D1/D1016/7905315>
134. Crist AM, Hinkle KM, Wang X, Moloney CM, Matchett BJ, Labuzan SA, et al. Transcriptomic analysis to identify genes associated with selective hippocampal vulnerability in Alzheimer's disease. *Nat Commun* [Internet]. 2021 Apr 19;12(1):2311. Available from: <https://www.nature.com/articles/s41467-021-22399-3>
135. Zhao Y, Garcia BA. Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harb Perspect Biol* [Internet]. 2015 Sep 1;7(9):a025064. Available from: <http://cshperspectives.cshlp.org/lookup/doi/10.1101/cshperspect.a025064>
136. Xia W, Xu J, Yu G, Yao G, Xu K, Ma X, et al. Resetting histone modifications during human parental-to-zygotic transition. *Science* (80- ) [Internet]. 2019 Jul 26;365(6451):353–60. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31273069>
137. Lu X, Zhang Y, Wang L, Wang L, Wang H, Xu Q, et al. Evolutionary epigenomic analyses in mammalian early embryos reveal species-specific innovations and conserved principles of imprinting. *Sci Adv* [Internet]. 2021 Nov 26;7(48). Available from: <https://www.science.org/doi/10.1126/sciadv.abi6178>
138. Frost JM, Amante SM, Okae H, Jones EM, Ashley B, Lewis RM, et al. Regulation of human trophoblast gene expression by endogenous retroviruses. *Nat Struct Mol Biol* [Internet]. 2023 Apr 3;30(4):527–38. Available from: <https://www.nature.com/articles/s41594-023-00960-6>
139. Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, et al. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* [Internet]. 2016;537(7621):548–52. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27626377>
140. Zhang B, Zheng H, Huang B, Li W, Xiang Y, Peng X, et al. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature*. 2016;537(7621):553–7.
141. Laubert SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, et al. H3K4me3 Interactions with TAF3 Regulate Preinitiation Complex Assembly and Selective Gene Activation. *Cell* [Internet]. 2013 Feb;152(5):1021–36. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S009286741300144X>
142. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* [Internet]. 2007 Aug 1;448(7153):553–60. Available from: <https://www.nature.com/articles/nature06008>
143. Lee BK, Salamah J, Cheeran E, Adu-Gyamfi EA. Dynamic and distinct histone modifications facilitate human trophoblast lineage differentiation. *Sci Rep* [Internet]. 2024 Feb 24;14(1):4505. Available from: <https://doi.org/10.1038/s41598-024-55189-0>
144. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* [Internet]. 2006 Apr;125(2):315–26. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867406003801>
145. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* [Internet]. 2008 Aug 7;454(7205):766–70. Available from: <https://www.nature.com/articles/nature07107>
146. Zhang Y, Jurkowska R, Soeroes S, Rajavelu A, Dhayalan A, Bock I, et al. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res* [Internet]. 2010 Jul;38(13):4246–53. Available from: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkq147>
147. Singh P, Li AX, Tran DA, Oates N, Kang ER, Wu X, et al. De Novo DNA Methylation in the Male Germ Line Occurs by Default but Is Excluded at Sites of H3K4 Methylation. *Cell Rep* [Internet]. 2013 Jul;4(1):205–19. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2211124713002805>
148. Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* [Internet]. 2007 Aug;448(7154):714–7. Available from: <https://www.nature.com/articles/nature05987>

149. Hanna CW, Taudt A, Huang J, Gahurova L, Kranz A, Andrews S, et al. MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. *Nat Struct Mol Biol* [Internet]. 2018;25(1):73–82. Available from: <http://dx.doi.org/10.1038/s41594-017-0013-5>
150. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* [Internet]. 2009 May 18;459(7243):108–12. Available from: <https://www.nature.com/articles/nature07829>
151. Pott S, Lieb JD. What are super-enhancers? *Nat Genet* [Internet]. 2015 Jan 30;47(1):8–12. Available from: <https://www.nature.com/articles/ng.3167>
152. Shirane K, Miura F, Ito T, Lorincz MC. NSD1-deposited H3K36me2 directs de novo methylation in the mouse male germline and counteracts Polycomb-associated silencing. *Nat Genet* [Internet]. 2020 Oct 14;52(10):1088–98. Available from: <https://www.nature.com/articles/s41588-020-0689-z>
153. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, et al. Histone H3 Methylation by Set2 Directs Deacetylation of Coding Regions by Rpd3S to Suppress Spurious Intragenic Transcription. *Cell* [Internet]. 2005 Nov;123(4):581–92. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867405011566>
154. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. Regulation of Alternative Splicing by Histone Modifications. *Science* (80- ) [Internet]. 2010 Feb 19;327(5968):996–1000. Available from: <https://www.science.org/doi/10.1126/science.1184208>
155. Pfister SX, Ahrabi S, Zalmas LP, Sarkar S, Aymard F, Bachrati CZ, et al. SETD2-Dependent Histone H3K36 Trimethylation Is Required for Homologous Recombination Repair and Genome Stability. *Cell Rep* [Internet]. 2014 Jun;7(6):2006–18. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S221112471400401X>
156. Dukatz M, Holzer K, Choudalakis M, Emperle M, Lungu C, Bashtrykov P, et al. H3K36me2/3 Binding and DNA Binding of the DNA Methyltransferase DNMT3A PWWP Domain Both Contribute to its Chromatin Interaction. *J Mol Biol* [Internet]. 2019 Dec;431(24):5063–74. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0022283619305601>
157. Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, et al. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* [Internet]. 2015 Apr 21;520(7546):243–7. Available from: <https://www.nature.com/articles/nature14176>
158. Xu Q, Xiang Y, Wang Q, Wang L, Brind'Amour J, Bogutz AB, et al. SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. *Nat Genet* [Internet]. 2019 May 29;51(5):844–56. Available from: <https://www.nature.com/articles/s41588-019-0398-7>
159. Yano S, Ishiuchi T, Abe S, Namekawa SH, Huang G, Ogawa Y, et al. Histone H3K36me2 and H3K36me3 form a chromatin platform essential for DNMT3A-dependent DNA methylation in mouse oocytes. *Nat Commun* [Internet]. 2022 Aug 3;13(1):4440. Available from: <https://www.nature.com/articles/s41467-022-32141-2>
160. Shirane K, Toh H, Kobayashi H, Miura F, Chiba H, Ito T, et al. Mouse Oocyte Methylomes at Base Resolution Reveal Genome-Wide Accumulation of Non-CpG Methylation and Role of DNA Methyltransferases. Bartolomei MS, editor. *PLoS Genet* [Internet]. 2013 Apr 18;9(4):e1003439. Available from: <https://dx.plos.org/10.1371/journal.pgen.1003439>
161. Deniz Ö, Frost JM, Branco MR. Regulation of transposable elements by DNA modifications. *Nat Rev Genet* [Internet]. 2019;20(7):417–31. Available from: <http://dx.doi.org/10.1038/s41576-019-0106-6>
162. Lehnertz B, Ueda Y, Derijck AAHA, Braunschweig U, Perez-Burgos L, Kubicek S, et al. Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin. *Curr Biol* [Internet]. 2003 Jul;13(14):1192–200. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0960982203004329>
163. Maenohara S, Unoki M, Toh H, Ohishi H, Sharif J, Koseki H, et al. Role of UHRF1 in de novo DNA methylation in oocytes and maintenance methylation in preimplantation embryos. Cohen PE, editor. *PLOS Genet* [Internet]. 2017 Oct 4;13(10):e1007042. Available from: <https://dx.plos.org/10.1371/journal.pgen.1007042>
164. Xie S, Jakoncic J, Qian C. UHRF1 Double Tudor Domain and the Adjacent PHD Finger Act Together to Recognize K9me3-Containing Histone H3 Tail. *J Mol Biol* [Internet]. 2012 Jan;415(2):318–28. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0022283611012319>
165. Bostick M, Jong KK, Estève PO, Clark A, Pradhan S, Jacobsen SE. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* (80- ). 2007;317(5845):1760–4.

166. Sharif J, Muto M, Takebayashi S ichiro, Suetake I, Iwamatsu A, Endo TA, et al. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* [Internet]. 2007 Dec 11;450(7171):908–12. Available from: <https://www.nature.com/articles/nature06397>
167. Qin W, Wolf P, Liu N, Link S, Smets M, Mastra F La, et al. DNA methylation requires a DNMT1 ubiquitin interacting motif (UIM) and histone ubiquitination. *Cell Res* [Internet]. 2015 Aug 12;25(8):911–29. Available from: <https://www.nature.com/articles/cr201572>
168. Nishiyama A, Yamaguchi L, Sharif J, Johmura Y, Kawamura T, Nakanishi K, et al. Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature* [Internet]. 2013 Oct 8;502(7470):249–53. Available from: <https://www.nature.com/articles/nature12488>
169. Takeshita K, Suetake I, Yamashita E, Suga M, Narita H, Nakagawa A, et al. Structural insight into maintenance methylation by mouse DNA methyltransferase 1 (Dnmt1). *Proc Natl Acad Sci* [Internet]. 2011 May 31;108(22):9055–9. Available from: <https://pnas.org/doi/full/10.1073/pnas.1019629108>
170. Soufi A, Donahue G, Zaret KS. Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome. *Cell* [Internet]. 2012 Nov;151(5):994–1004. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867412012986>
171. van Steensel B, Belmont AS. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* [Internet]. 2017 May;169(5):780–91. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867417304737>
172. Montavon T, Shukeir N, Erikson G, Engst B, Onishi-Seebacher M, Ryan D, et al. Complete loss of H3K9 methylation dissolves mouse heterochromatin organization. *Nat Commun* [Internet]. 2021 Jul 16;12(1):4359. Available from: <https://www.nature.com/articles/s41467-021-24532-8>
173. Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, et al. The Dynamics of Genome-wide DNA Methylation Reprogramming in Mouse Primordial Germ Cells. *Mol Cell* [Internet]. 2012 Dec;48(6):849–62. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S109727651200932X>
174. Liu S, Brind'Amour J, Karimi MM, Shirane K, Bogutz A, Lefebvre L, et al. Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. *Genes Dev* [Internet]. 2014 Sep 15;28(18):2041–55. Available from: <http://genesdev.cshlp.org/lookup/doi/10.1101/gad.244848.114>
175. Gruhn WH, Tang WWC, Dietmann S, Alves-Lopes JP, Penfold CA, Wong FCK, et al. Epigenetic resetting in the human germ line entails histone modification remodeling. *Sci Adv* [Internet]. 2023;9(3):1–22. Available from: <https://doi.org/10.1126/sciadv.ade1257>
176. Monteagudo-Sánchez A, Hernandez Mora JR, Simon C, Burton A, Tenorio J, Lapunzina P, et al. The role of ZFP57 and additional KRAB-zinc finger proteins in the maintenance of human imprinted methylation and multi-locus imprinting disturbances. *Nucleic Acids Res* [Internet]. 2020 Nov 18;48(20):11394–407. Available from: <https://academic.oup.com/nar/article/48/20/11394/5923428>
177. Takahashi N, Coluccio A, Thorball CW, Planet E, Shi H, Offner S, et al. ZNF445 is a primary regulator of genomic imprinting. *Genes Dev* [Internet]. 2019 Jan 1;33(1–2):49–54. Available from: <http://genesdev.cshlp.org/lookup/doi/10.1101/gad.320069.118>
178. Piunti A, Shilatifard A. The roles of Polycomb repressive complexes in mammalian development and cancer. *Nat Rev Mol Cell Biol*. 2021;22(5):326–45.
179. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, et al. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science* (80- ) [Internet]. 2002 Nov;298(5595):1039–43. Available from: <https://www.science.org/doi/10.1126/science.1076997>
180. Conway E, Jerman E, Healy E, Ito S, Holloch D, Oliviero G, et al. A Family of Vertebrate-Specific Polycombs Encoded by the LCOR/LCORL Genes Balance PRC2 Subtype Activities. *Mol Cell* [Internet]. 2018 May;70(3):408–421.e8. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276518301862>
181. Du Z, Zheng H, Kawamura YK, Zhang K, Gassler J, Powell S, et al. Polycomb Group Proteins Regulate Chromatin Architecture in Mouse Oocytes and Early Embryos. *Mol Cell* [Internet]. 2020 Feb;77(4):825–839.e7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276519308408>
182. Parreno V, Martinez AM, Cavalli G. Mechanisms of Polycomb group protein function in cancer. *Cell Res* [Internet]. 2022 Mar 19;32(3):231–53. Available from: <https://www.nature.com/articles/s41422-021-00606-6>
183. Højfeldt JW, Hedehus L, Laugesen A, Tatar T, Wiehle L, Helin K. Non-core Subunits of the

- PRC2 Complex Are Collectively Required for Its Target-Site Specificity. *Mol Cell* [Internet]. 2019 Nov;76(3):423-436.e3. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276519305866>
184. van Mierlo G, Veenstra GJC, Vermeulen M, Marks H. The Complexity of PRC2 Subcomplexes. *Trends Cell Biol* [Internet]. 2019 Aug;29(8):660–71. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0962892419300832>
  185. Beringer M, Pisano P, Di Carlo V, Blanco E, Chammass P, Vizán P, et al. EPOP Functionally Links Elongin and Polycomb in Pluripotent Stem Cells. *Mol Cell* [Internet]. 2016 Nov;64(4):645–58. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276516306633>
  186. Davidovich C, Cech TR. The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2. *RNA* [Internet]. 2015 Dec 16;21(12):2007–22. Available from: <http://rnajournal.cshlp.org/lookup/doi/10.1261/rna.053918.115>
  187. Inoue A, Chen Z, Yin Q, Zhang Y. Maternal Eed knockout causes loss of H3K27me3 imprinting and random X inactivation in the extraembryonic cells. *Genes Dev* [Internet]. 2018 Dec 1;32(23–24):1525–36. Available from: <http://genesdev.cshlp.org/lookup/doi/10.1101/gad.318675.118>
  188. Tanay A, O'Donnell AH, Damelin M, Bestor TH. Hyperconserved CpG domains underlie Polycomb-binding sites. *Proc Natl Acad Sci* [Internet]. 2007 Mar 27;104(13):5521–6. Available from: <https://pnas.org/doi/full/10.1073/pnas.0609746104>
  189. Ku M, Koche RP, Rheinbay E, Mendenhall EM, Endoh M, Mikkelsen TS, et al. Genomewide Analysis of PRC1 and PRC2 Occupancy Identifies Two Classes of Bivalent Domains. van Steensel B, editor. *PLoS Genet* [Internet]. 2008 Oct 31;4(10):e1000242. Available from: <https://dx.plos.org/10.1371/journal.pgen.1000242>
  190. Zhang J, Zhang Y, You Q, Huang C, Zhang T, Wang M, et al. Highly enriched BEND3 prevents the premature activation of bivalent genes during differentiation. *Science* (80- ) [Internet]. 2022 Mar 4;375(6584):1053–8. Available from: <https://www.science.org/doi/10.1126/science.abm0730>
  191. Saha B, Home P, Ray S, Larson M, Paul A, Rajendran G, et al. EED and KDM6B Coordinate the First Mammalian Cell Lineage Commitment To Ensure Embryo Implantation. *Mol Cell Biol* [Internet]. 2013 Jul 1;33(14):2691–705. Available from: <https://www.tandfonline.com/doi/full/10.1128/MCB.00069-13>
  192. Illingworth RS, Hölzenspies JJ, Roske F V, Bickmore WA, Brickman JM. Polycomb enables primitive endoderm lineage priming in embryonic stem cells. *Elife* [Internet]. 2016 Oct 10;5. Available from: <https://elifesciences.org/articles/14926>
  193. Xiang Y, Zhang Y, Xu Q, Zhou C, Liu B, Du Z, et al. Epigenomic analysis of gastrulation identifies a unique chromatin state for primed pluripotency. *Nat Genet* [Internet]. 2020 Jan 1;52(1):95–105. Available from: <https://www.nature.com/articles/s41588-019-0545-1>
  194. Liu X, Wang C, Liu W, Li J, Li C, Kou X, et al. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* [Internet]. 2016 Sep 22;537(7621):558–62. Available from: <https://www.nature.com/articles/nature19362>
  195. Mas G, Blanco E, Ballaré C, Sansó M, Spill YG, Hu D, et al. Promoter bivalency favors an open chromatin architecture in embryonic stem cells. *Nat Genet* [Internet]. 2018 Oct 17;50(10):1452–62. Available from: <https://www.nature.com/articles/s41588-018-0218-5>
  196. Kumar D, Cinghu S, Oldfield AJ, Yang P, Jothi R. Decoding the function of bivalent chromatin in development and cancer. *Genome Res* [Internet]. 2021 Dec;31(12):2170–84. Available from: <http://genome.cshlp.org/lookup/doi/10.1101/gr.275736.121>
  197. Monk D, Arnaud P, Frost J, Hills FA, Stanier P, Feil R, et al. Reciprocal imprinting of human GRB10 in placental trophoblast and brain: evolutionary conservation of reversed allelic expression. *Hum Mol Genet* [Internet]. 2009 Aug 15;18(16):3066–74. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddp248>
  198. Maupetit-Méhouas S, Montibus B, Nury D, Tayama C, Wassef M, Kota SK, et al. Imprinting control regions (ICRs) are marked by mono-allelic bivalent chromatin when transcriptionally inactive. *Nucleic Acids Res* [Internet]. 2016 Jan 29;44(2):621–35. Available from: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv960>
  199. Geng Z, Gao Z. Mammalian PRC1 Complexes: Compositional Complexity and Diverse Molecular Mechanisms. *Int J Mol Sci* [Internet]. 2020 Nov 14;21(22):8594. Available from: <https://www.mdpi.com/1422-0067/21/22/8594>
  200. Gao Z, Zhang J, Bonasio R, Strino F, Sawai A, Parisi F, et al. PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. *Mol Cell* [Internet]. 2012 Feb;45(3):344–56. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276512000354>
  201. Farcas AM, Blackledge NP, Sudbery I, Long HK, McGouran JF, Rose NR, et al. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *Elife*

- [Internet]. 2012 Dec 18;1. Available from: <https://elifesciences.org/articles/00205>
202. Zhao J, Wang M, Chang L, Yu J, Song A, Liu C, et al. RYBP/YAF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2AK119ub1 during cell division. *Nat Cell Biol* [Internet]. 2020 Apr 23;22(4):439–52. Available from: <https://www.nature.com/articles/s41556-020-0484-1>
  203. Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS. Hierarchical Recruitment of Polycomb Group Silencing Complexes. *Mol Cell* [Internet]. 2004 Jun;14(5):637–46. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276504002928>
  204. Cooper S, Grijzenhout A, Underwood E, Ancelin K, Zhang T, Nesterova TB, et al. Jarid2 binds mono-ubiquitylated H2A lysine 119 to mediate crosstalk between Polycomb complexes PRC1 and PRC2. *Nat Commun* [Internet]. 2016 Nov 28;7(1):13661. Available from: <https://www.nature.com/articles/ncomms13661>
  205. Cooper S, Dienstbier M, Hassan R, Schermelleh L, Sharif J, Blackledge NP, et al. Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Rep* [Internet]. 2014 Jun;7(5):1456–70. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S221112471400299X>
  206. Pintacuda G, Wei G, Roustan C, Kirmizitas BA, Solcan N, Cerase A, et al. hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing. *Mol Cell* [Internet]. 2017 Dec;68(5):955–969.e10. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276517308729>
  207. Almeida M, Pintacuda G, Masui O, Koseki Y, Gdula M, Cerase A, et al. PCGF3/5-PRC1 initiates Polycomb recruitment in X chromosome inactivation. *Science* (80- ) [Internet]. 2017 Jun 9;356(6342):1081–4. Available from: <https://www.science.org/doi/10.1126/science.aal2512>
  208. Steven Ward W, Coffey DS. DNA Packaging and Organization in Mammalian Spermatozoa: Comparison with Somatic Cell. *Biol Reprod* [Internet]. 1991 Apr 1;44(4):569–74. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2043729>
  209. Balhorn R. The protamine family of sperm nuclear proteins. *Genome Biol* [Internet]. 2007;8(9):227. Available from: <http://genomebiology.biomedcentral.com/articles/10.1186/gb-2007-8-9-227>
  210. YAMAMOTO R, ABE K ichiro, SUZUKI Y, SUZUKI MG, AOKI F. Characterization of gene expression in mouse embryos at the 1-cell stage. *J Reprod Dev* [Internet]. 2016;62(1):87–92. Available from: [https://www.jstage.jst.go.jp/article/jrd/62/1/62\\_2015-131/\\_article](https://www.jstage.jst.go.jp/article/jrd/62/1/62_2015-131/_article)
  211. Braude P, Bolton V, Moore S. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature* [Internet]. 1988 Mar;332(6163):459–61. Available from: <http://www.nature.com/articles/332459a0>
  212. Probst AV, Okamoto I, Casanova M, El Marjou F, Le Baccon P, Almouzni G. A Strand-Specific Burst in Transcription of Pericentric Satellites Is Required for Chromocenter Formation and Early Mouse Development. *Dev Cell* [Internet]. 2010 Oct;19(4):625–38. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S153458071000420X>
  213. Hendrickson PG, Doráis JA, Grow EJ, Whiddon JL, Lim JW, Wike CL, et al. Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat Publ Gr*. 2017;49.
  214. De Iaco A, Planet E, Coluccio A, Verp S, Duc J, Trono D. DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat Genet* [Internet]. 2017 Jun 1;49(6):941–5. Available from: <http://www.nature.com/articles/ng.3858>
  215. Li L, Guo F, Gao Y, Ren Y, Yuan P, Yan L, et al. Single-cell multi-omics sequencing of human early embryos. *Nat Cell Biol* [Internet]. 2018 Jul 18;20(7):847–58. Available from: <https://www.nature.com/articles/s41556-018-0123-2>
  216. Zhou F, Wang R, Yuan P, Ren Y, Mao Y, Li R, et al. Reconstituting the transcriptome and DNA methylome landscapes of human implantation. *Nature* [Internet]. 2019;572(7771):660–4. Available from: <http://dx.doi.org/10.1038/s41586-019-1500-0>
  217. Tyser RC V., Mahammadov E, Nakanoh S, Vallier L, Scialdone A, Srinivas S. Single-cell transcriptomic characterization of a gastrulating human embryo. *Nature* [Internet]. 2021 Dec 9;600(7888):285–9. Available from: <https://www.nature.com/articles/s41586-021-04158-y>
  218. Bu G, Zhu W, Liu X, Zhang J, Yu L, Zhou K, et al. Coordination of zygotic genome activation entry and exit by H3K4me3 and H3K27me3 in porcine early embryos. *Genome Res* [Internet]. 2022 Aug;32(8):1487–501. Available from: <http://genome.cshlp.org/lookup/doi/10.1101/gr.276207.121>
  219. Zheng H, Huang B, Zhang B, Xiang Y, Du Z, Xu Q, et al. Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. *Mol Cell* [Internet]. 2016;63(6):1066–79. Available from: <http://dx.doi.org/10.1016/j.molcel.2016.08.032>
  220. Demond H, Khan S, Castillo-Fernandez J, Hanna CW, Kelsey G. Transcriptome and DNA

- methylation profiling during the NSN to SN transition in mouse oocytes. *BMC Mol Cell Biol*. 2025;26(1).
221. Veselovska L, Smallwood SA, Saadeh H, Stewart KR, Krueger F, Maupetit-Méhouas S, et al. Deep sequencing and de novo assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. *Genome Biol* [Internet]. 2015 Dec 25;16(1):209. Available from: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0769-z>
  222. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* [Internet]. 2009 Jul 14;460(7254):473–8. Available from: <https://www.nature.com/articles/nature08162>
  223. Kumar B, Navarro C, Winblad N, Schell JP, Zhao C, Weltner J, et al. Polycomb repressive complex 2 shields naïve human pluripotent cells from trophectoderm differentiation. *Nat Cell Biol* [Internet]. 2022 Jun 30;24(6):845–57. Available from: <https://www.nature.com/articles/s41556-022-00916-w>
  224. Yuan S, Gao L, Tao W, Zhan J, Lu G, Zhang J, et al. Allelic reprogramming of chromatin states in human early embryos. *Natl Sci Rev* [Internet]. 2024 Feb 1;11(3). Available from: <https://doi.org/10.1093/nsr/nwad328>
  225. Zhu Y, Yu J, Rong Y, Wu YW, Li Y, Zhang L, et al. Genomewide decoupling of H2AK119ub1 and H3K27me3 in early mouse development. *Sci Bull* [Internet]. 2021;66(24):2489–97. Available from: <https://doi.org/10.1016/j.scib.2021.06.010>
  226. Macrae TA, Fothergill-Robinson J, Ramalho-Santos M. Regulation, functions and transmission of bivalent chromatin during mammalian development. *Nat Rev Mol Cell Biol* [Internet]. 2023 Jan 26;24(1):6–26. Available from: <https://www.nature.com/articles/s41580-022-00518-2>
  227. Xu R, Li S, Wu Q, Li C, Jiang M, Guo L, et al. Stage-specific H3K9me3 occupancy ensures retrotransposon silencing in human pre-implantation embryos. *Cell Stem Cell* [Internet]. 2022 Jul;29(7):1051-1066.e8. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1934590922002508>
  228. Yu H, Chen M, Hu Y, Ou S, Yu X, Liang S, et al. Dynamic reprogramming of H3K9me3 at hominoid-specific retrotransposons during human preimplantation development. *Cell Stem Cell*. 2022;29(7):1031-1050.e12.
  229. Johnson WE. Endogenous Retroviruses. In: Mahy BW., Van Regenmortel MHV, editors. *Encyclopedia of Virology*. 3rd ed. Academic Press; 2008. p. 105–9.
  230. Jern P, Greenwood AD. Wildlife endogenous retroviruses: colonization, consequences, and cooption. *Trends Genet* [Internet]. 2024 Feb;40(2):149–59. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0168952523002536>
  231. Mager DL, Stoye JP. Mammalian Endogenous Retroviruses. Sandmeyer S, Craig N, editors. *Microbiol Spectr* [Internet]. 2015 Feb 27;3(1). Available from: <https://journals.asm.org/doi/10.1128/microbiolspec.MDNA3-0009-2014>
  232. Magiorkinis G, Gifford RJ, Katzourakis A, De Ranter J, Belshaw R. Env -less endogenous retroviruses are genomic superspreaders. *Proc Natl Acad Sci* [Internet]. 2012 May 8;109(19):7385–90. Available from: <https://pnas.org/doi/full/10.1073/pnas.1200913109>
  233. Lamprecht B, Walter K, Kreher S, Kumar R, Hummel M, Lenze D, et al. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nat Med* [Internet]. 2010 May 2;16(5):571–9. Available from: <https://www.nature.com/articles/nm.2129>
  234. Barau J, Teissandier A, Zamudio N, Roy S, Nalesso V, Héroult Y, et al. The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science* (80-) [Internet]. 2016 Nov 18;354(6314):909–12. Available from: <https://www.science.org/doi/10.1126/science.aah5143>
  235. Brind'Amour J, Kobayashi H, Richard Albert J, Shirane K, Sakashita A, Kamio A, et al. LTR retrotransposons transcribed in oocytes drive species-specific and heritable changes in DNA methylation. *Nat Commun*. 2018;9(1).
  236. Bogutz AB, Brind'Amour J, Kobayashi H, Jensen KN, Nakabayashi K, Imai H, et al. Evolution of imprinting via lineage-specific insertion of retroviral promoters. *Nat Commun* [Internet]. 2019 Dec 12;10(1):5674. Available from: <https://www.nature.com/articles/s41467-019-13662-9>
  237. Friedli M, Trono D. The Developmental Control of Transposable Elements and the Evolution of Higher Species. *Annu Rev Cell Dev Biol* [Internet]. 2015 Nov 13;31(1):429–51. Available from: <https://www.annualreviews.org/doi/10.1146/annurev-cellbio-100814-125514>
  238. Cordaux R, Batzer MA. The impact of retrotransposons on human genome evolution. *Nat Rev Genet* [Internet]. 2009 Oct;10(10):691–703. Available from: <https://www.nature.com/articles/nrg2640>

239. Franke V, Ganesh S, Karlic R, Malik R, Pasulka J, Horvat F, et al. Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes. *Genome Res* [Internet]. 2017 Aug;27(8):1384–94. Available from: <http://genome.cshlp.org/lookup/doi/10.1101/gr.216150.116>
240. Fang S, Chang KW, Lefebvre L. Roles of endogenous retroviral elements in the establishment and maintenance of imprinted gene expression. *Front Cell Dev Biol*. 2024;12(March):1–12.
241. Göke J, Ng HH. <sc>CTRL</sc> + <sc>INSERT</sc> : retrotransposons and their contribution to regulation and innovation of the transcriptome. *EMBO Rep* [Internet]. 2016 Aug 11;17(8):1131–44. Available from: <https://www.embopress.org/doi/10.15252/embr.201642743>
242. Mi S, Lee X, Li X ping, Veldman GM, Finnerty H, Racie L, et al. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* [Internet]. 2000 Feb;403(6771):785–9. Available from: <https://www.nature.com/articles/35001608>
243. Okae H, Toh H, Sato T, Hiura H, Takahashi S, Shirane K, et al. Derivation of Human Trophoblast Stem Cells. *Cell Stem Cell* [Internet]. 2018;22(1):50–63.e6. Available from: <https://doi.org/10.1016/j.stem.2017.11.004>
244. Creighton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci* [Internet]. 2010 Dec 14;107(50):21931–6. Available from: <https://pnas.org/doi/full/10.1073/pnas.1016071107>
245. Hamilton G, Lysiak J, Watson A, Lala P. Effects of colony stimulating factor-1 on human extravillous trophoblast growth and invasion. *J Endocrinol* [Internet]. 1998 Oct 1;159(1):69–77. Available from: <https://joe.bioscientifica.com/view/journals/joe/159/1/69.xml>
246. Leños-Miranda A, Navarro-Romero CS, Sillas-Pardo LJ, Ramírez-Valenzuela KL, Isordia-Salas I, Jiménez-Trejo LM. Soluble Endoglin As a Marker for Preeclampsia, Its Severity, and the Occurrence of Adverse Outcomes. *Hypertension* [Internet]. 2019 Oct;74(4):991–7. Available from: <https://www.ahajournals.org/doi/10.1161/HYPERTENSIONAHA.119.13348>
247. Mano Y, Kotani T, Shibata K, Matsumura H, Tsuda H, Sumigama S, et al. The Loss of Endoglin Promotes the Invasion of Extravillous Trophoblasts. *Endocrinology* [Internet]. 2011 Nov 1;152(11):4386–94. Available from: <https://academic.oup.com/endo/article/152/11/4386/2457098>
248. Pires ND, Grossniklaus U. Different yet similar: evolution of imprinting in flowering plants and mammals. *F1000Prime Rep* [Internet]. 2014 Aug 1;6. Available from: <https://facultyopinions.com/prime/reports/b/6/63/>
249. Zhou Y, Shearwin-Whyatt L, Li J, Song Z, Hayakawa T, Stevens D, et al. Platypus and echidna genomes reveal mammalian biology and evolution. *Nature* [Internet]. 2021 Apr 29;592(7856):756–62. Available from: <https://www.nature.com/articles/s41586-020-03039-0>
250. Renfree MB, Hore TA, Shaw G, Marshall Graves JA, Pask AJ. Evolution of Genomic Imprinting: Insights from Marsupials and Monotremes. *Annu Rev Genomics Hum Genet* [Internet]. 2009 Sep 1;10(1):241–62. Available from: <https://www.annualreviews.org/doi/10.1146/annurev-genom-082908-150026>
251. Surani MAH, Barton SC, Norris ML. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature*. 1984;308(5959):548–50.
252. McGrath J, Solter D. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell*. 1984;37(1):179–83.
253. Ghassemzadeh S, Farci F, Kang M. Hydatidiform Mole [Internet]. *StatPearls*. 2024. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31079718>
254. Ferguson-Smith AC, Sasaki H, Cattanaach BM, Surani MA. Parental-origin-specific epigenetic modification of the mouse H19 gene. *Nature* [Internet]. 1993 Apr;362(6422):751–5. Available from: <https://www.nature.com/articles/362751a0>
255. Barlow DP, Stöger R, Herrmann BG, Saito K, Schweifer N. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature* [Internet]. 1991 Jan;349(6304):84–7. Available from: <https://www.nature.com/articles/349084a0>
256. Ferguson-Smith AC, Cattanaach BM, Barton SC, Beechey C V., Surani MA. Embryological and molecular investigations of parental imprinting on mouse chromosome 7. *Nature* [Internet]. 1991 Jun;351(6328):667–70. Available from: <https://www.nature.com/articles/351667a0>
257. DeChiara TM, Robertson EJ, Efstratiadis A. Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* [Internet]. 1991 Feb;64(4):849–59. Available from: <https://linkinghub.elsevier.com/retrieve/pii/009286749190513X>
258. Bartolomei MS, Zemel S, Tilghman SM. Parental imprinting of the mouse H19 gene. *Nature*

- [Internet]. 1991 May;351(6322):153–5. Available from: <https://www.nature.com/articles/351153a0>
259. STOGER R. Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal. *Cell* [Internet]. 1993 Apr;73(1):61–71. Available from: <https://linkinghub.elsevier.com/retrieve/pii/009286749390160R>
  260. Bartolomei MS, Webber AL, Brunkow ME, Tilghman SM. Epigenetic mechanisms underlying the imprinting of the mouse *H19* gene. *Genes Dev* [Internet]. 1993 Sep;7(9):1663–73. Available from: <http://genesdev.cshlp.org/lookup/doi/10.1101/gad.7.9.1663>
  261. Giannoukakis N, Deal C, Paquette J, Goodyer CG, Polychronakos C. Parental genomic imprinting of the human *IGF2* gene. *Nat Genet* [Internet]. 1993 May;4(1):98–101. Available from: <https://www.nature.com/articles/ng0593-98>
  262. Kalscheuer VM, Mariman EC, Schepens MT, Rehder H, Ropers HH. The insulin-like growth factor type-2 receptor gene is imprinted in the mouse but not in humans. *Nat Genet* [Internet]. 1993 Sep;5(1):74–8. Available from: <https://www.nature.com/articles/ng0993-74>
  263. Zhang Y, Shields T, Crenshaw T, Hao Y, Moulton T, Tycko B. Imprinting of human *H19*: allele-specific CpG methylation, loss of the active allele in Wilms tumor, and potential for somatic allele switching. *Am J Hum Genet* [Internet]. 1993 Jul;53(1):113–24. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8391213>
  264. Ito S, D'Alessio AC, Taranova O V., Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* [Internet]. 2010 Aug 26;466(7310):1129–33. Available from: <http://www.nature.com/articles/nature09303>
  265. Inoue A, Zhang Y. Replication-Dependent Loss of 5-Hydroxymethylcytosine in Mouse Preimplantation Embryos. *Science* (80-) [Internet]. 2011 Oct 14;334(6053):194–194. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1212483>
  266. Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* [Internet]. 2011 Sep 4;477(7366):606–10. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21892189>
  267. Nakamura T, Liu YJ, Nakashima H, Umehara H, Inoue K, Matoba S, et al. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* [Internet]. 2012 Jun 3;486(7403):415–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22722204>
  268. Wang QQ, Zhang YM, Zhong X, Li JW, An XR, Hou J. Dimethylated histone H3 lysine 9 is dispensable for the interaction between developmental pluripotency-associated protein 3 (*Dppa3*) and ten-eleven translocation 3 (*Tet3*) in somatic cells. *Reprod Fertil Dev* [Internet]. 2019;31(2):347. Available from: <http://www.publish.csiro.au/?paper=RD18062>
  269. Li Y, Zhang Z, Chen J, Liu W, Lai W, Liu B, et al. Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. *Nature* [Internet]. 2018;564(7734):136–40. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30487604>
  270. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. *Nature* [Internet]. 1993 Dec;366(6453):362–5. Available from: <http://www.nature.com/articles/366362a0>
  271. Fedoriw AM. Transgenic RNAi Reveals Essential Function for CTCF in *H19* Gene Imprinting. *Science* (80-) [Internet]. 2004 Jan 9;303(5655):238–40. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1090934>
  272. Dhayalan A, Tamas R, Bock I, Tattermusch A, Dimitrova E, Kudithipudi S, et al. The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Hum Mol Genet* [Internet]. 2011 Jun 1;20(11):2195–203. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddr107>
  273. Tang WWC, Kobayashi T, Irie N, Dietmann S, Surani MA. Specification and epigenetic programming of the human germ line. *Nat Rev Genet* [Internet]. 2016 Oct 30;17(10):585–600. Available from: <https://www.nature.com/articles/nrg.2016.88>
  274. Kagiwada S, Kurimoto K, Hirota T, Yamaji M, Saitou M. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J* [Internet]. 2012 Dec 14;32(3):340–53. Available from: <http://emboj.embopress.org/cgi/doi/10.1038/emboj.2012.331>
  275. Kurimoto K, Yabuta Y, Ohinata Y, Shigetani M, Yamanaka K, Saitou M. Complex genome-wide transcription dynamics orchestrated by *Blimp1* for the specification of the germ cell lineage in mice. *Genes Dev* [Internet]. 2008 Jun 15;22(12):1617–35. Available from: <http://www.genesdev.org/cgi/doi/10.1101/gad.1649908>
  276. Dawlaty MM, Breiling A, Le T, Raddatz G, Barrasa MI, Cheng AW, et al. Combined Deficiency of *Tet1* and *Tet2* Causes Epigenetic Abnormalities but Is Compatible with

- Postnatal Development. *Dev Cell* [Internet]. 2013 Feb;24(3):310–23. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1534580712005898>
277. Vincent JJ, Huang Y, Chen PY, Feng S, Calvopiña JH, Nee K, et al. Stage-Specific Roles for Tet1 and Tet2 in DNA Demethylation in Primordial Germ Cells. *Cell Stem Cell* [Internet]. 2013 Apr;12(4):470–8. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1934590913000192>
  278. Gkoutela S, Li Z, Vincent JJ, Zhang KX, Chen A, Pellegrini M, et al. The ontogeny of cKIT+ human primordial germ cells proves to be a resource for human germ line reprogramming, imprint erasure and in vitro differentiation. *Nat Cell Biol* [Internet]. 2013 Jan 16;15(1):113–22. Available from: <http://www.nature.com/articles/ncb2638>
  279. Eguizabal C, Herrera L, De Oñate L, Montserrat N, Hajkova P, Izpisua Belmonte JC. Characterization of the Epigenetic Changes During Human Gonadal Primordial Germ Cells Reprogramming. *Stem Cells* [Internet]. 2016 Sep;34(9):2418–28. Available from: <http://doi.wiley.com/10.1002/stem.2422>
  280. Guo F, Yan L, Guo H, Li L, Hu B, Zhao Y, et al. The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. *Cell* [Internet]. 2015 Jun;161(6):1437–52. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867415005632>
  281. Davis TL. The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Hum Mol Genet* [Internet]. 2000 Nov 1;9(19):2885–94. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/9.19.2885>
  282. Smallwood SA, Tomizawa S ichi, Krueger F, Ruf N, Carli N, Segonds-Pichon A, et al. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat Genet* [Internet]. 2011 Aug 26;43(8):811–4. Available from: <http://www.nature.com/articles/ng.864>
  283. Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, et al. Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet* [Internet]. 2007 Oct 1;16(19):2272–80. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddm179>
  284. Renfree MB, Suzuki S, Kaneko-Ishino T. The origin and evolution of genomic imprinting and viviparity in mammals. *Philos Trans R Soc B Biol Sci* [Internet]. 2013 Jan 5;368(1609):20120151. Available from: <https://royalsocietypublishing.org/doi/10.1098/rstb.2012.0151>
  285. Moore T. Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet* [Internet]. 1991 Feb;7(2):45–9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/016895259190230N>
  286. Hanna CW. Placental imprinting: Emerging mechanisms and functions. Bartolomei MS, editor. *PLOS Genet* [Internet]. 2020 Apr 23;16(4):e1008709. Available from: <http://dx.doi.org/10.1371/journal.pgen.1008709>
  287. Ishida M, Moore GE. The role of imprinted genes in humans. *Mol Aspects Med* [Internet]. 2013 Jul;34(4):826–40. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0098299712000817>
  288. Barlow DP, Bartolomei MS. Genomic Imprinting in Mammals. *Cold Spring Harb Perspect Biol* [Internet]. 2014 Feb 1;6(2):a018382–a018382. Available from: <http://cshperspectives.cshlp.org/lookup/doi/10.1101/cshperspect.a018382>
  289. Nordin M, Bergman D, Halje M, Engström W, Ward A. Epigenetic regulation of the Igf2/H19 gene cluster. *Cell Prolif* [Internet]. 2014 Jun;47(3):189–99. Available from: <http://doi.wiley.com/10.1111/cpr.12106>
  290. Latos PA, Pauler FM, Koerner M V., Senergin HB, Hudson QJ, Stocsits RR, et al. Airn Transcriptional Overlap, But Not Its lncRNA Products, Induces Imprinted Igf2r Silencing. *Science* (80- ) [Internet]. 2012 Dec 14;338(6113):1469–72. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1228110>
  291. Mancini-DiNardo D. Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev* [Internet]. 2006 May 15;20(10):1268–82. Available from: <http://www.genesdev.org/cgi/doi/10.1101/gad.1416906>
  292. Beygo J, Citro V, Sparago A, De Crescenzo A, Cerrato F, Heitmann M, et al. The molecular function and clinical phenotype of partial deletions of the IGF2/H19 imprinting control region depends on the spatial arrangement of the remaining CTCF-binding sites. *Hum Mol Genet* [Internet]. 2013 Feb 1;22(3):544–57. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/dds465>
  293. Sparago A, Cerrato F, Riccio A. Is ZFP57 binding to H19/IGF2:IG-DMR affected in Silver-Russell syndrome? *Clin Epigenetics* [Internet]. 2018 Dec 21;10(1):23. Available from: <https://clinicaepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-018-0454-7>
  294. Eggermann T, Perez de Nanclares G, Maher ER, Temple IK, Tümer Z, Monk D, et al.

- Imprinting disorders: a group of congenital disorders with overlapping patterns of molecular changes affecting imprinted loci. *Clin Epigenetics* [Internet]. 2015 Dec 14;7(1):123. Available from: <http://www.clinicalepigeneticsjournal.com/content/7/1/123>
295. Monteagudo-Sánchez A, Sánchez-Delgado M, Mora JRH, Santamaría NT, Gratacós E, Esteller M, et al. Differences in expression rather than methylation at placenta-specific imprinted loci is associated with intrauterine growth restriction. *Clin Epigenetics* [Internet]. 2019 Dec 26;11(1):35. Available from: <https://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-019-0630-4>
  296. Allach El Khattabi L, Backer S, Pinard A, Dieudonné MN, Tsatsaris V, Vaiman D, et al. A genome-wide search for new imprinted genes in the human placenta identifies DSCAM as the first imprinted gene on chromosome 21. *Eur J Hum Genet* [Internet]. 2019;27(1):49–60. Available from: <http://dx.doi.org/10.1038/s41431-018-0267-3>
  297. Greenberg MVC, Glaser J, Borsos M, Marjou F El, Walter M, Teissandier A, et al. Transient transcription in the early embryo sets an epigenetic state that programs postnatal growth. *Nat Genet* [Internet]. 2017 Jan 14;49(1):110–8. Available from: <http://www.nature.com/articles/ng.3718>
  298. Lu F, Liu Y, Inoue A, Suzuki T, Zhao K, Zhang Y. Establishing chromatin regulatory landscape during mouse preimplantation development. *Cell* [Internet]. 2016;165(6):1375–88. Available from: <http://dx.doi.org/10.1016/j.cell.2016.05.050>
  299. Okae H, Hiura H, Nishida Y, Funayama R, Tanaka S, Chiba H, et al. Re-investigation and RNA sequencing-based identification of genes with placenta-specific imprinted expression. *Hum Mol Genet* [Internet]. 2012 Feb 1;21(3):548–58. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22025075>
  300. Gigante S, Gouil Q, Lucattini A, Keniry A, Beck T, Tinning M, et al. Using long-read sequencing to detect imprinted DNA methylation. *Nucleic Acids Res* [Internet]. 2019 May 7;47(8):e46–e46. Available from: <https://academic.oup.com/nar/article/47/8/e46/5356940>
  301. Moreira de Mello JC, de Araújo ÉSS, Stabellini R, Fraga AM, de Souza JES, Sumita DR, et al. Random X inactivation and extensive mosaicism in human placenta revealed by analysis of allele-specific gene expression along the X chromosome. *PLoS One*. 2010;5(6):1–8.
  302. Moreira de Mello JC, Fernandes GR, Vibranovski MD, Pereira L V. Early X chromosome inactivation during human preimplantation development revealed by single-cell RNA-sequencing. *Sci Rep* [Internet]. 2017 Sep 7;7(1):10794. Available from: <https://www.nature.com/articles/s41598-017-11044-z>
  303. Patrat C, Ouimette JF, Rougeulle C. X chromosome inactivation in human development. *Dev*. 2020;147(1).
  304. Loda A, Collombet S, Heard E. Gene regulation in time and space during X-chromosome inactivation. *Nat Rev Mol Cell Biol*. 2022;23(4):231–49.
  305. Sagi I, De Pinho JC, Zuccaro M V., Atzmon C, Golan-Lev T, Yanuka O, et al. Distinct Imprinting Signatures and Biased Differentiation of Human Androgenetic and Parthenogenetic Embryonic Stem Cells. *Cell Stem Cell* [Internet]. 2019;25(3):419–432.e9. Available from: <https://doi.org/10.1016/j.stem.2019.06.013>
  306. Yuan S, Zhan J, Zhang J, Liu Z, Hou Z, Zhang C, et al. Human zygotic genome activation is initiated from paternal genome. *Cell Discov*. 2023;9(1).
  307. Andergassen D, Dotter CP, Wenzel D, Sigl V, Bammer PC, Muckenhuber M, et al. Mapping the mouse Allelome reveals tissue-specific regulation of allelic expression. *Elife*. 2017;6(Xci):1–29.
  308. Baran Y, Subramaniam M, Biton A, Tukiainen T, Tsang EK, Rivas MA, et al. The landscape of genomic imprinting across diverse adult human tissues. *Genome Res*. 2015;25(7):927–36.
  309. Babak T, DeVeale B, Tsang EK, Zhou Y, Li X, Smith KS, et al. Genetic conflict reflected in tissue-specific maps of genomic imprinting in human and mouse. *Nat Genet* [Internet]. 2015 May 13;47(5):544–9. Available from: <https://www.nature.com/articles/ng.3274>
  310. MILLER R, PHILLIPS M, JO I, DONALDSON M, STUDEBAKER J, ADDLEMAN N, et al. High-density single-nucleotide polymorphism maps of the human genome. *Genomics* [Internet]. 2005 Aug;86(2):117–26. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0888754305001187>
  311. Takahashi S, Okae H, Kobayashi N, Kitamura A, Kumada K, Yaegashi N, et al. Loss of p57KIP2 expression confers resistance to contact inhibition in human androgenetic trophoblast stem cells. *Proc Natl Acad Sci U S A*. 2019;116(52):26606–13.
  312. KAJII T, OHAMA K. Androgenetic origin of hydatidiform mole. *Nature* [Internet]. 1977 Aug;268(5621):633–4. Available from: <https://www.nature.com/articles/268633a0>
  313. Linder D, McCaw BK, Hecht F. Parthenogenic Origin of Benign Ovarian Teratomas. *N Engl J Med* [Internet]. 1975 Jan 9;292(2):63–6. Available from:

- <http://www.nejm.org/doi/abs/10.1056/NEJM197501092920202>
314. Turco MY, Moffett A. Development of the human placenta. *Development* [Internet]. 2019 Nov 15;146(22):1–14. Available from: <https://journals.biologists.com/dev/article/146/22/dev163428/223131/Development-of-the-human-placenta>
315. Cindrova-Davies T, Sferruzzi-Perri AN. Human placental development and function. *Semin Cell Dev Biol* [Internet]. 2022 Nov;131(March):66–77. Available from: <https://doi.org/10.1016/j.semcdb.2022.03.039>
316. Wadhwa P, Buss C, Entringer S, Swanson J. Developmental Origins of Health and Disease: Brief History of the Approach and Current Focus on Epigenetic Mechanisms. *Semin Reprod Med* [Internet]. 2009 Sep 26;27(05):358–68. Available from: <http://www.thieme-connect.de/DOI/DOI?10.1055/s-0029-1237424>
317. Lapehn S, Paquette AG. The Placental Epigenome as a Molecular Link Between Prenatal Exposures and Fetal Health Outcomes Through the DOHAD Hypothesis. *Curr Environ Heal Reports* [Internet]. 2022;9(3):490–501. Available from: <https://doi.org/10.1007/s40572-022-00354-8>
318. Aplin JD, Ruane PT. Embryo–epithelium interactions during implantation at a glance. Ewald A, editor. *J Cell Sci* [Internet]. 2017 Jan 1;130(1):15–22. Available from: <https://journals.biologists.com/jcs/article/130/1/15/56308/Embryo-epithelium-interactions-during-implantation>
319. Genbacev OD, Prakobphol A, Foulk RA, Krtolica AR, Ilic D, Singer MS, et al. Trophoblast L-Selectin-Mediated Adhesion at the Maternal-Fetal Interface. *Science* (80- ) [Internet]. 2003 Jan 17;299(5605):405–8. Available from: <https://www.science.org/doi/10.1126/science.1079546>
320. Nejabatkhsh R, Kabir-Salmani M, Dimitriadis E, Hosseini A, Taheripanah R, Sadeghi Y, et al. Subcellular localization of L-selectin ligand in the endometrium implies a novel function for pinopodes in endometrial receptivity. *Reprod Biol Endocrinol* [Internet]. 2012 Dec 15;10(1):46. Available from: <https://rbej.biomedcentral.com/articles/10.1186/1477-7827-10-46>
321. Chobotova K, Spyropoulou I, Carver J, Manek S, Heath JK, Gullick WJ, et al. Heparin-binding epidermal growth factor and its receptor ErbB4 mediate implantation of the human blastocyst. *Mech Dev* [Internet]. 2002 Dec;119(2):137–44. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0925477302003428>
322. Raab G, Kover K, Paria BC, Dey SK, Ezzell RM, Klagsbrun M. Mouse preimplantation blastocysts adhere to cells expressing the transmembrane form of heparin-binding EGF-like growth factor. *Development* [Internet]. 1996 Feb 1;122(2):637–45. Available from: <https://journals.biologists.com/dev/article/122/2/637/38973/Mouse-preimplantation-blastocysts-adhere-to-cells>
323. Cheong ML, Wang LJ, Chuang PY, Chang CW, Lee YS, Lo HF, et al. A Positive Feedback Loop between Glial Cells Missing 1 and Human Chorionic Gonadotropin (hCG) Regulates Placental hCG $\beta$  Expression and Cell Differentiation. *Mol Cell Biol* [Internet]. 2016 Jan 1;36(1):197–209. Available from: <https://www.tandfonline.com/doi/full/10.1128/MCB.00655-15>
324. Liang CY, Wang LJ, Chen CP, Chen LF, Chen YH, Chen H. GCM1 Regulation of the Expression of Syncytin 2 and Its Cognate Receptor MFSD2A in Human Placenta1. *Biol Reprod* [Internet]. 2010 Sep 1;83(3):387–95. Available from: <https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.110.083915>
325. Lu X, Wang R, Zhu C, Wang H, Lin HY, Gu Y, et al. Fine-Tuned and Cell-Cycle-Restricted Expression of Fusogenic Protein Syncytin-2 Maintains Functional Placental Syncytia. *Cell Rep* [Internet]. 2017 Oct;21(5):1150–9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2211124717314511>
326. Aghababaei M, Hogg K, Perdu S, Robinson WP, Beristain AG. ADAM12-directed ectodomain shedding of E-cadherin potentiates trophoblast fusion. *Cell Death Differ* [Internet]. 2015 Dec 24;22(12):1970–84. Available from: <https://www.nature.com/articles/cdd201544>
327. Evain-Brion D, Malassine A. Human placenta as an endocrine organ. *Growth Horm IGF Res* [Internet]. 2003 Aug;13:S34–7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1096637403000534>
328. Vento-Tormo R, Efremova M, Botting RA, Turco MY, Vento-Tormo M, Meyer KB, et al. Single-cell reconstruction of the early maternal–fetal interface in humans. *Nature*. 2018;563(7731):347–53.
329. Wang M, Liu Y, Sun R, Liu F, Li J, Yan L, et al. Single-nucleus multi-omic profiling of human placental syncytiotrophoblasts identifies cellular trajectories during pregnancy. *Nat Genet*. 2024;

330. Ross C, Boroviak TE. Origin and function of the yolk sac in primate embryogenesis. *Nat Commun* [Internet]. 2020 Jul 28;11(1):3760. Available from: <http://dx.doi.org/10.1038/s41467-020-17575-w>
331. Pham TXA, Panda A, Kagawa H, To SK, Ertekin C, Georgolopoulos G, et al. Modeling human extraembryonic mesoderm cells using naive pluripotent stem cells. *Cell Stem Cell* [Internet]. 2022;29(9):1346-1365.e10. Available from: <https://doi.org/10.1016/j.stem.2022.08.001>
332. HAMILTON WJ, BOYD JD. Development of the human placenta in the first three months of gestation. *J Anat* [Internet]. 1960 Jul;94(Pt 3):297-328. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14399291>
333. Arutyunyan A, Roberts K, Troulé K, Wong FCK, Sheridan MA, Kats I, et al. Spatial multiomics map of trophoblast development in early pregnancy. *Nature* [Internet]. 2023 Apr 6;616(7955):143-51. Available from: <https://www.nature.com/articles/s41586-023-05869-0>
334. Hertig AT, Rock J, Adams EC. A description of 34 human ova within the first 17 days of development. *Am J Anat* [Internet]. 1956 May 3;98(3):435-93. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/aja.1000980306>
335. Burton GJ, Watson AL, Hempstock J, Skepper JN, Jauniaux E. Uterine Glands Provide Histiotrophic Nutrition for the Human Fetus during the First Trimester of Pregnancy. *J Clin Endocrinol Metab* [Internet]. 2002 Jun 1;87(6):2954-9. Available from: <https://academic.oup.com/jcem/article/87/6/2954/2847536>
336. Knöfler M, Haider S, Saleh L, Pollheimer J, Gamage TKJB, James J. Human placenta and trophoblast development: key molecular mechanisms and model systems. *Cell Mol Life Sci* [Internet]. 2019 Sep 3;76(18):3479-96. Available from: <http://link.springer.com/10.1007/s00018-019-03104-6>
337. Robin C, Bollerot K, Mendes S, Haak E, Crisan M, Cerisoli F, et al. Human Placenta Is a Potent Hematopoietic Niche Containing Hematopoietic Stem and Progenitor Cells throughout Development. *Cell Stem Cell* [Internet]. 2009 Oct;5(4):385-95. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1934590909004445>
338. Aplin JD, Whittaker H, Jana Lim YT, Swietlik S, Charnock J, Jones CJP. Hemangioblastic foci in human first trimester placenta: Distribution and gestational profile. *Placenta* [Internet]. 2015 Oct;36(10):1069-77. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400415300321>
339. Zhou Y, McMaster M, Woo K, Janatpour M, Perry J, Karpanen T, et al. Vascular Endothelial Growth Factor Ligands and Receptors That Regulate Human Cytotrophoblast Survival Are Dysregulated in Severe Preeclampsia and Hemolysis, Elevated Liver Enzymes, and Low Platelets Syndrome. *Am J Pathol* [Internet]. 2002 Apr;160(4):1405-23. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0002944010625679>
340. Wang Y, Zhao S. Chapter 3 Structure of the Placenta. In: *Vascular Biology of the Placenta*. Morgan & Claypool Life Sciences; 2010.
341. Aplin JD, Myers JE, Timms K, Westwood M. Tracking placental development in health and disease. *Nat Rev Endocrinol* [Internet]. 2020 Sep 29;16(9):479-94. Available from: <https://www.taylorfrancis.com/books/9781003297468/chapters/10.1201/9781003297468-13>
342. Velicky P, Meinhardt G, Plessl K, Vondra S, Weiss T, Haslinger P, et al. Genome amplification and cellular senescence are hallmarks of human placenta development. Mundlos S, editor. *PLOS Genet* [Internet]. 2018 Oct 12;14(10):e1007698. Available from: <https://dx.plos.org/10.1371/journal.pgen.1007698>
343. Pijnenborg R, Vercruysse L, Hanssens M. The Uterine Spiral Arteries In Human Pregnancy: Facts and Controversies. *Placenta* [Internet]. 2006 Sep;27(9-10):939-58. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400405003206>
344. Chang WL, Liu YW, Dang YL, Jiang XX, Xu H, Huang X, et al. PLAC8, a new marker for human interstitial extravillous trophoblast cells, promotes their invasion and migration. *Development* [Internet]. 2018 Jan 15;145(2). Available from: <https://journals.biologists.com/dev/article/145/2/dev148932/48814/PLAC8-a-new-marker-for-human-interstitial>
345. Al-Lamki RS, Skepper JN, Burton GJ. Are human placental bed giant cells merely aggregates of small mononuclear trophoblast cells? An ultrastructural and immunocytochemical study. *Hum Reprod* [Internet]. 1999 Feb 1;14(2):496-504. Available from: <https://academic.oup.com/humrep/article-lookup/doi/10.1093/humrep/14.2.496>
346. Aplin JD, Charlton AK, Ayad S. An immunohistochemical study of human endometrial extracellular matrix during the menstrual cycle and first trimester of pregnancy. *Cell Tissue Res* [Internet]. 1988 Jul;253(1). Available from: <http://link.springer.com/10.1007/BF00221758>

347. Choudhury RH, Dunk CE, Lye SJ, Aplin JD, Harris LK, Jones RL. Extravillous Trophoblast and Endothelial Cell Crosstalk Mediates Leukocyte Infiltration to the Early Remodeling Decidual Spiral Arteriole Wall. *J Immunol* [Internet]. 2017 May 15;198(10):4115–28. Available from: <https://journals.aai.org/jimmunol/article/198/10/4115/105915/Extravillous-Trophoblast-and-Endothelial-Cell>
348. Wallace AE, Fraser R, Cartwright JE. Extravillous trophoblast and decidual natural killer cells: a remodelling partnership. *Hum Reprod Update* [Internet]. 2012 Jul 1;18(4):458–71. Available from: <https://academic.oup.com/humupd/article-lookup/doi/10.1093/humupd/dms015>
349. Burton GJ, Woods AW, Jauniaux E, Kingdom JCP. Rheological and Physiological Consequences of Conversion of the Maternal Spiral Arteries for Uteroplacental Blood Flow during Human Pregnancy. *Placenta* [Internet]. 2009 Jun;30(6):473–82. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400409000666>
350. Roberts VHJ, Morgan TK, Bednarek P, Morita M, Burton GJ, Lo JO, et al. Early first trimester uteroplacental flow and the progressive disintegration of spiral artery plugs: new insights from contrast-enhanced ultrasound and tissue histopathology. *Hum Reprod* [Internet]. 2017 Dec 1;32(12):2382–93. Available from: <http://academic.oup.com/humrep/article/32/12/2382/4616537>
351. Pijnenborg R, Dixon G, Robertson WB, Brosens I. Trophoblastic invasion of human decidua from 8 to 18 weeks of pregnancy. *Placenta* [Internet]. 1980 Jan;1(1):3–19. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400480800129>
352. Jauniaux E, Watson A, Burton G. Evaluation of respiratory gases and acid-base gradients in human fetal fluids and uteroplacental tissue between 7 and 16 weeks' gestation. *Am J Obstet Gynecol* [Internet]. 2001 Apr;184(5):998–1003. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0002937801989768>
353. Burton GJ, Cindrova-Davies T, Yung H wa, Jauniaux E. HYPOXIA AND REPRODUCTIVE HEALTH: Oxygen and development of the human placenta. *Reproduction* [Internet]. 2021 Jan;161(1):F53–65. Available from: <https://rep.bioscientifica.com/view/journals/rep/161/1/REP-20-0153.xml>
354. Hempstock J, Bao YP, Bar-Issac M, Segaren N, Watson AL, Charnock-Jones DS, et al. Intralobular Differences in Antioxidant Enzyme Expression and Activity Reflect the Pattern of Maternal Arterial Bloodflow Within the Human Placenta. *Placenta* [Internet]. 2003 May;24(5):517–23. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400402909550>
355. Teasdale F, Jean-Jacques G. Morphometric evaluation of the microvillous surface enlargement factor in the human placenta from mid-gestation to term. *Placenta* [Internet]. 1985 Sep;6(5):375–81. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S014340048580014X>
356. Robinson JM, Ackerman WE, Tewari AK, Kniss DA, Vandre DD. Isolation of highly enriched apical plasma membranes of the placental syncytiotrophoblast. *Anal Biochem* [Internet]. 2009 Apr;387(1):87–94. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0003269709000281>
357. Roopenian DC, Akilesh S. FcRn: the neonatal Fc receptor comes of age. *Nat Rev Immunol* [Internet]. 2007 Sep 17;7(9):715–25. Available from: <https://www.nature.com/articles/nri2155>
358. Mori M, Ishikawa G, Luo SS, Mishima T, Goto T, Robinson JM, et al. The Cytotrophoblast Layer of Human Chorionic Villi Becomes Thinner but Maintains Its Structural Integrity During Gestation. *Biol Reprod* [Internet]. 2007 Jan 1;76(1):164–72. Available from: <https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.106.056127>
359. Seval Y, Korgun ET, Demir R. Hofbauer Cells in Early Human Placenta: Possible Implications in Vasculogenesis and Angiogenesis. *Placenta* [Internet]. 2007 Aug;28(8–9):841–5. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400407000136>
360. Romberg SI, Kreis NN, Friemel A, Roth S, Souto AS, Hoock SC, et al. Human placental mesenchymal stromal cells are ciliated and their ciliation is compromised in preeclampsia. *BMC Med* [Internet]. 2022 Dec 27;20(1):35. Available from: <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-021-02203-1>
361. Robinson WP, Peñaherrera MS, Konwar C, Yuan V, Wilson SL. Epigenetic modifications in the human placenta. In: *Human Reproductive and Prenatal Genetics* [Internet]. Elsevier; 2023. p. 289–310. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780323913805000319>
362. Schroeder DI, Blair JD, Lott P, Yu HOK, Hong D, Crary F, et al. The human placenta methylome. *Proc Natl Acad Sci U S A*. 2013;110(15):6037–42.

363. Schroeder DI, Jayashankar K, Douglas KC, Thirkill TL, York D, Dickinson PJ, et al. Early Developmental and Evolutionary Origins of Gene Body DNA Methylation Patterns in Mammalian Placentas. *PLoS Genet*. 2015;11(8):1–20.
364. Yuan V, Hui D, Yin Y, Peñaherrera MS, Beristain AG, Robinson WP. Cell-specific characterization of the placental methylome. *BMC Genomics*. 2021;22(1):1–20.
365. Zhang B, Kim MY, Elliot GN, Zhou Y, Zhao G, Li D, et al. Human placental cytotrophoblast epigenome dynamics over gestation and alterations in placental disease. *Dev Cell [Internet]*. 2021;56(9):1238–1252.e5. Available from: <https://doi.org/10.1016/j.devcel.2021.04.001>
366. Schroeder DI, LaSalle JM. How has the Study of the Human Placenta Aided Our Understanding of Partially Methylated Genes? *Epigenomics [Internet]*. 2013 Dec 28;5(6):645–54. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf>
367. Robinson WP, Price EM. The human placental methylome. *Cold Spring Harb Perspect Med*. 2015;5(5):1–16.
368. Schroeder DI, Lott P, Korf I, LaSalle JM. Large-scale methylation domains mark a functional subset of neuronally expressed genes. *Genome Res [Internet]*. 2011 Oct;21(10):1583–91. Available from: <http://genome.cshlp.org/lookup/doi/10.1101/gr.119131.110>
369. Berman BP, Weisenberger DJ, Aman JF, Hinoue T, Ramjan Z, Liu Y, et al. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat Genet [Internet]*. 2012 Jan 27;44(1):40–6. Available from: <https://www.nature.com/articles/ng.969>
370. Hon GC, Hawkins RD, Caballero OL, Lo C, Lister R, Pelizzola M, et al. Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. *Genome Res [Internet]*. 2012 Feb;22(2):246–58. Available from: <http://genome.cshlp.org/lookup/doi/10.1101/gr.125872.111>
371. Monk D. Genomic imprinting in the human placenta. *Am J Obstet Gynecol [Internet]*. 2015 Oct;213(4):S152–62. Available from: <http://dx.doi.org/10.1016/j.ajog.2015.06.032>
372. Frost JM, Moore GE. The Importance of Imprinting in the Human Placenta. Ferguson-Smith AC, editor. *PLoS Genet [Internet]*. 2010 Jul 1;6(7):e1001015. Available from: <https://dx.plos.org/10.1371/journal.pgen.1001015>
373. Kobayashi EH, Shibata S, Oike A, Kobayashi N, Hamada H, Okae H, et al. Genomic imprinting in human placentation. *Reprod Med Biol*. 2022;21(1):1–10.
374. Fogarty NME, Burton GJ, Ferguson-Smith AC. Different epigenetic states define syncytiotrophoblast and cytotrophoblast nuclei in the trophoblast of the human placenta. *Placenta [Internet]*. 2015 Aug;36(8):796–802. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400415009248>
375. Gong S, Gaccioli F, Dopierala J, Sovio U, Cook E, Volders PJ, et al. The RNA landscape of the human placenta in health and disease. *Nat Commun [Internet]*. 2021;12(1):1–17. Available from: <http://dx.doi.org/10.1038/s41467-021-22695-y>
376. Noguer-Dance M, Abu-Amero S, Al-Khtib M, Lefevre A, Coullin P, Moore GE, et al. The primate-specific microRNA gene cluster (C19MC) is imprinted in the placenta. *Hum Mol Genet [Internet]*. 2010 Sep 15;19(18):3566–82. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddq272>
377. Malnou EC, Umlauf D, Mouysset M, Cavallé J. Imprinted MicroRNA Gene Clusters in the Evolution, Development, and Functions of Mammalian Placenta. *Front Genet [Internet]*. 2019 Jan 18;9. Available from: <https://www.frontiersin.org/article/10.3389/fgene.2018.00706/full>
378. Griffin KM, Oxford-Horrey C, Bourjeily G. Obstetric Disorders and Critical Illness. *Clin Chest Med*. 2022;43(3):471–88.
379. Dimitriadis E, Rolnik DL, Zhou W, Estrada-Gutierrez G, Koga K, Francisco RPV, et al. Pre-eclampsia. *Nat Rev Dis Prim*. 2023;9(1):1–22.
380. Poon LC, Shennan A, Hyett JA, Kapur A, Hadar E, Divakar H, et al. The International Federation of Gynecology and Obstetrics ( <sc>FIGO</sc> ) initiative on pre-eclampsia: A pragmatic guide for first-trimester screening and prevention. *Int J Gynecol Obstet [Internet]*. 2019 May 20;145(S1):1–33. Available from: <https://obgyn.onlinelibrary.wiley.com/doi/10.1002/ijgo.12802>
381. Magee LA, Brown MA, Hall DR, Gupte S, Hennessy A, Karumanchi SA, et al. The 2021 International Society for the Study of Hypertension in Pregnancy classification, diagnosis & management recommendations for international practice. *Pregnancy Hypertens [Internet]*. 2022 Mar;27:148–69. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2210778921005237>
382. Abalos E, Cuesta C, Carroli G, Qureshi Z, Widmer M, Vogel J, et al. Pre-eclampsia, eclampsia and adverse maternal and perinatal outcomes: a secondary analysis of the

- <sc>W</sc> orld <sc>H</sc> ealth <sc>O</sc> rganization Multicountry  
<sc>S</sc> urvey on <sc>M</sc> aternal and <sc>N</sc> ewborn <sc>H</sc>  
ea. BJOG An Int J Obstet Gynaecol [Internet]. 2014 Mar 18;121(s1):14–24. Available from:  
<https://obgyn.onlinelibrary.wiley.com/doi/10.1111/1471-0528.12629>
383. Arechvo A, Voicu D, Gil MM, Syngelaki A, Akolekar R, Nicolaides KH. Maternal race and pre-eclampsia: Cohort study and systematic review with meta-analysis. BJOG An Int J Obstet Gynaecol [Internet]. 2022 Nov 21;129(12):2082–93. Available from:  
<https://obgyn.onlinelibrary.wiley.com/doi/10.1111/1471-0528.17240>
384. Elawad T, Scott G, Bone JN, Elwell H, Lopez CE, Filippi V, et al. Risk factors for pre-eclampsia in clinical practice guidelines: Comparison with the evidence. BJOG An Int J Obstet Gynaecol [Internet]. 2024 Jan 22;131(1):46–62. Available from:  
<https://obgyn.onlinelibrary.wiley.com/doi/10.1111/1471-0528.17320>
385. Sibai BM. Evaluation and management of severe preeclampsia before 34 weeks' gestation. Am J Obstet Gynecol [Internet]. 2011 Sep;205(3):191–8. Available from:  
<https://linkinghub.elsevier.com/retrieve/pii/S0002937811009185>
386. Lyall F, Robson SC, Bulmer JN. Spiral Artery Remodeling and Trophoblast Invasion in Preeclampsia and Fetal Growth Restriction. Hypertension [Internet]. 2013 Dec;62(6):1046–54. Available from:  
<https://www.ahajournals.org/doi/10.1161/HYPERTENSIONAHA.113.01892>
387. Hong K, Kim SH, Cha DH, Park HJ. Defective Uteroplacental Vascular Remodeling in Preeclampsia: Key Molecular Factors Leading to Long Term Cardiovascular Disease. Int J Mol Sci [Internet]. 2021 Oct 18;22(20):11202. Available from:  
<https://www.mdpi.com/1422-0067/22/20/11202>
388. Conrad KP, Rabaglino MB, Post Uiterweer ED. Emerging role for dysregulated decidualization in the genesis of preeclampsia. Placenta [Internet]. 2017 Dec;60:119–29. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400417302898>
389. Quinn KH, Lacoursiere DY, Cui L, Bui J, Parast MM. The unique pathophysiology of early-onset severe preeclampsia: role of decidual T regulatory cells. J Reprod Immunol [Internet]. 2011 Sep;91(1–2):76–82. Available from:  
<https://linkinghub.elsevier.com/retrieve/pii/S0165037811002415>
390. Tilburgs T, Roelen DL, van der Mast BJ, de Groot-Swings GM, Kleijburg C, Scherjon SA, et al. Evidence for a Selective Migration of Fetus-Specific CD4+CD25bright Regulatory T Cells from the Peripheral Blood to the Decidua in Human Pregnancy. J Immunol [Internet]. 2008 Apr 15;180(8):5737–45. Available from:  
<https://journals.aai.org/jimmunol/article/180/8/5737/38099/Evidence-for-a-Selective-Migration-of-Fetus>
391. Shevach EM. Mechanisms of Foxp3+ T Regulatory Cell-Mediated Suppression. Immunity [Internet]. 2009 May;30(5):636–45. Available from:  
<https://linkinghub.elsevier.com/retrieve/pii/S1074761309001976>
392. Sultana Z, Maiti K, Dedman L, Smith R. Is there a role for placental senescence in the genesis of obstetric complications and fetal growth restriction? Am J Obstet Gynecol [Internet]. 2018 Feb;218(2):S762–73. Available from:  
<https://linkinghub.elsevier.com/retrieve/pii/S0002937817323323>
393. Wójtowicz A, Zembala-Szczerba M, Babczyk D, Kołodziejczyk-Pietruszka M, Lewaczyńska O, Huras H. Early- and Late-Onset Preeclampsia: A Comprehensive Cohort Study of Laboratory and Clinical Findings according to the New ISHHP Criteria. Int J Hypertens [Internet]. 2019 Sep 17;2019:1–9. Available from:  
<https://www.hindawi.com/journals/ijhy/2019/4108271/>
394. Takahashi M, Makino S, Oguma K, Imai H, Takamizu A, Koizumi A, et al. Fetal growth restriction as the initial finding of preeclampsia is a clinical predictor of maternal and neonatal prognoses: a single-center retrospective study. BMC Pregnancy Childbirth [Internet]. 2021 Dec 6;21(1):678. Available from:  
<https://bmcpregnancychildbirth.biomedcentral.com/articles/10.1186/s12884-021-04152-2>
395. Roth CJ, Haeussner E, Ruebelmann T, Koch F v., Schmitz C, Frank HG, et al. Dynamic modeling of uteroplacental blood flow in IUGR indicates vortices and elevated pressure in the intervillous space – a pilot study. Sci Rep [Internet]. 2017 Jan 19;7(1):40771. Available from: <https://www.nature.com/articles/srep40771>
396. Redman CWG, Staff AC, Roberts JM. Syncytiotrophoblast stress in preeclampsia: the convergence point for multiple pathways. Am J Obstet Gynecol [Internet]. 2022 Feb;226(2):S907–27. Available from:  
<https://linkinghub.elsevier.com/retrieve/pii/S0002937820311157>
397. Burton GJ, Yung HW, Murray AJ. Mitochondrial – Endoplasmic reticulum interactions in the trophoblast: Stress and senescence. Placenta [Internet]. 2017 Apr;52:146–55. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400416300522>

398. Burton GJ, Yung HW, Cindrova-Davies T, Charnock-Jones DS. Placental Endoplasmic Reticulum Stress and Oxidative Stress in the Pathophysiology of Unexplained Intrauterine Growth Restriction and Early Onset Preeclampsia. *Placenta* [Internet]. 2009 Mar;30:43–8. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400408003755>
399. Vennou KE, Kontou PI, Braliou GG, Bagos PG. Meta-analysis of gene expression profiles in preeclampsia. *Pregnancy Hypertens* [Internet]. 2020 Jan;19:52–60. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2210778919304799>
400. Farina A, Sekizawa A, De Sanctis P, Purwosunu Y, Okai T, Cha DH, et al. Gene expression in chorionic villous samples at 11 weeks' gestation from women destined to develop preeclampsia. *Prenat Diagn* [Internet]. 2008 Oct 15;28(10):956–61. Available from: <https://obgyn.onlinelibrary.wiley.com/doi/10.1002/pd.2109>
401. Ren Z, Gao Y, Gao Y, Liang G, Chen Q, Jiang S, et al. Distinct placental molecular processes associated with early-onset and late-onset preeclampsia. *Theranostics* [Internet]. 2021;11(10):5028–44. Available from: <https://www.thno.org/v11p5028.htm>
402. Verlohren S, Galindo A, Schlembach D, Zeisler H, Herraiz I, Moertl MG, et al. An automated method for the determination of the sFlt-1/PIGF ratio in the assessment of preeclampsia. *Am J Obstet Gynecol* [Internet]. 2010 Feb;202(2):161.e1-161.e11. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0002937809010230>
403. Blair JD, Yuen RKC, Lim BK, McFadden DE, von Dadelszen P, Robinson WP. Widespread DNA hypomethylation at gene enhancer regions in placentas associated with early-onset pre-eclampsia. *Mol Hum Reprod* [Internet]. 2013 Oct 1;19(10):697–708. Available from: <https://academic.oup.com/molehr/article-lookup/doi/10.1093/molehr/gat044>
404. Yeung KR, Chiu CL, Pidsley R, Makris A, Hennessy A, Lind JM. DNA methylation profiles in preeclampsia and healthy control placentas. *Am J Physiol Circ Physiol* [Internet]. 2016 May 15;310(10):H1295–303. Available from: <https://www.physiology.org/doi/10.1152/ajpheart.00958.2015>
405. Rahat B, Najar RA, Hamid A, Bagga R, Kaur J. The role of aberrant methylation of trophoblastic stem cell origin in the pathogenesis and diagnosis of placental disorders. *Prenat Diagn* [Internet]. 2017 Feb;37(2):133–43. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/pd.4974>
406. Lim JH, Kang YJ, Bak HJ, Kim MS, Lee HJ, Kwak DW, et al. Epigenome-wide DNA methylation profiling of preeclamptic placenta according to severe features. *Clin Epigenetics*. 2020;12(1):1–15.
407. Romanelli V, Belinchón A, Campos-Barros A, Heath KE, García-Miñaur S, Martínez-Glez V, et al. CDKN1C Mutations in HELLP/Preeclamptic Mothers of Beckwith–Wiedemann Syndrome (BWS) Patients. *Placenta* [Internet]. 2009 Jun;30(6):551–4. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400409000988>
408. Hromadnikova I, Kotlabova K, Ivankova K, Krofta L. First trimester screening of circulating C19MC microRNAs and the evaluation of their potential to predict the onset of preeclampsia and IUGR. Afink GB, editor. *PLoS One* [Internet]. 2017 Feb 9;12(2):e0171756. Available from: <https://dx.plos.org/10.1371/journal.pone.0171756>
409. Darendeliler F. IUGR: Genetic influences, metabolic problems, environmental associations/triggers, current and future management. *Best Pract Res Clin Endocrinol Metab* [Internet]. 2019 Jun;33(3):101260. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1521690X1930003X>
410. Russell Z, Quintero RA, Kontopoulos E V. Intrauterine growth restriction in monozygotic twins. *Semin Fetal Neonatal Med* [Internet]. 2007 Dec;12(6):439–49. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1744165X07000807>
411. Sharma D, Shastri S, Sharma P. Intrauterine Growth Restriction: Antenatal and Postnatal Aspects. *Clin Med Insights Pediatr* [Internet]. 2016 Jan 14;10:CMPed.S40070. Available from: <http://journals.sagepub.com/doi/10.4137/CMPed.S40070>
412. Chew LC, Osuchukwu OO, Reed DJ, Verma RP. Fetal Growth Restriction [Internet]. *StatPearls*. 2024. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/33972065>
413. Sancakli O, Darendeliler F, Bas F, Gokcay G, Disci R, Aki S, et al. Insulin, adiponectin, IGFBP-1 levels and body composition in small for gestational age born non-obese children during prepubertal ages. *Clin Endocrinol (Oxf)* [Internet]. 2008 Jul 28;69(1):88–92. Available from: <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2265.2007.03138.x>
414. Hofman PL, Cutfield WS, Robinson EM, Bergman RN, Menon RK, Sperling MA, et al. Insulin Resistance in Short Children with Intrauterine Growth Retardation 1. *J Clin Endocrinol Metab* [Internet]. 1997 Feb;82(2):402–6. Available from: <https://academic.oup.com/jcem/article-lookup/doi/10.1210/jcem.82.2.3752>
415. de Onis M, Blössner M, Villar J. Levels and patterns of intrauterine growth retardation in developing countries. *Eur J Clin Nutr* [Internet]. 1998 Jan;52 Suppl 1:S5-15. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9511014>

416. Dapkekar P, Bhalerao A, Kawathalkar A, Vijay N. Risk Factors Associated With Intrauterine Growth Restriction: A Case-Control Study. *Cureus* [Internet]. 2023 Jun 9; Available from: <https://www.cureus.com/articles/160961-risk-factors-associated-with-intrauterine-growth-restriction-a-case-control-study>
417. De Farias Aragão VM, Barbieri MA, Moura Da Silva AA, Bettiol H, Ribeiro VS. Risk Factors for Intrauterine Growth Restriction: A Comparison between Two Brazilian Cities. *Pediatr Res* [Internet]. 2005 May;57(5 Part 1):674–9. Available from: <https://www.nature.com/doi/10.1203/01.PDR.0000156504.29809.26>
418. Suhag A, Berghella V. Intrauterine Growth Restriction (IUGR): Etiology and Diagnosis. *Curr Obstet Gynecol Rep* [Internet]. 2013 Jun 23;2(2):102–11. Available from: <http://link.springer.com/10.1007/s13669-013-0041-z>
419. Roifman M, Choufani S, Turinsky AL, Drewlo S, Keating S, Brudno M, et al. Genome-wide placental DNA methylation analysis of severely growth-discordant monochorionic twins reveals novel epigenetic targets for intrauterine growth restriction. *Clin Epigenetics* [Internet]. 2016 Dec 21;8(1):70. Available from: <http://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-016-0238-x>
420. Eggermann T, Monk D, de Nancraes GP, Kagami M, Giabicani E, Riccio A, et al. Imprinting disorders. *Nat Rev Dis Prim*. 2023;9(1):1–19.
421. Lee CQE, Gardner L, Turco M, Zhao N, Murray MJ, Coleman N, et al. What Is Trophoblast? A Combination of Criteria Define Human First-Trimester Trophoblast. *Stem Cell Reports* [Internet]. 2016;6(2):257–72. Available from: <http://dx.doi.org/10.1016/j.stemcr.2016.01.006>
422. Karlsson M, Zhang C, Méar L, Zhong W, Digre A, Katona B, et al. A single-cell type transcriptomics map of human tissues. *Sci Adv* [Internet]. 2021 Jul 30;7(31). Available from: <https://www.science.org/doi/10.1126/sciadv.abh2169>
423. Ramaekers F, Huysmans A, Schaart G, Moesker O, Vooijs P. Tissue distribution of keratin 7 as monitored by a monoclonal antibody. *Exp Cell Res* [Internet]. 1987 May;170(1):235–49. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0014482787901339>
424. Apps R, Murphy SP, Fernando R, Gardner L, Ahad T, Moffett A. Human leucocyte antigen (HLA) expression of primary trophoblast cells and placental cell lines, determined using single antigen beads to characterize allotype specificities of anti-HLA antibodies. *Immunology* [Internet]. 2009 May 31;127(1):26–39. Available from: <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2567.2008.03019.x>
425. Sheridan MA, Zhao X, Fernando RC, Gardner L, Perez-Garcia V, Li Q, et al. Characterization of primary models of human trophoblast. *Dev*. 2021;148(21).
426. Suh MR, Lee Y, Kim JY, Kim SK, Moon SH, Lee JY, et al. Human embryonic stem cells express a unique set of microRNAs. *Dev Biol* [Internet]. 2004 Jun;270(2):488–98. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0012160604001381>
427. Toh H, Okae H, Shirane K, Sato T, Hamada H, Kikutake C, et al. Epigenetic dynamics of partially methylated domains in human placenta and trophoblast stem cells. *BMC Genomics*. 2024;25(1):1050.
428. Lea G, Doria-Borrell P, Ferrero-Micó A, Varma A, Simon C, Anderson H, et al. Ectopic expression of DNMT3L in human trophoblast stem cells restores features of the placental methylome. *Cell Stem Cell*. 2025;32(2):276–292.e9.
429. Dong C, Fu S, Karvas RM, Chew B, Fischer LA, Xing X, et al. A genome-wide CRISPR-Cas9 knockout screen identifies essential and growth-restricting genes in human trophoblast stem cells. *Nat Commun*. 2022;13(1):1–16.
430. Shimizu T, Oike A, Kobayashi EH, Sekiya A, Kobayashi N, Shibata S, et al. CRISPR screening in human trophoblast stem cells reveals both shared and distinct aspects of human and mouse placental development. *Proc Natl Acad Sci* [Internet]. 2023 Dec 19;120(51):2017. Available from: <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2216830120/-/DCSupplemental>. <https://doi.org/10.1073/pnas.2216830120>
431. Li X, Li ZH, Wang YX, Liu TH. A comprehensive review of human trophoblast fusion models: recent developments and challenges. *Cell Death Discov*. 2023;9(1):1–14.
432. Wilkinson AL, Zorzan I, Rugg-Gunn PJ. Epigenetic regulation of early human embryo development. *Cell Stem Cell* [Internet]. 2023;30(12):1569–84. Available from: <http://dx.doi.org/10.1016/j.stem.2023.09.010>
433. Taniguchi K, Kawai T, Hata K. Placental Development and Nutritional Environment. In 2018. p. 63–73. Available from: [http://link.springer.com/10.1007/978-981-10-5526-3\\_7](http://link.springer.com/10.1007/978-981-10-5526-3_7)
434. Coppède F. Genetics and Epigenetics of Parkinson's Disease. *Sci World J* [Internet]. 2012;2012:1–12. Available from: <http://www.hindawi.com/journals/tswj/2012/489830/>
435. Mastroeni D, Grover A, Delvaux E, Whiteside C, Coleman PD, Rogers J. Epigenetic changes in Alzheimer's disease: Decrements in DNA methylation. *Neurobiol Aging* [Internet]. 2010

- Dec;31(12):2025–37. Available from:  
<https://linkinghub.elsevier.com/retrieve/pii/S0197458008004211>
436. Ferguson-Smith AC. Genomic imprinting: The emergence of an epigenetic paradigm. *Nat Rev Genet*. 2011;12(8):565–75.
  437. Jain M, Singh M. Assisted Reproductive Technology (ART) Techniques [Internet]. *StatPearls*. 2024. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/79723>
  438. Matteo M. Assisted Reproductive Technology. In: *Practical Clinical Andrology* [Internet]. Cham: Springer International Publishing; 2023. p. 237–50. Available from:  
[https://link.springer.com/10.1007/978-3-031-11701-5\\_18](https://link.springer.com/10.1007/978-3-031-11701-5_18)
  439. World Health Organization. Infertility [Internet]. 2024. Available from:  
<https://www.who.int/news-room/fact-sheets/detail/infertility#:~:text=Infertility affects millions of people,universal health coverage benefit packages>.
  440. Peyser A, Goldman R. Reassuring outcomes for single and coupled gay men using assisted reproductive technology. *F&S Reports* [Internet]. 2022 Dec;3(4):300. Available from:  
<https://linkinghub.elsevier.com/retrieve/pii/S266633412200085X>
  441. Niakan KK, Han J, Pedersen RA, Simon C, Pera RAR. Human pre-implantation embryo development. *Development*. 2012;139(5):829–41.
  442. Gliozheni O, Hambartsoumian E, Strohmer H, Petrovskaya E, Tishkevich O, De Neubourg D, et al. ART in Europe, 2019: results generated from European registries by ESHRE. *Hum Reprod* [Internet]. 2023 Dec 4;38(12):2321–38. Available from:  
<https://academic.oup.com/humrep/article/38/12/2321/7320081>
  443. Kupka MS, Chambers GM, Dyer S, Zegers-Hochschild F, de Mouzon J, Ishihara O, et al. International Committee for Monitoring Assisted Reproductive Technology world report: assisted reproductive technology, 2015 and 2016. *Fertil Steril* [Internet]. 2024 Nov;122(5):875–93. Available from:  
<https://linkinghub.elsevier.com/retrieve/pii/S0015028224006022>
  444. Cortessis VK, Azadian M, Buxbaum J, Sanogo F, Song AY, Sriprasert I, et al. Comprehensive meta-analysis reveals association between multiple imprinting disorders and conception by assisted reproductive technology. *J Assist Reprod Genet* [Internet]. 2018 Jun 25;35(6):943–52. Available from: <http://link.springer.com/10.1007/s10815-018-1173-x>
  445. Ye M, Reyes Palomares A, Iwarsson E, Oberg AS, Rodriguez-Wallberg KA. Imprinting disorders in children conceived with assisted reproductive technology in Sweden. *Fertil Steril* [Internet]. 2024;122(4):706–14. Available from:  
<https://doi.org/10.1016/j.fertnstert.2024.05.168>
  446. Novakovic B, Lewis S, Halliday J, Kennedy J, Burgner DP, Czajko A, et al. Assisted reproductive technologies are associated with limited epigenetic variation at birth that largely resolves by adulthood. *Nat Commun*. 2019;10(1).
  447. Armes JE, McGown I, Williams M, Broomfield A, Gough K, Lehane F, et al. The placenta in Beckwith-Wiedemann syndrome: Genotype-phenotype associations, excessive extravillous trophoblast and placental mesenchymal dysplasia. *Pathology*. 2012;44(6):519–27.
  448. Soejima H, Hara S, Ohba T, Higashimoto K. Placental Mesenchymal Dysplasia and Beckwith–Wiedemann Syndrome. *Cancers (Basel)* [Internet]. 2022 Nov 12;14(22):5563. Available from: <https://www.mdpi.com/2072-6694/14/22/5563>
  449. Court F, Tayama C, Romanelli V, Martin-Trujillo A, Iglesias-Platas I, Okamura K, et al. Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. *Genome Res* [Internet]. 2014 Apr;24(4):554–69. Available from:  
<http://genome.cshlp.org/lookup/doi/10.1101/gr.164913.113>
  450. Richard Albert J, Kobayashi T, Inoue A, Monteagudo-Sánchez A, Kumamoto S, Takashima T, et al. Conservation and divergence of canonical and non-canonical imprinting in murids. *Genome Biol* [Internet]. 2023 Mar 14;24(1):48. Available from:  
<https://doi.org/10.1186/s13059-023-02869-1>
  451. Xie Z, Zhang W, Zhang Y. Loss of Slc38a4 imprinting is a major cause of mouse placenta hyperplasia in somatic cell nuclear transferred embryos at late gestation. *Cell Rep* [Internet]. 2022 Feb;38(8):110407. Available from:  
<https://linkinghub.elsevier.com/retrieve/pii/S2211124722001310>
  452. Matoba S, Nakamuta S, Miura K, Hirose M, Shiura H, Kohda T, et al. Paternal knockout of Slc38a4/SNAT4 causes placental hypoplasia associated with intrauterine growth restriction in mice. *Proc Natl Acad Sci U S A*. 2019;116(42):21047–53.
  453. Wang LY, Li ZK, Wang L Bin, Liu C, Sun XH, Feng GH, et al. Overcoming Intrinsic H3K27me3 Imprinting Barriers Improves Post-implantation Development after Somatic Cell Nuclear Transfer. *Cell Stem Cell* [Internet]. 2020;27(2):315–325.e5. Available from:  
<https://doi.org/10.1016/j.stem.2020.05.014>
  454. Dünzinger U, Haaf T, Zechner U. Conserved synteny of mammalian imprinted genes in

- chicken, frog, and fish genomes. *Cytogenet Genome Res* [Internet]. 2007;117(1–4):78–85. Available from: <https://karger.com/CGR/article/doi/10.1159/000103167>
455. Nagan N, Faulkner NE, Curtis C, Schrijver I. Laboratory Guidelines for Detection, Interpretation, and Reporting of Maternal Cell Contamination in Prenatal Analyses. *J Mol Diagnostics* [Internet]. 2011 Jan;13(1):7–11. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1525157810000279>
456. Muren NB, Barton JK. Electrochemical Assay for the Signal-On Detection of Human DNA Methyltransferase Activity. *J Am Chem Soc* [Internet]. 2013 Nov 6;135(44):16632–40. Available from: <https://pubs.acs.org/doi/10.1021/ja4085918>
457. Fulneček J, Kovařík A. How to interpret Methylation Sensitive Amplified Polymorphism (MSAP) profiles? *BMC Genet* [Internet]. 2014;15(1):2. Available from: <http://bmcgenet.biomedcentral.com/articles/10.1186/1471-2156-15-2>
458. Darst RP, Pardo CE, Ai L, Brown KD, Kladde MP. Bisulfite Sequencing of DNA. *Curr Protoc Mol Biol* [Internet]. 2010 Jul 15;91(1). Available from: <https://currentprotocols.onlinelibrary.wiley.com/doi/10.1002/0471142727.mb0709s91>
459. Aird D, Ross MG, Chen WS, Danielsson M, Fennell T, Russ C, et al. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol*. 2011;12(2).
460. Kozarewa I, Ning Z, Quail MA, Sanders MJ, Berriman M, Turner DJ. Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat Methods* [Internet]. 2009 Apr 15;6(4):291–5. Available from: <https://www.nature.com/articles/nmeth.1311>
461. Findlay I, Ray P, Quirke P, Rutherford A. Allelic drop-out and preferential amplification in single cells and human blastomeres: Implications for preimplantation diagnosis of sex and cystic fibrosis. *Mol Hum Reprod*. 1995;1(4):209–18.
462. Szilagyí A, Gelencser Z, Romero R, Xu Y, Kiraly P, Demeter A, et al. Placenta-Specific Genes, Their Regulation During Villous Trophoblast Differentiation and Dysregulation in Preterm Preeclampsia. *Int J Mol Sci* [Internet]. 2020 Jan 17;21(2):628. Available from: <https://www.mdpi.com/1422-0067/21/2/628>
463. Iglesias-Platas I, Martín Trujillo A, Court F, Monk D. Distinct promoter methylation and isoform-specific expression of RASFF1A in placental biopsies from complicated pregnancies. *Placenta* [Internet]. 2015 Apr;36(4):397–402. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400415000417>
464. Epigenetic Symmetry of DLGAP2: Pre-Implantation Maternal Methylation Switches to a Random Monoallelic Profile in Somatic Tissues. *OBM Genet* [Internet]. 2018 Jul 30;2(3):1–1. Available from: <http://www.lidsen.com/journals/genetics/genetics-02-03-026>
465. PROMEGA corporation. pGEM(R)-T and pGEM(R)-T Easy Vector Systems Technical Manual [Internet]. pGEM(R)-T and pGEM(R)-T Easy Vector Systems Technical Manual TMO42 - pgem-t and pgem-t easy vector systems protocol.pdf. USA; 2021. p. 11. Available from: [https://www.promega.co.uk/~media/files/resources/protocols/technical\\_manuals/o/pgem-t\\_and\\_pgem-t\\_easy\\_vector\\_systems\\_protocol.pdf](https://www.promega.co.uk/~media/files/resources/protocols/technical_manuals/o/pgem-t_and_pgem-t_easy_vector_systems_protocol.pdf)
466. KLIMAN HJ, NESTLER JE, SERMASI E, SANGER JM, STRAUSS JF. Purification, Characterization, and in vitro Differentiation of Cytotrophoblasts from Human Term Placentae\*. *Endocrinology* [Internet]. 1986 Apr;118(4):1567–82. Available from: <https://academic.oup.com/endo/article-lookup/doi/10.1210/endo-118-4-1567>
467. Maxian T, Prandstetter AM, Waldhäusl H, Höbner AL, Meinhardt G, Wächter J, et al. Step-by-step protocol for isolating the entire repertoire of human first trimester placental cells. *Placenta* [Internet]. 2024 Jun; Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400424002741>
468. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* [Internet]. 2008;3(6):1101–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18546601>
469. Busato F, Tost J. DNA Methylation Analysis by Pyrosequencing. *Encyclopedia of Metagenomics*. 2014. p. 1–11.
470. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods* [Internet]. 2016 Mar 11;13(3):229–32. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26752769>
471. Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods* [Internet]. 2015 Jun 27;12(6):519–22. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25915121>
472. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*. 2014;9(1):171–81.
473. Picelli S, Björklund ÅK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods*. 2013;10(11):1096–

- 100.
474. Raney BJ, Barber GP, Benet-Pagès A, Casper J, Clawson H, Cline MS, et al. The UCSC Genome Browser database: 2024 update. *Nucleic Acids Res* [Internet]. 2024 Jan 5;52(D1):D1082–8. Available from: <https://academic.oup.com/nar/article/52/D1/D1082/7416382>
475. Wang L, Zhang J, Duan J, Gao X, Zhu W, Lu X, et al. Programming and Inheritance of Parental DNA Methylation in Mammals. *Cell* [Internet]. 2014 May;157(4):979–91. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867414005340>
476. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* [Internet]. 2011 Dec 22;480(7378):490–5. Available from: <https://www.nature.com/articles/nature10716>
477. Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD, et al. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nat Genet* [Internet]. 2013 Oct 1;45(10):1198–206. Available from: <https://www.nature.com/articles/ng.2746>
478. Renwick JH. THE MAPPING OF HUMAN CHROMOSOMES. *Annu Rev Genet* [Internet]. 1971 Dec;5(1):81–120. Available from: <https://www.annualreviews.org/doi/10.1146/annurev.ge.05.120171.000501>
479. Passarge E, Horsthemke B, Farber RA. Incorrect use of the term synteny. *Nat Genet* [Internet]. 1999 Dec;23(4):387–387. Available from: [https://www.nature.com/articles/ng1299\\_387a](https://www.nature.com/articles/ng1299_387a)
480. Liu D, Hunt M, Tsai IJ. Inferring synteny between genome assemblies: a systematic evaluation. *BMC Bioinformatics* [Internet]. 2018 Dec 30;19(1):26. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-018-2026-4>
481. Xiao Z, Lam HM. ShinySyn: a Shiny/R application for the interactive visualization and integration of macro- and micro-synteny data. *Bioinformatics*. 2022;38(18):4406–8.
482. Tang H, Wang X, Bowers JE, Ming R, Alam M, Paterson AH. Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. *Genome Res* [Internet]. 2008 Dec;18(12):1944–54. Available from: <http://genome.cshlp.org/lookup/doi/10.1101/gr.080978.108>
483. Molinari NA, Petrov DA, Price HJ, Smith JD, Gold JR, Vassiliadis C, et al. Synteny and Collinearity in Plant Genomes. *Science* (80- ) [Internet]. 2008;(April):486–9. Available from: <http://www.sciencemag.org/content/320/5875/486.full.pdf>
484. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res* [Internet]. 2012 Apr 1;40(7):e49–e49. Available from: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkr1293>
485. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* [Internet]. 2010 Sep;20(9):1297–303. Available from: <http://genome.cshlp.org/lookup/doi/10.1101/gr.107524.110>
486. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* [Internet]. 2013 Jan 1;29(1):15–21. Available from: <https://academic.oup.com/bioinformatics/article/29/1/15/272537>
487. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* [Internet]. 2009 Aug 15;25(16):2078–9. Available from: <https://academic.oup.com/bioinformatics/article/25/16/2078/204688>
488. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinforma* [Internet]. 2013;43(1110):11.10.1–11.10.33. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25431634>
489. DePristo MA, Banks E, Poplin R, Garimella K V, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* [Internet]. 2011 May 10;43(5):491–8. Available from: <https://www.nature.com/articles/ng.806>
490. Brouard JS, Schenkel F, Marete A, Bissonnette N. The GATK joint genotyping workflow is appropriate for calling variants in RNA-seq experiments. *J Anim Sci Biotechnol* [Internet]. 2019 Dec 21;10(1):44. Available from: <https://jasbsci.biomedcentral.com/articles/10.1186/s40104-019-0359-0>
491. Brouard JS, Bissonnette N. Variant Calling from RNA-seq Data Using the GATK Joint Genotyping Workflow. In 2022. p. 205–33. Available from: [https://link.springer.com/10.1007/978-1-0716-2293-3\\_13](https://link.springer.com/10.1007/978-1-0716-2293-3_13)
492. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for

- multiple tools and samples in a single report. *Bioinformatics* [Internet]. 2016 Oct 1;32(19):3047–8. Available from: <https://academic.oup.com/bioinformatics/article/32/19/3047/2196507>
493. Baruzzo G, Hayer KE, Kim EJ, DI Camillo B, Fitzgerald GA, Grant GR. Simulation-based comprehensive benchmarking of RNA-seq aligners. *Nat Methods*. 2017;14(2):135–9.
494. Corchete LA, Rojas EA, Alonso-López D, De Las Rivas J, Gutiérrez NC, Burguillo FJ. Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. *Sci Rep* [Internet]. 2020 Nov 12;10(1):19737. Available from: <https://www.nature.com/articles/s41598-020-76881-x>
495. Du Y, Huang Q, Arisdakessian C, Garmire LX. Evaluation of STAR and Kallisto on Single Cell RNA-Seq Data Alignment. *G3 Genes|Genomes|Genetics* [Internet]. 2020 May 1;10(5):1775–83. Available from: <https://academic.oup.com/g3journal/article/10/5/1775/6026272>
496. Dufke A, Eggermann T, Kagan KO, Hoopmann M, Elbracht M. Prenatal testing for Imprinting Disorders: A clinical perspective. *Prenat Diagn* [Internet]. 2023 Jul 5;43(8):983–92. Available from: <https://obgyn.onlinelibrary.wiley.com/doi/10.1002/pd.6400>
497. Elbracht M, Mackay D, Begemann M, Kagan KO, Eggermann T. Disturbed genomic imprinting and its relevance for human reproduction: causes and clinical consequences. *Hum Reprod Update* [Internet]. 2020 Feb 28;26(2):197–213. Available from: <https://academic.oup.com/humupd/article/26/2/197/5739589>
498. Higashimoto K, Soejima H, Saito T, Okumura K, Mukai T. Imprinting disruption of the CDKN1C/KCNQ1OT1 domain: the molecular mechanisms causing Beckwith-Wiedemann syndrome and cancer. *Cytogenet Genome Res* [Internet]. 2006;113(1–4):306–12. Available from: <https://karger.com/CGR/article/doi/10.1159/000090846>
499. Chilosi M, Piazzola E, Lestani M, Benedetti A, Guasparri I, Granchelli G, et al. Differential expression of p57kip2, a maternally imprinted cdk inhibitor, in normal human placenta and gestational trophoblastic disease. *Lab Invest* [Internet]. 1998 Mar;78(3):269–76. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9520940>
500. Li C, Fan Y, Li G, Xu X, Duan J, Li R, et al. DNA methylation reprogramming of functional elements during mammalian embryonic development. *Cell Discov* [Internet]. 2018 Aug 7;4(1):41. Available from: <https://www.nature.com/articles/s41421-018-0039-9>
501. Smith ZD, Chan MM, Humm KC, Karnik R, Mekhoubad S, Regev A, et al. DNA methylation dynamics of the human preimplantation embryo. *Nature* [Internet]. 2014 Jul 31;511(7511):611–5. Available from: <https://www.nature.com/articles/nature13581>
502. Herse F, LaMarca B, Hubel CA, Kaartokallio T, Lokki AI, Ekholm E, et al. Cytochrome P450 Subfamily 2J Polypeptide 2 Expression and Circulating Epoxyeicosatrienoic Metabolites in Preeclampsia. *Circulation* [Internet]. 2012 Dec 18;126(25):2990–9. Available from: <https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.112.127340>
503. Gascoin-Lachambre G, Buffat C, Rebouret R, Chelbi ST, Rigourd V, Mondon F, et al. Cullins in Human Intra-Uterine Growth Restriction: Expressional and Epigenetic Alterations. *Placenta* [Internet]. 2010 Feb;31(2):151–7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0143400409003634>
504. Arai T, Kasper JS, Skaar JR, Ali SH, Takahashi C, DeCaprio JA. Targeted disruption of p185/Cul7 gene results in abnormal vascular morphogenesis. *Proc Natl Acad Sci* [Internet]. 2003 Aug 19;100(17):9855–60. Available from: <https://pnas.org/doi/full/10.1073/pnas.1733908100>
505. Suryawanshi H, Morozov P, Straus A, Sahasrabudhe N, Max KEA, Garzia A, et al. A single-cell survey of the human first-trimester placenta and decidua. *Sci Adv* [Internet]. 2018 Oct 5;4(10). Available from: <https://www.science.org/doi/10.1126/sciadv.aau4788>
506. Spahn L, Barlow DP. An ICE pattern crystallizes. *Nat Genet* [Internet]. 2003 Sep;35(1):11–2. Available from: <https://www.nature.com/articles/ng0903-11>
507. Dahlet T, Argüeso Lleida A, Al Adhami H, Dumas M, Bender A, Ngondo RP, et al. Genome-wide analysis in the mouse embryo reveals the importance of DNA methylation for transcription integrity. *Nat Commun* [Internet]. 2020 Jun 19;11(1):3153. Available from: <https://www.nature.com/articles/s41467-020-16919-w>
508. Uehara R, Au Yeung WK, Toriyama K, Ohishi H, Kubo N, Toh H, et al. The DNMT3A ADD domain is required for efficient de novo DNA methylation and maternal imprinting in mouse oocytes. Sproul D, editor. *PLOS Genet* [Internet]. 2023 Aug 1;19(8):e1010855. Available from: <https://dx.plos.org/10.1371/journal.pgen.1010855>
509. Hirasawa R, Chiba H, Kaneda M, Tajima S, Li E, Jaenisch R, et al. Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev* [Internet]. 2008 Jun 15;22(12):1607–16. Available from: <http://genesdev.cshlp.org/lookup/doi/10.1101/gad.1667008>

510. Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, et al. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* [Internet]. 2004 Jun 24;429(6994):900–3. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15215868>
511. Santini L, Halbritter F, Titz-Teixeira F, Suzuki T, Asami M, Ma X, et al. Genomic imprinting in mouse blastocysts is predominantly associated with H3K27me3. *Nat Commun*. 2021;12(1):1–16.
512. Higgs MJ, Hill MJ, John RM, Isles AR. Systematic investigation of imprinted gene expression and enrichment in the mouse brain explored at single-cell resolution. *BMC Genomics* [Internet]. 2022 Nov 17;23(1):754. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-022-08986-8>
513. Okae H, Matoba S, Nagashima T, Mizutani E, Inoue K, Ogonuki N, et al. RNA sequencing-based identification of aberrant imprinting in cloned mice. *Hum Mol Genet*. 2014;23(4):992–1001.
514. Lewis A, Mitsuya K, Umlauf D, Smith P, Dean W, Walter J, et al. Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat Genet*. 2004;36(12):1291–5.
515. Wagschal A, Sutherland HG, Woodfine K, Henckel A, Chebli K, Schulz R, et al. G9a Histone Methyltransferase Contributes to Imprinting in the Mouse Placenta. *Mol Cell Biol*. 2008;28(3):1104–13.
516. Regha K, Sloane MA, Huang R, Pauler FM, Warczok KE, Melikant B, et al. Active and Repressive Chromatin Are Interspersed without Spreading in an Imprinted Gene Cluster in the Mammalian Genome. *Mol Cell* [Internet]. 2007 Aug;27(3):353–66. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276507004145>
517. Hanna CW, Kelsey G. Features and mechanisms of canonical and noncanonical genomic imprinting. *Genes Dev*. 2021;38(11–12):821–34.
518. Tycko B. Allele-specific DNA methylation: beyond imprinting. *Hum Mol Genet* [Internet]. 2010 Oct 15;19(R2):R210–20. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddq376>
519. Hashimoto K, Jouhilahti EM, Töhönen V, Carninci P, Kere J, Katayama S. Embryonic LTR retrotransposons supply promoter modules to somatic tissues. *Genome Res*. 2021;31(11):1983–93.
520. Wang Q, Chow J, Hong J, Smith AF, Moreno C, Seaby P, et al. Recent acquisition of imprinting at the rodent Sfmt2 locus correlates with insertion of a large block of miRNAs. *BMC Genomics* [Internet]. 2011 Dec 21;12(1):204. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2164-12-204>
521. Zeng TB, Pierce N, Liao J, Szabó PE. H3K9 methyltransferase EHMT2/G9a controls ERVK-driven noncanonical imprinted genes. *Epigenomics*. 2021;13(16):1299–314.
522. Lindblad-Toh K. Three's company. *Nature* [Internet]. 2004 Apr 1;428(6982):475–6. Available from: <https://www.nature.com/articles/428475a>
523. Kobayashi H. Canonical and Non-canonical Genomic Imprinting in Rodents. Vol. 9, *Frontiers in Cell and Developmental Biology*. 2021.
524. Kimura Y, Hawkins MTR, McDonough MM, Jacobs LL, Flynn LJ. Corrected placement of *Mus-Rattus* fossil calibration forces precision in the molecular tree of rodents. *Sci Rep* [Internet]. 2015 Sep 28;5(1):14444. Available from: <https://www.nature.com/articles/srep14444>
525. Argelaguet R, Clark SJ, Mohammed H, Stapel LC, Krueger C, Kapourani CA, et al. Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* [Internet]. 2019;576(7787):487–91. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31827285>
526. Bansal V. A computational method for estimating the PCR duplication rate in DNA and RNA-seq experiments. *BMC Bioinformatics* [Internet]. 2017 Mar 14;18(S3):43. Available from: <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-017-1471-9>
527. Lake BB, Ai R, Kaeser GE, Salathia NS, Yung YC, Liu R, et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* (80-) [Internet]. 2016 Jun 24;352(6293):1586–90. Available from: <https://www.science.org/doi/10.1126/science.aaf1204>
528. Torre D, Francoeur NJ, Kalma Y, Gross Carmel I, Melo BS, Deikus G, et al. Isoform-resolved transcriptome of the human preimplantation embryo. *Nat Commun* [Internet]. 2023 Oct 30;14(1):6902. Available from: <https://www.nature.com/articles/s41467-023-42558-y>
529. Goto T. Paternal X-chromosome inactivation in human trophoblastic cells. *Mol Hum Reprod* [Internet]. 1997 Jan 1;3(1):77–80. Available from: <https://academic.oup.com/molehr/article-lookup/doi/10.1093/molehr/3.1.77>
530. Harrison KB. X-chromosome inactivation in the human cytotrophoblast. *Cytogenet Genome Res* [Internet]. 1989;52(1–2):37–41. Available from:

- <https://karger.com/CGR/article/doi/10.1159/000132835>
531. Peñaherrera MS, Jiang R, Avila L, Yuen RKC, Brown CJ, Robinson WP. Patterns of placental development evaluated by X chromosome inactivation profiling provide a basis to evaluate the origin of epigenetic variation. *Hum Reprod* [Internet]. 2012 Jun;27(6):1745–53. Available from: <https://academic.oup.com/humrep/article-lookup/doi/10.1093/humrep/des072>
532. Phung TN, Olney KC, Pinto BJ, Silasi M, Perley L, O'Bryan J, et al. X chromosome inactivation in the human placenta is patchy and distinct from adult tissues. *Hum Genet Genomics Adv* [Internet]. 2022;3(3):100121. Available from: <https://doi.org/10.1016/j.xhgg.2022.100121>
533. Initial sequencing and comparative analysis of the mouse genome. *Nature* [Internet]. 2002 Dec;420(6915):520–62. Available from: <https://www.nature.com/articles/nature01262>
534. Carrion SA, Michal JJ, Jiang Z. Imprinted Genes: Genomic Conservation, Transcriptomic Dynamics and Phenomic Significance in Health and Diseases. *Int J Biol Sci*. 2023;19(10):3128–42.
535. Edwards CA, Ferguson-Smith AC. Mechanisms regulating imprinted genes in clusters. *Curr Opin Cell Biol* [Internet]. 2007 Jun;19(3):281–9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0955067407000658>
536. Yamazawa K, Kagami M, Nagai T, Kondoh T, Onigata K, Maeyama K, et al. Molecular and clinical findings and their correlations in Silver-Russell syndrome: implications for a positive role of IGF2 in growth determination and differential imprinting regulation of the IGF2–H19 domain in bodies and placentas. *J Mol Med* [Internet]. 2008 Oct 8;86(10):1171–81. Available from: <http://link.springer.com/10.1007/s00109-008-0377-4>
537. Xie G, Si Q, Zhang G, Fan Y, Li Q, Leng P, et al. The role of imprinting genes' loss of imprints in cancers and their clinical implications. *Front Oncol* [Internet]. 2024 May 15;14. Available from: <https://www.frontiersin.org/articles/10.3389/fonc.2024.1365474/full>
538. Steenman MJC, Rainier S, Dobry CJ, Grundy P, Horon IL, Feinberg AP. Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. *Nat Genet* [Internet]. 1994 Jul 1;7(3):433–9. Available from: <https://www.nature.com/articles/ng0794-433>
539. Kasprzak A, Adamek A. Insulin-Like Growth Factor 2 (IGF2) Signaling in Colorectal Cancer—From Basic Research to Potential Clinical Applications. *Int J Mol Sci* [Internet]. 2019 Oct 3;20(19):4915. Available from: <https://www.mdpi.com/1422-0067/20/19/4915>
540. Cui H, Cruz-Correa M, Giardiello FM, Hutcheon DF, Kafonek DR, Brandenburg S, et al. Loss of IGF2 Imprinting: A Potential Marker of Colorectal Cancer Risk. *Science* (80-) [Internet]. 2003 Mar 14;299(5613):1753–5. Available from: <https://www.science.org/doi/10.1126/science.1080902>
541. Sandell LL, Guan XJ, Ingram R, Tilghman SM. Gtm, a creatine synthesis enzyme, is imprinted in mouse placenta. *Proc Natl Acad Sci* [Internet]. 2003 Apr 15;100(8):4622–7. Available from: <https://pnas.org/doi/full/10.1073/pnas.0230424100>
542. Monk D, Arnaud P, Apostolidou S, Hills FA, Kelsey G, Stanier P, et al. Limited evolutionary conservation of imprinting in the human placenta. *Proc Natl Acad Sci* [Internet]. 2006 Apr 25;103(17):6623–8. Available from: <https://pnas.org/doi/full/10.1073/pnas.0511031103>
543. Wang X, Soloway PD, Clark AG. A Survey for Novel Imprinted Genes in the Mouse Placenta by mRNA-seq. *Genetics* [Internet]. 2011 Sep 1;189(1):109–22. Available from: <https://academic.oup.com/genetics/article/189/1/109/6063983>
544. Proudhon C, Bourc'his D. Identification and resolution of artifacts in the interpretation of imprinted gene expression. *Brief Funct Genomics* [Internet]. 2010 Dec 1;9(5–6):374–84. Available from: <https://academic.oup.com/bfg/article-lookup/doi/10.1093/bfgp/elq020>
545. Inukai K, Funaki M, Ogiwara T, Katagiri H, Kanda A, Anai M, et al. p85 $\alpha$  gene generates three isoforms of regulatory subunit for phosphatidylinositol 3-kinase (PI 3-kinase), p50 $\alpha$ , p55 $\alpha$ , and p85 $\alpha$ , with different PI 3-kinase activity elevating responses to insulin. *J Biol Chem*. 1997;272(12):7873–82.
546. Inukai K, Anai M, Van Breda E, Hosaka T, Katagiri H, Funaki M, et al. A novel 55-kDa regulatory subunit for phosphatidylinositol 3-kinase structurally similar to p55PIK is generated by alternative splicing of the p85 $\alpha$  gene. *J Biol Chem*. 1996;271(10):5317–20.
547. Antonetti DA, Algenstaedt P, Kahn CR. Insulin receptor substrate 1 binds two novel splice variants of the regulatory subunit of phosphatidylinositol 3-kinase in muscle and brain. *Mol Cell Biol*. 1996;16(5):2195–203.
548. Stelzer Y, Bar S, Bartok O, Afik S, Ronen D, Kadener S, et al. Differentiation of Human Parthenogenetic Pluripotent Stem Cells Reveals Multiple Tissue- and Isoform-Specific Imprinted Transcripts. *Cell Rep* [Internet]. 2015 Apr;11(2):308–20. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2211124715002922>
549. Chotalia M, Smallwood SA, Ruf N, Dawson C, Lucifero D, Frontera M, et al. Transcription is

- required for establishment of germline methylation marks at imprinted genes. *Genes Dev* [Internet]. 2009 Jan 1;23(1):105–17. Available from: <http://genesdev.cshlp.org/lookup/doi/10.1101/gad.495809>
550. Dedeurwaerder S, Defrance M, Bizet M, Calonne E, Bontempi G, Fuks F. A comprehensive overview of Infinium HumanMethylation450 data processing. *Brief Bioinform* [Internet]. 2014 Nov 1;15(6):929–41. Available from: <https://academic.oup.com/bib/article-lookup/doi/10.1093/bib/bbt054>
  551. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. *Epigenomics* [Internet]. 2011 Dec 25;3(6):771–84. Available from: <https://www.tandfonline.com/doi/full/10.2217/epi.11.105>
  552. Choux C, Petazzi P, Sanchez-Delgado M, Hernandez Mora JR, Monteagudo A, Sagot P, et al. The hypomethylation of imprinted genes in IVF/ICSI placenta samples is associated with concomitant changes in histone modifications. *Epigenetics*. 2020;15(12):1386–95.
  553. Rahat B, Mahajan A, Bagga R, Hamid A, Kaur J. Epigenetic modifications at DMRs of placental genes are subjected to variations in normal gestation, pathological conditions and folate supplementation. *Sci Rep* [Internet]. 2017 Jan 18;7(1):40774. Available from: <https://www.nature.com/articles/srep40774>
  554. Özçelik T, Leff S, Robinson W, Donlon T, Lalonde M, Sanjines E, et al. Small nuclear ribonucleoprotein polypeptide N (SNRPN), an expressed gene in the Prader–Willi syndrome critical region. *Nat Genet* [Internet]. 1992 Dec;2(4):265–9. Available from: <https://www.nature.com/articles/ng1292-265>
  555. Ono R, Nakamura K, Inoue K, Naruse M, Usami T, Wakisaka-Saito N, et al. Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nat Genet* [Internet]. 2006 Jan 11;38(1):101–6. Available from: <https://www.nature.com/articles/ng1699>
  556. Mayer W, Hemberger M, Frank HG, Grümmer R, Winterhager E, Kaufmann P, et al. Expression of the imprinted genes MEST/Mest in human and murine placenta suggests a role in angiogenesis. *Dev Dyn* [Internet]. 2000 Jan;217(1):1–10. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10679925>
  557. Schertzer MD, Bracer KCA, Starmer J, Cherney RE, Lee DM, Salazar G, et al. IncRNA-Induced Spread of Polycomb Controlled by Genome Architecture, RNA Abundance, and CpG Island DNA. *Mol Cell* [Internet]. 2019 Aug;75(3):523–537.e10. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276519304009>
  558. Weaver JR, Bartolomei MS. Chromatin regulators of genomic imprinting. *Biochim Biophys Acta - Gene Regul Mech* [Internet]. 2014 Mar;1839(3):169–77. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1874939913001776>
  559. Li X, Wang CY. From bulk, single-cell to spatial RNA sequencing. *Int J Oral Sci* [Internet]. 2021 Dec 15;13(1):36. Available from: <https://www.nature.com/articles/s41368-021-00146-0>
  560. Choux C, Carmignac V, Bruno C, Sagot P, Vaiman D, Fauque P. The placenta: phenotypic and epigenetic modifications induced by Assisted Reproductive Technologies throughout pregnancy. *Clin Epigenetics* [Internet]. 2015 Dec 21;7(1):87. Available from: <http://dx.doi.org/10.1186/s13148-015-0120-2>
  561. Iglesias-Platas I, Martin-Trujillo A, Petazzi P, Guillaumet-Adkins A, Esteller M, Monk D. Altered expression of the imprinted transcription factor PLAGL1 deregulates a network of genes in the human IUGR placenta. *Hum Mol Genet* [Internet]. 2014 Dec 1;23(23):6275–85. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddu347>
  562. Arima T, Wake N. Establishment of the primary imprint of the HYMAI/PLAGL1 imprint control region during oogenesis. *Cytogenet Genome Res* [Internet]. 2006;113(1–4):247–52. Available from: <https://karger.com/CGR/article/doi/10.1159/000090839>
  563. Kamiya M. The cell cycle control gene ZAC/PLAGL1 is imprinted--a strong candidate gene for transient neonatal diabetes. *Hum Mol Genet* [Internet]. 2000 Feb 12;9(3):453–60. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/9.3.453>
  564. Mackay DJG, Temple IK. Transient neonatal diabetes mellitus type 1. *Am J Med Genet Part C Semin Med Genet*. 2010;154(3):335–42.
  565. Varrault A, Gueydan C, Delalbre A, Bellmann A, Houssami S, Aknin C, et al. Zac1 Regulates an Imprinted Gene Network Critically Involved in the Control of Embryonic Growth. *Dev Cell*. 2006;11(5):711–22.
  566. Fortier AL, Lopes FL, Darricarrère N, Martel J, Trasler JM. Superovulation alters the expression of imprinted genes in the midgestation mouse placenta. *Hum Mol Genet* [Internet]. 2008 Jun 1;17(11):1653–65. Available from: <https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddn055>
  567. Fortier AL, McGraw S, Lopes FL, Niles KM, Landry M, Trasler JM. Modulation of imprinted

- gene expression following superovulation. *Mol Cell Endocrinol* [Internet]. 2014 May;388(1–2):51–7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0303720714000859>
568. Moore GE, Ishida M, Demetriou C, Al-Olabi L, Leon LJ, Thomas AC, et al. The role and interaction of imprinted genes in human fetal growth. *Philos Trans R Soc B Biol Sci* [Internet]. 2015 Mar 5;370(1663):20140074. Available from: <https://royalsocietypublishing.org/doi/10.1098/rstb.2014.0074>
569. López-Tello J, Pérez-García V, Khaira J, Kusinski LC, Cooper WN, Andreani A, et al. Fetal and trophoblast PI3K p110 $\alpha$  have distinct roles in regulating resource supply to the growing fetus in mice. *Elife*. 2019;8:1–25.
570. Tsay A, Wang JC. The Role of PIK3R1 in Metabolic Function and Insulin Sensitivity. *Int J Mol Sci*. 2023;24(16):1–20.
571. Panwar V, Singh A, Bhatt M, Tonk RK, Azizov S, Raza AS, et al. Multifaceted role of mTOR (mammalian target of rapamycin) signaling pathway in human health and disease. *Signal Transduct Target Ther* [Internet]. 2023 Oct 2;8(1):375. Available from: <https://www.nature.com/articles/s41392-023-01608-z>
572. Zhang X, Heckmann BL, Campbell LE, Liu J. GoS2: A small giant controller of lipolysis and adipose-liver fatty acid flux. *Biochim Biophys Acta - Mol Cell Biol Lipids* [Internet]. 2017;1862(10):1146–54. Available from: <http://dx.doi.org/10.1016/j.bbalip.2017.06.007>
573. Heckmann BL, Zhang X, Xie X, Liu J. The Go/G1 switch gene 2 (GoS2): Regulating metabolism and beyond. *Biochim Biophys Acta - Mol Cell Biol Lipids* [Internet]. 2013 Feb;1831(2):276–81. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf>
574. RUSSELL L, FORSDYKE DR. A Human Putative Lymphocyte G 0 /G 1 Switch Gene Containing a CpG-Rich Island Encodes a Small Basic Protein with the Potential to Be Phosphorylated. *DNA Cell Biol* [Internet]. 1991 Oct;10(8):581–91. Available from: <http://www.liebertpub.com/doi/10.1089/dna.1991.10.581>
575. SIDEROVSKI DP, BLUM S, FORSDYKE RE, FORSDYKE DR. A Set of Human Putative Lymphocyte G 0 /G 1 Switch Genes Includes Genes Homologous to Rodent Cytokine and Zinc Finger Protein-Encoding Genes. *DNA Cell Biol* [Internet]. 1990 Oct;9(8):579–87. Available from: <http://www.liebertpub.com/doi/10.1089/dna.1990.9.579>
576. Yamada T, Park CS, Burns A, Nakada D, Lacorazza HD. The Cytosolic Protein GoS2 Maintains Quiescence in Hematopoietic Stem Cells. Ivanovic Z, editor. *PLoS One* [Internet]. 2012 May 31;7(5):e38280. Available from: <https://dx.plos.org/10.1371/journal.pone.0038280>
577. Cerk IK, Salzburger B, Boeszoermenyi A, Heier C, Phillip C, Romauch M, et al. A Peptide Derived from Go/G1 Switch Gene 2 Acts as Noncompetitive Inhibitor of Adipose Triglyceride Lipase. *J Biol Chem* [Internet]. 2014 Nov;289(47):32559–70. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0021925820474617>
578. Heckmann BL, Zhang X, Saarinen AM, Liu J. Regulation of Go/G1 Switch Gene 2 (GoS2) Protein Ubiquitination and Stability by Triglyceride Accumulation and ATGL Interaction. Oberer M, editor. *PLoS One* [Internet]. 2016 Jun 1;11(6):e0156742. Available from: <https://dx.plos.org/10.1371/journal.pone.0156742>
579. Welch C, Santra MK, El-Assaad W, Zhu X, Huber WE, Keys RA, et al. Identification of a Protein, GoS2, That Lacks Bcl-2 Homology Domains and Interacts with and Antagonizes Bcl-2. *Cancer Res* [Internet]. 2009 Sep 1;69(17):6782–9. Available from: <https://aacrjournals.org/cancerres/article/69/17/6782/549930/Identification-of-a-Protein-GoS2-That-Lacks-Bcl-2>
580. Liang Y, Lai S, Huang L, Li Y, Zeng S, Zhang S, et al. JAZF1 safeguards human endometrial stromal cells survival and decidualization by repressing the transcription of GoS2. *Commun Biol*. 2023;6(1):1–17.
581. Koldehoff M, Cierna B, Steckel NK, Beelen DW, Elmaagacli AH. Maternal molecular features and gene profiling of monocytes during first trimester pregnancy. *J Reprod Immunol* [Internet]. 2013;99(1–2):62–8. Available from: <http://dx.doi.org/10.1016/j.jri.2013.07.001>
582. Barrett HL, Kubala MH, Scholz Romero K, Denny KJ, Woodruff TM, McIntyre HD, et al. Placental lipases in pregnancies complicated by gestational diabetes mellitus (GDM). *PLoS One*. 2014;9(8):1–8.
583. Uhl O, Lewis RM, Hirschmugl B, Crozier S, Inskip H, Gazquez A, et al. Placental polar lipid composition is associated with placental gene expression and neonatal body composition. *Biochim Biophys Acta - Mol Cell Biol Lipids* [Internet]. 2021;1866(9):158971. Available from: <https://doi.org/10.1016/j.bbalip.2021.158971>
584. Bellazi L, Mornet E, Meurice G, Pata-Merci N, De Mazancourt P, Dieudonné MN. Genome wide expression profile in human HTR-8/Svneo trophoblastic cells in response to overexpression of placental alkaline phosphatase gene. *Placenta* [Internet].

- 2011;32(10):771–7. Available from: <http://dx.doi.org/10.1016/j.placenta.2011.06.029>
585. Vanhaesebroeck B, Guillermet-Guibert J, Graupera M, Bilanges B. The emerging mechanisms of isoform-specific PI3K signalling. *Nat Rev Mol Cell Biol.* 2010;11(5):329–41.
586. Huang CH, Mandelker D, Schmidt-Kittler O, Samuels Y, Velculescu VE, Kinzler KW, et al. The Structure of a Human p110 $\alpha$ /p85 $\alpha$  Complex Elucidates the Effects of Oncogenic PI3K $\alpha$  Mutations. *Science (80- )* [Internet]. 2007 Dec 14;318(5857):1744–8. Available from: <https://www.science.org/doi/10.1126/science.1150799>
587. Choi E, Duan C, Bai X chen. Regulation and function of insulin and insulin-like growth factor receptor signalling. *Nat Rev Mol Cell Biol* [Internet]. 2025 Feb 10; Available from: <https://www.nature.com/articles/s41580-025-00826-3>
588. Mann G, Riddell MC, Adegoke OAJ. Effects of Acute Muscle Contraction on the Key Molecules in Insulin and Akt Signaling in Skeletal Muscle in Health and in Insulin Resistant States. *Diabetology* [Internet]. 2022 Jul 28;3(3):423–46. Available from: <https://www.mdpi.com/2673-4540/3/3/32>
589. Edick AM, Auclair O, Burgos SA. Role of Grb10 in mTORC1-dependent regulation of insulin signaling and action in human skeletal muscle cells. *Am J Physiol Metab* [Internet]. 2020 Feb 1;318(2):E173–83. Available from: <https://www.physiology.org/doi/10.1152/ajpendo.00025.2019>
590. Chu CH, Tzang BS, Chen LM, Kuo CH, Cheng YC, Chen LY, et al. IGF-II/mannose-6-phosphate receptor signaling induced cell hypertrophy and atrial natriuretic peptide/BNP expression via G $\alpha$ q interaction and protein kinase C- $\alpha$ /CaMKII activation in H9c2 cardiomyoblast cells. *J Endocrinol* [Internet]. 2008 May;197(2):381–90. Available from: <https://joe.bioscientifica.com/view/journals/joe/197/2/381.xml>
591. Harris LK, Westwood M. Biology and significance of signalling pathways activated by IGF-II. *Growth Factors.* 2012;30(1):1–12.
592. Ballou LM, Lin HY, Fan G, Jiang YP, Lin RZ. Activated G $\alpha$ q Inhibits p110 $\alpha$  Phosphatidylinositol 3-Kinase and Akt. *J Biol Chem* [Internet]. 2003 Jun;278(26):23472–9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S002192582086147X>
593. DeChiara TM, Efstratiadis A, Robertsen EJ. A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature* [Internet]. 1990 May;345(6270):78–80. Available from: <https://www.nature.com/articles/345078a0>
594. Hitchins MP, Monk D, Bell GM, Ali Z, Preece MA, Stanier P, et al. Maternal repression of the human GRB10 gene in the developing central nervous system; evaluation of the role for GRB10 in Silver-Russell syndrome. *Eur J Hum Genet* [Internet]. 2001 Feb 25;9(2):82–90. Available from: <https://www.nature.com/articles/5200583>
595. Dymant DA, Smith AC, Alcantara D, Schwartzentruber JA, Basel-Vanagaite L, Curry CJ, et al. Mutations in PIK3R1 cause SHORT syndrome. *Am J Hum Genet.* 2013;93(1):158–66.
596. Klatka M, Rysz I, Kozyra K, Polak A, Kollataj W. SHORT syndrome in a two-year-old girl - Case report. *Ital J Pediatr.* 2017;43(1):1–7.
597. Innes AM, Dymant DA. SHORT Syndrome [Internet]. *GeneReviews*®. 1993. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/6407320>
598. Wang X, He A, Yip KC, Liu X, Li R. Diagnostic signature and immune characteristic of aging-related genes from placentas in Preeclampsia. *Clin Exp Hypertens.* 2022;44(8):687–94.
599. Rosario FJ, Kelly AC, Gupta MB, Powell TL, Cox L, Jansson T. Mechanistic Target of Rapamycin Complex 2 Regulation of the Primary Human Trophoblast Cell Transcriptome. *Front Cell Dev Biol.* 2021;9(November):1–17.
600. Cooper JC, Dealtry GB, Ahmed MA, Arck PC, Klapp BF, Blois SM, et al. An Impaired Breeding Phenotype in Mice with a Genetic Deletion of Beta-2 Microglobulin and Diminished MHC Class I Expression: Role in Reproductive Fitness<sup>1</sup>. *Biol Reprod* [Internet]. 2007 Aug 1;77(2):274–9. Available from: <https://academic.oup.com/biolreprod/article-lookup/doi/10.1095/biolreprod.106.057125>
601. Yamasaki K, Joh K, Ohta T, Masuzaki H, Ishimaru T, Mukai T, et al. Neurons but not glial cells show reciprocal imprinting of sense and antisense transcripts of Ube3a. *Hum Mol Genet.* 2003;12(8):837–47.
602. Judson MC, Sosa-Pagan JO, Del Cid WA, Han JE, Philpot BD. Allelic specificity of Ube3a Expression In The Mouse Brain During Postnatal Development. *J Comp Neurol* [Internet]. 2014 Jun 3;522(8):1874–96. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/cne.23507>
603. Furukawa S, Kuroda Y, Sugiyama A. A Comparison of the Histological Structure of the Placenta in Experimental Animals. *J Toxicol Pathol* [Internet]. 2014;27(1):11–8. Available from: [https://www.jstage.jst.go.jp/article/tox/27/1/27\\_2013-0060/\\_article](https://www.jstage.jst.go.jp/article/tox/27/1/27_2013-0060/_article)
604. Crossley BM, Bai J, Glaser A, Maes R, Porter E, Killian ML, et al. Guidelines for Sanger sequencing and molecular assay monitoring. *J Vet Diagnostic Investig.* 2020;32(6):767–75.

605. Al-Shuhaib MBS, Hashim HO. Mastering DNA chromatogram analysis in Sanger sequencing for reliable clinical analysis. *J Genet Eng Biotechnol* [Internet]. 2023;21(1). Available from: <https://doi.org/10.1186/s43141-023-00587-6>
606. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol* [Internet]. 2008 Oct 9;26(10):1135–45. Available from: <https://www.nature.com/articles/nbt1486>
607. Arteche-López A, Ávila-Fernández A, Romero R, Riveiro-Álvarez R, López-Martínez MA, Giménez-Pardo A, et al. Sanger sequencing is no longer always necessary based on a single-center validation of 1109 NGS variants in 825 clinical exomes. *Sci Rep* [Internet]. 2021;11(1):1–7. Available from: <https://doi.org/10.1038/s41598-021-85182-w>
608. Tukiainen T, Villani AC, Yen A, Rivas MA, Marshall JL, Satija R, et al. Landscape of X chromosome inactivation across human tissues. *Nature* [Internet]. 2017;550(7675):244–8. Available from: <http://dx.doi.org/10.1038/nature24265>
609. Leon LJ, Solanky N, Stalman SE, Demetriou C, Abu-Amero S, Stanier P, et al. A new biological and clinical resource for research into pregnancy complications: The Baby Bio Bank. *Placenta* [Internet]. 2016 Oct;46:31–7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S014340041630491X>
610. Yu SCY, Jiang P, Peng W, Cheng SH, Cheung YTT, Tse OYO, et al. Single-molecule sequencing reveals a large population of long cell-free DNA molecules in maternal plasma. *Proc Natl Acad Sci* [Internet]. 2021 Dec 14;118(50). Available from: <https://pnas.org/doi/full/10.1073/pnas.2114937118>
611. Akram KM, Kulkarni NS, Brook A, Wyles MD, Anumba DOC. Transcriptomic analysis of the human placenta reveals trophoblast dysfunction and augmented Wnt signalling associated with spontaneous preterm birth. *Front Cell Dev Biol* [Internet]. 2022 Oct 24;10. Available from: <https://www.frontiersin.org/articles/10.3389/fcell.2022.987740/full>
612. Das R, Lee YK, Strogantsev R, Jin S, Lim YC, Ng PY, et al. DNMT1 and AIM1 Imprinting in human placenta revealed through a genome-wide screen for allele-specific DNA methylation. *BMC Genomics* [Internet]. 2013 Dec 5;14(1):685. Available from: <https://bmcbgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-685>
613. Syding LA, Nickl P, Kasperek P, Sedlacek R. CRISPR/Cas9 Epigenome Editing Potential for Rare Imprinting Diseases: A Review. *Cells* [Internet]. 2020 Apr 16;9(4):993. Available from: <https://www.mdpi.com/2073-4409/9/4/993>
614. Kojima S, Shiochi N, Sato K, Yamaura M, Ito T, Yamamura N, et al. Epigenome editing reveals core DNA methylation for imprinting control in the Dlk1-Dio3 imprinted domain. *Nucleic Acids Res* [Internet]. 2022 May 20;50(9):5080–94. Available from: <https://academic.oup.com/nar/article/50/9/5080/6584436>
615. Policarpi C, Munafò M, Tsagkris S, Carlini V, Hackett JA. Systematic epigenome editing captures the context-dependent instructive function of chromatin modifications. *Nat Genet* [Internet]. 2024 Jun 9;56(6):1168–80. Available from: <https://www.nature.com/articles/s41588-024-01706-w>
616. Onyango P. Sequence and Comparative Analysis of the Mouse 1-Megabase Region Orthologous to the Human 11p15 Imprinted Domain. *Genome Res* [Internet]. 2000 Nov 1;10(11):1697–710. Available from: <http://www.genome.org/cgi/doi/10.1101/gr.161800>
617. Luo ZX, Yuan CX, Meng QJ, Ji Q. A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature* [Internet]. 2011 Aug 24;476(7361):442–5. Available from: <https://www.nature.com/articles/nature10291>
618. Inoue K, Ogonuki N, Kamimura S, Inoue H, Matoba S, Hirose M, et al. Loss of H3K27me3 imprinting in the Sfbmt2 miRNA cluster causes enlargement of cloned mouse placentas. *Nat Commun* [Internet]. 2020 May 1;11(1):2150. Available from: <https://www.nature.com/articles/s41467-020-16044-8>
619. Matoba S, Wang H, Jiang L, Lu F, Iwabuchi KA, Wu X, et al. Loss of H3K27me3 Imprinting in Somatic Cell Nuclear Transfer Embryos Disrupts Post-Implantation Development. *Cell Stem Cell* [Internet]. 2018;23(3):343–354.e5. Available from: <https://doi.org/10.1016/j.stem.2018.06.008>
620. TAKAGI N, SASAKI M. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* [Internet]. 1975 Aug;256(5519):640–2. Available from: <https://www.nature.com/articles/256640a0>
621. Auclair G, Borgel J, Sanz LA, Vallet J, Guibert S, Dumas M, et al. EHMT2 directs DNA methylation for efficient gene silencing in mouse embryos. *Genome Res* [Internet]. 2016 Feb;26(2):192–202. Available from: <http://genome.cshlp.org/lookup/doi/10.1101/gr.198291.115>
622. Chuong EB, Rumi MAK, Soares MJ, Baker JC. Endogenous retroviruses function as species-specific enhancer elements in the placenta. *Nat Genet*. 2013;45(3):325–9.
623. Sun MA, Wolf G, Wang Y, Senft AD, Ralls S, Jin J, et al. Endogenous Retroviruses Drive Lineage-Specific Regulatory Evolution across Primate and Rodent Placenta. *Mol Biol Evol*.

- 2021;38(11):4992–5004.
624. Wood AJ, Roberts RG, Monk D, Moore GE, Schulz R, Oakey RJ. A Screen for Retrotransposed Imprinted Genes Reveals an Association between X Chromosome Homology and Maternal Germ-Line Methylation. *Ferguson-Smoth AC, editor. PLoS Genet* [Internet]. 2007 Feb 9;3(2):e20. Available from: <https://dx.plos.org/10.1371/journal.pgen.0030020>
  625. Cowley M, Oakey RJ. Retrotransposition and genomic imprinting. *Brief Funct Genomics* [Internet]. 2010 Jul 1;9(4):340–6. Available from: <https://academic.oup.com/bfg/article-lookup/doi/10.1093/bfgp/elq015>
  626. Shiura H, Kitazawa M, Ishino F, Kaneko-Ishino T. Roles of retrovirus-derived PEG10 and PEG11/RTL1 in mammalian development and evolution and their involvement in human disease. *Front Cell Dev Biol* [Internet]. 2023 Sep 29;11. Available from: <https://www.frontiersin.org/articles/10.3389/fcell.2023.1273638/full>
  627. Kobayashi H, Igaki T, Kumamoto S, Tanaka K, Takashima T, Suzuki S, et al. Post-fertilization transcription initiation in an ancestral LTR retrotransposon drives lineage-specific genomic imprinting of ZDBF2 [Internet]. *eLife*. 2024. p. 2023.10.30.564869. Available from: <https://www.biorxiv.org/content/10.1101/2023.10.30.564869v2%0Ahttps://www.biorxiv.org/content/10.1101/2023.10.30.564869v2.abstract>
  628. Kindlova M, Byrne H, Kubler JM, Steane SE, Whyte JM, Borg DJ, et al. An allele-resolved nanopore-guided tour of the human placental methylome Comprehensive allele-resolved methylation profiling. 2023;(Clifton 2010).
  629. Brekke TD, Henry LA, Good JM. Genomic imprinting, disrupted placental expression, and speciation. *Evolution (N Y)* [Internet]. 2016 Dec 28;70(12):2690–703. Available from: <https://academic.oup.com/evolut/article/70/12/2690/6852334>
  630. Shi H, Strogantsev R, Takahashi N, Kazachenka A, Lorincz MC, Hemberger M, et al. ZFP57 regulation of transposable elements and gene expression within and beyond imprinted domains. *Epigenetics Chromatin* [Internet]. 2019 Dec 9;12(1):49. Available from: <https://epigeneticsandchromatin.biomedcentral.com/articles/10.1186/s13072-019-0295-4>
  631. Gopalan S, Wang Y, Harper NW, Garber M, Fazzio TG. Simultaneous profiling of multiple chromatin proteins in the same cells. *Mol Cell* [Internet]. 2021 Nov;81(22):4736–4746.e5. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S109727652100753X>
  632. Bartosovic M, Kabbe M, Castelo-Branco G. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. *Nat Biotechnol* [Internet]. 2021 Jul 12;39(7):825–35. Available from: <https://www.nature.com/articles/s41587-021-00869-9>

# Appendix

## Appendix 1. Samples used to characterise polymorphic placental imprinted genes with placenta-specific mDMRs and candidate non-canonical imprinted genes.

Term (37 > weeks), PTL – preterm late (35 to 37 weeks), PTM – preterm moderate (32 to 35 weeks), PTE – preterm extreme (<32 weeks).

PL_ID	Collection country	Gestational age	Child's phenotype	Sex	Birth weight (g)	Mother's age	Mother's height (mm)	Mother's weight (kg)	Mother's gained weight (kg)	<i>PIK3R1</i> and <i>G0S2</i> screen	Non-canonical imprinting screen
1	Spain	Term	AGA	Female	3390	33	156	60	23	Yes	No
4	Spain	Term	AGA	Female	3180	34	153	70.3	2.2	Yes	No
5	Spain	Term	AGA	Female	3355	39	172	75	9.5	Yes	Yes
6	Spain	Term	AGA	Female	3300	19	-	-	-	Yes	Yes
7	Spain	Term	AGA	Female	2800	26	154	48	14	Yes	Yes
8	Spain	Term	AGA	Female	3670	17	174	68	32	Yes	Yes
10	Spain	PTL	AGA	Female	2420	25	168	79	8.5	Yes	No
11	Spain	Term	SGA	Female	2330	34	155	65	-	Yes	No
12	Spain	Term	AGA	Male	3250	26	-	-	-	Yes	Yes
14	Spain	Term	IUGR	Female	2510	31	172	65	14.1	Yes	No
15	Spain	Term	AGA	Female	2800	36	157	68	12	No	Yes
16	Spain	Term	AGA	Female	3840	35	170	69	8	No	Yes
17	Spain	Term	AGA	Female	3300	33	154	58	16	Yes	Yes
19	Spain	Term	AGA	Male	3490	27	160	55	11	Yes	No
21	Spain	Term	AGA	Female	3280	25	157	58	10	Yes	Yes
24	Spain	Term	AGA	Male	2480	31	163	45	16	Yes	No
25	Spain	PTM	IUGR	Male	1410	35	155	55	17	No	Yes
26	Spain	PTM	AGA	Male	1980	35	155	55	17	Yes	No
30	Spain	Term	AGA	Female	3000	27	169	76	17	Yes	No
31	Spain	PTM	IUGR	Female	1730	28	157	54	16	Yes	No
33	Spain	Term	AGA	Male	3360	37	167	54	9	Yes	No

PL_ID	Collection country	Gestational age	Child's phenotype	Sex	Birth weight (g)	Mother's age	Mother's height (mm)	Mother's weight (kg)	Mother's gained weight (kg)	<i>PIK3R1</i> and <i>G0S2</i> screen	Non-canonical imprinting screen
34	Spain	Term	AGA	Male	3670	24	159	85	7	Yes	No
36	Spain	Term	AGA	Male	3480	30	167	54	10	Yes	No
38	Spain	Term	AGA	Male	3350	32	154	58	12	Yes	No
44	Spain	Term	IUGR	Male	1700	31	176	110	8	Yes	No
45	Spain	Term	SGA	Female	2740	22	169	60	21	Yes	No
46	Spain	Term	SGA	Female	2600	29	155	58	21	Yes	Yes
50	Spain	PTM	AGA	Female	2105	32	163	70	14	Yes	Yes
60	Spain	PTE	IUGR	Female	1150	46	163	61	8	Yes	No
61	Spain	Term	IUGR	Female	2330	35	163	108	9	Yes	No
62	Spain	PTE	AGA	Female	1425	37	165	82	9	Yes	No
64	Spain	Term	IUGR	Male	2160	34	167	92	4	Yes	No
65	Spain	PTL	IUGR	Male	1620	39	171	60	6	Yes	No
70	Spain	Term	AGA	Female	3170	37	165	71	18	Yes	Yes
77	Spain	PTL	AGA	Female	2070	34	171	75	15	Yes	No
88	Spain	Term	IUGR	Male	2490	22	143	47	13	Yes	No
89	Spain	PTM	AGA	Male	2410	34	164	50	11	Yes	No
90	Spain	PTM	IUGR	Male	1730	34	164	50	11	Yes	No
91	Spain	PTE	IUGR	Female	1330	34	160	80	5	Yes	No
92	Spain	Term	IUGR	Female	2410	36	161	62	4	Yes	No
93	Spain	PTM	AGA	Female	2600	35	162	50	-	Yes	No
94	Spain	PTM	AGA	Female	2030	36	165	57	12	Yes	No
95	Spain	PTM	AGA	Female	1815	35	152	62	2	Yes	Yes
96	Spain	PTM	SGA	Female	1255	38	163	53	12	No	Yes
97	Spain	PTM	AGA	Male	1690	43	159	60	14	Yes	Yes
98	Spain	PTL	IUGR	Male	1920	37	162	65	6	Yes	No
146	Spain	Term	AGA	Male	3100	34	-	-	-	Yes	No
199	Spain	PTE	IUGR	Female	-	-	-	-	-	Yes	No

PL_ID	Collection country	Gestational age	Child's phenotype	Sex	Birth weight (g)	Mother's age	Mother's height (mm)	Mother's weight (kg)	Mother's gained weight (kg)	<i>PIK3R1</i> and <i>G0S2</i> screen	Non-canonical imprinting screen
216	Spain	Term	AGA	Male	2720	33	160	62	17	Yes	No
21BR 12	UK	Term	-	-	-	-	-	-	-	Yes	No
21BR 19	UK	Term	-	-	-	-	-	-	-	No	Yes
21BR 20	UK	Term	-	-	-	-	-	-	-	No	Yes
21BR 21	UK	Term	-	-	-	-	-	-	-	No	Yes
21BR 306	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 307	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 308	UK	Term	-	-	-	-	-	-	-	No	Yes
21BR 309	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 310	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 311	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 312	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 313	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 430	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 431	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 432	UK	Term	-	-	-	-	-	-	-	Yes	Yes
21BR 433	UK	Term	-	-	-	-	-	-	-	Yes	Yes
22BR 160	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 161	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 162	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 163	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 164	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 165	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 166	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 168	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 169	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 539	UK	Term	-	-	-	-	-	-	-	Yes	No

PL_ID	Collection country	Gestational age	Child's phenotype	Sex	Birth weight (g)	Mother's age	Mother's height (mm)	Mother's weight (kg)	Mother's gained weight (kg)	<i>PIK3R1</i> and <i>G0S2</i> screen	Non-canonical imprinting screen
22BR 540	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 541	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 542	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 543	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 544	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 545	UK	Term	-	-	-	-	-	-	-	Yes	No
22BR 700	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 128	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 130	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 131	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 132	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 193	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 196	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 245	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 247	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 249	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 291	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 293	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 294	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 357	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 361	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 363	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 365	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 708	UK	Term	-	-	-	-	-	-	-	Yes	No
23BR 710	UK	Term	-	-	-	-	-	-	-	Yes	No

**Appendix 2. Placental samples used for pyrosequencing and qRT-PCR of PIK3R1 and G0S2.**

Term (37 > weeks), PTL – preterm late (35 to 37 weeks), PTM – preterm moderate (32 to 35 weeks), PTE – preterm extreme (<32 weeks).

PL_ID	Collection country	Gestational age	Child's phenotype	Sex	Birth weight (g)	Mother's age	Mother's height (mm)	Mother's weight (kg)	Mother's gained weight (kg)	Pyrosequencing	qRT-PCR
1	Spain	Term	AGA	Female	3100	33	156	60	23	Yes	Yes
2	Spain	Term	AGA	Female	3180	32	165	79	8	Yes	No
3	Spain	Term	AGA	Male	2555	25	165	88	9	No	Yes
6	Spain	Term	AGA	Female	2200	19	NA	NA	NA	Yes	Yes
7	Spain	Term	AGA	Female	2240	26	154	48	14	Yes	Yes
8	Spain	Term	AGA	Female	3215	17	174	68	32	Yes	Yes
9	Spain	Term	AGA	Male	3840	35	160	108.6	0.9	Yes	Yes
13	Spain	Term	AGA	Female	2720	36	160	66	4	Yes	Yes
14	Spain	Term	IUGR	Female	3630	31	172	65	14.1	Yes	Yes
15	Spain	Term	AGA	Female	3380	36	157	68	12	Yes	Yes
16	Spain	Term	AGA	Female	1980	35	170	69	8	Yes	Yes
18	Spain	Term	AGA	Female	3760	28	150	45	12	Yes	Yes
20	Spain	Term	AGA	Male	3670	34	170	74	15	Yes	Yes
21	Spain	Term	AGA	Female	3280	25	157	58	10	Yes	Yes
22	Spain	Term	AGA	Male	3480	25	NA	NA	NA	Yes	Yes
23	Spain	Term	AGA	Male	1535	34	167	68	15	Yes	Yes
26	Spain	PTM	AGA	Male	184	35	155	55	17	Yes	Yes
34	Spain	Term	AGA	Male	1830	24	159	85	7	Yes	Yes
36	Spain	Term	AGA	Male	940	30	167	54	10	Yes	Yes
37	Spain	Term	IUGR	Female	3590	34	156	73	11	Yes	Yes
41	Spain	PTE	AGA	Male	1700	31	162	112	4	Yes	Yes
42	Spain	PTE	AGA	Male	1620	34	176	62	11	Yes	Yes
43	Spain	PTE	AGA	Male	860	34	176	62	11	Yes	Yes
44	Spain	Term	IUGR	Male	990	31	176	110	8	Yes	Yes
45	Spain	Term	SGA	Female	2410	22	169	60	21	Yes	Yes
47	Spain	PTE	IUGR	Male	2790	37	165	56	8	Yes	Yes

PL_ID	Collection country	Gestational age	Child's phenotype	Sex	Birth weight (g)	Mother's age	Mother's height (mm)	Mother's weight (kg)	Mother's gained weight (kg)	Pyrosequencing	qRT-PCR
49	Spain	PTE	AGA	Male	1690	24	170	50	6	Yes	Yes
50	Spain	PTM	AGA	Female	3390	32	163	70	14	Yes	Yes
51	Spain	Term	IUGR	Female	3250	40	151	40	7	Yes	Yes
52	Spain	PTL	IUGR	Female	2240	39	NA	NA	NA	Yes	Yes
53	Spain	PTE	PE	Male	1120	27	NA	NA	NA	Yes	Yes
54	Spain	Term	SGA	Male	2000	33	159	54	13	Yes	Yes
55	Spain	Term	IUGR	Male	2350	30	161	63	-1	Yes	Yes
56	Spain	Term	AGA	Male	2800	32	164	57	11	Yes	Yes
58	Spain	PTE	AGA	Male	2150	46	163	61	8	Yes	Yes
59	Spain	PTE	IUGR	Male	3840	46	163	61	8	Yes	Yes
62	Spain	PTE	AGA	Female	1870	37	165	82	9	Yes	Yes
66	Spain	Term	SGA	Male	NA	39	NA	NA	NA	Yes	Yes
67	Spain	PTM	AGA	Male	3325	39	NA	NA	NA	No	Yes
68	Spain	PTE	IUGR	Female	700	26	NA	NA	NA	No	Yes
69	Spain	NA	AGA	Female	565	40	170	62	8	Yes	Yes
71	Spain	PTL	PE	Female	2105	25	157	53	13	Yes	Yes
72	Spain	PTM	AGA	Female	3300	28	171	62	16	Yes	Yes
73	Spain	PTL	IUGR	Female	1425	24	171	69	11	Yes	Yes
74	Spain	PTE	AGA	Male	1290	34	158	54	9	Yes	Yes
75	Spain	PTE	AGA	Male	2380	24	156	73	9	Yes	Yes
78	Spain	PTL	IUGR	Female	2800	34	171	75	15	Yes	Yes
79	Spain	Term	PE	Female	3670	25	164	55	21	Yes	Yes
80	Spain	PTM	PE	Female	1800	28	163	58	10	Yes	Yes
87	Spain	PTE	PE	Female	960	31	171	70	16	Yes	Yes
88	Spain	Term	IUGR	Male	2030	22	143	47	13	Yes	Yes
89	Spain	PTM	AGA	Male	NA	34	164	50	11	Yes	Yes
90	Spain	PTM	IUGR	Male	NA	34	164	50	11	Yes	Yes
91	Spain	PTE	IUGR	Female	NA	34	160	80	5	Yes	Yes

PL_ID	Collection country	Gestational age	Child's phenotype	Sex	Birth weight (g)	Mother's age	Mother's height (mm)	Mother's weight (kg)	Mother's gained weight (kg)	Pyrosequencing	qRT-PCR
93	Spain	PTM	AGA	Female	NA	35	162	50	NA	Yes	Yes
94	Spain	PTM	AGA	Female	NA	36	165	57	12	Yes	Yes
97	Spain	PTM	AGA	Male	NA	43	159	60	14	Yes	Yes
98	Spain	PTL	IUGR	Male	1680	37	162	65	6	Yes	Yes
142	Spain	PTL	IUGR	Female	2030	41	166	49	20	No	Yes
143	Spain	PTM	SGA	Female	1700	32	NA	NA	NA	Yes	Yes
144	Spain	PTM	PE	Female	620	32	NA	NA	NA	Yes	Yes
146	Spain	Term	AGA	Male	2320	34	NA	NA	NA	Yes	Yes
147	Spain	PTM	AGA	Female	1530	34	160	67	8	No	Yes
152	Spain	PTL	PE	Female	1920	41	161	100	6	Yes	Yes
154	Spain	PTL	IUGR	Male	2510	31	158	44	12	No	Yes
155	Spain	PTL	IUGR	Female	1960	35	NA	NA	NA	Yes	Yes
158	Spain	PTL	IUGR	Female	1340	32	172	60	22.7	Yes	Yes
159	Spain	PTL	AGA	Female	1770	32	172	60	22.7	Yes	Yes
160	Spain	Term	IUGR	Female	1450	38	154	46	8.2	Yes	Yes
161	Spain	Term	IUGR	Female	1670	31	160	68	11.7	Yes	Yes
162	Spain	PTM	AGA	Female	2210	37	166	56	9	Yes	Yes
163	Spain	PTL	AGA	Female	2490	30	177	78	16.5	Yes	Yes
164	Spain	PTL	AGA	Male	2520	30	177	78	16.5	Yes	Yes
165	Spain	PTL	AGA	Female	2120	30	177	78	16.5	Yes	Yes
166	Spain	PTE	AGA	NA	1150	NA	NA	NA	NA	Yes	Yes
167	Spain	PTM	PE	NA	810	NA	NA	NA	NA	Yes	Yes
170	Spain	PTL	AGA	Male	1345	44	NA	NA	NA	Yes	Yes
186	Spain	PTE	PE	Female	1330	25	173	NA	NA	Yes	Yes
191	Spain	Term	AGA	Male	2740	41	NA	NA	NA	No	Yes
216	Spain	Term	AGA	Male	2370	33	160	62	17	Yes	Yes
217	Spain	PTE	IUGR	NA	2270	NA	NA	NA	NA	Yes	Yes
222	Spain	PTE	AGA	NA	1980	NA	NA	NA	NA	Yes	Yes

<b>PL_ID</b>	<b>Collection country</b>	<b>Gestational age</b>	<b>Child's phenotype</b>	<b>Sex</b>	<b>Birth weight (g)</b>	<b>Mother's age</b>	<b>Mother's height (mm)</b>	<b>Mother's weight (kg)</b>	<b>Mother's gained weight (kg)</b>	<b>Pyrosequencing</b>	<b>qRT-PCR</b>
225	Spain	PTE	AGA	NA	2740	NA	NA	NA	NA	Yes	Yes
226	Spain	Term	AGA	NA	2690	NA	NA	NA	NA	Yes	Yes

### Appendix 3. Placental samples used for pyrosequencing and qRT-PCR - cohort summary.

Term (37 > weeks), PTL – preterm late (35 to 37 weeks), PTM – preterm moderate (32 to 35 weeks), PTE – preterm extreme (<32 weeks).

Pyrosequencing placental cohort summary:			qRT-PCR placental cohort summary:		
Placenta			Placenta		
Gestational age	PTE	18	Gestational age	PTEExt	19
	PTM	13		PTMod	15
	PTL	13		PTLeve	15
	Term	32		Term	33
	NA	1		NA	1
Child's phenotype	AGA	45	Child's phenotype	AGA	48
	IUGR	19		IUGR	22
	SGA	4		SGA	4
	PE	9		PE	9
Sex	Male	31	Sex	Male	35
	Female	40		Female	42
	NA	6		NA	6
Birth weight (g)	Min	184	Birth weight (g)	Min	184
	Mean	2194		Mean	2182
	Median	2135		Median	2135
	Max	3840		Max	3840
	NA	7		NA	7
Mother			Mother		
Mother's age	Min	17	Mother's age	Min	17
	Mean	32.21		Mean	32.36
	Median	33		Median	33
	Max	46		Max	46
	NA	6		NA	6
Mother's height	Min	143	Mother's height	Min	143
	Mean	164.1		Mean	164
	Median	164		Median	163.5
	Max	177		Max	177
	NA	16		NA	19
Mother's weight	Min	40	Mother's weight	Min	40
	Mean	64.33		Mean	64.9
	Median	62		Median	62
	Max	112		Max	112
	NA	17		NA	20
Mother's gained weight	Min	-1	Mother's gained weight	Min	-1
	Mean	11.83		Mean	11.92
	Median	11		Median	11
	Max	32.00		Max	32.00
	NA	18		NA	21

#### Appendix 4. Reagents for the media

Reagent	Supplier	Cat. No. <sup>1</sup>	Concentration / Form	Quantity	Purpose
A 83-01	Merck Life Science UK Ltd.	SML0788	powder	5 mg	Inhibits phosphorylation of SMAD2/3 (inhibits TGF- $\beta$ pathway)
CHIR99021	Merck Life Science UK Ltd.	SML1046	powder	5 mg	Promotes self-renewal of ESCs, WNT activator, (inhibits TGF- $\beta$ pathway)
Y-27632 dihydrochloride	Merck Life Science UK Ltd.	Y0503	powder	1 mg	ROCK-I and ROCK-II inhibitor, enhance ESCs survival
L-Ascorbic acid	Merck Life Science UK Ltd.	A4403	powder	100 mg	Antioxidant (hydroxyl radicals, superoxide, singlet oxygen)
EGF human	Merck Life Science UK Ltd.	SRP3027	powder	500 $\mu$ g	Stimulates growth of epidermal and epithelial cells
SB 431542 hydrate	Merck Life Science UK Ltd.	S4317	powder	5 mg	Inhibits SMAD proteins (inhibits TGF- $\beta$ pathway)
2-Mercaptoethanol	Merck Life Science UK Ltd.	M3148	14.3 M	25 mL	Reduces disulphite bonds
Valproic Acid (Sodium Salt)	Fisher Scientific	1.6E+07	powder	500 mg	Inhibiting Histone Deacetylases (HDACs)
Insulin-Transferrin-Selenium-Ethanolamine (ITS - X) (100X)	Fisher Scientific	1.1E+07	100x (liquid)	10 mL	Basel medium supplement to reduce the amount of FBS in a medium, contains insulin, transferrin, selenium and ethanolamine

<b>Reagent</b>	<b>Supplier</b>	<b>Cat. No.<sup>1</sup></b>	<b>Concentration / Form</b>	<b>Quantity</b>	<b>Purpose</b>
DMEM/F-12	Fisher Scientific	1.2E+07	RTU liquid	500 mL	Basal medium for mammalian cells (glucose, amino acids, vitamins, no growth factors)
FBS Supreme	PAN Biotech UK Ltd.	P30-3031HI	RTU liquid	500 mL	Basal medium supplement for hormone factors (improves growth proliferation and survival)
Penicillin-Streptomycin	PAN Biotech UK Ltd.	P06-07100	10,000 U/ml (penicillin)	100 mL	Prevents bacterial and fungal growth
Dimethyl sulfoxide (DMSO)	Merck Life Science UK Ltd.	D2650		100 mL	Polar solvent and cryoprotectant agent
Bovine Serum Albumin (BSA)	PAN Biotech UK Ltd.	P06-1403100	30% solution	100 mL	Improves cell growth and survival
PBS, pH 7.4	Fisher Scientific	11503387		500 mL	Salt solution to wash cells
iMatrix-511	Takara Bio Inc.	T303	0.5 mg/ml (liquid)	175 µg	Provides greater adhesion, self-renewal, promotes expression of pluripotency markers

(1) Catalog numbers

**Appendix 5. CT Basal Medium**

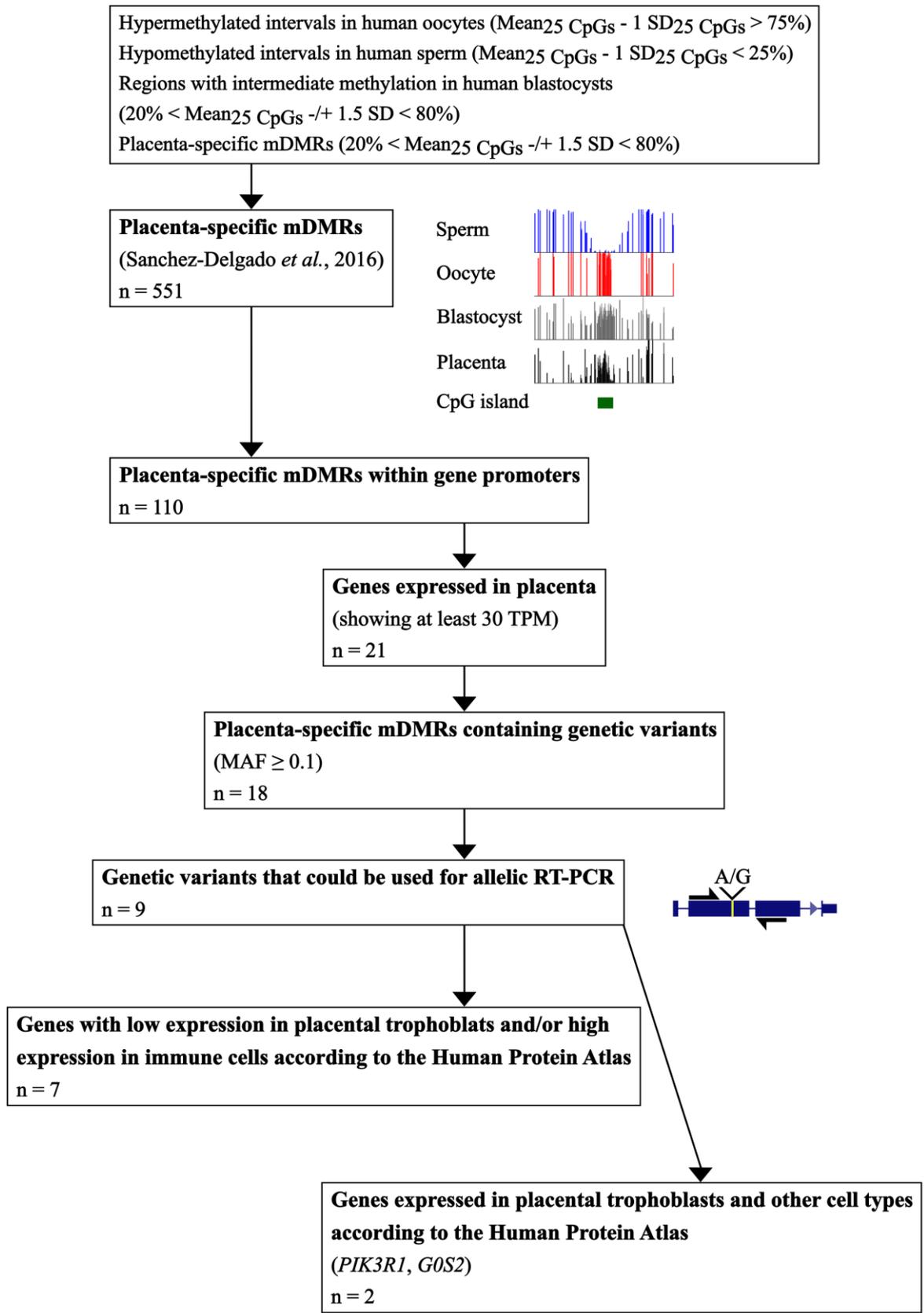
Reagent	Stock conc. <sup>1</sup>	Stock preparation	Final conc. <sup>2</sup>	Amt. for 50 mL <sup>3</sup>
DMEM/F12	n/a	500 mL (sold)	n/a	48.7 ml
BSA	30%	100 mL (sold)	0.30%	500 µl
Penicillin-Streptomycin	100%	100 mL (sold)	0.50%	250 µl
ITS-X	100%	10 mL (sold)	1.00%	500 µl
FBS	100%	500 mL (sold)	0.20%	100 µl
L-Ascorbic acid	50 mg/ml	Add 2ml H <sub>2</sub> O to whole 100mg bottle & filter sterilise	0.0015 mg/ml	1.5 µl

(1) Stock concentration; (2) Final concentration; (3) Amount for 50 mL

**Appendix 6. CT Working Medium**

Reagent	Stock conc. <sup>1</sup>	Stock preparation	Final conc. <sup>2</sup>	Amt. to add to make 40 mL <sup>3</sup>
CT Basal Medium	n/a		n/a	39.9 ml
Y-27632 Dihydrochloride	10 mM	Add 0.296ml H <sub>2</sub> O to a whole 1mg bottle	0.005 mM	20 µl
Epidermal growth factor	100 µg/mL	Add 5ml DDW to a whole 500µg bottle	0.05 µg/mL	20 µl
Valproic Acid (Sodium Salt)	1.156 M	Add 3ml DMSO to a whole 500mg bottle	0.0008 M	27.7 µl
A83-01	10 mM	Add 1.185ml DMSO to a whole 5mg bottle	0.0005 mM	2 µl
CHIR99021	4 mM	Add 2.69ml DMSO to a whole 5mg bottle	0.002 mM	20 µl
S8431542 hydrate	10 mM	Add 1.301ml DMSO to a whole 5mg bottle	0.001 mM	4 µl
2-mercaptoethanol	143 mM	Dilute: 10µl B-ME + 1ml PBS	0.1 mM	28 µl

(1) Stock concentration; (2) Final concentration; (3) Amount to add to make 40 mL



**Appendix 7. Selection of candidate placenta-specific mDMRs for further comprehensive methylation and expression characterisation in the placental cohort.**

**Appendix 8. Analysis summary for all informative samples identified for the *G0S2* placenta-specific mDMR**

Variant	Sample	Genotype	Mother's genotype	Father's genotype	Methylation-sensitive genotyping (HpaII)		Allelic expression	
rs1815548	BCN 5	C/T	C/T	-	Pref. C	Pref. monoallelic	-	-
	BCN 7	C/T	C/T	-	T	Monoallelic	-	-
	BCN 44	C/T	C/T	-	T	Monoallelic	-	-
	BCN 70	C/T	C/T	-	Pref. C	Pref. monoallelic	-	-
	BCN 95	C/T	T	-	T	Maternal	-	-
	22BR 162	C/T	-	-	T	Monoallelic	-	-
	21BR 311	C/T	C/T	-	C	Monoallelic	-	-
rs932375	BCN 12	C/G	C	-	C	Maternal	C	Maternal
	BCN 31	C/G	C	-	C	Maternal	G	Paternal
	23BR 128	C/G	C/G	C	G	Maternal	C	Paternal
	22BR 162	C/G	-	-	C	Monoallelic	Pref. C	Pref. monoallelic
	23BR 294	C/G	C	-	Pref. C	Pref. maternal	C	Maternal
	21BR 311	C/G	C	-	Pref. C	Pref. maternal	C	Maternal
	21BR 430	C/G	C/G	-	G	Monoallelic	Pref. C	Pref. monoallelic
21BR 432	C/G	C	-	C	Maternal	Pref. C	Maternal	

**Appendix 9. Analysis summary for all informative samples identified for *PIK3R1* isoform 3 placenta-specific mDMR (n - number of repeat copies)**

Variant	Sample	Genotype	Mother's genotype	Methylation-sensitive genotyping (HpaII)		Allelic expression	
rs138814985	BCN 5	3n/4n	3n/4n	3n	Monoallelic	-	-
	BCN 8	3n/4n	3n	3n	Maternal	3n/4n	Biallelic
	BCN 60	3n/4n	-	3n	Monoallelic	4n	Monoallelic
	BCN 65	3n/4n	3n/4n	3n	Monoallelic	N/A	-

Variant	Sample	Genotype	Mother's genotype	Methylation-sensitive genotyping (HpaII)		Allelic expression	
	BCN 70	3n/4n	3n/4n	3n	Monoallelic	3n/4n	Biallelic
	BCN 92	3n/4n	3n	3n	Maternal	3n/4n	Biallelic
	BCN 93	3n/4n	3n/4n	3n	Monoallelic	3n	Monoallelic
	BCN 95	3n/4n	3n	3n	Maternal	-	-
	22BR 160	3n/4n	3n/4n	4n	Monoallelic	-	-
	22BR 161	3n/4n	3n/4n	4n	Monoallelic	3n/4n	Biallelic
	22BR 162	3n/4n	-	3n	Monoallelic	4n	Monoallelic
	22BR 546	3n/4n	-	-	-	3n/4n	Biallelic
	22BR 700	3n/4n	-	4n	Monoallelic	4n	Monoallelic
	rs2888323	BCN 6	A/G	A	Pref. G	Pref. paternal	G/A
BCN 45		A/G	A/G	A/G	Biallelic	-	-
BCN 46		A/G	A	A	Maternal	A	Maternal
BCN 77		A/G	-	Pref. A	Pref. monoallelic	G	Monoallelic
22BR 161		A/G	A/G	A	Monoallelic	A/G	Biallelic
22BR 293		A/G	A/G	A	Monoallelic	A/G	Biallelic
22BR 546		A/G	-	-	-	A/G	Biallelic
22BR 548		A/G	-	A/G	Biallelic	A/G	Biallelic
22BR 700		A/G	-	A	Monoallelic	A	Monoallelic
rs3730089	BCN 6	G/A	G/A	-	-	G/A	Biallelic
	BCN 8	G/A	G	-	-	Pref. G	Pref. maternal
	BCN 21	G/A	G/A	-	-	Pref. G	Pref. monoallelic
	BCN 26	G/A	G/A	-	-	Pref. G	Pref. monoallelic
	BCN 44	G/A	G/A	-	-	G/A	Biallelic
	BCN 64	G/A	G	-	-	G/A	Biallelic
	BCN 95	G/A	G/A	-	-	G/A	Biallelic
	22BR 161	G/A	G/A	-	-	Pref. G	Pref. monoallelic

<b>Variant</b>	<b>Sample</b>	<b>Genotype</b>	<b>Mother's genotype</b>	<b>Methylation-sensitive genotyping (HpaII)</b>			<b>Allelic expression</b>	
	22BR 162	G/A	-	-	-	A	Monoallelic	
	22BR 701	G/A	G/A	-	-	G	Monoallelic	

**Appendix 10. Methylation levels of *PIK3R1* isoform 3 and *G0S2* placenta-specific mDMRs in the Infinium MethylationEPIC array datasets generated by Yuan and colleagues (364).**

*G0S2* region: chr1:209847279-209850881 (GRCh37 regions)

GEO accession: GSE159526

Cell types:	PM365					PM369				
	Endothelial	Hofbauer	Stromal	Trophoblast	Villi	Endothelial	Hofbauer	Stromal	Trophoblast	Villi
	GSM48317 78.118748	GSM4831947.1 18951	GSM483182 7.118797	GSM4831804. 118774	GSM483180 3.118773	GSM4831800. 118770	GSM4831850. 118820	GSM4831844. 118814	GSM4831953. 118957	GSM4831848. 118818
Probes:										
cg23646375	0.848	0.791	0.853	0.679	0.671	0.769	0.804	0.879	0.636	0.658
cg09886578	0.432	0.407	0.472	0.260	0.302	0.381	0.412	0.526	0.231	0.290
cg06616057	0.599	0.463	0.564	0.519	0.532	0.461	0.426	0.507	0.390	0.485
cg09666230	0.519	0.374	0.547	0.435	0.510	0.547	0.326	0.532	0.489	0.454
cg24933191	0.516	0.246	0.515	0.502	0.555	0.542	0.257	0.527	0.502	0.525
cg14824901	0.509	0.260	0.494	0.554	0.527	0.553	0.142	0.504	0.523	0.531
cg27176828	0.391	0.131	0.377	0.409	0.439	0.453	0.080	0.434	0.422	0.407
cg17710021	0.489	0.205	0.503	0.399	0.440	0.530	0.173	0.496	0.445	0.421
cg10671306	0.471	0.227	0.487	0.468	0.489	0.510	0.187	0.487	0.456	0.461
cg13460643	0.585	0.401	0.576	0.670	0.632	0.606	0.318	0.582	0.620	0.628
cg08185241	0.547	0.380	0.523	0.569	0.563	0.549	0.307	0.534	0.545	0.544
cg26050864	0.569	0.327	0.603	0.453	0.499	0.571	0.302	0.583	0.512	0.482
Average:	0.539	0.351	0.543	0.493	0.513	0.539	0.311	0.549	0.481	0.491

DMR/ promoter  
CpGs in bisulphite PCR

PIK3R1 region: chr5:67493244 - 67601815 (GRCh37 regions)

GEO accession: GSE159526

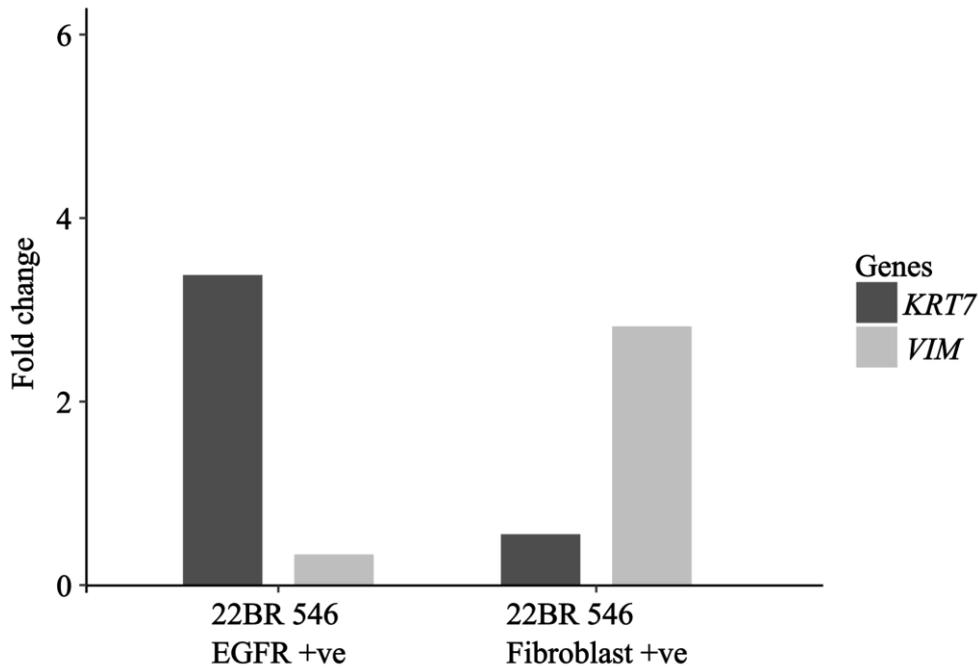
Cell types:	PM365					PM369				
	Endothelial	Hofbauer	Stromal	Trophoblast	Villi	Endothelial	Hofbauer	Stromal	Trophoblast	Villi
	GSM4831778. 118748	GSM4831947. 118951	GSM4831827. 118797	GSM4831804. 118774	GSM4831803. 118773	GSM4831800. 118770	GSM4831850. 118820	GSM4831844. 118814	GSM4831953. 118957	GSM4831848. 118818
Probes:										
cg03796175	0.926	0.932	0.929	0.944	0.917	0.930	0.934	0.940	0.944	0.945
cg05092819	0.027	0.031	0.026	0.035	0.034	0.030	0.040	0.025	0.028	0.028
cg19425035	0.027	0.033	0.043	0.036	0.042	0.030	0.035	0.032	0.036	0.024
cg18344938	0.054	0.065	0.053	0.037	0.058	0.072	0.058	0.061	0.056	0.062
cg19358016	0.046	0.055	0.061	0.053	0.054	0.071	0.040	0.086	0.081	0.046
cg14514263	0.043	0.043	0.037	0.022	0.045	0.038	0.044	0.042	0.041	0.027
cg02024097	0.042	0.044	0.042	0.024	0.045	0.045	0.045	0.042	0.047	0.028
cg25664275	0.033	0.034	0.035	0.020	0.036	0.032	0.037	0.033	0.034	0.023
cg03331123	0.102	0.106	0.120	0.069	0.146	0.165	0.139	0.117	0.116	0.092
cg23036683	0.008	0.011	0.012	0.008	0.010	0.015	0.014	0.011	0.016	0.012
Average:	0.131	0.135	0.136	0.125	0.139	0.143	0.139	0.139	0.140	0.129
cg01239651	0.088	0.289	0.265	0.668	0.647	0.154	0.153	0.230	0.875	0.673
cg24797508	0.227	0.263	0.342	0.781	0.743	0.295	0.159	0.368	0.842	0.815
cg09101894	0.363	0.406	0.470	0.897	0.857	0.495	0.306	0.503	0.922	0.892
cg11342429	0.035	0.112	0.060	0.388	0.423	0.040	0.033	0.041	0.532	0.408
cg01893041	0.033	0.118	0.074	0.337	0.376	0.070	0.041	0.067	0.521	0.360
cg03239914	0.023	0.104	0.072	0.314	0.373	0.045	0.028	0.056	0.446	0.340
cg05608159	0.035	0.100	0.061	0.246	0.330	0.058	0.049	0.066	0.409	0.299
cg25195415	0.024	0.090	0.054	0.298	0.381	0.046	0.037	0.046	0.492	0.380
cg08945395	0.021	0.085	0.039	0.280	0.345	0.042	0.020	0.046	0.414	0.352
cg20439288	0.016	0.058	0.029	0.344	0.453	0.026	0.015	0.042	0.568	0.530
cg01105385	0.054	0.123	0.070	0.553	0.547	0.075	0.040	0.111	0.619	0.565

PIK3R1 region: chr5:67493244 - 67601815 (GRCh37 regions)

GEO accession: GSE159526

PM365						PM369				
Cell types:	Endothelial	Hofbauer	Stromal	Trophoblast	Villi	Endothelial	Hofbauer	Stromal	Trophoblast	Villi
	GSM4831778. 118748	GSM4831947. 118951	GSM4831827. 118797	GSM4831804. 118774	GSM4831803. 118773	GSM4831800. 118770	GSM4831850. 118820	GSM4831844. 118814	GSM4831953. 118957	GSM4831848. 118818
Probes:										
cg20474370	0.035	0.126	0.069	0.340	0.384	0.085	0.049	0.102	0.485	0.372
cg16333716	0.814	0.947	0.851	0.935	0.930	0.928	0.946	0.907	0.949	0.951
cg02271687	0.872	0.898	0.876	0.823	0.869	0.908	0.893	0.875	0.917	0.904
cg16664523	0.242	0.853	0.101	0.802	0.729	0.538	0.817	0.099	0.893	0.802
cg07872489	0.502	0.926	0.202	0.885	0.825	0.792	0.908	0.095	0.928	0.879
cg25091228	0.874	0.898	0.818	0.858	0.857	0.892	0.900	0.812	0.877	0.878
cg10887670	0.790	0.865	0.700	0.797	0.763	0.869	0.838	0.634	0.836	0.832
cg06445944	0.803	0.903	0.816	0.909	0.871	0.802	0.892	0.791	0.852	0.858
Average:	0.308	0.430	0.314	0.603	0.616	0.377	0.375	0.310	0.704	0.636

	DMR/ promoter
	CpGs in bisulphite PCR
	Promoter



**Appendix 11. Expression of placental cell marker genes in placental cell fractions obtained by MACS.**

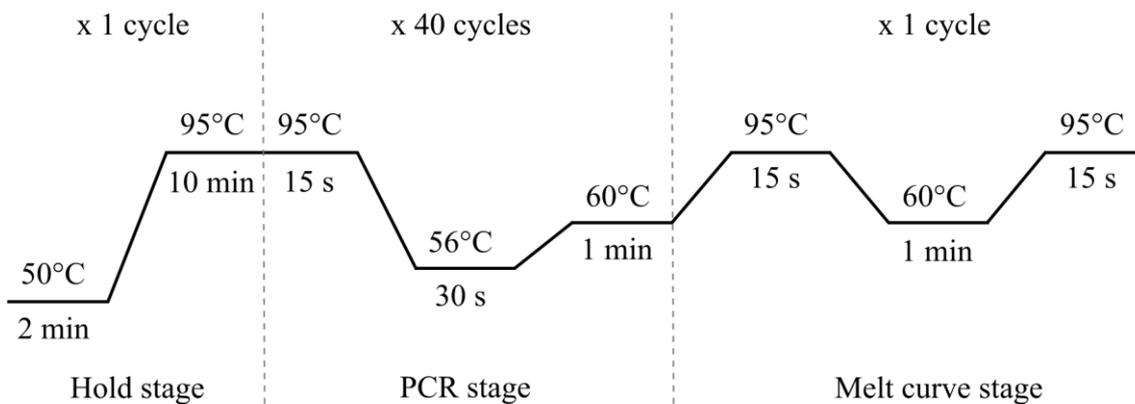
EGFR-positive cells represent placental trophoblasts, which show high expression of *KRT7* (a trophoblast marker). Fibroblast-positive cells represent placental stromal cells, which exhibit high expression of *VIM* (a stromal mesenchymal marker). Expression levels were normalized to the endogenous controls *RPL19* and *ACTB*.

**Appendix 12. Methylation levels of *PIK3R1* isoform 3 and *G0S2* placenta-specific mDMRs in the Illumina Infinium Human Methylation450 BeadChip array datasets generated by our group (295).**

<i>G0S2</i> DMR	GEO accession: GSE120981			
PL_ID	placenta 5	placenta 9	IUGR placenta 55	IUGR placenta 109
ID_REF	GSM3423383	GSM3423385	GSM3423392	GSM3423394
<b>Probes:</b>				
cg09666230	0.289	0.493	0.368	0.665
cg14824901	0.537	0.319	0.445	0.649
cg27176828	0.645	0.495	0.556	0.776
cg17710021	0.439	0.450	0.499	0.514
cg13460643	0.688	0.548	0.624	0.799
cg08185241	0.578	0.474	0.573	0.685
cg08158408	0.620	0.508	0.608	0.836
cg19534438	0.228	0.151	0.320	0.349
Average:	0.503	0.430	0.499	0.659

<i>PIK3R1</i> DMR	GEO accession: GSE120981			
PL_ID	placenta 5	placenta 9	IUGR placenta 55	IUGR placenta 109
ID_REF	GSM3423383	GSM3423385	GSM3423392	GSM3423394
<b>Probes:</b>				

PL_ID	placenta 5	placenta 9	IUGR placenta 55	IUGR placenta 109
ID_REF	GSM3423383	GSM3423385	GSM3423392	GSM3423394
<b>Probes:</b>				
cg01239651	0.759	0.767	0.693	0.751
cg24797508	0.833	0.712	0.641	0.709
cg09101894	0.880	0.823	0.770	0.821
cg01893041	0.447	0.209	0.401	0.407
cg05608159	0.373	0.069	0.322	0.367
cg25195415	0.374	0.064	0.349	0.402
cg08945395	0.379	0.031	0.357	0.385
cg20439288	0.443	0.019	0.365	0.356
cg01105385	0.439	0.077	0.437	0.477
cg20474370	0.462	0.080	0.447	0.493
cg07208333	0.583	0.059	0.554	0.539
cg02271687	0.948	0.821	0.883	0.907
Average:	0.577	0.311	0.518	0.551



**Appendix 13. qRT-PCR cycling conditions for *PIK3R1* with the Power SYBR™ Green PCR Master Mix.**

**Appendix 14. Methylation-sensitive genotyping results reported by Sanchez-Delgado *et al.*, 2016 (20).**

(A) Observations recorded for imprinted placenta-specific mDMRs. (B) Observations recorded for non-imprinted genomic intervals with oocyte-derived methylation maintained in term placenta. (C) Chi-squared test comparing methylation-sensitive genotyping results for imprinted placenta-specific mDMRs versus non-imprinted genomic regions.

(A)

gDMR	Biallelic methylation		Monoallelic methylation		Maternal methylation		Row sum
<i>SPHKAP</i>	5	50%	1	10%	4	40%	10
<i>EFCCI</i>	0	0%	3	100%	0	0%	3

gDMR	Biallelic methylation		Monoallelic methylation		Maternal methylation		Row sum
<i>FGF12</i>	3	50%	1	17%	2	33%	6
<i>PDE6B</i>	0	0%	1	25%	3	75%	4
<i>STX18-AS1</i>	0	0%	1	25%	3	75%	4
<i>SFRP2</i>	0	0%	4	57%	3	43%	7
<i>R3HCC1</i>	3	38%	1	13%	4	50%	8
<i>OPCML</i>	0	0%	1	50%	1	50%	2
<i>CACNA1C</i>	0	0%	3	60%	2	40%	5
<i>PAPLN-AS</i>	0	0%	3	60%	2	40%	5
<i>GRP78</i>	1	25%	2	50%	1	25%	4
<i>GRID2</i>	0	0%	1	50%	1	50%	2
<i>BOD1L2</i>	0	0%	4	100%	0	0%	4
<i>TPTEP1</i>	0	0%	3	38%	5	63%	8
<i>FRMD3</i>	3	43%	2	29%	2	29%	7
chr18 region	0	0%	4	57%	3	43%	7
<i>CACNA1L</i>	0	0%	4	67%	2	33%	6
<i>CACNA1E</i>	0	0%	0	0%	1	100%	1
<i>ZNF385D</i>	0	0%	2	50%	2	50%	4
<i>C3ORF62</i>	1	25%	1	25%	2	50%	4
<i>SH3BP2</i>	0	0%	2	100%	0	0%	2
<i>RYR3</i>	0	0%	1	50%	1	50%	2
<i>CACNA1A</i>	2	29%	1	14%	4	57%	7
<b>Column sum</b>	18		46		48		112
Biallelic methylation							
Mean:	11.304%						
SD:	18.396%						
Mean ± 2SD:	48.1%						

(B)

gDMR	Biallelic methylation		Monoallelic methylation		Maternal methylation		Row sum
<i>TMEM247</i>	0	0%	0	0%	2	100%	2
<i>DPP6</i>	11	92%	1	8%	0	0%	12
Chr. 1	16	100%	0	0%	0	0%	16
<i>THSD7B</i>	10	100%	0	0%	0	0%	10
<i>SLC2A2</i>	10	100%	0	0%	0	0%	10
<i>RPS6KAL</i>	11	100%	0	0%	0	0%	11
<i>OPRM1</i>	7	100%	0	0%	0	0%	7
<i>RADIL</i>	8	100%	0	0%	0	0%	8
<i>NTNG2</i>	7	100%	0	0%	0	0%	7
Chr. 10	3	100%	0	0%	0	0%	3
<i>UNC79</i>	12	100%	0	0%	0	0%	12
<i>OCA2</i>	5	100%	0	0%	0	0%	5
<i>FHOD3</i>	1	100%	0	0%	0	0%	1
<b>Column sum</b>	101		1		2		104

(C)

**Observed values**

<b>Imprinted gDMR</b>	<b>Biallelic methylation</b>	<b>Monoallelic methylation</b>	<b>Maternal methylation</b>	<b>Row sum</b>
Yes	18	46	48	112
No	101	1	2	104
<b>Column sum</b>	119	47	50	216

**Expected values**

<b>Imprinted gDMR</b>	<b>Biallelic methylation</b>	<b>Monoallelic methylation</b>	<b>Maternal methylation</b>	<b>Row sum</b>
Yes	62	24	26	112
No	57	23	24	104
<b>Column sum</b>	119	47	50	216

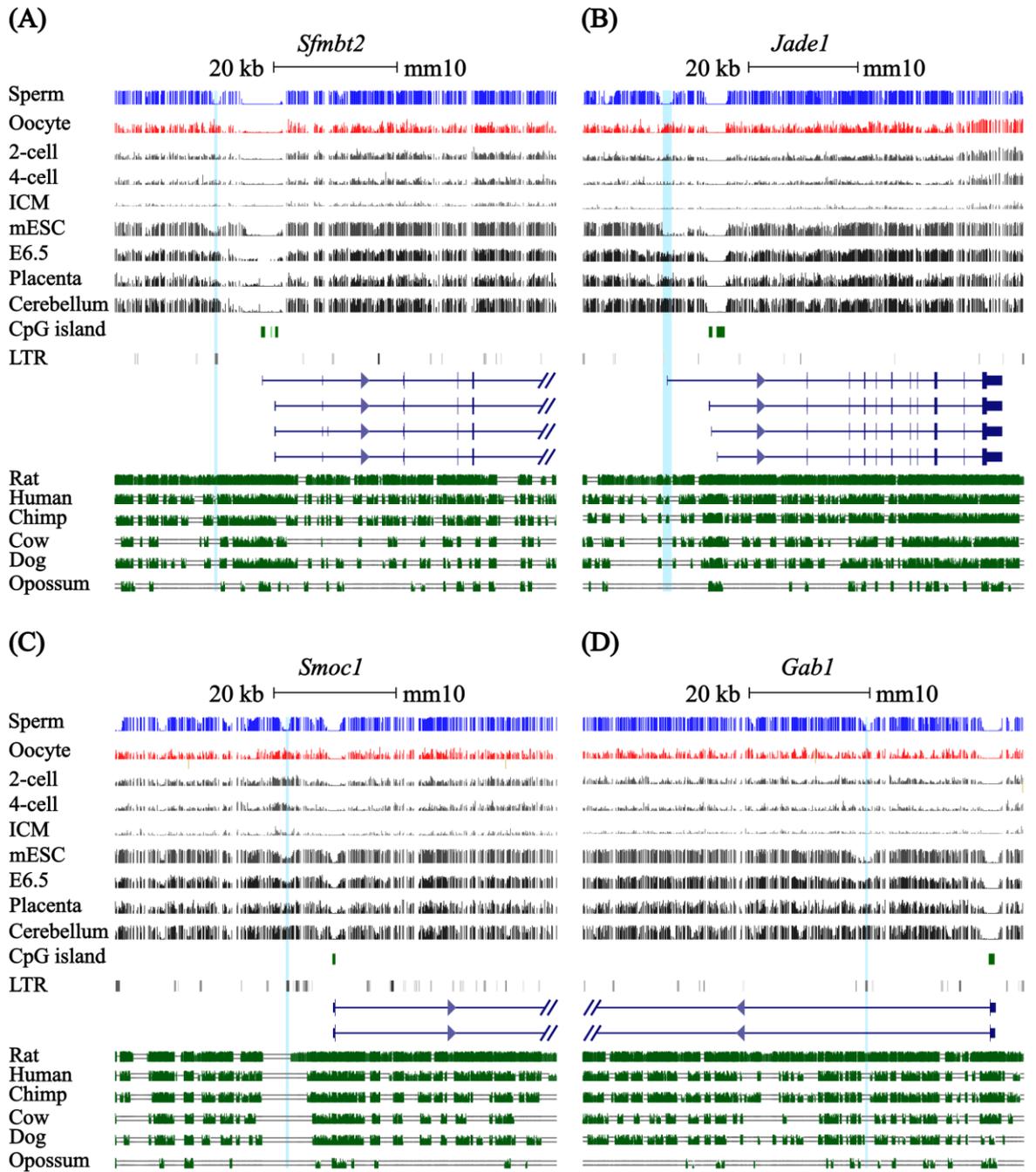
**Chi-Squared test**

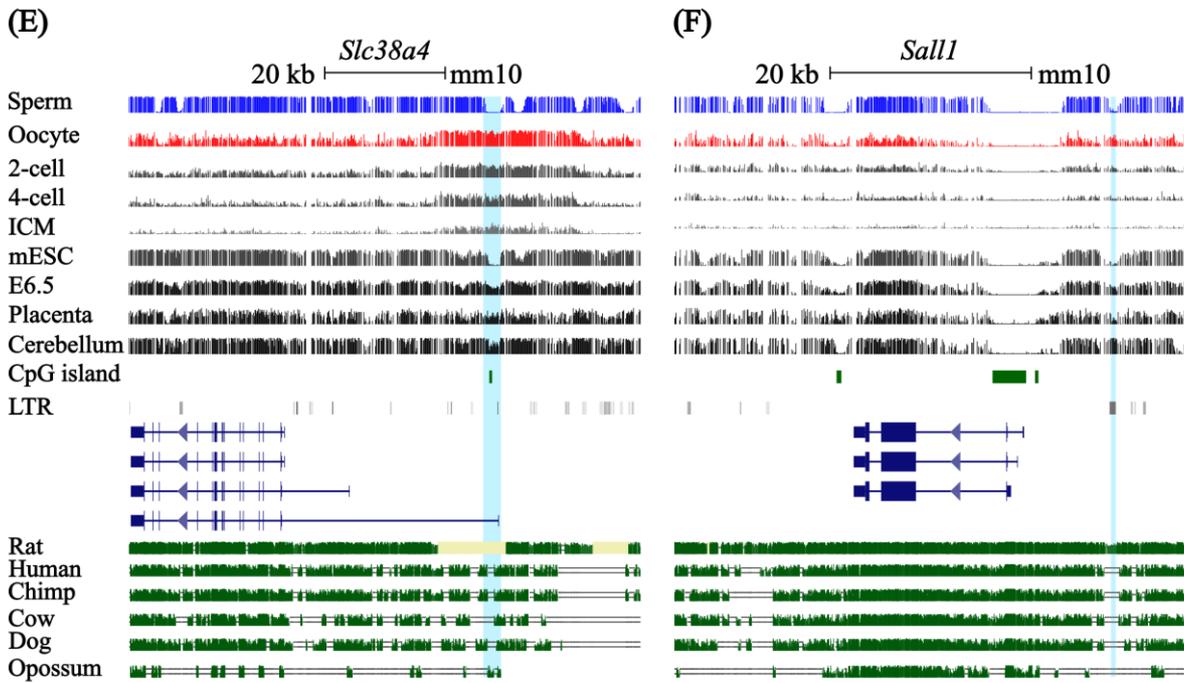
	<b>Observed</b>	<b>Expected</b>	<b>(Obs. - Exp.)<sup>2</sup>/Exp.</b>
Yes   biallelic	18	62	31
Yes   monoallelic	46	24	19
Yes   maternal	48	26	19
No   biallelic	101	57	33
No   monoallelic	1	23	21
No   maternal	2	24	20

X<sup>2</sup>: 143

df.:2

p < 2.2 × 10<sup>-16</sup>





**Appendix 15. DNA methylation profiles of mouse non-canonical imprinted genes across diverse mouse methyl-seq datasets.**

Genomic maps display mouse placental sDMRs, with some containing ERVK LTRs (highlighted in light blue), for the following genes: (A) *Sfmbt2*, (B) *Jade1*, (C) *Smoc1*, (D) *Gab1*, and (F) *Sall1*, while (E) *Slc38a4* contains an oocyte-derived gDMR. Gene transcripts are shown in dark blue, with thicker bars representing exons, while CpG islands are shown as dark green bars and ERV LTRs as grey bars (UCSC RepeatMasker track). Vertical lines in the methyl-seq tracks represent the mean methylation levels for individual CpG dinucleotides. Dark green tracks (from the UCSC Vertebrate Multiz Alignment & Conservation) indicate conserved genomic regions among mouse, rat, human, chimp, cow, dog, and opossum.

**Appendix 16. Detailed analysis results for the human orthologues of mouse and rat non-canonical imprinted genes**

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Methylation - cloning
3	<i>RPL39L</i>	1	-	-	-	-	BCN 8 - not informative, low, mosaic methylation
3	<i>RPL39L</i>		rs141173382	1	-	-	BCN 8 -informative, both alleles mostly methylated
4	<i>JADE1/ PHF17</i>		rs62317870	-	-	-	BCN 26 & BCN 17 - not informative, low methylation
4	<i>JADE1/ PHF17</i> (NM_001287441; NM_024900)	4	rs13114904	4	-	4 - biallelic	-
4	<i>JADE1/ PHF17</i> (NM_001287437)		rs11933240	4	-	4 - biallelic	-
4	<i>GABI</i> (NM_002039)	4	-	-	-	-	BCN 6 - not informative, unmethylated
4	<i>GABI</i>		rs1397529	4	-	4 - biallelic	-
4	<i>GABI</i> (AK295684)		rs62337524	1	-	1 - biallelic	21BR 309 - informative, mosaic methylation
10	<i>SFMBT2</i>	5	-	-	-	-	BCN 8 - not informative, unmethylated
			rs719809	1	-	-	21BR 307 - informative, biallelic methylation

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Methylation - cloning
			rs10795530	4	-	1 - pref. monoallelic, 3 - biallelic	-
12	<i>SLC38A1</i>	6	-	-	-	-	BCN 8 - not informative, mosaic methylation
			rs1045278	6	-	6 - biallelic	-
			-	-	-	-	BCN8 - not informative, mostly unmethylated
			rs4994910	10	6 - pref. monoallelic, 4 - biallelic	-	-
12	<i>SLC38A4</i>	14	rs74851348	3	1 - monoallelic, 1 - pref. monoallelic, 1 - biallelic	-	BCN 8 - not informative, mostly methylated; 21BR 309 - not informative mosaic methylation; 21BR 19 - informative, mosaic methylation, mostly methylated
			rs2429467	3	-	3 - biallelic	-
14	<i>SMOC1</i>	4	rs146095118	-	-	-	BCN 26 - not informative, low methylation
			rs3742909	4	-	4 - biallelic	-

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Methylation - cloning
16	<i>SALL1</i>	4	rs11643654	1	-	-	21BR 307 - not informative, mosaic methylation; BCN 5 - informative, mosaic methylation
			rs11645288	4	-	4 - biallelic	-
18	<i>ZNF516</i>	1	-	-	-	-	BCN 8 - not informative, mosaic methylation
			rs690353	1	-	1 - biallelic	-
20	<i>ZFP64</i>	2	-	-	-	-	BCN 8 - not informative, mosaic methylation
			rs3746413	2	-	2 - biallelic	-
X	<i>XIST</i>	2	-	-	-	-	Male sample fully methylated, female sample showed 50% methylation
			rs1894271	2	-	2 - biallelic	-

**Appendix 17. Summary of scRNA-seq aligned reads from human pre-implantation embryos.**

Green indicates good samples with high alignment, yellow indicates samples with intermediate alignment, and red indicates samples with poor alignment.

Raw reads				Aligned reads after trimming with TrimGalore					
Original name	Duplicated reads (%)	GC content (%)	Millions total sequences	Duplicated reads (%)	GC content (%)	Median Read Length	Aligned reads (%)	Millions uniquely mapped reads	Millions total sequences
1_12a.1	67.10%	48%	11.9	63.90%	47%	147 bp	85.00%	10.1	11.8
1_12a.10	10.50%	48%	0.1	9.50%	47%	147 bp	83.00%	0	0.1
1_12a.11	11.90%	48%	0.1	10.60%	47%	147 bp	82.70%	0	0.1
1_12a.12	69.30%	49%	15.4	64.50%	48%	147 bp	85.70%	13.2	15.4
1_12a.2	66.00%	49%	12.9	63.10%	48%	147 bp	87.60%	11.3	12.9
1_12a.3	65.40%	48%	12.2	62.10%	47%	147 bp	85.60%	10.4	12.2
1_12a.4	60.30%	48%	9.5	57.50%	47%	147 bp	86.80%	8.2	9.5
1_12a.5	46.30%	48%	4.8	44.80%	48%	147 bp	87.40%	4.2	4.8
1_12a.6_1	59.70%	48%	8.7	57.10%	47%	147 bp	84.20%	7.3	8.7
1_12a.6_2	56.50%	48%	8.3	54.50%	47%	147 bp	86.80%	7.2	8.3
1_12a.7	47.90%	47%	1.2	44.20%	44%	147 bp	69.30%	0.8	1.2
1_12a.8	62.60%	48%	4.3	58.40%	47%	147 bp	74.00%	3.2	4.3
1_12a.9	63.60%	48%	10.2	61.60%	47%	147 bp	86.00%	8.7	10.2
2_10a.1	57.00%	48%	13.8	54.20%	47%	147 bp	84.60%	11.6	13.7
2_10a.10	60.70%	48%	10.9	57.90%	47%	147 bp	85.80%	9.3	10.9
2_10a.2	87.80%	46%	9.4	85.60%	44%	142 bp	82.40%	7.8	9.4
2_10a.3	85.00%	46%	6.9	82.20%	44%	147 bp	79.90%	5.5	6.9
2_10a.4	86.00%	46%	6.2	83.60%	44%	147 bp	83.20%	5.1	6.2
2_10a.5	58.00%	49%	10.9	55.30%	48%	147 bp	83.90%	9.1	10.9
2_10a.6	87.10%	46%	6.8	84.30%	44%	142 bp	80.90%	5.5	6.8
2_10a.7	55.30%	47%	8.8	52.60%	47%	147 bp	83.30%	7.3	8.8
2_10a.8	86.50%	47%	4.5	82.50%	44%	142 bp	76.80%	3.4	4.4
2_10a.9	62.00%	48%	12.4	59.40%	47%	147 bp	84.80%	10.5	12.3
3_5a.1	62.70%	47%	7.8	61.10%	46%	147 bp	82.50%	6.4	7.8

Raw reads				Aligned reads after trimming with TrimGalore					
Original name	Duplicated reads (%)	GC content (%)	Millions total sequences	Duplicated reads (%)	GC content (%)	Median Read Length	Aligned reads (%)	Millions uniquely mapped reads	Millions total sequences
3_5a.2	59.20%	47%	9.2	57.00%	47%	147 bp	84.80%	7.8	9.2
3_5a.3	76.60%	48%	10.3	74.20%	47%	147 bp	72.20%	7.5	10.3
3_5a.4	79.20%	47%	5	75.90%	45%	147 bp	74.90%	3.7	5
3_5a.5	83.80%	46%	4.4	80.90%	44%	147 bp	83.20%	3.7	4.4
3_7a.1	58.00%	48%	10	55.90%	47%	147 bp	84.30%	8.4	9.9
3_7a.2	72.10%	47%	12.5	70.00%	46%	147 bp	81.30%	10.1	12.5
3_7a.3	56.60%	48%	9.6	55.00%	47%	147 bp	86.30%	8.3	9.6
3_7a.4	72.40%	48%	8.3	70.00%	47%	147 bp	72.30%	6	8.3
3_7a.5	59.70%	52%	0.9	52.20%	50%	127 bp	5.00%	0	0.9
3_7a.6	72.80%	48%	9.9	70.00%	47%	147 bp	74.70%	7.4	9.9
3_7a.7	55.60%	47%	6.3	53.90%	46%	147 bp	84.20%	5.3	6.3
3_Bla.1	57.90%	47%	6.8	56.00%	46%	147 bp	82.30%	5.6	6.8
3_Bla.2	64.40%	49%	7.5	61.80%	48%	147 bp	74.30%	5.6	7.5
3_Bla.3	51.70%	48%	9.9	50.40%	47%	147 bp	84.50%	8.3	9.9
3_Bla.4	56.10%	47%	5.5	53.80%	46%	147 bp	81.90%	4.5	5.5
3_Bla.5	60.60%	47%	10.1	58.20%	47%	147 bp	80.40%	8.1	10.1
3_Bla.6	56.10%	47%	11.4	53.90%	46%	147 bp	84.50%	9.6	11.4
3_Bla.7	62.50%	53%	0.9	54.40%	51%	122 bp	23.40%	0.2	0.9
3_Blb.1	72.10%	48%	7.8	69.40%	47%	147 bp	75.70%	5.9	7.7
3_Blb.2	52.60%	48%	8.8	50.90%	47%	147 bp	85.20%	7.5	8.8
3_Blb.3	54.70%	48%	6.7	53.30%	48%	147 bp	88.00%	5.9	6.7
3_Blb.4	58.00%	48%	9.4	56.10%	47%	147 bp	86.50%	8.1	9.3
3_Blb.5	64.10%	50%	8.7	62.70%	49%	147 bp	76.00%	6.6	8.7
3_Blb.6	52.60%	46%	3.3	51.40%	45%	150 bp	82.60%	2.8	3.3
3_Blb.7	65.50%	46%	5.8	63.10%	45%	147 bp	75.40%	4.4	5.8
3_Blb.8	61.10%	52%	0.6	52.90%	47%	107 bp	9.90%	0.1	0.6
3_TEa.1	62.60%	46%	13.4	60.50%	45%	147 bp	86.00%	11.5	13.4

Raw reads				Aligned reads after trimming with TrimGalore					
Original name	Duplicated reads (%)	GC content (%)	Millions total sequences	Duplicated reads (%)	GC content (%)	Median Read Length	Aligned reads (%)	Millions uniquely mapped reads	Millions total sequences
3_TeA.10	61.00%	48%	8.9	59.30%	47%	147 bp	86.30%	7.7	8.9
3_TeA.11	61.30%	49%	10.3	58.30%	48%	147 bp	86.00%	8.8	10.3
3_TeA.12	79.30%	48%	10.1	76.70%	47%	147 bp	78.70%	7.9	10.1
3_TeA.13	69.10%	49%	14.8	66.60%	48%	147 bp	79.20%	11.7	14.7
3_TeA.14	87.10%	46%	8.1	84.70%	44%	147 bp	85.50%	6.9	8
3_TeA.15	79.00%	46%	7	77.10%	45%	147 bp	80.30%	5.7	7
3_TeA.16	88.10%	45%	6	85.90%	44%	147 bp	88.50%	5.3	6
3_TeA.17	60.50%	49%	10.5	58.50%	48%	147 bp	85.10%	8.9	10.5
3_TeA.18	63.60%	52%	1.2	55.70%	49%	122 bp	28.10%	0.3	1.2
3_TeA.19	59.10%	48%	8.4	57.40%	47%	147 bp	85.50%	7.2	8.4
3_TeA.2	63.30%	49%	11.3	61.00%	48%	147 bp	85.40%	9.6	11.2
3_TeA.20	87.70%	46%	6.3	85.70%	44%	147 bp	84.70%	5.3	6.3
3_TeA.21	60.40%	49%	11.6	58.70%	49%	147 bp	76.90%	8.9	11.6
3_TeA.22	57.40%	48%	6.3	55.80%	47%	147 bp	86.60%	5.4	6.3
3_TeA.23	64.00%	49%	11	62.20%	49%	147 bp	87.70%	9.6	11
3_TeA.24	71.40%	49%	4.2	67.30%	47%	147 bp	64.70%	2.7	4.2
3_TeA.25	63.70%	48%	13.5	61.50%	48%	147 bp	84.50%	11.4	13.5
3_TeA.26	61.40%	48%	11.6	58.90%	47%	147 bp	83.60%	9.7	11.6
3_TeA.27	74.30%	47%	6.3	71.30%	46%	147 bp	76.40%	4.8	6.3
3_TeA.28	54.60%	48%	8.8	51.90%	47%	147 bp	82.60%	7.3	8.8
3_TeA.29	61.00%	49%	10.3	58.80%	48%	147 bp	83.40%	8.5	10.2
3_TeA.3	54.20%	48%	8.2	52.80%	47%	147 bp	87.10%	7.1	8.2
3_TeA.30	56.40%	48%	8.9	54.60%	47%	147 bp	84.90%	7.5	8.8
3_TeA.31	60.50%	48%	6.3	58.70%	47%	147 bp	84.70%	5.3	6.2
3_TeA.32	83.20%	47%	9.5	81.20%	46%	147 bp	79.20%	7.5	9.5
3_TeA.33	61.00%	47%	7.8	59.40%	46%	147 bp	81.80%	6.4	7.8
3_TeA.34	64.60%	49%	11.4	62.50%	49%	147 bp	79.20%	9	11.4

Raw reads				Aligned reads after trimming with TrimGalore					
Original name	Duplicated reads (%)	GC content (%)	Millions total sequences	Duplicated reads (%)	GC content (%)	Median Read Length	Aligned reads (%)	Millions uniquely mapped reads	Millions total sequences
3_TeA.35	59.60%	49%	11.4	57.70%	48%	147 bp	84.80%	9.6	11.3
3_TeA.36	75.90%	47%	4.5	72.40%	45%	147 bp	75.30%	3.4	4.5
3_TeA.37	61.90%	48%	9.9	59.70%	48%	147 bp	85.70%	8.4	9.9
3_TeA.38	58.20%	48%	7.6	56.70%	48%	147 bp	87.30%	6.6	7.5
3_TeA.39	57.20%	49%	6.7	55.20%	48%	147 bp	84.30%	5.7	6.7
3_TeA.4	55.40%	49%	9.3	53.20%	48%	147 bp	86.60%	8	9.2
3_TeA.40	77.00%	46%	2.6	74.30%	45%	147 bp	81.90%	2.1	2.6
3_TeA.41	55.60%	48%	6.6	53.90%	48%	147 bp	82.70%	5.4	6.6
3_TeA.5	58.70%	49%	10.4	56.90%	48%	147 bp	86.20%	8.9	10.4
3_TeA.6	67.70%	53%	1.9	62.20%	50%	142 bp	31.60%	0.6	1.9
3_TeA.7	78.90%	53%	4.2	76.70%	52%	147 bp	17.60%	0.7	4.2
3_TeA.8	76.10%	54%	2.1	71.10%	52%	142 bp	9.90%	0.2	2.1
3_TeA.9	75.40%	48%	5.1	72.20%	46%	147 bp	71.80%	3.7	5.1
4_BLa.1	66.50%	44%	9.8	61.70%	41%	137 bp	72.40%	7.1	9.8
4_BLa.10	64.40%	48%	1.7	58.40%	45%	142 bp	57.20%	1	1.7
4_BLa.11	78.10%	46%	5.9	75.10%	44%	147 bp	81.50%	4.8	5.9
4_BLa.2	74.70%	49%	4.2	69.20%	46%	122 bp	57.40%	2.4	4.2
4_BLa.3	77.80%	50%	1.5	71.00%	47%	117 bp	38.00%	0.6	1.5
4_BLa.4	58.30%	51%	0.8	49.50%	47%	102 bp	34.80%	0.3	0.8
4_BLa.5	63.80%	48%	4.3	58.50%	47%	147 bp	77.60%	3.4	4.3
4_BLa.6	61.90%	48%	4.3	57.50%	46%	147 bp	75.10%	3.2	4.3
4_BLa.7	51.50%	47%	1.4	46.00%	44%	127 bp	65.30%	0.9	1.4
4_BLa.8	70.00%	51%	0.7	60.60%	45%	92 bp	22.20%	0.2	0.7
4_BLa.9	75.00%	47%	3.3	70.20%	43%	127 bp	71.80%	2.4	3.3
4_TeA.1	78.70%	49%	5.8	72.70%	45%	107 bp	48.60%	2.8	5.8
4_TeA.10	68.30%	48%	12.9	65.00%	47%	147 bp	83.40%	10.7	12.8
4_TeA.11	60.30%	48%	10	57.90%	48%	147 bp	86.80%	8.7	10

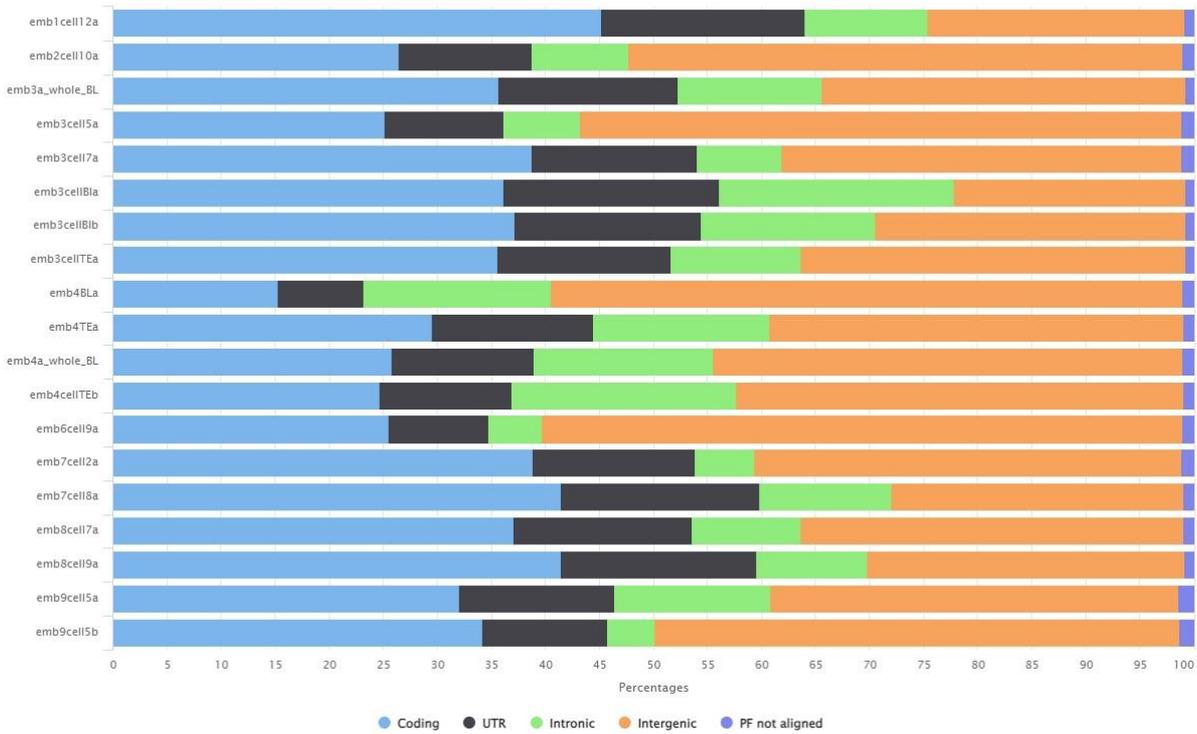
Raw reads				Aligned reads after trimming with TrimGalore					
Original name	Duplicated reads (%)	GC content (%)	Millions total sequences	Duplicated reads (%)	GC content (%)	Median Read Length	Aligned reads (%)	Millions uniquely mapped reads	Millions total sequences
4_TeA.12	53.10%	48%	5.6	51.30%	47%	147 bp	87.00%	4.9	5.6
4_TeA.13	80.30%	49%	10.2	75.80%	47%	137 bp	74.00%	7.5	10.1
4_TeA.14	60.90%	47%	14.8	57.90%	46%	147 bp	87.00%	12.8	14.8
4_TeA.15	76.70%	49%	8.5	72.50%	48%	142 bp	66.90%	5.7	8.4
4_TeA.2	70.40%	50%	3.1	61.70%	46%	107 bp	50.00%	1.6	3.1
4_TeA.3	72.90%	51%	1.6	61.70%	43%	67 bp	14.20%	0.2	1.6
4_TeA.4	69.00%	51%	1.1	59.00%	46%	82 bp	21.00%	0.2	1.1
4_TeA.5	71.70%	51%	1.1	63.10%	47%	97 bp	23.90%	0.3	1.1
4_TeA.6	67.20%	47%	8.6	64.70%	45%	147 bp	83.10%	7.1	8.5
4_TeA.7	64.40%	51%	1	55.30%	46%	112 bp	43.80%	0.5	1
4_TeA.8	51.20%	48%	7.4	48.80%	47%	147 bp	85.40%	6.3	7.4
4_TeA.9	83.50%	47%	4.8	80.80%	45%	147 bp	73.50%	3.5	4.8
4_TeB.1	73.30%	50%	5	67.70%	48%	147 bp	56.40%	2.8	5
4_TeB.10	64.40%	49%	13.4	62.20%	48%	147 bp	84.90%	11.4	13.4
4_TeB.11	44.60%	45%	3.9	43.10%	43%	147 bp	84.20%	3.2	3.8
4_TeB.12	76.60%	50%	2.1	69.70%	47%	127 bp	31.40%	0.7	2.1
4_TeB.2	57.10%	48%	14.3	54.90%	47%	147 bp	89.90%	12.8	14.2
4_TeB.22	56.00%	51%	0.7	50.70%	48%	147 bp	26.00%	0.2	0.7
4_TeB.23	68.70%	51%	1.5	62.00%	48%	132 bp	10.90%	0.2	1.5
4_TeB.24	66.60%	51%	1.8	58.30%	47%	112 bp	20.40%	0.4	1.8
4_TeB.25	69.20%	52%	2	61.70%	49%	122 bp	20.00%	0.4	2
4_TeB.26	72.00%	49%	2.1	66.80%	47%	147 bp	37.10%	0.8	2.1
4_TeB.27	63.20%	49%	1.4	56.50%	46%	122 bp	43.50%	0.6	1.4
4_TeB.28	68.00%	47%	2.2	65.00%	46%	147 bp	67.70%	1.5	2.2
4_TeB.29	63.70%	49%	7	61.70%	48%	147 bp	66.30%	4.6	6.9
4_TeB.3	76.80%	48%	9.4	74.40%	47%	147 bp	74.00%	6.9	9.4
4_TeB.30	66.80%	49%	1.8	61.70%	46%	147 bp	48.20%	0.9	1.8

Raw reads				Aligned reads after trimming with TrimGalore					
Original name	Duplicated reads (%)	GC content (%)	Millions total sequences	Duplicated reads (%)	GC content (%)	Median Read Length	Aligned reads (%)	Millions uniquely mapped reads	Millions total sequences
4_TEb.31	62.00%	51%	0.8	54.40%	48%	122 bp	20.70%	0.2	0.8
4_TEb.32	54.30%	48%	8.6	52.40%	47%	147 bp	87.60%	7.5	8.6
4_TEb.33	57.60%	46%	4.8	53.40%	44%	147 bp	78.30%	3.8	4.8
4_TEb.34	70.40%	48%	1.3	66.70%	46%	147 bp	52.40%	0.7	1.3
4_TEb.35	67.80%	50%	0.9	58.70%	45%	97 bp	15.50%	0.1	0.9
4_TEb.4	75.30%	49%	1.1	69.80%	45%	117 bp	61.10%	0.7	1.1
4_TEb.5	78.50%	48%	5.9	76.50%	46%	147 bp	65.40%	3.8	5.9
4_TEb.6	78.70%	49%	5.7	75.90%	48%	147 bp	54.90%	3.2	5.7
4_TEb.7	66.20%	53%	1	59.20%	51%	137 bp	6.80%	0.1	1
4_TEb.8	67.90%	48%	1.6	63.00%	46%	142 bp	52.80%	0.8	1.6
4_TEb.9	66.30%	52%	1.4	59.50%	50%	137 bp	11.10%	0.1	1.4
6_9a.1	73.90%	47%	14.1	71.70%	46%	147 bp	83.00%	11.6	14
6_9a.2	76.80%	47%	15.6	74.00%	46%	147 bp	84.30%	13.1	15.5
6_9a.3	75.20%	47%	13.3	72.60%	46%	147 bp	82.90%	11	13.3
6_9a.4	72.50%	47%	8.7	69.20%	45%	147 bp	79.80%	6.9	8.7
6_9a.5	65.60%	47%	8.8	63.10%	46%	147 bp	86.20%	7.5	8.7
6_9a.6	62.10%	48%	10.2	60.50%	48%	147 bp	85.20%	8.7	10.2
6_9a.7	81.40%	47%	10.9	78.40%	45%	147 bp	81.50%	8.9	10.9
6_9a.8	72.90%	47%	7.7	70.10%	45%	147 bp	79.00%	6	7.7
6_9a.9	85.30%	46%	9.8	82.60%	44%	147 bp	81.00%	7.9	9.8
7_2a.1	69.60%	49%	12.1	65.70%	48%	147 bp	75.80%	9.2	12.1
7_2a.2	71.40%	49%	6.7	66.80%	47%	147 bp	66.80%	4.5	6.7
7_8a.1	52.70%	47%	4.1	51.00%	46%	147 bp	85.20%	3.5	4.1
7_8a.2	62.80%	48%	9.3	59.80%	47%	147 bp	84.20%	7.8	9.3
7_8a.3	55.90%	48%	5.3	54.10%	47%	147 bp	84.10%	4.4	5.3
7_8a.4	57.80%	48%	6.4	55.70%	47%	147 bp	83.90%	5.4	6.4
7_8a.5	59.90%	48%	8.9	57.80%	48%	147 bp	86.00%	7.6	8.9

Raw reads				Aligned reads after trimming with TrimGalore					
Original name	Duplicated reads (%)	GC content (%)	Millions total sequences	Duplicated reads (%)	GC content (%)	Median Read Length	Aligned reads (%)	Millions uniquely mapped reads	Millions total sequences
7_8a.6	57.70%	48%	9.4	56.00%	47%	147 bp	86.40%	8.1	9.4
7_8a.7	54.40%	47%	6.5	52.90%	47%	147 bp	86.00%	5.6	6.5
7_8a.8	61.20%	48%	9.8	58.50%	47%	147 bp	84.80%	8.3	9.8
8_7a.1	70.30%	48%	14.2	67.00%	47%	147 bp	79.50%	11.3	14.2
8_7a.2	59.20%	48%	10.5	56.80%	47%	147 bp	86.00%	9	10.4
8_7a.3	56.30%	47%	11.5	54.30%	47%	147 bp	86.80%	10	11.5
8_7a.4	47.60%	48%	3.4	46.50%	48%	147 bp	85.90%	2.9	3.4
8_7a.5	86.60%	47%	8.8	84.30%	45%	142 bp	78.00%	6.8	8.8
8_7a.6	72.40%	49%	4.5	66.40%	47%	142 bp	62.90%	2.9	4.5
8_7a.7	65.70%	49%	4.7	61.00%	48%	147 bp	69.70%	3.3	4.7
8_7a.9	49.90%	47%	6.4	48.10%	47%	147 bp	86.50%	5.6	6.4
8_9a.1	53.50%	47%	8.4	51.80%	46%	147 bp	85.10%	7.1	8.4
8_9a.2	66.60%	47%	7.5	63.50%	46%	147 bp	79.80%	6	7.5
8_9a.3	64.60%	51%	0.6	53.90%	46%	77 bp	8.50%	0.1	0.6
8_9a.4	79.80%	46%	6.9	76.70%	45%	147 bp	83.00%	5.7	6.9
8_9a.5	56.50%	48%	8.6	55.20%	47%	147 bp	88.00%	7.5	8.6
8_9a.6	60.20%	48%	8.3	58.10%	47%	147 bp	87.00%	7.2	8.3
8_9a.8	64.00%	48%	8.9	61.70%	47%	147 bp	86.50%	7.6	8.8
9_5a.1	68.90%	48%	8.8	65.00%	46%	147 bp	77.50%	6.8	8.8
9_5a.2	59.30%	47%	6.8	56.90%	46%	147 bp	84.10%	5.7	6.7
9_5a.3	60.40%	47%	10.1	57.40%	45%	147 bp	80.70%	8.2	10.1
9_5a.4	53.30%	48%	6.9	51.60%	47%	147 bp	82.60%	5.7	6.9
9_5a.5	55.70%	47%	5.7	53.60%	46%	147 bp	80.50%	4.6	5.7
9_5b.1	69.50%	47%	9.6	66.10%	46%	147 bp	80.90%	7.8	9.6
9_5b.2	17.80%	49%	0.1	15.90%	46%	147 bp	70.20%	0.1	0.1
9_5b.3	64.30%	50%	3.4	58.20%	47%	132 bp	62.30%	2.1	3.4
9_5b.4	75.00%	49%	5.2	70.00%	47%	127 bp	60.30%	3.1	5.2

Raw reads				Aligned reads after trimming with TrimGalore					
Original name	Duplicated reads (%)	GC content (%)	Millions total sequences	Duplicated reads (%)	GC content (%)	Median Read Length	Aligned reads (%)	Millions uniquely mapped reads	Millions total sequences
9_5b.5	76.80%	48%	13.7	73.10%	46%	147 bp	75.60%	10.3	13.7
Empty	69.50%	54%	1.3	59.80%	52%	97 bp	4.80%	0.1	1.3
Empty	69.70%	55%	1.4	59.80%	50%	87 bp	5.50%	0.1	1.4
Empty	60.50%	55%	0.6	51.40%	52%	102 bp	6.60%	0	0.6
Empty	16.70%	50%	0.1	15.50%	46%	117 bp	45.80%	0.1	0.1
Empty	61.30%	55%	1.1	53.80%	54%	122 bp	6.50%	0.1	1.1
Average excluding empty wells			7.054					5.613	7.039

Picard: RnaSeqMetrics Base Assignments



Created with MultiQC

**Appendix 18. Percentages of aligned reads from pseudo bulk samples located in different genomic regions.**

Percentages were generated using Picard v3.1.1, and the plot was created with MultiQC v1.17. Each bar represents a pseudo bulk sample generated by merging scRNA-seq datasets of single cells derived from an individual embryo. Different colours represent reads aligned to coding regions (blue), UTRs (black), intronic regions (green), intergenic regions (orange), and unaligned reads (violet).

**Appendix 19. Identified SNP variants in pseudo bulk samples. Some SNPs are found in multiple genomic locations, so the numbers in genomic regions may exceed the total number of SNPs reported in VCF files**

Embryo name	Number of cells sequenced (scRNA-seq)	No. of positions in VCF	Annotated SNPs	%	Not annotated SNPs	%	Splice site	%	Intron	%	5' UTR	%	3' UTR	%	Coding	%	Intergenic	%	Promoter	%	ASE positions
1_12a	13	171582	110815	69	50416	31	6007	0	945651	74	17734	1	81377	6	93012	7	25045	2	110101	9	43015
2_10a	10	133640	86027	69	38845	31	4506	1	625942	70	14683	2	56952	6	67636	8	25264	3	93379	11	28242
3_5a	5	54744	33760	65	17788	35	3299	1	233702	64	7590	2	28919	8	36223	10	12001	3	40824	11	10906
3_7a	7	116266	75143	68	35269	32	5060	1	575017	69	14495	2	59893	7	71436	9	18324	2	86565	10	30266
3_Bla	7	101194	66701	70	28431	30	3612	0	576522	75	10460	1	50941	7	49886	6	13561	2	64403	8	19580
3_TEa	41	482860	286621	63	166466	37	9248	0	2848862	77	39434	1	206372	6	235278	6	61360	2	279426	8	136799
3a whole BL	48	548115	323801	63	189903	37	9746	0	3236725	78	42691	1	227689	5	260398	6	71470	2	306755	7	154291
3_Blb	8	141165	93546	69	41241	31	4291	0	863303	77	14799	1	65678	6	68823	6	15940	1	91113	8	34991
4_BLa	11	47402	30010	68	13949	32	2247	1	187928	72	3037	1	14822	6	15935	6	15386	6	20414	8	4437
4_TEa	15	138404	87368	67	43636	33	4110	0	782738	75	13779	1	60784	6	67793	6	21965	2	93249	9	24819
4a whole BL	26	185951	115456	66	59555	34	5236	0	978646	75	16139	1	72607	6	81975	6	37601	3	110701	8	32092
4_TEb	26	171351	113295	70	48241	30	4747	0	912099	78	12715	1	53523	5	60496	5	38219	3	91801	8	29075
6_9a	9	110984	66601	64	37730	36	4864	1	466732	64	15550	2	59161	8	78332	11	21539	3	85435	12	24982
7_2a	2	15250	9034	63	5328	37	1937	2	51246	51	3545	4	10602	11	15529	15	2556	3	15323	15	2743
7_8a	8	160460	104576	69	45989	31	4928	0	823692	73	17134	2	68889	6	78252	7	25211	2	112111	10	39541
8_7a	8	153547	96967	67	47502	33	4656	0	818757	73	16544	1	64559	6	75703	7	24658	2	112719	10	39541

Embryo name	Number of cells sequenced (scRNA-seq)	No. of positions in VCF	Annotated SNPs	%	Not annotated SNPs	%	Splice site	%	Intron	%	5' UTR	%	3' UTR	%	Coding	%	Intergenic	%	Promoter	%	ASE positions
8_9a	7	111373	70604	67	34392	33	4307	1	569832	71	13828	2	53283	7	62572	8	16544	2	87898	11	29944
9_5a	5	121710	75975	66	38356	34	4193	0	679895	75	12747	1	42764	5	55924	6	21995	2	86589	10	25780
9_5b	5	25090	14804	62	8901	38	3001	2	93293	55	5357	3	15755	9	25136	1 5	4309	3	24241	14	4020
Averages per embryo:		148081.9	92693.6	67	46630.4	33	4718. 5	1	791635.5	70	15036 .7	2	64110.1	7	75429.8	8	24045.1	2	97037	10	35295.3

#### Appendix 20. Detailed analysis results for human genes with LTR-associated promoters

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Methylation - cloning
2	<i>GALNT13</i>	10	rs62174125	6	2 - pref. monoallelic, 3 - biallelic, 1 - fully digested	-	BCN 8 & BCN 5 - not informative, low, mosaic methylation
			rs12999856	6	1 - monoallelic, 3 - biallelic, 2 - fully digested	-	BCN 8 - not informative, mosaic methylation; BCN 5 - informative, low, mosaic methylation
			rs10194599	5	2 - monoallelic, 2 - biallelic, 1 - fully digested	-	-
2	<i>SCHLAPI</i>	5	rs144415983	2	1 - pref. monoallelic, 1 - biallelic	-	21BR 21 - informative, mosaic methylation
			rs148398319	2	2 - monoallelic	-	21BR 21 - not informative, mosaic methylation

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Methylation - cloning
4	<i>SLC7A11 - ASI</i>	4	rs7560378	3	1 - pref. monoallelic, 2 - biallelic	-	21BR 21 - not informative, mosaic methylation
			rs7693285	4	1 - pref. maternal (HpaII), 3 - biallelic (HpaII); 2 - pref. maternal (BstUI), 1 - pref. monoallelic (BstUI), 1 - biallelic (BstUI)	-	BCN 8 - informative, mosaic methylation
			rs7699108	4	1 - pref. maternal (HpaII), 3 - biallelic (HpaII); 2 - pref. maternal (BstUI), 1 - pref. monoallelic (BstUI), 1 - biallelic (BstUI)	-	BCN 8 - not informative, mosaic methylation
17	<i>LOC33916 6</i>	10	rs12450161	9	2 - maternal, 2 - paternal, 4 - monoallelic, 1 - pref. monoallelic	-	-
			rs12450165	9	3 - maternal, 2 - paternal, 4 - monoallelic	-	-
			rs12453225	9	2 - pref. maternal, 1 - pref. paternal, 4 - pref. monoallelic methylation, 2 - biallelic	1 - maternal, 3 - monoallelic, 1 - pref. maternal, 1 - pref. paternal, 1 - pref. monoallelic, 2 - biallelic	21BR 430 - informative, low, mosaic methylation

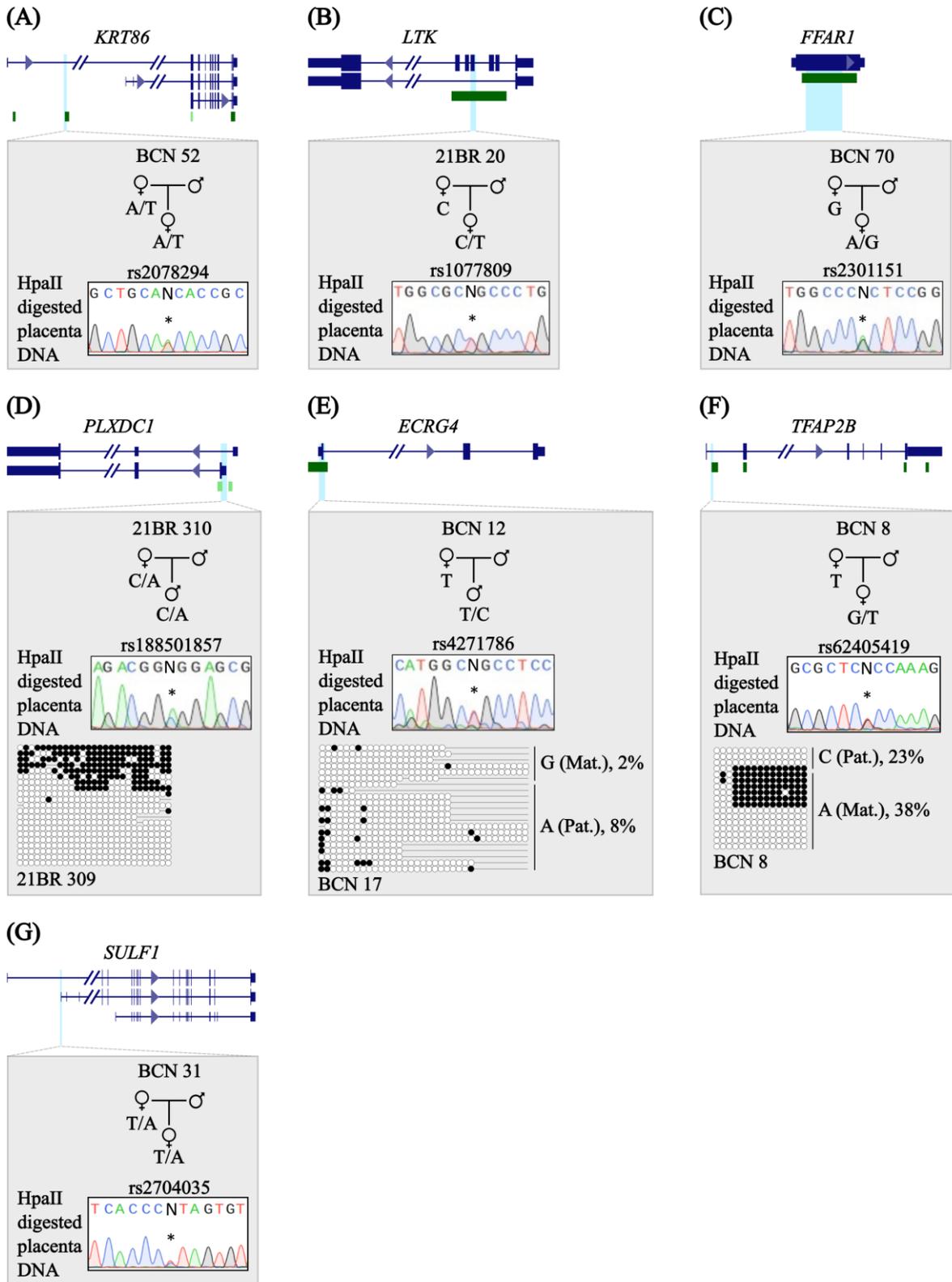
**Appendix 21. Detailed analysis results for human genes with placental sDMRs**

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Methylation - cloning
1	<i>DNAJC6</i>	6	rs577841	6	1 - monoallelic, 1 - pref. maternal, 2 - pref. monoallelic, 2 - biallelic methylation	-	-
			rs4271786	13	1 - maternal, 1 - paternal, 1 - pref. paternal, 4 - pref. monoallelic, 5 - biallelic	-	BCN 17 - informative, both alleles mostly unmethylated
2	<i>C2ORF40 &amp; ECRG4</i>	12	rs4266035	13	1 - maternal methylation, 1 - paternal, 3 - pref. monoallelic, 7 - biallelic	-	BCN 17 - informative, both alleles mostly unmethylated
			rs73949223	1	1 - pref. monoallelic	-	-
			rs4477942	9	1 - maternal, 1 - pref. monoallelic, 7 - biallelic	-	-
4	<i>CRMP1</i>	2	rs139357095	2	2 - biallelic	-	-
4	<i>CWH43</i>	14	rs3747690	14	1 - monoallelic, 1 - pref. maternal, 1 - pref. paternal, 1 - pref. monoallelic, 8 - biallelic, 2 - fully digested	-	-
5	<i>C5ORF38</i>	3	rs62333235	2	2 - biallelic	-	-
			rs76652220	3	3 - biallelic	-	-
			rs72633976	6	1 - paternal, 4 - monoallelic, 1 - pref. monoallelic	-	-
5	<i>ANKDD1B</i>	6	rs1489	2	1 - monoallelic, 1 - pref. monoallelic	-	-
			rs61516153	2	2 - monoallelic	-	-

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Methylation - cloning
			rs4628086	1	1 - biallelic	-	BCN 8 - not informative, low methylation
6	<i>TFAP2B</i>	3					BCN 8 - informative, both alleles
			rs62405419	3	3 - biallelic	-	methylated, low methylation, maternal allele more methylated
8	<i>SULFI</i>	4	rs2704035	4	4 - biallelic	-	-
8	<i>RGS22</i>	12	rs2453627	12	1 - pref. monoallelic, 11 - biallelic	-	-
12	<i>KRT86</i>	2	rs2078294	2	2 - biallelic	-	-
12	<i>FAM101A</i> (all isoforms)		rs12318072	7	-	2 - pref. monoallelic, 5 - biallelic	-
12	<i>FAM101A</i> (NM_001204299)	7	rs12823740	1	-	-	21BR 308 - informative, unmethylated
12	<i>FAM101A</i> (NM_001365156)		-	-	-	-	BCN 5 - not informative, low, mosaic methylation
			rs12318072	5	-	3 - maternal, 2 - biallelic	-

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Methylation - cloning
12	<i>FAM101A</i> (NM_181709)		rs12318072	6	-	1 - pref. maternal, 1 - pref. monoallelic, 4 - biallelic	-
15	<i>LTK</i>	4	rs1077809	4	4 - biallelic	-	-
17	<i>PLXDC1</i>	4	rs188501857	4	1 - monoallelic, 1 - pref. monoallelic, 2 - biallelic	-	21BR 309 - not informative mosaic methylation, lower methylation; BCN 50 - not informative mosaic methylation.
19	<i>NUDT19</i>	15	rs8108621	6	1 - maternal, 1 - paternal, 1 - monoallelic, 3 - biallelic	1 - pref. maternal, 1 - pref. monoallelic, 1 - biallelic	BCN 17 - not informative, mosaic methylation
			rs8109823	6	1 - maternal, 1 - paternal, 1 - pref. maternal, 1 - monoallelic, 2 - biallelic	1 - pref. maternal, 2 - biallelic	BCN 17 - not informative, mosaic methylation
			rs61732600	11	3 - monoallelic, 3 - pref. monoallelic, 5 - biallelic	2 - biallelic	BCN 17 - informative, mosaic methylation, paternal allele more methylated

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Methylation - cloning
19	<i>FFAR1</i>	8	rs2301151	8	1 - pref. maternal, 1 - pref. monoallelic, 6 - biallelic	-	-
20	<i>TSPY26P &amp; PLAGL2</i>	12	rs11907716	9	1 - maternal, 7 - monoallelic, 1 - biallelic	-	-
			rs11907235	12	2 - maternal, 2 - paternal, 8 - monoallelic	-	-



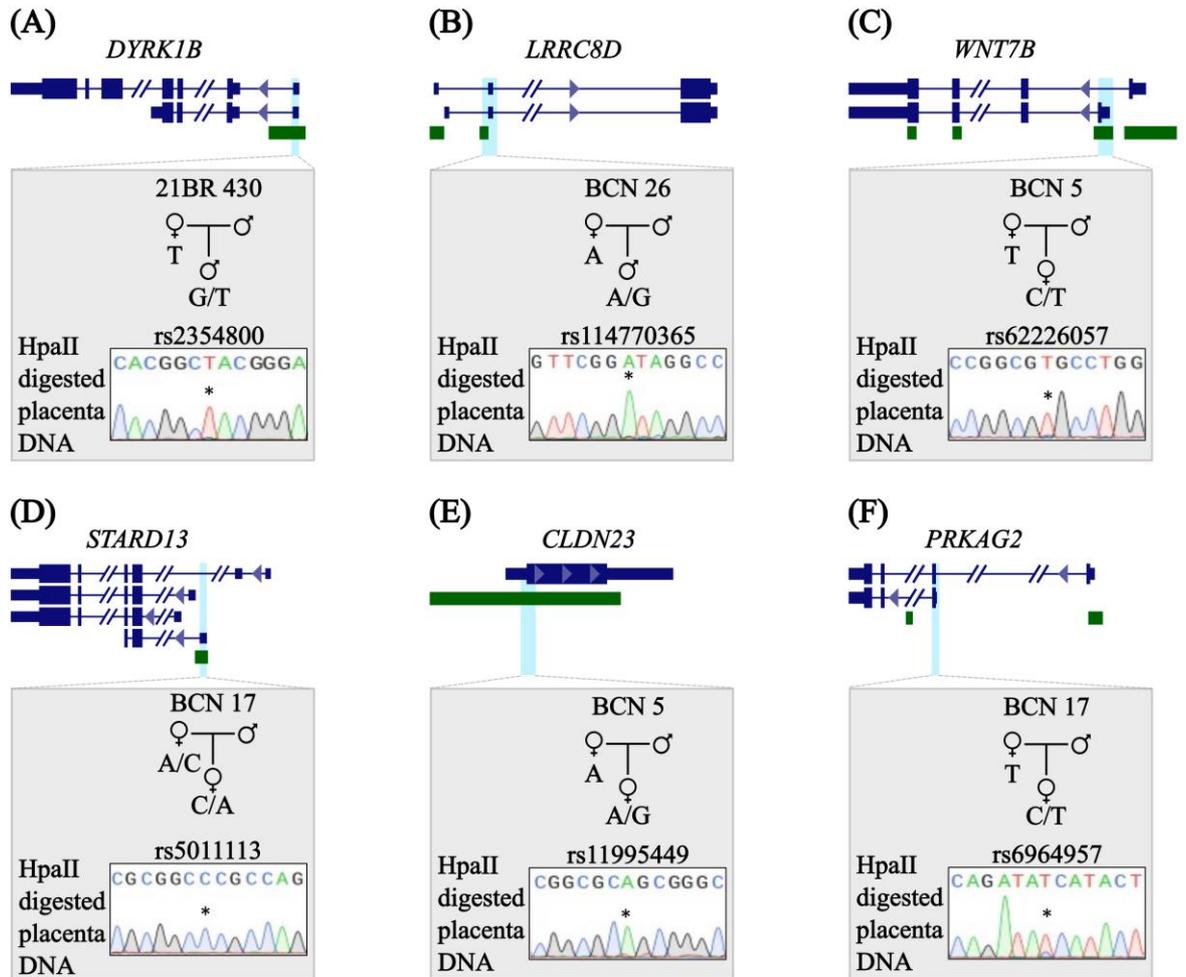
**Appendix 22. Allelic methylation of human candidate genes with placental sDMRs. For each gene, allelic methylation was investigated using methylation-sensitive genotyping.**

For (D) *PLXDC1*, (E) *ECRG4*, and (F) *TFAP2B*, DNA methylation was further analysed by bisulphite PCR and sub-cloning of placenta-derived DNA. Gene transcripts are shown in dark blue, with thicker bars representing exons, while CpG islands are shown in dark green. Methylated cytosines are indicated by (●), and unmethylated cytosines by (○). Each row corresponds to an individual cloned sequence, with the parent-of-origin inferred through SNP genotyping if the placental sample was heterozygous.

**Appendix 23. Detailed analysis results for human genes with placenta-specific mDMRs**

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Bisulphite PCR
1	<i>LRRC8D</i>	4	rs114770365	4	2 - maternal, 1 - monoallelic, 1 - biallelic	-	-
			rs115363384	4	2 - maternal, 1 - monoallelic, 1 - pref. monoallelic	-	-
			rs114208181	4	2 - maternal, 1 - monoallelic, 1 - pref. monoallelic	-	-
			rs113834473	4	2 - maternal, 1 - monoallelic, 1 - biallelic	-	-
7	<i>PRKAG2</i>	11	rs6964957	8	3 - pref. maternal, 4 - pref. monoallelic, 1 - biallelic	-	-
7	<i>PRKAG2</i> (NM_024429)		rs6964957	8	3 - pref. maternal, 4 - pref. monoallelic, 1 - biallelic	-	-
			rs8961	10	-	1 - maternal, 1 - paternal, 1 - monoallelic, 2 - pref. maternal, 1 - pref. paternal, 2 - pref. monoallelic, 2 - biallelic	-
8	<i>CLDN23</i>	9	rs9644774	8	2 - maternal, 2 - monoallelic, 4 - fully digested	Not expressed	-
			rs11995449	3	2 - maternal, 1 - fully digested	Not expressed	-
12	<i>EID3</i>	5	rs7488680	5	1 - pref. monoallelic, 3 - biallelic, 1 - fully digested	-	-
			rs58078551	1	1 - biallelic	-	-
13	<i>STARD13</i>	13	rs5011113	8	2 - maternal, 3 - monoallelic, 3 - fully digested	-	-
13	<i>STARD13</i> (AK308453)		rs495680	9	-	1 - monoallelic, 1 - pref. paternal, 7 - biallelic	-

Chr.	Gene	Total no. of informative samples	SNP	Number of heterozygous samples	Methylation-sensitive genotyping (HpaII & BstUI)	Allelic RT-PCR	Bisulphite PCR
			rs5011113	8	2 - maternal, 3 - monoallelic, 3 - fully digested	-	-
19	<i>MBD3</i> (all isoforms)	12	rs8104174	12	4 - maternal, 4 - monoallelic, 1 - pref. monoallelic, 3 - biallelic	-	-
			rs190802753	2	-	1 - pref. paternal, 1 - biallelic	-
19	<i>MBD3</i> (AK001474)	12	rs8104174	12	4 - maternal, 4 - monoallelic, 1 - pref. monoallelic, 3 - biallelic	-	BCN 6 & BCN 8 - informative, maternal methylation
			rs190802753	2	-	1 - pref. monoallelic, 1 - biallelic	-
19	<i>DYRK1B</i>	6	rs2354800	6	2 - maternal, 1 - monoallelic, 1 - pref. maternal, 2 - biallelic	-	-
22	<i>WNT7B</i>	10	rs62226057	10	4 - maternal, 1 - monoallelic, 5 - fully digested	Not expressed	-



#### Appendix 24. Characterisation of allelic methylation at human candidate loci with placenta-specific mDMRs.

Maternal methylation was determined for (A) *DYRK1B*, (B) *LRRC8D*, (C) *WNT7B*, (E) *CLDN23*, and (F) *PRKAG2* through methylation-sensitive genotyping, while monoallelic methylation was observed for (D) *STARD13*. The promoters of these genes (highlighted in light blue) were hypermethylated in human oocytes, hypomethylated in sperm, and partially methylated in blastocysts and placental tissues, based on methyl-seq datasets. Gene transcripts are shown in dark blue, with thicker bars representing exons, and CpG islands are depicted in dark green.

#### Appendix 25. Previously reported placenta-specific imprinted genes exhibiting polymorphic imprinting

<i>THAP3</i> region: chr1:6684860-6685996 (GRCh37 regions)					GEO accession: GSE159526	
PM374						
Cell types:	Endothelial	Hofbauer	Stromal	Trophoblast	Villi	
	GSM4831792-118762	GSM4831975-119000	GSM4831859-118829	GSM4831791-118761	GSM4831970-118974	
Probes:						
cg11688219	0.038	0.015	0.073	0.241	0.192	
cg24312985	0.345	0.469	0.322	0.543	0.500	
cg02357257	0.112	0.047	0.172	0.404	0.327	
cg24635178	0.164	0.042	0.203	0.426	0.396	

PM374					
Cell types:	Endothelial	Hofbauer	Stromal	Trophoblast	Villi
	GSM4831792-118762	GSM4831975-119000	GSM4831859-118829	GSM4831791-118761	GSM4831970-118974
Probes:					
cg14848383	0.257	0.033	0.301	0.549	0.521
cg27300967	0.233	0.027	0.289	0.494	0.505
cg06826730	0.076	0.019	0.126	0.278	0.243
cg08250099	0.164	0.036	0.250	0.529	0.466
Average:	0.174	0.086	0.217	0.433	0.394
DMR/ promoter					

PM374					
Cell types:	Endothelial	Hofbauer	Stromal	Trophoblast	Villi
	GSM4831792-118762	GSM4831975-119000	GSM4831859-118829	GSM4831969-118973	GSM4831970-118974
Probes:					
cg20015272	0.633	0.163	0.468	0.302	0.466
cg02829743	0.676	0.154	0.666	0.525	0.572
cg02391713	0.640	0.050	0.639	0.514	0.611
cg22620090	0.419	0.031	0.491	0.343	0.374
cg09723833	0.503	0.074	0.518	0.412	0.407
Average:	0.574	0.094	0.557	0.419	0.486
DMR/ promoter					

### Appendix 26. Different PCR primers for human *PIK3R1* and *G0S2*

Gene	Variants	Name	Sequence (5' -> 3')
<b>Genotyping (DNA)</b>			
<i>PIK3R1</i>	rs138814985, rs2888323	F	GTTGGCTTCTCAATGAGGAG
		R	AATCCCCAAAGCTGTTCTTCCA
	rs3730089	DMR seq R	GCTGTTCTTCCACCAAGTG
		PIK 089 F	TCCATATTGCATGGAATTGTGAACT
<i>G0S2</i>	rs1815548, rs932375	PIK 089 R	CTCCCCAGTACCATTTCAGCATC
		F	TGGGACCTTCGCGTGCACACT
		R	GCTCTCCCAGTTGGAGACTCCG
<b>Bisulphite PCR</b>			
<i>PIK3R1</i>	rs138814985, rs2888323	Bis F	AGTTGGTTTTTTAATGAGGA
		Bis in R	CCACCAAATAAACCAAACCCC
		Bis out R	CCCTTAAAATACCTATATCC
<i>G0S2</i>	rs1815548, rs932375	Bis F	GTTGTAGTTTTTTAGTTGGAG
		Bis in R	TACACACTAACCTTCCCAC
		Bis out R	TCCCTAAACTCCGAATCCTCCCCT
		Bis in seq F	GATTGTGTGAGTTAGGGGGT
<b>Pyrosequencing</b>			

Gene	Variants	Name	Sequence (5' -> 3')
<i>PIK3R1</i>	rs138814985, rs2888323	Bis F	AGTTGGTTTTTTAATGAGGA
		Bis out R Bio	[BTN]CCCTTTAAAATACCTATATCC
<i>G0S2</i>	-	Bis F	GTTGTAGTTTTTTTAGTTGGAG
		Bis in seq F	GATTGTGTGAGTTAGGGGGT
		Bis out R Bio	[BTN]TCCCTAAACTCCGAATCCTCCCT
<b>RT-PCR</b>			
<i>PIK3R1</i> (isoform 3)	rs138814985, rs2888323, rs3730089	New RT F	CAATGAGGAGCCGGCAGTGAGC
		RT R	AGATATCTCCCCAGTACCATTCA
<i>PIK3R1</i> (isoform 3) nested RT-PCR (Out/Out)	rs138814985, rs2888323, rs3730089	New RT F	CAATGAGGAGCCGGCAGTGAGC
		PIK 089 R	CTCCCCAGTACCATTACAGCATC
<i>PIK3R1</i> (isoform 3) nested RT-PCR (In/Out)	rs3730089	q in F	GGGAAACCGTTGAAATGCATAACCTG
		PIK 089 R	CTCCCCAGTACCATTACAGCATC
<i>PIK3R1</i> (isoform 3)	rs3730089	Trans 2 seq F	TTTTTCATTGTTCGGATACAGGCATT
		PIK 089 R	CTCCCCAGTACCATTACAGCATC
<i>PIK3R1</i> (isoform 3) MACS fractions	rs138814985, rs2888323, rs3730089	F	GTTGGCTTCTCAATGAGGAG
		RT R	AGATATCTCCCCAGTACCATTCA
<i>PIK3R1</i> (isoform 3) sequencing primer for the common SNP	rs3730089	Trans 2 seq F	TTTTTCATTGTTCGGATACAGGCATT
<i>PIK3R1</i> (isoform 3) sequencing primer for the DMR variants	rs138814985, rs2888323	R	AATCCCCAAAGCTGTTCTTCCA
<i>PIK3R1</i> (isoform 1)	rs3730089	Trans 1 RT F	TCTCTGAAATTTTCAGCCCTATGCT
		PIK 089 R	CTCCCCAGTACCATTACAGCATC
<i>PIK3R1</i> (isoform 1) sequencing primer	rs3730089	Trans 1 RT F	TCTCTGAAATTTTCAGCCCTATGCT
<i>PIK3R1</i> (isoform 2)	rs3730089	Trans 2 q out F	AACTGAGCTCAGCCAAGGAA
		PIK 089 R	CTCCCCAGTACCATTACAGCATC
<i>PIK3R1</i> (isoform 2) sequencing primer	rs3730089	Iso 2 new q F	ACTTGATGTTTTATATAGAAATGGA
<i>G0S2</i> nested RT-PCR (Out/ Out)	rs932375	RT F	GCTCTGACCGCGCTGGCCTGG
		RT R	GAGGCGGGAATGACCTTAGTGG
<i>G0S2</i> nested RT-PCR (Out/ In)	rs932375	RT F	GCTCTGACCGCGCTGGCCTGG
		In RT R	GAATGACCTTAGTGGCACGGCGCGAG
<b>qRT-PCR SYBR™ Green</b>			
<i>ACTB</i>	-	m-h B-actin 1F	CCGGCTTCGCGGGCGACGAT
		m-h B-actin 1R	CTCCATGTCGTCCAGTTGG
<i>RPL19</i>	-	m-h L19 q RT F	AATCGCCAATGCCAACTCCCGTCA
		m-h L19 q RT R	CCTATGCCCATGTGCCTGCCCTTC
<i>PIK3R1</i> isoform 3	-	q In F	GGGAAACCGTTGAAATGCATAACCTG
		All new q R	GTTTTTCATTCACTTCTCCCTCGAG
<i>PIK3R1</i> isoform 1	-	Trans 1 RT F	TCTCTGAAATTTTCAGCCCTATGCT
		Iso 1 q RT R	GTCGTTCACTCCATTACAGTTGAG
<i>PIK3R1</i> isoform 2	-	Iso 2 q RT F	ATGTTTTATATAGAAATGGACCCA
		All new q R	GTTTTTCATTCACTTCTCCCTCGAG
<i>PIK3R1</i> all isoforms	-	All iso q F	AGCTATTGAAGCATTTAATGAAACCA
		All iso q R	CACTGATTTCGAGACTTCAACTTATC

Gene	Variants	Name	Sequence (5' -> 3')
<i>KRT7</i>	-	q RT F	CAGGCTGAGATCGACAACATC
		q RT R	CTTGGCACGAGCATCCTT
<i>VIM</i>	-	q RT F	GGCTCAGATTCAGGAACAGC
		q RT R	AGCCTCAGAGAGGTCAGCAA
<i>CGB3</i>	-	q RT F	GTGTCGAGCTCACCCCAGCATCCTA
		q RT R	AGCAGCCCCTGGAACATCT
<i>COL3A1</i>	-	q RT F	GGAGCTGGCTACTTCTCGC
		q RT R	GGGAACATCCTCCTTCAACAG
<i>CD45</i>	-	CD45/PTPRC F	AGCTAAGGCGACAGAGATGCCTGA
		CD45/PTPRC R	CTCACTGGGTGGATCCCTTTTCTTC
<i>CD14</i>	-	q RT F	CGGAAGACTTATCGACCATGGAGC
		q RT R1	AAGGCTTCGGACCAGTCGGGCTGA
<b>TaqMan™</b>			
<i>G0S2</i>	-	Assay Id: Hs00377852_g1 ( <i>G0S2</i> ) TaqMan Gene Expression Assay (FAM)	
		<i>RPL19</i>	-
<b>Cloning</b>			
pGEM®-T Easy Vector	-	pGEMt out F	GATGGTGCTGCAAGGCGATTAAGTTG
		pGEMt out R	ATGTTGTGTGGAATTGTGAGCGGA
		Sp6 primer	ATTTAGGTGACACTATAG
		MF13 F	GTAAAACGACGGCCAG
		Seq(S) T7	TAATACGACTCACTATAGGG

#### Appendix 27. Different PCR primers for mouse *Pik3r1* and *G0s2*

Gene	Variants	Name	Sequence (5' -> 3')
<b>Bisulphite PCR</b>			
<i>Pik3r1</i>	-	Bis out F	TATTAAGTGGTTTTAGTTTTTGGAG
		Bis in F	GTAAAGAATTTAGTTGGAGGAGAG
		Bis in R	TAACTCAACAAATATTTAAACCT
		Bis out R	TAACTTAAATACCCCTCCCCCT
<i>G0s2</i>	rs31626975	Bis out F	AGTTAAGAAAGTAGTATTTTGGGAAGA
		Bis in F	TTTTGATTGGTGAGAGGTGATTTTT
		Bis in R	CTAAAAACCCAAAACACCACTTC
		Bis out R	CCAAAAAATAACCACRAATAATAC
<b>RT-PCR</b>			
<i>Pik3r1</i>	rs37236366, rs13463306	RT 1F	ATTATGCATAACCATGATAAGCTGA
		RT 1R	CGTTGCTGCTCCCGACATTCCAC
<i>G0s2</i>	No expression in the placenta	RT F	AACGCCAAAGCCAGTCTGACGCA
		RT R	GATCTGTGTGGGGTCAGTTCTGG

**Appendix 28. Different primers used for non-canonical imprinting analysis**

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
<b>Genotyping (DNA)</b>					
1	<i>LRRC8D</i>	-	rs114770365, rs115363384, rs114208181, rs113834473	Forward	TCTATAACGTGCTGCCGGGTCT
				Reverse	CAGCTCCAGCGCAGCCCGGGGC
1	<i>ERO1B</i>	-	rs557205, rs73117239	Forward	AAACGAAACGAAGCCAAACAGA
				Reverse	CGGTGTCAGTGTGACTACATTTC
				Sequencing Reverse	GTGTTACATTTACATAGTGG
2	<i>ECRG4/ C2ORF40</i>	-	rs4271786, rs4266035, rs73949223, rs4477942	Forward	GAGAGAGGACCTCGGTGGTACT
				Reverse	CACCCATCACCGATCGCTCT
				Sequencing Forward	GGCAGCGACGCAGGGATAAC
2	<i>GALNT13</i>	-	rs62174125, rs12999856, rs10194599	Forward	TTGATCTGAGGCTGAATCCCGT
				Reverse	CAGAAAGTTCCGCGCCACGCGGTC
2	<i>SCHLAP1</i>	-	rs144415983, rs148398319, rs7560378	Outer Forward	CACTCACCGCGAAGGTCCGCAGC
				Outer Reverse	TTTCAGTCTGACCAATCAGGAGT
				Inner Forward	GAGGAACGAACAACTCCCCGAC
				Inner Reverse	CAGCCAGCACAGTGTTACCTAGA
3	<i>RASSF1</i>	HpaII Control gene	rs4688725	Forward	ATGCGCAGCGCGTTGGCAGCTCC A
				Reverse	GATCCTGGGGGAGGCGCTGAAG
4	<i>CRMP1</i>	-	rs139357095	Forward	GTACCTGGCCATTGTCCCGGCCGA G
				Reverse	AGGGCGCTACGAGAACAAGACC A
4	<i>CWH43</i>	-	rs3747690	Forward	AGGAGGCAAAGGCGGGGACCAGA
				Reverse	CAAGAGGATTCTCTCCACAGC
4	<i>JADE1</i>	NM_001287 437	rs11933240	Forward	AGGAGAAGCATCTTGCTTCTTGA
				Reverse	CAAATAACACAACTTCTCAC
				Sequencing Forward	TGAGTAGCTGGGATTACAGGCGT
4	<i>SLC7A11- AS1</i>	-	rs7693285, rs7699108	Forward	TCACTGCCCCGGTGCTTGCGGGCT
				Reverse	GTAAAACAAATACTTCTTCG
4	<i>GAB1</i>	-	rs1397529	Forward	AGATGAATTGTAGACTAGTAACA
				Reverse	GATAGTTTAGGCACATTTTCAGG
				Forward	TTAGAAGCCTGCCCCAGAGTCT
5	<i>C5ORF38</i>	-	rs62333235, rs76652220	Reverse	CTTCTCTGTACCTCTGACTTC
				Forward	TCAGGAGTCGCTTAGGTTTT
5	<i>ANKDD1B</i>	-	rs72633976,	Reverse	GGCAGTTTCAGGTTCCCTGGTG
5				Forward	CAGGTCTCCCTGAGACCCTT

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
			rs1489, rs61516153	Reverse	GATTATCCCAGGCCAGCCCAAGTC
6	<i>TFAP2B</i>	-	rs4628086, rs62405419	Forward	TAGCAGTTTATTAGTTTCTGTTTTCT T
				Reverse	GGAGCCGTCTGGCCGCGTCAG
6	<i>SMOC2</i>	-	rs76776636, rs73270928	Forward	AGGAGCGAGGGCGGACGCAAAGA
				Reverse	GCAAGAGGCGGCACCACTTGAG
7	<i>PRKAG2</i>	5'UTR/ DMR	rs6964957	Forward	CCCATCCCTGCAGAGTGCAC
				Reverse	GCCTGGTTTCTGAACTTCATAG
		Exonic SNP	rs8961	Forward	ACAAAAGGAGACAGAAACGGA
				Reverse	CAACATCACTGGAAGAAATAC
8	<i>CLDN23</i>	-	rs9644774, rs11995449	Forward	TGACTTCGGGTCCCCGGAGCCT
				Reverse	GTCCACTGGCTGGTTCAGGAAG
8	<i>DLGAP2</i>	HpaII Control gene	rs36018196	Forward	AGTAAGATTTTGTGTTGGAGAAAG TTAYG
				Reverse	CRTCCTTATCRAACAAAAACCRG
8	<i>KLF10</i>	HpaII Control gene	-	Forward	AGGAAGTATAGGGGTATTTTTAAA TGA
				Reverse	CTCACACACCTTTACCGTTAATTA AC
8	<i>SULF1</i>	-	rs2704035, rs2725092	Forward	AGTTTGTGGCCGAGGTTTGCA
				Reverse	CTCTGATCCTCGCTGCCCTCGC
8	<i>RGS22</i>	-	rs2453627	Forward	ACCCCAGCGCGGTCACCCGGAA
				Reverse	GCATTCATACAACGTGTGATG
10	<i>SFMBT2</i>	-	rs10795530	Forward	AGGTTGTAATGCAGTGGCGCA
				Reverse	CAGATGTTCTAGGCTTCAATC
				Sequencing Forward	TCTCAGCTCACTGCAACCTCT
10	<i>GSTO1</i>	-	rs4925	Forward	TCCTTGGTAGGAAGCTTTAT
				Reverse	GAATTTTCCTAATTACCTTAAAG
12	<i>SLC38A1</i>	-	rs1045278	Forward	CATCCCTGTTGTCTGCTCAGTC
				Reverse	ATATGATTGTATGAAATTTGAAAA A
				Forward	GTTCTGAACATCAACACAAAG
		Exonic SNP	rs2429467	Reverse	TGCTCATTGCTGCCTTTTCT
				Sequencing Forward	GGAAGAACCCTTAAGCTGAAGG
12	<i>SLC38A4</i>	Upstream 5' UTR/ CpG island	rs4994910, rs74851348	Forward	GCCACCTCTCCTGGACTCAAGGGT G
				Reverse	AAGGGAGAAGGCGAGAGCAGA
				Sequencing Reverse	CGGTTCCGAGGGCGGCTTAC
12	<i>KRT86</i>	-	rs117031005, rs2078294	Forward	GTGAGGCCGCGGTAGCAGGAG
				Reverse	TGGCTCGCTTTCATTCCCGGCT
12	<i>EID3</i>	-	rs7488680, rs58078551	Forward	CCAAACACCACCTTGCAAAGAAGC
				Reverse	AAGTGGCGGCAGTTAGAGCCGA
12	<i>FAM101A</i>	-	rs12318072	Forward	CAACTCTGAGGTCAAGTACGCC
				Reverse	TCCGAAAGTGCTCCTGGCAT
				Sequencing Reverse	GGCATGCTTGGGGAAGATG

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
13	<i>STARD13</i>	DMR / 5' UTR/ CpG island	rs5011113	Outer Forward	CAGCGGAGCAGGGGACAGCC
				Outer Reverse	CAGGTCAGTGCCCCGGAGAC
				Inner Forward	CAGCCCCTCCAGGTAACCCGTC
				Sequencing Inner	AGGGCATGAGTTTCAGAGCCCA
				Reverse	
				Forward	CTCCAGAATTCGCCGCCACC
		Exonic SNP	rs495680	Reverse	ATCCTTGGGAATATGCACCA
14	<i>SMOC1</i>	-	rs3742909, rs146095118	Forward	TTCTAATAAGTGACCGTGAC
				Reverse	CATCTACCTCGATGCACCACGC
15	<i>LTK</i>	-	rs1077809	Forward	CCCACTGGCTGCGCTCACTCC
				Reverse	TAGCCCTTACCCGGAACCTCTT
15	<i>SNURF</i>	HpaII Control gene	rs4906939	Forward	ACTGCGCCACAACCGGAAAGGA
				Reverse	GTAGAGCCGCCAGTGGGGAGG
16	<i>SALL1</i>	-	rs11645288	Forward	GCTGATGACTCTGGGGGCATG
				Reverse	TGTGGCAAACCTTCTCCTCAT
17	<i>LOC339166</i>	-	rs12450161, rs12450165, rs12453225	Forward	CTCCAGACGCGCCGCCTTAAG
				Reverse	ATATGGAGGGACTGCCCTGTAGA
17	<i>PLXDC1</i>	-	rs188501857	Forward	CCCGCCAGTCCTACCTGCTCC
				Reverse	TCGCGCTCTCGCCGCTCCT
18	<i>ZNF516</i>	-	rs690353	Forward	ACATCTTACCTCTGTGCTCCA
				Reverse	CGTCTTACACTCCATCAAAC
19	<i>MDB3</i>	DMR/ close to CpG island/ downstream 5' UTR	rs8104174	Forward	TGGCACCAATACCCTGCACATT
				Reverse	CCAGGCCGGACTGCATATCC
				Forward	AGGAGGAGGAGCCCGACCCGGA
				Reverse	GCCAGGAGCACGGCTTCCTCCTG
19	<i>NUDT19</i>	-	rs8108621, rs8109823, rs61732600	Forward	AAGGCTTCATGCCGGGCGCGCA
				Reverse	CGCAGGAAGTGGCGCGGGTCC
19	<i>DYRK1B</i>	-	rs2354800	Forward	TCCCTTGGCCTCGTGCTAAGTCT
				Reverse	GGGGCGGAGTCCAGGGCGTGG
20	<i>TSPY26P</i>	-	rs11907716, rs11907235	Forward	TCTTGAAGATGGCGCCCTCCTCCT
				Reverse	CAGAGGCTCCCGCAGGCGATGGC
				Sequencing Reverse	GAGGGCGGGGACCCCAGAAG
20	<i>ZFP64</i>	-	rs3746413	Forward	GTCGGAGCATCCTGAGAAGTG
				Reverse	TCTAGAGCCTCAGTCTTAACCAT
22	<i>WNT7B</i>	-	rs62226057	Forward	GAGCCTGTTTACGCCCCGCCAG
				Reverse	GTGCTCCACCTCGGCAGCTTAG
X	<i>XIST</i>	-	rs1894271	Forward	TGAAGGACAGCATGGTTGGT
				Reverse	CACATGGAATGAGCAGTGTGC

**Bisulphite PCR**

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
2	<i>ECRG4/ C2ORF40</i>	-	rs4271786, rs4266035	Outer Forward	GGTTTTAGTATAGGAGTAGGAGTA G
				Outer Reverse	TTAACCTTAAAACCCAAAAACT
				Inner Forward	GGTTAGGGTTAGGATAGTAGG
				Inner Reverse	CTTAACCCTCAACCCTCTAA
2	<i>GALNT13</i>	-	rs62174125, rs12999856	Outer Forward	GGGGTTGGTYGAGGTTGGA
				Inner Forward	GGGTAAGTGTGAAGAGAGAGG
				Reverse	AAAACCTACTATCCTAACCA
2	<i>SCHLAP1</i>	-	rs144415983, rs148398319, rs7560378	Outer Methylated Forward	GAGTTGTAATATTTATCGCGAAGG
				Outer Unmethylate d Forward	GAGTTGTAATATTTATTGTGAAGG
				Outer Reverse	ACTCCTATAACTAATTTATATTCTC A
				Inner Forward	GAGTTTATTGGGAGGAA
				Inner Reverse	AAAAAACTCACCCATAAAACTTA
3	<i>RASSF1</i>	Control gene	rs4688725	Forward	GTTTTAGATGAAGTCGTTATAGAG
				Reverse	TAAACTACGAAAATAACACCC
3	<i>RPL39L</i>	LTR promoter (MER4E1, MER61C)	rs141173382	Outer Forward	GTGTAAGTTATAGGGGATGTGATG
				Outer Reverse	ACATATTCAATATAAACCAACCA
				Inner Forward	GTTTGGTTTGGGTTTAGAGGTTTG
		Inner Reverse	AACCCTAACTACATTATCTACA		
		Promoter/ 5' UTR	-	Outer Forward	GTTGTTTAGTTGTTGTTGG
				Inner Forward	GTTGTTTGGTTAYGGTATTTAG
Reverse	AACCCAACTATAAACCTCTAA				
4	<i>JADE1</i>	-	rs62317870	Outer Forward	ATTAGTTYGGTGTAGTGA
				Outer Reverse	CTACCRCTCCCATCTTAAAAC
				Inner Forward	TTTTATTTGAAAGTGTTATTT
				Inner Reverse	TTTAATTTACAACCTAAAACC
				Outer Forward 2	TTTTATTTGAAAGTGTTATTT
				Outer Reverse 2	CTTCCTATAACAAAAATAATAC

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
				Inner Forward 2	GGTTTTAGTTGTAAATTA
				Inner Reverse 2	ACAATTTCCAAAACATC
4	<i>SLC7A11-ASI</i>	-	rs7693285, rs7699108	Forward	AGTTTTTTATTGTTYGGTGT
				Inner Reverse	CTTACCAAACAAAATAAATCC
				Outer Reverse	ATAAATTAATACCAATTCCTATTA AA
4	<i>GAB1</i>	AK295684 promoter	rs62337524	Outer Forward	GAATAGTTTTTGGGAGGTGG
				Outer Reverse	TAACTAACCTACACCCAAAT
				Inner Forward	GTTATAGGGAGGATTATTTG
				Inner Reverse	ATAACTTCAACTACTCCACATTA
		Major promoter	-	Outer Forward	TGGAGTTTGTTYGTTTAGTT
				Outer Reverse	CAACTCTACTTACATAAC
				Inner Forward	ATTAGGAGAGTTAGGTTTT
				Inner Reverse	AAACAAACCACTTCACCACC
6	<i>TFAP2B</i>	-	rs4628086, rs62405419	Outer Methylated Forward	GTTTCGAGTCGGAAAAGGGTTTTG
				Outer Unmethylated Forward	GTTTGAGTTGGAAAAGGGTTTTG
				Outer Reverse	ACTTCCTTAAAAATCACTA
				Inner Reverse	CCTCCTATATAAACATCTTTCA
8	<i>DLGAP2</i>	Control gene	rs36018196	Forward	AGTAAGATTTTGTGTTGGAGAAAG TTAYG
				Reverse	CRTCCTTATCRAACAAAACCRG
8	<i>KLF10</i>	Control gene	-	Forward	AGGAAGTATAGGGTATTTTTAAA TGA
				Reverse	CTCACACACCTTTACCGTTAATTA AC
10	<i>GSTO1</i>	-	-	Outer Forward	TGTAAATTTTAGAGGAGTT
				Outer Reverse	AAACCCCCRTATCCCA
				Inner Forward	TGGGGAAGGGTGAGGTTTGT
				Inner Reverse	CCCACTACAACCTCCRACC
12	<i>SLC38A1</i>	-	-	Outer Forward	GAGGGGTAGAGTATTAGGAAGG

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
				Outer Reverse	ATATATTTTCCTATAACTATAACA
				Inner Forward	TAAATATGTTTCGGTTTAGTGG
				Inner Reverse	ACCAAATCTAATTCATTTTFA
				Outer Forward	GTTGGATGTGGGTTTTTGGTTTTTG
		Upstream 5' UTR/ CpG island	rs74851348	Outer Reverse	TCTCACTTTCTTCCTTCATT
				Inner Forward	GTTGGAGTGAAGGGTAGGG
				Inner Reverse	CCATCAACTCTAACCTATAATCA
12	<i>SLC38A4</i>			Outer Forward	GTTGTTTTGTTTGTAAATGTTGG
		Promoter	-	Outer Reverse	ATCTTTTTCTCTTTCCCTTCCA
				Inner Forward	GGTTGGGGGTTTTTTAGTG
				Inner Reverse	CCTAAACCCTCTAACCAAATTACC A
				Outer Forward	GTATTTGTGAGGTGTTTTTGAGG
		NM_001204 299	rs12823740	Outer Reverse	ACTCCAACAACAATAAAACCTAA
				Inner Forward	GGTTAGGGAGYGGGTTGGG
				Inner Reverse	TACCCRCTACCCAACCCT
12	<i>FAM101A</i>			Outer Forward	TGGTTGTTATTGGTTATTTT
		NM_001365 156	-	Outer Reverse	AAAAAACTTACTAAACAACCCC
				Inner Forward	GGGTAGGTAGAGGAGGAGGTA
				Inner Reverse	CCCGCCCTTCTACTCCCTAAC
				Outer Forward	GTTTAGGYGTTTAATTTGTTG
				Outer Reverse	ACAACACTACCAACACCAACAA
				Inner Forward	GTTTATGATTGTGTTTTTTG
				Inner Reverse	AACATAATACCAACCAACTA
				Forward	GTTGTTGTATTAGTTAGGTGAAGG
15	<i>SNURF</i>	Control gene	rs4906939	Reverse	AATAATATATATTCAACTTCTACT A
16	<i>SALL1</i>	-	rs11643654	Outer Forward	ATATTAGGGGTAAAGGGA

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
				Outer Reverse	ATACCACTCRAAATACCCA
				Inner Forward	TATTGTGTTTTAGTTTTAT
				Inner Reverse	CTAAACCCCRACAAAACCTC
17	<i>LOC33916</i> 6	-	rs12450161, rs12450165, rs12453225	Forward	GTTAAAAGGTAATTTGTAATTTGA GG
				Outer Reverse	ACTTCTCTAAACCAACCTCTCTAA AACTAA
				Inner Reverse	TAACTAAATTTAAAAAACCTAC
17	<i>PLXDC1</i>	-	rs188501857	Forward	GAGGTYGTAGTTTTTAGTTT
				Outer Reverse	AATCCTACCTACTCCRAACTAAAA
				Inner Reverse	AAACACCAACACCAAAAAACCAA
18	<i>ZFP516</i>	-	-	Forward	TAAGGTTTAAGGTTGTTGTAGTTT
				Outer Reverse	CTTAATATATTAATCCTACATC
				Inner Reverse	CTAAACACCCCAAAAACATTTAC C
19	<i>MBD3</i>	-	rs8104174	Outer Forward	GAGGGGATYGTAGGATTGGGTTTT
				Outer Reverse	ACCACCCCAACAACAAAATCAAA
				Inner Forward	TGTGATTATAGTTTATTGTAGT
				Inner Reverse	CACAACRACCCCAACCTTCCC
				Unmethylate d Inner Reverse	CAACAACCCCAACCTTCCCAACCA
19	<i>NUDT19</i>	-	rs8108621, rs8109823, rs61732600	Forward	GYGGGAGGTTTTTGAGGAGG
				Reverse	CAACAAAACCCTAAACAAAC
20	<i>ZFP64</i>	-	-	Outer Forward	TGTAAAGTAAGTTGTATTT
				Outer Reverse	AAATCTCCCRCAAACCACCC
				Inner Forward	TTTTAATTTGGTTTTGTAGT
				Inner Reverse	CCTAAAATTACAAATACAAAAAAC
X	<i>XIST</i>	-	rs41305409	Outer Forward	GGTTAGTATGGTGGTGGATATGT
				Outer Reverse	AAATTATACAACAATCCAACACTA TCC
				Inner Forward	GTAGGGATAATATGGTAG
				Inner Reverse	CACTATCCATCCCACCTTTTC

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')		
<b>Allelic RT-PCR</b>							
4	<i>JADE1</i>	NM_001287 441, NM_024900	rs13114904	Forward	AGGCTGGAGTGCAGTGGCGTGA		
				Reverse	CAAATAACTGCAACTCTCTGGGC		
		NM_001287 437	rs11933240	Forward	AGGAGAAGCATCTTGGCTTCTTGA		
				Reverse	CAAATAACACAAACTTCTCAC		
		New embryos	Very weak signal, rs1391181754	Forward	GCCTGGAGATGAAGACCATCTTAG		
				Reverse	TCCTCCCGGTTCTGCTCAAGGCT		
		New embryos	rs11933240	Forward	TGATGCGGAAGCCCTTGGGCT		
				Reverse	CAAATAACACAAACTTCTCAC		
		4	<i>GAB1</i>	All isoforms	rs1397529	Forward	AGATGAATTGTAGACTAGTAACA
						Reverse	GATAGTTTAGGCACATTTTCAGG
AK295684 nested RT- PCR (Out/Out)	rs62337524			Forward	TTAGAAGCCTGCCCCAGAGTCT		
				Outer Reverse	TGGATCTCCAGTTAAACGGCCACT		
AK295684 nested RT- PCR (Out/In)				Forward	TTAGAAGCCTGCCCCAGAGTCT		
				Inner Reverse	CTTCTCTGTACCTCTGACTTC		
New embryos	rs1269389517, rs1360288278			Forward	CCAGGAACATTTGATTTTTCC		
				Reverse	AAAGTCACCTATGGTTTGTGA		
New embryos	rs28924077			Forward	AAAGGACCTTTCTGACATAATC		
				Reverse	TTGGGAAAACCCAGAACAATG		
7	<i>PRKAG2</i>	NM_024429 nested RT - PCR (Out/Out)	rs8961	Outer Forward	TCATGCTGATCGCTGTCCTCCTCCT		
				Reverse	CAACATCACTGGAAGAAATAC		
		NM_024429 nested RT - PCR (In/Out)		Inner Forward	ACAAAAGGAGACAGAAACGGA		
				Reverse	CAACATCACTGGAAGAAATAC		
		Not nested		Inner Forward	AGGTTGTAATGCAGTGGCGCA		
				Outer Reverse	CAGATGTTCTAGGCTTCAATC		
10	<i>SFMBT2</i>	Nested RT - PCR (Out/Out)	rs10795530	Outer Forward	CTTGGCCAAGATATTTTCAGGAGC		
				Inner Reverse	GTGGCTCACACCTGTAATCCCAG		
		Nested RT - PCR (In/Out)		Inner Forward	AGGTTGTAATGCAGTGGCGCA		
				Inner Reverse	GTGGCTCACACCTGTAATCCCAG		
		LTR	rs719809	Outer Forward	ATTTAAGTAAGAAGTGTTAG		
				Outer Reverse	TATAAAAACTCTCCTCCTT		
				Inner Forward	TATAGAGTGGTTAGTTTAAAT		

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
				Inner	ATTTAAACTAACCACTCTATA
				Reverse	
10	<i>GSTO1</i>	-	rs4925	Forward	TCCTTGGTAGGAAGCTTTAT
				Reverse	CATAGAGATAGAATTGCCAC
		Placenta samples	rs1045278	Forward	GGAGATAAAGGAACTCAAAG
				Reverse	ATATGATTGTATGAAATTTGAAAA A
12	<i>SLC38A1</i>	New embryos	rs1045278	Forward	CATGTCCCTCCAAGATTTGAGATC
				Reverse	TGTATAATAAATAAACATTATTGT
		New embryos	rs3498, rs61923106	Forward	AGGAGGAGGGTGAAGGAGGGTGA
				Reverse	CTATGCAGCAGCATCCTTTTC
		New embryos	rs1938843414	Forward	CAAGTAAGGAATATTTAGAC
				Reverse	TTCTTCTCCCAGCTTCTGT
12	<i>SLC38A4</i>	-	rs2429467	Forward	CCTCGGGACACCCCACTCACAC
				Reverse	TGCTCATTGCTGCCTTTTCT
		All isoforms		Forward	GAGCATCAAGGTGAACCCGG
				Reverse	TCCGGAAAGTGCTCCTGGCAT
		NM_001365 156 nested RT – PCR (Out/Out)		Outer Forward 2	AGACATGGTGGGCCACCTGCA
				Outer Reverse 2	CTTGGGGAAGATGATGGTGGTGC
		NM_001365 156 nested RT – PCR (In/Out)	rs12318072	Inner Forward 2	TCTACTCCCTGGCGCCCGGCAT
12	<i>FAM101A</i>			Outer Reverse 2	CTTGGGGAAGATGATGGTGGTGC
		NM_181709 nested RT – PCR (Out/Out)		Outer Forward 3	GAAGCCTCTCAGCCGTAGGCG
				Outer Reverse 3	TCCGGAAAGTGCTCCTGGCAT
		NM_181709 nested RT – PCR (In/In)		Inner Forward 3	TGCAACTCTGAGGTCAAGT
				Inner Reverse 3	CTTGGGGAAGATGATGGTGGTGC
		AK308453 nested RT – PCR (Out/Out)	rs5011113	Outer Forward	GGATCTGCTGTGGAAGAACG
13	<i>STARD13</i>			Reverse	ATCCTTGGGAATATGCACCA
		AK308453 nested RT – PCR (In/Out)		Inner Forward	CTCCAGAATTCGCCGCCACC
				Reverse	ATCCTTGGGAATATGCACCA
14	<i>SMOC1</i>	-	rs3742909	Forward	ACTTGCTGCTGGTGTGGTGCA
				Reverse	CATCTACCTCGATGCACCACGC
16	<i>SALL1</i>	-	rs11645288	Forward	TGCCACATCCCAGTTCTGCT
				Reverse	CATGGGGCCATCCACAGAGAGC
17	<i>LOC33916 6</i>	-	rs12453225	Forward	GCTGGATTTGAGGAGCCTGCATG
				Reverse	AGGAGATGGCCAAAACACTGA
		Placenta samples	rs690353	Forward	GTCCAGGGGCGACGCGCCTTG
18	<i>ZNF516</i>			Reverse	TGTGCTCCAACCCAGGGCCGCT
		New embryos	rs72973711	Forward	TGTTAGGAATGTCAGGGACT
				Reverse	CAGCTCCAAGGCCAACTGCAC

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
		New embryos	rs2074488845	Forward	GGAAGGACGGCATTACATAG
				Reverse	TGGCCTGTGGTTGTTTCATCTGTTT
		AK001474 nested RT – PCR (Out/Out)		Outer Forward	GAGGCCTGGGTTTGGGGTCTG
				Reverse	GCCAGGAGCACGGCCTTCCTCCTG
19	<i>MBD3</i>	AK001474 nested RT – PCR (In/Out)	rs190802753	Inner Forward	AGGAGGAGGAGCCCGACCCGGA
				Reverse	GCCAGGAGCACGGCCTTCCTCCTG
		All isoforms		Forward 2	AGGAGGAGGAGCCCGACCCGGA
				Reverse	GCCAGGAGCACGGCCTTCCTCCTG
		All isoforms		Forward 3	AAAGCCTTCATGGTGACCGA
				Reverse	GCCAGGAGCACGGCCTTCCTCCTG
19	<i>NUDT19</i>	-	rs8108621, rs8109823, rs61732600	Forward	GCGCCGTGCGGGAGGCCTTTG
				Reverse	TCAGTTGCCTCTGATGGAGAT
20	<i>ZFP64</i>	-	rs3746413	Forward	CGATCCACACGGGGGACGCC
				Reverse	TCTAGAGCCTCAGTCTTAACCAT
<b>qRT-PCR SYBR™ Green</b>					
		AK295684 - test gene		Forward	GAAGTCAGAGGTACAGAGAAG
				Reverse	CGGCCACTGCGTAACACGAACC
		NM_002039		Forward 2	GAAAAAGTTGAAGCGTTATGCATG
4	<i>GAB1</i>	NM_207123 - test gene	-	Reverse	CGGCCACTGCGTAACACGAACC
		All isoforms; test gene		Forward 5	GAACCCAAACCTGTCCAGTGAAG
				Reverse 3	ATCATAGGGCTGCTTCCTCCATCA
7	<i>ACTB</i>	Endogenous control	-	Forward	CCGGCTTCGCGGGGCGACGAT
				Reverse	CTCCATGTTCGTCCTCCAGTTGG
7	<i>PEG10</i>	Canonical imprinting - control gene	-	Forward	GGACCCCATCCTTCCTGT
				Reverse	TTCAAACCCGCTTATTTTCG
10	<i>SFMBT2</i>	Test gene	-	Forward	CAGCAGAGGAAGGGGAGAAGTGC
				Reverse	ACGTCGGTGACCGTCCACTCCA
				Forward 2	GCTGTGCCATGCAGATTCTTTG
				Reverse 2	TGCTGCTTGTGATAATCTGCCAGT
12	<i>SLC38A1</i>	Test gene	-	Forward	CAAATCCCTGCATTGTTCCAGAGC
				Reverse	TGGCAAACAAATGCAAATGCAATGGT
12	<i>SLC38A4</i>	Test gene	-	Forward	TCATGGTTCGCTGGCAGT
				Reverse	GCAATAAGCACAGCTGCAATCAG
17	<i>RPL19</i>	Endogenous control	-	Forward	AATCGCCAATGCCAACTCCCCTCA
				Reverse	CCTATGCCCATGTGCCTGCCCTTC
18	<i>ZNF516</i>	Test gene	-	Forward	CGAAGACAGTGGTGAGGAGGG
				Reverse	AAGTCACCTCTTCGGAAAAGCA
19	<i>MBD3</i>	AK001474 - test gene	-	Forward	TTCCGCAGCAAGCCGCAGCT
				Reverse	GAAGATGGACGCCGTCTGGCG

Chr.	Gene	Comments	Variants	Oligo name	Sequence (5' -> 3')
		NM_001281 453, NM_001281 454 - test gene		Forward 3	AGCCGGGCGCAATGGAGCGGAA
				Reverse	GAAGATGGACGCCGTCTGGCG
19	<i>DNMT1</i>	Placenta imprint - control gene	-	Forward	ACCAAGAACGGCATCCTGTA
				Reverse	CACGGGACTGGACAGCTT
20	<i>ZFP64</i>	Test gene	-	Forward	CGATCCCACACGGGGGACGCC
				Reverse	AGCGGACATTGCAGAACTCGCA
X	<i>XIST</i>	Test gene	-	Forward	TTGCCCTACTAGCTCCTCGGAC
				Reverse	TTCTCCAGATAGCTGGCAACC
X	<i>XIST</i>	Test gene	rs1894271	Forward	TGAAGGACAGCATGGTTGGT
				Reverse	CACATGGAATGAGCAGTGTGC
<b>Mouse RT - PCR</b>					
			rs239651917, rs3401378677, rs260488938, rs212289214, rs237566955, rs582356198, rs264173787	Forward	AGAAGAAGTGGTCAGGACCATGG A
10	<i>Mbd3</i>	3' UTR		Reverse	CTGAAGGCAGTCTGCAGCCCAGAC
			rs239670010, rs252459742, rs253068355, rs217960000	Forward	ACGCCTGCGCAGACGAGCCCCA
		5' UTR		Reverse	CATAGCGCACACGCTGGCGACTC
<b>Mouse Bisulphite PCR</b>					
			rs585492322, rs233179056, rs250957823, rs8256341	Forward	GTTTAGTTAGAGTTGAATGGTG
10	<i>Mbd3</i>	-		Inner Reverse	TCCATAAACCTCAACACCTT
				Outer Reverse	CAACAAACCACAACCTAACAC
<b>Cloning</b>					
				Outer Forward	GATGGTGCTGCAAGGCGATTAAGT TG
				Outer Reverse	ATGTTGTGTGGAATTGTGAGCGGA
pGEM®-T Easy Vector	-	-		Sp6	ATTTAGGTGACACTATAG
				MF13 Forward	GTAACACGACGGCCAG
				Seq(S) T7	TAATACGACTCACTATAGGG