The role of NLRP7 in maternal imprinting and early embryonic development

Huihan Zhi

Thesis submitted towards the fulfilment of the degree of

Doctor of Philosophy

Institute of Reproductive and Developmental Biology

Department of Metabolism, Digestion and Reproduction

Faculty of Medicine

Imperial College London

October 2020

Abstract

NLRP7 is a maternal effect gene in humans whose mutations are responsible for a rare abnormal pregnancy with familial predisposition, biparental complete hydatidiform mole (BiCHM). Given BiCHM exhibits a similar pathology to androgenetic CHM, including overgrowth of trophoblast tissues and dysregulated maternal imprinting, NLRP7 is thought to play an important role in meditating maternal imprinting and early embryonic development in humans. However, due to the lack of rodent models and limited human materials, it has not been possible to reveal the mechanistic function of NLRP7 in maternal imprinting. In this project, hESCs are used as a model system, to investigate the role of NLRP7 in maternal imprinting and its potential relationship with other imprinting modulators, particularly DNMT3L.

Firstly, SNPs were identified in selected imprinted genes in hESCs, so that it was feasible to distinguish their two alleles in imprinting studies. In addition, naïve conversion of hESC was achieved to mimic, at least partly, the process of global DNA demethylation, which would enable them to be used to investigate the possible role of NLRP7 in maintaining maternal imprinting during this process. When NLRP7 was overexpressed in hESCs, it showed partial protection of maternal imprinting from global demethylation after naïve conversion, indicating that NLRP7 may be involved in protecting maternally imprinted genes from global demethylation. Furthermore, by analysing gene expression profiles of imprinting related genes, the expression of NLRP7 and DNMT3L were found to be correlated during early embryonic development and the naïve conversion process; and overexpression of NLRP7 can upregulate DNMT3L mRNA expression, which implies a potential link between them. Consequently, my findings demonstrate the possibility of employing hESCs as an alternative and unlimited resource to study NLRP7 in future. More importantly, this is the first time NLRP7 has been shown to directly protect maternally imprinted genes from global demethylation.

Statement of Originality

All the experiments in this thesis were carried out by myself, Huihan Zhi, unless otherwise stated in the text.

Copyright Declarations

Acknowledgements

First and foremost, I would like to express my sincere gratefulness to my supervisors, Dr. Wei Cui and Dr. Rosemary Fisher. Their invaluable experience, immense knowledge and great passion for science guided me throughout the four-year of my PhD. I would appreciate and treasure the skills and knowledge that they passed unto me.

I also would like to thank all the members of our lab, past and present, Dr. Shuchen Zhang, Dr. Siti Aminah Muhammad Imran, Dr. Jen Chiu, Manjari Triverdi, Xiaolun Ma, Li Tong, Faiza Batool and Faiza Farhan. Their help and encouragement make this difficult journey full of joy and motivation. My gratitude also goes to Prof. Simak Ali, Dr. Luca Magnani, Dr. Veronique Azuara and Dr. Nick Dibb for their valuable insights and suggestions in my early and late stage assessments.

Words could never describe my miss and love to my best friends in my home town, with all their care, accompany and encouragement in the past 15 years.

Finally, this thesis is dedicated to my beloved family. Thank you, my dearest parents Mr. Hongfa Zhi, Ms. Hua Wang and my husband Mr. Yu Jiang for your support to me all the way.

Table of contents

List of Tables and Figures

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Abbreviations

2iL	Cocktail of MEKi, GSK3i and LIF
5acC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5iLAF	Cocktail of MEKi, GSK3i, SRCi, bRAFi, ROCKi with LIF, Activin and FGF2
5mC	5-methylcytosine
APS	Ammonium persulfate
ART	Assisted reproductive technologies
b-act	Beta-actin
BCA	Bbicinchoninic acid
BER	Base excision repair
BiCHM	Biparental complete hydatidiform mole
BMP4	Bone morphogenetic protein 4
BSA	Bovine serum albumin
CAPS	Cryopyrin-associated periodic syndrome
CARD	Caspase activation and recruitment domain
cDNA	Complementary DNA
CHM	Complete hydatidiform mole
CM	Conditioned medium
Co-IP	Co-immunoprecipitation
CPL	Cytoplasmic lattices
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco's Modified Eagle Media
DMR	Differentially methylated region
DMSO	Dimethyl Sulfoxide
DNMT	DNA methyltransferase
DPBS	Dulbecco’s Phosphate Buffered Saline
ECM	Extra cellular matrix
EDTA	Ethylenediaminetetraacetic acid
EPI	Epiblast
EpiSC	Epiblast stem cell
ER	Endoplasmic reticulum
ESC	Embryonic stem cell
FGF2/bFGF	Fibroblast growth factor 2
FIIND	Domain with function to find
FLOPED	Factor located in oocytes permitting embryonic development

gDNA	Genomic DNA
H3K4	Histone 3 Lysine 4
H3K9	Histone 3 Lysine 9
HEK293T	Human embryonic kidney cell line
hESC	Human embryonic stem cell
hPSC	Human pluripotent stem cells
ICM	Inner cell mass
ICR	Imprinting control region
IG-DMR	Intergenic DMR
Igf2r	Insulin-like growth factor 2 receptor
iMEF	Irradiated-MEF
iPSC	Induced pluripotent stem cell
IVF	In vitro fertilization
KHDC3	KH domain-containing protein 3
KHDC3L	KH Domain Containing 3-Like
KSR	Knockout Serum Replacement
KSR-M	Knockout Serum Replacement medium
LB	Lysogeny broth
LIF	Leukaemia inhibitory factor
lncRNA	Long non-coding RNA
LRR	Leucine rich repeats
LTR	Long terminal repeats
MAF	Minor allele frequency
MATER	Maternal antigen that embryos require
MEF	Mouse embryonic fibroblasts
mEpiSC	Mouse epiblast stem cell
mESC	Mouse embryonic stem cell
MLID	Multi-locus imprinting disturbance
mRNA	messenger RNA
MUSCLE	Multiple sequence alignment tool
NACHT	Nucleotide-binding oligomerization domain
NAD	NACHT-associated domain
NHSM	Naïve Human embryonic Stem cell Media
NLR	Nucleotide-binding domain and leucine-rich repeat-containing receptor
NLRP	NACHT, LRR, and PYD containing protein
NLRP7-OE	NLRP7-overexpressed
NS	non-significant
OOEP	Oocyte-expressed protein

PADI6	Peptidyl arginine deiminase 6
PAMP	Pathogen associated molecular pattern
PCR	Polymerase chain reaction
PE	Primitive endoderm
Pen/Strep	Penicillin-Streptomycin
PGC	Primordial germ cells
PMSF	Phenylmethanesulfonyl Fluoride
PSC	Pluripotent stem cell
PVDF	Polyvinylidene fluoride
PYD	Pyrin domain
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
RIPA	Radio immunoprecipitation assay
RNA-seq	RNA-sequencing
rNLRP	Reproductive NLRP
ROS	Reactive oxygen species
RPKM	Reads Per Kilobase of transcript per Million mapped reads
RT	Reverse Toggle media
RT-PCR	Reverse transcription polymerase chain reaction
SCMC	Subcortical maternal complex
SDS	Sodium Dodecyl Sulfate
shRNA	Short-hairpin RNA
SNP	Single nucleotide polymorphisms
STAT3	Signal transducer and activator of transcription 3
t2iLGö	Cocktail of MEKi, GSK3i and PKCi with LIF
TAE	Tris-Acetate-EDTA Buffer
TBST	Tris Buffered Saline with Tween-20
TDG	Thymine-DNA glycosylase
TE	Trophectoderm
Tem	T-associated maternal effect
TEMED	Tetramethylethylenediamine
TET	Ten-eleven translocation methylcytosine dioxygenase
TLE6	Transducin-Like Enhancer of Split 6
Trim28	Tripartite motif-containing 28
UPD	Uniparental disomy
WT	Wild-type
ZGA	Zygotic genome activation
Zfp57	Zinc-finger protein 57

Chapter 1.

Introduction

Genomic imprinting is an epigenetic phenomenon whereby a subset of genes are activated or silenced in a parental-origin-specific manner (Reik and Walter, 2001). Most of the imprinted genes in mammals are involved in regulating the development of embryo and placenta, while a small group of them are associated with post-natal development (Tycko and Morison, 2002). Therefore, dysregulated genomic imprinting can lead to aberrant embryonic development, diseases and infertility. For example, complete hydatidiform mole (CHM), an abnormal human pregnancy characterised by excessive placental tissues and absent foetal development, is thought to be related to abnormal genomic imprinting during early embryogenesis resulting from the complete absence of the maternal genome and doubling of the paternal genome (Fisher and Hodges, 2003). However, there are some unique cases of familial recurrent CHM, that have both paternal and maternal genomes, but have widespread loss of maternal imprinting and intact paternal imprinting, known as biparental CHM (BiCHM) (Sanchez-Delgado et al., 2015).

NLRP7 is the major causative gene of BiCHM and is also the first maternal effect gene identified in humans (Slim and Wallace, 2013). NLRP7 protein is highly expressed in oocytes, fertilized eggs and preimplantation human embryos, which indicates that NLRP7 may have a crucial effect on embryo development (Wang et al, 2009; Akoury et al., 2015). However, little is known of the mechanisms by which NLRP7 affects maternal imprinting and the functions of NLRP7 in early embryogenesis. Since common laboratory rodent models do not have this gene (Tian et al., 2009), human embryonic stem cells (hESCs) derived from early human embryos might be the most closely associated model available to study the role of NLRP7 in this project. Especially, different methods have been developed in the last few years to convert routinely cultured hypermethylated primed hESCs to a naïve state with hypomethylated genome (Theunissen et al., 2014), which may provide an alternative model for the study of global DNA demethylation during early embryogenesis.

This chapter will provide updated information in all these aspects in detail to highlight

the importance and timely manner of this project.

1.1 NLRP7

1.1.1 NLRP7 and the NLRP protein family

NLRP7, also known as NALP7 and PYPAF3, is a member of the NLRP protein family that is characterised by containing a nucleotide-binding oligomerization domain (NACHT), a cluster of leucine rich repeats (LRR), and a pyrin domain (PYD) (Figure 1.1). The NLRP protein family consists of 14 members in humans (NLRP1-14), the most well characterised being NLRP1, NLRP3 and NLRP6 that are conserved in all mammals and play critical roles in innate immunity and inflammation (Tschopp et al., 2003). However, studies in the last decade have shown that NLRP family proteins can actually be divided into two subgroups, one subgroup of NLRP proteins (NLRP1, 3, 6, 10 and 12) are conserved, ubiquitously expressed in almost all tissues and their functions are mainly in immune responses, whilst the other subgroup of NLRP proteins, including NLRP 2, 4, 5, 7, 8, 9, 11, 13 and 14, are predominantly expressed in reproductive organs and early embryos, called reproductive NLRPs (rNLRP). The rNLRP proteins are evolutionally less conserved and are all present in humans. Thus, their functions are thought to be related to human reproduction and early embryonic development (Tian et al., 2009).

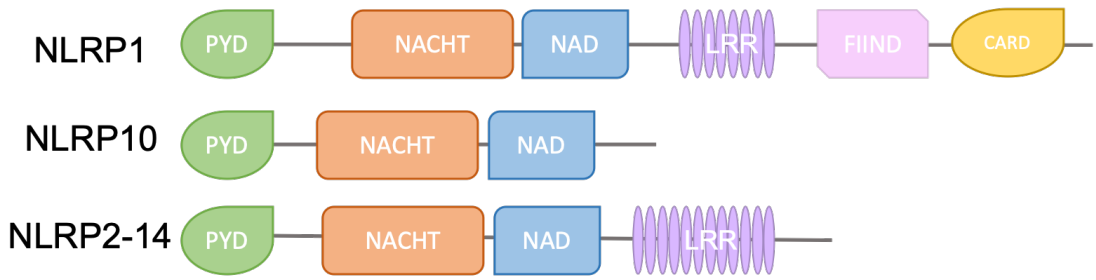

Figure 1.1 Structure of NLRP family members in humans.
All NLRPs contain a PYD domain (green), a NACHT domain (orange), followed by a NAD domain (blue) and an LRR region (purple), except NLRP10 lacks the LRR region and NLRP1 possesses an extra FIIND (pink) and CARD domain (yellow).

i. Structure of NLRP proteins

Most of the human NLRPs contain an N-terminal PYD, a central NACHT/NAD (NACHT-associated domain) and a C-terminal LRR, except that NLRP1 has an additional FIIND (domain with function to find) and a CARD (caspase activation and recruitment domain), while NLRP10 does not contain a LRR (Fig 1.1) (Schroder and Tschopp, 2010). All the rNLRPs are composed of the three key components: PYD, NACHT and LRR domains. PYD is an effector domain that is crucial for inflammasome formation and downstream signal transduction (Schroder and Tschopp, 2010; MacDonald et al., 2013). Different NLRPs may have some structural variations on PYDs, such as the length of α-helices that are essential for its folding structure and the charged surface residues that allow interactions with different protein partners (MacDonald et al., 2013); NACHT/NAD is an evolutionarily conserved domain essential for ATP binding and hydrolysis, which may result in the conformational change of NLRP proteins and affect their oligomerization (MacDonald et al., 2013; Singer et al., 2014; Proell et al., 2008); LRR domains are highly diverged in different NLRP proteins, and are mainly responsible for pathogen associated molecular pattern (PAMP) recognition, providing a horseshoe-shaped binding scaffold for protein-protein interaction and autoregulation (Ye and Ting, 2008; Proell et al., 2008).

ii. Divergence of NLRP genes

Innate immunity is an ancestral system to defend against microbial infection, thus most inflammation related NLRPs are expressed in various tissues and conserved from *C. elegans*, through rat and mouse to humans (Zhang et al., 2008). However, several rNLRPs have highly diverged between primates and lower species. All 14 human *NLRP* gene homologues can be identified in rhesus macaque monkeys, with 90% amino acid identity. Among them, *NLRP11* is a primate-specific gene expanded from *NLRP4* or *NLRP9*, *NLRP4* only exists in rodents and primates (Tian et al., 2009), while, in rodents, such as mouse and rat, *NLRP7, 8, 11* and*13* are absent from their genome (McDaniel and Wu, 2009; Tian et al., 2009). Human *NLRP7* is thought to have diverged from the duplication of *NLRP2*, thus *Nlrp2* is the closest mouse gene to human *NLRP7* in terms of protein homology (Duenez-Guzman and Haig, 2014; Slim and Wallace, 2013) (Fig. 1.2).

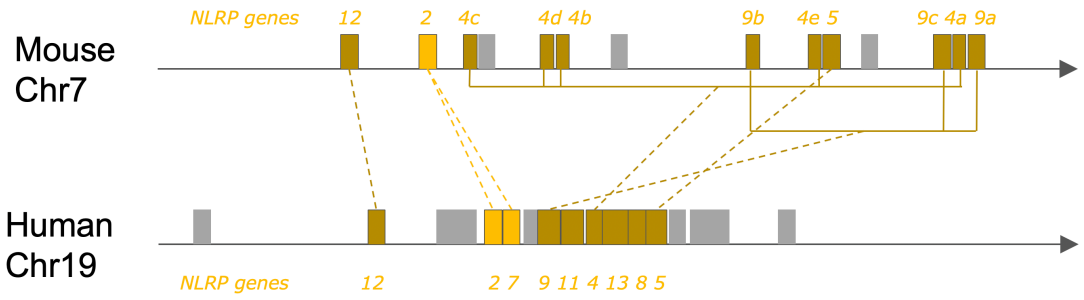

Figure 1.2 Syntenic comparison of NLRP genes in Human (chromosome 9) and mouse (chromosome 7). Human chromosome 9 and mouse chromosome 7 contain some lineage-specific duplications of NLRP genes, such as *NLRP7* and *NLRP11* in humans and *Nlrp4a/b/c/d/e* and *Nlrp9a/b/c* in mouse. The coloured bars represent NLRP genes and the orthologues are indicated by dashed lines (figure adapted from Tian et al., 2009).

iii. Functions of NLRP proteins

As a subfamily of the nucleotide-binding domain and leucine-rich repeat-containing

receptors (NLRs) superfamily that regulate antibacterial inflammation, the NLRP protein family is mainly involved in the formation of inflammasome complexes that mediate the activation of caspase-1 and induce cell apoptosis (Velloso et al., 2019). NLRP1 and NLRP3 are the most well characterized members that participate in the assembly of inflammasome complexes, while NLRP2, 6, 7 and 12 are less well characterized inflammasome components (Velloso et al., 2019). NLRP1 was the first member of the NLRP family confirmed to form an inflammasome (Martinon et al., 2002). NLRP1 deficiency in human is associated with vitiligo-related autoimmune disease (Jin et al., 2007). NLRP3 inflammasomes can be stimulated in response to pathogens, ion fluxes and reactive oxygen species (ROS) (Latz et al., 2013). NLRP3 mutations can lead to a dominantly inherited syndrome, named cryopyrin-associated periodic syndrome (CAPS), which can cause a wide range of autoimmune diseases (Turunen et al., 2018). NLRP7 also has been shown to be involved in inflammasome assembly activated by bacterial infection and microbial acylated lipopeptides, thus regulating IL-1b secretion (Radian et al., 2015; Messaed et al., 2011).

However, increasing studies have revealed that a subset of NLRPs, including NLRP7, are highly expressed in mammalian gonads and oocytes, and are also important to reproductive biology. Phylogenetic studies showed these *rNLRP* genes, comprising *NLRP 2, 4, 5, 7, 8, 9, 11, 13* and *14*, are less evolutionally conserved than other NLRPs and may have important functions in oogenesis and early embryogenesis (Tian et al., 2009). Maternal deficiency of some *rNLRPs*, such as *Nlrp2* and *Nlrp5*, in mouse can lead to embryo lethality (Peng et al., 2012; Tong et al., 2000). Similarly, several *NLRP2* and *NLRP5* mutations in human have been found to be responsible for female infertility due to early embryonic arrest (Mu et al., 2019). In addition, *NLRP2* mutations can be responsible for a congenital foetal overgrowth syndrome, Beckwith-Wiedemann syndrome (Meyer et al., 2009); *NLRP5* mutations can cause offspring with multi-locus imprinting disturbance (MLID) (Docherty et al., 2015). Particularly, biallelic mutations of *NLRP7* in women can lead to a rare autosomal recessive condition (Murdoch et al., 2006), which will be discussed in detail in the next section.

1.1.2 NLRP7 and hydatidiform moles

Molar pregnancy, also known as a hydatidiform mole, is a type of abnormal human pregnancy characterised by excessive trophoblastic/placental tissues with partial or completely absent foetal development, and in the latter case, is called complete hydatidiform mole (CHM). The incidence of CHM in UK is about 1 in 1,400 pregnancies, with a much higher prevalence in Asia (Savage et al., 2013). Most CHM are non-hereditary and androgenetic in origin with a diploid genome that is entirely paternally derived, caused by a single sperm which duplicates, or two sperm, fertilizing an egg without its maternal genome, and consequently show aberrant imprinting (Fisher and Hodges, 2003) (Fig 1.3). This loss of the maternal genome and doubling of the paternal genome results in the loss of expression of the maternally transcribed genes and excessive expression of paternally transcribed genes after fertilization, thus leading to loss of fetal development and trophoblastic hyperplasia of the placenta. However, there are a rare group of patients with CHM who exhibit recurrent molar pregnancies in whom the molar tissues contain both maternal and paternal genomes, suggesting a normal sperm-oocyte fertilisation. Furthermore, this type of CHM, also called biparental CHM (BiCHM), shows familial association, indicating a hereditary genetic defect might be involved (Fisher and Hodges, 2003).

Indeed, genetic linkage analysis in these families identified a causative gene, *NLRP7*, located at chromosome 19q13.3-13.4 (Moglabey et al., 1999; Murdoch et al., 2006). Although women who carry homozygous or compound heterozygous mutations in their *NLRP7* genes can develop normally without any phenotype in their somatic cells, they are rarely able to have normal pregnancies but instead have recurrent BiCHM. By contrast, biallelic mutations of NLRP7 in men neither show any abnormal phenotype nor do they have any effect on their offspring (Van den Veyver and Al-Hussaini, 2006; Qian et al., 2007). This indicates that *NLRP7*-associated BiCHM is a maternal effect, which refers to a phenomenon where the phenotype is determined by the mother's

genotype (Hager et al., 2008). *NLRP7* is the first maternal effect gene identified in humans (Slim and Wallace, 2013).

However, the mechanisms by which *NLRP7* mutations induce this abnormal pregnancy remains largely ambiguous. It was first proposed that mutations in *NLRP7* of pregnant women might result in an abnormal immune response of the mother toward the embryo, which then leads to a CHM. This hypothesis is based on the facts that several members of the NLRP family play crucial roles in inflammation and the immune response (Murdoch et al., 2006) and that NLRP7 was considered to be a feedback regulator of IL-1b secretion (Kinoshita et al., 2005). However, a woman who has biallelic mutations in NLRP7 and a previous history of recurrent BiCHM was able to successfully give birth to a healthy baby with the donation of an oocyte from a woman having normal NLRP7. This suggests that BiCHM may be due to an abnormal oocyte, rather than maternal immune rejection (Fisher et al., 2011).

Given that BiCHM exhibit similar pathology to androgenetic CHM, in which imprinting is dysregulated, it has been proposed that NLRP7 may function in the regulation of maternal imprinting. In the absence of normal NLRP7, the oocytes or their fertilised embryos lose maternal imprinting, leading to an androgenetic-like phenotype (Van den Veyver and Al-Hussaini, 2006). Indeed, a recent genome-wide methylation study of molar tissues collected from patients with *NLRP7* mutations confirmed that the maternally imprinted genes lost methylation on the maternal allele, leading to their expression from both alleles, while the paternally imprinted genes were unaffected (Fig 1.3) (Sanchez-Delgado et al., 2015). Therefore, loss of maternal imprinting resulting from *NLRP7* mutations is more likely to account for BiCHM, than abnormal host immunity. However, it is essential to provide direct evidence of the role of NLRP7 in establishing/maintaining normal maternal imprinting.

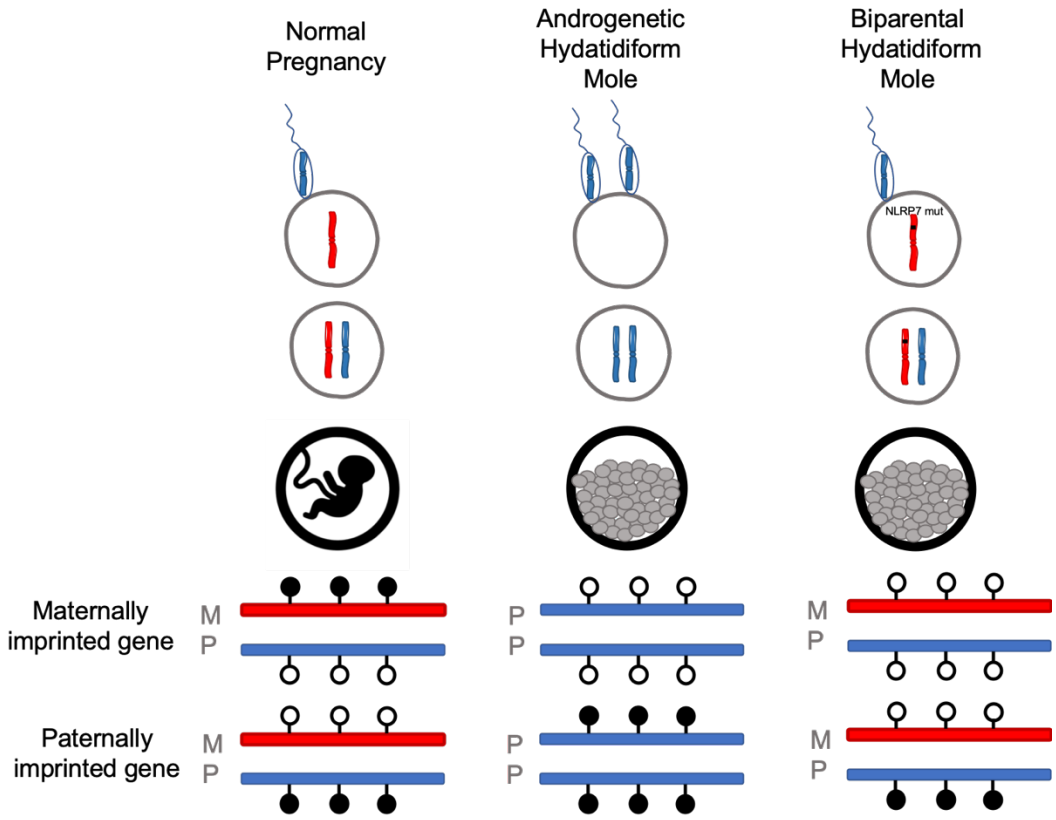

Figure 1.3 Methylation status of imprinted genes in normal pregnancy, androgenetic CHM and BiCHM.

Androgenetic CHM and BiCHM result from different mechanisms. Normal embryos and molar tissues resulting from androgenetic HM and BiHM have different methylation profiles at differentially methylated regions (DMRs) of their maternally and paternally imprinted genes. 'M' denotes maternal alleles, while 'P' denotes paternal alleles. Open and closed circles represent unmethyalted and methylated CpGs respectively.

1.1.3 NLRP7, a potential member of the subcortical maternal complex (SCMC)

i. SCMC

The subcortical maternal complex (SCMC) is a protein complex composed of several maternal effect proteins that are distinctively expressed in mammalian early embryos and oocytes and play essential roles in zygote progression to the two-cell stage (Yu et al., 2014; Bebbere et al., 2016). The SCMC is characterised by its subcortical localization in oocytes and early cleavage-stage embryos, while expelled from cell-cell contacting regions in early embryos (Li et al., 2008a). The complex was first identified in mouse preimplantation embryos with four protein members confirmed: Mater/Nlrp5 (Maternal Antigen That Embryos Require), Filia/Khdc3 (KH domain-containing protein 3), Floped/Ooep (Factor Located in Oocytes Permitting Embryonic Development/ oocyte-expressed protein) and Tle6 (Transducin-Like Enhancer of Split 6) (Li et al., 2008a; Ohsugi et al., 2008). It was then reported that Padi6 (peptidyl arginine deiminase 6), another maternal-effect protein, is able to interact with Floped and Mater, two members of SCMC, thereby also a member of the complex (Li et al., 2010; Yurttas et al., 2008; Kim et al., 2010). Recently, more SCMC members have been identified in mouse, such as Nlrp2, Zbed3 and Nlrp4f (Mahadevan et al., 2017; Gao et al., 2018; Qin et al., 2019) (Fig 1.4). In human, although only four core members of SCMC have been clearly identified, NLRP5, KHDC3L (KH Domain Containing 3-Like), OOEP, and TLE6 (Zhu et al. 2015), their combined molecular weight is much less than that of the isolated SCMC, suggesting that more proteins are involved in this complex in humans which have yet to be identified. NLRP7, NLRP2 and PADI6 are all potential members, due to their similar subcortical localisation and the phenotype of offspring in women with mutations in these genes (Mahadevan et al., 2017; Akoury et al., 2015; Qian et al., 2018).

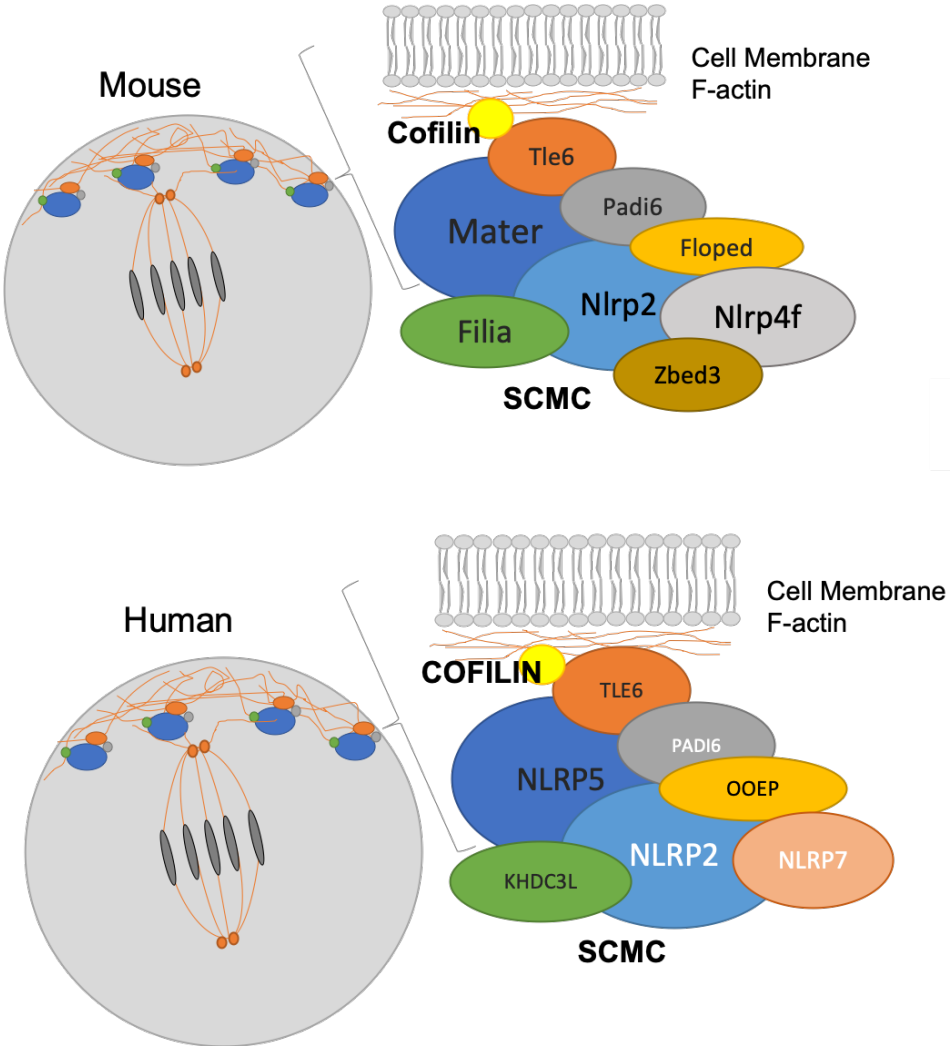

Figure 1.4 Comparison of mouse and human SCMC composition.
At least 7 members have been identified in the mouse SCMC (upper diagram), while only 4 members has been confirmed in the human SCMC (lower diagram). Confirmed SCMC members are shown in black lettering, potential members in white lettering. TLE6 is thought to link the SCMC to the F-actin network through a cytoskeleton regulator, Cofilin.

ii. Functions of the SCMC

Currently, most studies on the SCMC are performed in the mouse system due to limited availability of human materials. Mutations in any of the SCMC members in mouse exhibit similar phenotypes in the offspring, including asymmetric division of the zygote, embryo arrest at the two-cell stage, defective distribution of organelle and lack of cytoplasmic lattices (CPLs) in zygotes and early embryos (Lu et al., 2017; Mahadevan et al., 2017; Gao et al., 2018; Qin et al., 2019; Yurttas et al., 2008). Some mutations

also result in abnormal DNA methylation at imprinted loci (Mahadevan et al., 2017). All of these phenotypes indicate the importance of the SCMC in oocyte-to-embryo transition as well as in preimplantation development of the mouse embryo. Although it is far from clear how the SCMC regulates the zygotic function and early embryogenesis, studies have shown some progress. Firstly, the SCMC was thought to regulate spindle assembly via interaction with Cofilin, a major regulator of F-actin cytoskeleton formation. *Tle6*-muations in female mice can disturb the normal localization and activity of Cofilin, thus leading to asymmetric zygote cleavage (Yu et al., 2014). Secondly, the SCMC has been shown to have a close relationship with the CPLs, a microtubule network uniquely distributed in oocytes and preimplantation embryos, which is thought to stock maternal ribosomes and assist de novo protein synthesis (Yurttas et al., 2008). They share several common components with the SCMC, such as Mater, Floped and Padi6. Depletion of the SCMC can cause dissociation of CPLs (Tashiro et al., 2010; Kim et al., 2010; Yurttas et al., 2008; Qin et al., 2019). Thus, SCMC may also regulate the CPL network formation. Thirdly, the SCMC plays an essential role in organelle distribution during oocyte-to-embryo transition. Disorganization of mitochondria and/or endoplasmic reticulum (ER) was observed in mouse oocytes that are deficient for one of the following proteins: Mater-, Padi6-, Zbed3-, Nlrp2- and Nlrp4f (Kim et al., 2014; Fernandes et al., 2012; Qin et al., 2019; Gao et al., 2018; Mahadevan et al., 2017; Kan et al., 2011). Mitochondrial and ER abnormality can lead to increased oxidative stress and Ca^{2+} oscillation defects respectively in oocytes, and thus might affect developmental competence (Lu et al., 2017). Finally, the SCMC may be crucial for maintaining normal zygotic DNA methylation of imprinted regions. *Nlrp2* depletion in female mice not only affects their fertility, but also produces stillborn pups with DNA methylation abnormalities of imprinted loci (Mahadevan et al., 2017).

According to the limited number of clinical studies, mutation of SCMC members in humans is thought to cause early embryonic arrest, BiCHM and MLID (Qian et al., 2018; Parry et al., 2011; Murdoch et al., 2006; Docherty et al., 2015). In fact, NLRP7 is not the only gene associated with BiCHM. NLRP7 deficiency contributes to about

75% cases of BiCHM, while mutations in KHDC3L (also known as C6orf221, ECAT1), a member of the SCMC, leads to 5-10% of cases (Demond et al., 2019). PADI6 biallelic mutations have also been reported to be associated with embryonic lethality and recurrent pregnancy loss (Qian et al., 2018). Maternal NLRP2 deficiency can also lead to imprinting disorders, such as Beckwith-Wiedemann Syndrome, resulting from abnormal DNA methylation at the *KvDMR1* and *PEG1* imprinted regions (Meyer et al., 2009). In addition, women with NLRP5 mutation can experience recurrent miscarriages and live born babies with MLID (Docherty et al., 2015). All of these observations suggest that the SCMC may be involved in the establishment of genomic imprinting in the oocyte or the maintenance of imprinting in the post-zygotic embryo.

1.2 Genomic imprinting

Genomic imprinting is an inheritance process independent of Mendelian inheritance, which can control the expression of genes dependent on parent-of-origin, via *cis*-regulation through epigenetic mechanisms without changing the DNA sequence (Ferguson-Smith, 2011). Maternal imprinting occurs when the maternally inherited allele of a gene is suppressed, while the paternally inherited allele is expressed. In contrast, a paternally imprinted gene is one where only the paternally inherited allele is silenced (Hanin & Ferguson-Smith, 2020). Most of the imprinted genes in mammals are involved in controlling embryonic growth and development of placenta, while a small group of them are associated with post-natal development, such as suckling and metabolism (Piedrahita, 2011). 'Parental conflict hypothesis' is now a generally accepted hypothesis to explain the origin of imprinting: paternally expressed genes (maternally imprinted) often tend to promote the fitness of the offspring, while maternally expressed genes (paternally imprinted) often limit fetus size to conserve nutrients for the mother's own survival (Hanin & Ferguson-Smitdh, 2020). Therefore, the balance between paternal and maternal imprinted genes are essential for the wellbeing of mothers and the success of the fetus.

1.2.1 Importance of genomic imprinting

The first imprinted gene locus was identified in 1974, when Johnson discovered the viability of mice with deletions on *T-associated maternal effect* (*Tme*) locus is dependent on parent-of-origin. Maternal deletion of *Tme* is lethal while paternal deletion does not show any effect (Johnson, 1974). Further studies discovered that the *Tme* locus is a paternally imprinted locus, which contains an *Igf2r* (*insulin like growth factor 2 receptor*) coding gene that is normally expressed only from the maternal chromosome, to regulate fetal growth (Kalscheuer, 1993; Kaku et al., 2007). Furthermore, mouse embryos containing two sets of chromosomes entirely from one parent will fail to develop, indicating mouse development needs genetic materials from

both paternal and maternal sides (McGrath and Solter, 1984). Consequently, the correct establishment of genomic imprinting is essential for embryonic development.

In humans, abnormal imprinting not only leads to embryonic lethality but also causes many human diseases. The human imprinting-associated diseases are often caused by uniparental disomy (UPD), which refers to the inheritance of two copies of a chromosome or part-of a chromosome from only one parent (Engel and Antonarakis, 2004). For example, Prader-Willi syndrome results from the loss of paternally expressed genes on 15q11-13 (eg. *SNRPN* and *NDN*); while Angelman syndrome is caused by the loss of the maternally expressed gene, *UBE3A*, on the same region (Buiting, 2010). Silver-Russell syndrome, a congenital growth disorder, is another rare disease caused by imprinting deficiency. It results from the low expression of growth promoting *IGF2* and overexpressed growth suppressing *H19*, due to maternal UPD on chromosome 11 (Bartholdi et al., 2009). Apart from these rare syndromes, imprinting is also associated with some cancers. For example, Wilm's tumour, also called nephroblastoma, is a kidney cancer that often develops in children. It has been reported that nephroblastoma tissue can show reduced maternal expression of *H19*, thus leading to the overexpression of *IGF2* and uncontrolled cell proliferation (Dome and Coppes, 2002). Now with the application of modern technology, such as assisted reproductive technologies (ART), cases of different imprinting disorders are increasingly occurring in babies conceived via in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI) (Hattori et al., 2019). Many studies have suggested this may result from superovulation of oocytes, mechanical damage to germ cells caused by the IVF and ICSI procedure, in vitro culture of fertilized egg or lack of natural selection that should happen during natural conception (Hiura et al., 2014; Uyar & Seli, 2014). Since our understanding of the control and regulation of imprinting establishment and maintenance is far from clear, it is important to do more research in this area to prevent or identify suitable treatments for such diseases.

1.2.2 Molecular mechanism of Imprinting

i. Molecular control of imprinted genes

Imprinted genes are often located in clusters on chromosomes with each cluster containing several imprinted genes that can be regulated in an allele specific manner, via shared regulatory elements, called imprinting control regions (ICRs). These ICRs are normally marked by parental-specific DNA methylation, revealing different methylation patterns on paternal and maternal alleles, called differentially methylated regions (DMRs) (MacDonald and Mann, 2020). DNA methylation refers to the biological process by which methyl ($-CH_3$) groups are covalently added to cytosines of DNA at the C-5 position (Jin et al., 2011). Notably, DNA methylation often occurs on CpG dinucleotides in mammals, because it allows perfectly paired symmetrical methylation on both DNA strands (Greenberg and Bourc'his, 2019). There are more than 30000 CpG islands in the human genome. About 60% of them are associated with promotors/enhancers, while many others are located in gene bodies (Jeziorska et al., 2017; Li and Zhang, 2014). DNA methylation of promoters/enhancers often prevents the binding of transcription factors, thereby suppressing gene transcription, whereas high methylation levels on gene bodies is often associated with active expression even though the underlying mechanism is unclear (Moore et al., 2013).

Another common feature of imprinted gene clusters is that at least one of the imprinted genes in each cluster encodes a long noncoding RNA (lncRNA), which is usually parental-specifically regulated via differential methylation on its promoter and functions as a cis-acting regulator of other neighbouring imprinted genes (MacDonald and Mann, 2020). For example, paternally imprinted *H19,* the first identified lncRNA-coding gene in the mammalian genome, can regulate the expression of another maternally imprinted gene *IGF2* within the same imprinting locus (Brannan, 1990; Kaffer et al., 2001). H19 lncRNA is exclusively expressed from the maternal chromosome due to the hypomethylation of maternal DMR on its promoter region. The actively expressed H19 can act as a transcriptional insulator to suppress the activation of the *IGF2*

promoter. In contrast, *H19*-DMR is hypermethylated on the paternal chromosome, thereby blocking the insulator function of H19 and allowing IGF2 expression from the same chromosome (Kaffer et al., 2001).

ii. DNA methylation by DNA methyltransferases (DNMTs)

Differential DNA methylation is established in the paternal and maternal genome during gametogenesis and maintained throughout embryonic development (Barlow and Bartolomei, 2014). Acquisition of DNA methylation requires enzymes called DNA methyltransferases (DNMTs) that catalyse the transfer of methyl groups to DNA. There are 4 DNMTs (DNMT1, 3A, 3B and 3L) identified in mammals, which are grouped into two main categories: maintenance and *de novo* DNMTs (Li and Zhang, 2014).

Maintenance DNMT has only one protein, DNMT1, which acts with its functional partner UHRF1 during DNA replication to add a methyl group to the newly synthesised DNA strand when the template strand is methylated (Bronner et al., 2019). The maintenance process is essential for retaining the DNA methylation pattern on newly synthesized DNA after replication and cell division to confer cellular memory. Depletion of DNMT1 in mouse can cause loss of DNA methylation in the whole genome upon cell division, including imprinted genes, thus leading to embryonic death at E9.5 (Li et al., 1992; Li and Zhang, 2014)

The other group of *de novo* DNMTs contains the rest of the DNMT proteins. DNMT3A and 3B are mainly expressed during early embryonic development and can methylate hemimethylated or unmethylated cytosines in CpG islands to establish new DNA methylation patterns (Okano et al., 1999; Kato et al., 2007). DNMT3A deficiency in mouse can cause loss of imprinting in both male and female germ cells and post-natal death. DNMT3B deficiency has a more severe phenotype, in that mouse embryos die at E14.5, and causes demethylation of some centromeric satellite DNA (Kaneda et al., 2004; Kato et al., 2007). Double knockout of DNMT3A/3B leads to mouse embryo

lethality at E9.5 due to lack of de novo methylation (Okano et al., 1999). The DNMT3L protein shows high similarity to DNMT3A and DNMT3B but lacks the amino acid residues necessarily required for DNA cytosine methyltransferase activity. In mice, DNMT3L exhibits similar expression patterns as DNMT3A/3B and is considered to act as a cofactor for DNMT3A/B mediated DNA methylation (Suetake et al., 2004). Dnmt3l-knockout mice show normal development to adult. However, both male and female mice reveal defects in their reproduction (Hata et al., 2002). The male mice are infertile due to the failure in differentiation of spermatogonia to functional spermatocytes; while the female mice cannot produce any live pups as a result of embryonic lethality at mid-gestation caused by failed maternal imprinting (Hata et al., 2002). These results suggest that DNMT3L has important functions in reproductive and developmental biology and plays a critical role in the regulation and establishment of DNA methylation on maternally imprinted genes in oocytes or preimplantation embryos.

However, most of the research on DNMTs has been carried out in mice. Limited studies in hESCs show DNMT1 deletion can lead to global demethylation and rapid cell death; DNMT3A and DNMT3B single knockout only mildly affect genomic DNA methylation, but DNMT3A/3B double knockout causes progressive loss of CpG methylation with hESC propagation (Liao et al., 2015). Since imprinting is largely regulated by differential DNA methylation and the phenotype of DNMT3L-deficiency in female mouse is very similar to the loss of maternal imprinting in human embryos resulting from NLRP7-deficiency women, this project will try to explore the potential relationship between NLRP7 and DNMT3L.

iii. DNA demethylation via passive and active approaches

DNA methylation can also be reversed through both passive and active pathways. Passive DNA demethylation refers to replication dependent dilution due to the lack of maintenance DNMT1 (Rougier et al., 1998), whereas active demethylation is an enzymatic process to actively remove methyl groups from cytosine in CpG

dinucleotides. This process requires ten-eleven translocation methylcytosine dioxygenase (TETs) and three of such proteins have been identified, TET1, TET2 and TET3 (Wu and Zhang, 2017). All three TET proteins contain a common core catalytic region composed of a double stranded beta-helix domain, a cytosine-rich domain and cofactor binding sites for Fe(II) and 2-oxoglutarate. In addition, TET1 and TET3, but not TET2, also have a N-terminal CXXC zinc finger domain that can bind to CpG rich DNA regions (Pastor et al., 2013). All three TET proteins can oxidize 5-methylcytosine (5mC) to different forms of 5mC, such as 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Fig 1.5), which can then be removed via replication-dependent dilution as DNMT1 is unable to recognize oxidized methyl groups during DNA replication (Wu and Zhang, 2017; Valinluck and Sowers, 2007). Additionally, 5fC and 5caC can be recognized and removed by thymine-DNA glycosylase (TDG)-dependent base excision repair (BER) in the absence of DNA replication (Fig 1.5) (He et al., 2011; Wu and Zhang, 2017). *Tet1* deficiency in mouse is associated with abnormal hypermethylation of a number of ICR in germ cells and offspring (SanMiguel et al., 2018). Triple knockout of three Tets in mESC causes complete loss of global 5hmC and extensive hypermethylation predominantly on DNase I-hypersensitive sites, bivalent promotors and distal enhancers (Lu et al., 2014). Similarly, in hESCs, triple knockout of TETs causes locus-specific hypermethylation, particularly on bivalent promoters and enhancers. Gain of DNA methylation on these regulatory elements is catalysed by DNMT3B, which represses gene expression upon differentiation (Verma et al., 2018). Therefore, TETs are essential to maintain hypomethylation on gene regulatory elements and may affect cell differentiation during embryogenesis.

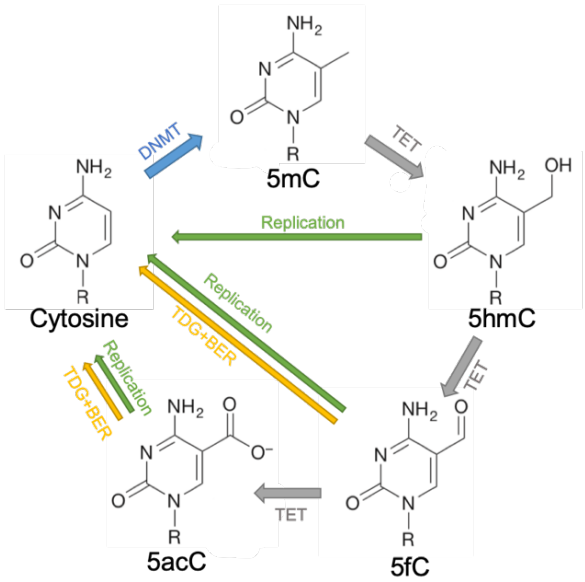

Figure 1.5 Active DNA demethylation mechanism.
Unmodified cytosines can be methylated to 5mC via DNA methyltransferases (DNMT; blue arrow); 5mC can be oxidized to different forms (5hmC, 5fC and 5acC) via ten-eleven translocation methylcytosine dioxygenase (TET; grey arrows), then the oxidized 5mCs can be converted back to unmethylated cytosines via replication-dependent dilution (green arrows), or like 5fC and 5acC can be directly cleaved and repaired via TDG-BER method (yellow arrows). (Figure adapted from Wu and Zhang, 2017)

iv. Histone modification in the regulation of DNA methylation and genomic imprinting

Increasing evidence has shown histone modifications also can modulate DNA methylation as they are able to interact with either DNMT directly (e.g. H3K4-DNMT3) or DNMT partners (e.g. H3K9me2-UHRF1) to affect DNA methylation, which has a critical role in development (Cedar and Bergman, 2009). Suv39h is a histone 3 lysine 9 (H3K9) methyltransferase and it has been shown that Suv39h-deficiency in mouse ESCs leads to the loss of DNA methylation on pericentric satellite repeats (Lehnertz et al., 2003); Moreover, H3K4 methyltransferase, SETD1, can be recruited to unmethylated CpG by binding to a CpG recognition protein CFP1, which in turn promotes H3K4-trimethylation at these CpG islands to reduce their accessibility to de novo DNMTs, thereby preventing them from being methylated (Li and Zhang, 2014; Thomson et al., 2010). In addition, histone modifications may also affect genomic imprinting. A recent study found that depletion of H3K4 demethylase KDM1B can

reduce maternally imprinted DMR methylation in mouse oocytes (Stewart et al., 2015). These studies suggest that factors that regulate histone modifications may function indirectly to control DNA methylation.

1.2.3 Establishment and maintenance of genomic imprinting during germ cell development and early embryogenesis

The global DNA methylation pattern is dynamic during mammalian embryonic development and has been intensively studied in mouse. In embryonic development, there are two waves of genome-wide DNA demethylation and each one is followed by a global de novo methylation (Fig 1.6). The first wave of DNA demethylation occurs in primordial germ cells (PGCs) after they first emerge and during their proliferation and migration to the genital ridge, the precursor of gonads. This demethylation results in an almost complete erasure of DNA methylation, including the original ICRs. In gonads, a de novo methylation process follows during gametogenesis. The second wave of global demethylation begins after fertilization and continues to the blastocysts stage. Then the global DNA methylation pattern is reestablished in embryos after implantation. In this second wave of global demethylation, imprinting ICRs are not affected (Smallwood and Kelsey, 2012).

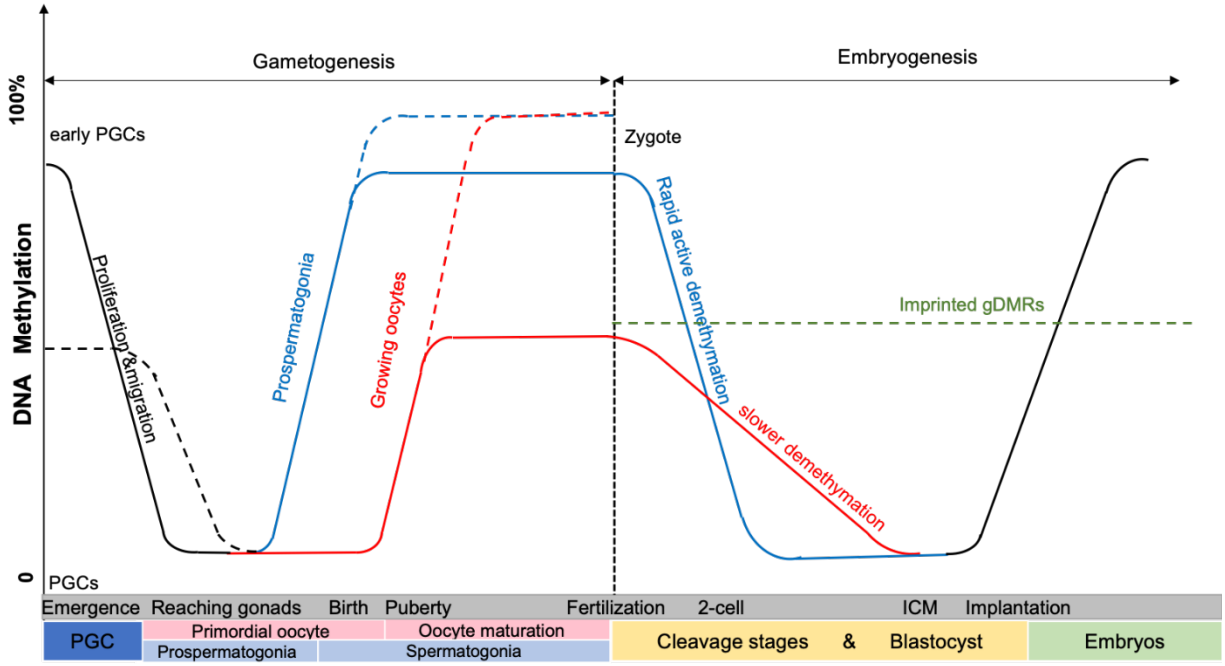

Figure 1.6 DNA methylation changes during human early embryonic development. Global demethylation during PGC migration and remethyaltion after implantation are shown in black. The differential *de novo* methylation and demethylation process between male and female are shown in blue and red respectively. Dashed lines represent methylation status of the ICRs. PGC, Primordial Germ Cell (Figure adapted from Monk et al., 2019).

Although only limited data are available from human resources, human DNA methylation dynamics show a large resemblance to that of mouse. In humans, the specification of PGC starts from developmental week 2. By week 4, DNA methylation level in PGC is already low compared with somatic cells (von Meyenn and Reik, 2015; Gkountela et al., 2015). Nevertheless, the imprinted methylation is mainly erased after the PGCs enter the gonads (Monk et al., 2019; White et al., 2016). A recent publication shows that genome-wide demethylation in human male PGCs is completed around 12 weeks of gestation with only about 7% of the methylation left (Guo et al., 2017b). However, in human female PGCs, methylation is initially erased rapidly, but then becomes progressively reduced to 17% until at least week-16 of gestation (Gkountela et al., 2015). The global demethylation of 5mC in humans was thought to be attributed to passive dilution along with the expansion of PGCs due to the reduced level of DNMT3A and UHRF1 although active demethylation may also be involved because

TET1 and BER members are highly expressed in hPGCs (White et al., 2016). However, it is worth noting that in mPGCs the abundant expression of Tet1 mainly facilitates the maintenance of the hypomethylated state of gonadal PGCs from de novo methylation, rather than contributes to the active global demethylation (Hill et al., 2018).

Following erasure, global DNA methylation is re-established asymmetrically in male and female germ cells. In male hPGCs, de novo methylation begins immediately in mitotically arrested prespermatogonia and is completed before birth, reaching ~ 80% in mature sperms (Wermann et al., 2010; Zhu et al., 2018). In contrast, *de novo* methylation of female hPGCs takes place after birth during the oocyte growth phase, until the germinal vesicle stage. MII oocytes harbour ~ 55% DNA methylation levels, but are globally less methylated than sperms (White et al., 2016; Zhu et al., 2018). In addition, the oocyte DNA methylation pattern features distinctive alternating hyper- and hypo-methylated regions, which are correlated to transcription-active and inactive regions, with transcription helping guide DNA methylation establishment in the oocytes (Veselovska et al., 2015; Pastor et al., 2016). In mouse, de novo methylation is known to be established by Dnmt3a/3l in oocytes, but Dnmt3a/3b in spermatogonia. However, DNMT3L is not detectable in human oocytes during oocyte development (Kelsey and Feil, 2013; Kato et al., 2007; Petrussa et al., 2014), suggesting a potential difference between mouse and human.

Following fertilization, another wave of DNA demethylation occurs, which plays an essential role in zygotic genome activation (ZGA) and contributes to the initiation of the early embryonic transcriptional network (Schulz and Harrison, 2019). In both humans and mice, paternal methylation is more rapidly erased by an active mechanism in a replication-independent manner (Gu et al., 2011; Lee et al., 2014; Zhu et al., 2015); while a large amount of maternal DNA methylation is lost through a passive replication-dependent dilution (White et al., 2016). According to the studies in mice, the passive demethylation of the maternal genome is due to less accessibility of Dnmt1 and Uhrf1 during replication and cell division (Hirasawa et al., 2008), even

though active demethylation is also reported to be involved in maternal demethylation (Wang et al., 2014). With the predominant passive demethylation in the early embryos, genome methylation reaches the lowest level of ~ 30% in human early blastocysts (Zhu et al., 2018). Additionally, a more detailed analysis of early human embryos found that there are two small but distinct gains of methylation occurring during preimplantation embryonic development, between the early- and mid-paternal pronuclear stage and between the 4-cell to 8-cell stages (Zhu et al., 2018). This study claimed the reprogramming of DNA methylome during early human embryonic development may be achieved by a dynamic balance between robust genome-wide demethylation and drastic focused remethylation. Nevertheless, differential methylation of imprinted regions is protected from the post-fertilization demethylation, resulting in parental-allele-specific methylation of these elements and consequent parental-allele-specific expression of imprinted genes (Guo et al., 2014). Mouse studies show the protection of imprinting may involve binding of proteins (eg. Zfp57 and Stella) to specific regions of the chromosome and recruitment of Dnmts and Uhrf1 to the nucleus (Li et al., 2008b; Hirasawa et al., 2008; Nakamura et al., 2012).

Concomitant with subsequent blastocyst implantation and cell-lineage determination, new methylation landscapes become established (Zhu et al., 2018), mainly via combined activity of DNMT3A/3B (Auclair et al., 2014).

1.3 Human pluripotent stem cells as a model to study genomic imprinting

1.3.1 Human pluripotent stem cells

Pluripotent stem cells (PSCs) are characterized by their ability to produce unlimited daughter cells equivalent to themselves, named indefinite self-renewal, and to generate all of the body's and some extraembryonic cell types, named pluripotency (Smith, 2006). The main PSCs include embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs). ESCs are derived from the inner cell mass (ICM) of preimplantation embryos, while iPSCs are generated from various somatic cell types by pluripotency factor-mediated reprogramming (Zhu and Huangfu, 2013). Owning to their unique properties, PSCs are now becoming a widely utilized resource to study mechanisms of embryonic development, modelling of different human diseases and discovering therapies. Given the absence of *NLRP7* in rodents and limited availability of human embryonic materials, hPSC are probably the closest cell model for the study of NLRP7 function in early human embryos. In the current project, hESCs have been used as a cell model for NLRP7 function.

i. Derivation and culture of hESCs

The first ESCs were derived in 1981 from the ICM of mouse preimplantation blastocysts, and required mitotically inactivated mouse embryonic fibroblasts (MEF) in serum containing medium (Evans and Kaufman, 1981). The MEF-feeders were then replaced by supplementation of culture medium with leukaemia inhibitory factor (LIF), a secreted cytokine that can stimulate signal transducer and activator of transcription 3 (STAT3), in combination with gelatin as extra-cellular matrix (ECM) (Smith, 2001). Thereafter, chemically defined medium supplemented with bone morphogenetic protein 4 (BMP4) and LIF was developed (Ying et al., 2003). However, these culture conditions cause a metastable state of cells with a large degree of heterogeneity, particularly regarding the expression of Nanog, a key pluripotent marker (Williams et al., 1988; Smith, 2013; Marks and Stunnenberg, 2014). This heterogeneity of Nanog

expression can be dramatically mitigated by replacing BMP and LIF with MEK and GSK3 inhibitors (Ying et al., 2008; Silva et al., 2008).

In 1998, the first human ESCs (hESCs) were successfully established from the ICM of human blastocysts using a similar method to that of mouse, which are cultured in serum-containing medium on MEF (Thomson et al., 1998). However, the morphology of hESCs is tightly packed flattened colonies, different from the domed colonies of mESCs, even though hESCs also express high levels of key pluripotent transcription factors (eg. OCT4 and SOX2) and cell surface markers (eg. SSEA4 and TRA-1-81) (Vazin and Freed, 2010; Thomson et al., 1998). Further optimisation of culture condition for hESCs replaced serum with knockout serum replacement and supplementation with FGF2 in the presence of MEF feeders (Amit et al., 2000). Soon, a feeder-free culture condition was developed by replacing MEF feeders with Matrigel and culturing hESCs in MEF-conditioned medium (CM) (Xu et al., 2001). After several years, a chemically defined and animal product-free medium, TeSR1 was developed (Ludwig et al., 2006). TeSR1 contains multiple key components, such as LiCl, γ-aminobutyric acid, pipecolic acid, FGF2 and TGFβ, to maintain hESC pluripotency (Ludwig et al., 2006).

It is apparent that mESCs and hESCs require different culture conditions: mESCs are dependent on LIF and BMP, while hESCs cultures requires activation of FGF2 and TGFβ/Activin signalling. Multiple factors are considered to account for these differences, such as developmental stage and species specificity (see next sections).

ii. Pluripotent status of hESCs

In the last two decades, considerable progress has been made to enrich our understanding of pluripotency and the relationship between PSCs and in vivo embryonic development. With the establishment of ESCs and EpiSCs in mouse, it became clear that pluripotent cells in vivo are not restricted to the ICM but include cells

from ICM of preimplantation blastocyst to epiblast of post-implantation embryos, although they are not exactly the same in gene expression patterns and epigenome profiles (Brons et al., 2007; Ying et al., 2008; Tesar et al., 2007). Moreover, ESCs derived from the ICM do not necessarily reflect the cells of ICM but depend on the culture conditions (Ying et al., 2008).

mESCs, particularly those cultured under 2iL condition, closely resemble the cells in the ICM of preimplantation blastocysts as they share similar gene expression and epigenomic profiles and can be incorporated into mouse blastocyst to form chimeras (Silva et al., 2008) (Table 1.1). As such, the mESCs cultured in 2iL medium were named initially as ground state then, more accepted, as naïve PSCs (Nichols and Smith, 2009). On the other hand, mEpiSCs derived from the post-implantation mouse egg cylinder epiblasts in culture medium containing FGF and Activin cannot yield chimeras when injected into preimplantation blastocysts even though they are capable of generating the three germ layers in vitro and form teratoma when injected into immune-deficient mice (Tesar et al., 2007; Brons et al., 2007). This is probably due to their more advanced developmental status, named primed state (Huang et al., 2012). Although both naïve and primed PSCs are able to generate the cells of the three germ layers and express high levels of key pluripotent factors Oct4 and Sox2, they also have their distinct gene expression and epigenomic characteristics. For example, naïve cells express some distinct pluripotency factors (eg. Nanog, Rex1, Klfs and Esrrβ), are characterized by genome-wide hypomethylation and retain two active X chromosomes in female cells. In contrast to naïve cells, primed PSCs start to express lineage commitment factors (eg. Otx2, Zic2 and Brachyury), although they remain undifferentiated, and they have upregulated global DNA methylation and one X chromosome inactivated in female cells (Weinberger et al., 2016). Table 1.1 summarizes the characteristics of naïve and primed pluripotent stem cells in mouse.

Table 1.1 Differences between naïve and primed pluripotent stem cells in mouse.

Properties	***Naïve state* (mESCs)**	***Primed state* (mEpiSC)**
Corresponding embryonic cells	Preimplantation blastocyst (ICM)	Post-implantation Epiblast
Culture conditions	2iL*	FGF2+Activin
Dominant OCT4 enhancer	Distal	Proximal
Pluripotency markers	Oct4, Sox2, Nanog, Rex1, Klf2/4, Esrrβ	Oct4, Sox2
Blastocyst chimaeras	Yes	No
XX status	XaXa	XaXi
Genome-wide DNA methylation	Hypomethylation	Hypermethylation
Colony morphology	Domed	Flattened
Metabolism	Oxidative Phosphorylation	Glycolysis

*2iL (dual inhibition of MEK and GSK3 signalling pathways with the supplement of LIF)

Interestingly, although hESCs are derived from the ICM rather than post-implantation epiblast, they require similar culture conditions to that of mEpiSCs and both cells exhibit similar flattened morphology of their colonies. Moreover, hESCs share several molecular and epigenetic characteristics with mEpiSCs, such as loss of pluripotency upon inhibition of the MEK-ERK pathway, usage of the OCT4 proximal enhancer and a hypermethylated global genome (De Los Angeles et al., 2012; Xiao et al., 2016; Weinberger et al., 2016). Given these similarities between hESCs and mEpiSCs, hESCs and hiPSCs are considered to be in a 'primed' state (De Los Angeles et al., 2012; Pastor et al., 2016). However, hESCs are not totally identical to mouse EpiSCs as several naïve markers (eg. Nanog, Rex1 And Prdm14) that are not expressed in mEpiSCs are positive in hESCs (Weinberger et al., 2016). Therefore, although hESCs are thought to be at primed state, their pluripotency could be mapped earlier than mEpiSCs on the developmentally pluripotent timeline, thereby between mESCs and mEpiSCs. The differences between species are supported by the study that compares gene expression profiles of mouse and monkey embryos at various developmental stages, as well as mESCs, monkey ESCs and mEpiSCs (Nakamura et al., 2016).

1.3.2 Conversion of hESC to the naïve state

It is generally accepted that pluripotency of hESCs is the primed state and following the successful conversion of mEpiSCs to naïve mESCs by altering culture conditions, many laboratories have been attracted to develop culture conditions that can convert hESCs to naïve hESCs. As a result of this, several protocols have been developed using transgene-dependent or independent methods (Hanna et al., 2010; Theunissen et al., 2014; Gafni et al., 2013; Chan et al., 2013).

Unlike mESCs, 2iL condition alone is not sufficient to either maintain hESCs or convert them to the naïve state, but with exogenous transgene expression of KLF2, KLF4 and OCT4, or NANOG and KLF2, it can induce hESC to naïve pluripotency. (Hanna et al., 2010; Theunissen et al., 2014). In addition, several transgene-independent conversion methods have also been established in the past 10 years (Theunissen et al., 2014; Chan et al., 2013; Gafni et al., 2013; Guo et al., 2017a; Ware et al., 2014), which are summarised in Table 1.2.

Table 1.2 Naïve conversion methods without transgenesis

Methods	*Basal Medium*	*Growth Factors/Inhibitors*	*References*
NHSM*	KO-DMEM	MEKi, GSK3i, p38i, JNKi, aPKCi, ROCKi, LIF, TGFβ or Activin A, FGF2	Gafni et al., Nature (2013)
3iL	TeSR1	MEKi, GSK3i, BMPi, LIF, FGF2, TGFβ	Chan et al., Cell stem cell (2013)
5iLAF	N2B27	MEKi, GSK3i, SRCi, bRAFi, ROCKi, LIF, Activin A, FGF2	Theunissen et al., Cell Stem Cell (2014)
RT**	DMEM/F12 +KSR	1. HDACi for 2 passages 2. MEKi, GSK3i, FGF2	Ware et al., PNAS (2014)
t2iLGö	N2B27	1. HDACi, MEKi, LIF for 3 days 2. WNTi, MEKi, PKCi, LIF for 6 days 3. MEKi, GSK3i, PKCi, LIF, ROCKi	Guo et al., Development (2017a)

*NHSM (Naïve Human embryonic Stem cell Media)
**RT (Reverse toggle media)

Naïve human embryonic stem cell media (NHSM) was the first published method, without genetic manipulation of hESCs, reported to be capable of converting hESCs to naïve hESCs in both feeder-containing and feeder-free conditions. NHSM contains 6 inhibitors and 3 growth factors (Table 1.2), which include not only 2iL, used in mESCs, but also inhibitors of p38, JNK, aPKC (atypical PKC) and ROCK together with low dosages of TGFβ/Activin A and FGF2. The resulting hESCs exhibit several features that are closely associated with the cells in the ICM, including mild decreases of genome-wide DNA methylation, increased oxidative phosphorylation, reactivation of X chromosomes and preservation of genomic imprinting (Gafni et al., 2013; Weinberger et al., 2016; Pastor et al., 2016). More importantly, these naïve hESCs are reported to have a capacity to generate interspecies blastocyst chimeras in mouse (Gafni et al., 2013). However, other groups have shown NHSM is not capable of inducing OCT4-distal enhancer and the global DNA methylation level of NHSM-converted hESCs is

still very similar to that of primed hESCs (Theunissen et al., 2014; Liu et al., 2017).

Another more studied protocol is 5iLAF medium that was developed by screening small molecules for the ability to activate the OCT4-distal enhancer (Theunissen et al., 2014). This medium contains similar components to that of NHSM except for replacing p38i, JUNKi and aPKCi with bRAFi and SRCi (Table 1.2). The hESCs resulting from this method reveal much higher upregulation of naïve markers and stronger reduction in global DNA methylation than that of NHSM-derived hESCs. However, they lose genomic imprinting (Theunissen et al., 2014; Pastor et al., 2016). Unfortunately, the attempt to generate interspecies chimera by Theunissen's group was not successful due to technical issues, whether using NHSM-converted cells provided by Hanna's group or 5iLAF-converted cells generated on their own (Theunissen et al., 2014).

While in most of the naïve conversion protocols, FGF2 or serum factors are indispensable, Smith's group developed a new method t2iLGö, which does not require FGF2 or Activin/TGFβ, but needs an initial treatment of HDACi together with MEKi and LIF, followed by a resetting step of WNTi, MEKi, PKCi and LIF treatment before the final stabilisation of reset naïve hPSCs (Table 1.2). The resulting naïve hPSCs produced by this protocol exhibit many naïve characteristics and largely resemble 5iLAF-converted naïve cells in terms of their transcriptome profile, DNA methylation profile, X-chromosome status and imprinting status (Guo et al., 2017a; Vallot et al., 2017).

Although multiple protocols have been developed for the naïve conversion of hESCs, it is difficult to verify whether the resulting cells are indeed naïve hESCs that resemble the cells in the human ICM as the gold-standard chimera formation assay is not applicable in humans for the ethical reasons. It is noticeable that both 5iLAF- and t2iLGö- converted naïve hESCs have been reported to lose their primary imprinting pattern, in contrast to that of the cells of the ICM (Guo et al., 2017a; Pastor et al., 2016). Interestingly, when naïve hESCs are converted back to the primed state, the imprinting

pattern is not reestablished (Pastor et al., 2016). This phenomenon has also been observed in naïve mESCs, which appears as progressive loss and is irreversible (Choi et al., 2017). These new findings inspired me to investigate whether hESCs could be used as a model system to explore the function of NLRP7 in the establishment and maintenance of human genomic imprinting.

1.4 Aims and Objectives

Since NLRP7 does not exist in lower mammals and hESCs are the most closely associated cell type to early human embryonic cells, I hypothesise that hESC may be a suitable model system to explore the role of NLRP7 in maternal imprinting during early human embryogenesis. The ultimate goal of this project is to investigate the role of NLRP7 in the establishment or/and maintenance of maternal imprinting during early human embryonic development, using hESCs as a model. In particular, I am interested in utilising the hESC prime-to-naïve conversion system that leads to considerable global demethylation as a model to explore the effect of NLRP7 on maternal imprinting. I aim to achieve the overall goal through:

1. Identification of single nucleotide polymorphisms (SNPs) in imprinted genes in H1 and H7 hESCs in order to investigate the role of NLRP7 in imprinting.

2. Conversion of primed hESC to the naïve state and validation of the changes in expression and DNA methylation profiles of the imprinted genes.

3. Investigation of the role of NLRP7 in the maintenance of human imprinting during the naïve conversion of hESCs.

4. Exploration of whether NLRP7 might affect maternal imprinting by modulating the function of *de novo* DNA methyltransferases, particularly DNMT3L.

Chapter 2.

Materials and Methods

2.1 Cell Culture

2.1.1 List of materials

Name	Supplier	Catalogue Number
Cell Culture Medium Components		
200 mM L-Glutamine (LG)	Sigma-Aldrich	G7513
100x N2 Supplement	Life Technologies	17502048
100x Non-Essential Amino Acids (NEAA)	Sigma-Aldrich	M7145
100x Penicillin-Streptomycin (Pen/Strep)	Sigma-Aldrich	P0781
30% Bovine Serum Albumin (BSA) Solution	Sigma-Aldrich	A9576
50x B27 Supplement	Life Technologies	17504044
DMEM/F12	Thermo Fisher	21331020
Dulbecco's Modified Eagle's Medium (DMEM)	Sigma-Aldrich	D0819
Heat-Inactivated Fetal Bovine Serum (FBS)	Sigma-Aldrich	10500064
Knockout DMEM	Life Technologies	10829018
Knockout Serum Replacement (KSR)	Life Technologies	10828028
Neurobasal medium	Life Technologies	21103-049
β-Mercaptoethanol For Tissue Culture (50 mM)	Thermo Fisher	31350-010
Growth Factors and Inhibitors		
Activin A	Peprotech	120-14E
Chloroquine Disulphate Cystalline	Sigma-Aldrich	C6628
Human Basic Fibroblast Growth Factor (bFGF)	Peprotech	100-18B
Human Leukemia Inhibitory Factor (hLIF)	Peprotech	300-05
IM-12 (GSK3i)	Sigma-Aldrich	SML0084
PD0325901 (MEKi)	Sigma-Aldrich	PZ0162
Puromycin	Sigma-Aldrich	P8833
SB590885 (BRAFi)	Bio-Teche	2650
WH-4-023 (SRCi)	Bio-Teche	5413
Y-27632 (ROCKi)	R&D	1254
Dissociation Enzymes		
Accutase	Sigma	A6964
Collagenase IV	Life Technologies	17104019
Trypsin-EDTA	Sigma	T3924

Other Cell Culture Associated Reagents		
2% Gelatin in PBS	Sigma-Aldrich	G1393
Calcium Phosphate Transfection Kit	Sigma-Aldrich	CAPHOS-1KT
Chondroitin Sulphate	Sigma-Aldrich	C4384
Dimethyl Sulphoxide (DMSO)	Sigma-Aldrich	D2650
Dulbecco's Phosphate Buffered Saline (DPBS)	Sigma-Aldrich	D8537
Hexadimethrine Bromide (Polybrene)	Sigma-Aldrich	H9268
Lipofectamine LTX with PLUS Reagent	Thermo Fisher	15338030
Matrigel® Matrix Growth Factor Reduced	SLS	354230
OptiMEM	Thermo Fisher	11058021

Cell Lines	
Name	**Notes**
H1 hESC	Male Hesc Line (WA-01) Distributed By Wicell
H7 hESC	Female Hesc Line (WA-07) Distributed By Wicell
HEK293T	Immortalized Human Embryonic Kidney Cell Line (Cat#ATCC® CRL-3216TM)

2.1.2 Stock Solutions and Medium

DMEM medium recipe

Reagent	Volume	Final Conc.
DMEM	500 ml	
FBS	50 ml	10%
100x Pen/Strep	5 ml	1X

KSR medium recipe (KSR-M)

Reagent	Volume	Final Conc.
Knockout™ DMEM	400 ml	80%
Knockout™ Serum Replacement (KSR)	100 ml	20%
200 mM L-Glutamine	2.5 ml	1 mM
100x Non-Essential Amino Acids	5 ml	1 x
50 mM ß-Mercaptoethanol	1 ml	0.1 mM
100x Pen/Strep	5 ml	1 x

N2B27 base medium recipe (Theunissen et al., 2014)

Components	Volume	Final Conc.
DMEM/F12	240 ml	48%
Neurobasal	240 ml	48%
100x N2 Supplement	5 ml	1 x
50x B27 Supplement	10 ml	1 x
200 mM L-Glutamine	5 ml	2 mM
Non-Essential Amino Acids	5 ml	1%
50mM ß-Mercaptoethanol	1 ml	0.1 mM
30%BSA	87 µl	50 µg/mL
100 x Pen/Strep	5 ml	1 x
KSR	2.5 ml	0.5%

5iLAF cocktail

Reagent Name	Solvent	Stock conc.	Working Conc.
PD0325901 (MEKi)	DMSO	1 mM	1 µM
IM-12 (GSK3i)	DMSO	1 mM	1 µM
SB590885 (bRAFi)	DMSO	1 mM	0.5 µM
WH-4-023 (SRCi)	DMSO	1 mM	1 µM
Y-27632 (ROCKi)	dH2O	10 mM	10 µM
Activin A	0.2% BSA in dH2O	10 µg/ml	20 ng/ml
rhLIF	0.2% BSA in dH2O	10 µg/ml	20 ng/ml
bFGF	0.2% BSA in dH2O	10 µg/ml	10 ng/ml

Dissociation enzymes

Reagent Name	Solvent	Final Conc.
Collagenase IV	Knockout DMEM	200 unit/ml

2.1.3 Methods

2.1.2.1 hESC culture and propagation

i. Generation of irradiated MEFs (iMEF) and MEF-Conditioned Medium (MEF-CM)

MEFs were prepared from the carcass of CD-1 mouse foetus at E13.5. Each embryo was washed in DPBS containing 2X Pen/Strep and then the viscera was removed. The remaining carcass was finely minced and trypsinised at 37 °C to obtain single cells. All the cells were transferred to a T75 flask with 15 ml MEF medium and grown for 2 to 3 days at 37 °C in a 5% CO2 incubator before they were frozen in MEF medium containing 10% DMSO at 1 x 10^7 cells/vial, defined as passage 0, and stored in liquid nitrogen tank until use or up to one year.

For preparation of MEF-CM, one vial of MEF was recovered into a T75 flask and split in 1:3-6 ratio when the cells got 90% confluency. After 3 - 4 passages, they were collected into 50ml tubes by trypsin dissociation and counted for cell number. MEFs were mitotically inactivated by irradiation at 40 grays and then plated into gelatin-precoated flasks at ~80,000 cells /cm^2 or 2 - 2.5 x 10^7 per T225 flask in 40 ml MEF medium. On the following day, MEF medium was replaced by 150ml KSR-M supplemented with 4 ng/ml bFGF, which was collected ~24 hours later, named MEF-CM. Following the collection, fresh 150 ml KSR-M was added into the flask for the collection on the following day. This cycle was repeated up to 7 days and all the collected MEF-CM were stored at -80 °C until use. Frozen CM was thawed in fridge one day before use. 1x Pen/Strep and 2 mM LG was added to CM before going through a Corning® Filter system (0.22 μm).

Mitotically inactivated iMEF can also be stored in liquid nitrogen for future use as feeder cells. In this case, 2.5 x 10^6 iMEFs were frozen in each vials.

ii. Matrigel preparation and coating

Matrigel® stock solution (10 ml) should be slowly thawed on ice at cold room overnight to avoid solidification, and then diluted in 10 ml cold Knockout™ DMEM to prepare 20 tubes of 1 ml aliquots, which should be stored at -20 ºC. To prepare the working solution, each aliquot of Matrigel should be defrosted on ice or at 4 ºC overnight and then diluted again in 17 ml cold Knockout™ DMEM. 1 ml diluted Matrigel could be used to coat each well of a 6-well plate. The coated plates were incubated overnight at 4 ºC or 3 - 4 hours at room temperature. Matrigel solution must be removed immediately before use.

iii. Culture and passage of hESC

hESCs (H1 and H7) were routinely cultured in Matrigel-coated plates in MEF-CM with 10 ng/ml bFGF as previously described (Gerrard et al., 2005). Medium was changed every day and hESCs were passaged every 5-7 days in a ratio of 1:3 after treatment with 200 unit/ml collagenase to remove spontaneously differentiated cells. The hESC colonies were mechanically cut into small clumps with 5-ml pipette in 1 ml MEF-CM and then transferred to a new Matrigel-coated plate.

In the routine culture condition, hESCs were passaged as small clumps, because they were likely to undergo apoptosis and differentiation when they were dissociated into single cells. To split them in single cells, hESCs were dissociated with Accutase at 37 ºC for 5 to 10 min and split at a ratio of 1:3 in MEF-CM media with the supplementation of 10 μM ROCKi and 10 ng/ml bFGF. The ROCKi was removed from the culture medium from the following day.

iv. Freezing hESCs

hESCs were treated with collagenase IV for about 5 to 10 min at 37 °C as for cell

passaging except colonies were scraped mechanically in 1 ml cold KSR. Small hESC clumps were then transferred to a labelled cryotube and 0.1 ml DMSO was added dropwise into the cryotube. The tubes were placed in a Mr. Frosty freezing container and transferred into a -80 °C freezer overnight. The frozen vials can then be transferred to liquid nitrogen tank for long-term storage.

2.1.2.2 Naïve conversion and propagation of naïve hESC

i. MEF feeder coating

Unlike the routinely cultured hESCs, all the naïve cell cultures need to grow on the iMEF-feeder layer. Firstly, 6-well plates were coated with 0.25% Gelatin at room temperature for 1 hour. Then, 1.5 x 10^5 – 2.5 x 10^5 iMEFs were seeded to each well of a 6-well plate. The cells were plated down evenly after overnight incubation.

ii. Naïve conversion and passaging

To achieve naïve conversion, primed hESCs were dissociated into single cells via Accutase-Rocki split. The cell suspension was passed through a 40 µm cell strainer (Corning 352340) and plated at a density of 1.5-2x10^5 cells in each well of a 6-well plate in the MEF-CM. the medium was changed on the following day using MEF-CM in the absence of ROCKi. Two days after plating the hESCs, medium was switched to 5iLAF medium which is N2B27 supplemented with 5iLAF cocktail. The medium was changed daily and the cells were passaged again ~10 days later with Accutase as the first time. The domed naïve hESC colonies can be observed in this passage. The naïve hESCs were then split in 1:3-1:5 ratio every 5-7 days.

iii. Freezing naïve hESCs

Naïve hESCs should be treated with Accutase for about 5 min at 37 °C and Accutase

was inactivated and removed by centrifugation. Cell pellet was resuspended in 1 ml cold KSR and transferred to a cryotube. 0.1 ml DMSO was added dropwise to the cryotube. The tubes were then placed in a Mr. Frosty and transferred into a -80 °C freezer overnight before transferred to the liquid nitrogen tank.

2.1.2.3 HEK293T cell culture

HEK293T cells were routinely cultured at 37 °C in DMEM medium and were split in a ratio of 1:10 using Trypsin-EDTA when they were ~90% confluent.

2.1.2.4 Genetic manipulation of cell lines

i. Calcium Phosphate Transfection

Calcium phosphate transfection was routinely used to transfect HEK293T cells for transient or stable expression, via forming a calcium phosphate-DNA precipitate and allowing the binding of the DNA to the cell membrane. To transfect a 10 cm dish of HEK293T cells, cells should be about 70% confluent at the time of transfection and 6 ml of fresh medium should be replaced 2 hours before that. The transfection reaction mix were prepared in two tubes (A and B). In tube B, 30 µg DNA should be mixed with 62 µl $CaCl_2$ and H_2O (added up to total volume of 500 µl). Tube A contained 500 µl of 2x HeBS buffer. Then, tube B mixture should be slowly added to tube A while making bubbles with a 2-ml pipette. After incubated at room temperature for 20 min, the transfection mixture was dropped evenly over the cell culture. On the following day, the medium was replaced with 10 ml fresh medium. Cells could be collected 48 or 72h post-transfection. The calcium phosphate transfection protocol could be scaled up or down accordingly.

ii. Lipofectamine LTX transfection

Lipofectamine transfection is another method to introduce foreign DNA into different cell lines with high efficiencies and viabilities. To transfect hESCs, 1.6 μg of purified plasmid DNA was suspended in 50 μl of OptiMEM with 1.6 μl of PLUS Reagent in an Eppendorf tube; meanwhile, 4 μl of Lipofectamine LTX was suspended in 50 μl of OptiMEM in another tube. Both tubes were incubated at room temperature for 5 min before mixing them. The mixture was then incubated 30 min at room temperature to form the transfection complex. hESCs were split according to the Accutase–Rocki split protocol and about 1 million cells were resuspended in the transfection mix. Cell mixture were incubated at room temperature for another 15 min. The reaction was then neutralized by adding MEF-CM supplemented with Rocki and bFGF. Finally, the cells were plated into three wells and incubated overnight. Regular MEF-CM was used to replace the medium on the following day and the cells were normally collected 48 or 72 hours post-transfection for further analysis.

iii. Lentiviral production and titration

Lentiviral transduction is a potent way to mediate stable expression in both dividing and non-dividing cells. Once the target cells were infected, the viral RNA would be reverse transcribed and incorporated into the host genome. For safety reasons, the essential components for viral production are separated into three plasmids: a lentiviral vector to express the gene of interest, a helper construct pCMVΔ8.91 and an envelope construct VSV-G. To produce lentivirus, $7x10^6$ HEK293T cells were seeded to a T225 flask the day before the transfection and the medium was changed on the following day. The cells were then calcium phosphate transfected with DNA as follows: 45 μg lentiviral vector, 30 μg pCMVΔ8.91 and 15 ug VSV-G. The transfected cells were incubated overnight at 37 °C. 24 ml of DMEM was replaced to the flask after 18 hours. After that, the supernatant was collected after 24-hour incubation and centrifuged at 1,500 g for 5 min. The supernatant was filtered through a 0.45 μm sterile filter unit.

Then, 120 µl Polybrene (20 mg/ml) and 120 µl Chondroitin Sulphate (20mg/ml) were added to the supernatant and incubated at 37 °C for 20 min. The virus particles were collected by centrifugation at 10,000 g for 20 min at room temperature and a white pellet would be visible at the bottom of the tube. The pellet containing virus is dissolved by appropriate volume of DPBS.

GFP lentiviral particles were titrated by making serial dilution of the concentrated virus (1:1, 1:10, 1:100, 1:1000) in PBS. 20 µl of each viral dilution were added to each well of HEK293T cells in a 6-well Plate. The supernatant was removed and replaced by fresh medium on the following day. After 2 days, the cells were observed under the fluorescent microscope to determine the percentage of GFP expression and the amount of viral particles can be calculated via the formula listed below. To transduce 2.5 $x10^5$ single cell hESC in each well of 6-well plate, 1.25 x 10^6 virus can be directly added and transduce the cells overnight. All procedures were performed in biosafety level 2 cabinet and waste was disposed based on the regulations.

virus titer calculation: $Titer = [(F \times Cn) / V] \times DF$

F: The frequency of GFP-positive cells

Cn: The total number of HEK293T cells infected.

V: The volume of the inoculum.

DF: The virus dilution factor.

2.2 DNA and RNA techniques

2.2.1 List of materials

Reagent name	Supplier	Catalogue number
Kits, Chemicals and Reagents		
2-Log DNA Ladder	NEB	N3200S
50 bp DNA Ladder	NEB	N3236S
6x Loading Dye	NEB	B7025
Agarose	Sigma-Aldrich	A9539
Ampicillin	Sigma-Aldrich	A0166
DH-5a Competent *E. Coli*	Thermo fisher	18265017
dNTP Mix (10 mM each)	Thermo Fisher	R0192
Ethanol Absolute	VWR	20821.330
Ethylenediaminetetraacetic Acid (EDTA)	Sigma-Aldrich	E9884
EZ DNA Methylation-Gold ™ Kit	Zymo	D5005
Hispeed Midiprep Kit	Qiagen	12643
Isopropanol	Thermo Fisher	BP2618212
Molecular Grade Water	Thermo Fisher	R0581
Monarch® DNA Gel Extraction Kit	NEB	T1020S
Murine RNase Inhibitor	Thermo Fisher	EO0381
Oligo(dT)18 100 mM	Thermo Fisher	10753741
pGEM-T Easy Vector System	Promega	A1360
Phenol:Chloroform:Isoamyl Alcohol 25:24:1	Sigma-Aldrich	P2069
Q5® Site-directed mutagenesis kit	NEB	E0554S
SafeGreen Nucleic Acid Stain	NBS biologicals	NBS-SG1
Sodium Chloride (NaCl)	Sgima-Aldrich	S3014
Sodium Dodecyl Sulfate (SDS)	Sgima-Aldrich	L3771
SYBR Green Jumpstart Taq Ready Mix	Sgima-Aldrich	S4438
TRI-Reagent	Sgima-Aldrich	T9424
X-Gal/IPTG Solution	Thermo Fisher	10225163

Enzymes		
Jumpstart™ Taq Polymerase	Sigma-Aldrich	D9307
BamHI-HF	NEB	R3136S
DNase I	Thermo Fisher	EN0521
FseI-HF	NEB	R0588S
KpnI-HF	NEB	R3142S
NotI-HF	NEB	R3189S
Protoscript II Reverse Transcriptase	NEB	M0368
PvuII-HF	NEB	R3151S
Q5® High Fidelity Dna Polymerase	NEB	M0493S
SpeI-HF	NEB	R3133S
T4 Ligase	NEB	M0202S
Rnase A	Fisher	10174711
Proteinase K	Sgima-Aldrich	P2308
Plasmids		
Name	**Source**	**Purpose**
pBluescriptII-KS (+)	Addgene-212207	cDNA Cloning vector
pGEMT-Easy-NLRP7 Variant 2	Made in this project; NLRP7 isoform2 cDNA insertion to pGEMT-easy	cDNA isolation; NLRP7 expression vector construction
pCAG-Puro-2A-HA-NLRP7	Made in this project; derived from pCAG-puro-2A-HA-Sox2 (Zhang et al., 2019)	NLRP7 transient overexpression
pCAG-Puro	Lab plasmid	control plasmid
pCAG-dsRed	Lab plasmid	control plasmid
pCAG-HA-NLRP7	Made in this project; derived from pCAG-dsRed	NLRP7 transient overexpression
pCAG-HA-NLRP7-E570del	Made in this project via site directed mutagenesis	NLRP7 patient mimic mutant
pCAG-HA-NLRP7-R693P	Made in this project via site directed mutagenesis	NLRP7 patient mimic mutant
pCMV-Flag-DNMT3L	Made in this project; DNMT3L cDNA insertion to pGEMT-easy	DNMT3L transient expression
pCMV-Flag-EGFP	Lab plasmid	negative control for

		co-IP
pLVTHM-CAG-Puro-2A-HA-NLRP7	Made in this project; derived from pLVTHM-puro-2A-HA-Pax6 (Zhang et al., 2019)	NLRP7 stable overexpression
pLVTHM-CAG-EGFP	Lab plasmid	control plasmid
pLKO.1-NLRP7-shA	Made in this project; insert to pLKO.1-TRC	NLRP7 knockdown
pLKO.1-NLRP7-shD	Made in this project; insert to pLKO.1-TRC	NLRP7 knockdown
pLKO.1-TRC Cloning System	Addgene-10878	shRNA cloning

2.2.2 Buffers and solutions

Buffers	**Components**
TNE lysis buffer	100 mM Tris-HCl pH 7.4 5 mM EDTA 0.2% SDS 200 mM NaCl
50X Tris-Acetate-EDTA Buffer (TAE)	242 g Tris Base 57.1 ml Acetic Acid 18.6 g EDTA Fill dH_2O up to 1L
Lysogeny Broth (LB) Medium	10 g Bacto-tryptone 5 g yeast extract 10 g NaCl Fill dH_2O up to 1L
LB Agar	10 g Bacto-tryptone 5 g yeast extract 10 g NaCl 15 g Agar Fill dH_2O up to 1L
P1 Resuspension Buffer (Qiagen)	100 mM Tris-HCl pH 8.0 10 mM EDTA 100 µg/ml RNase A
P2 Lysis Buffer (Qiagen)	200 mM NaOH 1% SDS
P3 Neutralization Buffer (Qiagen)	3 M Potassium Acetate pH 5.5

2.2.3 List of Primers

i. Bisulphite sequencing primers

Name	Primer	5'→3'	Location
LIN28B-Met[1]	F	ATATTTTGAAGTGTTTTTGTTGTAA	Chr6:104952962-104953307
	original	ATATTCTGAAGTGCCTCTGCTGCAA	
	R	CATACCTAAAACTAAACTATCTTCC	
	original	CATGCCTAAAACTGAGCTGTCTTCC	
NAP1L5-Met[1]	F	GGGGTTTTTTAGTTATTTGATTAG	Chr4:88697589-88697830
	original	GGGGCTCCTCAGCCATCTGACCAG	
	R	CAAAATCTCTCTAAACCAACTCT	
	original	CAAGATCTCTCTGGACCAGCTCT	
PEG10-Met[2]	F	GTGTTATGTTTTATAAATAGATAAG	Chr7:94656871-94657245
	original	GTGCCATGCTTTACAAACAGATAAG	
	R	AACTCATATACCTCTACAATTC	
	original	AGCTCATGTACCTCTGCAGTTC	
IG-DMR-Met[2]	F	GGGTTGGGTTTTGTTAGTT	Chr14:100810889-100811147
	original	GGGCTGGGCCTTGCCAGTT	
	R	CCAATTACAATACCACAAAATTAC	
	original	CCAGTTACAGTACCACAGGATTAC	
MEG3-DMR-Met[2]	F	GTAAGTTTTATAGGTTGTAAAGGGGGT	Chr14:100824582-100824797
	original	GCAAGCTCCACAGGCTGTAAAGGGGGT	
	R	CCACAACTAATAACTAAAAAAATAAACAT	
	original	CCACAGCTAATGACTAGGGAGGTGAACAT	

[1] primers obtained from Sanchez-Delgado et al. (2015).

[2] primers obtained from Frost et al. (2010).

ii. Genotyping and allelic-expression analysis primers

NAME	SNP	Primer 5'→3'	LENGTH	MAF
LIN28B-1634-F	rs221634	GTATTGGTCCTGTTAGGTTTCGG	220 bp	A=0.4858
LIN28B-1634-R		TGTCTCATTTGAGTCATGCTATT		
NAP1L5-0834-F	rs710834	GGTGAGCTCTTGGATCTTGG	389 bp	T=0.4435
NAP1L5-0834-R		GCGGCTTCTCCTCTAACATG		
PEG10-3073-F	rs13073	ACAGAGATGTAAGAGGCAGGC	228 bp	T=0.2344
PEG10-3073-R		CTAGTCACCACTTCAAAACACAC		
ZDBF2-2864-F	rs7582864	TCTGAACAACTTCAGGAAGCG	309 bp	A=0.4707

ZDBF2-2864-R		TTCAATGTGCTCCTGCTCCA		
MEST-0582-F	rs1050582	TGACCACATTAGCCACTATCCA	247 bp	G=0.4837
MEST-0582-R		CCTGCTGGCTTCTTCCTATACA		
SNRPN-705T-F	rs705-Transcript	CAGGCATTCTTAGCTGAGAC	294 bp	T=0.4782
SNRPN-705T-R		CATCTTGCAGGATACATCTC		
SNRPN-705G-F	rs705-Genotyping	CATCAGTCCTAAGTGTGTC	423 bp	
SNRPN-705G-R		GATCACTGCACATGCTGGCAAAC		
MEG3-0608-F	rs11160608	CTGTGGGCTGTGTGTACCTT	442 bp	C=0.4433
MEG3-0608-R		TGCACTTCACTACTCATACATCCA		
H19-9698-F	rs2839698	AACTGGGGAAGTGGGGAAC	278 bp	A=0.2929
H19-9698-R		AAAAGTGACCGGGATGAATG		

*MAF (minor allele frequency) is the second most common allele of a SNP that occurs in a given population.

iii. Primer sets used in RT-PCR

Gene	**Primer**	**5'→3'**
NLRP7	F-start	ATACTCGGAGCACTATGACATCGC
	R-stop	CTCAGCAAAAAAAGTCACAGCACGG
DNMT3L	F-start	AAGACTAGTATGGCGGCCATC
	R-stop	GTCGCATGCTTATAAAGAGGAAG

iv. Primer sets used for site directed mutagenesis

NLRP7 mutation	**Primer**	**5'→3'**
E570del	F	TGAGAATCACTAGTGAATTCGCGGCC
	R	CTTCAGGTCGGTCACGGATAAGGG
R693P	F	CTCTTCTGTGccgATTCTTTGTG
	R	TCACTCAGGAAGCTTTGTTTC

v. Oligo sequence for NLRP7 shRNA

	Primer	**5'→3'**
shRNA-A	F	CCGGCCGTTCAAGGAAATTTCTATTCTCGAGAATAGAAATTTCCTTGAACGGTTTTTG
	R	AATTCAAAAACCGTTCAAGGAAATTTCTATTCTCGAGAATAGAAATTTCCTTGAACGG
shRNA-D	F	CCGGGCAATGCAAAGCACATCTTCACTCGAGTGAAGATGTGCTTTGCATTGCTTTTTG
	R	AATTCAAAAAGCAATGCAAAGCACATCTTCACTCGAGTGAAGATGTGCTTTGCATTGC

iv. qRT-PCR primers

Primer name	5'→ 3'
q*b-actin-F*	TGTCTGGCGGCACCACCATG
q*b-actin-R*	AGGATGGAGCCGCCGATCCA
q*DNMT1-F*	GGTTCTTCCTCCTGGAGAATGTC
q*DNMT1-R*	GGGCCACGCCGTACTG
q*DNMT3A-F*	CAATGACCTCTCCATCGTCAAC
q*DNMT3A-R*	CATGCAGGAGGCGGTAGAA
q*DNMT3B-F*	CCATGAAGGTTGGCGACAA
q*DNMT3B-R*	TGGCATCAATCATCACTGGATT
q*DNMT3L-F*	GGGACAACTGAAGCATGTGGT
q*DNMT3L-R*	AAGATCGAAGGGTCCCCACT
qTET1-F	ACCCCCTGTCACCTGCTGAGG
q*TET1-R*	GCGATGGCCACCCCACCAAT
q*TET2-F*	TCACACCAGGTGCACTTCTC
q*TET2-R*	GGATGGTTGTGTTTGTGCTG
q*TET3-F*	TCTCCCCAGTCTTACCTCCG
q*TET3-R*	CCAGGCTTCAGGGAACTCAG
q*OCT4-F*	TCGAGAACCGAGTGAGAGGC
q*OCT4-R*	CACACTCGGACCACATCCTTC
q*REX1-F*	TCACAGTCCAGCAGGTGTTT
q*REX1-R*	GCCATCACATAAGGCCCACA
q*KLF2-F*	CACACAGGTGAGAAGCCCTA
q*KLF2-R*	CAGTCACAGTTTGGGAGGGG
q*NANOG-F*	TGATTTGTGGGCCTGAAGAAAA
q*NANOG-R*	GAGGCATCTCAGCAGAAGACA
q*PRDM14-F*	TCTGTGATGTGGGAGATCTTTGA
q*PRDM14-R*	ACTCTTCAGAGGGCCCAGAT
q*TCL1B-F*	GGGCAGCAGATATGAACCCA
q*TCL1B-R*	CCCAGAAACTGGAATCCGCT
q*NLRP7-F*	GACGGACAGGTGCAAGAAAT
q*NLRP7-R*	AGAATGGAATGAACCGTTGG
q*NLRP2-F*	TGAGGAAACCACTGTGCAACTT
q*NLRP2-R*	AACTGAACGGAGGGATGGAA
qOOEP-F	ACTTCCGCCGCCACAGATT

qOOEP-R	GGCTCGTCCACTCCATTTC
qKHDC3L-F	CTAACGGCGAGGCTGAGAT
qKHDC3L-R	GCCTTTCCTGAGCCTTTCG
qNLRP5-F	GTGGTCCCTCTATGGATGCG
qNLRP5-R	CCCGCTCTGTCAGGATGCT
qTLE6-F	AGCGTTGGAATGGACGACTT
qTLE6-R	TGGAAGAGACGTCACAGCAC
qPADI6-F	GTGGCATCTGTCTATGAGGACC
qPADI6-R	CTGAGGTGTGTCGAGGATCAAG

vii. Semi-quantitative PCR primers

Gene	**Primer**	**5'→3'**
NLRP7-exon10	F	CAGCCTCACAAACCTGGACT
	R	CCTCCAACAGCTTCTTGATTTCC
	mid-R	GCTGAGGAGAGCAGATCCAA

2.2.4 Methods

2.2.4.1. Genomic DNA (gDNA) extraction

Cells in a well of 6-well plate were lysed in 700 µl TNE lysis buffer with proteinase K (100 µg/ml) and RNase A (40 µg/ml) by overnight incubation at 37°C. Then, samples were pulse-spun for 30 sec in a microcentrifuge and the supernatant were transferred to a 1.5 ml Eppendorf tube followed by the addition of equal volume of phenol:chloroform:isoamyl alcohol solution. The tube was vigorously shaken and spun at 13,000 g for 5 min. The aqueous phase was transferred to a new tube and mixed with equal volume of isopropanol. The sample was further incubated at -20°C for 2 hours and the gDNA precipitation was pelleted by centrifugation at 13,000 g for 20 min. The DNA pellet was washed with 70% ethanol, followed by 5 min centrifugation at 13,000 g. Finally, the DNA pellet was air dried. 50 µl DNase-free distilled water was added to resuspend the DNA and its concentration was measured by a NanoDrop® Spectrophotometer.

2.2.4.2. Bisulphite conversion

Bisulphite conversion is the most popular way to determine and quantify the DNA methylation. Treating DNA with bisulphite can convert unmethylated C to U, while the methylated C remain unchanged. EZ DNA Methylation-Gold™ Kit was used here according to the manufacture's instruction to convert gDNA from hESCs.

2.2.4.3. Polymerase Chain Reaction (PCR)

After bisulphite conversion, the DNA methylation profile can be determined by PCR amplification using Taq polymerase and specific primers targeting the DMR of the imprinted genes and followed by Sanger sequencing. Reagents in the following table were always kept on ice before mixing them in a PCR tube. Bio-Rad Dyad DNA Engine thermocycler was set up according to different annealing temperature of primers.

Unsaturated PCR amplification (25 cycles) was used for semi-quantitative PCR, otherwise 30-35 cycles were used for normal PCR.

Component	Amount
10x Taq Buffer with $MgCl_2$	2.5 µl
dNTP	0.5 µl
Forward primer	0.5 µl
Reverse primer	0.5 µl
***Taq* Polymerase (2.5 unit/µl)**	0.25 µl
DNA template	2 µl
Nuclease-free Water	Up to 25 µl

Taq Polymerase PCR cycle

Initial Denaturation	95 °C	30 sec	
Denaturation	95 °C	10 sec	25-35 cycles
Annealing	Ta	30 sec	
Extension	72 °C	30 sec	
Final Extension	72 °C	5 min	
Hold at	10 °C	∞	

The PCR products were run in a 1.5% agarose gel to analyse the size of the PCR products.

2.2.4.4 RNA extraction from adherent cells

Culture medium was removed from 6-well plate and cells were washed with PSB before extraction. Then, 1 ml of TRI reagent was added into the well to lyse the cells. The cell lysate was then transferred to a DNase-free and RNase-free Eppendorf tube. 200µl of chloroform: isoamyl alcohol was added into the tube and vigorously shaken. The mixture was centrifuged at 13,000 g for 15 min at 4 °C and the samples would be separated into 3 layers. The top aqueous layer containing RNA was transferred into another new Eppendorf tube, followed by addition of 500 µl of isopropanol. The tube was incubated at room temperature for 10 min to allow nucleic acid precipitation. The

tube was centrifuged at 13,000 g for 10 min at 4 °C to pellet down the RNA and the pellet was washed once by 1ml 75% ethanol, followed by a final centrifugation at 13,000 g for 10 min. Finally, the RNA pellet was air dried for 10 min and resuspended in 20 µl of RNase-free water. DNase I was used to treat the RNA sample to remove trace remaining genomic DNA as shown below:

Component	Amount
RNA	20 µl
10X Reaction buffer with $MgCl_2$	2.5 µl
DNaseI, RNase-free	2.5 µl

After the mixture was incubated at 37 °C for 30 min, 2.5 µl 50 mM EDTA was added to the tube and heat-inactivated at 65 °C for 10 min. The concentration of the RNA can be measured by NanoDrop® Spectrophotometer. Extracted RNA should be stored at -80 °C freezer to prevent degradation.

2.2.4.5 Cloning cDNA in to pGEM-T easy

i. cDNA generation by Reverse transcription polymerase chain reaction (RT-PCR)

The extracted RNA was reverse transcribed to cDNA by NEB ProtoscriptII Reverse Trasncriptase. 1.3 µg RNA was diluted in RNase-free water to obtain total volume of 11 µl. It was mixed with 1 µl 10mM dNTP and 1 µl oligo d(T) and incubated at 65 °C for 5 min. Then, the PCR tube was cooled down on ice and the mixture of 4µl 5X ProtoscriptII buffer, 1 µl DTT, 1 µl RNase inhibitor and 1 µl reverse transcriptase was added to the tube. Finally, the tube was incubated at 42°C for 1 hour, followed with heat inactivation (65°C, 20 min). The 20 µl cDNA can be diluted in 180 µl molecular grade water for further analysis.

The NLRP7 and DNMT3L cDNAs were then amplified by PCR using NEB Q5® High-Fidelity DNA polymerase to reduce mis-reading errors. The PCR reaction was gently

mixed in a 0.2 ml PCR tube and underwent PCR cycles in a thermocycler.

Component	Amount
5x Q5 Reaction buffer	10 µl
10mM dNTP	1 µl
10µM Forward Primer	2.5 µl
10µM Reverse Primer	2.5 µl
Reverse transcription mix (diluted)	1 µl
Q5 High-Fidelity DNA polymerase	0.5 µl
Nuclease-free Water	Up to 50 µl

Q5 high-fidelity DNA polymerase PCR cycle

Initial Denaturation	98°C	30 sec	
Denaturation	98°C	10 sec	30-35 cycles
Annealing	Ta	30 sec	
Extension	72°C	1-2 min	
Final Extension	72°C	10 min	
Hold at	10°C	∞	

ii. A-tailing

Since Q5® High-Fidelity DNA polymerase produces blunt-ended PCR products, the NLRP7 and DNMT3L cDNA PCR products were A-tailed and then ligated to pGEMT-Easy vector.

Component	Amount
10x Taq Buffer with $MgCl_2$	1 µl
Kit purified PCR products	200 ng
Taq Polymerase (2.5 unit/µl)	2 µl
dATP (1mM)	2 µl
Nuclease-free Water	Up to10 µl

The A-tailing reaction was incubated at 70 °C for 30 min.

iii. Cloning DNA into pGEM-T easy vector and sequencing

The A-tailed PCR products were ligated to pGEM-T-Easy vectors as described in the following table.

Component	Amount
2x Ligation Buffer	5 µl
A-tailing Reaction	1-2 µl
pGEMT-Easy vector(50 ng/µl)	1 µl
T4 DNA Ligase (3 units/ µl)	1 µl
Nuclease-free Water	Up to10 µl

The reaction mix was incubated overnight at 4°C to maximize the number of transformants.

2.2.4.6. Construction of expression vectors

NLRP7 transient expression vector (pCAG-2A-HA-NLRP7) was generated via 2 steps. First, HA-NLRP7 DNA fragment was cut from pCAG-Puro-2A-HA-NLRP7 donor plasmid by BamHI and NotI and inserted into an intermediate vector (pBluescriptII KS (+)) cut with the same restriction enzymes. This step provided an additional restriction cut site, KpnI, at the upstream of HA-NLRP7 fragment. Second, the resulting recombined plasmid was then double digested by NotI and KpnI, so that the HA-NLRP7 fragment could be inserted to the backbone of pCAG-dsRed which was also generated by NotI/KpnI double digestion. Construction of other NLRP7 expression vectors is going to be described in Chapter 4.2.

i. Restriction digest and gel extraction

Restriction reactions were carried out according to the NEB restriction enzyme instructions (https://nebcloner.neb.com/#!/redigest).

Component	Amount
10x CutSmart Buffer	2 µl
Plasmid DNA	1-2 µg
Restriction enzyme 1	1 µl
Restriction enzyme 2	1 µl
Nuclease-free Water	Up to 20 µl

All digestion products were separated by gel electrophoresis. The band with desired size was cut out under the UV light. The DNA was extracted and purified by Monarch® DNA Gel Extraction Kit.

ii. Ligation of inserts to backbone plasmids

To ligate the cleaned up insert DNA to backbone plasmid, 50 ng backbone plasmid (4 kb) and 37.5 ng insert (1 kb) was mixed at a molar ratio of 1:3 with 2 µl T4 DNA ligase buffer (10x), 1 µl T4 DNA ligase and nuclease free water up to 20 µl total volume. The reaction mix was incubated at 16 °C overnight before heat inactivation (65 °C, 10 min). The reaction can be stored at -20 °C for further usage.

2.2.4.7. Generation of shRNA Lentivector

pLKO1 is a commonly used lentivector for the expression of shRNAs. First, the top-scoring target of NLRP7 was chosen by using an online siRNA selection tool (http://sirna.wi.mit.edu). To generate oligos that can be cloned into the pLKO1 vector, the sense and antisense sequence of the selected siRNA was inserted into the oligos below, without changing the ends.

Forward oligo: 5’ CCGG—sense—CTCGAG—antisense—TTTTTG 3

Reverse oligo: 5’ AATTCAAAAA—sense—CTCGAG—antisense 3’

There are two unique cloning sites in pLKO1 cloning vector, EcoRI and AgeI, for inserting shRNA. The annealed oligos would generate two ends compatible with EcoRI and AgeI restriction site. To anneal the oligos, 5 µl forward oligo (20µM), 5 µl reverse oligo (20µM), 5 µl NEB buffer and 35 µl ddH_2O were mixed together and incubated at 95 °C in a heating block for 10 min. Then, the heating block was slowly cooled down to room temperature over a few hours. Therefore, the annealed shRNA oligos could be ligated into the pLKO1 cloning vector by T4 ligation.

2.2.4.8. Site-directed mutagenesis

Q5® Site-directed mutagenesis kit was used to generate the NLRP7 transient expression vectors (pCAG-HA-NLRP7) with patient mimic mutations: E570 nonsense mutation and R693P missense mutation. NEBaseChanger® was used to design the primers and the appropriate annealing temperature was suggested by NEB Tm Calculator (http://tmcalculator.neb.com). Firstly, the listed reagents were mixed in a PCR tube and transferred to a thermocycler with the following cycling conditions:

Component	Amount
2x Q5 Hot Start High-Fidelity Master Mix	12.5 µl
10µM Forward Primer	1.25 µl
10µM Reverse Primer	1.25 µl
DNA Template	1 µl (1-25 ng)
Nuclease-free Water	9 µl

Initial Denaturation	98 °C	30 sec	
Denaturation	98 °C	10 sec	25 cycles
Annealing	Ta	30 sec	
Extension	72 °C	20-30 sec/kb	
Final Extension	72 °C	10 min	
Hold at	10 °C	∞	

Secondly, kinase, ligase and DpnI (KLD) treatment was performed to clean up the parental plasmid DNA and ligate the newly amplified linearized plasmids.

Component	Amount
PCR product	1 µl
2X KLD Reaction Buffer	5 µl
10X KLD Enzyme Mix	1 µl
Nuclease-free Water	9 µl

After 5 min incubation at room temperature, the reaction mix was ready to be transformed into DH5a competent *E. coli* cells.

2.2.4.9. Bacterial transformation

5-10 µl of the ligation mix was added to 100 µl DH5α *E.coli* after it was slowly thawed on ice. The tube with DNA-bacterial mixture was gently flicked and incubated on ice for 30 min. The mixture was heat-shocked in a 42 °C water bath for 40 sec and then immediately put on ice for another 2 min, followed by addition of 200 µl LB broth and incubation for 1 hour at 37 °C with shaking (250 rpm). The bacterial mix was then plated onto a LB agar plate containing 100 ng/ml Ampicillin and X-gal/IPTG to select transformants. The plate was incubated at 37 °C overnight.

2.2.4.10. Plasmid amplification and extraction

Blue-white screening is a simple and rapid way to select the successful insertion of DNA into a vector. To amplify the successfully inserted bacterial colonies, each single white colony was picked and grown in 5ml LB containing 100 ng/ml Ampicillin at 37 °C, overnight on 250 rpm shaker. 1.5 ml of bacterial culture was pelleted by centrifugation at 6,000 g for 15 min and the pellet was resuspended in 250 µl P1 buffer and vortexed vigorously. Then, an equal volume of P2 lysis buffer was added followed by an

incubation for 5 min before adding 250 µl of P3 neutralisation buffer. Bacterial debris was removed by centrifugation at 13,000 g for 5 min and the supernatant was transferred to a new tube, mixed vigorously with an equal volume of isopropanol. Plasmid DNA was pelleted down by centrifugation at 13,000 g for 5 min and air dried, finally resuspended in 40 µl DNase- and Rnase-free H_2O. The plasmid DNA concentration was measured by a NanoDrop® spectrometer. Those extracted plasmids were sent for Sanger sequencing and the sequencing results were compared with the original sequence to determine the methylation status of the imprinted genes.

2.2.4.11 Genotyping and imprinting analysis

Potential SNPs in imprinted genes were chosen through the NCBI Variation viewer (https://www.ncbi.nlm.nih.gov/variation/view/overview/). PCR primers were designed to amplify desired regions with potential SNPs from the genomic DNA of H1 and H7 hESCs. PCR products were purified by separation in and extraction from agarose gels using NEB Monarch gel extraction kit. And those PCR products were sent for Sanger sequencing directly to identify the genotypes of potential SNPs. Sequencing results were interrogated using SnapGene Viewer. Heterozygous samples were further analysed for allelic expression using RT-PCR and direct Sanger sequencing.

2.2.4.12. Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

qRT-PCR was a convenient way to check the RNA expression level of different genes in hESCs. qRT-PCR were performed in Frame Star 96 full skirted PCR plates. For each well, the reaction mix was listed in the following table and the qPCR programme was setup as described in the table below.

component	Volume
SYBR	10 µl
Forward primer	0.5 µl
Reverse primer	0.5 µl
cDNA template	Δ
Nuclease-free Water	Up to 20 µl

qPCR cycle set up:

Initial denaturation	10 min	94 °C	
denaturation	30 sec	94 °C	40 cycles
Annealing	30 sec	58 °C	
Extension	30 sec	72 °C	
Plate read	1 sec	80 °C	
Plate read	1 sec	82 °C	
Plate read	1 sec	85 °C	
Melting curve from 70 °C to 95 °C, read every 0.5 °C, hold 1 sec			

Two housekeeping genes, β-actin and RPL22, were included in the qRT-PCR experiments, and always showed similar expression trends. Therefore, only β-actin was used as the normalizer for analysis. The ΔCt and Δ(ΔCt) method was used to analyse qRT-PCR data:

$$\text{Expression of target gene relative to b-actin} = 2^{-\Delta CT}$$

$$\text{Fold change of gene expression level after naïve conversion} = 2^{-\Delta(\Delta CT)}$$

*$\Delta Ct = Ct_{\text{(target gene)}} - Ct_{\text{(housekeeping gene: b-actin)}}$

**$\Delta(\Delta Ct) = \Delta Ct_{\text{(naïve)}} - \Delta Ct_{\text{(primed)}}$

2.3 Protein techniques

2.3.1 List of Materials

Reagent	Supplier	Catalogue Number
30% Acrylamide/Bis Solution-37.5:1	BioRad	1610159
4',6-Diamidino-2-Phenylindole (DAPI)	Sigma-Aldrich	D9542
Ammonium Persulfate (APS)	Sigma-Aldrich	A3678
Anti-Flag M2 Agarose Beads	Sigma-Aldrich	A2220
Anti-HA Agarose Beads	Sigma-Aldrich	A2095
Anti-Mouse IgG Agarose Beads	Sigma-Aldrich	A6531
Bicinchoninic Acid (BCA) Assay Protein Quantification Kit	Thermo Fisher	23227
Bovine Serum Albumin (BSA)	Sigma-Aldrich	A5611
Bromophenol Blue	Sigma-Aldrich	B1026
Clarity ECL Western Blotting Substrate	BioRad	1706061
CL-Xposure Film	SLS	MOL7016
Glycerol	Sigma-Aldrich	G5516
Goat Serum	Sigma-Aldrich	G9023
Immobilon® Forte Western HRP Substrate	Millipore	WBLUF0500
Immobilon® Polyvinylidene Fluoride (PVDF) Membrane	Millipore	431175
Methanol	VWR	85800
Mowiol 4-88	Calbiochem	475904
Nonidet P-40 (NP-40)	Anachem	E109-100ML-R
Phenylmethanesulfonyl Fluoride (PMSF)	NEB	8553S
Pierce™ 16% Formaldehyde	Thermo Fisher	28908
Prestained Protein Ladder	NEB	p7712L
Protease Inhibitor Cocktail	Sigma-Aldrich	P8340
Skimmed Milk Powder	Sigma-Aldrich	70166
Sodium Azide	Sigma-Aldrich	S2002
Sodium Dodecyl Sulphate (SDS)	Sigma-Aldrich	L3771
Tetramethylethylenediamine (TEMED)	Sigma-Aldrich	T7024
Triton X-100	Sigma-Aldrich	T8787
Tween-20	Sigma-Aldrich	P9416
β-Mercaptoethanol 14.3M	Sigma-Aldrich	M3148

2.3.2 List of Antibodies

Antibody	Company	Cat. number	Dilution
Mouse anti-NLRP7	Santa Cruz	sc-377190	IB: 1:500 IF: 1:100
Rabbit anti-DNMT3L	Abcam	ab3493	IB: 1:1000 IF: 1:500
Mouse anti-β-actin	Sigma-Aldrich	A5316	1:5000
Rabbit anti-HA	Sigma-Aldrich	H6908	1:1000
Mouse anti-Flag M2	Sigma-Aldrich	F1804	1:1000
Mouse anti-GAPDH	Santa Cruz	sc-365062	1:500
Mouse anti-laminB	Santa Cruz	sc-365962	1:500
Mouse anti-a-tubulin	Cell signalling	3873S	1:1000
Goat-anti-GFP	AbD Serotec	AHP975	1:500
Anti-Goat IgG-HRP	Santa Cruz	2020	1:5000
Anti-mouse IgG-HRP	Startech	115-035-174	1:5000
Anti-rabbit IgG (H+L)-HRP	Thermo	31460	1:5000
Mouse IgG-Alexa Fluor 488	Life Technologies	A11001	1:400
Rabbit IgG-Alexa Fluor 568	Life Technologies	A11011	1:400

2.3.3 Buffers and Solutions

i. Western Blotting

Buffers/Solutions	Components	Final conc.
Radio Immunoprecipitation Assay (RIPA) Buffer	25 ml 1M Tris-HCl pH 8.0 15 ml 5M NaCl 5 ml NP-40 2.5 g Sodium deoxycholate 0.5 g SDS Add ddH_2O up to 500 ml	50 mM 150 mM 1% v/v 0.5% w/v 0.1% w/v
5 x Sample Buffer	1.25 ml 0.5M Tris-HCl pH 6.8 2.5 ml glycerol 2 ml 10% SDS 400 µl 0.5% bromophenol blue 500 µl β-mercaptoethanol (14.3M) Autoclaved water to 10 ml	50 mM 25% v/v 2% w/v 0.02% w/v 715 mM
4% Polyacrylamide Stacking Gel	2.5 ml 0.5 M Tris-HCl pH 6.8 1.3 ml 30% Acrylamide/Bis 100 µl 10% SDS 50 µl 12% APS	125 mM 4% v/v 0.1% w/v 0.06 % w/v

	12 µl TEMED add ddH_2O up to 10 ml	0.12% v/v
7.5% Polyacrylamide Running Gel	2.5 ml 0.5M Tris-HCl pH 8.8 2.5 ml 30% Acrylamide/Bis 100 µl 10% SDS 50 µl 12% APS 12 µl TEMED add ddH_2O up to 10 ml	125 mM 7.5% v/v 0.1% w/v 0.05% w/v 0.12% v/v
Transfer Buffer	5.82 g Tris base 2.93 g Glycine 3.75 ml 10% SDS 200 ml methanol add ddH_2O up to 1 L	48 mM 39 mM 0.04% w/v 20% v/v
10X SDS Running Buffer	30.3 g Tris base 144.2 g Glycine 50 ml 20% SDS add ddH_2O up to 1 L	250 mM 1.9 M 1% w/v
Tris Buffered Saline with Tween-20 (TBS-T)	20 ml 1M Tris-HCl pH 7.6 26 ml NaCl 1 ml Tween-20 add ddH_2O water to 1 L	20 mM 130 mM 0.1% v/v

ii. Co-immunoprecipitation

Buffers	Components	Final conc.
NP40 IP buffer	50 ml of 1 M Tris pH 8.0 30 ml 5 M NaCl 100 ml NP-40 add ddH_2O water up to 1L	50 mM 150 mM 10% v/v
2 x Sample Buffer	2.5 ml 0.5M Tris-HCl pH 6.8 2 ml glycerol 4 ml 10 % SDS 160 µl 1.25% bromophenol blue 500 µl β-mercaptoethanol add ddH_2O up to 10ml	125 mM 20% v/v 4% w/v 0.02% w/v 715 mM

iii. Immunocytochemistry

Buffers	Components	Final conc.
Fixation Buffer	10 ml 16% PFA solution 30 ml DPBS	4% w/v
Blocking Buffer	1 ml goat serum 0.25 g BSA 30 µl Triton X-100 add PBS up to 10ml	10% v/v 2.5% w/v 0.3% v/v
Mowiol mounting solution	12 ml 200 mM Tris-HCl pH 8.5 6 g glycerol 2.4 g Mowiol 4-88 6 ml ddH_2O Heat at 60°C to dissolve Mowiol	135 mM

iv. Nuclear Cytoplasmic Fractionation

Buffers	Components	Final conc.
0.5% TritonX-100 Lysis Buffer	50 ml 1M Tris-HCl pH7.5 5 ml TritonX-100 27.5 ml 5M NaCl 100 ml Glycerol 10 ml 0.5M EDTA add ddH_2O up to 1 L	50 mM 0.5% v/v 137.5 mM 10% v/v 5mM
0.5% SDS Nuclear Lysis Buffer	380 ml 0.5% TritonX-100 Lysis Buffer 20 ml 10% SDS	

2.3.4 Methods

2.3.2.1 Western Blotting

Cells were dissociated by trypsin-EDTA and collected from a well of 6-well plate into a pre-chilled Eppendorf tube. The cell pellet was then lysed in 400 µl ice-cold RIPA buffer, containing 1% Phosphatase inhibitor cocktail and 1%PMSF. The ice-cold RIPA buffer could also be directly added into culture dish after medium was removed and cells were washed in DPBS. Cell lysate was centrifuged at 13,000 g for 20 min at 4°C. Supernatant containing proteins was taken, avoiding the cell debris. BCA assay protein quantification kit was used to quantify the protein concentration based on manufacturer's instruction.

Normally, 15 to 20 µg protein from each sample was mixed with 5x sample buffer to get 1x sample buffer. The sample was then incubated at 95 °C for 5 min and the proteins were separated in a 7.5% Bis-Tris polyacrylamide gel (180 v, 60 min). The proteins were then transferred to a PVDF membrane via electroblotting (semi-dry, 20 v, 60 min). The membrane was immediately blocked by 5% skimmed-milk and then incubated with primary antibody at 4°C overnight. On the following day, the membrane was washed 3 times with 1% TBS-T and incubated with secondary antibody for 1h. Repeated the washing steps as above, the signals were detected using ECL substrate followed by image development onto a CL-Xposure film in Optimax X-ray film processor.

2.3.2.2 Co-immunopricipitation (co-IP)

Co-IP is widely used to identify the physical interaction between different proteins, by using a protein-specific antibody to capture the specific protein as well as other proteins that bind to the targeted protein. The protein complex can be pulled down and used to identify those new binding proteins. Firstly, both pCAG-HA-NLRP7 and pCMV-Flag-DNMT3L were co-expressed in a 10 cm dish of HEK293T cells for 48 hours and

then lysed in 1 ml pre-chilled NP-40 IP buffer with 1:100 protease inhibitor and 1:100 PMSF. After 2 min incubation, the cells were scraped and collected into a clean 1.5ml tube and the cells were passed through a 23 G needle 25 times to help break cells and shear DNA. The supernatant was transferred to a new tube after centrifugation at 13,000 g for 10 min at 4 °C.

Agarose beads used here are already covalently linked to anti-HA or anti-Flag M2 antibody to pull down the fused proteins. To prepare the protein lysate and eliminate the non-specific binding, the cell lysate were pre-cleaned with normal anti-mouse IgG agarose beads on shaker for an hour at 4°C. Then, the cell lysate and IgG agarose beads were separated by centrifugation at 5000 rpm for 1 min. Normally, 1 mg of protein was used for Co-IP. Either anti-Flag or anti-HA agarose beads were used for precipitation and rotated at 4°C overnight. On the following day, the agarose beads were pelleted via centrifugation at 5000 rpm for 40 sec and washed 5 times with 1 ml pre-chilled IP buffer. Finally, the precipitated proteins were eluted from agarose beads with 2 x sample buffer. The eluted samples were boiled at 95 ºC for 5 min and subjected to immunoblotting.

2.3.2.3 Immunostaining of hESCs

Coverslips were placed in a 12-well plate and pre-coated with Matrigel. hESCs were seeded onto the coverslips and expanded a couple of days, Then medium was removed and cells were washed with PBS before they were fixed in 4% formaldehyde at room temperature for 10 min. After washing with PBS to remove the fixative, the cells were incubated in blocking buffer for 1 hour, followed by incubation with diluted primary antibody overnight at 4ºC. On the following day, the cells were washed 3 times with PBS and incubated with fluorophore-conjugated secondary antibody in dark for 1 hour and then incubated in 1ng/ml DAPI for 5 min. Finally, coverslips were mounted with Mowiol and can be observed under the Nikon inverted fluorescent microscope.

2.3.2.4 Nuclear-Cytoplasmic Fractionation

To separate the nuclear and cytoplasmic protein of Naïve hESCs, cells were seeded in a 10 cm dish. 500 µl 0.5% Tritonx-100 lysis buffer supplemented with 1% protease inhibitors and 1% PMSF was added and incubated on ice for 5 min to lyse the cells. The insoluble nuclei was pelleted by centrifugation at 13,000 g for 15 min at 4^{o}C. Then, the supernatant containing cytoplasmic protein was transferred to a new tube. The nuclei pellet was washed 3 times with Tritonx-100 lysis buffer and resuspended in 0.5% SDS Nuclear Lysis Buffer. The gDNA was sheared by passing through a syringe. The nuclear lysate was centrifuged at 13,000 g for 15 min at 4^{o}C and the supernatant was transferred into a new tube. The cytoplasmic and nuclear proteins were subjected to immunoblotting.

2.4 Equipment, software and online tools

Equipment
Nanodrop® Spectrophotometer
Bio-Rad Dyad DNA Engine Thermocycler
Bio-Rad Mini Trans-Blot® Cell
Bio-Rad Opticon2™ DNA Engine Real-Time Fluorescence Thermocycler
Nikon Inverted Fluorescence Microscope
Optimax X-Ray Film Processor

Software
Image J
Opticon Monitor 3
Serial Cloner
SnapGene

Online tools	
T_m Calculator	http://tmcalculator.neb.com
Basic Local Alignment Search Tool (BLAST)	http://blast.ncbi.nlm.nih.gov/Blast.cgi
NCBI Variation viewer	https://www.ncbi.nlm.nih.gov/variation/view/overview/
NEB Double Digest Finder	https://nebcloner.neb.com/#!/redigest
NEBcutter V2.0	http://nc2.neb.com/NEBcutter2/
Primer-BLAST	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
siRNA selection tool	http://sirna.wi.mit.edu

Chapter 3

Characterization of Human Embryonic Stem Cell (hESC) as a Model System to Study NLRP7 and Genomic Imprinting

3.1 Introduction

Biallelic *NLRP7* mutations in humans have been identified to be responsible for recurrent BiCHM (Murdoch et al., 2006; Wang et al., 2009). However, how mutations in NLRP7 induce this abnormal pregnancy remains largely unknown. Two different hypotheses have been proposed: 1) Defective NLRP7 might give rise to an abnormal immune response which could lead to BiCHM. This hypothesis is supported by the fact that several members of NLRP protein family have crucial roles in inflammation and the immune response (Murdoch et al., 2006). 2) Defective NLRP7 might generate deficient maternal imprinting, leading to molar pregnancy. This presumption is based on the phenotypic similarity between BiCHM, and typical androgenic CHM (Mahadevan et al., 2013). With more research in this area, the later theory has become more accepted. For example, a recent study on the molar tissues collected from patients with NLRP7 mutations provided direct evidence that methylation of maternally imprinted genes is indeed absent in those molar tissues (Sanchez-Delgado et al., 2015). These results support the hypothesis that NLRP7 has an essential role in the establishment or maintenance of maternal imprinting during oogenesis or early embryogenesis. As stated in Chapter 1, maternal imprinting is established by *de novo* methylation during oocyte maturation and this maternal methylation pattern is protected from the second wave of global demethylation after fertilisation (Figure1.5). It remains unknown what role NLRP7 plays in this process. This is largely attributed to a lack of research materials as it is extremely difficult to access human embryonic materials and none of the laboratory rodent animals have the *NLRP7* gene (Tian et al., 2009). Although some farm animals have the *NLRP7* gene, it seems to function differently from that in human, as knockdown of NLRP7 in sheep leads to embryonic arrest at the 8 cell stage, a time when zygotic genes are activated, rather than a molar pregnancy. This phenotype is similar to that of NLRP2-knockdown in mice (Li et al., 2019; Peng et al., 2012; Mahadevan et al., 2017). Therefore, it is essential to find a suitable human or non-human primate model system to study NLRP7.

hESCs derived from the inner cell mass (ICM) of preimplantation human embryos may be a potential model system to explore the role of NLRP7 in maternal imprinting during embryogenesis as they are the only *in vitro* human materials that are close to human embryonic tissues. Particularly since the recent development of the conversion of hESCs to the naïve state (Theunissen et al., 2014; Guo et al., 2017), in which process genomic DNAs become hypomethylated, to a level similar to that of human ICM. Interestingly, imprinted genes were also reported to lose their imprinting status, becoming hypomethylated in both alleles and are therefore different from those in the ICM (Pastor et al., 2016). Nonetheless, this process, to an extent, recapitulates the *genome wide* demethylation during early development, and thereby the hESC conversion system might be a suitable model to study whether NLRP7 plays any role in the establishment or maintenance of human maternal imprinting during the global demethylation process.

In order to test this hypothesis, it is necessary that the two alleles of the imprinted gene are distinguishable in our hESC lines and that the conversion process can be established with those hESCs in our laboratory. Hence, in this chapter I aim to address the following questions:

i. Can single nucleotide polymorphisms (SNPs) be identified in the imprinted genes of H1 and H7 hESCs?
ii. What are the expression profiles of the imprinted genes in the primed state, i.e. under routine culture condition?
iii. Can hESCs in our lab be converted from primed to the naïve state via the 5iLAF method?
iv. Do imprinted genes alter their expression pattern upon conversion from primed to naïve states? If so, changes in DNA methylation related factors will be investigated thereafter.

3.2 Results

3.2.1 identification of allele-specific SNPs in the imprinted genes of H1 and H7 hESC lines.

To investigate whether hPSCs could be used for human imprinting studies, I first analysed the genomic DNA (gDNA) in H1 and H7 hESCs to determine whether there are allele-specific SNPs in imprinted genes. SNPs are the most common genetic variation that occur in human, and refers to the substitution of a single nucleotide at a specific site of the genome (Syvänen, 2001). If allele-specific SNPs could be identified, it would enable the determination of allele-specific gene expression of these imprinted genes. Six maternally imprinted genes (*LIN28B, PEG10, NAP1L5, ZDBF2, MEST* and *SNRPN*) and 2 paternally imprinted genes (*H19* and *MEG3*) were initially selected for this study as these maternally imprinted genes have been shown to exhibit aberrant imprinting in the molar tissue of BiCHM in women with NLRP7 mutations, even though aberration is not limited to these genes (Sanchez-Delgado et al., 2015).

Figure 3.1A illustrates the process to identify the SNPs in hESCs. First, gDNA was isolated from these cells and PCR was carried out with primers that flank the regions of imprinted gene described above. The amplified PCR products were then purified by gel extraction and sent for Sanger sequencing. Sequencing results showing two different nucleotides were equally detected at the same site, indicated that the paternal and maternal allele of this gene were different at this position. If only one nucleotide was detected at the site, it indicated that the paternal and maternal alleles of this gene were the same at this specific position, thus no SNP is identified. The results showed that maternally imprinted genes *LIN28B*, *ZDBF2, PEG10* and *NAP1L5* exhibited two peaks at one position, labelled N, rather than one of the 4 nucleotides: A, T, C and G, and that the same pattern appeared in both H1 and H7 hESC lines (Fig 3.1B). For example, in *LIN28B* gene nucleotides T & A were detected at the same position in H1 and H7 DNAs. These results indicate that SNPs in these genes exist in both hESC

lines. Similarly, a SNP was identified in the paternally imprinted gene *MEG3* in both H1 and H7 genome (Fig 3.1C). By contrast, SNPs were only detected in one hESC line for the maternally imprinted genes *MEST* (H7 hESCs) and *SNRPN* (H1 hESCs), and the paternally imprinted gene *H19* (H7 hESCs) (Fig 3.1B & C). These data are summarised in Table 3.1.

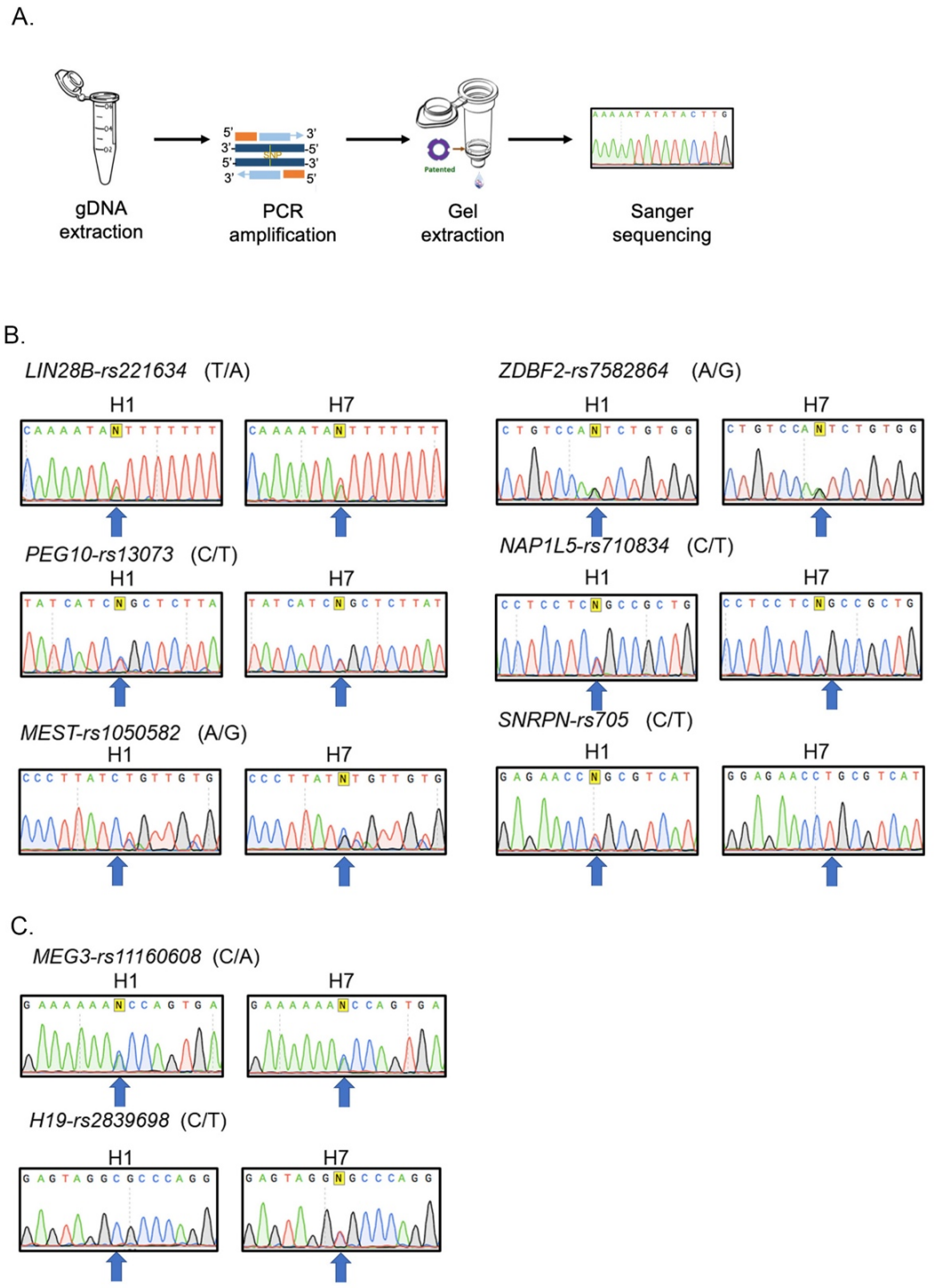

Figure 3.1 SNP identification of imprinted genes in H1 and H7 hESCs.
(A) Schematic description of the process to identify the SNPs in hESCs.
(B & C) Electropherogram of genomic DNA sequencing showing SNP (blue arrow) in maternally imprinted genes (B) and paternally imprinted genes (C) in H1 and H7 hESCs.

Table 3.1 Summary of identified SNPs in H1 and H7 hESCs

	GENE	Location	SNP	H1 hESC	H7 hESC
MATERNALLY IMPRINTED GENES	*LIN28B*	6q16.3-q21	*rs221634*	✔	✔
	PEG10	7q21.3	*rs13073*	✔	✔
	NAP1L5	4q21-q22	*rs710834*	✔	✔
	ZDBF2	2q33.3	*rs7582864*	✔	✔
	MEST	7q32.2	*rs1050582*	-	✔
	SNRPN	15q11.2	*rs705*	✔	-
PATERNALLY IMPRINTED GENES	*MEG3*	14q32.2	*rs11160608*	✔	✔
	H19	11p15.5	*rs2839698*	-	✔

"✔":presence of SNP
"-": no SNP identified

3.2.2 Expression pattern of imprinted genes in hESCs under routine culture condition

Currently, several methods have been developed for the culture of hESCs, which can be summarised into three groups: 1) culture on feeder cells with KSR medium; 2) feeder free culture with MEF-conditioned medium; 3) feeder free culture with commercial mTeSR or E8 media. Previous reports have shown that the majority of human imprinted genes have monoallelic expression in hESCs, supporting the fact that hESCs preserve genomic imprinting (Frost et al., 2011; Pastor et al., 2016). In our laboratory, both H1 and H7 hESCs are routinely cultured with the method of MEF-conditioned medium and propagated by collagenase-assisted mechanical dissociation (Gerrard et al., 2005). Thus, I thought to examine the expression profile of imprinted genes and to ask whether one allele or both alleles are expressed in our cultured hESCs.

Total RNA was collected when cells were nearly confluent in a 6-well plate and cDNA generated by reverse transcription. The allelic expression of imprinted genes can be identified by RT-PCR amplification of the regions containing SNPs using the same primers as described in the previous section with the exception of *SNRPN* (see section 2.2.3ii). Where the Sanger sequencing detected only one nucleotide at the SNP site, it means only one allele, either the paternal allele or maternal allele, was expressed in this gene, confirming that the gene is imprinted. If two different nucleotides were detected at the same site, it means both the paternal and maternal genome were expressed, thus an absence of imprinting. The sequencing results showed that maternally imprinted gene *PEG10*, *NAP1L5, SNRPN* and paternally imprinted gene *MEG3* were monoallelically expressed in hESCs that contain SNPs in these genes (Fig 3.2A & B). This indicates that these genes were all imprinted as expected. Nevertheless, *LIN28B* was shown to be biallelically expressed in both hESC lines, which is not surprising, given that this gene is a trophoblast-specific maternally imprinted gene (Barbaux et al., 2012). *ZDBF2* was expressed from both alleles in H1 and H7 albeit one allele appeared to be expressed more than the other in H1 hESCs, representing absence of *ZDBF2* imprinting in H7 and partial imprinting in H1. (Fig 3.1 A). Another maternally imprinted gene *MEST* and paternally imprinted *H19* were biallelically expressed in H7 hESCs, indicating that neither *MEST* nor *H19* were imprinted in H7 (Fig 3.2A & B).

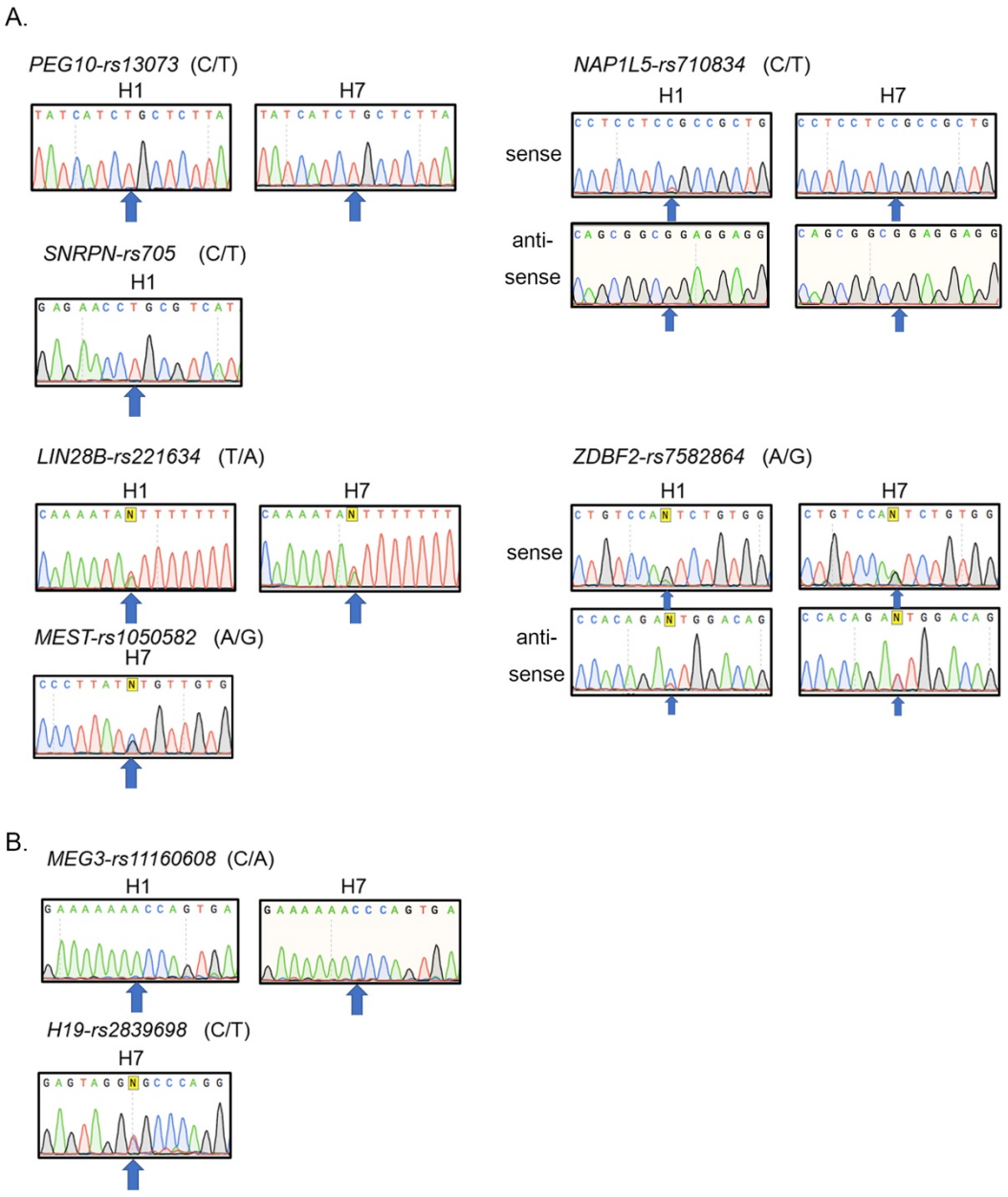

Figure 3.2 Allelic expression of imprinted genes in H1 and H7 hESCs.

Electropherogram of RT-PCR fragments of imprinted genes in hESC lines. The position for the SNPs are indicated by blue arrows. **(A)** Maternally imprinted genes; both sense and antisense strand of *NAP1L5* and *ZDBF2* were sequenced **(B)** Paternally imprinted genes.

The expression of imprinted genes is controlled by differential DNA methylation of CpG dinucleotides in the ICR, so called a DMR (Delaval & Feil, 2004). Thus, I asked whether differential expression of above imprinted genes were also attributed to the different methylation status of their ICR. Two maternally imprinted genes, *PEG10* and *LIN28B,*

and one paternally imprinted gene *MEG3* were selected for this analysis. *PEG10* and *MEG3* are stably imprinted genes, exhibiting monoallelic expression in hESCs, while *LIN28B* is a trophoblast-specific imprinted gene and shows biallelic expression in hESCs (Figure 3.2). Thus, they should show different methylation pattern at their ICRs in hESCs. To examine methylation of the ICR, the gDNA of hESCs was isolated and subjected to bisulphite sequencing with 8-10 clones being sequenced for each DMR, containing 10 CpG dinucleotides.

The sequencing results revealed that in the *LIN28B-DMR,* 97.5% (78/80) and 100% (80/80) CpG dinucleotides were unmethylated in H1 and H7 hESCs, respectively (Figure 3.3A, left), indicating that both alleles are hypomethylated on this gene in hESCs, which is consistent with their active expression in these cells. By contrast, for *PEG10-DMR*, 4 of the 10 clones in H1 and 5 of the 10 clones in H7 showed over 90% methylation on CpG dinucleotides (38/40 and 46/50 for H1 and H7, respectively), while the other ~50% clones revealed at least 95% unmethylation both H1 (57/60) and H7 (48/50) hESCs (Figure 3.3A, right). These data corresponded well with the single allele expression detected in the previous assay (Figure 3.2A), supporting the fact that methylation at this ICR region represses the expression of this gene. For the paternally imprinted gene *MEG3,* imprinting is controlled by two DMRs, the primary *DLK-MEG3* intergenic DMR (*IG-DMR)* and the secondary *MEG3-DMR* (Kagami et al, 2010) (Fig 3.3B, upper panel). Thus, primers were designed to examine both DMRs. Surprisingly, the majority of the CpG dinucleotides in both the *IG-DMR* and *MEG3-DMR* were methylated in both H1 and H7 hESCs (Figure 3.3B lower panel) with 97% (97/100) in the *IG-DMR* of H1 and H7 hESCs and 88% (88/100) and 91% (91/100) in the *MEG3-DMR* of H1 and H7 hESCs, respectively. These results seemed to be contradictory to the paternal imprint of the gene in the hESCs (Figure 3.2B). One explanation could be that only a small section of the DMR was examined in this experiment, which may not represent the methylation status of the full-length *MEG3-ICR* and may not be the region which regulates imprinting in ESCs. This is supported by the work of McMurry and Schmidt (2012), which revealed that in mouse ESCs the upstream region of

MEG3-DMR was hypermethylated on both alleles, while the downstream region of *MEG3-DMR*, only the paternal allele was methylated. Another possibility may be the potential biases in preferential amplification of CG-rich amplicons, but this was unlikely to be the case here, unless it occurred coincidentally with these two pairs of MEG3 primers. Nonetheless, this possibility could be avoided by applying an equimolar mixture of unmethylated and fully-methylated samples of DNA as a proper technical control (Darst et al., 2010).

Taken together (Table 3.2), these results demonstrated that H1 and H7 exhibited a similar expression pattern of the imprinted genes, regardless of their maternal or paternal imprinting status. They also showed similar methylation status in the DMRs of these genes, which is consistent with their expression pattern except for *MEG3*. These results also suggest that most imprinted genes are monoallelically expressed in hESCs and that those which are expressed biallelically could be due to tissue-specific imprinting.

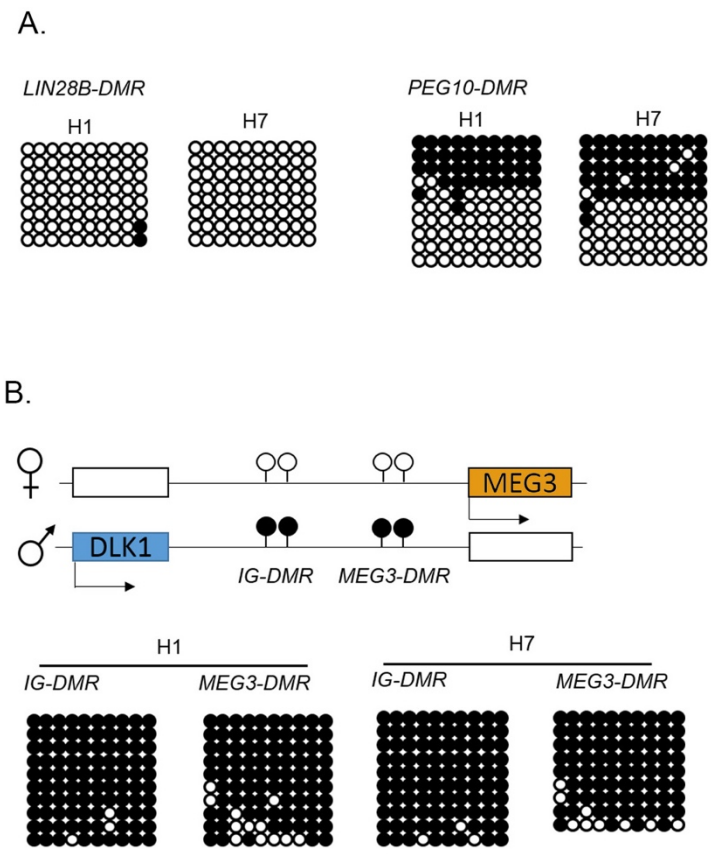

Figure 3.3 the methylation pattern of imprinted genes in H1 and H7 hESCs. **(A)** the CpG island methylation status of the differential methylated region (DMR) of the maternally imprinted genes (*LIN28B* and *PEG10*) were identified by bisulphite sequencing. **(B)** Paternally imprinted gene *MEG3* expression is controlled by two DMR: primary *IG-DMR* and secondary *MEG3-DMR* (upper panel). The CpG island methylation status on *IG-DMR* and *MEG3-DMR* were identified by bisulphite sequencing (lower panel). Open and closed circles represent unmethylated and methylated CpGs respectively.

Table 3.2 Summary of imprinting status of genes in H1 and H7 hESCs

	GENES	Location	H1 hESC	H7 hESC	DMR-Met Status
MATERNALLY IMPRINTED GENES	*LIN28B*	6q16.3-q21	✘	✘	Hypomethylated
	PEG10	7q21.3	✔	✔	Hemi-methylated
	NAP1L5	4q21-q22	✔	✔	
	ZDBF2	2q33.3	partial	✘	
	MEST	7q32.2	-	✘	
	SNRPN	15q11.2	✔	-	
PATERNALLY IMPRINTED GENES	*H19*	11p15.5	-	✘	
	MEG3	14q32.2	✔	✔	Hypermethylated

"✔" : monoallelic expression

"✘" :biallelic expression

"Partial": one allele expressed more than another

"-": no SNP identified

3.2.3 Conversion of hESCs to naïve state using 5iLAF culture condition

Compared with routinely cultured hESCs, naïve hESCs have more similarities with the cells in the ICM, in terms of their morphology, gene expression and epigenetic features (Nichols and Smith, 2009). 5iLAF (MEKi, GSK3i, SRCi, bRAFi, ROCKi, LIF, Activin, and bFGF) is a commonly used and well-studied naïve conversion method for hESCs. However, naïve hESCs resulting from this conversion method have shown not only decreased global DNA methylation but also erasure of DNA methylation at imprinted regions (Pastor et al., 2016).

Currently, all reported naïve conversion of hESCs with 5iLAF has been carried out on MEF-feeders. However, in our lab, hESCs are routinely cultured in feeder-free conditions with MEF-CM. Thus, I initially tested whether naïve conversion was feasible under a feeder-free condition. Since the 5iLAF method uses N2B27 as its base medium, I tried to generate MEF-conditioned N2B27 medium for the conversion, which was similar to that of MEF-CM for hESCs. Basically, N2B27 medium was added into culture vessels containing iMEF and then collected after 24h incubation. This MEF-conditioned N2B27 medium supplemented with 5iLAF was used for the naïve conversion of hESC in a matrigel-coated plate as illustrated in Fig 3.4A. Under these conditions massive cell death was observed within 24 hours. Although domed colonies emerged after 11 days, the colonies grew very slowly and almost all cells were dead after 3 passages (Fig. 3.4B). Gene expression analysis of these domed colonies showed that the expression of pluripotency markers (OCT4 and NANOG) was dramatically decreased, even though the expression of naïve markers, REX1, KLF2 and TCL1B were slightly increased (Fig. 3.4C). These data indicate that these cells have lost their pluripotency and suggest that MEF feeders may be integral for naïve cell conversion, as they can provide an undefined and complex mixture of growth factors and extracellular matrix (Wang et al., 2014).

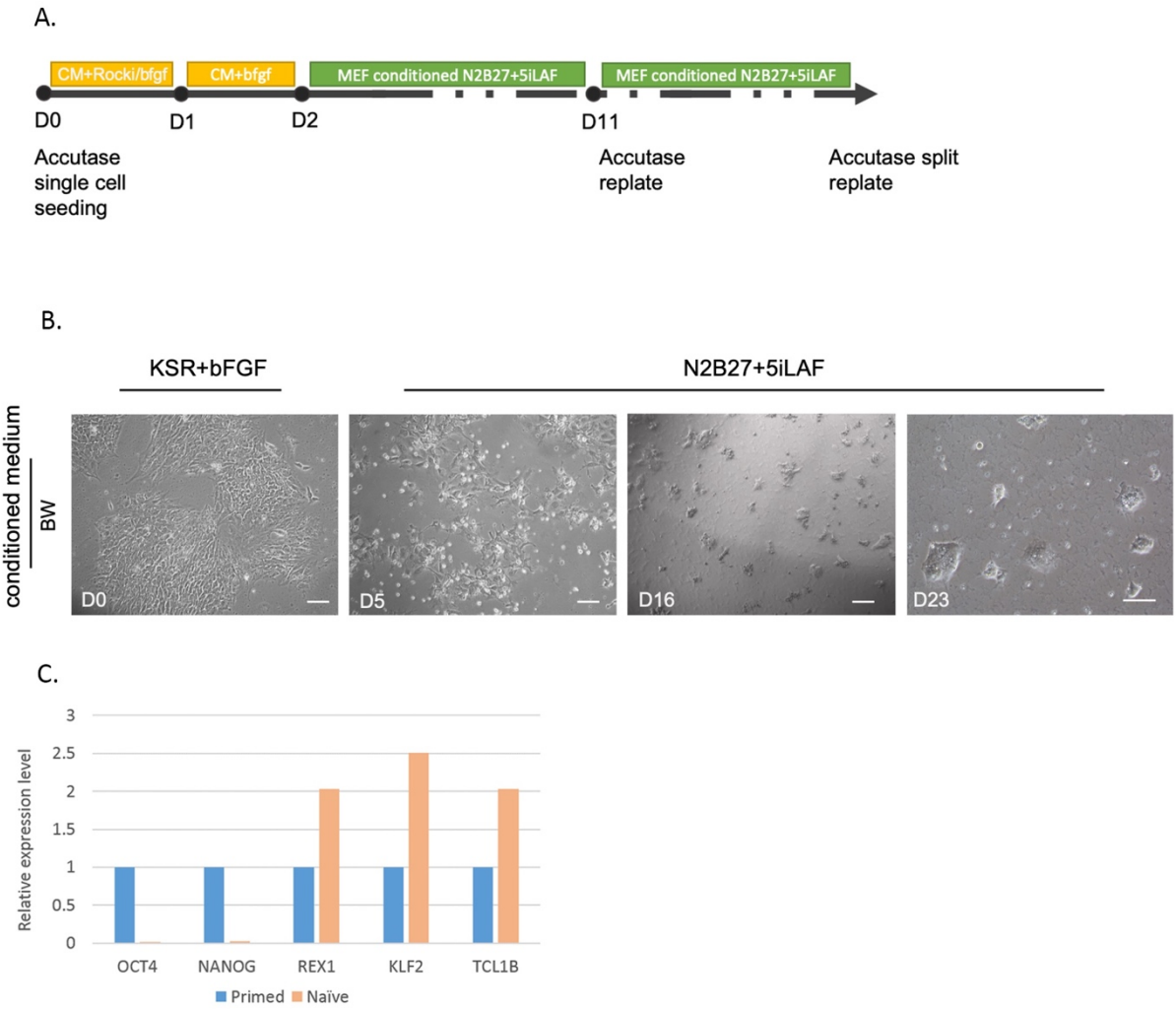

Figure 3.4 Conversion of hESC to naïve state without MEF feeders.
(A) Diagram depicting the experimental process to generate naïve hESC under feeder-free condition. **(B)** Phase-contrast images showing morphological changes of hESCs during naïve conversion with days in conversion indicated (Scale bar, 100 μm). **(C)** qRT-PCR analysis of mRNA expression in pluripotency (OCT4 and NANOG) and naïve markers (REX1, KLF2 and TCL1B) after day 23 of naïve conversion. Data are presented as relative expression to hESCs (primed) (n=1).

Since the trial of feeder-free naïve conversion was unsuccessful, MEF feeders were employed thereafter for the naïve conversion. For direct and quick assessment of naïve conversion, an OCT4-GFP reporter hESC line (T7) was employed, in which GFP expression is driven by the 4 kb upstream region of the human *POU5F1* gene containing distal enhancer, proximal enhancer and promoter (Gerrard et al., 2005). To

prepare for the naïve conversion, $2x10^5$ iMEFs were seeded on a gelatin-coated plate one day before single cell-dissociated hESCs were plated onto MEF feeders via Accutase-Rocki split. Two days later, medium of MEF-CM supplemented with bFGF was replaced by N2B27 supplemented with 5iLAF (Fig 3.5A). Under these conditions, a small number of dome-shaped colonies emerged around day 12 after widespread cell death in the first two days, which was similar to that observed in feeder-free conversion. However, these colonies could be propagated and retained similar if not higher, levels of OCT4-GFP expression (Fig. 3.5B), indicating that OCT4 was still expressed at least at similar levels after conversion. Furthermore, the colonies showed increased mRNA expression in pluripotency (OCT4, NANOG and DPPAs) and naïve markers (REX1, KLF2, TCL1B and PRDM14), the increases being particularly significant for NANOG and REX1 (Fig. 3.5C). In addition, genes that are highly expressed in the ICM, such as DPPA3 and DPPA5, also exhibited significant elevation after naïve conversion (Fig. 3.5D) (Tang et al, 2010). These results indicate the successful conversion of hESCs to the naïve state using this method.

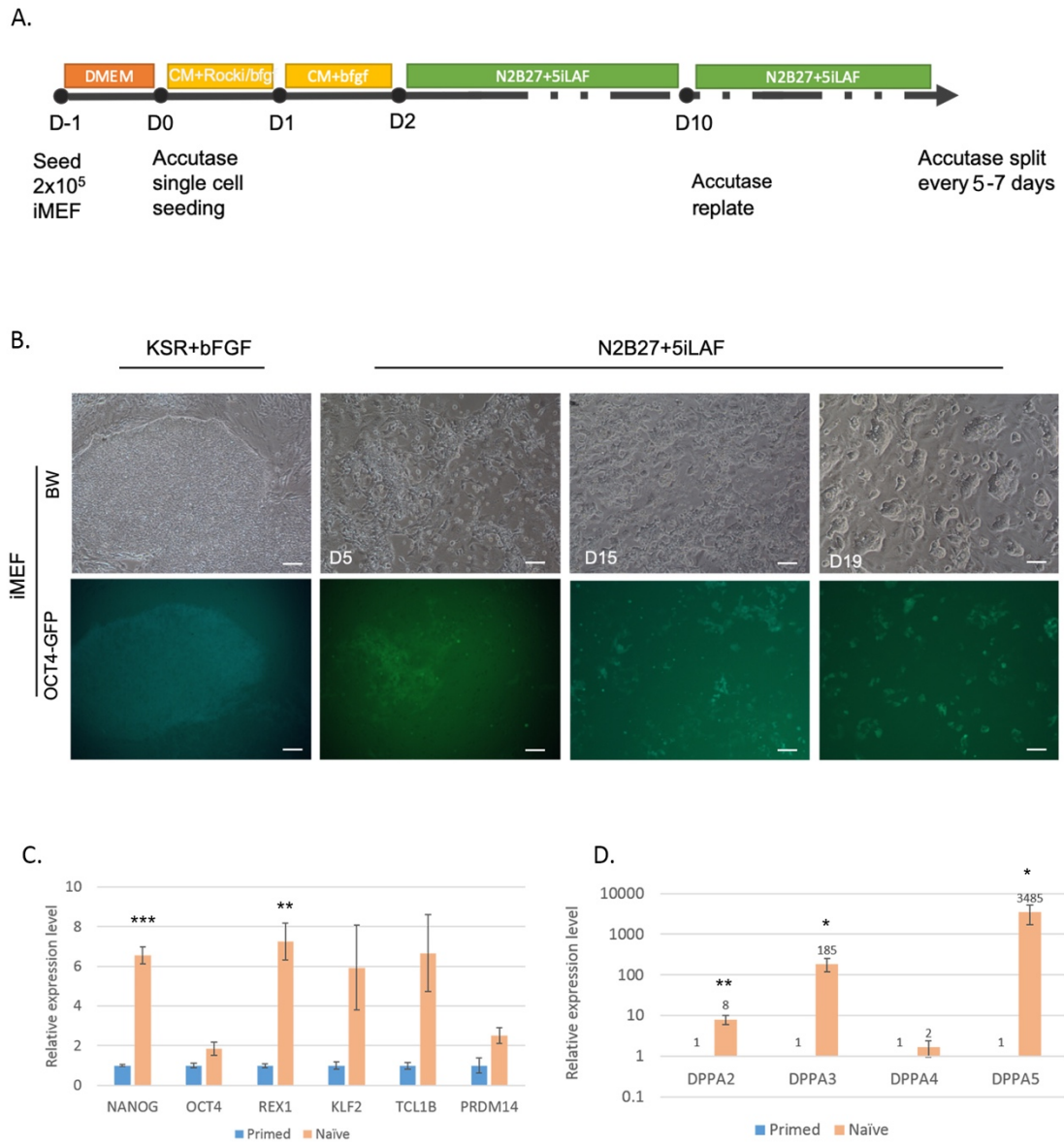

Figure 3.5 Conversion of hESC to naïve state with MEF feeders.
(A) Experimental process to generate naïve hESC with iMEF feeder cells. **(B)** Morphological change of hESC during naïve conversion (Scale bar, 100 μm). **(C).** RT-qPCR analysis of pluripotency and naïve markers in naïve converted hESC with iMEF feeders. **(D)** RT-qPCR analysis of developmental pluripotency associated genes (DPPAs) in naïve converted hESC (*p<0.05,**p<0.01, ***p < 0.001 by two-tail unpaired t-test, n=3).

3.2.4. Imprinted gene expression after naïve conversion

Next, I undertook to analyse whether there are any changes in imprinting status, since previous reports claimed loss of imprinting after naïve conversion of hESCs by 5iLAF (Pastor et al., 2016). Allelic expression of the imprinted genes containing SNPs in our hESCs were analysed after naïve conversion by direct Sanger sequencing following RT-PCR as previously described. All maternally and paternally imprinted genes examined (*LIN28B*, *PEG10*, *ZDBF2, NAP1L5* and *MEG3*) exhibited biallelic expression, suggesting that those genes that were monoallelically expressed before the conversion had lost their imprint after the conversion (Fig 3.6). Correspondingly, their methylation status was also changed in naïve cells. For maternally imprinted genes, *PEG10* had its hypermethylated clones reduced from 40% (4/10) in primed state to 11% (1/9) after naïve conversion; *NAP1L5* methylation rate dropped from 69% (62/90) in primed cells to 35% (28/80) in naïve cells. Paternally imprinted *MEG3* also revealed a significant loss of methylation in all 10 clones with 97% (97/100) of CpGs being hypermethylated in primed cells to only 1/9 clones being hypermethylated with a total of 12% (11/90) CpGs methylated in all clones after conversion (Fig. 3.6). These data confirm the previous finding that naïve conversion by 5iLAF indeed erases all the imprinting, possibly through demethylation of the majority of the CpGs in the ICR.

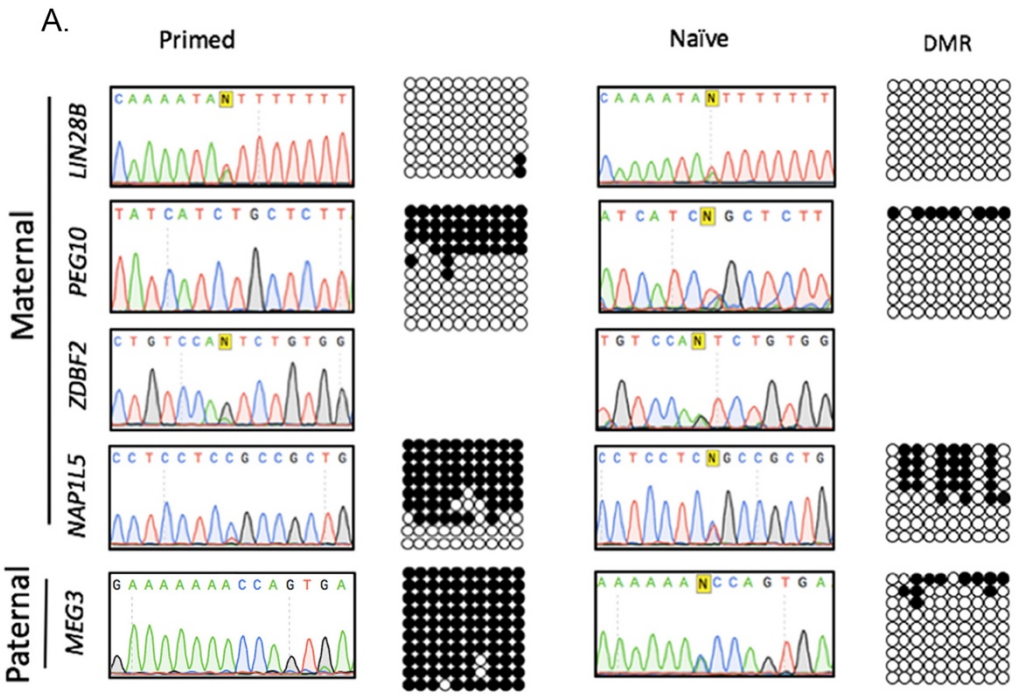

Figure 3.6 Allelic expression and methylation status of imprinted genes in primed and naïve hESCs cultured with iMEF feeders.

Comparison of allelic expression and DMR methylation pattern of the maternal imprinted genes (*LIN28B*, *PEG10*, *ZDBF2* and *NAP1L5*) and paternal imprinted gene (*MEG3*) between primed and naïve hESCs. Open and closed circles represent unmethylated and methylated CpGs respectively.

3.2.5 Naïve conversion alters the mRNA expression of DNMTs, TETs and SCMC members.

Genomic imprinting can be established through differential DNA methylation of CpG islands in the regulatory regions of imprinted genes (Reik and Walter, 2001). This CpG DNA methylation is catalysed by DNA methyltransferases (DNMTs): DNMT1, 3A, 3B and 3L (Okano et al., 1999; Kato et al., 2007; Suetake et al., 2004). Methylation of CpGs can also be erased through either passive demethylation, e.g. DNA replication dependent dilution or active demethylation by TET dioxygenase-mediated oxidation and DNA glycosylase (TDG)-dependent base excision repair (Wu and Zhang, 2017) (see section 1.2.2. for detailed review). Thus, examining the expression of DNMTs and TETs in naïve cells could provide information about the mechanisms underlying the DNA methylation change during hESC naïve conversion.

In both H1 and H7 hESCs, DNMT1, the methyltransferase for hemimethylated DNA, had a lower expression level than *de novo* methyltransferases DNMT3A and DNMT3B, while the *de novo* methyltransferase DNMT3L exhibited the lowest expression (Fig 3.7A). This pattern is consistent with the RNA-Seq data of H1 hESCs (Wu et al., 2010). Comparing the mRNA expression between primed and naïve hESCs, it revealed that DNMT1 had an around a 2-fold and 1.5-fold increase in H1 and H7 naïve cells respectively. The *de novo* DNMT3A did not change significantly in H1, but had a 3-fold increase in H7; and DNMT3B exhibited a slight decrease in H1, while showing a 1.5-fold increase in H7 after the naïve conversion (Fig 3.7A). Nonetheless, the relative expression pattern of these 3 DNMTs did not change, with DNMT1 still lower than DNMT3A and 3B. Remarkably, DNMT3L exhibited a more than 5000-fold increase in both H1 and H7 naïve cells, from the lowest DNMT in primed state leaping to the highest expressed one in naïve state (Fig. 3.7A).

For enzymes involved in active DNA demethylation, TET1 had much higher expression level (~10-fold) than TET2 and TET3 in primed state of both H1 and H7 hESCs. After

naïve conversion, TET1 and TET3 both showed approximately 3 to 4-fold increase in both types of naïve cells, while TET2 had the biggest increase (more than 30-fold) albeit still lower than TET1 (Fig. 3.7 B). Nevertheless, their overall expression pattern remained the same before and after naïve conversion, with TET1 still the highest expressed. Therefore, almost all the DNMT and TET enzymes showed increased expression after naïve conversion, which was particularly dramatic for DNMT3L and TET2. These changes may play a role in the turnover of epigenetic profiles in naïve hESCs.

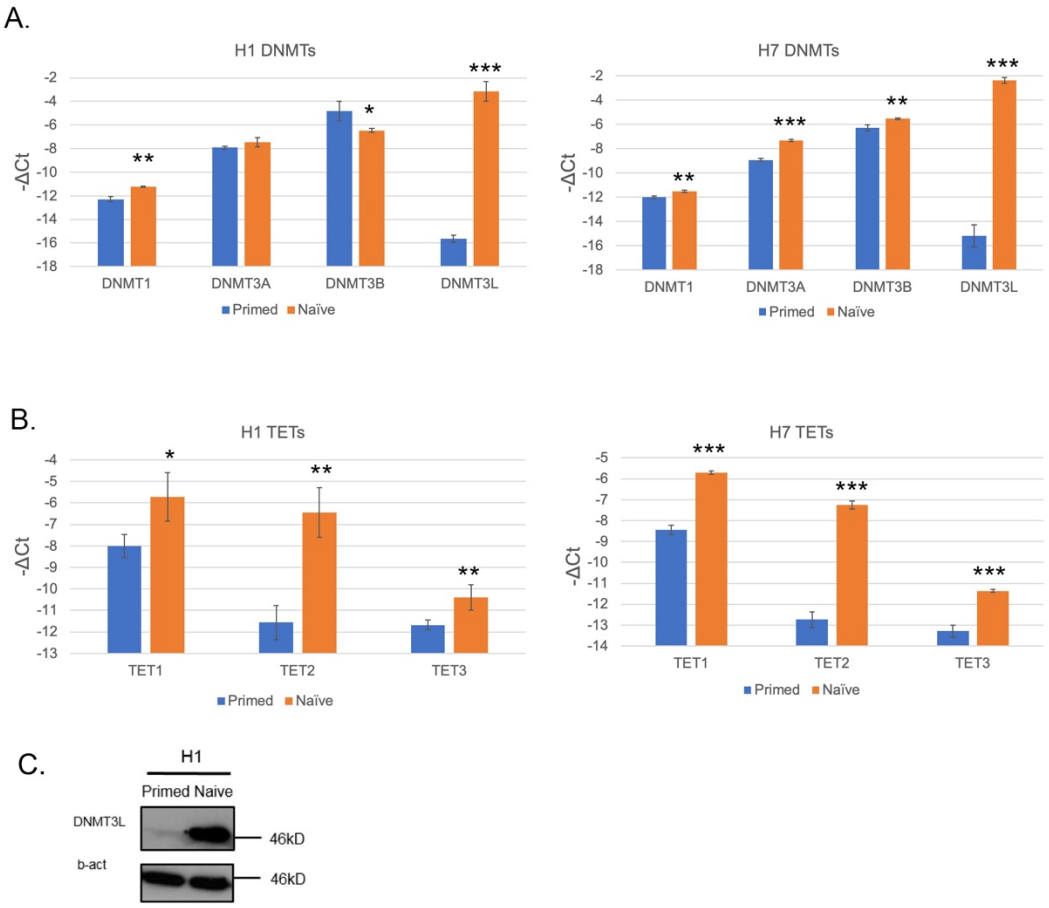

Figure 3.7 Expression analysis for imprinting-related enzymes in primed and naïve hESCs.

Comparison of DNMT1, 3A, 3B, 3L **(A)** and TET1, 2, 3 **(B)** mRNA expression between primed and naïve hESCs in both H1 and H7 hESC line via qRT-PCR; Data are presented as mean ± SD (n=3) of -ΔCt= $\log_2$(expression level relative to beta-actin). (H7 qPCR was done by Xiaolun Ma) (*p<0.05, **p<0.01 and ***p < 0.001 by two-tail unpaired *t*-test). **(C)**. Western blot analysis on DNMT3L in H1 primed and naïve hESCs.

Since NLRP7 has been hypothesized to be a member of the SCMC and this complex is considered to be involved in the establishment and maintenance of genomic imprinting during early embryogenesis (Monk et al., 2017), I also examined the mRNA expression of all known and potential members of the SCMC complex. In hESCs, TLE6, NLRP7 and NLRP2 showed similar and relatively high expression levels, while PADI6, NLRP5, KHDC3L and OOEP had a very low expression. Noticeably, almost all known members of SCMC were shown to have significantly increased their mRNA expression after naïve conversion in both H1 and H7 hESCs. NLRP2, NLRP7, KHDC3L and OOEP all exhibited more than 5000-fold increased expression, although NLRP5 and PADI6 only showed approximately a 30-fold increase. Only TLE6 showed downregulation of almost 80% and 30% in H1 and H7 naïve hESCs, respectively. (Fig 3.8 A). Given DNMT3L is the most dramatically elevated DNMT and shows similar level of increase compared with some SCMC members, it is interesting to examine whether this increase also occurs at the protein level. Western blotting confirmed this dramatic upregulation of NLRP7 and DNMT3L at protein levels after hESCs were converted from primed to naïve state, which is consistent with the change of their mRNA expression (Fig 3.7 B and 3.8 B).

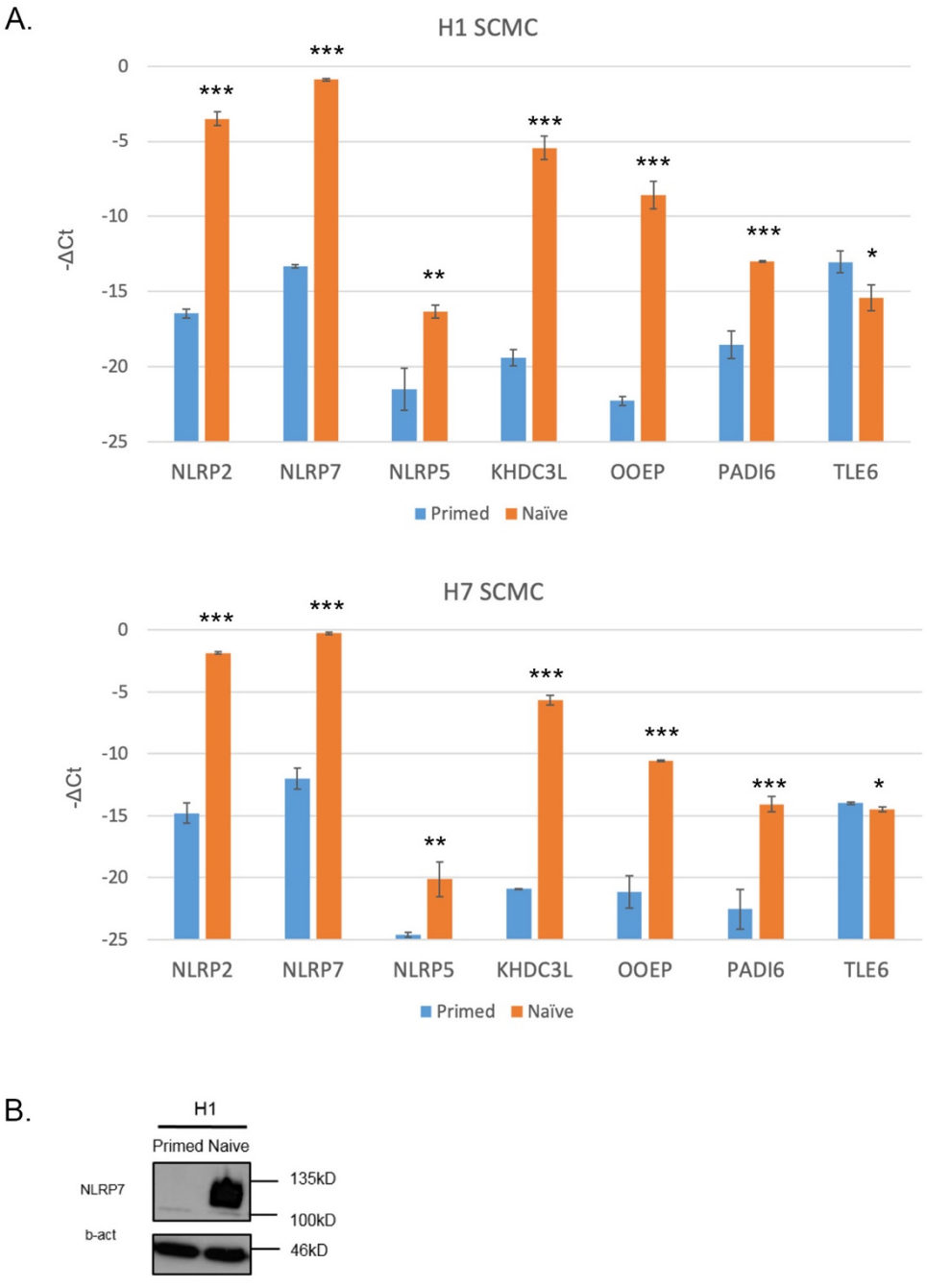

Figure 3.8 Expression analysis of SCMC members in primed and naïve hESCs. **(A)** Comparison of NLRP7, NLRP2, OOEP, KHDC3L, PADI6, NLRP5 and TLE6 mRNA expression between primed and naïve hESCs in both H1 and H7 hESC line via qRT-PCR. Data are presented as mean ± SD (n=3) of -ΔCt = $\log_2$(expression level relative to beta-actin) (H7 qPCR was done by Xiaolun Ma). (*p<0.05, **p<0.01 and ***p < 0.001, by two-tail unpaired t-test) **(B)** Western blot analysis on NLRP7 in H1 primed and naïve hESCs.

3.3 Conclusion and Discussion

In this chapter, I have successfully identified SNPs in selected genes for which imprinting has been shown to be lost in molar tissues of women with NLRP7 recessive mutations and subsequently, determined the imprinting status in both hESC lines. This revealed that all the imprinted genes examined had similar allelic expression patterns in H1 and H7 hESCs and that the methylation pattern of their DMR corresponded to their expression. Most of those genes showed the expected imprinting and methylation status.

Although SNPs were identified in most imprinted genes (*LIN28B, PEG10, NAP1L5, ZDBF2* and *MEG3*) in both H1 and H7 hESCs, *MEST* and *H19* SNPs were only identified in H7 while *SNRPN* SNP was only identified in H1. This is not surprising since H1 and H7 are hESCs derived from embryos of different individuals with unique genetic identities (Abeyta et al, 2004) (Table 3.1). These two cell lines revealed very similar expression pattern for those imprinted genes, which mostly showed monoallelic expression as expected, indicating that these genes are imprinted in the hESCs. In addition, the trophoblast-specific maternally imprinted gene, *LIN28B*, showed biallelic expression as expected. Three genes, maternally imprinted genes *ZDBF2*, *MEST* and paternal imprinted gene *H19*, exhibited aberrant loss of imprinting in our hESCs. Among them, biallelic expression of *MEST* and *H19* was only tested in H7 hESC line due to lack of SNPs in the other line (Table 3.2). The loss of allele-specific expression of some imprinted genes in hESCs has also been reported in several previous studies, implicating epigenetic instability. For example, *H19* and *MEG3* were shown to be monoallelically expressed in most hESC lines, but they did occasionally have biallelic expression, whereas *MEST* frequently exhibited biallelic expression in hESCs (Rugg-Gunn et al., 2007). The exact cause of the epigenetic instability is still unknown, even though several possible mechanisms have been proposed. Firstly, all the hESCs are derived from surplus embryos of IVF clinic and IVF procedures may result in epigenetic instability. A) Superovulation: oocytes in IVF are often obtained by superovulation,

which has been shown to contribute to the instability of maternal *Mest* methylation in mouse embryos (Velker et al., 2017); B) *In vitro* culture: all the fertilization and early embryo development in the IVF clinic are performed in culture dishes, where cells experience more environmental stress compared with the cells *in vivo,* which could lead to epigenetic changes, including imprinted genes (Ventura-Juncá et al., 2015); C) The patients seeking IVF treatment may already have some epigenetic problems causing infertility. These points are supported by the evidence of increased risk of epigenetic disorders in ART children, such as Beckwith-Wiedemann syndrome, Angelman syndrome, Prader-Willi syndrome, and Silver-Russell syndrome (Hattori et la., 2019). Secondly, the hESCs were derived by the expansion of individual undifferentiated colonies without clonal selection (Thomson et al., 1998), thereby loss of imprinting stability in a few single cells due to environmental stress may lead to a growth advantage to take over the population. For example, *IGF2*, a paternally expressed growth factor sharing an ICR with *H19*, is important for hESC proliferation and survival. Increased *IGF2* biallelic expression due to due to imprinting instability may result in a selective advantage during routine culture (Adewumi et al., 2007). Nonetheless, the majority of the imprinted genes retained the expected imprinting status in the hESCs of our culture.

Interestingly, after naïve conversion by 5iLAF, all the imprinted genes tested lost their imprinted status even though the converted cells acquired several characteristics of naïve ESCs, such as morphology, transcriptional profile and epigenetic profile, including global hypomethylation. hESCs were flattened large colonies, with high nuclear-cytoplasmic ratio, while naïve hESCs appeared to be small dome-shaped colonies, resembling mouse ESCs (De Los Angeles et al., 2012) (Fig 3.5B). The gene expression analysis also revealed an elevated expression of naïve markers, such as REX1, NANOG, KLF2, TCL1B and PRDM14, as well as pluripotency genes associated with early embryos (*OCT4*, *DPPAs*) (Fig3.5 C & D). The increase of NANOG and KLF2 seems to be important for successful naïve conversion, because only a short-term overexpression of *NANOG* and *KLF2* transgene in hESCs is sufficient to trigger their

transition from primed to naïve state in 2iL/PKCi condition (Takashima et al., 2014). This loss of imprinting status by 5iLAF-induced naïve conversion of hESCs is consistent with the previous report (Fig 3.6) (Pastor et al., 2016), which clearly showed the global hypomethylated genome in naïve hESCs, similar to human preimplantation blastocyst cells, except loss of genomic imprinting.

Although it remains unclear why the naïve conversion leads to the loss of imprinting, it was found by studying the components in the 5iLAF medium that many of them may have a positive role in DNA demethylation, leading to global DNA hypomethylation in naïve hESCs. The MEK inhibitor, PD0325901, a component of 5iLAF, has been reported to enhance the expression of *Prdm14* in mESCs (Li et al., 2016). Prdm14 can promote active demethylation via recruiting Tet1 and Tet2 at targeted genes and impair de novo methylation via downregulating Dnmt3a/3b (Okashita et al., 2014; Leitch et al., 2013). Indeed, my results also showed that PRDM14 was increased in naïve cells (Fig 3.5C). Furthermore, GSK3 inhibitor, another component of 5iLAF, may also result in the loss of imprinting, as double deletion of *Gsk-3a* and *Gsk-3b* in mESC reduced the expression of Dnmt3a2, leading to hypomethylation at ICR of imprinted genes (Popkie et al., 2010). Moreover, vitamin A in B27 supplement may also enhance the mRNA expression of TET2 by binding to its retinoic acid receptor element (RARE) and may contribute to the active erasure of DNA methylation during naïve conversion (Hore et al., 2016).

Indeed, my results showed that after naïve conversion, all three TET genes were elevated in naïve cells, especially TET2 (Fig 3.7B). However, a previous study in mESCs indicated that deletion of Tet genes in primed mESCs did not affect the demethylation dynamics during transition to naïve pluripotency in 2iL condition and it was proposed that the major contributor to the global demethylation is the impaired maintenance DNA methylation (von Meyenn et al., 2016). Regarding DNA methyltransferases, all the catalytic active DNMTs, DNMT1, 3A and 3B mRNA exhibited some subtle changes after naïve conversion (Fig 3.7A), while the protein

level of DNMT1 and its functional partner UHRF1 both showed significant reductions, of ~ 50%, in naïve hESCs (Pastor et al., 2016). This may cause the reduced maintenance of global DNA methylation. The most dramatic alteration of DNMT mRNA expression came from the catalytic inactive DNMT3L, which revealed five thousand-fold increase in naïve cells (Fig 3.7A). DNMT3L plays an essential role in the regulation of gene imprinting in gametes and early embryos of mice (Hata et al., 2002), but its role in imprinting in humans is not clear. In terms of the genes may associated with global DNA demethylation, these dramatic changes in the expression of DNMTs and TETs in my study are also generally consistent with the published RNA-seq data (Pastor et al., 2016).

This raises a question whether naïve hESCs could be used as a model resembling preimplantation human ICM. Although naïve hESCs can recapitulate many aspects of ICM cells *in vivo*, they still have some differences in terms of their imprinting status. 5iLAF converted naïve cells clearly showed loss of imprinting; and this similar phenomenon is also observed in 2iL-cultured mESCs (Yagi et al., 2017, Choi et al., 2017). Despite the differences in genomic imprinting, the global demethylation process during naïve conversion may still enables the simulation of post-zygotic demethylation *in vivo*, so hESCs may be a suitable model to study the role of NLRP7 on genomic imprinting during global demethylation.

As a maternal protein complex, SCMC is abundantly expressed in oocytes and preimplantation embryos in human, and indicated to be related to imprinting establishment and maintenance during *de novo* methylation and demethylation (Poli et al., 2015; Demond et al., 2019). By contrast, my results show that SCMC members, including NLRP7, were expressed at a very low level in hESCs, and only become highly expressed in naïve hESCs, except TLE6 (Fig 3.8A). This raises a question whether the low level of NLRP7 expression may be associated with the loss of maternal imprinting resulting from global demethylation during naïve conversion. Even though NLRP7 was dramatically upregulated after naïve conversion, this upregulation

might occur after global demethylation. Therefore, further studies are required to determine the dynamic change of NLRP7 during naïve conversion in relation to the global DNA methylation level, and investigate whether high level of maternal NLRP7 may protect maternally imprintied genes from 5iLAF-indcued global demethylation.

Taken together, this chapter confirmed that both H1 and H7 hESCs have correct imprinting state for the majority of imprinted genes, which were examined via SNPs, indicating that hESCs are a good model to study genomic imprinting. Additionally, conversion of hESCs to naïve state has resulted in a dramatic change in epigenetic and gene expression profile. The erasure of global DNA methylation by naïve conversion may be used to simulate the global demethylation during early embryonic development. Additionally, the high expression level of DNMT3L and SCMC members, including NLRP7, in naïve hESCs may also give us a chance to study their relationships.

Chapter 4.

Investigating the role of NLRP7 in maternal imprinting during naïve conversion of hESCs

4.1 Introduction

Genomic imprinting is established by de novo methylation during gametogenesis after the demethylation when the PGCs first emerge. The imprinted differentially methylated regions inherited from parental gametes are protected from another wave of global demethylation post-fertilization, resulting in parental-allele-specific methylation of these elements (Smallwood & Kelsey, 2012). Although it is not clear what role NLRP7 plays in the establishment and maintenance of maternal imprinting, NLRP7 has been shown to be dynamically expressed during oogenesis and development of preimplantation human embryos (Zhang et al., 2008; Yan et al., 2013; Petropoulos et al., 2016). Particularly, NLRP7 protein is abundantly expressed in human early embryos from the zygote to the 8-cell stage (Akoury et al., 2015). Thus, it raised a question whether the high level of maternal NLRP7 protein in the fertilised zygote protects maternally imprinted genes from the global demethylation after fertilization, when the majority of the CpG nucleotides become demethylated.

Given that NLRP7 mRNA and protein are both shown to be expressed at low levels in hESCs (Chapter 3), which is different from the high NLRP7 protein levels in the oocytes and early embryos, a question is raised whether this might lead to a loss of imprinting, particularly maternal imprinting, during naïve conversion when global demethylation occurs. Although NLRP7 ultimately becomes highly expressed after naïve conversion, this upregulation might occur simultaneously or after the demethylation of the maternally imprinted genes. As a consequence, even though the naïve hESCs largely resemble the ICM cells regarding their transcriptional and epigenetic profiles, they differ from the ICM cells by surprisingly losing the genomic imprinting. Therefore, I hypothesized that a high level of NLRP7 might be able to protect the maternally imprinted genes from global demethylation.

To test this hypothesis and investigate whether NLRP7 can protect maternal imprinting during naïve conversion, this chapter is going to focus on the following objectives:

i. To examine dynamic expression of NLRP7 during naïve conversion, in particular the onset of its upregulation in this process.

ii. To clone NLRP7 cDNA and establish hESCs that overexpress NLRP7.

iii. To investigate whether NLRP7-overexpression in hESCs has any protective role for maternally imprinted genes from demethylation upon naïve conversion.

4.2 Results

4.2.1 Characterising NLRP7 expression during naïve conversion

In the previous chapter, I demonstrated a dramatic increase of NLRP7 in the naïve hESCs after naïve conversion. However, the dynamic progression of this increase and its relationship with the global demethylation during the naïve conversion was unknown. Thus, it is interesting to investigate when this change of NLRP7 expression started to appear during naïve conversion and how rapidly it occurred.

To gather the information, H1 hESCs were collected every two days for further analysis after replacing the MEF-CM medium with the N2B27 naïve conversion medium supplemented with 5iLAF (Fig 4.1A). Cell morphology was continuously monitored as one of the parameters of successful naïve conversion and the collection of cells was stopped when the dome-shaped colony clearly visible on the 10^{th} day of the conversion (Fig 4.1B). In addition to morphological changes, the increase of REX1 and NANOG expression is also used as another important indicator of naïve conversion. REX1 mRNA was only slightly decreased in the first 4 days ($p > 0.05$), but significantly upregulated to 2.5 fold thereafter; NANOG expression showed an approximately 2-fold increase from day 2 and persisted to day 6, then was rapidly increased to around 5 fold by day 10 (Fig 4.1C). Interestingly, NLRP7 mRNA initially exhibited a gradual increase from the beginning of naïve conversion, showing 5-fold increase on day 2 and 32-fold increase on day 4 ($p < 0.05$). However, the expression was exponentially elevated after day 6 of the conversion, becoming more than 5000-fold at day 10 (Fig 4.1D). Given that NLRP7 expression is very low in hESCs (Fig 3.8A), it remained at a low level in the first few days of naïve conversion despite increasing several-fold. Thus, a functional level of NLRP7 may only be reached at a later stage of the naïve conversion, probably after global demethylation has taken place. This suggests that NLRP7 would not be able to protect maternally imprinted genes from this demethylation if it had such a function. Therefore, in order to test my hypothesis, I

intended to overexpress NLRP7 in hESCs and investigate the effect of its overexpression on maternally imprinted genes upon naïve conversion.

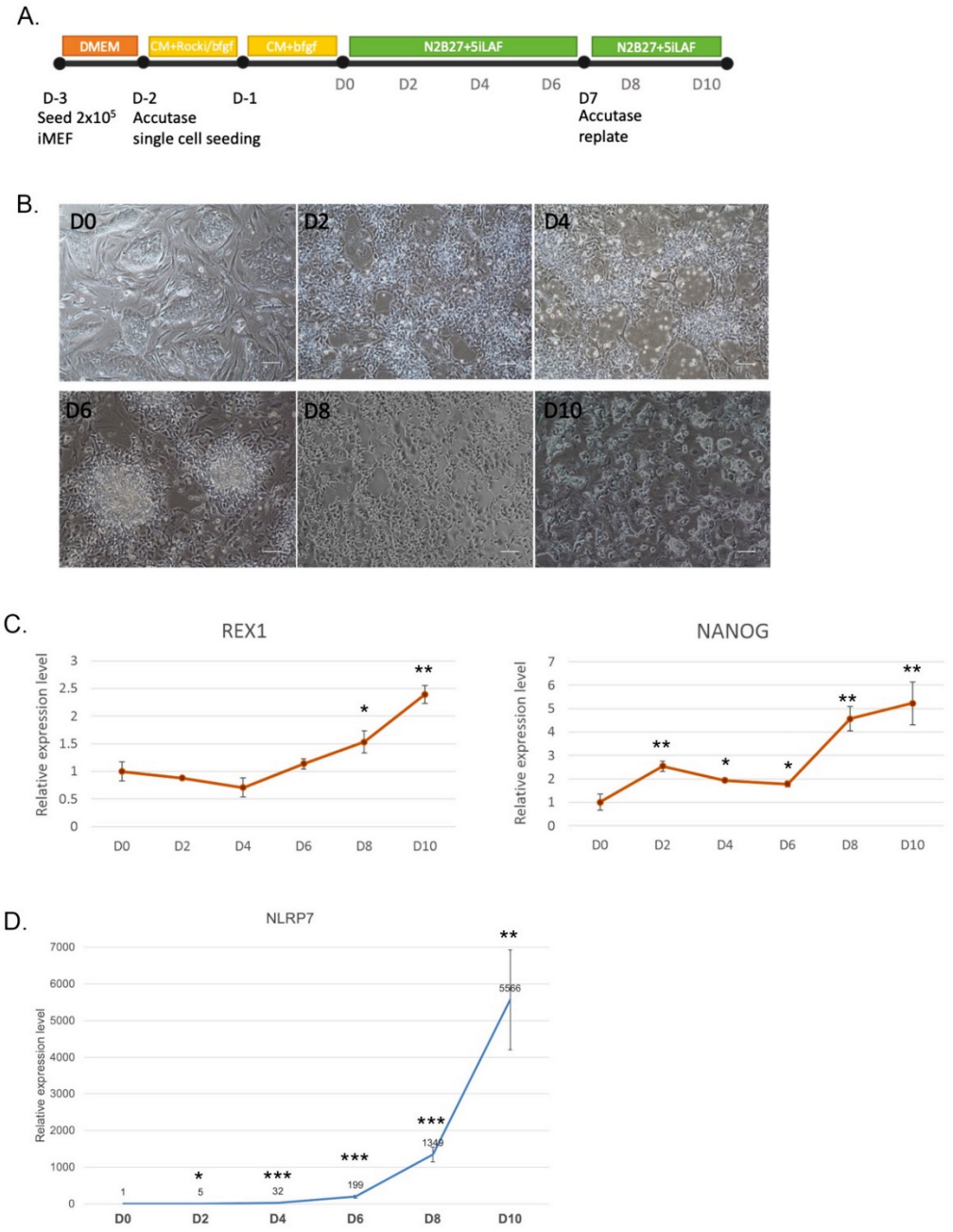

Figure 4.1 Monitoring NLRP7 expression during naïve conversion.
(A) The experimental process to generate naïve hESCs (H1). **(B)** Representative phase-contrast images showing the morphological changes of hESCs during the naïve conversion (Scale bar, 100 μm). **(C)** mRNA expression of the naïve makers, REX1 and NANOG, during naïve conversion by RT-qPCR analysis. **(D)** RT-qPCR analysis of NLRP7 mRNA expression during naïve conversion. In all the RT-qPCR, data are presented as mean ± SD from three independent conversion experiments (*p<0.05,**p<0.01, ***p < 0.001 by two-tail unpaired t-test, n=3)

4.2.2 Isolation and cloning of NLRP7 cDNA

In order to overexpress NLRP7 in hESCs, it is necessary to obtain NLRP7 cDNA first. I selected hESCs for the isolation and cloning of NLRP7 cDNA despite the low expression levels as NLRP7 is almost undetectable in other cell types available in our lab, such as HEK293T cells and PC3 cells. The NLRP7 cDNA coding region was amplified by RT-PCR from H7 hESCs using primers designed from the start codon to the stop codon. The PCR product was checked in agarose gel, which showed a single band with the predicted size (~3.1kb) (Fig 4.2B). The PCR product was then purified, A-tailed and ligated into the pGEMT-easy cloning vector (pGEMT easy-NLRP7) (Fig 4.2A). Five colonies were picked from the plate and amplified, then the plasmid DNA was cut by restriction enzyme PvuII to check if NLRP7 cDNA had been successfully inserted. The restriction digest results gave the expected number of bands, but the size of the second bands varied. In clone 2, 4 and 5, the second band was slightly larger in size than that in clone 1 and 3 (Fig 4.2C), thereby clone 2 and clone 3 were sent for sequencing to determine the differences between them. The sequencing results revealed that clone 2 contains a 171 bp fragment at the 3' end of the mRNA, which is missing from clone 3 (Fig 4.2D). The sequences were subjected to BLAST in the NCBI database, which showed that they represent two different NLRP7 isoforms: clone 3 is isoform 2, while clone 2 is isoform 3. Isoform 2 lacks a 171 bp (57 amino acids) exon 10 compared with isoform 3. However, there are 3 NLRP7 isoforms in total. The undetected isoform 1 is 84 bp shorter than isoform 3 at the 5' end of exon 5 via using an alternative splice site (Fig 4.2E).

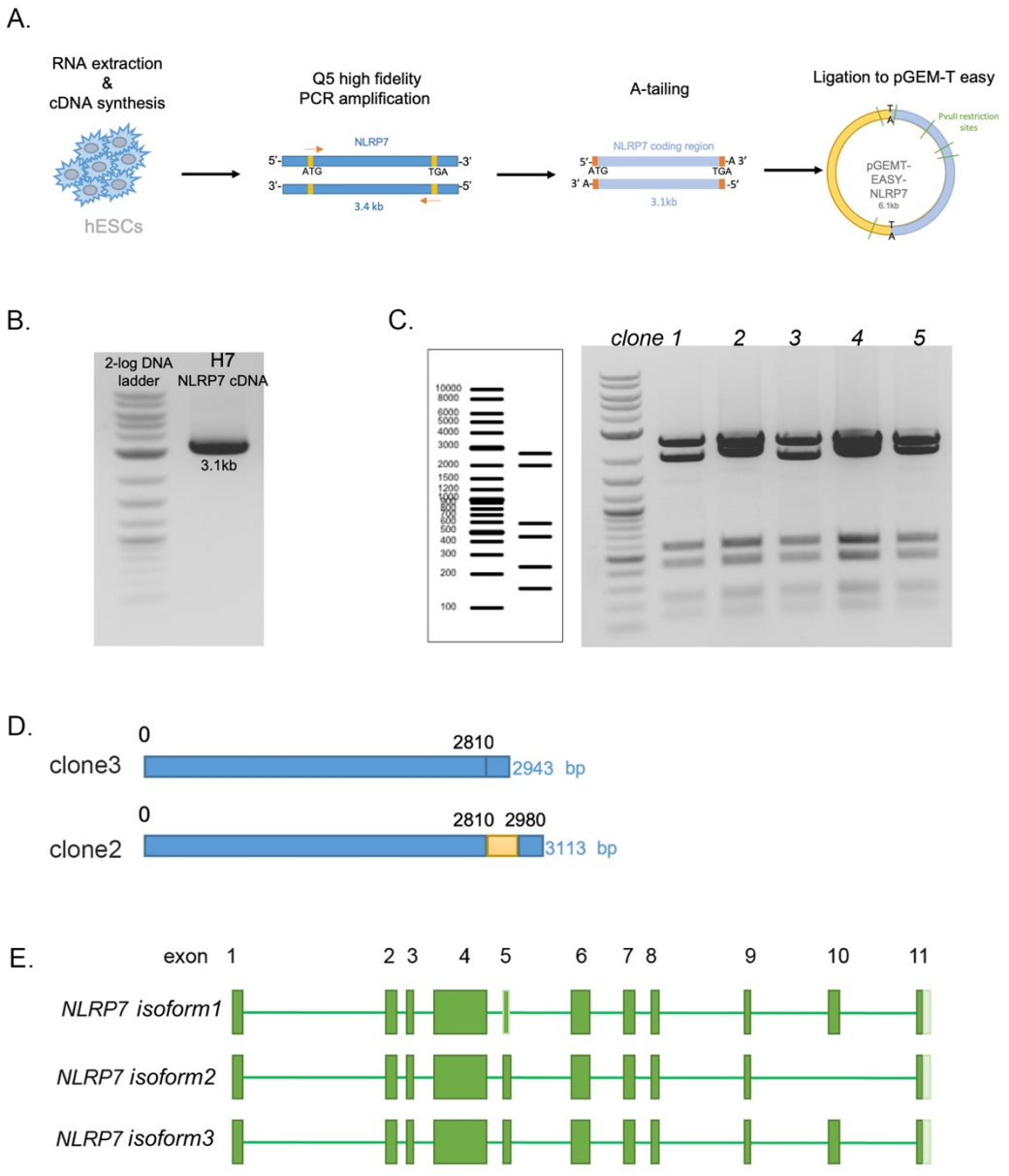

Figure 4.2 Isolation and cloning of NLRP7 cDNA

(A) Experimental process of NLRP7 cloning; dark blue bars represent template DNA; yellow bars represent start and stop codons; orange arrows represent primers and orange bars are primer sequences which are amplified in NLRP7 cDNA; light blue bars represent newly synthesized NLRP7 coding region; green lines indicate PvuII cutting sites on pGEMT-easy-NLRP7 plasmid. **(B)** Gel electrophoresis of NLRP7 cDNA PCR product amplified by Q5 high-fidelity DNA polymerases. **(C)** Gel electrophoresis of plasmids digested by PvuII. The image on the left showed the expected number and size of bands after PvuII digestion. **(D)** Schematic diagram showing sequencing results of clone 2 and 3 with the 171 bp insert in clone 2 highlighted in the yellow box. **(E)** Diagram depicting three isoforms of NLRP7. The boxes represent exons while lines represent introns. The numbers above each box indicate the exon number.

In order to estimate the expression ratio of isoform 2 to 3 in both hESC lines, two primers (exon10-F and exon10-R) were designed flanking exon 10 for the semi-quantitative PCR (Fig 4.3A). Surprisingly, no NLRP7 isoforms containing exon 10 was detected in H1 cDNA, while in H7 cDNA the ratio of NLRP7 isoform with and without exon 10 was about 1 to 3.8 (Fig 4.3B). To preclude the bias of PCR to amplify smaller fragment that results in the absence or low expression of NLRP7 isoform 1 and 3 detected in hESCs, another primer inside exon 10 (exon10-mid-R) was designed to specifically amplify exon 10 (Fig 4.3A). The result confirmed that H1 hESCs do not express NLRP7 isoform 1 or 3 containing exon 10 (Fig 4.3C). It is also likely that in H7, NLRP7 isoform 2 was more abundantly expressed than isoform 3 (Fig 4.3B). Nonetheless, these results show that H1 and H7 hESCs have different preferences for NLRP7 alternative splicing although it is unknown whether this might be due to the gender difference of the two hESC lines, since H1 hESCs contain chromosomes XY while H7 hESCs have XX chromosomes, or whether the three NLRP7 isoforms have different functions.

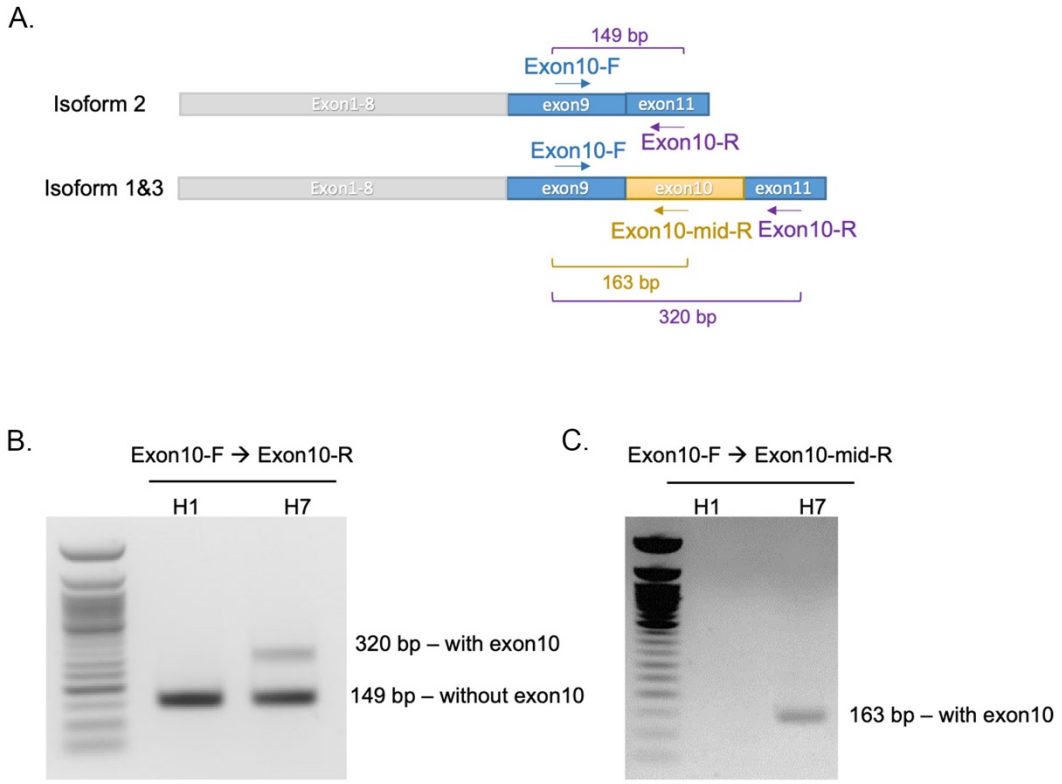

Figure 4.3 Different isoforms of NLRP7 expressed in H1 and H7 hESCs.

(A) Diagram demostrating the position of different designed primers (Exon10-F, Exon10-R and Exon10-mid-R) and the size of their amplification products in different NLRP7 isoforms. **(B)** Semi-quantitative PCR of the cDNA of H1 and H7 hESCs to determine the ratio of transcripts with and without NLRP7 exon 10. **(C)** Semi-quantitative PCR of the cDNA of H1 and H7 hESCs to determine the presence of exon 10.

4.2.3. Establishment of hESC line that overexpress NLRP7 cDNA.

Given the high expression of NLRP7 isoform 2 in both H1 and H7 hESCs and its mutations associated with BiCHM, I decided to construct a vector to express isoform 2 of NLRP7. The NLRP7-isoform 2 (NLRP7 thereafter) cDNA fragment was digested by FseI and NotI from pGEMT-easy-NLRP7 cloning vector after the insertion of an FseI restriction site at the 5' end via site directed mutagenesis. Then the cDNA fragment was subcloned into pCAG-Puro-2A-HA-Sox2 plasmid cut with the same restriction enzymes, to replace SOX2, so that the NLRP7 cDNA would be in-frame downstream of the Puro-HA to generate the pCAG-Puro-2A-HA-NLRP7 expression construction (Fig 4.4 A).

The resulting vector was initially introduced into H1 hESCs by lipofectamine transfection and NLRP7 expression was examined 48 h post-transfection via western blotting in the absence of puromycin selection (Fig 4.4B). The results showed a clear expression of the HA-NLRP7 band of the expected size around 110 kD and a non-specific band detected by the NLRP7 antibody at 100kD (Fig 4.4C), confirming the successful transient expression of NLRP7 protein and the viability of NLRP7 antibody.

However, after puromycin selection for one week (Fig 4.4B), NLRP7 antibody could not detect any transgene expression in the stably transfected hESCs despite a faint band of HA-NLRP7 detected by HA antibody on the western blotting, which might be due to 3 HA peptides attached upstream of NLRP7 transgene, This result implicated a much lower expression of transgene in stable-transfected cell line than that of transient transfection within 48h (Fig 4.4 C). The low efficiency of generating stable clones is probably due to the fact that circular plasmids will be randomly linearized after transfection and the transgene may be destroyed by linearization (Stuchbury and Münch, 2010). NLRP7 is not a very small protein, with cDNA of more than 3 kb and the whole pCAG-Puro-2A-HA-NLRP7 promotor-transgene fragment size is almost 6 kb, so there is a high chance that the plasmid was cut at the transgene fragment.

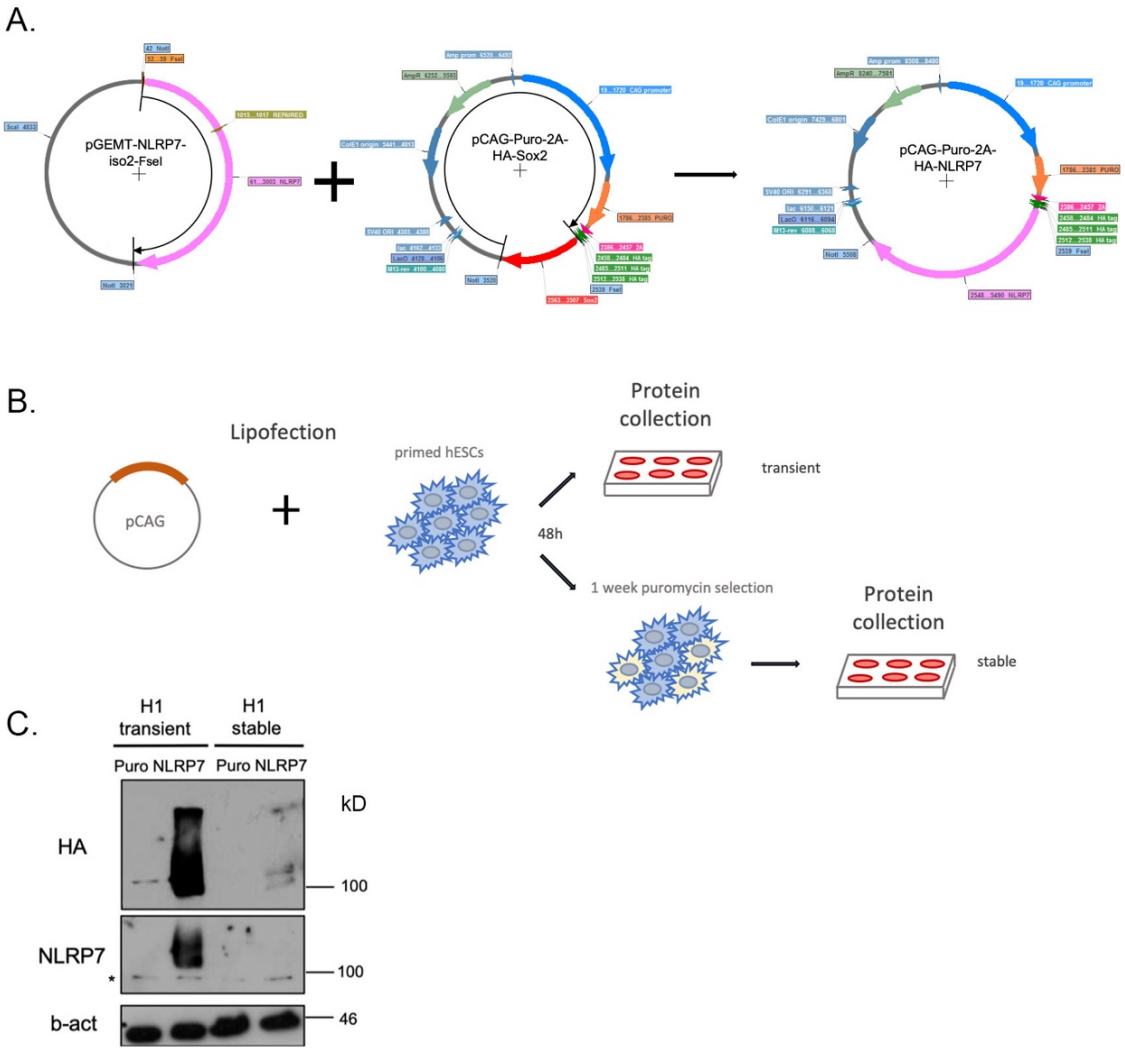

Figure 4.4 NLRP7 overexpression in hESCs via lipofectamine transfection.

(A). Generation of NLRP7 expression vector via recombination of pGEMTeasy-NLRP7- iso2-FseI and pCAG-Puro-2A-HA-Sox2. **(B)**. Experimental process of transient and stable NLRP7-OE in primed hESCs. **(C)**. Western blot analysis comparing NLRP7-OE and control hESCs (H1) after transient and stable transfection. * represents the position of non-specific bands at 100 kD.

Due to the large size of the transgene to be integrated into the host genome, it proved difficult to get the intact fragment inserted via lipofectamine transfection, thus lentiviral transduction was then employed in this experiment. NLRP7 lentiviral expression vector (pLVTHM-CAG-Puro-2A-NLRP7) was constructed via generating two cutting sites on pCAG-Puro-2A-HA-NLRP7 and pLVTHM-CAG-Puro-2A-TRF2 by SpeI (Fig4.5A). The recombinant lentivector with correct orientation was selected and amplified thereafter. The lentivector was co-expressed with envelope vector (VSV-G plasmid) and packaging vector (Δ8.91 plasmid) in HEK293T cells to produce lentivirus, which was then cleaned up and concentrated to directly transduce hESCs in single cell suspension (Fig4.5B). After puromycin selection, only ~10% NLRP7-tranduced hESCs survived, which is much lower than the puro-control group (approximately 40% survival rate) (result not shown), which might also be due to the larger size of the NLRP7 plasmid than the control plasmid. According to Addgene lentiviral guidance, the maximum cloning capacity between long terminal repeats (LTRs) is about 8.5 kb and inserts larger than 3 kb will have less packaging efficiency, while our NLRP7 lentivector possessed an 8.2 kb insert. Immunoblotting of the NLRP7-transduced cell lysate detected the HA-NLRP7 transgene by both NLRP7 and HA antibodies at 110 kD, albeit the NLRP7 antibody only showed a weak band, indicating that HA-NLRP7 was ectopically expressed at detectable level after lentiviral transduction and selection (Fig 4.5C).

The NLRP7-OE hESCs exhibit the same cell morphology as control and parental hESCs, with a large nuclei-cytoplasmic ratio, tightly compacted in the center of the colonies and loose on the periphery. The expression of pluripotent genes is similar to the control and parental hESCs as well (Fig 4.6B left). Therefore, the results demonstrate that NLRP7-overexpressed (NLRP7-OE) hESC lines have been successfully established and that ectopic expression of NLRP7 in hESCs does not show a clear effect on cell morphology.

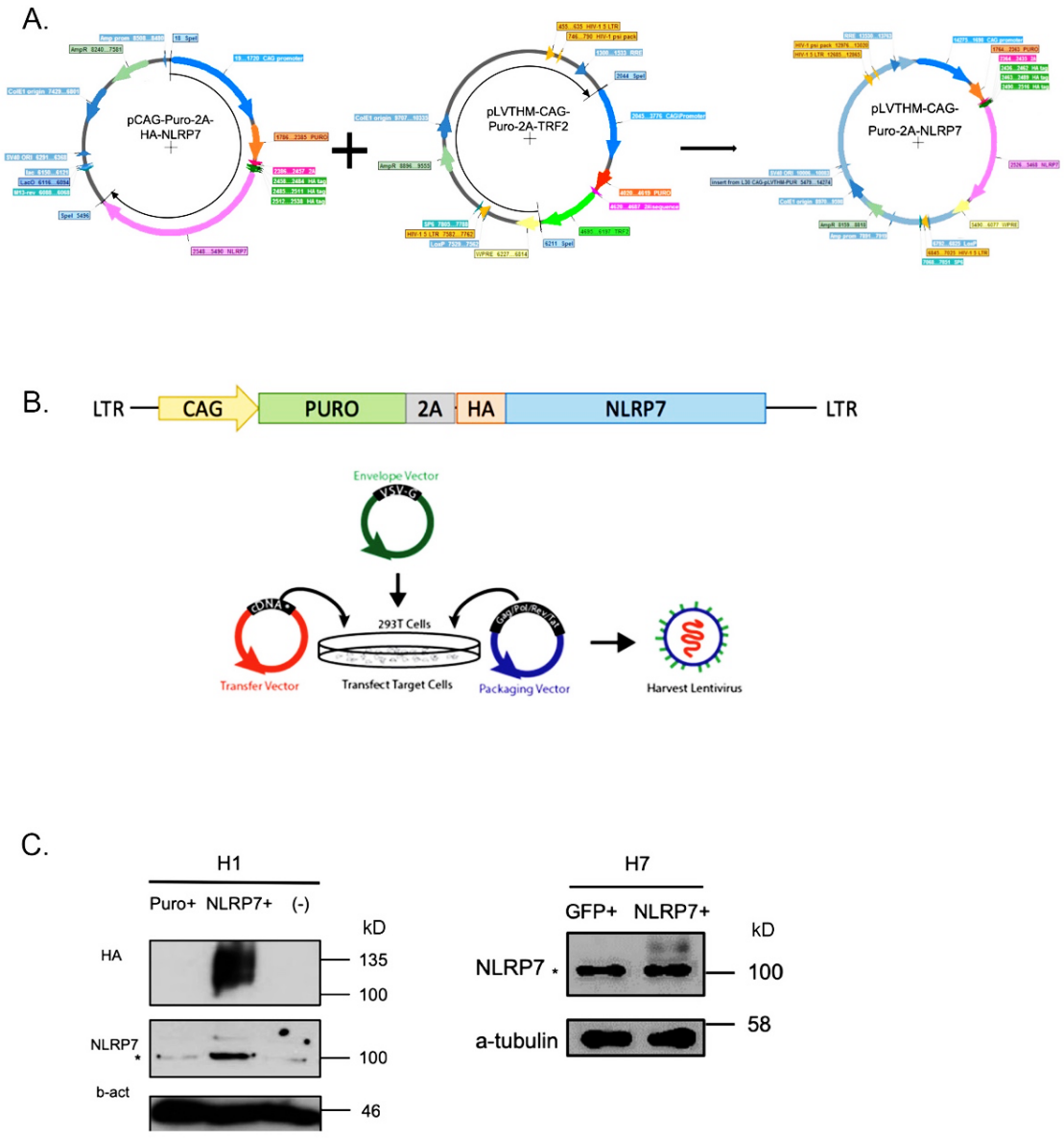

Figure 4.5 NLRP7 overexpression in hESCs via lentiviral transduction.

(A). Generation of NLRP7 lentivector via recombination of pCAG-Puro-2A-HA-NLRP7 and pLVTHM-CAG-Puro-2A-TRF2. **(B)**. Schematic figure showing the construct carried between pLVTHM lentivrial LTR and the experimental process demonstrating the production of lentivirus particles via HEK293T host cells. **(C).** Western blot analysis comparing NLRP7-OE and control hESCs (H1 and H7) after lentiviral stable integration. * represents the position of non-specific bands at 100 kD.

4.2.4 NLRP7-overexpression partially protected maternally imprinted genes in the naïve conversion

To investigate whether promoting the NLRP7 expression level in primed hESCs can help the protection of maternal imprinting during the global demethylation process of naïve conversion, NLRP7-overexpressed (NLRP7-OE) hESCs were converted to the naïve state via the 5iLAF method. Before naïve conversion, NLRP7-OE hESCs (H7) clearly showed a more than 70-fold increase in NLRP7 mRNA expression compared with the control hESCs (Fig 4.6A). Similar to control hESCs, the NLRP7-OE hESCs exhibited dome-shaped colonies within two passages of naïve conversion (Fig 4.6B). Furthermore, wild-type (WT) and NLRP7-OE hESCs also showed similar level of increase in the expression of naïve markers (REX1 and NANOG) after naïve conversion, indicating that the NLRP7-OE hESCs had been successfully converted to the naïve state (Fig 4.6C).

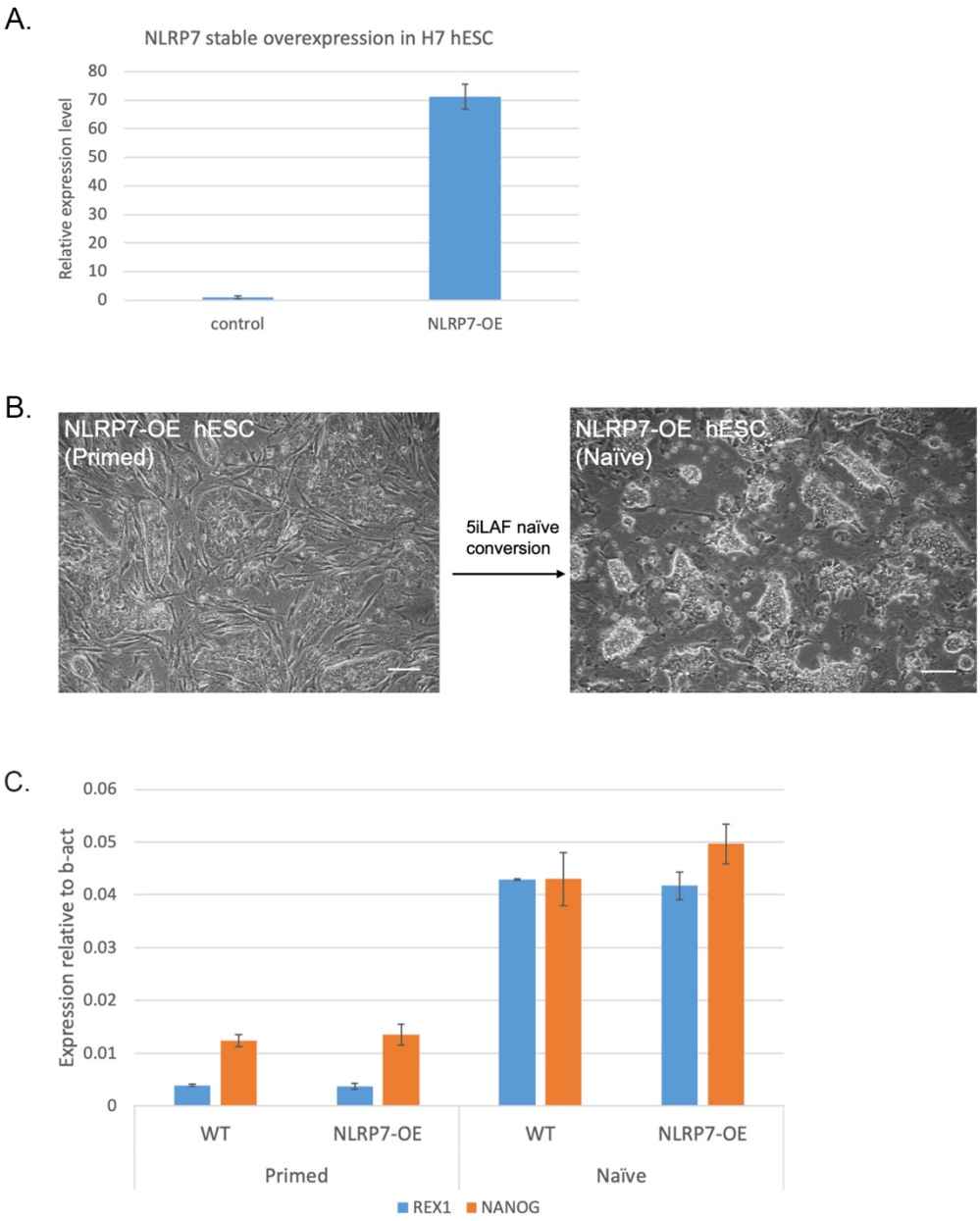

Figure 4.6 Naïve conversion of NLRP7-overexpressed (NLRP7-OE) hESCs.

(A) mRNA expression of NLRP7 in control and NLRP7-OE hESCs (H7). **(B)** Phase contrast images of NLRP7-OE hESCs before and after 5iLAF naïve conversion (scale bar, 100 μm). **(C)** Comparison of REX1 and NANOG mRNA expression between wild-type and NLRP7-OE in hESCs before and after naïve conversion (n=3).

To explore whether NLRP7-overexpression would affect the imprinting status of the primed hESCs before I proceeded to naïve conversion, allelic expression pattern of the selected imprinted genes were analyzed in WT and NLRP7-OE hESCs. In primed state, the imprinting status of none of the tested imprinted genes, *LIN28B*, *PEG10*, *NAP1L5* and *MEG3*, was influenced by NLRP7 overexpression. For example, maternally imprinted *PEG10*, *NAP1L5* and paternally imprinted *MEG3* were still monoallelically expressed while trophoblast-specific maternally imprinted gene *LIN28B* retained biallelic expression (Fig 4.7A left).

To answer the question whether NLRP7 overexpression can protect the loss of maternal imprinting after naïve conversion, the imprinting status of naïve converted NLRP7-OE cells were investigated. Interestingly, the ubiquitous maternally imprinted genes, *PEG10* and *NAP1L5* that exhibit the loss of maternally imprinted status after naïve conversion in the parental cell line, showed partial retention of their imprinted status, manifested by the apparent unequal expression of two alleles. In *PEG10*, paternally expressed “T” was the only nucleotide detected at SNP position in primed hESCs but became equal levels of “T” and “C” after naïve conversion, indicating an equal expression of both maternal and paternal alleles. Whereas, with overexpression of NLRP7 the “T” was expressed at a much higher level than the “C”. A similar situation was also identified at *NAP1L5*, the paternally expressed “C” was higher than maternally expressed “T” in NLRP7-OE hESCs after naïve conversion (Fig 4.7A right). By contrast, NLRP7-OE exhibited no effect on the paternally imprinted gene *MEG3,* which still revealed a change from monoallelic in primed hESCs to equal biallelic expression after naïve conversion, the same as in WT naïve hESCs. Additionally, NLRP7-OE did not show any effect on trophoblast-specific maternally imprinted gene *LIN28B,* which showed biallelic expression in both primed and naïve hESCs. The unequal expression of imprinted genes is specific to ubiquitous maternally imprinted *PEG10* and *NAP1L5*, rather than the paternally imprinted *MEG3*, which indicates these results are unlikely to be a technical artifact. However, since this experiment was only done once in one hESC line, more experiments are required to validate these results.

If the results could be repeated in more cell lines, they would suggest that a high level of NLRP7 may have a protective role on maternally imprinted genes from de novo DNA demethylation.

To further investigate whether different methylation pattern on the ICR account for the different allelic expression of the maternally imprinted genes, bisulfite sequencing was used to identify the methylation status of the ICR in the maternally imprinted *PEG10*. As stated in Chapter 3, the DMR of *PEG10* was hemimethylated in the primed hESCs (Fig. 3.3), but lost its methylation after naïve conversion and became hypomethylated (Fig 3.6). *PEG10*-DMR in NLRP7-OE hESCs was hemimethylated (3 out of 8 clones) at primed state, similar to the control hESCs. However, its methylation rate remained at around 50% (5 out of 10 clones) after naïve conversion (Fig 4.7B), which was different from the almost complete loss of DNA methylation identified in WT naïve hESCs, but consistent to the partial protection of *PEG10* maternal imprinting status. Additionally, the methylation rate of another maternally imprinted gene *NAP1L5* also remained at about 50% (5 out of 9 clones) after naïve conversion, despite its methylation status in primed NLRP7-OE cells was not determined due to insufficient clones. Nevertheless, previous result obtained in primed WT hESCs showed NAP1L5-DMR was indeed hemi-methylated (Fig 3.6). All the results of bishulphite analyses on DMRs in different cell lines (H1 and H7) and different conditions (Naïve/Primed or WT/NLRP7-OE) were summarized in Table 4.1. Here, only 8 to 10 clones were randomly picked up after bisulfite conversion, which could result in this slight discrepancy in DMR methylation between primed and naïve states. Picking more clones can help to better estimate the accurate methylation ratio on DMR. These results suggest that NLRP7 protein might have a role to protect maternally imprinted genes from global demethylation, thus retaining their imprinting status. This is only a preliminary experiment and more experiments are required to validate this.

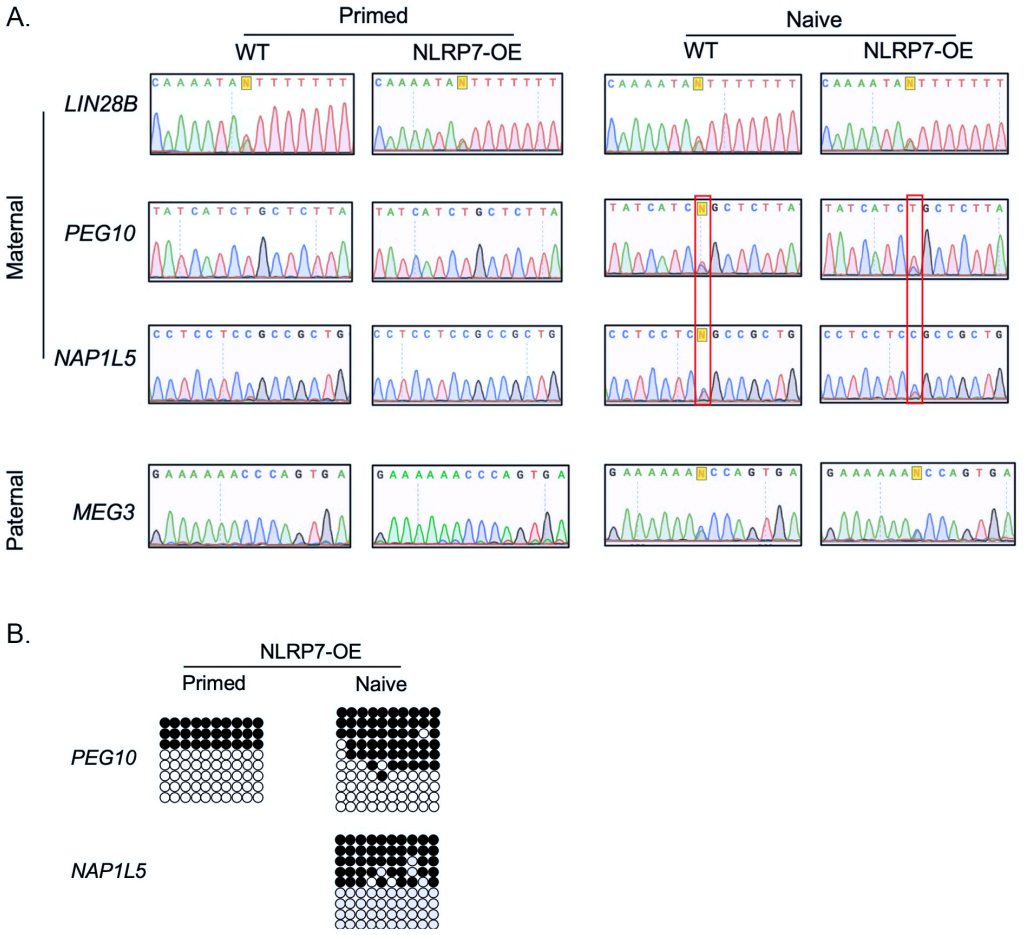

Figure 4.7 Imprinting status of wild-type and NLRP7-overxpressed hESC (H7) before and after naïve conversion.

(A) Allelic expression of maternally (*LIN28B*, *PEG10* and *NAP1L5*) and paternally (*MEG3*) imprinted genes in wild-type and NLRP7-overxpressed hESC before and after naïve conversion. **(B)** the CpG island methylation status on the DMR of the maternally imprinted gene *PEG10* and *NAP1L5* in NLRP7-overxpressed hESC before and after naïve conversion.

Table 4.1 Summary of bisulphite analyses on DMRs in this project.

	Wild type			***NLRP7-OE***	
	Primed		**Naïve**	**Primed**	**Naïve**
Cell line	**H1**	**H7**	**H1**	**H7**	**H7**
LIN28B	Hypomethylated (2.5%)	Hypomethylated (0%)	Hypomethylated (0%)	N/A	N/A
PEG10	Hemimethylated (41%)	Hemimethylated (48%)	Hypomethylated (11%)	Hemimethylated (38%)	Hemimethylated (55%)
NAP1L5	Hemimethylated (69%)	Hemimethylated (62%)	Hypomethylated (35%)	N/A	Hemimethylated (49%)
MEG3 IG DMR	Hypermethylated (97%)	Hypermethylated (88%)	Hypomethylated (12%)	N/A	N/A

NOTE: the methylation level of DMR is indicated by CpG methylation percentage.

N/A: bisulphite sequencing not performed or failed due to technical issues.

4.3 Conclusion and Discussion

In this chapter, I have confirmed that, as an important indicator of naïve pluripotency, NANOG had an over 2-fold upregulation within 2 days after changing to 5iLAF condition and eventually increased to 5-fold when naïve conversion was almost completed (Fig 4.1C). Since NANOG mRNA is already expressed at an easily detectable level (~1% of β-actin) in hESCs (Fig 4.6B), a few folds increase could make a considerable functional differences. Nanog is a master transcription factor for pluripotency and has an important function in regulating epigenetic configuration. A recent study has found that Nanog interacts with Tet1 and Tet2 to potentiate global 5-hydroxymethylcytosine (5hmC) to assist the iPSC reprograming in mouse (Costa et al., 2013). Thus, the prompt upregulation of NANOG upon naïve conversion may trigger the global demethylation in the first place to induce the expression of other naïve associated factors. Another important naïve marker REX1 was only significantly upregulated from D8 (Fig 4.1C), and might be triggered by the global demethylation at the later stage of naïve conversion.

In contrast, NLRP7 exhibited a slow increase initially after hESCs were subjected to naïve conversion, followed by a dramatic upsurge of 5000-fold when hESCs were

completely converted to naïve state. Given that NLRP7 expression is extremely low in hESCs, only ~1% of NANOG (0.01% of β-actin) (my data and Yan et al., 2013), it is anticipated that several-fold increase may not reach a functional level. Thus, functional NLRP7 may only be reached at the later stage of conversion, after or at least the same time as global DNA demethylation takes place. Due to the unforeseen situation, I did not get time to perform an immunostaining experiment with antibody against 5mC to monitor the dynamic changes of global methylation levels during the naïve conversion, which could provide further evidence to support this claim. Nonetheless, these results suggested that NLRP7 levels were still low when global demethylation took place during naïve conversion, thereby might be unable to protect maternally imprinted genes from the demethylation.

Indeed, after converting NLRP7-OE hESCs back to the naïve state, maternally imprinted genes, *PEG10 and NAP1L5*, were shown not to exhibit an equal biallelic expression, rather the allele that was active in primed hESCs revealed a higher expression than the allele that was silenced in primed hESCs. This result indicates a partial rescue of maternal imprinting, rather than the complete loss of imprinting as in WT cells (Fig 4.7A). Furthermore, the methylation on the ICR of *PEG10* and *NAP1L5* revealed ~50% DNA methylation after naïve conversion, similar to that of primed hESCs before naïve conversion (Fig 4.7B). The explanation for partial rescue of maternal imprinting by NLRP7-OE before naïve conversion might be attributed to the fact that transgenic NLRP7 was not expressed at a sufficiently high level at primed state. In fact, NLRP7 mRNA was increased about 70 fold after stable overexpression (Fig 4.6A), which is close to the endogenous NLRP7 expression level between D4 and D6 of naïve conversion, but still much lower than endogenous NLRP7 level when the naïve conversion was completed at D10 (Fig 4.1D). This might be due to some other factors, e.g. miRNAs in primed hESCs that inhibit NLRP7 expression; or the NLRP7 transgene may get truncated when inserted into the host genome, due to its large size, during lentiviral infection. Another possible reason could be that NLRP7 may require other factors, e.g. KHDC3L, which together efficiently protect maternal imprinting. A

recent study of KHDC3L-deficient human oocytes showed that KHDC3L affects imprinting not only through the de novo methylation in oocytes, but also via maintenance of DMR during early embryonic development (Demond et al., 2019). Thus, the two genes may have a similar role in maintaining maternal imprinting during global demethylation of early human embryos. Interestingly, NLRP7-OE did not seem to affect paternally imprinted gene *MEG3,* which still revealed a loss of imprinting upon naïve conversion. Therefore, these results suggest that NLRP7 may specifically support the protection of DNA methylation on the DMR of maternally imprinted genes during global demethylation. However, this experiment was only done once due to the unusual circumstances. Ideally, more repeats should be carried out in different cell lines. In addition, since only two maternally imprinted genes and one paternally imprinted gene were analysed in this experiment, it is essential to examine more imprinted genes to validate the results, or perform genome-wide bisulfite sequencing to efficiently determine the specific protection on maternal imprinting. Furthermore, different patient mimicking mutants could be introduced into the NLRP7 transgene to see whether the rescue of maternal imprinting would occur or not.

Another interesting result from my experiments is that I incidentally identified that H1 and H7 had differential expression of NLRP7 isoforms. H1 hESCs only expressed the NLRP7 isoform 2 without exon 10, while H7 expressed predominantly isoform 2 as well as other isoforms (Fig 4.3). When Slim's group studied NLRP7 mutations related to different recurrent reproductive wastages, they found mutations in NLRP7 exon 10 were very rare, but they did identify one patient with an L946P homozygous mutation in exon 10 who experienced two partial hydatidiform mole and two spontaneous abortions (Deveault et al., 2009), which may implicate the importance of exon 10 in NLRP7 function. However, the exact functions of different isoforms are not clear and the reason why H1 and H7 express different isoforms is also unknown, so further studies with more hESC lines are required for better understanding of the different isoforms of NLRP7.

In conclusion, results obtained in this chapter revealed that NLRP7 expression was very low at the early stage of naïve conversion, while overexpression of NLRP7 in hESCs may partially protect maternal imprinting from global demethylation caused by naïve conversion. However, the exact mechanism of how NLRP7 may affect maternal imprinting remains to be explored.

Chapter 5.

Investigating whether NLRP7 regulates maternal imprinting through DNMT3L

5.1 Introduction

Autosomal recessive mutations of maternal effect gene, NLRP7, have been identified to be responsible for BiCHM in humans through dysregulation of maternal imprinting (Sanchez-Delgado et al., 2015; Reddy et al., 2013). In the last chapter, it was shown that overexpression of NLRP7 in hESCs might be able to partially protect maternal imprinting from erasure during naïve conversion. However, the underlying mechanisms by which NLRP7 modifies imprinting remain largely unknown. There are two possible mechanisms: 1) NLRP7 may interact with epigenetic factors that regulate DNA methylation, to affect their subcellular localization and nuclear translocation (Sanchez-Delgado et al., 2015). 2) Given that NLRP7 is also found in the nuclear fraction of decidualized human immortalized endometrial stromal cells (Huang et al., 2017), NLRP7 itself may translocate to the nucleus to interact with epigenetic factors and assist the protection of maternal imprinting.

Imprinting is mainly established through differential DNA methylation of the ICR of paternal and maternal alleles (Reik and Walter, 2001). Four DNMTs are essential to the establishment and maintenance of DNA methylation. DNMT1 functions to maintain DNA methylation after replication, whereas DNMT3A and 3B are de novo DNMTs, mainly expressed in early embryonic development and methylate CpG islands to establish the methylation pattern (Kato et al., 2007; Okano et al., 1999). DNMT3L is catalytically inactive, although having high homology to DNMT3A and 3B protein sequences, and is therefore considered to act as a cofactor for DNMT3A/B mediated DNA methylation. Specifically, DNMT3L has been shown to facilitate DNMT3A to stimulate DNA methylation on maternally imprinted genes in a human kidney cell line (Chedin et al., 2002). Interestingly, targeted disruption of the *Dnmt3l* gene in mouse resulted in Dnmt3l-null females that produced embryos failing to grow beyond 9.5 days, due to embryonic defects, as a result of biallelic expression of maternally imprinted genes (Bourc'his et al., 2001). This phenomenon is similar to the loss of maternal imprinting in humans due to NLRP7-deficiency. Although human DNMT3L transcripts

and proteins, unlike murine Dnmt3l, are only detected after fertilization and not during female oogenesis (Huntriss et al., 2004; Petrussa et al., 2014), it is not known whether DNMT3L plays a role in protecting maternal imprinting from DNA demethylation after fertilisation in humans.

In chapter 3, it have been shown that both NLRP7 and DNTM3L are exponentially increased in naïve hESCs, while other DNMTs are not significantly affected. This, in combination with the phenotype produced by Dnmt3l-knockout mouse, raised the question as to whether NLRP7 modifies genomic imprinting by modulating the function of DNMT3L. Thus, this chapter is focused on studying the relationship between NLRP7 and DNMT3L via:

i. Comparing the expression pattern of imprinting-related genes (eg. DNMTs, TETs and SCMC members) using single-cell RNA-seq profiles from different resources, such as human oocytes, preimplantation embryos and hESCs undergoing naïve conversion.
ii. Investigating whether NLRP7 expression affects DNMT3L expression, and vice versa.
iii. Identifying the subcellular localization of NLRP7 and the possibility of interaction between NLRP7 and DNMT3L.

5.2 Results

5.2.1 NLRP7 and DNMT3L have similar expression patterns during early embryonic development and naïve conversion.

From gametogenesis to fertilization and embryonic development, the temporal and spatial pattern of gene expression in the cells undergoes a dramatic transformation, as well as the global DNA methylation and other epigenetic profiles. Nowadays, several single-cell RNA-seq profiles have been obtained from oocytes, preimplantation embryos and hESCs (Yan et al., 2013; Petropoulos et al., 2016). These datasets enabled me to analyse the expression of genes that may be involved in regulating maternal imprinting during the early embryonic development, such as DNMTs, TETs and SCMC members, including NLRP7.

In the dataset of Yan et al., (2013), they examined the gene expression profile in metaphase II mature oocytes before fertilisation, zygotes and multiple-cell embryos until the blastocyst stage where cells become trophectoderm (TE), epiblast (EPI) and primitive endoderm (PE). Firstly, the expression of *DNMT* genes was analysed. DNMT1 exhibits a highly dynamic expression pattern during early embryogenesis, being highly expressed in oocytes followed by an increase after fertilisation, then peaking at the 2-cell stage. After that, it dramatically declines and almost disappears in all cell lineages of blastocysts, including TE, EPI and PE (Fig 5.1A). By contrast, the expression of de novo DNMTs are very low in oocytes and stay at these levels until the blastocyst stage where DNMT3A retains its expression at low levels while DNMT3B is slightly increased (Fig 5.1A). Noticeably, DNMT3L considerably increased its expression in all the lineages of blastocysts, with the highest expression in trophoblast cells (TE), which makes it the highest expressed *DNMT* genes (Fig 5.1A).

These dynamic expression patterns of DNMT genes are supported by another study that inspected gene expression in embryos from day-3 (E3) after fertilisation when

embryos are around 8 to 16 cells to day-7 (E7) when the embryo has progressed to the late blastocyst (Sahakyan and Plath, 2016; Petropoulos et al., 2016). In this study, the trends in expression patterns of DNMT genes are similar to Yan et al.'s study despite there being some subtle differences in the absolute expression levels. This study extends to the late blastocyst where DNMT3L exhibits downregulation in all three lineages of the late blastocysts, at E6 and E7, after the initial increase at E5 (Fig 5.1B).

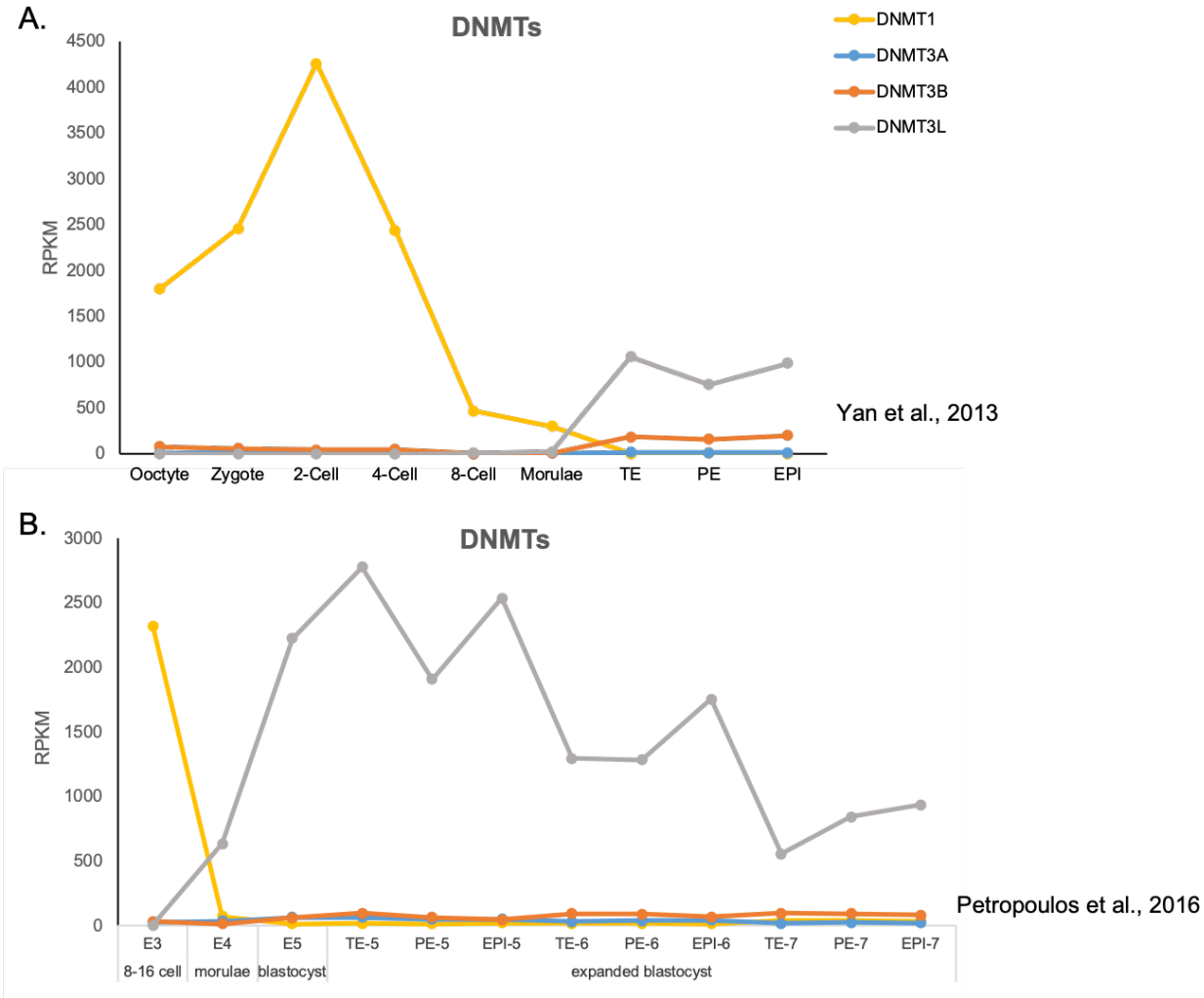

Figure 5.1 Dynamic expression of DNMTs in human oocytes and preimplantation embryos. Plot of average mRNA expression levels of DNMT genes from single-cell RNA-seq data from **(A)** Yan et al, 2013 and **(B)** Petropoulos et al., 2016. RPKM, Reads Per Kilobase of transcript per Million mapped reads.

Next, expression of *TET* genes were analysed as TETs are proteins responsible for active DNA demethylation. TET3 is relatively highly expressed in oocytes compared with other TETs and shows an upsurge after fertilization. It is gradually decreased in 2- and 4-cell embryos and then rapidly declines to a very low level from the 8-cell stage onwards (Fig 5.2A). Both datasets show the expression of all three TETs were very low between the 8-cell stage and the morula stage, and that TET1 and TET2 suddenly increase in the TE of blastocyst, although TET3 expression in all 3 lineages is not consistent (Fig 5.2A and B). According to a recent study in mouse zygotes, Tet3 may predominantly contribute to the maintenance of the global hypomethylated state via hydroxylation of newly formed 5mC, avoiding Dnmt1-driven methylation maintenance (Amouroux et al., 2016), while TET1 and TET2 might have a more important role in the blastocyst.

Towards the end of oogenesis, chromatin of full-mature oocytes undergoes substantial condensation and transcription is globally silenced (Schultz et al., 2018), therefore it is interesting to observe a drastic upregulation of DNMT1 and TET3 in zygotes compared with mature oocytes in the absence of transcription (Fig 5.1A and 5.2A). Recent studies showed DNMT1 and TET3 mRNA are indeed detectable in human sperm (Rahiminia et al., 2019; Ni et al., 2016). Although there might be some contribution from the sperm, it is very unlikely to explain this observation. However, sample variation may be a possible explanation.

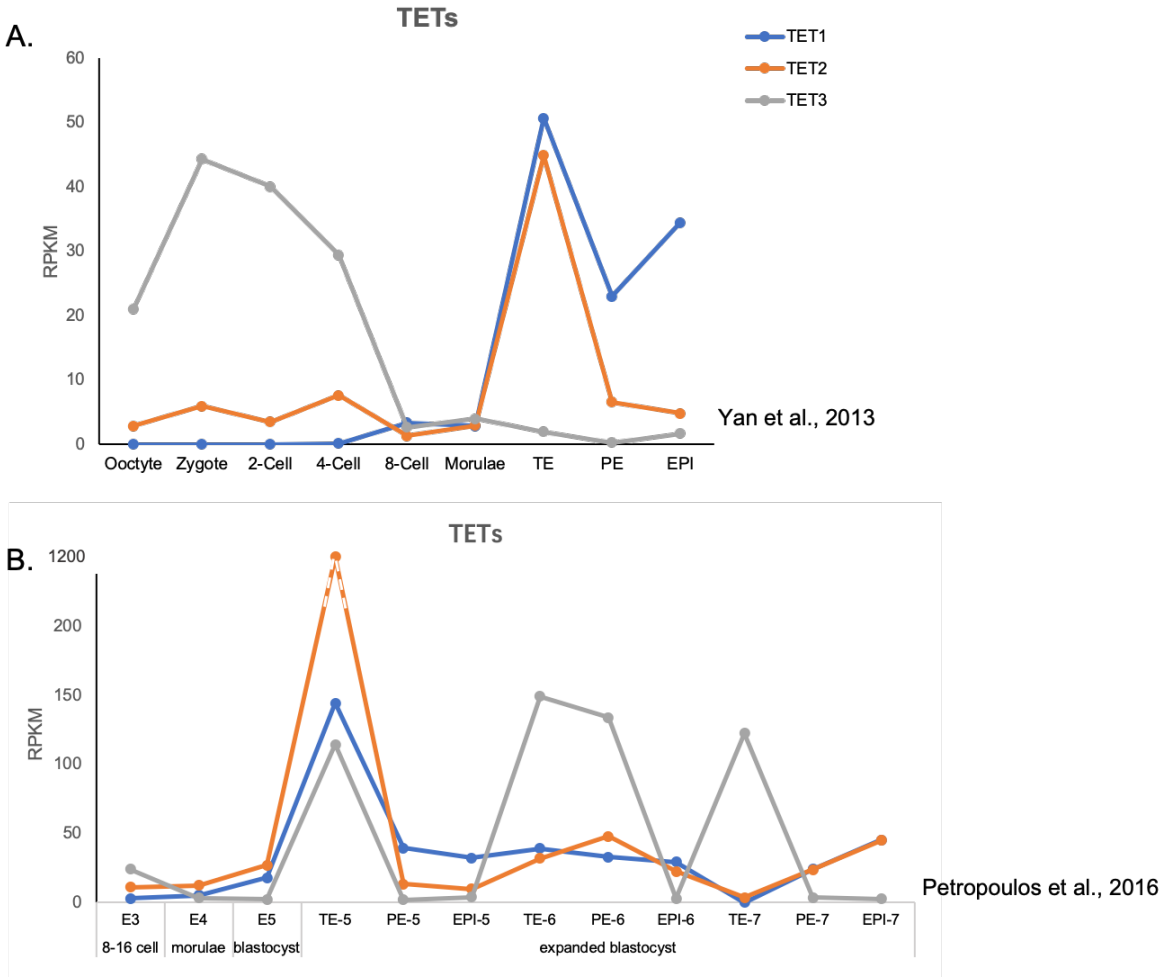

Figure 5.2 Dynamic expression of TETs in human oocytes and preimplantation embryos. Plot of average mRNA expression levels of TET genes from single-cell RNA-seq data from **(A)** Yan et al, 2013 and **(B)** Petropoulos et al., 2016. RPKM, Reads Per Kilobase of transcript per Million mapped reads.

Moreover, given that NLRP7 is a potential member of the SCMC (Akoury et al., 2015; Monk et al., 2017) and that mutations in human SCMC members, e.g. *KHDC3L* and *NLRP5*, have also been reported to cause CHM, early embryonic arrest and widespread MLID (Parry et al., 2011; Docherty et al., 2015), the expression of SCMC members were also analysed. NLRP2 and NLRP7 showed a similar trend of expression during embryogenesis. Both are continuously expressed at relative low levels in oocytes and before the 8-cell stage, although NLRP7 had a lower expression than NLRP2, both of them become highly expressed in blastocysts (Fig 5.3 A and B). This may be due to the divergence of NLRP7 from the duplication of NLRP2 and their adjacent location in the genome, thus they may share similar functions and regulation mechanism (Tian et al., 2009). KHDC3L reveals a different pattern of expression from NLRP2 and NLRP7, being highly expressed in the oocyte and early cleavage stages, peaking at the 8-cell stage, then dramatically declines at the morula stage and remains undetectable in blastocysts. OOEP is the highest expressed SCMC gene in the oocyte and early embryo before the 4-cell stage, and then rapidly reduces to undetectable level after the morula stage (Fig 5.3 A & B). The other SCMC members (NLRP5, PADI6 and TLE6) are generally expressed at a relatively low level in human oocyte and throughout the early embryonic stages (Fig 5.3 A and B).

Noteworthy, Yan's RNA-seq data shows that most SCMC members have relatively low mRNA expression in mature oocytes, except OOEP and KHDC3L, which is consistent with another transcriptome data obtained in preovulatory oocytes (Zhang et al., 2018). However, it is contradictory to the high SCMC protein expression observed in oocytes (Zhu et al., 2015; Akoury et al., 2015). By studying human fetal ovaries and female PGCs around gestational age 20 weeks, it was found that the SCMC transcripts are strongly detected (Zhu et al., 2015; Guo et al., 2015), which may contribute to the translation of SCMC proteins in primordial oocytes. These proteins may be stored in oocytes even after birth in spite of the low mRNA level. This contradiction between the mRNA and protein level may therefore be due to the short half-life of the transcripts and high stability of the proteins.

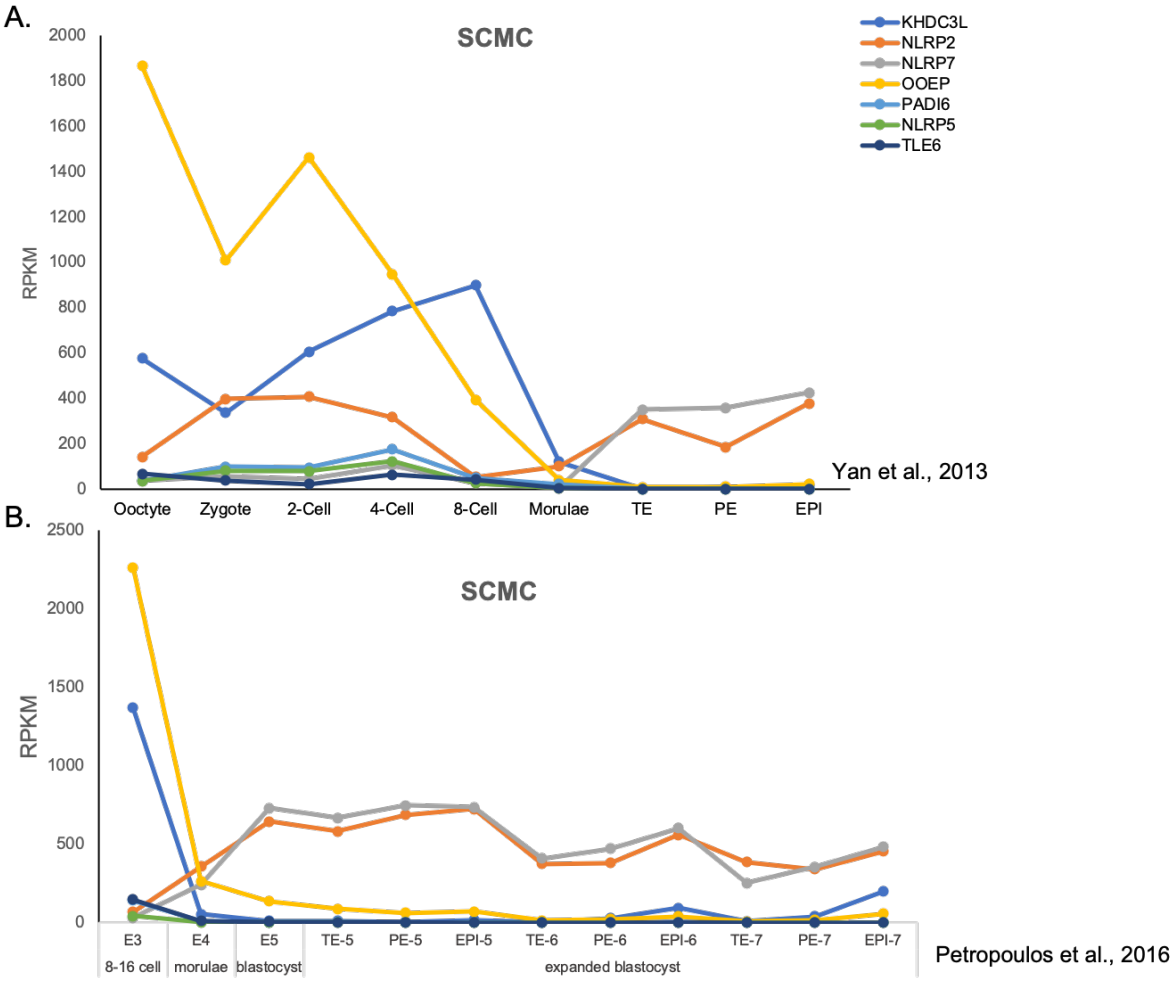

Figure 5.3 Dynamic expression of SCMC members in human oocytes and preimplantation embryos.

Plot of average mRNA expression levels of SCMC genes from single-cell RNA-seq data from **(A)** Yan et al, 2013 and **(B)** Petropoulos et al., 2016. RPKM, Reads Per Kilobase of transcript per Million mapped reads.

In this analysis, it has attracted my attention that, among all the DNMT and TET genes, the expression of DNMT3L exhibits a very similar pattern to that of NLRP7 (Fig 5.4 A & B). Together with the similar phenotypic consequences of Dnmt3l-deficiency in mouse and NLRP7-mutations in humans, it raised a question whether NLRP7 affects genomic imprinting via interacting with de novo DNMTs, particularly DNMT3L.

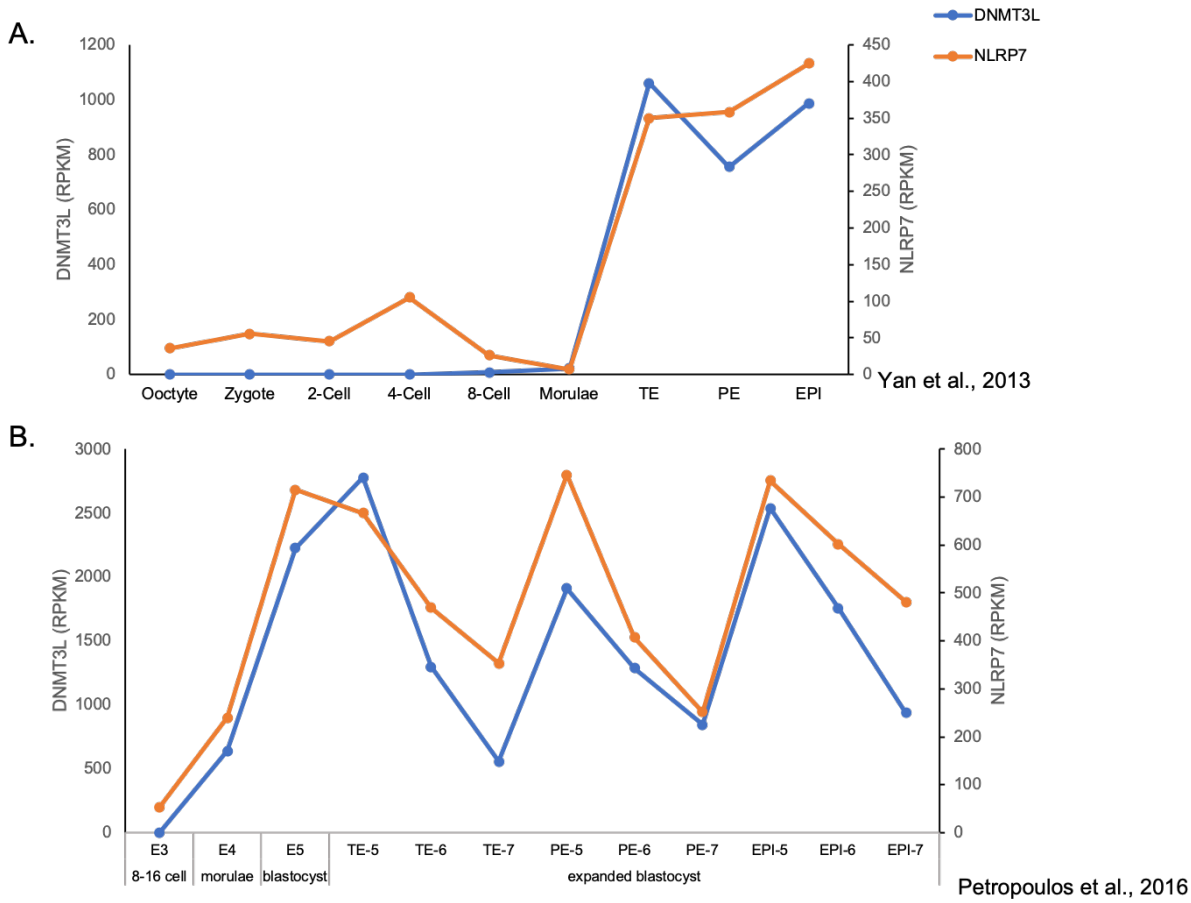

Figure 5.4 Comparison of DNMT3L and NLRP7 mRNA expression in human oocytes and preimplantation embryos.

Plot of average mRNA expression levels of DNMT3L and NLRP7 from single-cell RNA-seq data from **(A)** Yan et al, 2013 and **(B)** Petropoulos et al., 2016. RPKM, Reads Per Kilobase of transcript per Million mapped reads.

In Chapter 3, both NLRP7 and DNTM3L were shown to be expressed at low levels in hESCs but dramatically upregulated in naïve hESCs, while the other *DNMT* genes do not show such drastic increase in their expression. Thus, it is interesting to further explore whether NLRP7 and DNMT3L also have the same trend of increase during the process of naïve conversion. cDNA samples collected every two days after switching to 5iLAF naïve medium (section 4.2.1) were analysed via qRT-PCR. Interestingly, both NLRP7 and DNMT3L slowly increased in the first 6 days of naïve conversion and exponentially increased at the later period. Noticeably, the rising trend was almost identical (Fig 5.5). However, it was unknown whether it was due to the global demethylation process during naïve conversion or any regulatory relationships between NLRP7 and DNMT3L.

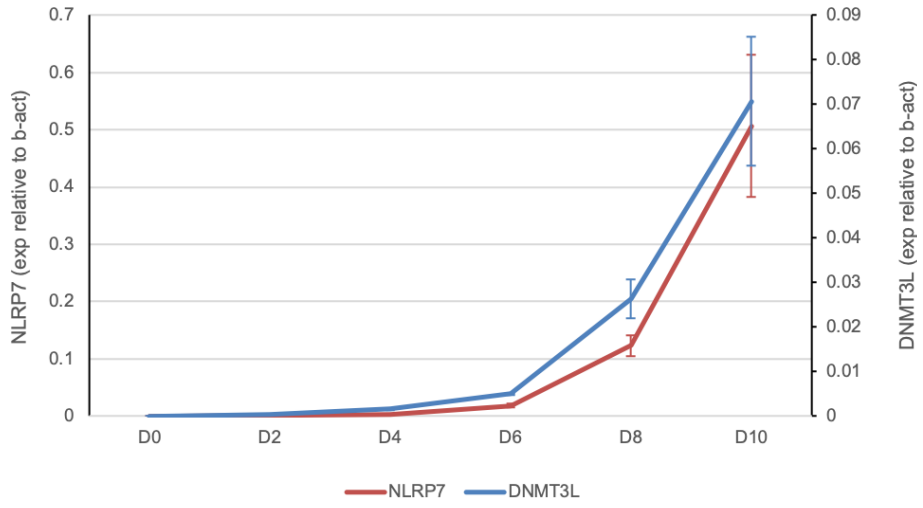

Figure 5.5 Comparing the trend of NLRP7 and DNMT3L mRNA expression during naïve conversion.

RNA expression analysis on NLRP7 and DNMT3L every 2 days after replacing to 5iLAF naïve medium via qRT-PCR. Data are presented as mean ± SD (n=3)

5.2.2 Overexpression of NLRP7 in hESCs upregulates DNMT3L mRNA expression.

As is known, DNMT3L-deficient female mice exhibit a similar phenotype to NLRP7-deficinet women, both giving rise to embryos with maternal imprinting deficiency (Bourc'his et al., 2001; Sanchez-Delgado et al., 2015). Additionally, NLRP7 and DNMT3L transcripts show a correlated expression pattern during human early embryogenesis and naïve conversion. It is interesting to investigate whether NLRP7 can affect DNMT3L expression. To address this, an HA-tagged NLRP7 expression vector was constructed as described in section 4.2.3 (Fig 5.6 A). NLRP7 was overexpressed in hESCs by lipofectamine transfection. qRT-PCR confirmed the overexpression of the NLRP7 transgene, with more than 200-fold increase in H1 and 100-fold increase in H7 (Fig.5.6 B & C). The DNMT3L mRNA expression exhibited a 2-fold increase ($p<0.05$) in NLRP7-OE hESCs in both H1 and H7 cell lines, while other DNMTs were not significantly affected (Fig. 5.6 B & C). Although NLRP7OE-induced increase of DNMT3L is small, it is statistically significant and this result has been validated 3 times in each cell line.

Based on the phenomenon that NLRP7 and DNMT3L synchronously increased during naïve conversion, it is also possible that NLRP7 overexpression may be involved in the transition of hESCs from primed to naïve state, thus indirectly leading to the increase of DNMT3L, rather than directly regulating its expression. However, the NLRP7-OE hESCs did not exhibit any morphological changes that resembled domed naïve hESCs. To further validate their pluripotent status, the mRNA expression of pluripotency and naïve markers (OCT4, NANOG and REX1) were examined to determine if the cells still remained in the primed state. The result revealed expression of OCT4, NANOG and REX1 remained the same in control and NLRP7-OE hESCs, which means that NLRP7 overexpression may not affect the developmental status of the hESCs (Fig 5.6 D). Since NLRP7 is not a transcription factor, it may affect DNMT3L expression via an indirect mechanism subtly tuning DNMT3L expression levels.

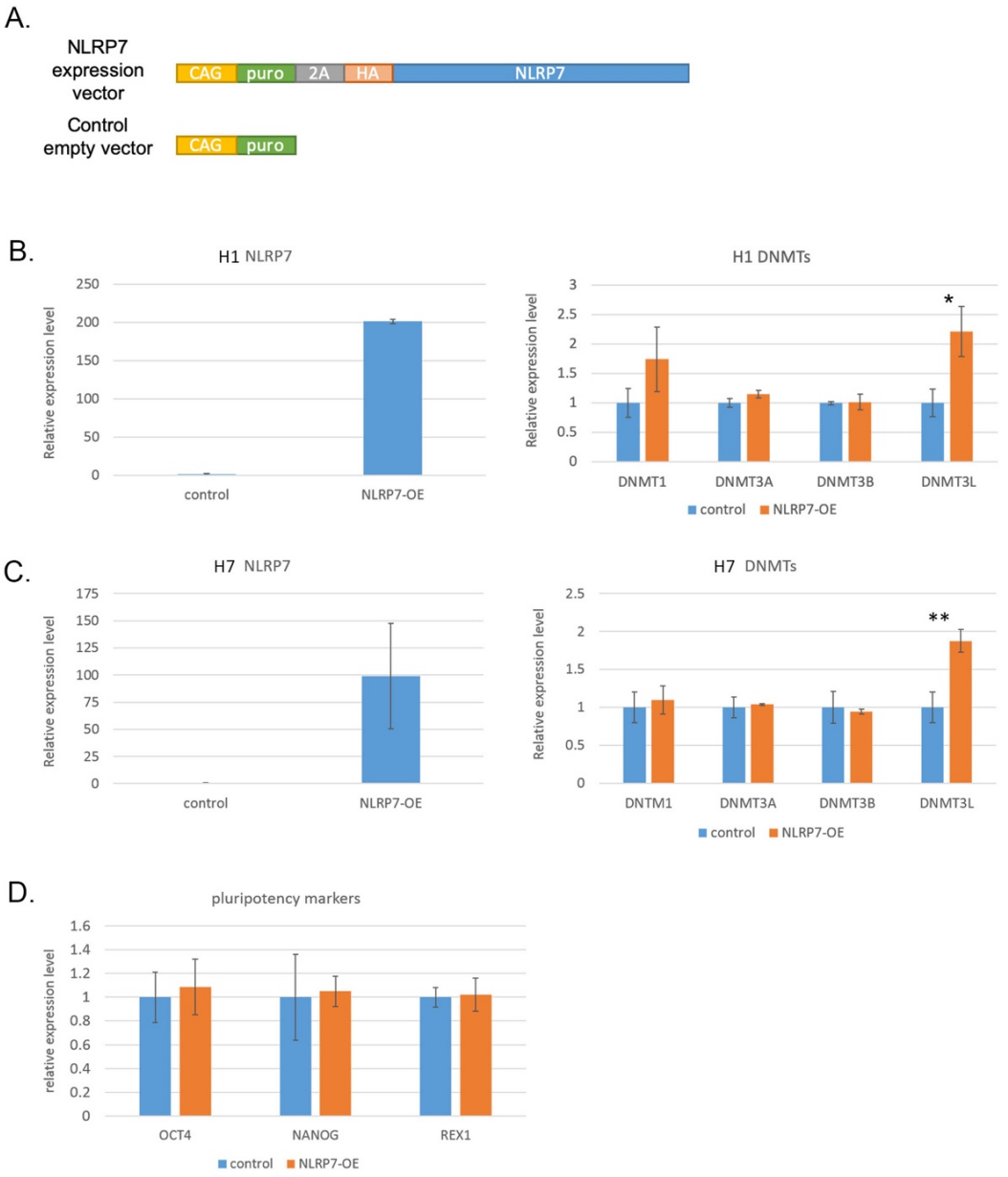

Figure 5.6 The effect of NLRP7 overexpression in hESCs

(A). Schematic figure of NLRP7 expression vector and empty control vector; qRT-PCR analysis of the expression of NLRP7 and DNMTs in NLRP7-OE and control hESCs in H1 **(B)** and H7 **(C)** cell lines.(*p<0.05, **p<0.01, by two-tail unpaired t-test; n=3) **(D)**. qRT-PCR analysis of pluripotency and naïve marker expression in NLRP7-OE and control H1 hESCs.

Since NLRP7 was identified as the major causative gene of BiCHM (Murdoch et al., 2006), a number of different NLRP7 mutations have been found in BiCHM patients (Fig 5.7 A). As illustrated in the schematic, mutations found in the PYD and NACHT domains are usually nonsense mutations, while missense mutations are generally located in the LRR, which indicates that the LRR may have an important function in the NLRP7 protein. Two representative NLRP7 mutations, E570*, a nonsense mutation, and R693P, the most frequent missense mutation identified in NLRP7 (Wang et al., 2009) (Fig 5.7 B), were chosen to investigate whether the LRR region plays any role in the regulation of DNMT3L expression. E570* will produce a truncated NLRP7 without the entire LRR region and R693P a missense mutation in the LRR. Western blotting demonstrated that E570* and R693P NLRP7 had been successfully constructed via site directed mutagenesis of pCAG-HA-NLRP7 wild-type (WT) plasmid: E570* produced a 67kD truncated protein and R693P gave a band with the same size as the WT NLRP7 (Fig 5.7 C).

After transfection of WT and mutant NLRP7s into hESCs, the WT and R693P NLRP7 transcripts showed similar levels of expression, which were much lower than that of the E570* truncated NLRP7 (Fig 5.7D left). This is probably due to the E570* plasmid being smaller in size than the WT and R693P NLRP7 plasmids, making it easier to transfect, leading to a higher transfection efficiency. Interestingly, the DNMT3L mRNA expression was not affected in hESCs transfected with either mutant NLRP7 (Fig 5.7 D right), while the hESCs overexpressed with WT NLRP7 showed an almost 2-fold increase. These transfection experiments have been repeated three times and similar results were obtained (Fig 5.7). These results indicate that the LRR region of NLRP7 may play a role in regulating DNMT3L expression.

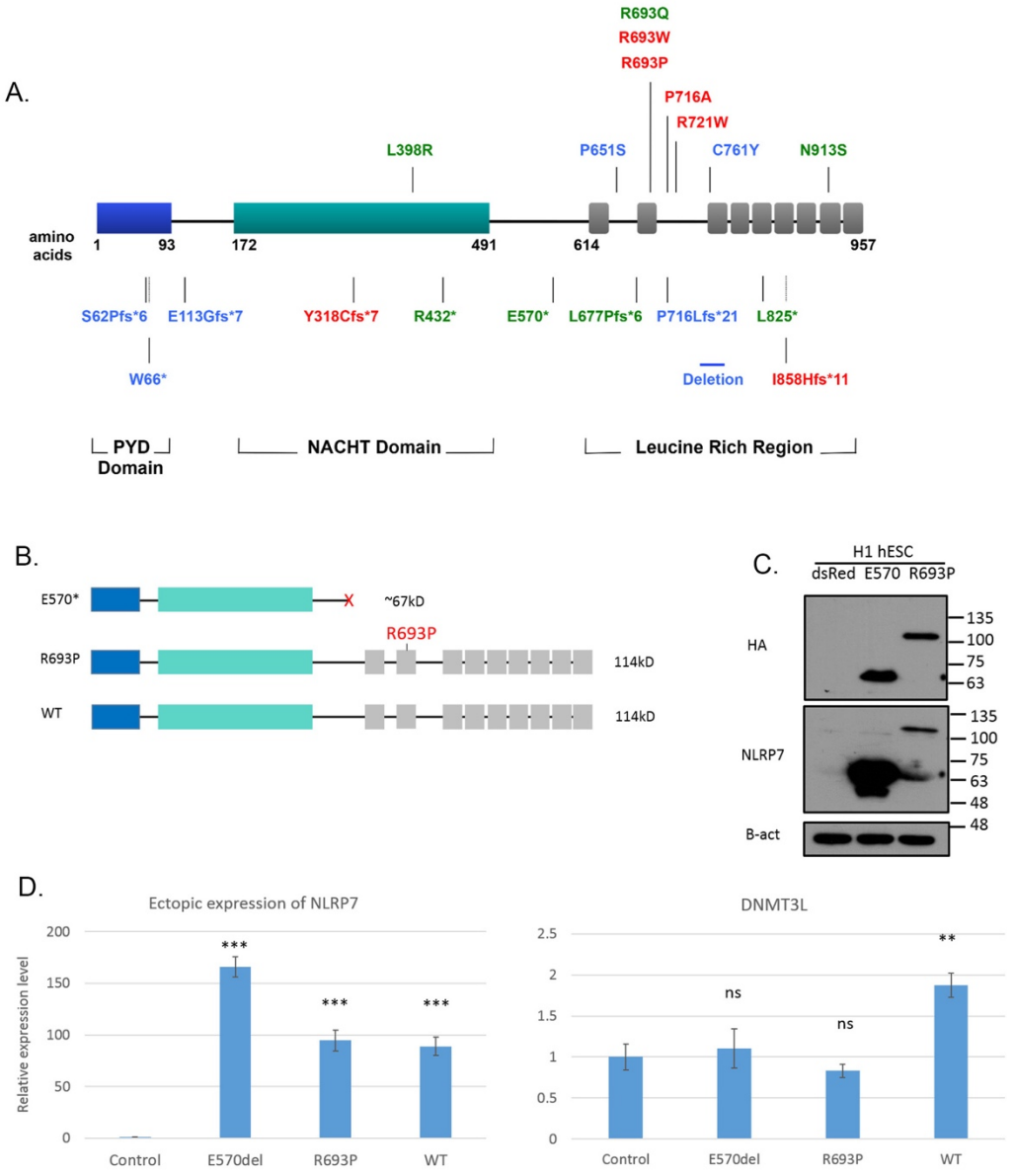

Figure 5.7 NLRP7 LRR region may have a crucial role in regulating DNMT3L expression. **(A)** Schematic representation of NLRP7 mutations identified in BiCHM patients presenting at Imperial College Healthcare NHS (Wang et al; 2009; Dixon et al; 2012). Red represents mutations found in more than one family in these series, green represents mutations also found in other series and blue represents mutations currently reported in only a single family. **(B)** Schematic map of NLRP7 with patient mimicking mutations E570* and R693P compared to wild-type NLRP7. **(C)** Western blot analysis on H1 hESCs transfected by mutant NLRP7s. **(D)** RT-qPCR analysis on H1 hESCs transfected by wild-type and mutant NLRP7. (ns= not significant, $p>0.05$, **$p<0.01$, ***$p<0.001$, by two-tail unpaired t-test; n=3)

5.2.3 NLPR7-knockdown showed no effect on DNMT3L expression in naïve hESCs, but slightly reduced DNMT3L expression in primed hESCs

To further validate the hypothesis that NLRP7 may regulate the expression of DNMT3L, knockdown of NLRP7 was performed in hESCs. Since the EMBL-MUSCLE (multiple sequence alignment tool) revealed that human NLRP7 and NLRP2 share more than 70% homology in their protein sequence and the most divergent sequences are in the PYD domain and LRR region, two shRNAs located in the PYD domain, shRNA-A and shRNA-D, were selected by using an siRNA selection program (http://sirna.wi.mit.edu/) and cloned in to the expression vector (Fig 5.8 A). Since NLRP7 and DNMT3L are highly expressed in naïve cells, each shRNA was introduced into naïve hESCs (Fig 5.8 B). Many naïve hESCs showed green fluorescence after lipofection, indicating a good transfection efficiency in naïve hESCs (Fig 5.8 C). NLRP7 mRNA expression was reduced more than 20% and 40% by shRNA-A and -D respectively, while NLRP2 revealed almost no change (Fig 5.8 D), indicating that both shRNAs specifically target NLRP7 and that shRNA-D showed a more efficient knockdown than shRNA-A. This has also been confirmed by NLRP7 western blotting (Fig 5.9 A).

Surprisingly, different to what has been observed in NLRP7 overexpressing primed hESCs, NLRP7 knockdown in naïve hESCs did not show any effect on DNMT3L expression at either transcriptional or translational level (Fig 5.9 A & B); neither did these affect the expression of other DNMTs, including DNMT1, DNMT3A and DNMT3B (Fig 5.9 B). In an attempt to increase the efficiency of knockdown, shRNA-A and shRNA-D were co-introduced in the same cells, however, the knockdown level was still similar to shRNA-D and the DNMT3L mRNA expression was not affected (Fig 5.9 C). The reason for conflicting results in NLRP7 knockdown cells was not clear, but could be due to the different nature of primed and naïve hESCs.

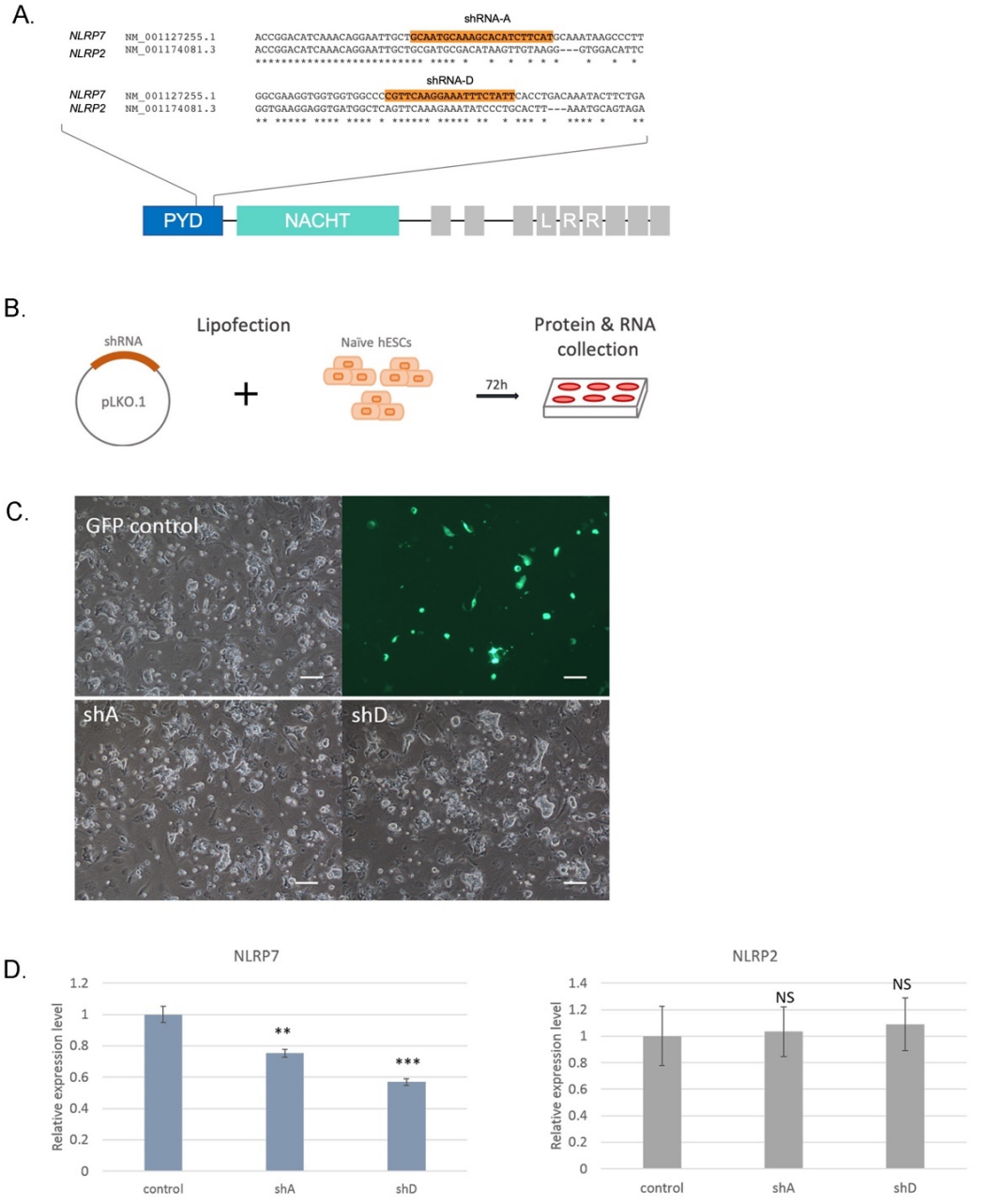

Figure 5.8 NLRP7 knockdown in naïve hESCs.

(A). Schematic figure shows the targeting of sequence shRNA-A and shRNA-D. **(B)**. experimental process of NLRP7 knockdown in naïve hESCs. **(C)**. Phase contrast and green fluorescent images of naïve hESCs 72h post-transcription (scale bar, 100 μm). **(D)**. RT-qPCR analysis on NLRP7 and NLRP2 in knockdown cell lines compared with control cells; (NS=non-significant, **p<0.01,***p<0.001, by two-tail unpaired t-test; n=3).

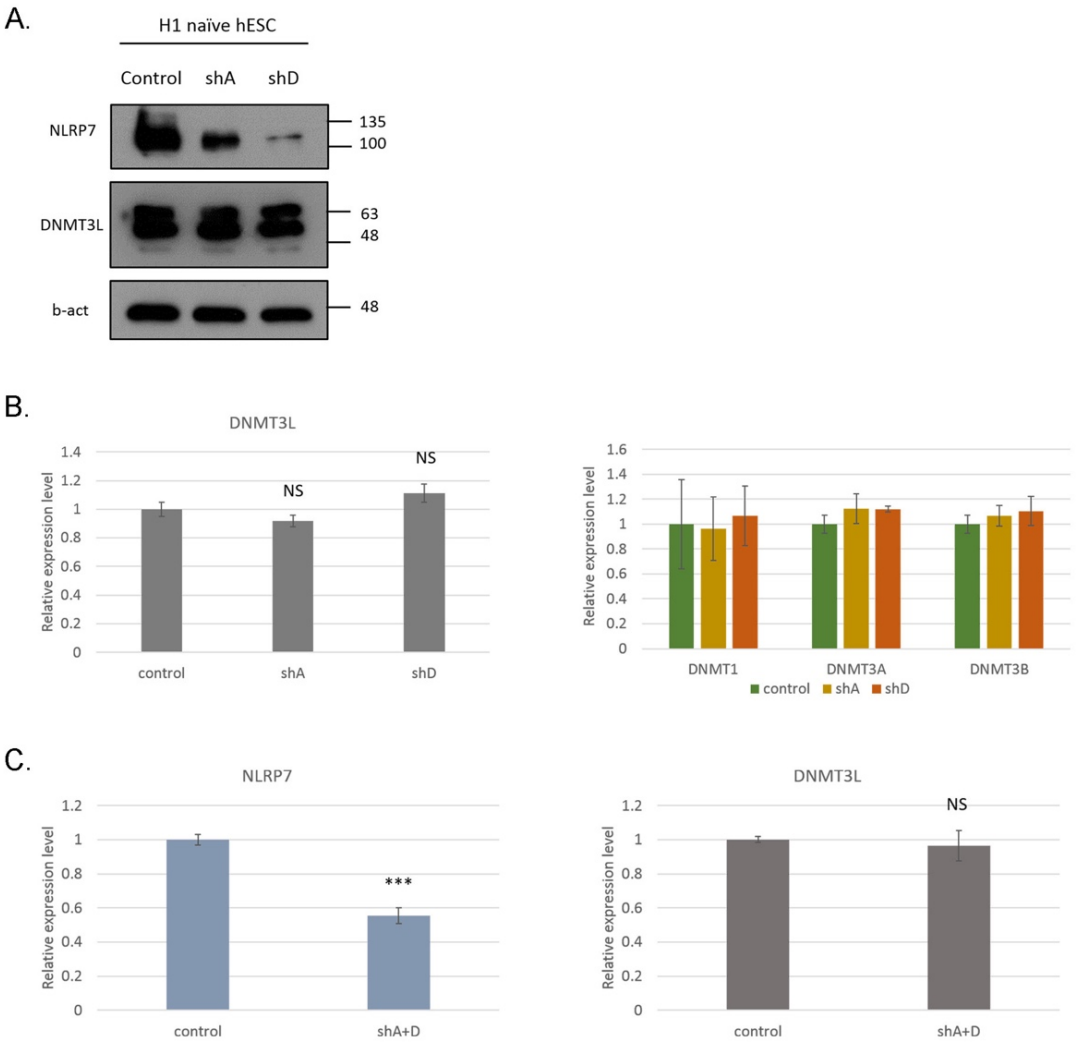

Figure 5.9 The effect of NLRP7 knockdown on DNMT expression in naïve hESCs.
(A) Western blot analysis of NLRP7 and DNMT3L in GFP control and NLRP7 knockdown cell lines. (B) qRT-PCR analysis on DNMT1, 3A, 3B and 3L in knockdown cell lines compared with control cells. (C) NLRP7 knockdown in naïve hESCs via combining shRNA-A and shRNA-D, as well as its effect on DNMT3L mRNA expression. (NS=not significant, ***$p<0.001$, by two-tail unpaired t-test; n=3)

Therefore, I tried knockdown of NLRP7 in primed hESCs even though both NLRP7 and DNMT3L are expressed at low levels in these cells (Fig 5.10 A). The qRT-PCR result revealed that NLRP7 was significantly decreased after the knockdown, with 50% reduction by shRNA-A and 80% reduction by shRNA-D. However, DNMT3L was only slightly decreased and not statistically significant in knockdown cells (Fig 5.10 B). This result was not surprising, because comparing the NLRP7 expression level in primed and naïve hESCs, primed hESCs possess very little NLRP7 (Fig 5.10 C); and based on Mahadevan et al., (2014)'s study, only a small fraction of NLRP7 could be detected in the nucleus even when NLRP7 expression level was high after overexpression. Therefore, most of the NLRP7 was probably located in the cytoplasm in hESCs rather than in the nucleus participating in transcriptional events. As a result it was difficult to determine the regulatory role of NLRP7 on DNMT3L transcription by knocking it down it in primed hESCs.

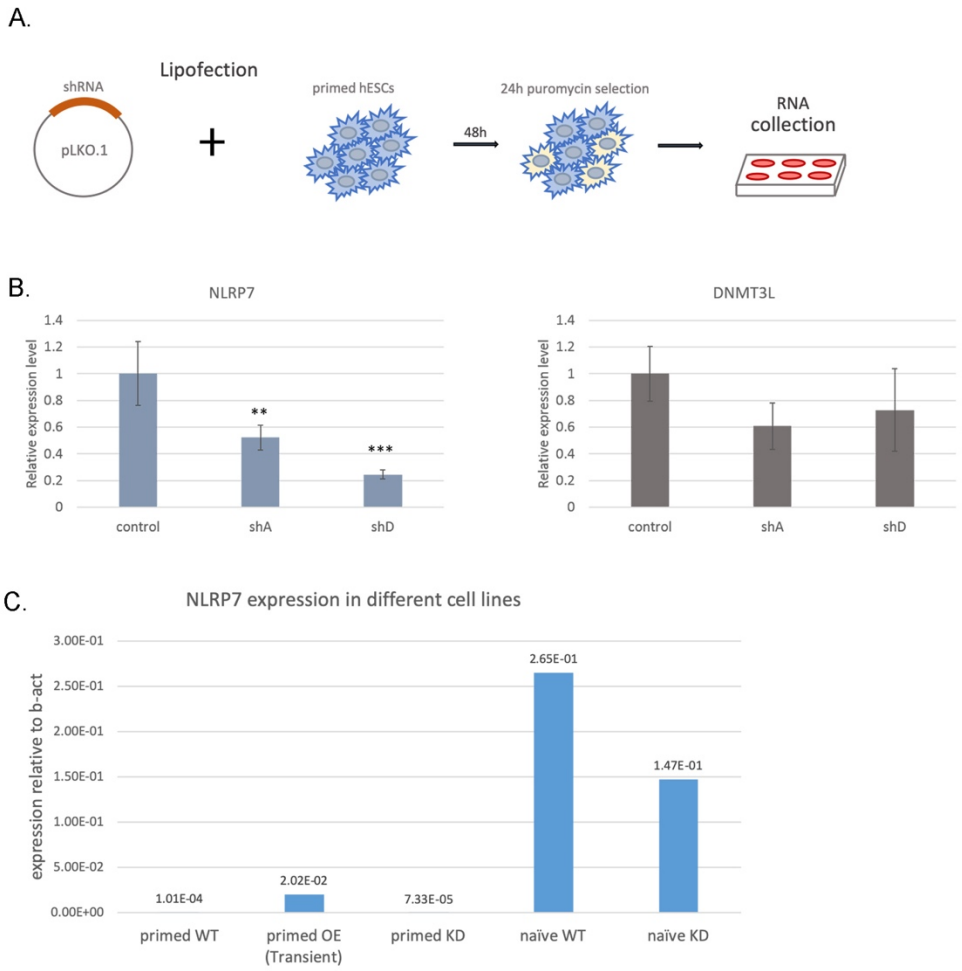

Figure 5.10 The effect of NLRP7 knockdown on DNMT expression in primed hESCs. **(A)** Experimental process of NLRP7 knockdown in primed hESCs. **(B)** qRT-PCR analysis on NLRP7 and DNMT3L in knockdown cell lines compared with control cells. (*$p<0.05$, **$p<0.005$, by two-tail unpaired t-test; n=3) **(C)** Comparison of NLRP7 mRNA expression in different WT (wild-type), NLRP7-OE (overexpression) and NLRP7-KD (knockdown) cell lines

5.2.4 Overexpression of DNMT3L in hESCs did not affect NLRP7 expression.

To further explore whether DNMT3L can regulate NLRP7 mRNA expression in hESCs, DNMT3L cDNA was isolated and cloned into an expression vector (Fig 5.11A). The FLAG-tagged DNMT3L expression vector was transfected into hESCs. qRT-PCR results showed that DNMT3L mRNA expression was elevated more than 9000 fold after overexpression, while NLRP7 mRNA expression did show any significant changes (Fig 5.11 B). Therefore, DNMT3L appears not to have a role in the regulation of NLRP7 expression.

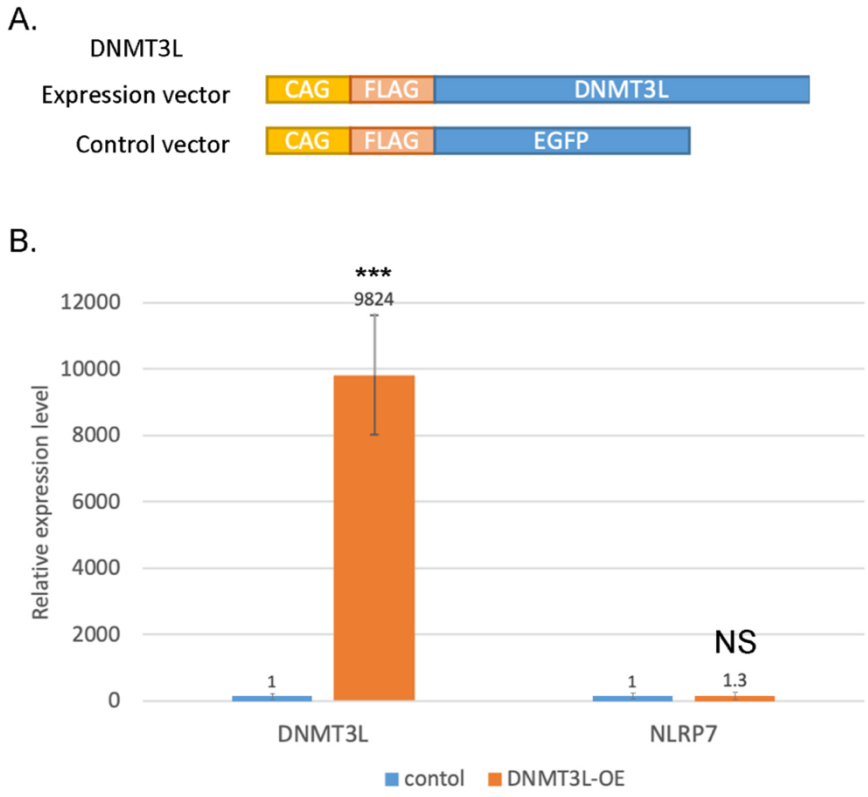

Figure 5.11 the effect of DNMT3L overexpression in hESCs.
(A). Schematic figure of DNMT3L expression vector and GFP control vector; **(B)**. qRT-PCR analysis of DNMTs and NLRP7 gene expression in DNMT3L-OE and control H1 hESCs (*** $p<0.01$; NS=non-significant; by two-tail unpaired t-test; n=3).

5.2.5 NLRP7 may also be expressed in the nucleus despite its predominant subcortical localisation.

Although NLRP7 is mainly detected in the subcortical region of human oocytes and preimplantation embryonic cells (Akoury et al., 2015), one report also detected its expression in the nucleus when overexpressed in HEK293T cells (Mahadevan et al., 2014). Thus, it is possible that a proportion of NLRP7 might be located in the nucleus where it regulates DNMT3L expression and DNA methylation of the maternally imprinted genes. However, Mahadevan et al.'s report was based on NLRP7 overexpression, not endogenous NLRP7, and thereby not a normal physiological situation. Furthermore, it is generally accepted that DNMT3L is mainly expressed in the nucleus (Guenatri et al., 2013). However, it has been reported that DNMT3L could also be expressed cytosolically in humans (Petrussa et al., 2014). Thus, I thought to further investigate the subcellular localisation of NLRP7 and DNMT3L in naïve hESCs using immunostaining since I have shown that both NLRP7 and DNMT3L are highly expressed in the naïve hESCs (Fig 3.8).

To ensure the specificity of the antibodies against NLRP7 and DNMT3L, hESCs in primed state were used as a negative control and showed no signals (Fig 5.12 A & B upper panel). This is consistent with their low expression shown by qRT-PCR and western blotting (described in chapter 3). By contrast, in naïve hESC colonies, but not the surrounding cells, NLRP7 signals were clearly detected. Close inspection revealed that the signals were mainly located in the cytoplasm of naïve hESCs, particularly accumulating at the subcortical regions. Interestingly, the signals were not evenly distributed in the cytoplasm, but often concentrated at certain subcortical areas of each cell. In approximately 10% of cells, the signals can be detected at juxtanuclear regions while ~ 5% of cells have nuclear signal (Fig 5.12 A lower panel). Similarly, DNMT3L proteins were also detected in the naïve hESC colonies, but the signals were predominantly localised in the nucleus of each naïve hESCs, with some expression in the cytoplasm (Fig 5.12 B lower panel).

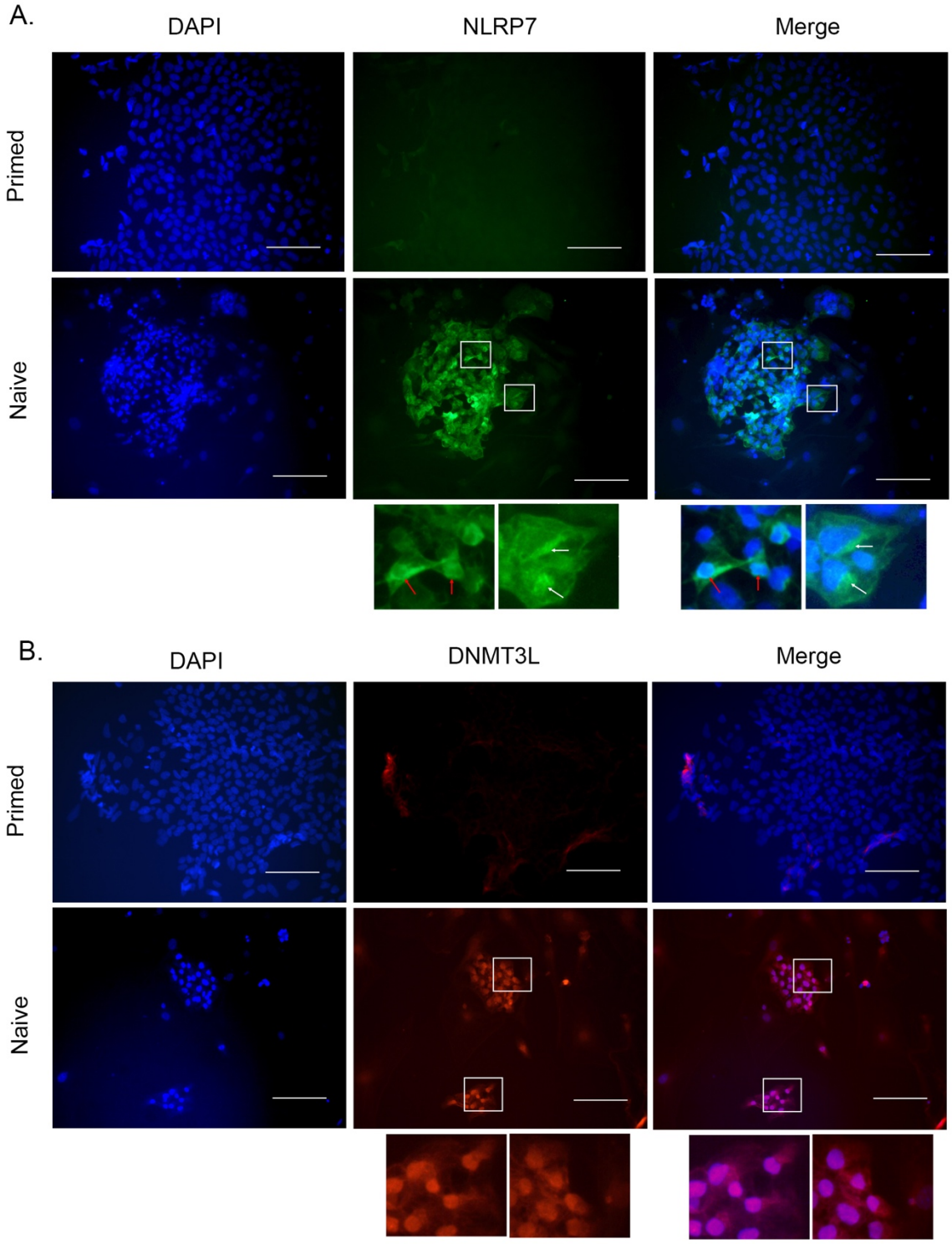

Figure 5.12 Immunostaining of NLRP7 and DNMT3L in primed and naïve hESCs. Immunostaining of NLRP7 (**A**) and DNMT3L **(B)** in primed (upper) and naïve (lower) hESCs. White arrow points to juxtanuclear localized NLRP7 staining; red arrow points to nuclear localized NLRP7 staining (Scale bar, 100 μm).

To further validate the subcellular localization of these two proteins, immunoblotting of cell fractions was also performed, which showed that NLRP7 was indeed also expressed in the nucleus although the majority of the protein was in the cytoplasmic fraction; whereas DNMT3L was mainly expressed in the nucleus, but a small amount could also be detected in the cytoplasm (Fig 5.13). Therefore, it is possible that NLRP7 may also perform some of its functions in the nucleus.

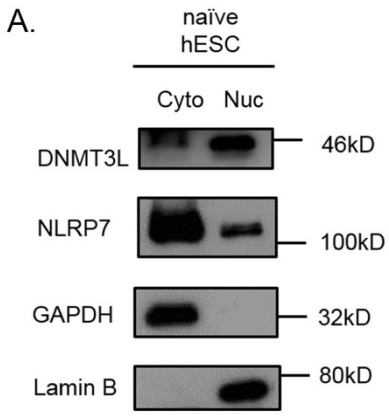

Figure 5.13 Identifying the localization of NLRP7 and DNMT3L in naïve hESCs via cytoplasmic-nuclear fractionation.

Naïve hESCs were subjected to cytoplasmic-nuclear fractionation and analyzed by immunoblotting. GAPDH and Lamin B were used to assess levels of cytoplasmic and nuclear contamination respectively.

5.2.6 NLRP7 and DNMT3L may have a physical interaction

Given NLRP7 does not contain a DNA binding domain, it is proposed that NLRP7 may recruit other epigenetic factors to achieve the protective role on maternally imprinted genes, so in order to investigate whether NLRP7 interacts with DNMT3L to assist the protection of maternal imprinting, I co-expressed HA-NLRP7 and FLAG-DNMT3L in HEK293T cells and then performed co-immunoprecipitation (Co-IP) (Fig 5.14 A). I applied HA and FLAG agarose beads for the Co-IPs and used HA-NLRP7/FLAG-EGFP co-transfection as a negative control. In the experimental group, a thin band of FLAG-DNMT3L can be detected in HA pull-down. FLAG pull-down also contained increased amount of HA-NLRP7 compared with the IgG control (Fig 5.14 B left). In the control group, HA-NLRP7 is precipitated while FLAG-EGFP cannot be detected (Fig 5.14 B right). In addition, in FLAG-IP, the HA-NLRP7 level is similar to the IgG control, which means NLRP7 was not precipitated when FLAG-EGFP was pulled down. Therefore, NLRP7 may have physical interaction with DNMT3L.

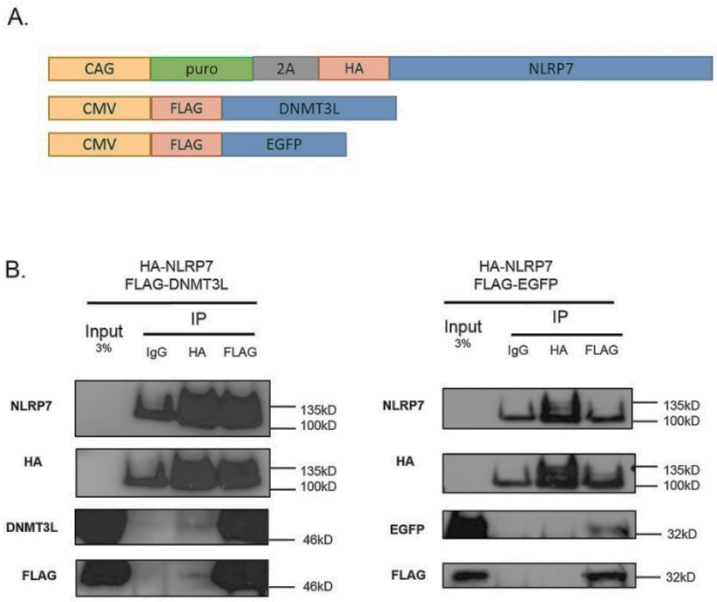

Figure 5.14. Detecting the physical interaction between NLRP7 and DNMT3L via co-IP
(A). Schematic of HA-NLRP7, HA- dsRed, FLAG-DNMT3L and FLAG-EGFP expression vectors. **(B)**. Co-IP analysis on cell lysates from HA- NLRP7/FLAG-DNMT3L co-transfected (left) or HA-NLRP7/FLAG-EGFP co-transfected (right) HEK293T cells with HA or FLAG antibodies as indicated.

5.3 Conclusion and Discussion

By analyzing two published RNA seq datasets from various stages of preimplantation human embryos, it has been shown that the genes directly involved in the regulation of DNA methylation, *DNMT*s and *TET*s, exhibit distinct expression patterns during this developmental process. The mRNA levels of both de novo DNMTs, DNMT3A and 3B, are very low throughout the preimplantation period; while the mRNA of maintenance DNMT1 is highly expressed in oocytes and early cleavage-stage embryos, particularly at the 2-cell stage, but progressively downregulated afterward (Fig 5.1). This DNMT1 is likely to be mainly the oocyte-specific isoform, DNMT1O (Hayward et al., 2003). In mouse, Dnmt1o, together with somatic Dnmt1, is thought to maintain differential methylation on imprinted genes (Cirio et al., 2008).

On the other hand, the expression pattern of *TET* genes is different from that of DNMTs. TET3 is highly expressed in embryos before the 8-cell stage, while TET1 and 2 are dramatically upregulated after the morula stage. This expression pattern makes the overall expression of TETs relatively high in preimplantation embryos, except for a transient trough between the 8-cell and the morula stage. It is important to note that before the 8-cell stage in humans, the majority of the mRNAs are maternal mRNAs as ZGA mainly starts to emerge from the 8-cell stage which is different to mice where it occurs from the 2-cell stage (Jukam et al., 2017). These maternal mRNAs are gradually degraded during cell division and replaced by embryonic mRNAs (Jukam et al., 2017; Schulz and Harrison, 2019). The clear differential expression patterns between DNMTs and TETs, particularly in early preimplantation embryos may account for the loss of global DNA methylation and the maintenance of hypomethylated state.

It is interesting to note that the methylation on the DMRs of imprinting genes is not disturbed in the preimplantation embryos where dramatic demethylation occurs (Guo et al., 2014). During this early embryonic development, NLRP7 mRNA exhibits a dynamic expression pattern: low expression before the blastocyst stage followed by a

considerable upregulation in the blastocyst, indicating that maternal NLRP7 mRNA levels are low until zygotic NLRP7 is activated. However, the maternal NLRP7 protein is abundantly expressed in oocytes, zygotes and cleaved early embryos (Akoury et al., 2015). Therefore, with the upregulation of zygotic NLRP7 mRNA in blastocysts, NLRP7 proteins are maintained at a high expression level throughout the genomic demethylation in early human embryonic development (Akoury et al., 2015). Given the facts that in the absence of normal NLRP7 in the human oocyte, the embryo will lose maternal imprinting and therefore develop as a CHM (Sanchez-Delgado et al., 2015) and that high expression of NLRP7 in hESCs is shown to partially rescue maternally imprinted genes from loss of imprinting during naïve conversion (Chapter 4), it is likely that this high level of NLRP7 may also have a role in protecting maternal imprinting during global demethylation.

Another interesting finding from the RNA-seq analysis is that the DNMT3L expression pattern correlates with the expression of NLRP7 (Fig 5.4), which shows clear upregulation in blastocysts even though other DNMT genes are still expressed at low levels. Consistently, it was also found that NLRP7 and DNMT3L showed very similar mRNA expression patterns during hESC naïve conversion (Fig 5.5). One explanation is that NLRP7 may affect DNMT3L expression as a transcription co-regulator or mRNA stabilizer. Remarkably, NLRP7-OE did promote DNMT3L mRNA expression in hESCs, although only two fold (Fig 5.6). Moreover, it may be the LRR domain of NLRP7 protein that plays a role in regulating DNMT3L expression as overexpressing LRR-deficient NLRP7 did not affect DNMT3L mRNA expression (Fig 5.7). There is already evidence that NLRP7 can interact with transcription factors, such as YY1 and ZBTB16 (Mahadevan et al., 2014; Singer et al., 2015). More specifically, YY1 has revealed a direct interaction with the PYD domain and LRR domain of NLRP7 (Mahadevan et al., 2014). Alternatively, NLRP7 may be able to stabilize DNMT3L mRNA, since the LRR domain is not only identified in NLRP proteins, but is also the main structural unit of the RNase inhibitor proteins that can bind to RNase A to regulate RNA degradation (Slim & Wallance, 2013). However, there is no direct evidence showing NLRP7 plays

any role in mRNA stabilization. Another possible explanation is that both NLRP7 and DNMT3L are controlled by the same regulatory machinery. Little is yet known about the factors that are involved in the regulation of their transcription, neither have their enhancer sequences been identified. Thus, more studies are needed in this area.

NLRP7 has been well characterised for its subcortical localisation in the oocytes and early embryonic cells. However, my study has clearly demonstrated that NLRP7 is not exclusively located at the subcortical region of a cell, but can also be detected in the nuclear and juxtanuclear regions (Fig 5.12 A). The nuclear localisation of NLRP7 has also been reported in at least two other reports. For example, a recent study in endometrial stromal cells reported that NLRP7 was aggregated at juxtanuclear region before in vitro decidualization, while it was translocated to the nucleus after decidualization (Huang et al., 2017). This suggests that the NLRP7 can be stored at the juxtanuclear region ready to be transported to the nucleus. Additionally, NLRP7 has also been shown to interact with the transcription factor YY1, in the cell nucleus, to mediate DNA methylation of specific genes in hESCs (Mahadevan et al., 2014). Structurally, NLRP7 protein possesses a nuclear localization signal in the NACHT domain, which may enable its translocation into the nucleus (Li et al., 2019). Since NLRP7 can translocate into the cell nucleus where it has the opportunity to interact with other transcription factors, it is also possible that NLRP7 interacts with DNMT3L in the nucleus to modify genomic imprinting. The result of co-IP in HEK293T cells indicates the potential physical interaction between these two proteins (Fig 5.14), although this may be due to the non-specific binding caused by abundantly overexpressed NLRP7 and DNMT3L in HEK293T cells. Therefore, more experiments, such as co-IP with endogenous proteins and co-immunostaining of NLRP7 and DNMT3L in naïve hESCs, are required to validate the result. Although DNMT3L does not have a DNA binding domain or functional catalytic domain (Suetake et al., 2004), DNMT3L has been shown to interact with unmethylated H3K4, to induce de novo DNA methylation via recruiting DNMT3A, in the mouse (Ooi et al., 2007). According to another study, demethylation of H3K4 via KDM1B lysine demethylase is essential for

the establishment of maternal imprinting during mouse oogenesis (Ciccone et al., 2009), which suggests a potential mechanism in that DNMT3L may recognize maternal imprinted loci and thereby affect the establishment or maintenance of maternal imprinting. Taken together, it was hypothesized that DNMT3L may interact with NLRP7 to form a complex and guide it to a maternally imprinted DMR with particular chromatin architecture (eg. demethylated H3K4) to protect it from global demethylation.

In summary, this chapter demonstrated that DNMTs and TETs, with different temporal expression, coordinating with each other may contribute to the hypomethylated state in the preimplantation blastocyst. The correlation of NLRP7 and DNMT3L expression in preimplantation embryos and during naïve conversion may be due to the regulatory role of NLRP7 on DNMT3L mRNA expression. Moreover, as a subcortically localized protein, NLRP7 was shown to be able to translocate to the cell nucleus, and may therefore affect gene transcription as well as imprinting probably via interaction with DNMT3L.

Chapter 6

General discussion

NLRP7, as a maternal effect protein, has now been thought to play an important role in meditating maternal imprinting and early embryonic development in humans (Sanchez-Delgado et al., 2015). However, to date most studies of NLRP7 in relation to maternal imprinting are based on the investigations of the expression of imprinted genes in molar tissue samples (Kou et al., 2008; Hayward et al., 2009; Sanchez-Delgado et al., 2015). Only a few functional studies of NLRP7 have been carried out in the field (Akoury et al., 2015; Singer et al., 2015; Mahadevan et al., 2014), while none of these studies provide direct evidence that NLRP7 plays a role in maternal imprinting. This is probably due to the expression of NLRP7 occurring mainly in human germ cells and early embryos and the limited availability of these human materials. In this project, I have successfully identified SNPs in imprinted genes in two hESC lines, applied naïve conversion in these hESCs as a cell model to mimic the process of global DNA demethylation and investigated the possible role of NLRP7 in protecting maternal imprinting. It has also been demonstrated that NLRP7 is abundantly expressed in naïve hPSCs, which may be employed as an alternative and unlimited resource to study NLRP7. In addition, this project found that overexpression of NLRP7 in primed hPSC may partially rescue maternal imprinting from global demethylation after naïve conversion, directly showing the protective role of NLRP7 on maternal imprinted genes in in vitro studies, for the first time. Finally, I have demonstrated a correlation between the expression of NLRP7 and DNMT3L during early embryonic development and the process of naïve conversion, which implies a potential link between them and that NLRP7 may mediate maternal imprinting via DNMT3L. This work also raises several interesting issues which could be useful to facilitate future research in this area.

6.1 hPSC naïve conversion as a model to study the DNA methylation reprogramming during embryonic development

To date, many groups have reported different culture systems to derive naïve hPSCs from routinely cultured primed hPSCs. Among them, the naïve cells maintained in 5iLAF and t2iLGo have recapitulated the most similarities to the ICM in preimplantation blastocysts, in terms of their response to different signaling activities, transcription profiles and epigenetic features (Collier and Rugg-Gunn, 2018). Noticeably, their global DNA methylation is dramatically erased during this process, including the DMRs of imprinted genes. Moreover, these hypomethylated DMRs cannot be restored when naïve cells are returned in culture to primed hESCs, even though global methylation becomes high again (Guo et al., 2017; Pastor et al., 2016). This is similar to a previous study which found the loss of genome-wide methylation in Dnmt1-KO mESC together with the global methylation could be rescued after re-expression of Dnmt1, except for the imprinted regions (Tucker et al, 1996). Although the exact demethylation mechanisms of hESC naïve conversion are not fully elucidated, a study on the transition of mESC from serum to 2iL condition showed the global demethylation was driven by the impaired maintenance DNA methylation due to reduced Uhrf1 and its binding marker H3K9me2, rather than Tet-dependent active demethylation (von Meyenn et al., 2016). However, it is worth noting that mESCs cultured in serum are in the metastable state (Kumar & Ivanova, 2015), which may not be the same as primed hPSCs, and there might be interspecies differences between human and mouse in terms of their DNA methylation/demethylation mechanisms. Similar to the mouse naïve conversion study, UHRF1 and DNMT1 protein levels have been shown to be reduced in 5iLAF naïve cells (Pastor et al., 2016), which indicates that a passive demethylation may be involved. However, it is unknown whether the upregulation of TETs in naïve hESCs only leads to the locus-specific demethylation or may also contribute to active global demethylation.

Even though it is not clear whether this culture-induced demethylation process is

similar to the PGC or postzygotic demethylation in vivo, they do share some common features, such as upregulation of TETs and reduced expression/accessibility of UHRF1 (Pastor et al., 2016; Yang et al., 2020; Gkountela et al., 2015; Maenohara et al, 2017). Therefore, the similarities and differences between naïve conversion and in vivo demethylation are going to be evaluated in detail below.

Since the differential methylation on imprinted genes is not protected during the naïve conversion, I first considered whether this process might resemble the deeper demethylation that occurs in hPGC. After migration and colonization of gonads, hPGCs are drastically demethylated to a very low level with only 7% methylation remaining at about 11 weeks after gestation (Guo et al., 2015), much lower than the global DNA methylation level (~30%) of naïve cells (Pastor et al., 2016). This wave of demethylation in hPGCs is considered to be a passive process, which results from diminished protein level of UHRF1 and DNMT3A (Monk et al., 2019; Gkountela et al., 2015). In mouse, when PGCs migrate towards the gonad (between E7.5 and E9.5), PGC methylation is thought to be passively erased due to the reduced level of Dnmt1 and Uhrf1 (Kurimoto et al., 2008; Hill et al., 2014). However, there seems to be another wave of demthylation after PGC entering into the gonad; it is currently unclear what mechanism is implicated in this process (Hill et al., 2018). Tet1 predominantly regulates locus-specific methylation level via maintaining the DNA demethylation status in gonadal PGCs (Hill et al., 2018). Additionally, by comparing the whole genome bisulfite sequencing data obtained in t2iLGo-converted naïve cells and hPGCs, naïve cells reveal unbridled demethylation on transposable elements (L1HS and L1PA2), which are normally maintained at high methylation level in hPGCs at different stages (Gkountela et al., 2015). Consequently, the reprogramming from hypermethylated primed hPSCs to hypomethylated naïve hPSCs may share similar passive demethylation mechanism during hPGC development with the loss of imprinting, while still differing in the extent of global demethylation and their targeting regions.

Noteworthy, naïve hESCs have been developed based on naïve mESCs which largely

resemble the ICM of mouse preimplantation blastocyst, so it is interesting to evaluate whether the global demethylation induced by naïve conversion mimics the global demethylation in the early cleavage stage of human embryos. Immediately after fertilization, the paternal genome is hypermethylated (>80%), and the maternal genome is moderately (~50%) methylated. Their genomes are erased asymmetrically in that the paternal genome is radically demethylated by a mechanism that is replication independent, while the maternal genome is gradually demethylated (Zhu et al., 2018; Monk et al., 2019). However, due to limited human embryonic materials, the mechanism of this wave of demethylation in human is less understood. According to mouse studies, after the first cell division, parallel demethylation of both parental genomes is mainly dependent on a passive dilution (Messerschmidt et al., 2014; Shen et al., 2014), despite active demethylation having also been reported to be involved (Wang et al., 2014). The high expression of Tet3 in mouse zygote predominantly helps to protect the newly acquired hypomethylation state from de novo methylation (Amouroux et al., 2016). Additionally, it has been shown that the DNA methylation landscape of human blastocysts is characterized by distinctive alternating hyper- and hypomethylated regions, similar to that of oocytes. However, 5iLAF and t2iLGo-converted naïve hESCs show a very different methylation pattern compared with human blastocyst, although both human blastocyst and naïve hESCs are hypomethylated with ~30% CG methylation (Pastor et al., 2013). This loss of oocyte-specific memory is consistent with the defective inheritance of imprinting status. Nonetheless, naïve conversion may still to some extent share similarity to post-fertilization demethylation, such as the predominant passive demethylation process and upregulated TET3. Therefore, even though naïve conversion may not be able to fully mimic the post-zygotic global demethylation on a physiological level, it may recapitulate some molecular mechanisms of DNA demethylation in vivo.

Based on the considerations above, the current naïve-to-primed model may provide a useful model to study the role of NLRP7 in maternal imprinting during postzygotic demethylation, rather than establishment of imprinting during oocyte maturation.

6.2 Role of NLRP7 in establishing/maintaining maternal imprinting

Analysis of the methylome of molar tissues collected from females with NLRP7 mutations has revealed a widespread loss of methylation at the DMRs of maternal imprinted genes (Sanchez-Delgado et al., 2015), which provides a strong link between NLRP7 and maternal imprinting. During oogenesis and early embryonic development, there are several timepoints when this maternal imprinting defect could originate: 1) an abnormality in de novo methylation during oogenesis that affects maternal imprinting establishment. 2) a defect in maintenance of methylation on the maternal DMRs during postzygotic demethylation in preimplantation embryos. 3) loss of protection on paternal alleles against post-implantation de novo methylation (Proudhon et al., 2012). However, the third point is not applicable to the proposed role of NLRP7 in recurrent BiCHM, because loss of NLRP7 does not affect methylation on paternal alleles, rather it exhibits hypomethylation on the DMR of maternally imprinted genes (Sanchez-Delgado et al., 2015). Therefore, NLRP7 might regulate maternal imprinting by facilitating de novo methylation during oogenesis or protecting maternal DMR from global demethylation in preimplantation embryos or via both mechanisms. Moreover, a recent study on KHDC3L, another minor gene involved in recurrent BiCHM, has recently been shown to play a role in the establishment of the maternal imprint in oocytes and in the protection of maternal imprinting after fertilization (Demond et al., 2019). This suggests that NLRP7 may have similar roles.

In order to investigate whether this methylation defect originates from abnormal maternal imprinting establishment or results from the lack of protection on maternal imprinted genes from post-fertilization demethylation, I studied the expression profile of NLRP7 throughout the different periods of human germ cell and embryonic development to explore the potential timepoint that NLRP7 may be involved.

In humans, the specification of germ cells starts from 2 weeks post-fertilization, followed by migration of hPGCs towards the genital ridge from week 3 when de novo DNA demethylation starts to occur (von Meyenn and Reik, 2015; Gkountela et al., 2015). According to the limited single-cell RNA-seq data obtained from human germline cells (Guo et al., 2015; Gkountela et al., 2015), NLRP7 mRNA expression in female hPGCs is extremely low when they are migrating to the developing gonads (week 4), then NLRP7 starts to increase when female hPGC are colonizing the genital ridge (week 7-8). Moreover, after the female germ cells start sex-specific differentiation and initiate meiosis (week 11), NLRP7 expression is rapidly increased and remains at a high level at least until week 17. Distinct from the significant increase of NLRP7 in female germ line, NLRP7 levels in male germ cells remain very low from week 4 to week 17, which suggests that NLRP7 may play a more important role in female germ cells.

During folliculogenesis in adult women, NLRP7 mRNA is stably expressed from primordial follicles to preovulatory follicles (Zhang et al., 2018; Yan et al., 2013). Immunostaining of oocytes acquired through superovulation reveals abundant expression of NLRP7 protein in the oocytes, even though its mRNA is expressed at a relatively low level (Akoury et al., 2015). Since the reestablishment of global DNA methylation in female germ cells takes place during the oocyte growth phase, the accumulation of NLRP7, a crucial maternal effect protein, in growing oocytes may assist the formation of oocyte-specific methylation pattern, including the DNA methylation on maternally imprinted genes. However, this proposed role of NLRP7 has not been proved by any research yet.

Although the upregulation of zygotic NLRP7 mRNA occurs in blastocysts (section 5.2.1), NLRP7 proteins generated during oocyte development are sustained at a high level throughout the post-fertilisation genomic demethylation in early human embryos (Akoury et al., 2015). Thus, NLRP7 may play an important role to protect methylation on the DMR of maternally imprinted genes from global demethylation in pre-

implantation embryos. My present study shows that overexpression of NLRP7 in hESCs lacking maternally inherited factors could partially rescue the loss of maternal imprinting (section 4.2.4). Consequently, NLRP7 is also likely to be involved in the maintenance of maternal DMRs during post-fertilization demethylation.

6.3 Potential molecular mechanisms by which NLRP7 may be involved in regulating maternal imprinting.

Given that the maternal deficiency of some SCMC members, including NLRP7, KHDC3L (Sanchez-Delgado et al., 2015; Demond et al., 2019), can cause BiHM due to the loss of maternal imprinting during early embryonic development and that *NLRP5* and *NLRP2* mutations are also associated with cases of imprinting disorders (Docherty et al., 2015; Meyer et al., 2009), NLRP7 has been thought to affect maternal imprinting through the SCMC. Particularly, studies of the genome-wide methylation profiles of KHDC3L-mutant oocytes and the resulting preimplantation embryo have shown that not only the establishment of methylation pattern is disturbed in mutant oocyte, but the preservation of maternal imprinted DMRs is also attenuated in the preimplantation blastocyst (Demond et al., 2019). Therefore, it is possible that both KHDC3L and NLRP7 affect maternal imprinting through similar mechanism that involve the SCMC. Although the SCMC is a subcortically localized complex, it may interact with epigenetic factors, such as DNMTs and histone modifiers, in the cytoplasm of oocytes or early embryonic cells to mediate their transportation into the nucleus thereby having a trans-acting effect on genomic imprinting.

Another possibility is that NLRP7 may regulate maternal imprinting in ways that are independent of the SCMC. After fertilization maternal transcripts and proteins are degraded or diluted along with zygotic cell cleavage, which may lead to the reduction of SCMC members, unless they can be compensated by zygotic gene activation after the 4-8 cell stage (Jukam et al., 2017). However, the transcription profiles of human

preimplantation embryos show that most transcripts of SCMC members are not increased upon ZGA, except NLRP7 and NLRP2 (Fig. 5.3, Yan et al., 2013; Petropoulos et al., 2016). Thus, NLRP7 may have a unique role independent of the SCMC, particularly in the early preimplantation human embryos. To date, a few studies have shown that NLRP7 is able to translocate to the cell nucleus which is consistent with my finding (section 5.2.5) (Mahadevan et al., 2014; Huang et al., 2017). It provides the possibility for NLRP7 to interact with epigenetic factors, such as the histone modifier YY1 (Mahadevan et al., 2014), as well as other putative partners DPPA3, ZFP57, TRIM28 and DNMT3L, to protect maternal imprinting during the post-fertilization demethylation process. In mice, maternally methylated DMRs can be protected by Dppa3 via binding H3K9me2 and blocking Tet3 activity (Nakamura et al., 2012). Furthermore, protection of both paternally and maternally imprinted DMRs from reprogramming requires Zinc finger protein 57 (Zfp57) that recruits another factor Tripartite motif-containing 28 (Trim28) to guide Dnmt1/3a/3b to specific loci via interacting with the methylated allele of ICRs (Li et al., 2008b; Quenneville et al., 2011). In addition, my study has indicated the potential interaction between NLRP7 and DNMT3L (section 5.2.6). DNMT3L, a significant co-factor for DNMT3A, is highly expressed in mouse oocytes, although it is not detected in human oocytes. However, DNMT3L may still possess the ability to recognize maternal imprint-specific histone markers, such as unmethylated H3K4, and guide NLRP7 to maternal imprinted loci in humans, thereby maintaining DNA methylation at the maternal DMR in human preimplantation embryos via either expelling TET enzymes or recruiting other DNMTs.

6.4 Future work

6.4.1 Validation of current findings

The immediate future work would be to validate the current findings. 1) As was mentioned in Chapter 4, the dynamic changes of global DNA methylation level during naïve conversion should be mapped by immunostaining or a DNA dot plot with

antibody against 5mC to obtain detailed information on the dynamic timeline between global demethylation and the increase of NLRP7. This would provide stronger evidence that dramatic NLRP7 upregulation is caused by the genomic DNA demethylation, so that loss of maternal imprinting in naïve conversion may be attributed, at least in part, to the low level of NLRP7 in primed hESCs. 2) Even though I have shown that NLRP7-OE in hESCs can partially rescue the maternal imprinting loss after their naïve conversion, only one experiment was completed in one hESC line (H7), therefore these experiments need to be repeated in more cell lines and more maternally imprinted genes needs to be recruited to validate this result.

6.4.2 Investigate whether NLRP7 affects the establishment of maternal imprinting during oogenesis

My current study has indicated that NLRP7 may play a role in maintaining maternal imprinting during zygotic genome-wide demethylation, but it is not clear whether NLRP7 may also affect the establishment of maternal imprinting during oogenesis. In the long term, there are two possible ways to further explore this question. Firstly, we could carry out a genome-wide analysis on the methylation landscape of oocytes produced by patients with *NLRP7*-mutations compared with normal oocytes to identify whether there is any global or imprint-specific reduction of DNA methylation in NLRP7-defective oocytes. A similar experiment has been performed in the oocytes of a woman with homozygous *KHDC3L* mutations (Demond et al., 2019). This approach requires identification of patients and ethical approval. Alternatively, NLRP7 could be knocked out in hESCs that can be reprogrammed to PGC-like cells (PGCLCs) (Irie and Surani, 2017), and further induced to oocytes to mimic the de novo methylation process in vivo, to investigate whether the establishment of maternal imprinting is affected. However, unlike mouse PGCLCs that can contribute to gametogenesis, human PGCLCs have not successfully been developed to mature germ cells, even though they share many gene-expression properties with human PGCs (Yamashiro and Sasaki, 2018).

Therefore, this idea could only be achieved when human in vitro oogenesis become fully developed and relative ethical regulation has been established.

6.4.3 Explore the mechanisms by which NLRP7 affects maternal imprinting

Up to the present time, most publications on NLRP7 are case reports, rather than mechanistic studies, due to the limited availability of human materials. Since almost all members of the SCMC, including NLRP7, show a high expression level in naïve hESCs, the naïve cells can be used as an unlimited resource to explore the interaction between NLRP7 and other proteins, such as SCMC members and DNMT3L, through NLRP7 immunoprecipitation, mass spectrometry analysis and further validation via co-IP, co-immunostaining and proximity ligation assay. Finally the role of NLRP7 could be investigated by generating NLRP7-knockout hESCs and using their derived naïve hESCs, or PGCLCs if possible, to identify differential gene expression and epigenetic profiling compared with WT cells.

References

Abeyta, M. J., Clark, A. T., Rodriguez, R. T., Bodnar, M. S., Pera, R. A. & Firpo, M. T. 2004. Unique gene expression signatures of independently-derived human embryonic stem cell lines. *Hum Mol Genet,* 13(6)**,** pp 601-608.

Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P. W., Beighton, G., Bello, P. A., Benvenisty, N., Berry, L. S. & Bevan, S. 2007. Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat biotechnol,* 25(7)**,** pp 803-816.

Akoury, E., Zhang, L., Ao, A. & Slim, R. 2015. NLRP7 and KHDC3L, the two maternal-effect proteins responsible for recurrent hydatidiform moles, co-localize to the oocyte cytoskeleton. *Hum Reprod,* 30(1)**,** pp 159-169.

Amit, M., Carpenter, M. K., Inokuma, M. S., Chiu, C. P., Harris, C. P., Waknitz, M. A., Itskovitz-Eldor, J. & Thomson, J. A. 2000. Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. *Dev Biol,* 227(2)**,** pp 271-278.

Amouroux, R., Nashun, B., Shirane, K., Nakagawa, S., Hill, P.W., D'Souza, Z., Nakayama, M., Matsuda, M., Turp, A., Ndjetehe, E. and Encheva, V., 2016. De novo DNA methylation drives 5hmC accumulation in mouse zygotes. *Nature cell biology*, *18*(2), pp.225-233.

Auclair, G., Guibert, S., Bender, A. & Weber, M. 2014. Ontogeny of CpG island methylation and specificity of DNMT3 methyltransferases during embryonic development in the mouse. *Genome Biol,* 15(12)**,** p 545.

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

Barlow, D. P. & Bartolomei, M. S. 2014. Genomic imprinting in mammals. *Cold Spring Harb Perspect Biol,* 6(2)**,** pp a018382.

Bartholdi, D., Krajewska-Walasek, M., Ounap, K., Gaspar, H., Chrzanowska, K. H., Ilyana, H., Kayserili, H., Lurie, I. W., Schinzel, A. & Baumer, A. 2009. Epigenetic mutations of the imprinted IGF2-H19 domain in Silver-Russell syndrome (SRS): results from a large cohort of patients with SRS and SRS-like phenotypes. *J Med Genet,* 46(3)**,** pp 192-197.

Bebbere, D., Masala, L., Albertini, D. F. & Ledda, S. 2016. The subcortical maternal complex: multiple functions for one biological structure? *J Assist Reprod Genet,* 33(11), pp 1431-1438.

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

Bradley, A., Evans, M., Kaufman, M.H., Robertson, E., 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature,* 309(5965), pp 255-256.

Brannan, C. I., Dees, E.C., Ingram, R.S., Tilghman, S.M., 1990. The product of the H19 gene may function as an RNA. *Mol cell biol,* 10(1), pp 28-36.

Bronner, C., Alhosin, M., Hamiche, A. & Mousli, M. 2019. Coordinated Dialogue between UHRF1 and DNMT1 to Ensure Faithful Inheritance of Methylated DNA Patterns. *Genes,* 10(1), p 65

Brons, I. G., Smithers, L. E., Trotter, M. W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. A. & Vallier, L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature,* 448(7150), pp 191-195.

Buiting, K. 2010. Prader-Willi syndrome and Angelman syndrome. *Am Journal of Med Genet C: Seminars in Medical Genetics,* 154C(3), pp 365-376.

Cedar, H. & Bergman, Y. 2009. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet,* 10(5), pp 295-304.

Chan, Y. S., Goke, J., Ng, J. H., Lu, X., Gonzales, K. A., Tan, C. P., Tng, W. Q., Hong, Z. Z., Lim, Y. S. & Ng, H. H. 2013. Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell,* 13(6), pp 663-675.

Chedin, F., Lieber, M. R. & Hsieh, C. L. 2002. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc Natl Acad Sci U S A,* 99(26), pp 16916-16921.

Choi, J., Huebner, A. J., Clement, K., Walsh, R. M., Savol, A., Lin, K., Gu, H., Di Stefano, B., Brumbaugh, J., Kim, S. Y., Sharif, J., Rose, C. M., Mohammad, A., Odajima, J., Charron, J., Shioda, T., Gnirke, A., Gygi, S., Koseki, H., Sadreyev, R. I., Xiao, A., Meissner, A. & Hochedlinger, K. 2017. Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature,* 548(7666), pp 219-223.

Ciccone, D. N., Su, H., Hevi, S., Gay, F., Lei, H., Bajko, J., Xu, G., Li, E. & Chen, T. 2009. KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature,* 461(7262)**,** pp 415-418.

Cirio, M. C., Martel, J., Mann, M., Toppings, M., Bartolomei, M., Trasler, J. & Chaillet, J. R. 2008. DNA methyltransferase 1o functions during preimplantation development to preclude a profound level of epigenetic variation. *Dev Biol,* 324(1)**,** pp 139-150.

Collier, A. J. & Rugg-Gunn, P. J. 2018. Identifying Human Naïve Pluripotent Stem Cells - Evaluating State-Specific Reporter Lines and Cell-Surface Markers. *Bioessays,* 40(5)**,** pp e1700239.

Costa, Y., Ding, J., Theunissen, T.W., Faiola, F., Hore, T.A., Shliaha, P.V., Fidalgo, M., Saunders, A., Lawrence, M., Dietmann, S. and Das, S., 2013. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*, 495(7441), pp370-374.

Darst, R.P., Pardo, C.E., Ai, L., Brown, K.D. and Kladde, M.P., 2010. Bisulfite sequencing of DNA. *Current protocols in molecular biology*, *91*(1), pp.7-9.

De Los Angeles, A., Loh, Y.-H., Tesar, P. J. & Daley, G. Q. 2012. Accessing naïve human pluripotency. *Curr Opin Genet Dev,* 22(3)**,** pp 272-282.

Delaval, K. & Feil, R. 2004. Epigenetic regulation of mammalian genomic imprinting. *Curr Opin Genet Dev,* 14(2)**,** pp 188-195.

Demond, H., Anvar, Z., Jahromi, B. N., Sparago, A., Verma, A., Davari, M., Calzari, L., Russo, S., Jahromi, M. A., Monk, D., Andrews, S., Riccio, A. & Kelsey, G. 2019. A KHDC3L mutation resulting in recurrent hydatidiform mole causes genome-wide DNA methylation loss in oocytes and persistent imprinting defects post-fertilisation. *Genome Med,* 11(1)**,** pp 1-14.

Deveault, C., Qian, J. H., Chebaro, W., Ao, A., Gilbert, L., Mehio, A., Khan, R., Tan, S. L., Wischmeijer, A., Coullin, P., Xie, X. & Slim, R. 2009. NLRP7 mutations in women with diploid androgenetic and triploid moles: a proposed mechanism for mole formation. *Hum Mol Genet,* 18(5)**,** pp 888-897.

Dixon, P. H., Trongwongsa, P., Abu-Hayyah, S., Ng, S. H., Akbar, S. A., Khawaja, N. P., Seckl, M. J., Savage, P. M. & Fisher, R. A. 2012. Mutations in NLRP7 are associated with diploid biparental hydatidiform moles, but not androgenetic complete moles. *J Med Genet,* 49(3)**,** pp 206-211.

Docherty, L. E., Rezwan, F. I., Poole, R. L., Turner, C. L., Kivuva, E., Maher, E. R., Smithson, S. F., Hamilton-Shield, J. P., Patalan, M., Gizewska, M., Peregud-

Pogorzelski, J., Beygo, J., Buiting, K., Horsthemke, B., Soellner, L., Begemann, M., Eggermann, T., Baple, E., Mansour, S., Temple, I. K. & Mackay, D. J. 2015. Mutations in NLRP5 are associated with reproductive wastage and multilocus imprinting disorders in humans. *Nat Commun,* 6(1), pp 1-7.

Dome, J. S. & Coppes, M. J. 2002. Recent advances in Wilms tumor genetics. *Curr Opin Pediatr,* 14(1)**,** pp 5-11.

Duenez-Guzman, E. A. & Haig, D. 2014. The evolution of reproduction-related NLRP genes. *J Mol Evol,* 78(3-4)**,** pp 194-201.

El-Maarri, O., Seoud, M., Coullin, P., Herbiniaux, U., Oldenburg, J., Rouleau, G. & Slim, R. 2003. Maternal alleles acquiring paternal methylation patterns in biparental complete hydatidiform moles. *Hum Mol Genet,* 12(12)**,** pp 1405-1413.

Engel, E. and Antonarakis, S.E., 2004. *Genomic imprinting and uniparental disomy in medicine: clinical and molecular aspects*. John Wiley & Sons.

Evans, M. J. & Kaufman, M. H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature,* 292(5819)**,** pp 154-156.

Ferguson-Smith, A. C. (2011). Genomic imprinting: the emergence of an epigenetic paradigm. *Nat Rev Genetic*s, 12(8), 565-575.

Fernandes, R., Tsuda, C., Perumalsamy, A. L., Naranian, T., Chong, J., Acton, B. M., Tong, Z. B., Nelson, L. M. & Jurisicova, A. 2012. NLRP5 mediates mitochondrial function in mouse oocytes and embryos. *Biol Reprod,* 86(5)**,** pp 138, 1-10.

Fisher, R. A. & Hodges, M. D. 2003. Genomic Imprinting in Gestational Trophoblastic Disease—A Review. *Placenta,* 24, pp S111-S118.

Fisher, R. A., Hodges, M. D., Rees, H. C., Sebire, N. J., Seckl, M. J., Newlands, E. S., Genest, D. R. & Castrillon, D. H. 2002. The maternally transcribed gene p57KIP2 (CDNK1C) is abnormally expressed in both androgenetic and biparental complete hydatidiform moles. *Hum Mol Genet,* 11(26)**,** pp 3267-3272.

Fisher, R. A., Lavery, S. A., Carby, A., Abu-Hayyeh, S., Swingler, R., Sebire, N. J. & Seckl, M. J. 2011. What a difference an egg makes. *The Lancet,* 378(9807)**,** p 1974

Frost, J. M., Monk, D., Moschidou, D., Guillot, P. V., Stanier, P., Minger, S. L., Fisk, N. M., Moore, H. D. & Moore, G. E. 2011. The effects of culture on genomic imprinting profiles in human embryonic and fetal mesenchymal stem cells. *Epigenetics,* 6(1)**,** pp 52-62.

Frost, J.M., Monk, D., Stojilkovic-Mikic, T., Woodfine, K., Chitty, L.S., Murrell, A., Stanier, P. and Moore, G.E., 2010. Evaluation of allelic expression of imprinted genes in adult human blood. *PloS one*, *5*(10), p.e13556

Gafni, O., Weinberger, L., Mansour, A. A., Manor, Y. S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., Rais, Y., Shipony, Z., Mukamel, Z., Krupalnik, V., Zerbib, M., Geula, S., Caspi, I., Schneir, D., Shwartz, T., Gilad, S., Amann-Zalcenstein, D., Benjamin, S., Amit, I., Tanay, A., Massarwa, R., Novershtern, N. & Hanna, J. H. 2013. Derivation of novel human ground state naïve pluripotent stem cells. *Nature,* 504(7479)**,** pp 282-286.

Gao, Z., Zhang, X., Yu, X., Qin, D., Xiao, Y., Yu, Y., Xiang, Y., Nie, X., Lu, X., Liu, W., Yi, Z. & Li, L. 2018. Zbed3 participates in the subcortical maternal complex and regulates the distribution of organelles. *J Mol Cell Biol,* 10(1)**,** pp 74-88.

Gerrard, L., Rodgers, L. & Cui, W. 2005. Differentiation of human embryonic stem cells to neural lineages in adherent culture by blocking bone morphogenetic protein signaling. *Stem Cells,* 23(9)**,** pp 1234-1241.

Gkountela, S., Zhang, K. X., Shafiq, T. A., Liao, W. W., Hargan-Calvopina, J., Chen, P. Y. & Clark, A. T. 2015. DNA Demethylation Dynamics in the Human Prenatal Germline. *Cell,* 161(6)**,** pp 1425-1436.

Greenberg, M. V. C. & Bourc'his, D. 2019. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol,* 20(10)**,** pp 590-607.

Gu, T. P., Guo, F., Yang, H., Wu, H. P., Xu, G. F., Liu, W., Xie, Z. G., Shi, L., He, X., Jin, S. G., Iqbal, K., Shi, Y. G., Deng, Z., Szabo, P. E., Pfeifer, G. P., Li, J. & Xu, G. L. 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature,* 477(7366)**,** pp 606-610.

Guenatri, M., Duffie, R., Iranzo, J., Fauque, P. & Bourc'his, D. 2013. Plasticity in Dnmt3L-dependent and -independent modes of de novo methylation in the developing mouse embryo. *Development,* 140(3)**,** pp 562-572.

Guibert, S., Forne, T. & Weber, M. 2012. Global profiling of DNA methylation erasure in mouse primordial germ cells. *Genome Res,* 22(4)**,** pp 633-641.

Guo, F., Yan, L., Guo, H., Li, L., Hu, B., Zhao, Y., Yong, J., Hu, Y., Wang, X., Wei, Y., Wang, W., Li, R., Yan, J., Zhi, X., Zhang, Y., Jin, H., Zhang, W., Hou, Y., Zhu, P., Li, J., Zhang, L., Liu, S., Ren, Y., Zhu, X., Wen, L., Gao, Y. Q., Tang, F. & Qiao, J. 2015. The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. *Cell,* 161(6)**,** pp 1437-1452.

Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W., Plath, K. & Smith, A. 2017a. Epigenetic resetting of human pluripotency. *Development,* 144(15), pp 2748-2763.

Guo, H., Hu, B., Yan, L., Yong, J., Wu, Y., Gao, Y., Guo, F., Hou, Y., Fan, X., Dong, J., Wang, X., Zhu, X., Yan, J., Wei, Y., Jin, H., Zhang, W., Wen, L., Tang, F. & Qiao, J. 2017b. DNA methylation and chromatin accessibility profiling of mouse and human fetal germ cells. *Cell Res,* 27(2)**,** pp 165-183.

Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., Yan, J., Ren, X., Lin, S., Li, J., Jin, X., Shi, X., Liu, P., Wang, X., Wang, W., Wei, Y., Li, X., Guo, F., Wu, X., Fan, X., Yong, J., Wen, L., Xie, S. X., Tang, F. & Qiao, J. 2014. The DNA methylation landscape of human early embryos. *Nature,* 511(7511)**,** pp 606-610.

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

Hager, R., Cheverud, J. M. & Wolf, J. B. 2008. Maternal effects as the cause of parent-of-origin effects that mimic genomic imprinting. *Genetics,* 178(3)**,** pp 1755-1762.

Hanin, G., & Ferguson-Smith, A. C. (2020). The evolution of genomic imprinting: Epigenetic control of mammary gland development and postnatal resource control. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *12*(3), e1476.

Hanna, J., Cheng, A. W., Saha, K., Kim, J., Lengner, C. J., Soldner, F., Cassady, J. P., Muffat, J., Carey, B. W. & Jaenisch, R. 2010. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc Natl Acad Sci U S A,* 107(20)**,** pp 9222-9227.

Hata, K., Okano, M., Lei, H. & Li, E. 2002. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development,* 129(8)**,** pp 1983-1993.

Hattori, H., Hiura, H., Kitamura, A., Miyauchi, N., Kobayashi, N., Takahashi, S. & Arima, T. (2019). Association of four imprinting disorders and ART. *Clin epigenetics*, 11(1), 21.

Hayward, B. E., De Vos, M., Judson, H., Hodge, D., Huntriss, J. D., Picton, H. M., Sheridan, E. & Bonthron, D. T. 2003. Lack of involvement of known DNA methyltransferases in familial hydatidiform mole implies the involvement of other factors in establishment of imprinting in the human female germline. *BMC genetics,* 4(1)**,** pp1-8

Hayward, B.E., De Vos, M., Talati, N., Abdollahi, M.R., Taylor, G.R., Meyer, E., Williams, D., Maher, E.R., Setna, F., Nazir, K. and Hussaini, S., 2009. Genetic and epigenetic analysis of recurrent hydatidiform mole. *Hum Mutat*, 30(5), pp E629-E639.

He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L. & Sun, Y. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science,* 333(6047)**,** pp 1303-1307.

Hill, P.W., Leitch, H.G., Requena, C.E., Sun, Z., Amouroux, R., Roman-Trufero, M., Borkowska, M., Terragni, J., Vaisvila, R., Linnett, S. and Bagci, H., 2018. Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. *Nature*, *555*(7696), pp.392-396.

Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R. & Sasaki, H. 2008. Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev,* 22(12)**,** pp 1607-1616.

Hiura, H., Okae, H., Chiba, H., Miyauchi, N., Sato, F., Sato, A., & Arima, T. (2014). Imprinting methylation errors in ART. *Reproductive medicine and biology*, *13*(4), 193-202.

Hore, T. A., von Meyenn, F., Ravichandran, M., Bachman, M., Ficz, G., Oxley, D., Santos, F., Balasubramanian, S., Jurkowski, T. P. & Reik, W. 2016. Retinol and ascorbate drive erasure of epigenetic memory and enhance reprogramming to naïve pluripotency by complementary mechanisms. *Proc Natl Acad Sci U S A,* 113(43)**,** pp 12202-12207.

Huang, J. Y., Yu, P. H., Li, Y. C. & Kuo, P. L. 2017. NLRP7 contributes to in vitro decidualization of endometrial stromal cells. *Reprod Biol Endocrinol,* 15(1)**,** pp 1-10.

Huang, Y., Osorno, R., Tsakiridis, A. & Wilson, V. 2012. In Vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. *Cell Rep,* 2(6)**,** pp 1571-1578.

Huntriss, J., Hinkins, M., Oliver, B., Harris, S. E., Beazley, J. C., Rutherford, A. J., Gosden, R. G., Lanzendorf, S. E. & Picton, H. M. 2004. Expression of mRNAs for DNA methyltransferases and methyl-CpG-binding proteins in the human female germ line, preimplantation embryos, and embryonic stem cells. *Mol Reprod Dev,* 67(3)**,** pp 323-336.

Irie, N. & Surani, M. A. 2017. *Efficient Induction and Isolation of Human Primordial Germ*

Cell-Like Cells from Competent Human Pluripotent Stem Cells, New York: Humana Press.

Jeziorska, D. M., Murray, R. J. S., De Gobbi, M., Gaentzsch, R., Garrick, D., Ayyub, H., Chen, T., Li, E., Telenius, J., Lynch, M., Graham, B., Smith, A. J. H., Lund, J. N., Hughes, J. R., Higgs, D. R. & Tufarelli, C. 2017. DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. *Proc Natl Acad Sci U S A,* 114(36)**,** pp E7526-E7535.

Jin, B., Li, Y. & Robertson, K. D. 2011. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer,* 2(6)**,** pp 607-617.

Jin, Y., Mailloux, C. M., Gowan, K., Riccardi, S. L., LaBerge, G., Bennett, D. C., Fain, P. R. & Spritz, R. A. 2007. NALP1 in Vitiligo-associated autoimmune disease. *N Engl J Med,* 356(12)**,** pp 1216-1225.

Johnson, D. R. 1974. Hairpin tail: a case of post reductional gene action in the mouse EGG? *Genetics,* 76(4)**,** pp 795-805.

Jukam, D., Shariati, S. A. M. & Skotheim, J. M. 2017. Zygotic Genome Activation in Vertebrates. *Dev Cell,* 42(4)**,** pp 316-332.

Kaffer, C. R., Grinberg, A. & Pfeifer, K. 2001. Regulatory mechanisms at the mouse Igf2/H19 locus. *Mol Cell Biol,* 21(23)**,** pp 8189-8196.

Kagami, M., O'Sullivan, M. J., Green, A. J., Watabe, Y., Arisaka, O., Masawa, N., Matsuoka, K., Fukami, M., Matsubara, K., Kato, F., Ferguson-Smith, A. C. & Ogata, T. 2010. The IG-DMR and the MEG3-DMR at human chromosome 14q32.2: hierarchical interaction and distinct functional properties as imprinting control centers. *PLoS Genet,* 6(6)**,** pp e1000992.

Kaku, K., Osada, H., Seki, K. & Sekiya, S. 2007. Insulin-like growth factor 2 (IGF2) and IGF2 receptor gene variants are associated with fetal growth. *Acta Paediatr,* 96(3)**,** pp 363-367.

Kalscheuer, V. M., Mariman, E.C., Schepens, M.T., Rehder, H., Ropers, H.H., 1993. The insulin–like growth factor type–2 receptor gene is imprinted in the mouse but not in humans. *Nat Genet,* 5(1)**,** pp 74-78.

Kan, R., Yurttas, P., Kim, B., Jin, M., Wo, L., Lee, B., Gosden, R. & Coonrod, S. A. 2011. Regulation of mouse oocyte microtubule and organelle dynamics by PADI6 and the cytoplasmic lattices. *Dev Biol,* 350(2)**,** pp 311-322.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E. & Sasaki, H. 2004.

Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature,* 429(6994)**,** pp 900-903.

Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y., Okano, M., Li, E., Nozaki, M. & Sasaki, H. 2007. 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,* 16(19)**,** pp 2272-2280.

Kelsey, G. & Feil, R. 2013. New insights into establishment and maintenance of DNA methylation imprints in mammals. *Philos Trans R Soc Lond B Biol Sci,* 368(1609)**,** pp 20110336.

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

Kim, B., Zhang, X., Kan, R., Cohen, R., Mukai, C., Travis, A. J. & Coonrod, S. A. 2014. The role of MATER in endoplasmic reticulum distribution and calcium homeostasis in mouse oocytes. *Dev Biol,* 386(2)**,** pp 331-339.

Kinoshita, T., Wang, Y., Hasegawa, M., Imamura, R. & Suda, T. 2005. PYPAF3, a PYRIN-containing APAF-1-like protein, is a feedback regulator of caspase-1-dependent interleukin-1beta secretion. *J Biol Chem,* 280(23)**,** pp 21720-21725.

Kobayashi, H., Sakurai, T., Imai, M., Takahashi, N., Fukuda, A., Yayoi, O., Sato, S., Nakabayashi, K., Hata, K., Sotomaru, Y., Suzuki, Y. & Kono, T. 2012. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLoS Genet,* 8(1)**,** pp e1002440.

Kou, Y.C., Shao, L., Peng, H.H., Rosetta, R., Del Gaudio, D., Wagner, A.F., Al-Hussaini, T.K. and Van den Veyver, I.B., 2007. A recurrent intragenic genomic duplication, other novel mutations in NLRP7 and imprinting defects in recurrent biparental hydatidiform moles. *MHR: Basic Sci Reprod Med*, 14(1), pp 33-40.

Kumar, I. and Ivanova, N., 2015. Moving toward the ground state. *Cell stem cell*, *17*(4), pp.375-376.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K. and Saitou, M., 2008. Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. *Genes & development*, *22*(12), pp 1617-1635.

Latz, E., Xiao, T. S. & Stutz, A. 2013. Activation and regulation of the inflammasomes. *Nat Rev Immunol,* 13(6)**,** pp 397-411.

Lee, H. J., Hore, T. A. & Reik, W. 2014. Reprogramming the methylome: erasing memory and creating diversity. *Cell Stem Cell,* 14(6)**,** pp 710-719.

Lehnertz, B., Ueda, Y., Derijck, A. A. H. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. & Peters, A. H. F. M. 2003. Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin. *Curr Biol,* 13(14)**,** pp 1192-1200.

Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A. and Surani, M.A., 2013. Naive pluripotency is associated with global DNA hypomethylation. *Nat Struct Mol Biol*, *20*(3), p.311.

Li, C., Liu, B., Zhong, S. & Wang, H. 2016. MEK inhibitor PD0325901 and vitamin C synergistically induce hypomethylation of mouse embryonic stem cells. *Oncotarget,* 7(26)**,** pp 39730.

Li, E., Bestor, T.H., & Jaenisch, R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell,* 69(6)**,** pp 915-926.

Li, E. & Zhang, Y. 2014. DNA methylation in mammals. *Cold Spring Harb Perspect Biol,* 6(5)**,** pp a019133.

Li, G., Tian, X., Lv, D., Zhang, L., Zhang, Z., Wang, J., Yang, M., Tao, J., Ma, T., Wu, H., Ji, P., Wu, Y., Lian, Z., Cui, W. & Liu, G. 2019. NLRP7 is expressed in the ovine ovary and associated with in vitro pre-implantation embryo development. *Reproduction,* 158(5)**,** pp 415-427.

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

Li, L., Zheng, P. & Dean, J. 2010. Maternal control of early mouse development. *Development,* 137(6)**,** pp 859-870.

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P. & Ferguson-Smith, A. C. 2008b. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev Cell,* 15(4)**,** pp 547-557.

Liao, J., Karnik, R., Gu, H., Ziller, M. J., Clement, K., Tsankov, A. M., Akopian, V., Gifford, C. A., Donaghey, J., Galonska, C., Pop, R., Reyon, D., Tsai, S. Q., Mallard, W., Joung, J. K., Rinn, J. L., Gnirke, A. & Meissner, A. 2015. Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat Genet,* 47(5)**,** pp 469-478.

Liu, X., Nefzger, C. M., Rossello, F. J., Chen, J., Knaupp, A. S., Firas, J., Ford, E., Pflueger,

J., Paynter, J. M., Chy, H. S., O'Brien, C. M., Huang, C., Mishra, K., Hodgson-Garms, M., Jansz, N., Williams, S. M., Blewitt, M. E., Nilsson, S. K., Schittenhelm, R. B., Laslett, A. L., Lister, R. & Polo, J. M. 2017. Comprehensive characterization of distinct states of human naïve pluripotency generated by reprogramming. *Nat Methods,* 14(11), pp 1055-1062.

Lu, F., Liu, Y., Jiang, L., Yamaguchi, S. & Zhang, Y. 2014. Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev,* 28(19), pp 2103-2119.

Lu, X., Gao, Z., Qin, D. & Li, L. 2017. A Maternal Functional Module in the Mammalian Oocyte-To-Embryo Transition. *Trends Mol Med,* 23(11), pp 1014-1023.

Ludwig, T. E., Bergendahl, V., Levenstein, M. E., Yu, J., Probasco, M. D. & Thomson, J. A. 2006. Feeder-independent culture of human embryonic stem cells. *Nat Methods,* 3(8), pp 637-646.

MacDonald, J. A., Wijekoon, C. P., Liao, K. C. & Muruve, D. A. 2013. Biochemical and structural aspects of the ATP-binding domain in inflammasome-forming human NLRP proteins. *IUBMB Life,* 65(10), pp 851-862.

MacDonald, W. A. & Mann, M. R. W. 2020. Long noncoding RNA functionality in imprinted domain regulation. *PLOS Genet,* 16(8), pp e1008930

Maenohara, S., Unoki, M., Toh, H., Ohishi, H., Sharif, J., Koseki, H. and Sasaki, H., 2017. Role of UHRF1 in de novo DNA methylation in oocytes and maintenance methylation in preimplantation embryos. *PLoS genetics*, *13*(10), p.e1007042.

Mahadevan, S., Sathappan, V., Utama, B., Lorenzo, I., Kaskar, K. & Van den Veyver, I. B. 2017. Maternally expressed NLRP2 links the subcortical maternal complex (SCMC) to fertility, embryogenesis and epigenetic reprogramming. *Sci Rep,* 7, p 44667.

Mahadevan, S., Wen, S., Balasa, A., Fruhman, G., Mateus, J., Wagner, A., Al-Hussaini, T. & Van den Veyver, I. B. 2013. No evidence for mutations in NLRP7 and KHDC3L in women with androgenetic hydatidiform moles. *Prenat Diagn,* 33(13), pp 1242-1247.

Mahadevan, S., Wen, S., Wan, Y. W., Peng, H. H., Otta, S., Liu, Z., Iacovino, M., Mahen, E. M., Kyba, M., Sadikovic, B. & Van den Veyver, I. B. 2014. NLRP7 affects trophoblast lineage differentiation, binds to overexpressed YY1 and alters CpG methylation. *Hum Mol Genet,* 23(3), pp 706-716.

Maher, E. R., Afnan, M. & Barratt, C. L. 2003. Epigenetic risks related to assisted reproductive technologies: epigenetics, imprinting, ART and icebergs? *Hum Reprod,* 18(12), pp 2508-2511.

Marks, H. & Stunnenberg, H. G. 2014. Transcription regulation and chromatin structure in the pluripotent ground state. *Biochim Biophys Acta,* 1839(3)**,** pp 129-137.

Martinon, F., Burns, K. & Tschopp, J. 2002. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. *Mol Cell,* 10(2)**,** pp 417-426.

McDaniel, P. & Wu, X. 2009. Identification of oocyte-selective NLRP genes in rhesus macaque monkeys (Macaca mulatta). *Mol Reprod Dev,* 76(2)**,** pp 151-159.

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

McMurray, E. N. & Schmidt, J. V. 2012. Identification of imprinting regulators at the Meg3 differentially methylated region. *Genomics,* 100(3)**,** pp 184-194.

Messaed, C., Akoury, E., Djuric, U., Zeng, J., Saleh, M., Gilbert, L., Seoud, M., Qureshi, S. & Slim, R. 2011. NLRP7, a nucleotide oligomerization domain-like receptor protein, is required for normal cytokine secretion and co-localizes with Golgi and the microtubule-organizing center. *J Biol Chem,* 286(50)**,** pp 43313-43323.

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

Meyer, E., Lim, D., Pasha, S., Tee, L. J., Rahman, F., Yates, J. R., Woods, C. G., Reik, W. & Maher, E. R. 2009. Germline mutation in NLRP2 (NALP2) in a familial imprinting disorder (Beckwith-Wiedemann Syndrome). *PLoS Genet,* 5(3)**,** pp e1000423.

Moglabey, Y. B., Kircheisen, R., Seoud, M., El Mogharbel, N., Van den Veyver, I. & Slim, R. 1999. Genetic mapping of a maternal locus responsible for familial hydatidiform moles. *Hum Mol Genet,* 8(4)**,** pp 667-667.

Monk, D., Mackay, D. J. G., Eggermann, T., Maher, E. R. & Riccio, A. 2019. Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. *Nat Rev Genet,* 20(4)**,** pp 235-248.

Monk, D., Sanchez-Delgado, M. & Fisher, R. 2017. NLRPs, the subcortical maternal complex and genomic imprinting. *Reproduction,* 154(6)**,** pp R161-R170.

Moore, L. D., Le, T. & Fan, G. 2013. DNA methylation and its basic function. *Neuropsychopharmacology,* 38(1)**,** pp 23-38.

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

Mu, J., Wang, W., Chen, B., Wu, L., Li, B., Mao, X., Zhang, Z., Fu, J., Kuang, Y., Sun, X. & Li, Q. 2019. Mutations in NLRP2 and NLRP5 cause female infertility characterised by early embryonic arrest. *J Med Genet,* 56(7)**,** pp 471-480.

Murdoch, S., Djuric, U., Mazhar, B., Seoud, M., Khan, R., Kuick, R., Bagga, R., Kircheisen, R., Ao, A. & Ratti, B. 2006. Mutations in NALP7 cause recurrent hydatidiform moles and reproductive wastage in humans. *Nat Genet,* 38(3)**,** pp 300-302.

Nakamura, T., Liu, Y. J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S., Tachibana, M., Ogura, A., Shinkai, Y. & Nakano, T. 2012. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature,* 486(7403)**,** pp 415-419.

Nakamura, T., Okamoto, I., Sasaki, K., Yabuta, Y., Iwatani, C., Tsuchiya, H., Seita, Y., Nakamura, S., Yamamoto, T. & Saitou, M. 2016. A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature,* 537(7618)**,** pp 57-62.

Ni, K., Dansranjavin, T., Rogenhofer, N., Oeztuerk, N., Deuker, J., Bergmann, M., Schuppe, H.C., Wagenlehner, F., Weidner, W., Steger, K. and Schagdarsurengin, U., 2016. TET enzymes are successively expressed during human spermatogenesis and their expression level is pivotal for male fertility. *Hum Reprod*, 31(7), pp.1411-1424.

Nichols, J. & Smith, A. 2009. Naïve and primed pluripotent states. *Cell Stem Cell,* 4(6)**,** pp 487-492.

Ohsugi, M., Zheng, P., Baibakov, B., Li, L. & Dean, J. 2008. Maternally derived FILIA-MATER complex localizes asymmetrically in cleavage-stage mouse embryos. *Development,* 135(2)**,** pp 259-269.

Okano, M., Bell, D. W., Haber, D. A. & Li, E. 1999. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell,* 99(3), pp 247-257.

Okashita, N., Kumaki, Y., Ebi, K., Nishi, M., Okamoto, Y., Nakayama, M., Hashimoto, S., Nakamura, T., Sugasawa, K., Kojima, N. and Takada, T., 2014. PRDM14 promotes active DNA demethylation through the ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development*, 141(2), pp 269-280.

Ooi, S. K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S. P., Allis, C. D., Cheng, X. & Bestor, T. H. 2007. DNMT3L connects

unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature,* 448(7154), pp 714-717.

Parry, D. A., Logan, C. V., Hayward, B. E., Shires, M., Landolsi, H., Diggle, C., Carr, I., Rittore, C., Touitou, I., Philibert, L., Fisher, R. A., Fallahian, M., Huntriss, J. D., Picton, H. M., Malik, S., Taylor, G. R., Johnson, C. A., Bonthron, D. T. & Sheridan, E. G. 2011. Mutations causing familial biparental hydatidiform mole implicate c6orf221 as a possible regulator of genomic imprinting in the human oocyte. *Am J Hum Genet,* 89(3), pp 451-458.

Pastor, W. A., Aravind, L. & Rao, A. 2013. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Rev Mol Cell Biol,* 14(6), pp 341-356.

Pastor, W. A., Chen, D., Liu, W., Kim, R., Sahakyan, A., Lukianchikov, A., Plath, K., Jacobsen, S. E. & Clark, A. T. 2016. Naïve Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory. *Cell Stem Cell,* 18(3), pp 323-329.

Peng, H., Chang, B., Lu, C., Su, J., Wu, Y., Lv, P., Wang, Y., Liu, J., Zhang, B., Quan, F., Guo, Z. & Zhang, Y. 2012. Nlrp2, a Maternal Effect Gene Required for Early Embryonic Development in the Mouse. *PLoS One,* 7(1), p e30344.

Petropoulos, S., Edsgard, D., Reinius, B., Deng, Q., Panula, S. P., Codeluppi, S., Plaza Reyes, A., Linnarsson, S., Sandberg, R. & Lanner, F. 2016. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell,* 165(4), pp 1012-1026.

Petrussa, L., Van de Velde, H. & De Rycke, M. 2014. Dynamic regulation of DNA methyltransferases in human oocytes and preimplantation embryos after assisted reproductive technologies. *Mol Hum Reprod,* 20(9), pp 861-874.

Piedrahita, J. A. (2011). The role of imprinted genes in fetal growth abnormalities. *Birth Defects Res A: Clin Mol Teratol*, *91*(8), 682-692.

Poli, M., Ori, A., Child, T., Jaroudi, S., Spath, K., Beck, M. & Wells, D. 2015. Characterization and quantification of proteins secreted by single human embryos prior to implantation. *EMBO Mol Med,* 7(11), pp 1465-1479.

Popkie, A. P., Zeidner, L. C., Albrecht, A. M., D'Ippolito, A., Eckardt, S., Newsom, D. E., Groden, J., Doble, B. W., Aronow, B., McLaughlin, K. J., White, P. & Phiel, C. J. 2010. Phosphatidylinositol 3-kinase (PI3K) signaling via glycogen synthase kinase-3 (Gsk-3) regulates DNA methylation of imprinted loci. *J Biol Chem,* 285(53), pp 41337-41347.

Proell, M., Riedl, S. J., Fritz, J. H., Rojas, A. M. & Schwarzenbacher, R. 2008. The Nod-like receptor (NLR) family: a tale of similarities and differences. *PLoS One,* 3(4)**,** p e2119.

Proudhon, C., Duffié, R., Ajjan, S., Cowley, M., Iranzo, J., Carbajosa, G., Saadeh, H., Holland, M.L., Oakey, R.J., Rakyan, V.K. and Schulz, R., 2012. Protection against de novo methylation is instrumental in maintaining parent-of-origin methylation inherited from the gametes. *Mol cell*, *47*(6), pp 909-920.

Qian, J., Deveault, C., Bagga, R., Xie, X. & Slim, R. 2007. Women heterozygous for NALP7/NLRP7 mutations are at risk for reproductive wastage: report of two novel mutations. *Hum Mutat,* 28(7)**,** p 741.

Qian, J., Nguyen, N. M. P., Rezaei, M., Huang, B., Tao, Y., Zhang, X., Cheng, Q., Yang, H., Asangla, A., Majewski, J. & Slim, R. 2018. Biallelic PADI6 variants linking infertility, miscarriages, and hydatidiform moles. *Eur J Hum Genet,* 26(7)**,** pp 1007-1013.

Qin, D., Gao, Z., Xiao, Y., Zhang, X., Ma, H., Yu, X., Nie, X., Fan, N., Wang, X., Ouyang, Y., Sun, Q. Y., Yi, Z. & Li, L. 2019. The subcortical maternal complex protein Nlrp4f is involved in cytoplasmic lattice formation and organelle distribution. *Development,* 146(20)**,** p.dev183616.

Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P. V., Grimaldi, G., Riccio, A. & Trono, D. 2011. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell,* 44(3)**,** pp 361-372.

Radian, A. D., Khare, S., Chu, L. H., Dorfleutner, A. & Stehlik, C. 2015. ATP binding by NLRP7 is required for inflammasome activation in response to bacterial lipopeptides. *Mol Immunol,* 67(2 Pt B)**,** pp 294-302.

Rahiminia, T., Yazd, E.F., Ghasemi-Esmailabad, S. and Talebi, A.R., 2019. Relation between sperm protamine transcripts with global sperm DNA methylation and sperm DNA methyltransferases mRNA in men with severe sperm abnormalities. *Hum Fertil*. e1574032

Reddy, R., Akoury, E., Phuong Nguyen, N. M., Abdul-Rahman, O. A., Dery, C., Gupta, N., Daley, W. P., Ao, A., Landolsi, H., Ann Fisher, R., Touitou, I. & Slim, R. 2013. Report of four new patients with protein-truncating mutations in C6orf221/KHDC3L and colocalization with NLRP7. *Eur J Hum Genet,* 21(9)**,** pp 957-964.

Reik, W. & Walter, J. 2001. Genomic imprinting: parental influence on the genome *Nat*

Rev Genet, 2(1), pp 21-32.

Rougier, N., Bourc'his, D., Gomes, D. M., Niveleau, A., , Plachot, M., Pàldi, A. & Viegas-Péquignot, E. 1998. Chromosome methylation patterns during mammalian preimplantation development. *Genes Dev,* 12(14)**,** pp 2108-2113.

Rugg-Gunn, P. J., Ferguson-Smith, A. C. & Pedersen, R. A. 2007. Status of genomic imprinting in human embryonic stem cells as revealed by a large cohort of independently derived and maintained lines. *Hum Mol Genet,* 16(R2), pp.R243-R251.

Sahakyan, A. & Plath, K. 2016. Transcriptome Encyclopedia of Early Human Development. *Cell,* 165(4)**,** pp 777-779.

Sanchez-Delgado, M., Martin-Trujillo, A., Tayama, C., Vidal, E., Esteller, M., Iglesias-Platas, I., Deo, N., Barney, O., Maclean, K., Hata, K., Nakabayashi, K., Fisher, R. & Monk, D. 2015. Absence of Maternal Methylation in Biparental Hydatidiform Moles from Women with NLRP7 Maternal-Effect Mutations Reveals Widespread Placenta-Specific Imprinting. *PLoS Genet,* 11(11)**,** p e1005644.

SanMiguel, J. M., Abramowitz, L. K. & Bartolomei, M. S. 2018. Imprinted gene dysregulation in a Tet1 null mouse model is stochastic and variable in the germline and offspring. *Development,* 145(7)**,** pp. 1-11

Sasaki, H. & Matsui, Y. 2008. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat Rev Genet,* 9(2)**,** pp 129-140.

Sato, A., Otsu, E., Negishi, H., Utsunomiya, T. & Arima, T. 2007. Aberrant DNA methylation of imprinted loci in superovulated oocytes. *Hum Reprod,* 22(1)**,** pp 26-35.

Savage, P. M., Sita-Lumsden, A., Dickson, S., Iyer, R., Everard, J., Coleman, R., Fisher, R. A., Short, D., Casalboni, S., Catalano, K. & Seckl, M. J. 2013. The relationship of maternal age to molar pregnancy incidence, risks for chemotherapy and subsequent pregnancy outcome. *J Obstet Gynaecol,* 33(4)**,** pp 406-411.

Schroder, K. & Tschopp, J. 2010. The inflammasomes. *Cell,* 140(6)**,** pp 821-832.

Schultz, R.M., Stein, P. and Svoboda, P., 2018. The oocyte-to-embryo transition in mouse: past, present, and future. *Biol Reprod*, 99(1), pp.160-174.

Schulz, K. N. & Harrison, M. M. 2019. Mechanisms regulating zygotic genome activation. *Nat Rev Genet,* 20(4)**,** pp 221-234.

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C.,

Thienpont, B., Dean, W. & Reik, W. 2012. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol Cell,* 48(6)**,** pp 849-862.

Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M. & Matsui, Y. 2005. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev Biol,* 278(2)**,** pp 440-458.

Shen, L., Inoue, A., He, J., Liu, Y., Lu, F. and Zhang, Y., 2014. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell stem cell*, *15*(4), pp.459-471.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W. and Smith, A., 2008. Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol*, 6(10), p e253.

Singer, H., Biswas, A., Nuesgen, N., Oldenburg, J. & El-Maarri, O. 2015. NLRP7, Involved in Hydatidiform Molar Pregnancy (HYDM1), Interacts with the Transcriptional Repressor ZBTB16. *PLoS One,* 10(6)**,** p e0130416.

Singer, H., Biswas, A., Zimmer, N., Messaed, C., Oldenburg, J., Slim, R. & El-Maarri, O. 2014. NLRP7 inter-domain interactions: the NACHT-associated domain is the physical mediator for oligomeric assembly. *Mol Hum Reprod,* 20(10)**,** pp 990-1001.

Singh, P., Li, A. X., Tran, D. A., Oates, N., Kang, E. R., Wu, X. & Szabo, P. E. 2013. De novo DNA methylation in the male germ line occurs by default but is excluded at sites of H3K4 methylation. *Cell Rep,* 4(1)**,** pp 205-219.

Slim, R. & Wallace, E. P. 2013. NLRP7 and the Genetics of Hydatidiform Moles: Recent Advances and New Challenges. *Front Immunol,* 4, p 242.

Smallwood, S. A. & Kelsey, G. 2012. De novo DNA methylation: a germ cell perspective. *Trends Genet,* 28(1)**,** pp 33-42.

Smith, A. 2006. A glossary for stem-cell biology. *Nature,* 441(7097)**,** p 1060.

Smith, A. 2013. Nanog heterogeneity: tilting at windmills? *Cell Stem Cell,* 13(1)**,** pp 6-7.

Smith, A. G. 2001. Embryo-derived stem cells: of mice and men. *Annu Rev Cell Dev Biol* 17(1)**,** pp 435-462.

Stewart, K. R., Veselovska, L., & Kelsey, G. 2016. Establishment and functions of DNA methylation in the germline. *Epigenomics,* 8(10)**,** pp 1399-1413.

Stewart, K. R., Veselovska, L., Kim, J., Huang, J., Saadeh, H., Tomizawa, S., Smallwood, S. A., Chen, T. & Kelsey, G. 2015. Dynamic changes in histone modifications precede de novo DNA methylation in oocytes. *Genes Dev,* 29(23)**,** pp 2449-2462.

Stuchbury, G. & Munch, G. 2010. Optimizing the generation of stable neuronal cell lines via pre-transfection restriction enzyme digestion of plasmid DNA. *Cytotechnology,* 62(3)**,** pp 189-194.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H. & Tajima, S. 2004. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *J Biol Chem,* 279(26)**,** pp 27816-27823.

Syvänen, A.-C. 2001. Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat Rev Genet,* 2(12)**,** pp 930-942.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., Reik, W., Bertone, P. & Smith, A. 2014. Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell,* 158(6)**,** pp 1254-1269.

Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., Lao, K. & Surani, M. A. 2010. Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell,* 6(5)**,** pp 468-478.

Tashiro, F., Kanai-Azuma, M., Miyazaki, S., Kato, M., Tanaka, T., Toyoda, S., Yamato, E., Kawakami, H., Miyazaki, T. & Miyazaki, J. 2010. Maternal-effect gene Ces5/Ooep/Moep19/Floped is essential for oocyte cytoplasmic lattice formation and embryonic development at the maternal-zygotic stage transition. *Genes Cells,* 15(8)**,** pp 813-828.

Tesar, P. J., Chenoweth, J. G., Brook, F. A., Davies, T. J., Evans, E. P., Mack, D. L., Gardner, R. L. & McKay, R. D. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature,* 448(7150)**,** pp 196-199.

Theunissen, T. W., Powell, B. E., Wang, H., Mitalipova, M., Faddah, D. A., Reddy, J., Fan, Z. P., Maetzel, D., Ganz, K., Shi, L., Lungjangwa, T., Imsoonthornruksa, S., Stelzer, Y., Rangarajan, S., D'Alessio, A., Zhang, J., Gao, Q., Dawlaty, M. M., Young, R. A., Gray, N. S. & Jaenisch, R. 2014. Systematic identification of culture conditions for induction and maintenance of naïve human pluripotency. *Cell Stem Cell,* 15(4)**,** pp 471-487.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S. & Jones, J. M. 1998. Embryonic stem cell lines derived from human

blastocysts. *Science,* 282(5391), pp 1145-1147.

Thomson, J. P., Skene, P. J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A. R., Deaton, A., Andrews, R., James, K. D., Turner, D. J., Illingworth, R. & Bird, A. 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature,* 464(7291), pp 1082-1086.

Tian, X., Pascal, G. & Monget, P. 2009. Evolution and functional divergence of NLRP genes in mammalian reproductive systems. *BMC Evol Biol,* 9(1), p 202.

Tong, Z. B., Gold, L., Pfeifer, K. E., Dorward, H., Lee, E., Bondy, C. A., Dean, J. & Nelson, L. M. 2000. Mater, a maternal effect gene required for early embryonic development in mice. *Nat Genet,* 26(3), pp 267-268.

Tschopp, J., Martinon, F. & Burns, K. 2003. NALPs: a novel protein family involved in inflammation. *Nat Rev Mol Cell Biol,* 4(2), pp 95-104.

Tucker, K.L., Beard, C., Dausmann, J., Jackson-Grusby, L., Laird, P.W., Lei, H., Li, E. and Jaenisch, R., 1996. Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. Genes & development, 10(8), pp.1008-1020.

Turunen, J. A., Wedenoja, J., Repo, P., Jarvinen, R. S., Jantti, J. E., Mortenhumer, S., Riikonen, A. S., Lehesjoki, A. E., Majander, A. & Kivela, T. T. 2018. Keratoendotheliitis Fugax Hereditaria: A Novel Cryopyrin-Associated Periodic Syndrome Caused by a Mutation in the Nucleotide-Binding Domain, Leucine-Rich Repeat Family, Pyrin Domain-Containing 3 (NLRP3) Gene. *Am J Ophthalmol,* 188, pp 41-50.

Tycko, B. & Morison, I. M. 2002. Physiological functions of imprinted genes. *J Cell Physiol,* 192(3), pp 245-258.

Uyar, A., & Seli, E. (2014). The impact of assisted reproductive technologies on genomic imprinting and imprinting disorders. *Curr opin in obstet gynecol*, 26(3), 210.

Valinluck, V. & Sowers, L. C. 2007. Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res,* 67(3), pp 946-950.

Vallot, C., Patrat, C., Collier, A. J., Huret, C., Casanova, M., Liyakat Ali, T. M., Tosolini, M., Frydman, N., Heard, E., Rugg-Gunn, P. J. & Rougeulle, C. 2017. XACT Noncoding RNA Competes with XIST in the Control of X Chromosome Activity during Human Early Development. *Cell Stem Cell,* 20(1), pp 102-111.

Van den Veyver, I. B. & Al-Hussaini, T. K. 2006. Biparental hydatidiform moles: a maternal effect mutation affecting imprinting in the offspring. *Hum Reprod Update,* 12(3)**,** pp 233-242.

Vazin, T. & Freed, W. J. 2010. Human embryonic stem cells: derivation, culture, and differentiation: a review. *Restor Neurol Neurosci,* 28(4)**,** pp 589-603.

Velker, B. A., Denomme, M. M., Krafty, R. T., & Mann, M. R. (2017). Maintenance of Mest imprinted methylation in blastocyst-stage mouse embryos is less stable than other imprinted loci following superovulation or embryo culture. *Environmental Epigenetics*, *3*(3).

Velloso, F. J., Trombetta-Lima, M., Anschau, V., Sogayar, M. C. & Correa, R. G. 2019. NOD-like receptors: major players (and targets) in the interface between innate immunity and cancer. *Biosci Rep,* 39(4)**,** pp. 1-21

Ventura-Junca, P., Irarrazaval, I., Rolle, A. J., Gutierrez, J. I., Moreno, R. D. & Santos, M. J. 2015. In vitro fertilization (IVF) in mammals: epigenetic and developmental alterations. Scientific and bioethical implications for IVF in humans. *Biol Res,* 48(1), pp 1-13.

Verma, N., Pan, H., Dore, L. C., Shukla, A., Li, Q. V., Pelham-Webb, B., Teijeiro, V., Gonzalez, F., Krivtsov, A., Chang, C. J., Papapetrou, E. P., He, C., Elemento, O. & Huangfu, D. 2018. TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat Genet,* 50(1)**,** pp 83-95.

Veselovska, L., Smallwood, S. A., Saadeh, H., Stewart, K. R., Krueger, F., Maupetit-Mehouas, S., Arnaud, P., Tomizawa, S., Andrews, S. & Kelsey, G. 2015. Deep sequencing and de novo assembly of the mouse oocyte transcriptome define the contribution of transcription to the DNA methylation landscape. *Genome Biol,* 16(1), pp1-17.

Vincent, J.J., Huang, Y., Chen, P.Y., Feng, S., Calvopiña, J.H., Nee, K., Lee, S.A., Le, T., Yoon, A.J., Faull, K. and Fan, G., 2013. Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells. *Cell stem cell*, *12*(4), pp.470-478.

von Meyenn, F. & Reik, W. 2015. Forget the Parents: Epigenetic Reprogramming in Human Germ Cells. *Cell,* 161(6)**,** pp 1248-1251.

von Meyenn, F., Iurlaro, M., Habibi, E., Liu, N.Q., Salehzadeh-Yazdi, A., Santos, F., Petrini, E., Milagre, I., Yu, M., Xie, Z. and Kroeze, L.I., 2016. Impairment of DNA methylation maintenance is the main cause of global demethylation in naïve embryonic stem cells. *Molecular cell*, *62*(6), pp.848-861.

Wang, C. M., Dixon, P. H., Decordova, S., Hodges, M. D., Sebire, N. J., Ozalp, S., Fallahian, M., Sensi, A., Ashrafi, F., Repiska, V., Zhao, J., Xiang, Y., Savage, P. M., Seckl, M. J. & Fisher, R. A. 2009. Identification of 13 novel NLRP7 mutations in 20 families with recurrent hydatidiform mole; missense mutations cluster in the leucine-rich region. *J Med Genet,* 46(8)**,** pp 569-575.

Wang, L., Zhang, J., Duan, J., Gao, X., Zhu, W., Lu, X., Yang, L., Zhang, J., Li, G., Ci, W., Li, W., Zhou, Q., Aluru, N., Tang, F., He, C., Huang, X. & Liu, J. 2014. Programming and inheritance of parental DNA methylomes in mammals. *Cell,* 157(4)**,** pp 979-991.

Ware, C. B., Nelson, A. M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E. C., Jimenez-Caliani, A. J., Deng, X., Cavanaugh, C., Cook, S., Tesar, P. J., Okada, J., Margaretha, L., Sperber, H., Choi, M., Blau, C. A., Treuting, P. M., Hawkins, R. D., Cirulli, V. & Ruohola-Baker, H. 2014. Derivation of naïve human embryonic stem cells. *Proc Natl Acad Sci U S A,* 111(12)**,** pp 4484-4489.

Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. 2016. Dynamic stem cell states: naïve to primed pluripotency in rodents and humans. *Nat Rev Mol Cell Biol,* 17(3)**,** pp 155-169.

Wermann, H., Stoop, H., Gillis, A.J., Honecker, F., van Gurp, R.J., Ammerpohl, O., Richter, J., Oosterhuis, J.W., Bokemeyer, C. and Looijenga, L.H., 2010. Global DNA methylation in fetal human germ cells and germ cell tumours: association with differentiation and cisplatin resistance. *The Journal of pathology*, 221(4), pp.433-442.

White, C.R., MacDonald, W.A. and Mann, M.R., 2016. Conservation of DNA methylation programming between mouse and human gametes and preimplantation embryos. *Biology of Reproduction*, 95(3), pp.61-1.

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A. & Gough, N. M. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature,* 336(6200)**,** pp 684-687.

Wu, H. & Zhang, Y. 2014. Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell,* 156(1-2)**,** pp 45-68.

Wu, J.Q., Habegger, L., Noisa, P., Szekely, A., Qiu, C., Hutchison, S., Raha, D., Egholm, M., Lin, H., Weissman, S. and Cui, W., 2010. Dynamic transcriptomes during neural differentiation of human embryonic stem cells revealed by short, long, and paired-end sequencing. *Proc Natl Acad Sci U S A*, 107(11), pp 5254-5259.

Wu, X. & Zhang, Y. 2017. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet,* 18(9)**,** pp 517-534.

Xiao, J., Mai, D. H. & Xie, L. 2016. Resetting human naïve pluripotency. *Genet Epigenet,* 8, pp GEG-S38093.

Xu, C., Inokuma, M.S., Denham, J., Golds, K., Kundu, P., Gold, J.D., & Carpenter, M. K. 2001. Feeder-free growth of undifferentiated human embryonic stem cells. *Nat Biotechnol,* 19(10)**,** pp 971-974.

Yagi, M., Yamanaka, S. & Yamada, Y. 2017. Epigenetic foundations of pluripotent stem cells that recapitulate in vivo pluripotency. *Lab Invest,* 97(10)**,** pp 1133-1141.

Yamaguchi, S., Shen, L., Liu, Y., Sendler, D. & Zhang, Y. 2013. Role of Tet1 in erasure of genomic imprinting. *Nature,* 504(7480)**,** pp 460-464.

Yamashiro, C., Sasaki, K., Yabuta, Y., Kojima, Y., Nakamura, T., Okamoto, I., Yokobayashi, S., Murase, Y., Ishikura, Y., Shirane, K., & Sasaki, H. 2018. Generation of human oogonia from induced pluripotent stem cells in vitro. *Science,* 362(6412)**,** pp 356-360.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., Huang, J., Li, M., Wu, X., Wen, L., Lao, K., Li, R., Qiao, J. & Tang, F. 2013. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol,* 20(9)**,** pp 1131-1139.

Yang, J., Bashkenova, N., Zang, R., Huang, X. and Wang, J., 2020. The roles of TET family proteins in development and stem cells. *Development*, *147*(2).

Ye, Z. & Ting, J. P. 2008. NLR, the nucleotide-binding domain leucine-rich repeat containing gene family. *Curr Opin Immunol,* 20(1)**,** pp 3-9.

Ying, Q. L., Nichols, J., Chambers, I. & Smith, A. 2003. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell,* 115(3)**,** pp 281-292.

Ying, Q. L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P. & Smith, A. 2008. The ground state of embryonic stem cell self-renewal. *Nature,* 453(7194)**,** pp 519-523.

Yu, X. J., Yi, Z., Gao, Z., Qin, D., Zhai, Y., Chen, X., Ou-Yang, Y., Wang, Z. B., Zheng, P., Zhu, M. S., Wang, H., Sun, Q. Y., Dean, J. & Li, L. 2014. The subcortical maternal complex controls symmetric division of mouse zygotes by regulating F-actin dynamics. *Nat Commun,* 5(1), pp 1-12.

Yurttas, P., Vitale, A. M., Fitzhenry, R. J., Cohen-Gould, L., Wu, W., Gossen, J. A. & Coonrod, S. A. 2008. Role for PADI6 and the cytoplasmic lattices in ribosomal storage in oocytes and translational control in the early mouse embryo. *Development,* 135(15)**,** pp 2627-2636.

Zhang, S., Bell, E., Zhi, H., Brown, S., Imran, S.A., Azuara, V. and Cui, W., 2019. OCT4 and PAX6 determine the dual function of SOX2 in human ESCs as a key pluripotent or neural factor. *Stem Cell Res Ther* 10(1), p 122.

Zhang, P., Dixon, M., Zucchelli, M., Hambiliki, F., Levkov, L., Hovatta, O. & Kere, J. 2008. Expression analysis of the NLRP gene family suggests a role in human preimplantation development. *PLoS One,* 3(7)**,** pp e2755.

Zhang, Y., Yan, Z., Qin, Q., Nisenblat, V., Chang, H. M., Yu, Y., Wang, T., Lu, C., Yang, M., Yang, S., Yao, Y., Zhu, X., Xia, X., Dang, Y., Ren, Y., Yuan, P., Li, R., Liu, P., Guo, H., Han, J., He, H., Zhang, K., Wang, Y., Wu, Y., Li, M., Qiao, J., Yan, J. & Yan, L. 2018. Transcriptome Landscape of Human Folliculogenesis Reveals Oocyte and Granulosa Cell Interactions. *Mol Cell,* 72(6)**,** pp 1021-1034 e4.

Zhu, K., Yan, L., Zhang, X., Lu, X., Wang, T., Yan, J., Liu, X., Qiao, J. & Li, L. 2015. Identification of a human subcortical maternal complex. *Mol Hum Reprod,* 21(4)**,** pp 320-329.

Zhu, P., Guo, H., Ren, Y., Hou, Y., Dong, J., Li, R., Lian, Y., Fan, X., Hu, B., Gao, Y., Wang, X., Wei, Y., Liu, P., Yan, J., Ren, X., Yuan, P., Yuan, Y., Yan, Z., Wen, L., Yan, L., Qiao, J. & Tang, F. 2018. Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat Genet,* 50(1)**,** pp 12-19.

Zhu, Z. & Huangfu, D. 2013. Human pluripotent stem cells: an emerging model in developmental biology. *Development,* 140(4)**,** pp 705-717.