THE UNIVERSITY
of EDINBURGH

Genome Defence in Hypomethylated Developmental Contexts

Christopher J. Playfoot

Thesis presented for the degree of Doctor of Philosophy

The University of Edinburgh

September 2016

Declaration

I declare that the work presented in this PhD thesis is my own except where otherwise stated. This work has not been submitted for any other degree.

Christopher J. Playfoot

8th September, 2016

Acknowledgements

There are many people I wish to thank who have made this PhD both possible and enjoyable.

Firstly, Ian Adams. I could not have asked for a better mentor. His staggering amount of knowledge about, not only our labs interests, but also that of the wider field is quite awe inspiring. His door has, quite literally, always been open for any questions I might have, however small. He provides a passionate and stimulating environment to work in and the amount I have learnt from him during my studies is testament to the support he provides, whilst respecting and nurturing my freedom as a scientist to pursue interesting avenues in the lab. Thanks Ian.

I also wish to thank the members of the Adams lab, past and present, for all their help, support and humour over the years. James Crichton, Marie MacLennan, Diana Best, Abby Mann, Karen Dobie, Eleanor Raymond, Issy MacGregor, Sona Relovska and my favourite Brazilian, Renato Tesser. You all make the days fly by. Our friend David Read, tragically gone too soon, made all of our days that little bit brighter with his humour, sincere generosity and questionable internet searches, along with making all of our work possible. There are no words...

I also wish to thank Mark Ditzel, Bob Hill, Andrew Jackson and Didier Trono for mice and reagents which have made some of the experiments in this thesis possible. I owe a special thank you to Sam Corless and Nezha Benabdallah for all our adventures over the years, from the many pre / post work surfs, to the countless forays to the highlands.

Finally, I wish to thank my family. My parents, Alison and Martyn Playfoot for always being there and making me the person I am today. Your never-ending support has made this PhD possible. I can't describe how appreciative I am of what you have done for me in the past, present and no doubt, future. My siblings, Emma and Michael, with whom I can be a child again and lastly, but most importantly, my Grandma, Doreen Linfoot. You are an inspiration.

"When we try to pick out anything by itself, we find it hitched to everything else in the Universe"

- John Muir

Abstract

Retrotransposons constitute around 40% of the mammalian genome and their aberrant activation can have wide ranging detrimental consequences, both throughout development and into somatic lineages. DNA methylation is one of the major epigenetic mechanisms in mammals, and is essential in repressing retrotransposons throughout mammalian development. Yet during normal mouse embryonic development some cell lineages become extensively DNA hypomethylated and it is not clear how these cells maintain retrotransposon silencing in a globally hypomethylated genomic context.

In this thesis I determine that hypomethylation in multiple contexts results in the consistent activation of only one gene in the mouse genome - *Tex19.1*. Thus if a generic compensatory mechanism for loss of DNA methylation exists in mice, it must function through this gene. *Tex19.1*$^{-/-}$ mice de-repress retrotransposons in the hypomethylated component of the placenta and in the mouse germline, and have developmental defects in these tissues. In this thesis I examine the mechanism of TEX19.1 mediated genome defence and the developmental consequences upon its removal. I show that TEX19.1 functions in repressing retrotransposons, at least in part, through physically interacting with the transcriptional co-repressor, KAP1. *Tex19.1*$^{-/-}$ ES cells have reduced levels of KAP1 bound retrotransposon chromatin and reduced levels of the repressive H3K9me3 modification at these loci. Furthermore, these subsets of retrotransposon loci are de-repressed in *Tex19.1*$^{-/-}$ placentas. Thus, my data indicates that mouse cells respond to hypomethylation by activating expression of Tex19.1, which in turn augments compensatory, repressive histone modifications at retrotransposon sequences, thereby helping developmentally hypomethylated cells to maintain genome stability.

I next aimed to further elucidate the role of Tex19.1 in the developing hypomethylated placenta. I determine that *Tex19.1*$^{-/-}$ placental defects precede intrauterine growth restriction of the embryo and that alterations in mRNA abundance in E12.5 *Tex19.1*$^{-/-}$ placentas is likely in part due to genic transcriptional changes. De-repression of LINE-1 is evident in these placentas and elements of the de-repressed subfamily are

associated with significantly downregulated genes. If retrotransposon de-repression is contributing to developmental defects by interfering with gene expression remains to be determined, however I identify a further possible mechanism leading to placental developmental defects. I determine that *Tex19.1*$^{-/-}$ placentas have an increased innate immune response and I propose that this is contributing to the developmental defects observed.

Developmental defects and retrotransposon de-repression are also observed in spermatogenesis in *Tex19.1*$^{-/-}$ testes, the molecular basis for which is unclear. I therefore investigate the possibility that the TEX19.1 interacting partners, the E3 ubiquitin ligase proteins, may be contributing to the phenotypes observed in *Tex19.1*$^{-/-}$ testes. I show that repression of MMERVK10C in the testes is dependent on UBR2, alongside TEX19.1. Furthermore, I have identified a novel role for the TEX19.1 interacting partner, UBR5, in spermatogenesis, whose roles are distinct from those of TEX19.1.

The work carried out during the course of this thesis provides mechanistic insights into TEX19.1 mediated genome defence and highlights the importance of protecting the genome from aberrant retrotransposon expression.

Lay Abstract

Every cell in humans and mice contains DNA; precise instructions necessary for life. Staggeringly, almost half of human and mouse DNA is made up of old viruses, which can copy and paste themselves to new places in the DNA. In normal circumstances these cause no problems to how the cell works. This is because defence mechanisms have evolved alongside these to stop them becoming active. In the development of the mouse one of these major defence mechanisms called DNA methylation is lost at defined time points, making the developing cells vulnerable to activation of these old viruses. The consequences of this can be wide ranging, causing problems in development, through to adulthood.

Extra levels of defence, in addition to DNA methylation, ensure continued safety from the activation of these old viruses. One gene called *Tex19.1* is activated when DNA methylation is lost and limits the activation of some of these retrotransposons. How *Tex19.1* does this is unclear and is the primary focus of this thesis. Proteins often work in combination with other proteins to do a particular job in the cell. Because of this, I investigated a different protein called KAP1 that binds to TEX19.1 and showed that these work together to stop the activation of some of the old viruses. I also indicate that this may contribute to problems in the development of the placenta.

Along with problems in the placenta, development of sperm does not happen as normal when TEX19.1 is not present. It is not known how these problems occur. I therefore looked at another set of proteins binding to TEX19.1. I show that UBR5, is important for development of sperm but is unlikely to be contributing to the problems in animals without TEX19.1. I show that one of these other sets of proteins, UBR2, is important for stopping the activation of the same retrotransposon as TEX19.1 in testes.

Together, this thesis identifies how *Tex19.1* works to stop the activation of retrotransposons when the major layer of defence, DNA methylation, is reduced and highlights the problems arising when these defence mechanisms fail.

Table of Contents

List of Figures

List of Tables

Abbreviations

5-Aza	5-aza deoxycytodine
5mC	5-methyl cytosine
5hmC	5-hydroxymethyl cytosine
5fC	5-formylcytosine
5caC	5-carboxylcytosine
°C	Degrees centigrade
A	Adenine
Ac	Acetate
AGS	Aicardi Goutières Syndrome
BCIP	5-Bromo-4-Chloro-3'-Indolyphosphate p-Toluidine Salt
BER	Base excision repair
BLAST	Basic Local Alignment Search Tool
bp	Base-pair
BSA	Bovine serum albumin
C	Cytosine
cDNA	Complementary deoxyribonucleic acid
C. elegans	*Caenorhabtidis elegans*
CGI	CpG island
ChIP	Chromatin immunoprecipitation
CpG	Cytosine and guanine separated by a phosphate
C-terminal	Carboxy-terminal
dH_20	Distilled water
DAPI	4,6-diaminophenylindole
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulphoxide
DMR	Differentially methylated region
DNA	Deoxyribonucleic Acid
DNase	Deoxyribonuclease
Dnmt	Dna methyltransferase
dNTP	Deoxynucleotide triphosphate
dpp	Days post partum
Drosophila	*Drosophila melanogaster*
DTT	Dithiothreitol
E	Embryonic day
EB	Embryoid body
E. coli	*Escherichia coli*
EDTA	Ethylenediamineotetra-acetic acid
ERV	Endogenous retrovirus
ES	Embryonic stem
EtBr	Ethidium bromide
EtOH	Ethanol

ExE	Extra embryonic ectoderm
FCS	Foetal calf serum
g	Relative centrifugal force
G	Guanine
GFP	Green fluorescent protein
GMEM	Glasgow's modified eagle medium
H3K4me	Histone H3 lysine 4 monomethylation
H3K4me3	Histone H3 lysine 4 trimethylation
H3K9me3	Histone H3 lysine 9 trimethylation
H3K27ac	Histone H3 lysine 27 acetylation
H3K27me3	Histone H3 lysine 27 trimethylation
HAT	Histone acetyltransferase
Het	Heterozygous
HCl	Hydrochloric acid
HDAC	Histone deacetylase
HDM	Histone demethylase
HMT	Histone methyltransferase
HP1	Heterochromatin associated protein 1
IAP	Intracisternal A particle
ICC	Immunocytochemistry
ICM	Inner cell mass
IHC	Immunohistochemistry
IF	Immunofluorescence
IFN	Interferon
ISH	In situ hybridisation
ISRE	Interferon stimulated response element
IUGR	Intrauterine growth restriction
Kb	Kilobase pairs of DNA
kDa	KiloDaltons
KO	Knockout
LB	Luria-Bertani
LINE	Long interspersed nuclear element
LIF	Leukaemia inhibitory factor
LTR	Long terminal repeat
M	Molar
MBD	Methyl-CpG binding domain
MBP	Methyl-CpG binding protein
MEF	Mouse embryonic fibroblast
mES	Mouse embryonic stem (cell)
$MgCl_2$	Magnesium chloride
MI	Meiosis I

MII	Meiosis II
µg	Microgram
µl	Microlitre
min	Minute
miRNA	Micro RNA
mRNA	Messenger RNA
Mw	Molecular weight
NBT	Nitro-Blue Tetrazolium Chloride
NP-40	Nonidet NP-40
nt	Nucleotide
N-terminal	N amino-terminal
NTP	Nucleotide triphosphate
Oligo	Oligonucleotide
ORF	Open reading frame
P	Postnatal day
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline (plus 0.2% w/v Tween-20)
PCR	Polymerase chain reaction
pen	Penicillin
PFA	Paraformaldehyde
PGC	Primordial germ cell
piRNA	Piwi-interacting RNA
PRC	Polycomb repressive complex
qRT-PCR	Quantitative reverse-transcriptase PCR
RBCC	RING finger B boxes coiled coil motif
RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase polymerase chain reaction
rpm	Revolutions per minute
RT	Room temperature
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulphate
SINE	Short interspersed nuclear element
SSC	Spermatogonial stem cell
strep	Streptomycin
TAE	Tris-EDTA acetic acid buffer
TBE	Tris-EDTA boric acid buffer
TBS(t)	Tris buffered saline (plus 0.05% w/v Tween-20)
TDG	Thymine-DNA-glycosylase
TdT	Terminal deoxynucleotidyl transferase
TE	Trophectoderm

Tex19	*Tex19*.1
TKO	Triple knockout
TSS	Transcription start site
Ub	Ubiquitinated
UTR	Untranslated region
Vol	Volumes
WT	Wild-type
X-Gal	5-Bromo-4-Chloro-3-indolyl B-D-galactopyranoside

Chapter 1: Introduction

1.1 The *Tex19.1* Genome Defence Gene

Tex19.1 is a mammalian-specific gene of unknown biochemical function which is expressed in germ cells, pluripotent stem cells and the placenta (Hackett et al., 2012; Ollinger et al., 2008; Reichmann et al., 2013). *Tex19.1* was originally identified as a 'testis expressed' gene, following a screen to identify genes expressed in spermatogonia but not in somatic cells and is a putative target of DAZL, a germline specific RNA-binding protein (Reynolds et al., 2005; Wang et al., 2001). It is well conserved between mammals, with humans, macaque and chimpanzee having a single copy of *Tex19,* with a premature stop codon which truncates the protein to 164 amino acids in length when expressed (Kuntz et al., 2008)(figure 1.1). In mice, *Tex19* has undergone a duplication event to give rise to two *Tex* genes, *Tex19.1* and *Tex19.2* expressed in different tissues (Kuntz et al., 2008)(figure 1.1). *Tex19.1* expression is restricted to pluripotent or hypomethylated tissues and is primarily and causally regulated by promoter DNA methylation, as determined by a study to detect gene expression changes in somatic cells when treated with the DNA methyltransferase inhibitor, the cytidine analog 5aza-deoxycytidine (5-Aza) followed by recovery (Hackett et al., 2012). Genes retaining increased expression after recovery from 5-Aza treatment and in mouse embryonic fibroblasts (MEFs) lacking the maintenance methyltransferase, *Dnmt1* were largely germline genes with roles in spermatogenesis and defence against retrotransposons in the germline (see section 1.2)(Hackett et al., 2012). Previously in the Adams lab the association of *Tex19.1* expression with pluripotency, placental and germline development has been explored by the generation of *Tex19.1*$^{-/-}$ mice (Ollinger et al., 2008; Reichmann et al., 2012). This identified numerous roles for *Tex19.1* in spermatogenesis, placental development and retrotransposon repression (discussed further in section 1.12)(Ollinger et al., 2008). *Tex19.1*$^{-/-}$ males are sterile and exhibit severe testicular atrophy, meiotic chromosome asynapsis and meiotic pachytene arrest (Ollinger et al., 2008). Additionally, *Tex19.1*$^{-/-}$ pups exhibit intrauterine growth restriction (IUGR) with defective placental development (Reichmann et al., 2012). Alongside these developmental defects, both *Tex19.1* testis and placenta have increased retrotransposon de-repression (Ollinger et al., 2008; Reichmann et al., 2012). As *Tex19.1* is expressed at hypomethylated time

points permissive for the activation of retrotransposons, TEX19.1 is likely activated in response to reduced DNA methylation to protect the genome from excessive retrotransposon activation, which may have potentially deleterious consequences (Crichton et al., 2014; Hackett et al., 2012; Seisenberger et al., 2012). The potential TEX19.1 mediated transcriptional mechanism of retrotransposon repression remains unclear and despite being well conserved among mammals, *Tex19.1* does not share sequence similarities with any known protein domains, therefore imparting potential mechanistic functions is difficult (Kuntz et al., 2008). This thesis therefore investigates TEX19.1 interacting partners such as KAP1 and UBR5 to unravel the potential mechanisms contributing to *Tex19.1*$^{-/-}$ phenotypes, from retrotransposon de-repression to defective spermatogenesis and placental development.

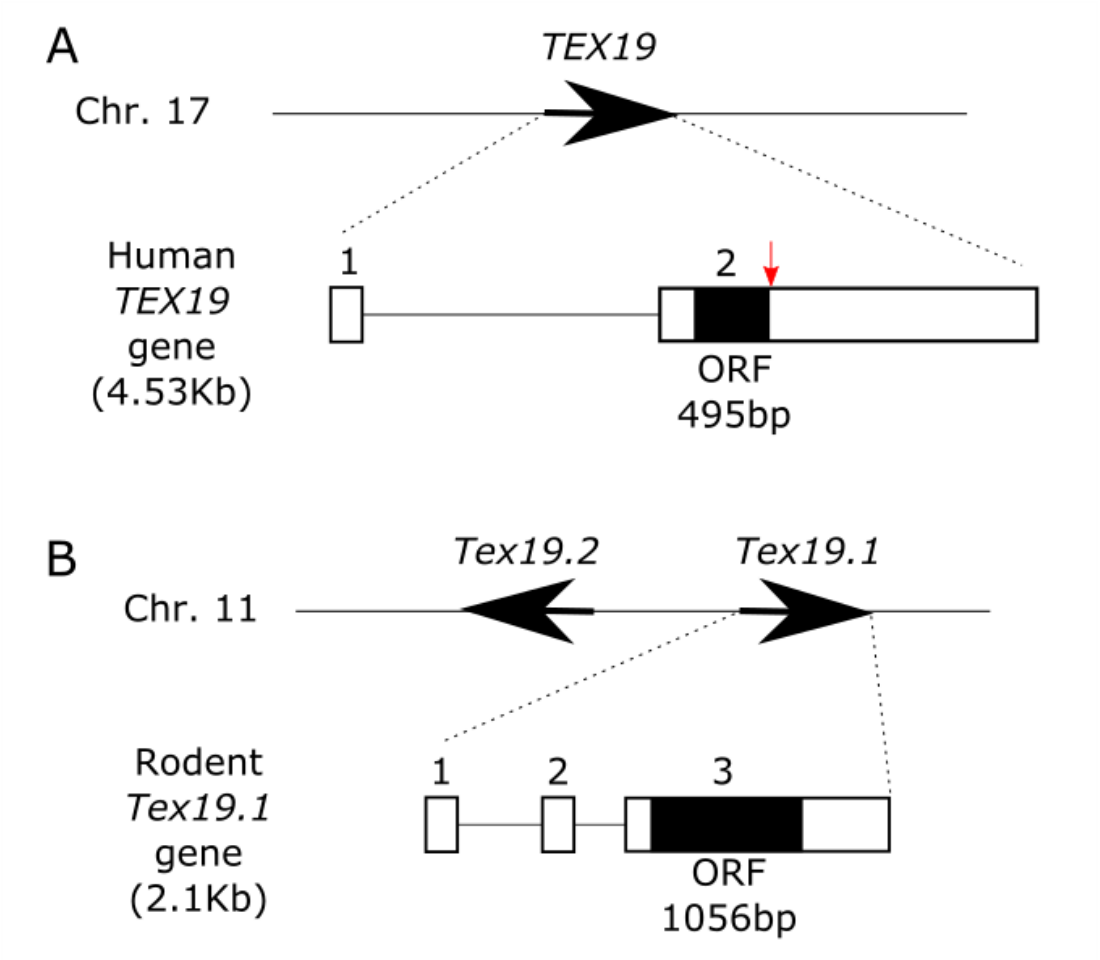


Figure 1.1 Schematic diagram of the human and rodent *Tex19* gene locus. (A) Human *TEX19* resides on chromosome 17 and has not undergone a gene duplication event unlike in rodents. Human *TEX19* is comprised of two exons (boxes) with a 459bp ORF in the second exon (bold box). The premature stop codon, when compared to rodent *Tex19.1,* is indicated by the red arrow. (B) Rodent *Tex19.1* and the duplicated gene, *Tex19.2* are transcribed in opposite directions (bold arrows). Rodent *Tex19.1* is comprised of three exons (boxes) with the 1056bp ORF encoded in the third exon.

1.2 Retrotransposons

First described by Barbara McClintock over 60 years ago (McClintock, 1950, 1956), transposable elements have shaped the mammalian genome throughout evolution (Cordaux and Batzer, 2009). Transposons are mobile genetic elements that comprise around 40% of the mammalian genome, a remarkable figure when considering that protein coding regions account for around 1.5% (Lander et al., 2001; Waterston et al., 2002). Transposable elements can be separated into two major classes; DNA transposons which mobilise via a 'cut and paste' mechanism and retrotransposons which mobilise using a 'copy and paste' mechanism, to allow integration at new genomic locations (Goodier 2016). Propagation of retrotransposons across generations relies on successful retrotransposition in the germline or pluripotent cells early in development (Zamudio and Bourc'his, 2010). The potentially high mutagenic consequences of novel insertions at these vulnerable time points has resulted in the evolution of a diverse array of host genome defence mechanisms that limit their retrotransposition capacity (Goodier 2016). Retrotransposons can be split into three main classes comprised of the long interspersed nuclear element (LINE), short interspersed nuclear element (SINE) and endogenous retroviruses (ERVs) also known as long terminal repeat (LTR) retrotransposons (referring to the 5' and 3' repetitive sequences flanking the internal sequence of the ERV). Each class are repressed by mechanisms fundamental to mammalian development (Goodier 2016).

1.1.1 LINE-1

LINE-1 is arguably the most successful retrotransposon in mammals as it accounts for around 17% of the sequenced genome (Waterston et al., 2002). Full length LINE-1 elements in both mouse and human are approximately 6 to 7kb long however their internal promoter differs both between species and between different subfamilies of LINE-1 within species (Khan et al., 2006). Many LINE-1 elements are also 5' truncated (Sookdeo et al., 2013). Of the half a million LINE-1 elements in the human genome only around 100 of the human specific L1H subfamily are thought to be

retrotranspositionally competent (Beck et al., 2010; Brouha et al., 2003; Solyom et al., 2012). Evidence of de novo spontaneous LINE-1 integrations have been reported to cause disease and phenotypic changes in both humans and mice with approximately 1:1000 spontaneous cases of human disease attributable to LINE-1 insertions (Levin and Moran, 2011). Mouse LINE-1 mRNA encodes two open reading frames (ORF)(figure 1.2A), whereas humans have an additional ORF (Denli et al., 2015; Moran et al., 1996). ORF1p encodes an RNA binding protein with nucleic acid chaperone activity and ORF2 encodes a protein with endonuclease and reverse transcriptase activities, both of which are required for autonomous retrotransposition (figure1.2A)(Moran et al., 1996). Recent evidence also points to a previously overlooked but highly important promoter of human LINE-1 elements. When analysing the human 5'UTR of LINE-1 for ORFs with an upstream promoter and a strong kozak sequence, Denli et al identified a primate specific antisense open reading frame named ORF0, which enhances LINE-1 mobility (Denli et al., 2015). Transcription of LINE-1 is mediated by RNA polymerase II from an internal L1 promoter in the 5'UTR which also produces antisense transcripts that can affect expression of neighbouring genes (figure 1.2A)(Cruickshanks et al., 2013; Mätlik et al., 2006; Speek, 2001). The youngest, most intact LINE-1 subfamilies appeared in the mouse genome less than 2 million years ago and are called L1MdTf, L1MdGf and L1MdA. These are defined by the monomeric repeats in their 5'UTRs with the many polymorphisms reflecting ongoing retrotransposition and evolution (Sookdeo et al., 2013). Older subfamilies such as L1MdF and L1MdF2 are estimated to be between 7.3 million and 3.8 million years old, with older, more degenerate LINE-1 subfamilies such as L1mus1 to L1Mus4 inserted into the mouse genome between 12 and 7 million years ago after the divergence between mouse and rat (Sookdeo et al., 2013).

1.1.2 SINEs

In mice short interspersed nuclear elements (SINEs) make up around 10% of the mammalian genome and rely on the endonuclease and reverse transcriptase activity of LINE-1 ORF2 to facilitate their retrotransposition (Dewannieux et al., 2005;

Waterston et al., 2002). The two main SINE elements are B1 and B2, derived from a 7SL RNA gene (figure 1.2B) and a tRNA gene respectively (Kramerov and Vassetzky, 2011). B1 is closely related to human Alu sequences which also rely on LINE-1 to catalyse their retrotransposition and are the most prevalent SINE subfamilies in human and mouse genomes (Dewannieux et al., 2003). Conversely, SINE B2 has a greater LINE-1 dependent retrotransposition capacity than B1, despite fewer genomic copies which is likely indicative of historic bursts of B1 transposition throughout evolution (Dewannieux et al., 2005). As with LINE-1, d*e novo* integrations of SINE and Alu elements in humans and mouse have been associated with disease but also have intriguing roles in regulation of gene expression (Hancks and Kazazian, 2016; Lunyak et al., 2007).

1.1.3 LTRs

LTRs comprise around 9% of the human and mouse genome and can be split into 4 families in the mouse based on their phylogenetic relationships (Reichmann et al., 2012; Waterston et al., 2002). The ERV1, ERVK, ERVL and MaLR families comprise around 400 diverse elements with differing evolutionary origins and genic composition (Reichmann et al., 2012). Generally, LTRs contain genes essential to facilitate their autonomous retrotransposition, which are usually found in exogenous retroviruses (Dewannieux et al., 2004). The *gag* gene encodes structural proteins to make virus like particles in the cytoplasm of the host cell, where reverse transcription occurs (figure 1.2C). The enzymatic proteins are provided by the *pol* gene encoding a protease, a reverse transcriptase and integrase, necessary for cDNA synthesis and insertion into the host genome (figure 1.2C)(Havecker et al., 2004). The major difference between retrotransposons and infectious retroviruses is the absence of an *env* gene which allows for viral infection of different cells, however some elements retain *env* like ORFs which have lost the capacity to form viral envelopes (figure 1.2C)(Havecker et al., 2004). This gene cluster is usually flanked by terminal repeat regions from which LTR elements get their name, with RNA transcribed from a promoter in the 5' LTR flanking the internal sequence (figure 1.2C). The LTR

elements in the mouse genome are retrotranspositionally active, whereas human LTR elements are thought to have lost this capacity with no *de novo* insertions detected in human disease (Bannert and Kurth, 2006; Maksakova et al., 2006). Around 600,000 solo LTRs exist throughout the human genome with recent evidence suggesting many have been co-opted as regulatory regions controlling neighbouring gene expression as either enhancers or alternative promoters (Thompson et al., 2016).

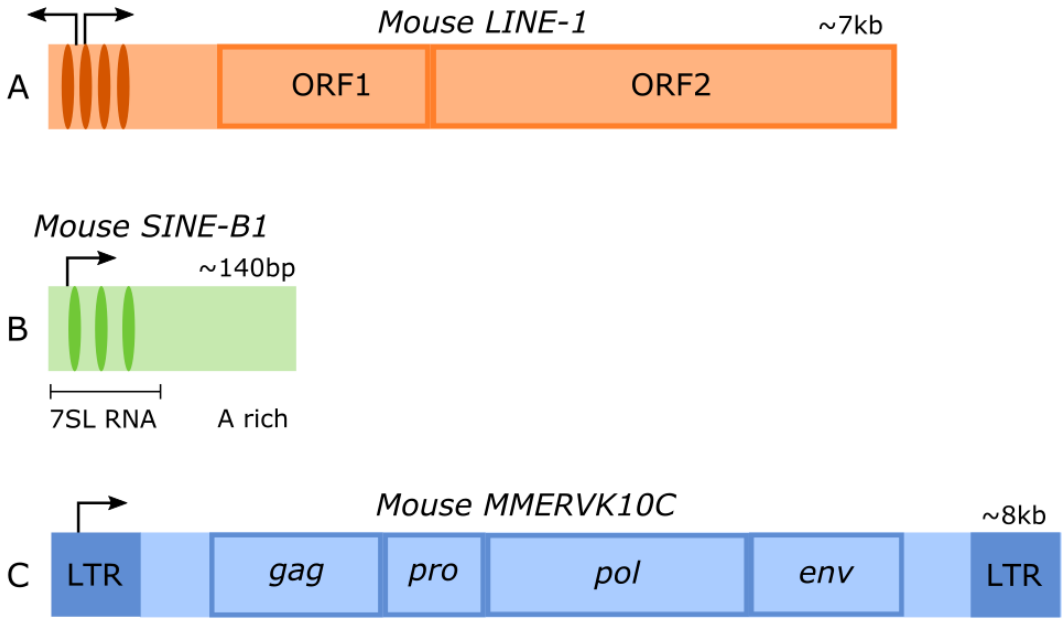


Figure 1.2: Schematic diagram of the structure of mouse LINE-1, SINE B1 and LTR MMERVK10C elements. Shaded areas represent transcriptional regulatory regions with transcriptional start sites indicated by arrows. The molecular function of each protein coding region are discussed in the main text. MMERVK10C represents the LTR class of retrotransposons.

1.3 DNA Methylation

In many species, the transcription of retrotransposons and selected genes is stably repressed in a tissue specific manner both throughout development and into terminally differentiated cell types (Bird, 2002). This control of stable transcriptional repression is mediated by DNA methylation, an epigenetic modification of the DNA whereby, a methyl group is physically added to the 5th position of the pyrimidine ring of a cytosine base when present in a CpG dinucleotide. This modified nucleotide is known as 5-methylcytosine (5mC) (Bird, 2002). Transcriptional regulation as a consequence of methylated CpGs can be explained by two primary models. In the first; methylation of cytosines can prevent transcription factors and regulatory proteins from binding to DNA (Bell and Felsenfeld, 2000), and in the second; DNA methylation can create sites for methyl-CpG binding proteins (MBPs), which recognise methylated CpGs and recruit co-repressors to modify chromatin structure. This in turn leads to the formation of inaccessible, compacted regions of DNA known as heterochromatin and transcriptional repression (Boyes and Bird, 1991).

In eukaryotes, DNA methylation is present in fungal, animal and plant species. CpG methylation plays an essential role in vertebrate development (Li et al., 1992), however some invertebrates such as *Caenorhabditis elegans* and *Drosophila melanogaster* and the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* have very low or largely undetectable levels of genomic DNA methylation (Zemach and Zilberman, 2010). This suggests that DNA methylation is not a prerequisite for development. However, important roles are being identified for DNA methylation in these organisms where the levels of DNA methylation are very low. For example, methylation of cytosines regulates retrotransposon silencing and integrity of telomeres in *Drosophila* (Phalke et al., 2009).

Most of the mammalian genome is heavily methylated, however CpG methylation is highly dynamic and undergoes extensive changes during differentiation, especially in regulatory regions outwith genic promoters (Meissner et al., 2008). Global methylation is interrupted by stretches of DNA with high densities of unmethylated CpGs, called CpG islands (CGIs). These CGIs overlap the promoter regions of around 60%-70% of

all human genes (Weber et al., 2007). CpG density affects the behaviour of these promoters and CGIs are generally associated with constitutively expressed genes (Zhu et al., 2008), however many CGI-associated genes remain transcriptionally repressed despite the absence of methylation, indicating additional regulatory mechanisms such as histone modifications (Weber et al., 2007). In mammals, DNA methylation is critical for embryonic development, genomic imprinting, X-inactivation, gene repression and repression of transposons (Bird, 2002). Furthermore, the existence of DNA methylation has been suggested to have primarily evolved to facilitate this transcriptional control of retrotransposons (Yoder et al., 1997). For DNA methylation to be a truly epigenetic system it needs to be mitotically and meiotically heritable, criteria that are achieved via distinct *de novo* and maintenance DNA methylation mechanisms throughout development and beyond (summarised in table 1.1).

Protein	Functions	Mutant Phenotype	Reference
Dnmt3a	*De novo* establishment of DNA methylation	Postnatal lethality, defective spermatogenesis and *de novo* DNA methylation at imprinted regions	Okano et al., 1999; Kaneda et al., 2004; Kato et al., 2007
Dnmt3b	*De novo* establishment of DNA methylation	Embryonic lethality by E10.5 with defective *de novo* methylation of satellite DNA	Okano et al., 1999
Dnmt3L	Co-factor required for maternal methylation by Dnmt3a and 3b	Abnormal imprinting and male sterility	Bourc'his et al., 2001
Dnmt1	Maintenance of DNA methylation	Embryonic lethality by E8.5. Hypomethylation, loss of imprinting, retrotransposon activation	Li et al., 1992 Lei et al., 1996 Walsh et al., 1998
Dnmt1o	Oocyte specific isoform of Dnmt1	Loss of maternal imprints	Howell et al., 2001

Table 1.1. DNA methyltransferases and their functions and mutant phenotypes in mouse

1.2.1 *De novo* DNA methylation

DNA methylation is added to CpG dinucleotides via two distinct but related DNA methyltransferases, DNMT3A and DNMT3B, whose activity is stimulated by the non-catalytic co-factor, DNMT3L (table 1.1). Expression of these enzymes is greatest during embryonic development, where their primary role is genome wide *de novo* methylation (Okano et al., 1999). The DNMT3 proteins both have essential but differing roles in development as determined by genetic knockouts of *Dnmt3a* or *Dnmt3b*. *Dnmt3a*$^{-/-}$ mice develop to term but are runted and die by 4 weeks post-partum, however global DNA methylation patterns are retained (table 1.1)(Okano et al., 1999). In contrast, no viable *Dnmt3b*$^{-/-}$ mice are born, due to embryonic lethality

at around E10.5 (table 1.1)(Okano et al., 1999). The more severe phenotype of *Dnmt3b*$^{-/-}$ mice suggests a more important role for DNMT3B in embryonic development compared to DNMT3A. Consistent with this, *Dnmt3b* is expressed earlier during embryonic development with DNMT3B protein detected at E4.5 in the inner cell mass at the implantation stage but being reduced in embryos by E10.5 and absent from the trophoblast at each stage (Watanabe et al., 2002). In contrast, DNMT3A is detectable in the embryo at E10.5 onwards with low levels of DNMT3A detectable in the trophoblast component (Oda et al., 2013; Watanabe et al., 2002). Despite these different requirements, both have overlapping functions as indicated by the more severe phenotype of double knockout *Dnmt3a*$^{-/-}$*Dnmt3b*$^{-/-}$ embryos. These embryos die around E8.5 due to growth arrest shortly after gastrulation (Okano et al., 1999).

It is also important to note that there are also some distinct differences in the loci targeted. For example, DNA methylation at centromeric minor satellite repeats is perturbed in *Dnmt3b*$^{-/-}$ mice but not in *Dnmt3a*$^{-/-}$ mice (table 1.1)(Okano et al., 1999). Interestingly, some genomic loci such as imprinted regions remain methylated even in the combined absence of both *Dnmt3a* and *Dnmt3b,* due to the activity of the maintenance methyltransferase, DNMT1 (Hirasawa et al., 2008; Li et al., 1993; Okano et al., 1999). Also DNA methylation at the Intracisternal A particle (IAP) family of LTRs is relatively resistant to this global hypomethylation (Okano et al., 1999). This supports a role for the maintenance of DNA methylation at these loci by the maintenance methyltransferase mechanisms (Okano et al., 1999; Walsh et al., 1998).

The differing targets of each of these proteins in embryonic development suggested that they may also have differing genomic targets in the germline. Conditional deletion of *Dnmt3a* or *Dnmt3b* in primordial germ cells (PGCs) using TNAP-Cre knock-in mice, which is expressed in PGCs from E9.5 to late gestation, showed that DNMT3A has distinct roles from DNMT3B in establishing *de novo* methylation at imprinted genes during male and female gametogenesis (Kaneda et al., 2004; Kato et al., 2007). Conditional *Dnmt3a*$^{-/-}$ mice are viable but offspring of conditional *Dnmt3a*$^{-/-}$ female mice exhibit embryonic lethality around E9.5 and lack methylation at maternally imprinted loci leading to loss of their allele specific expression (table 1.1)(Kaneda et al., 2004; Kato et al., 2007). DNMT3A is also necessary for some paternal imprints

(Kaneda et al., 2004; Kato et al., 2007). Despite the embryonic lethality of complete *Dnmt3b*$^{-/-}$ knockout mice, offspring from germline conditional *Dnmt3b*$^{-/-}$ exhibit no apparent phenotype, however the methylation of at least one paternally imprinted gene, *Rasgrf1,* is dependent on DNMT3B, along with DNMT3A (Kaneda et al., 2004; Kato et al., 2007). Interestingly, the defects observed in offspring of conditional *Dnmt3a*$^{-/-}$ females were very similar to the maternal-effect seen in offspring from *Dnmt3L*$^{-/-}$ females (Bourc'his et al., 2001). Both *Dnmt3a* and *Dnmt3L* are also individually required for maternal and paternal imprinting (table 1.1)(Bourc'his et al., 2001). *Dnmt3a* and *Dnmt3b* are also critical in the female germ line. Conditional double knockout oocytes null for both *Dnmt3a* and *Dnmt3b* give rise to offspring with defects in DNA methylation at imprinted regions and non-imprinted genes, resulting in defective gene expression driving loss of cell adhesion in methylation-deficient trophoblast cells (Branco et al., 2016). Thus *de novo* methylation is restricted to regulation of genomic imprinting but also plays a role in regulating gene expression during development.

The activity of DNMT3A and DNMT3B is stimulated by the catalytically inactive DNMT3L which is essential for *de novo* methylation and maternal imprints during spermatogenesis and oogenesis, but dispensable in somatic tissues (table 1.1)(Jaenisch and Bird, 2003). This requirement of DNMT3L for establishing *de novo* DNA methylation on paternally imprinted genes in the developing germline was indicated by a severe reduction in methylation at paternally imprinted loci in *Dnmt3l* conditional knockout prospermatogonia (Kato et al., 2007). Evidently, *de novo* DNA methylation is highly important for embryonic and germ cell development with many distinct but often overlapping genomic targets of DNMT3A and DNMT3B. The overlapping but often different requirement for each methyltransferase at particular loci in different developmental contexts demonstrates the importance of partially redundant *de novo* DNA methyltransferases. In summary, DNMT3A is globally dispensable for *de novo* methylation in the early embryo but is required for gametic methylation. In contrast, DNMT3B is generally not required for gametic methylation, however is required for *de novo* methylation in the early embryo. Both *Dnmt3a* and *Dnmt3b* play minor roles in *de novo* methylation of the trophoblast lineage, possibly leading to the general hypomethylation but maintenance of imprints this tissue (Branco et al., 2016).

1.2.2 Maintenance of DNA methylation

It is critical to faithfully reproduce DNA methylation patterns between cell generations, not only to maintain imprinting but also to facilitate faithful regulation of transcriptional networks and retrotransposons in daughter cells upon mitotic division. Maintenance of DNA methylation is performed by the DNMT1 maintenance methyltransferase, whereby hemi-methylated loci carrying a methyl group on the parental strand is perpetuated to the new strand (Cedar and Bergman, 2012). Recruitment of DNMT1 to these loci is facilitated by NP95 which binds to heterochromatin associated H3K9me histone modifications during DNA replication to mark hemi-methylated loci (Rothbart et al., 2012). *Dnmt1* is essential for embryonic development as mice hypomorphic for *Dnmt1* die around E8.5 (Li et al., 1992). *Dnmt1* is also essential for DNA methylation as *Dnmt1* hypomorphic mice exhibit a 95% reduction in genomic methylation levels, along with dramatic increases in transcripts of the LTR element IAP which is repressed by DNA methylation in developing and somatic cells (Lei et al., 1996; Li et al., 1992; Walsh et al., 1998). Global and conditional deletion of *Dnmt1* in maternal and zygotic stages of mouse development also results in demethylation of imprinted genes and their aberrant expression in the blastocyst and post-implantation embryos (table 1.1)(Hirasawa et al., 2008; Li et al., 1993). Similar phenotypes arise in embryos carrying catalytically inactive mutant versions of *Dnmt1* suggesting the defects observed are likely due to the failure to maintain DNA methylation (Takebayashi et al., 2007).

1.4 Global Epigenetic Reprogramming

Despite the highly important role for DNA methylation in a plethora of cellular processes, at two distinct time points in development global epigenetic reprogramming occurs where DNA methylation is globally removed from the genome (figure 1.3)(Cantone and Fisher, 2013). This provides a relatively DNA methylation free

genome ready for re-establishment of DNA methylation by *de novo* DNA methyltransferases (Smallwood and Kelsey, 2012).

1.3.1 Epigenetic reprogramming of primordial germ cells

During germ cell development, epigenetic reprogramming occurs leading to profound remodelling of the DNA methylation landscape to remove parental epigenetic marks. Existing DNA methylation patterns from the parental genomes are erased by E13.5 to around less than 10% of starting methylation (figure 1.3)(Popp et al., 2010). This global wave of DNA demethylation seemingly occurs in two phases (Seisenberger et al., 2012). Shortly after gastrulation around E7.25 approximately 50 primordial germ cells are specified from the developing epiblast (Smallwood and Kelsey, 2012). These then undergo extensive proliferation and migration to the genital ridge at around E10.5 where the gonads develop. The first phase of this demethylation is proposed to occur in a DNA replication dependent, passive manner during this proliferation event prior to E9.5, with bulk methylation levels already reduced to 30% at most loci (Seisenberger et al., 2012; Smith and Meissner, 2013). Interestingly, DNA methylation is maintained at imprinted genes at this time point (Seisenberger et al., 2012; Seki et al., 2005). The second wave of DNA demethylation is more likely an active process which contributes to reducing DNA methylation at imprinted regions around E10.5 to E12.5 (Seisenberger et al., 2012). This is presumably necessary to facilitate sex specific establishment of *de novo* methylation marks at imprinted loci in the sexually differentiating germ cells (Smallwood and Kelsey, 2012). A minority of sequences such as IAP retrotransposons are resistant to complete DNA methylation reprogramming as indicated by bisulphite sequencing on PGCs at E11.5 and E13.5, however more recent genome wide DNA methylation techniques has allowed detection of more specific subsets of retrotransposons resistant to these reprogramming events (Hajkova et al., 2002; Lane et al., 2003).

Genome wide bisulphite sequencing covering different time points of PGC DNA demethylation has provided further evidence of differing dynamics and resistance of some genomic loci to these events (Hackett et al., 2013, Popp et al., 2010, Seisenberger

et al., 2012). Retrotransposons such as LINE, SINE and the majority of LTR class are hypomethylated to around 20% in PGCs at E13.5 compared to 80% in the embryo (Popp et al., 2010). During this global reprogramming, one LTR family of elements called intracisternal A particles (IAP) remain heavily controlled by DNA methylation and retain around 60% of methylated CpGs (Popp et al., 2010; Seisenberger et al., 2012). Despite the dramatic loss in DNA methylation at the majority of retrotransposon loci creating a window permissive of their de-repression (figure 1.3), most retrotransposons are not de-repressed at E13.5. Interestingly, LINE-1 is specifically expressed in female PGCs at E16.5 (Seisenberger et al., 2012), however the reason for this specific window of de-repression is unclear. Together this data indicates additional genome defence mechanisms exist to limit any retrotransposition events in reprogrammed PGCs to protect future generations (Seisenberger et al., 2012).

Throughout the whole PGC reprogramming event the *de novo* methyltransferases are transcriptionally silenced (DNMT3A) or excluded from the nucleus (DNMT3B) supporting the absence of any *de novo* methyltransferase activity during PGC epigenetic reprogramming (Hajkova et al., 2002). The regulators of active DNA demethylation in PGCs are steadily being identified with the first protein identified being the deaminase protein, AID but also the Ten-eleven translocation (TET) enzymes (Hackett et al., 2013; Popp et al., 2010). No genetic mouse mutants have yet been identified which retain DNA methylation in PGCs, possibly indicating that there may be more, as yet unknown, mechanisms of DNA removal at work.

This genome wide DNA demethylation event precedes any alterations in histone modifications in post-migratory PGCs, suggesting that erasure of DNA methylation may be an initiating event for reprogramming at this stage. Global levels of many histone modifications associated with transcriptional repression or activation are reduced at E11.5 such as histone H3 lysine 9 trimethylation (H3K9me3), histone H3 lysine 27 trimethylation (H3K27me3), histone H3 lysine 9 acetylation (H3K9ac) and the linker histone H1 (Cantone and Fisher 2013). This reduction is transient as these marks are thought to be globally re-established by E12.5 (Hajkova et al., 2008). The reason for the highly regulated timing of this reduction and re-establishment is still unclear. The combined loss of DNA methylation and histone modifications acts to completely epigenetically reprogram the PGC genome at E11.5 (figure 1.3). In PGCs

specified for the male lineage, DNMT3A and DNMT3B dependent *de novo* methylation begins around E16.5 in prospermatogonia during mitotic arrest (figure 1.3)(Seisenberger et al., 2012). In contrast, developing female oocytes remain in a state of global hypomethylation during meiotic arrest in prophase I until after birth when growing oocytes establish *de novo* DNA methylation patterns until puberty (figure 1.3). Meiosis is resumed upon oocyte extrusion followed by consequent arrest in metaphase II until fertilisation (Smallwood and Kelsey 2012).

1.3.2 Epigenetic reprogramming post-fertilisation

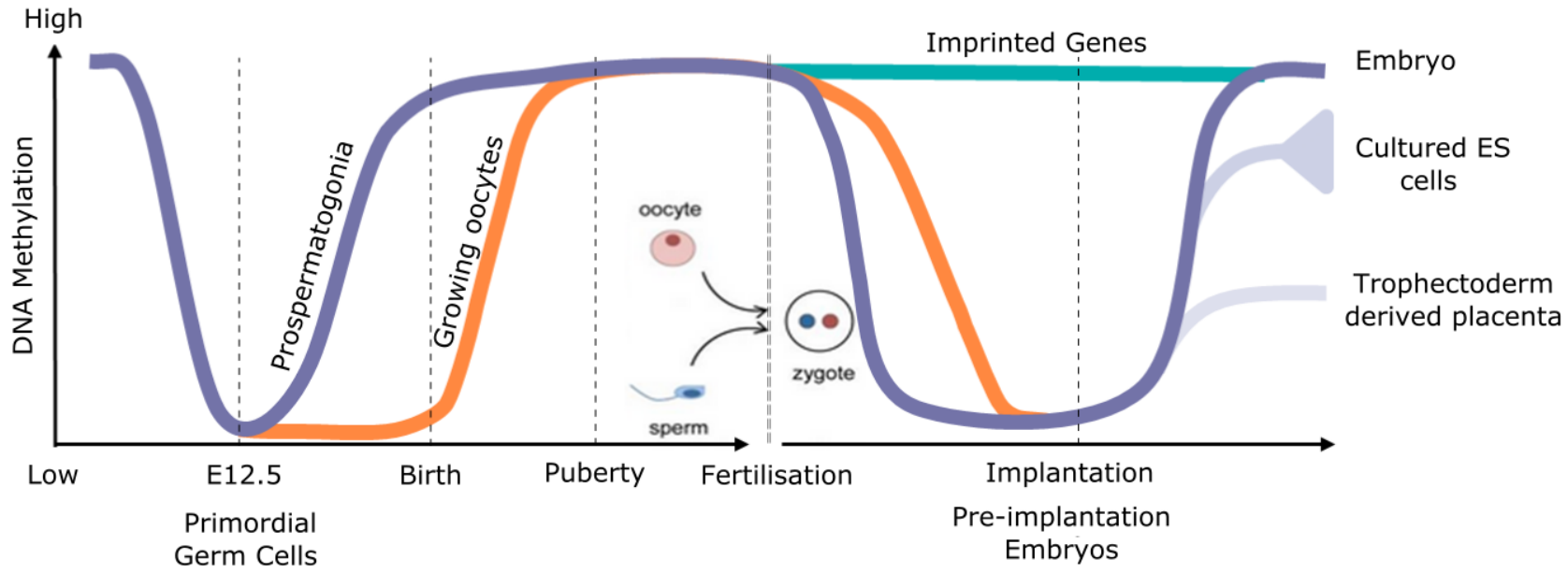
Upon fertilisation, parental genomes remain physically separated for around 24 hours in their own pronuclei until syngamy takes place. During this period the maternal and paternal genomes have asymmetric epigenetic marks and undergo distinct reprogramming events, with loss of DNA methylation occurring quicker in the male pronucleus compared to the female pronucleus (figure 1.2)(Cantone and Fisher, 2013). Two hypotheses as to the mechanism of removal of DNA methylation have recently been explored. The first; active DNA demethylation where 5mCs are removed without the requirement for DNA replication and the second; replication dependent demethylation where the 5mC on the parental strand is not copied to the newly synthesised daughter strand during S phase (Piccolo and Fisher, 2014). Demethylation of the paternal pronucleus was initially thought to be mainly performed by active DNA demethylation as it occurs before DNA replication and in the presence of DNA replication inhibitor (Oswald et al., 2000). More recently, the detection of 5-hydroxymethyl cytosine (5hmC) by immunostaining indicates the active conversion of 5mC to 5hmC via the TET3 enzyme (see section 1.4.2)(Gu et al., 2011; Wossidlo et al., 2011). This was originally not thought to be occurring in the female pronucleus, however low levels of 5hmC have recently been detected by more sensitive genome wide scale reduced representation bisulphite sequencing and a modified version to detect 5hmC and derivatives (Guo et al., 2014; Hu et al., 2014). This suggests that active DNA demethylation may play a role in the female pronucleus, albeit to a lesser

extent than in the male pronucleus (Guo et al., 2014). However, bulk DNA demethylation in both male and female pronuclei appears to be due to passive, replication dependent depletion of 5mC. This was shown after culture of male and female pronuclei with aphidicolin to block DNA replication, which prevented extensive demethylation (Guo et al., 2014; Shen et al., 2014). This replication dependent demethylation is also supported by the increase in hemimethylated DNA molecules after first round replication in the zygote (Arand et al., 2015). Throughout these demethylation events, DNA methylation is maintained at imprinted genomic differentially methylated regions (gDMRs) (figure 1.3). In addition to genome wide demethylation of DNA, the paternal spermatogenic chromatin is decondensed and protamines are exchanged for newly synthesised ooplasm derived histones (Santos et al., 2002).

By approximately E3.5 in mice the inner cell mass (ICM) is extensively globally hypomethylated (figure 1.3)(Cantone and Fisher 2013; Smallwood and Kelsey 2012). At E6.5, specification to PGC lineage can occur which primes these cells for another wave of global DNA demethylation. DNA methylation is fully re-established in the embryonic lineages but only partially in extraembryonic trophectoderm lineages of the blastocyst, creating a striking global epigenetic asymmetry (Chapman et al., 1984; Rossant et al., 1986). DNA methylation levels of ES cells in laboratory culture can vary dependent on the conditions used (Habibi et al., 2013).

Figure Overleaf

Figure 1.3 DNA methylation dynamics during the two phases of global epigenetic reprogramming. Primordial germ cells are specified around E7.25 and DNA methylation is erased until E12.5 where levels are extremely reduced. Germ cell sex determination occurs around E12.5 and the male and female germ cells undergo sex specific de novo DNA methylation with differing timings. Following fertilisation loss of DNA methylation occurs quicker in the male pronucleus than the female pronucleus to lowest levels in pre-implantation embryos. De novo DNA methylation is then re-established to varying extents in the embryo, ES cells and the trophectoderm derived cells of the placenta. The purple line denotes male patterns of DNA methylation erasure and the orange line represents where the female specific reprogramming differs from male. The green line indicates maintenance of DNA methylation at imprinted regions. Figure modified from Smallwood and Kelsey 2012.


High
DNA Methylation
Low
Prospermatogonia
Growing oocytes
oocyte
sperm
zygote
E12.5
Birth
Puberty
Fertilisation
Implantation
Primordial
Germ Cells
Pre-implantation
Embryos
Imprinted Genes
Embryo
Cultured ES
cells
Trophectoderm
derived placenta

1.5 Removing DNA Methylation

DNA methylation can be removed from the genome in a passive manner dependent on DNA replication or actively removed by a family of catalytic proteins called the Fe(II) 2-oxoglutarate dioxygenase Ten-eleven translocation (TETs) enzymes, which convert 5mC to 5hmC, 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (figure 1.4)(Seisenberger et al., 2013). Further potential mechanisms of active removal of DNA methylation exist such as deamination by AID/APOBEC DNA editing family enzymes (Morgan et al., 2004; Popp et al., 2010) which converts 5mC to thymine with the resulting T:G mismatches removed by thymine-DNA-glycosylase (TDG) and the base excision repair (BER) pathway (figure 1.4)(Hajkova et al., 2010; He et al., 2011; Yan et al., 2011).

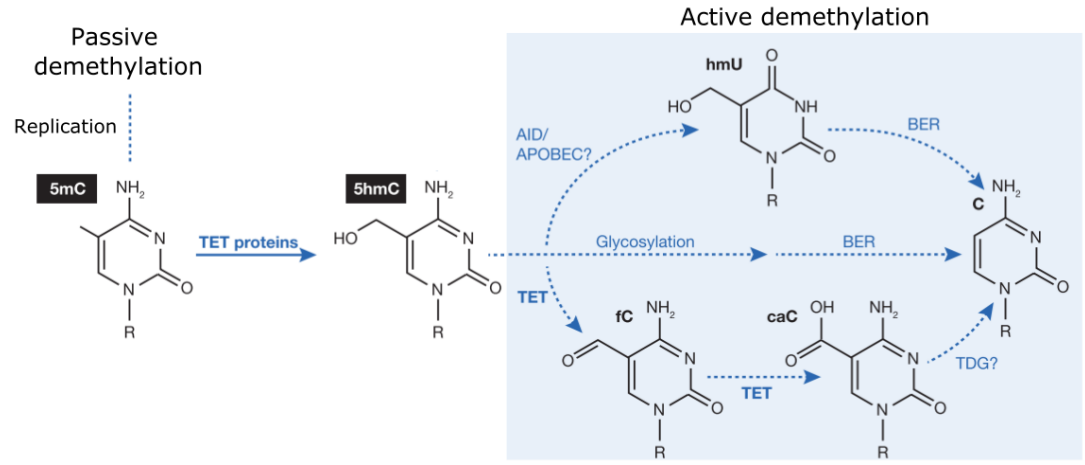


Figure 1.4: Schematic diagram of passive and active TET dependent DNA demethylation. 5mC can be removed in a passive, replication dependent manner or via activity of the TET enzymes which subsequently oxidise 5mC to 5hmC, 5fC and 5caC prior to removal via TDG and the BER pathway. AID/APOBEC provide an additional active removal mechanism whereby 5mC is deaminated and converted to unmethylated cytosine by the BER. (Diagram modified from Williams et al., 2012).

1.4.1 Passive DNA demethylation

Passive loss of DNA methylation can occur by blocking the maintenance machinery in dividing cells by exclusion, enzymatic inhibition or loss of expression of the maintenance methyl transferases (Piccolo and Fisher, 2014). Shortly after post fertilisation, the maternally derived genome of the zygote undergoes progressive demethylation upon cleavage divisions (Mayer et al., 2000). This may be due to the oocyte specific isoform of the maintenance methyltransferase, *Dnmt1o* being excluded from the nucleus of one cell, two cell and four cell embryos (Cardoso and Leonhardt, 1999; Meyenn et al., 2016). During mammalian development PGCs undergo several rounds of cell division where DNA methylation is not maintained, which has been proposed to be due to exclusion of the *Dnmt1* co-factor NP95 from the nucleus (Meyenn et al., 2016; Seisenberger et al., 2012).

1.4.2 Active removal of DNA methylation

Active DNA demethylation has been shown to be facilitated by a family of three catalytic TET enzymes which catalyse the conversion of 5mC to 5hmC and the more oxidised forms 5-fC and 5-caC (figure 1.4)(Ito et al., 2011**;** Tahiliani et al., 2009). Mammalian TET1 and TET2 have highly specific binding to 5mC in a CpG context with little specificity for flanking DNA sequences. Their enzymatic oxidative capacity was demonstrated by *in vitro* enzymatic assays which showed human and mouse TET proteins are significantly more active on 5mC-DNA substrates than 5hmC and 5fC substrates (Hu et al., 2015; Ito et al., 2011).

TET enzymatic function has common outcomes but genetic studies suggest that TET proteins are functionally non-redundant. For example, oocyte and early preimplantation embryonic expression of *Tet* enzymes is restricted to *Tet3* but its expression decreases as pre-implantation development progresses (Gu et al., 2011). *Tet1* and *Tet2* expression is high in the inner cell mass (ICM) and ESCs, but *Tet1* is downregulated upon differentiation, whilst *Tet3* is upregulated (Rasmussen and Helin, 2016). *Tet2* and modest expression of *Tet3* is present in a large number of adult tissues,

however the tissue specific importance of each TET enzyme is still under debate (Rasmussen and Helin, 2016; Szwagierczak et al., 2010; Tahiliani et al., 2009). 5mC levels are relatively constant between tissues, however 5hmC levels in the mouse vary drastically dependent on cell type, possibly as a consequence of differing activity of TET enzymes (Globisch et al., 2010).

Recently, mapping of 5hmC across the whole genome in human and mouse ESCs has provided insights into the genomic location and scale of TET protein function (Rasmussen and Helin, 2016). Traditional bisulphite sequencing cannot discriminate 5mC from 5hmC as both resist deamination upon bisulphite treatment (Yu et al., 2012). Therefore, a modified genome wide bisulphite sequencing approach was developed, whereby all 5mC is oxidised to 5-caC by an excess of recombinant TET1, but all 5hmC is protected from oxidation by the addition of a glucose to 5hmC (Yu et al., 2012). This showed that in ESCs 5hmC is enriched in regions with low CpG content, including promoters associated with lowly expressed genes, gene bodies, intragenic regions and at active enhancers with varying distributions in different tissues (Stroud et al., 2011; Yu et al., 2012). This distribution data provides further understanding of the potential roles for 5hmC in genome regulation, alongside it being a by-product of DNA demethylation.

The physiological developmental relevance of TET proteins has been investigated by genetic knockout of the *Tet* genes and further indicates distinct roles for the different TET proteins. *Tet3* constitutive deletion leads to death shortly after birth indicating an essential role in development (Gu et al., 2011). Maternally inherited TET3 also has roles in reprogramming the paternal pronucleus (Gu et al., 2011). In contrast, individual $Tet1^{-/-}$ or $Tet2^{-/-}$ mice are viable and fertile with no evident developmental phenotype, however these proteins were determined to have redundancy in development as double $Tet1^{-/-}Tet2^{-/-}$ mice exhibit some embryonic lethality with a spectrum of disorders such as exencephaly, growth retardation and defective imprinting (Dawlaty et al., 2013). Despite this, around 40% of $Tet1^{-/-}Tet2^{-/-}$ mice are born grossly normal and are fertile (Dawlaty et al., 2013). TET enzymes contribute to reprogramming methylation patterns at some sequences throughout development and recent findings indicate that addition of vitamin C to embryonic stem cells (ESCs)

promotes TET activity leading to a global increase in 5hmC and reductions in 5mC at different loci in mESCs (Blaschke et al., 2013; Walter et al., 2016).

Recent evidence indicates roles for TET enzymes in active removal of DNA methylation in developing PGCs. TET1 and TET2 enzymes peak in expression in PGCs between E10.5 and E11.5 at the height of DNA methylation removal (Hackett et al., 2013). This reduction in 5mC was accompanied by a global enrichment of 5hmC suggesting conversion by the TET proteins (Hackett et al., 2013). To confirm the conversion was due to TET1 and TET2, inducible knockout *Tet1/Tet2* PGC-like cells were specified from epiblast-like cells and indicated that upon knockdown of *Tet1* and *Tet2* there was a substantial inhibition of DNA demethylation, not only genome wide but at imprinted regions, LINE-1 elements and some IAP loci (Hackett et al., 2013). Furthermore, components of the BER pathway such as poly(ADP-ribose)-polymerase 1 (*Parp1*) and *Tet1* are upregulated in E11.5 PGCs (Hajkova et al., 2010) which indicates that reprogramming in PGCs may involve multiple redundant mechanisms.

1.4.3 TDG catalysed base excision and base excision repair

Another major driver of active DNA demethylation involves thymine-DNA-glycosylate (TDG) catalysed base excision and DNA base excision repair (BER). TDG is a mismatch repair enzyme which was found to bind and excises G:U and G:T base pairs (Lindahl and Wood, 1999). More recently, TDG has been shown to be a critical step in the TET mediated active DNA demethylation pathway (figure 1.4). No binding or catalytic activity of TDG was detected for 5hmC in vitro and indicated the specificity of TDG for 5fC and 5caC, the more highly oxidised forms of 5mC (He et al., 2011; Zhang et al., 2012). TDG is first recruited to 5fC and 5caC and then catalyses base excision of these oxidised cytosine bases, which are replaced by an unmodified cytosine by BER (Nabel et al., 2012). This mechanism was supported by experiments in ESCs null for *Tdg* which had an increase in 5fC and 5caC and HEK293T cells

overexpressing *Tdg* having reduced 5fC and 5caC levels (Nabel et al., 2012). Furthermore, the biological relevance of TDG in development was highlighted by *Tdg*$^{-/-}$ mice which died by E11.5 with haemorrhages and problems with vasculature development, among other defects (Cortázar et al., 2011). In support of the *in vivo* functional studies, these mice exhibited hypermethylation affecting the expression of developmental genes. Failure to establish and maintain correct DNA methylation patterns was proposed to be the likely cause of developmental defects (Cortázar et al., 2011).

Aside from TET and TDG/BER pathways, alternative mechanisms can exist to perform active DNA demethylation. For instance, DNA repair mechanisms such as deamination by the AID/APOBEC family of cytidine deaminases results in conversion of 5hmC to 5hmU followed by BER and replacement with an unmodified cytosine (figure 1.4)(Popp et al., 2010; Morgan et al., 2004). AID has clear activity in PGCs at E13.5, however the phenotype upon *Aid* deletion in PGCs is moderate, indicating the presence of additional mechanisms which actively remove DNA methylation in the absence of AID (Popp et al., 2010). Each of these components of the active DNA demethylation mechanisms have roles in the global epigenetic reprogramming events occurring shortly after fertilisation and in the developing primordial germ cells.

1.6 DNA Methylation Associated Interferon Response

Experimental removal of DNA methylation using DNA demethylating drugs has provided a method to induce retrotransposon de-repression in a variety of cell types and assess the molecular and physiological consequences for many cellular processes. Removing DNA methylation in differentiated cells using 5-Aza de-represses LTR retrotransposons (Chiappinelli et al., 2015; Davis et al., 1989; Roulois et al., 2015; Rowe et al., 2010), suggesting DNA methylation is playing an important role in repressing these retrotransposons in these cells. Interestingly, retrotransposon de-repression in cells treated with 5-Aza induces a type I interferon β response, leading to apoptosis of the treated cells, possibly through the accumulation of retrotransposon-encoded nucleic acids in the cytoplasm (figure 1.5). This mimicks infection by an

exogenous retrovirus (Chiappinelli et al., 2015; Kassiotis and Stoye 2016; Roulois et al., 2015). Repression of retrotransposons by DNA methylation and potentially also by other mechanisms may therefore be important to prevent inappropriate activation of innate immune responses in mammals.

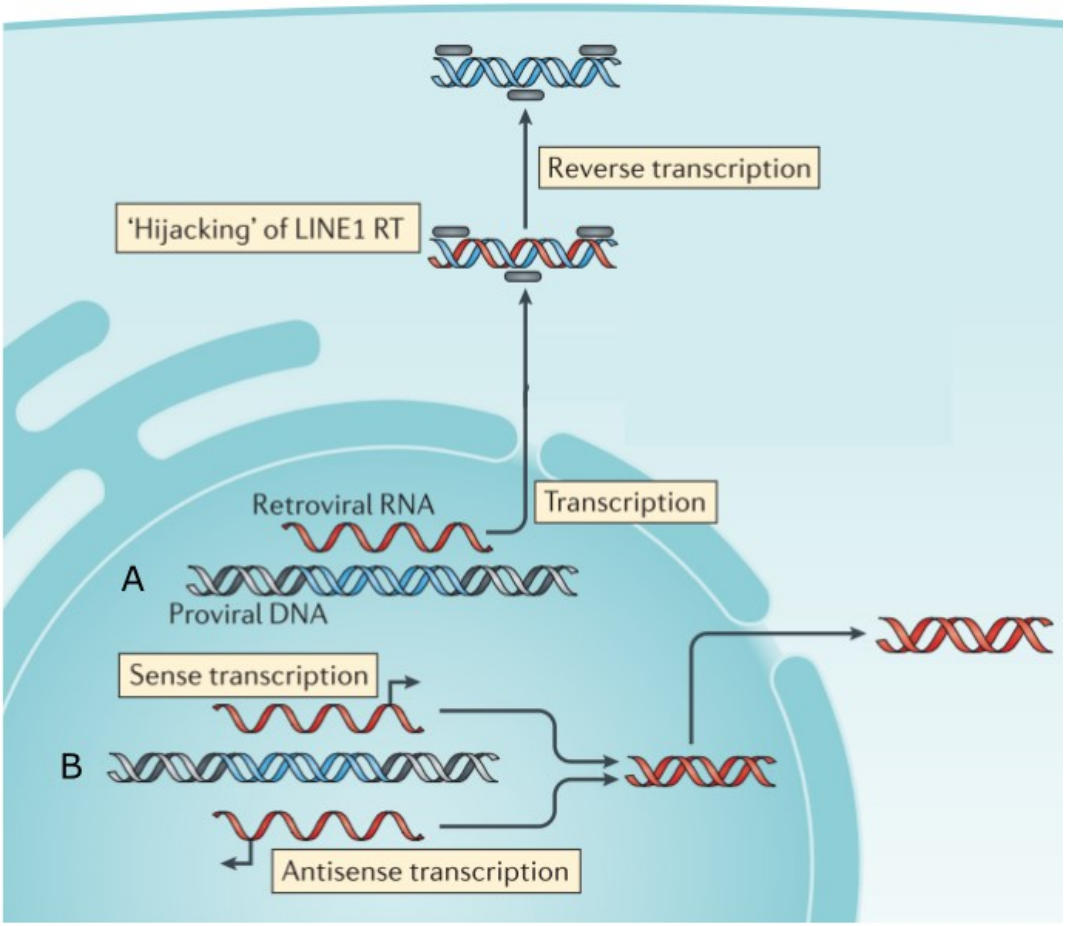


Figure 1.5: Nucleic acid intermediates of transcriptionally de-repressed retrotransposons. (A) Transcriptional de-repression of LINE-1 and LTR retrotransposons (proviral DNA, internal blue helix) can result in production of retroviral RNA (single red line) followed by reverse transcription to double stranded cytosolic DNA (dsDNA)(isolated blue helix), due to the increased availability or 'hijacking' of LINE-1 reverse transcriptase (RT). (B) Additionally, increased complementary sense and antisense single-stranded RNA (ssRNA)(single red lines) transcripts that may form double-stranded RNA (dsRNA)(red helix). Cytosolic dsRNA and dsDNA molecules are visible to RNA and DNA sensors respectively. Dark blue shading denotes the nucleus and light blue, the cytoplasm (Figure modified from Kassiotis and Stoye, 2016)

The immune system is frequently activated in response to invasion of cells via non-self-pathogens, resulting in host cell death (Platanias, 2005). One such mechanism of immune response is the interferon pathway (figure 1.6). The effects of the interferon pathway and activation of interferon stimulated genes are wide ranging as they are pro-inflammatory, enhance adaptive immunity, and are directly antiviral (Schneider et al., 2014). Type I interferons are largely comprised of interferon α (IFN) and IFNβ produced by a broad range of cell types (Platanias, 2005). The type II interferon response is reliant on IFNγ which is predominantly produced by T cells and natural killer cells. Type III interferons are comprised of IFNλ and are largely restricted to epithelial cell surfaces (McNab et al., 2015).

All type I interferons bind a receptor comprised of two subunits known as IFNAR1 and IFNAR2 on the surface of the cell (figure 1.6). The receptor is associated with the Janus kinase (JAK) / Tyrosine kinase (TYK) which when bound by IFNα/β results in tyrosine phosphorylation of STAT1 and STAT2 (Paludan and Bowie, 2013). These then form a trimeric complex, translocate to the nucleus with IRF9 and bind IFN-stimulated response elements (ISREs) in DNA, initiating the transcription of interferon stimulated genes (figure 1.6)(Platanias, 2005). The type II IFN pathway functions via a similar pathway mainly reliant on STAT1 (Platanias, 2005). Depending on the type of viral insult detected, from cytosolic dsRNA or cDNA, different proteins facilitate the immune response. RIG-1 and MDA5 are cytosolic dsRNA sensors and TMEM173 (STING) is a mediator of the response to cytosolic DNA (Paludan and Bowie, 2013).

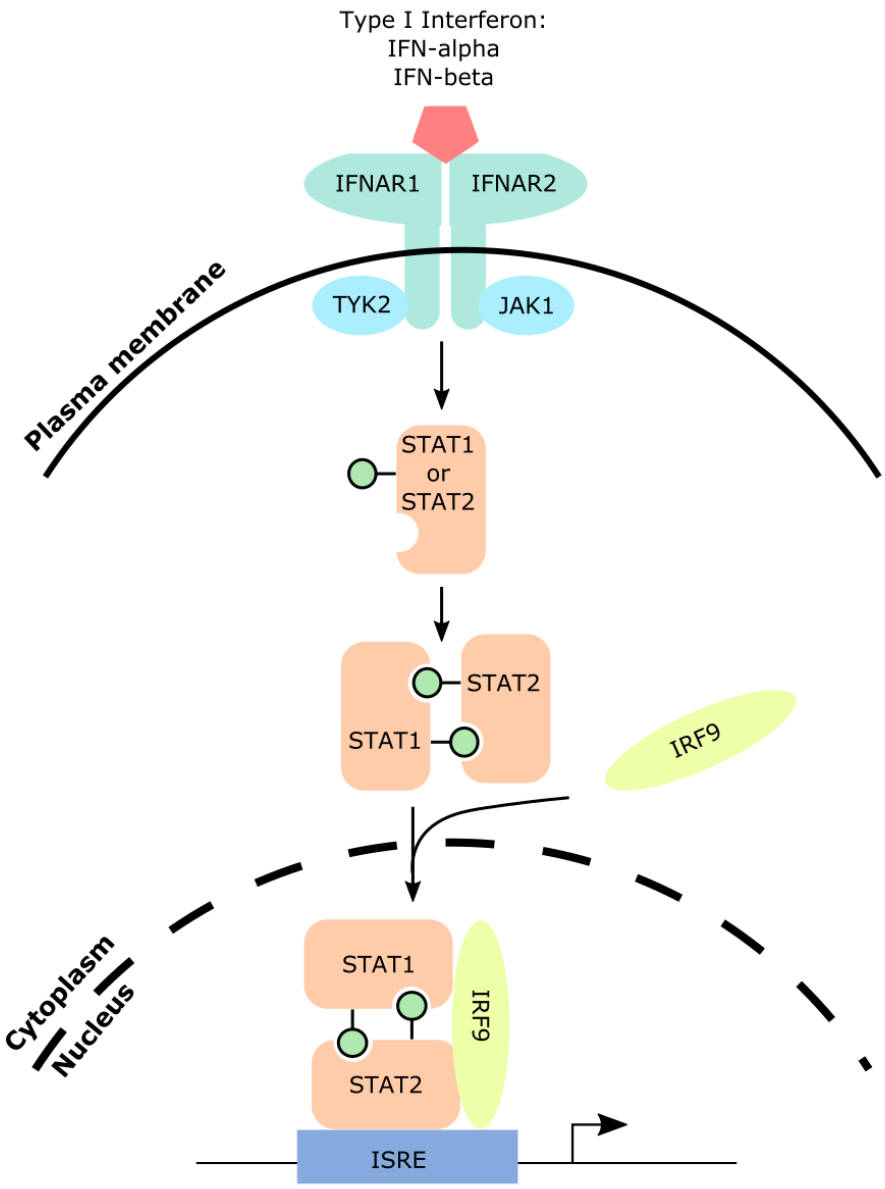


Figure 1.6: Schematic diagram of activation of the JAK-STAT pathway by **type-I interferon.** IFN-alpha or beta is sensed by IFNAR receptors which results in phosphorylation (green circles) of STAT1 or STAT2. STAT1 and STAT2 dimerise and translocate to the nucleus with IRF9 and bind to ISREs upstream of the transcriptional start sites of interferon stimulated genes, inducing transcriptional activation. Figure modified from Platanias 2005.

Activation of a type I interferon response is associated with the chronic autoimmune disease, Aicardi-Goutières Syndrome (AGS) in humans (Behrendt and Roers, 2014). The syndrome is caused by mutations in one of several enzymes involved in nucleic acid metabolism; *TREX1, ADAR1* and *SAMHD1* (Kassiotis and Stoye 2016). Information about the pathogenic mechanisms contributing to AGS has been gained from studies on mouse mutants in these key genes. For instance, *Trex1*$^{-/-}$ mice die around 6 months of age (Morita et al., 2004). They exhibit upregulation of interferon stimulated genes which suggested that loss of *Trex1* results in a spontaneous activation of an anti-viral type I IFN response (Morita et al., 2004; Stetson et al., 2008). This autoimmune response was found to be causative as the IFN response and disease pathogenesis was lost when crossing *Trex1*$^{-/-}$ mice with mice lacking the type-I IFN receptor, IFNAR (Stetson et al., 2008). The pathogenic type I IFN response is triggered by increased levels of endogenous nucleic acids in *Trex1*$^{-/-}$ mice (Stetson et al., 2008). Interestingly, ssDNA fragments from both LINE-1 and LTR retrotransposons were found to be in the cyotosol of *Trex1*$^{-/-}$ cells and were proposed to be endogenous substrates for TREX1 in the heart (Stetson et al., 2008)(figure 1.5). In general, excess cytosolic nucleotides are detected by a cytosolic DNA sensing pathway, reliant on IRF3 and STING (TMEM173). These sensors were deemed to be essential for auto-immune activation as the *Trex1*$^{-/-}$ phenotype was rescued by crossing with mice null for *Irf3* or *Sting* indicating the phenotype was likely arising due to the increase in cytosolic DNA (Gall et al., 2012; Stetson et al., 2008).

The association of AGS and retrotransposon de-repression was further confirmed as 7 out of the 8 known *SAMHD1* mutations in human AGS cause increased L1 retrotransposition when analysed with a GFP reporter screen (Zhao et al., 2013). Along with TREX1 and SAMHD1, ADAR1 protein is also associated with AGS, immune response and retrotransposon activity. The ADAR1 adenosine deaminase proteins bind dsRNA and convert adenosine to inosine to promote an antiviral response (Kim et al., 1994). To determine a role for ADAR1 in retrotransposon repression, transcript expression of livers null for *Adar* and the cytoplasmic dsRNA sensing pathway component, *Mavs* were compared to *Adar*$^{-/-}$ livers (Mannion et al., 2014). This abrogated the cytoplasmic double-stranded RNA removal common to MAVs and indicated increased transcripts of the LTR retrotransposon, *MMERVK10C* (Manion et

al., 2014). This indicates that ADAR is not only involved in AGS pathogenesis but also in retrotransposon repression in the absence of a type I interferon response (Manion et al., 2014). In summary, each of these studies demonstrate retrotransposon increases which are associated with increased type I interferon response contributing to disease state.

1.7 Histone Modifications

A fundamental principle of genome regulation is the ability to dynamically manipulate chromatin state. The first line mechanism of genome regulation is reliant on epigenetic modifications distinct from DNA methylation; modifications to the histones of the nucleosome core of DNA. DNA methylation has then been suggested as a secondary mechanism to 'lock in' transcriptional repression of genes and retrotransposons that is first set up by histone modifications (Cantone and Fisher, 2013).

The nucleosome core has 147 base pairs of DNA periodically wound around a histone octamer comprised of two H2A – H2B dimers and a tetramer of histone H3-H4 (Bannister and Kouzarides, 2011). This forms a 'beads-on-a-string' arrangement which is folded into more condensed fibres to facilitate packaging of around 2 meters of DNA into the nucleus of each cell (Bannister and Kouzarides, 2011). Conserved throughout evolution, histones have a 'tail' of amino acid residues which are subject to post translational modifications such as methylation, acetylation, phosphorylation, ubiquitination, SUMOylation and citrullination which mediate chromatin folding and histone function (Peterson and Laniel, 2004). Chromatin modifications are generated and removed by specific 'writer' and 'eraser' enzymes which can influence chromatin function by directly altering its structure or by recruiting 'reader' proteins that recognise these modifications. Such modifications play important roles in the regulation of gene expression, chromosome compaction and organisation, DNA repair in mitotic cells and transcriptional control of retrotransposons (Bannister and Kouzarides, 2011).

1.6.1 Methylation of histones

The most prevalent modification of core histones is methylation, which mainly occurs in combinations of mono, di or tri-methylation on lysine tails or mono or demethylation on arginine residues. These methylation marks can be deposited or removed by groups of enzymes referred to as histone methyltransferases (HMTs) and histone demethylases (HDMs) respectively. Varying combinations and numbers of methylated sites can be associated with repression of a gene or retrotransposon due to the formation of heterochromatin around those repressed loci (table 1.2)(Kouzarides, 2007). Histone methylation itself does not alter chromatin structure, as the charge of methylated histones is unaltered. Instead histone methylation recruits downstream protein complexes which recognise methylated residues through chromo-domains (Kouzarides, 2007). This is unlike acetylation and phosphorylation which alter the charge of modified histones disrupting their interaction with DNA and therefore the stability of the local chromatin architecture (Bannister and Kouzarides 2011). Importantly, histone methylation appears to be the primary mechanism of retrotransposon repression in ESCs (Goodier 2016). They have been shown to play a lesser role in differentiated somatic cells where retrotransposon repression is largely facilitated by DNA methylation, however a few recent examples appear to be exceptions to this rule (Bulut-Karslioglu et al., 2014; Ecco et al., 2016; Fasching et al., 2015; Karimi et al., 2011; Mikkelsen et al., 2007; Wolf et al., 2015a).

Chromatin Modification	Functions	Histone Methyltransferase	Retrotransposons targeted	Reference
H3K9me1/2	Transcriptional repression	G9a	Specific ERV (MERVL)	Maksakova et al., 2013
H3K9me3	Transcriptional repression	SUV39H1, SUV39H2	Many LINE-1 subtypes	Bulut-Karslioglu et al., 2014
H3K9me3	Transcriptional repression	SETDB1, KAP1 (co-factor),	Many ERVs and some LINE-1 subtypes	Karimi et al., 2011; Matsui et al., 2010; Rowe et al., 2010

Table 1.2: Repressive histone marks, histone methyltransferases and their association with retrotransposon repression

1.6.2 The SUVAR proteins

The first histone methyltransferase detected in mice and humans was of the SUVAR family of proteins and are homologues of *Drosophila Su(var)3-9.* This homologue contained a conserved SET domain known to methylate histones (Rea et al., 2000). The histone tail residues targeted by mammalian SUVAR proteins were unclear and in 2000, Rea et al. showed that in somatic cells lysine 9 of histone 3 was tri-methylated by SUV39H proteins in humans and mice (table 1.2)(Rea et al., 2000). This was demonstrated by i*n vitro* methylation assays using tagged human and mouse SUV39H1, free histones and a methyl donor (Rea et al., 2000). Recent genome wide chromatin immunoprecipitation sequencing experiments (ChIPseq) have revealed SUV39h proteins and SUV39h dependent H3K9me3 are enriched at many retrotransposon sequences in the genome (Bulut-Karslioglu et al., 2014). The majority of intact LINE-1 and some LTR elements are bound by SUV39H and marked for transcriptional repression by H3K9me3 in mouse embryonic stem cells (mESCs). The H3K9me3 was determined to be at least partly SUV39H dependent as mESCs null for SUV39H proteins exhibited reduced SUV39H and H3K9me3 at these loci (Bulut-Karslioglu et al., 2014). This was accompanied by de-repression of the younger L1MdA and L1MdT subfamilies of elements and to a far lesser extent some LTR elements in mESCs (Bulut-Karslioglu et al., 2014). Interestingly, upon differentiation of *Suv39h* null mESCs to neural progenitor cells, SUV39H dependent H3K9me3 was not diminished at LTR or LINE repeats excluding the SUV39H histone methyltransferases from a role in the regulation of retrotransposons in more committed cells, which is mainly provided by DNA methylation (Bulut-Karslioglu et al., 2014; Karimi et al., 2011). Furthermore, the SUV39H proteins play essential roles in development as double knockout mice lacking *Suv39h1* and *Suv39h2* are born at sub-Mendelian ratios with growth retardation of viable pups, sterility and perturbations in H3K9 methylation at pericentric heterochromatin made up of major satellite repeats (Peters et al., 2001a). This indicates functional histone methyltransferase activity of SUVAR proteins *in vivo* (Peters et al., 2001).

1.6.3 SETDB1

Another member of the SUV39 family of histone methyl transferase is SET domain, bifurcated 1 (SETDB1/ESET) which is related but distinct from SUV39H1 and SUV39H2 proteins. SETDB1 is involved in many aspects of development and genome regulation, indicated by its many interacting partners (Kang, 2014). Interactions with important regulators of chromatin state such as KAP1 and DNMT3A (Li et al., 2006; Schultz et al., 2002) along with the presence of a catalytic SET domain indicated the potential importance of SETDB1 in genome regulation (Kang, 2014; Schultz et al., 2002). In initial research aimed to determine the functionality of the SET domain of SETDB1, Schultz et al performed elegant experiments whereby the SET domain of SETDB1 was mutated and was shown to reduce its catalytic histone methyltransferase activity (Schultz et al., 2002). The methyl transferase activity was shown to be highly selective for H3K9 when using recombinant SETDB1 protein, histone tail proteins and methyl donors. Interestingly, unlike SUV39h, SETDB1 can catalyse mono, di or tri methylation of H3K9 (Schultz et al., 2002). This H3K9 methylation was then shown to recruit heterochromatin protein 1 (HP1) to facilitate heterochromatin formation and transcriptional repression (Schultz et al., 2002).

Setdb1 is expressed from the blastocyst and is ubiquitous during post implantation development of the mouse (Dodge et al., 2004). SETDB1 is critical in development as shown by genetic knockout studies which indicated that *Setdb1*$^{-/-}$ embryos arrest at the blastocyst stage of development around E4.5 (Dodge et al., 2004). This is far earlier than the arrest of other H3K9 methyltransferases such as perinatal death of SUV39H1 knockouts (Peters et al., 2001a) and lethality at E9.5 for GLP and G9A knockouts (Tachibana et al., 2002, 2005). Each of these HMTs are expressed at the blastocyst stage, however they fail to compensate for the loss of *Setdb1* (Cho et al., 2011). This highlights the significance of SETDB1 function in early mouse development and implies non-redundant roles for HMTs throughout gestation. No global alterations in H3K9me3 were detected in *Setdb1*$^{-/-}$ blastocysts by immunostaining and it was proposed that this was due to a maternal stock of SETDB1 rescuing global H3K9me3 (Dodge et al., 2004). Despite this, the role of SETDB1 in transcriptional control across

the genome in mESCs has been characterised using refined ChIPseq and RNAseq studies (Bilodeau et al., 2009; Maksakova et al., 2006).

Many loci are susceptible to SETDB1 dependent repression, from genic loci to retrotransposons. SETDB1 is enriched at some subtypes of LTR retrotransposons, along with H3K9me3 in mESCs (Matsui et al., 2010). This H3K9me3 deposition was shown to be SETDB1 dependent as it was reduced in conditional *Setdb1* null mESCs resulting in dramatic de-repression of many LTR and LINE-1 retrotransposons (table 1.2)(Karimi et al., 2011; Matsui et al., 2010). The DNA methylation state at these LINE-1 and LTR loci does not change, suggesting that silencing of most retrotransposons in ESCs occurs relatively independently of DNA methylation (Leung and Lorincz, 2012; Matsui et al., 2010). The overlap in SETDB1 and SUV39H binding and activity at some loci, such as LINE-1 elements, provides redundancy in retrotransposon repression in ESCs, despite SETDB1 having a far greater repertoire of retrotransposon targets (Leung and Lorincz, 2012). When the extent of upregulation of retrotransposons in *Setdb1* null mESC and *Suv39h*$^{-/-}$ mESC was compared by RNAseq, there was a far greater de-repression of LTRs in *Setdb1* null mESCs (Bulut-Karslioglu et al., 2014; Karimi et al., 2011). This indicates the SETDB1 repressive mechanism is far more widespread at LTRs but that LINE-1 is likely co-ordinately regulated by both proteins (Bulut-Karslioglu et al., 2014). Similarly to *Suv39h* null somatic cells, *Setdb1* null embryonic fibroblasts have only limited de-repression of a subset of MLV retrotransposons which is likely an indirect effect as no retrotransposon loci are bound by SETDB1 in this cell type, further indicating a lesser role of histone modifications in differentiated cell types (Hutnick et al., 2010; Matsui et al., 2010). Interestingly, recent studies are finding direct roles for SETDB1 in retrotransposon silencing in somatic cells such as the de-repression of LTRs in *Setdb1*$^{-/-}$ B-lymphocytes (Collins et al., 2015).

Recent evidence also indicates that SETDB1 is required for germline development and silencing of H3K9me3 marked LTRs in primordial germ cells (Liu et al., 2014b). The technical challenges associated with genome wide DNA methylation and ChIP analysis of histone marks during PGC reprogramming have recently been overcome and has provided extensive knowledge of the extent of SETDB1 mediated repression at specific genomic loci at E13.5 (Liu et al., 2014). This analysis misses the window

of lowest histone modifications at E11.5, therefore it is unclear if histone marks are retained at retrotransposon loci during reprogramming, however it indicates many retrotransposons from LTR (IAP, MMERVK10C) to LINE-1 elements (L1MdA, L1MdT, L1MdG and L1MdF) are heavily marked by H3K9me3 and H3K27me3 in E13.5 PGCs of both sexes (Liu et al., 2014). Global decreases in this mark were detected by refined small sample ChIPseq experiments for H3K9me3 in PGCs with conditional deletion of *Setdb1*. A reduction of this mark was detected at LTR elements such as IAP, along with some LINE-1 subfamilies (Liu et al., 2014). Interestingly, a decrease in H3K27me3 was also observed in conditional null *Setdb1* PGCs suggesting possible interdependence of SETDB1 and polycomb repressive mechanisms in PGCs (Liu et al., 2014). SETDB1 mediated repression appeared to be relatively sex specific as IAP elements are highly de-repressed in male *Setdb1* null PGCs and whereas MMERVK10C was most highly de-repressed in female PGCs, with only modest changes in IAP expression (Liu et al., 2014). The reason for this sex specific nature of repression remains unclear but may be reflective of the differing reprogramming dynamics associated with sex specific PGC development. Alternative transcriptional or post-transcriptional silencing pathways for LINE-1 in PGCs were proposed as LINE-1 elements exhibited very little reactivation in *Setdb1* null PGCs despite reduced H3K9me3 (Liu et al 2014).

1.6.4 KAP1 and KRAB-ZFP mediated retrotransposons repression

The binding of SETDB1 to retrotransposon loci is largely dependent on the transcriptional repressive co-factor, KRAB domain-associated protein 1 (KAP1/TRIM28). Potential roles for KAP1 in transcriptional repression were indicated when KAP1 was demonstrated to have an N-terminal RING finger B boxes coiled coil motif (RBCC) which facilitates targeting to or assembly of different protein complexes (Saurin et al., 1996) and a C-terminal region containing a PHD finger and a bromodomain, which are present in many transcriptional cofactors acting at the chromatin level (Aasland and Stewart, 1995; Friedman et al., 1996). Further early analysis indicated that the RBCC domain is necessary and sufficient for KAP1 binding as a homotrimer to the KRAB repression motif of KRAB-ZNFs (figure 1.4)(Peng et

al., 2000, 2002) and that the PHD finger and bromodomain are both required to obtain maximum levels of KAP1 mediated transcriptional repression via targeting to genic and retrotransposon loci by KRAB-ZFPs (figure 1.7)(Iyengar et al., 2011; Schultz et al., 2001; Sripathy et al., 2006).

After recruitment to the genome by KRAB-ZFPs, KAP1 induces repressive histone modifications by recruiting multiprotein chromatin modifying complexes including SETDB1 which adds the H3K9me3 mark to facilitate transcriptional repression (figure 1.7). SUV39H proteins are not thought to be recruited to retrotransposon loci by KAP1 (Bulut-Karslioglu et al., 2014). The interaction with SETDB1 is dependent on the sumoylation of the KAP1 bromodomain by its PHD domain (figure 1.7)(Ivanov et al., 2007). Heterochromatin protein 1 (HP1) can bind to the PxVxL motif of KAP1 and to H3K9me3 at repressed loci, stabilising the bound repressive complex (Schultz et al., 2002; Sripathy et al., 2006). KAP1 also binds to the NuRD histone deacetylase complex (HDAC) which removes histone acetylation, an active mark of transcription as an additional mechanism of repression (Schultz et al., 2001, 2002). Furthermore, KAP1 has been reported to interact with DNMT3A and DNMT3B and the KRAB-ZFP:KAP1:SETDB1 repressive complex can direct *de novo* DNA methylation to LTRs (Li et al., 2008; Quenneville et al., 2011; Rowe et al., 2013a; Zuo et al., 2012). As KAP1 is generally important to the function of each of these interacting partners, it is unsurprising that removal of KAP1 in many different contexts results in dramatic developmental phenotypes and extensive transcriptional alterations (Iyengar and Farnham, 2011).

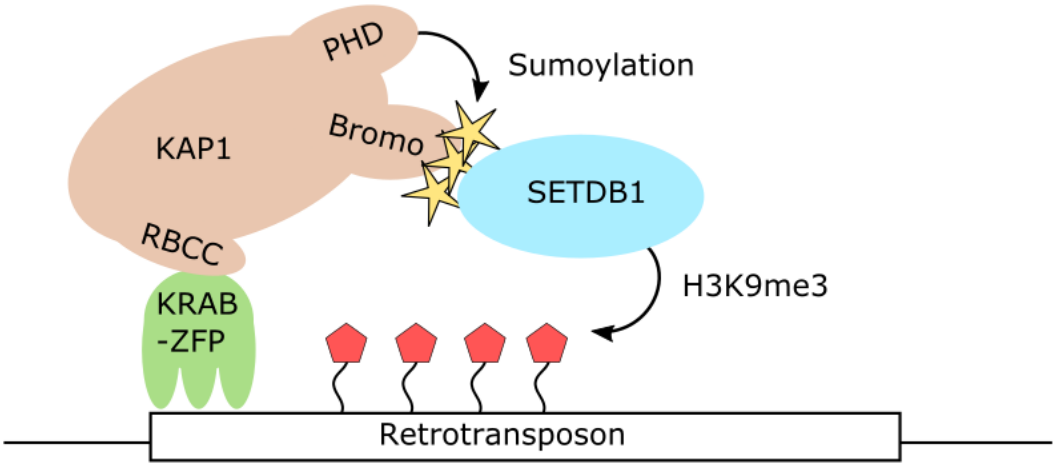


Figure 1.7: Schematic diagram of the mechanism of KRAB-ZFP:KAP1:SETDB1 mediated transcriptional repression. KRAB-ZFPs bind to distinct genomic loci in a sequence dependent manner and recruit KAP1 via binding to its RBCC domain. The PHD domain of KAP1 sumoylates the Bromodomain to facilitate the interaction between SETDB1. SETDB1 then trimethylated histone 3 lysine 9 (red pentagons) to transcriptionally repress the target loci. Interactions of KAP1 with HP1, NuRD or the DNMT proteins are not shown. Modified from Iyengar and Farnham 2011.

Kap1 is expressed ubiquitously and extensive roles for KAP1 have been characterised throughout development and beyond. The importance of the protein in early development was demonstrated by *Kap1*$^{-/-}$ mice which die at E5.5 (Cammas et al., 2000). *Kap1*$^{-/-}$ embryos did not undergo gastrulation and those embryos surviving longer exhibited severe defects in development of the 3 germ layers (Cammas et al., 2000). In addition to this important role in embryonic development, KAP1 exhibits diverse roles across many different aspects of physiology. Roles in cell differentiation, DNA damage response, virus replication, tumourigenesis and repression of retrotransposons have been determined in a variety of conditional *Kap1* knockout models in numerous cell types and tissues (Cammas et al., 2000; Cheng, 2014; Friedman et al., 1996; Iyengar and Farnham, 2011).

As genetic knockout of *Kap1* has a severe embryonic lethal phenotype the role of KAP1 in transcriptional regulation in ES cells was studied by conditional removal of *Kap1* (Rowe et al., 2010). *Kap1* null ES cells exhibited dramatically increased expression of many LTR retrotransposons such as IAP and MMERVK10C but more modest de-repression of LINE-1 elements (Rowe et al., 2010). To determine if this de-repression was a consequence of reduced H3K9me3, Rowe et al performed ChIPseq and indicated dramatic reductions at de-repressed retrotransposons (Rowe et al., 2010). De-repression of IAP was also detected in E5.5 – E6.5 *Kap1*$^{-/-}$ embryos and cultured *Kap1* depleted blastocysts by *in situ* hybridisation (ISH) and qRT-PCR (Rowe et al., 2010). This KAP1 dependent retrotransposon repression appears to be restricted to pluripotent cells as knockout of *Kap1* in murine embryonic fibroblasts (MEFs), hepatocytes and white blood cells has little effect on the expression of retrotransposons (Bojkowska et al., 2012; Rowe et al., 2010; Sio et al., 2013). Taken together, it was thought that the KAP1/SETDB1 mediated repression of retrotransposons in early embryogenesis leads to irreversible silencing, which is maintained independently of KAP1/SETDB1, by DNA methylation following differentiation (Matsui et al., 2010; Rowe et al., 2010; Wiznerowicz et al., 2007; Wolf et al., 2015a). Recently however, this hypothesis is being questioned by various studies which indicate roles for continued KAP1 mediated retrotransposon repression beyond development. For example, de-repression of IAP and MMERVK10C elements occurs in conditional *Kap1* null neural progenitor cells, accompanied by decreases in H3K9me3, indicating that histone mediated repression is more dynamic than previously thought (Fasching et al., 2015). This dynamic, tissue specific nature of repression is likely due to differing transcription factors or KRAB-ZFPs in differing tissues (Ecco et al., 2016; Fasching et al., 2015).

1.6.5 KRAB-ZFPs

The mechanism of KAP1-KRAB-ZFP mediated transcriptional repression was investigated by tethering KAP1 to DNA as a Gal4-KAP1 fusion protein (Sripathy et al., 2006). This facilitated repression of reporter genes only when bound, indicating

that for KAP1 to elicit its repressive capacity it must be bound to DNA (Sripathy et al., 2006). Furthermore, KRAB-ZFPs were proposed to be auto-regulated by KAP1 and sufficient to direct KAP1 and H3K9me3 to the genome in human cells (Iyengar et al., 2011; O'Geen et al., 2007). Recent evidence suggests the vastly widespread nature of this KAP1 dependent mechanism with examples of KRAB-ZFP KAP1 dependent retrotransposon repression in mouse and human ESCs and even somatic, terminally differentiated tissues, in a highly cell type specific manner (Castro-Diaz et al., 2014; Ecco et al., 2016; Wolf et al., 2015a).

Two studies present compelling evidence to suggest that early embryonic control of LINE-1 in humans and mice is an evolutionarily dynamic process in which the KRAB-ZFP KAP1 repressive mechanism has evolved alongside LINE-1 elements (Castro-Diaz et al., 2014; Jacobs et al., 2014). Both indicate, using genome wide chromatin immunoprecipitation experiments (ChIP), that KAP1 binds to particular subfamilies of LINE-1 in both human and mESCs. Using a candidate based (Jacobs et al., 2014) and a more global approach (Castro-Diaz et al., 2014) they further identify the KRAB-ZFPs responsible for targeting KAP1 and H3K9me3 to particular LINE-1 retrotransposon sequences. Generally the younger LINE-1 elements are regulated by DNA methylation until a specific KRAB-ZFP has sufficient time to evolve to target it with the KAP1 mediated H3K9me3 repressive mechanisms, which are mainly active on the intermediate aged LINE-1 elements (Castro-Diaz et al., 2014; Jacobs et al., 2014).

The role for KRAB-ZFPs in differentiated cells has, until recently, been unclear. As mentioned, retrotransposon silencing has been proposed to be established early in development by KAP1, which induces H3K9me3 repressive chromatin domains, followed by DNA methylation (Playfoot and Adams, 2016). It was thought that this silencing could then be maintained independently of KRAB-ZFP and KAP1 function in differentiated cells, however this analysis was based on low numbers of KRAB-ZFPs and retrotransposons (Rowe et al., 2010; Wolf et al., 2015). Recently, a targeted GFP repression-based functional screen of over 200 KRAB-ZFPs in somatic cell lines, identified KRAB-ZFPs that recognise selected retrotransposon derived sequences in the mouse somatic cells and are necessary for their repression (Ecco et al. 2016). Firstly, the retrotransposon repressive capacity of two of these, ZFP932 and its paralog

Gm15336 were demonstrated in mESCs. They then show that in multiple somatic contexts, depletion of the specific ZFPs or KAP1 are sufficient to de-repress their target retrotransposon loci, indicating that for at least some retrotransposons transcriptional repression in somatic cells is achieved through the persistent identification by KRAB-ZFPs and the continual recruitment of KAP1 (Ecco et al., 2016). The evolutionary drive to evolve ZFPs that target rapidly evolving retrotransposons has been proposed to contribute to the large number of species-specific ZFPs typically found in mammalian genomes (Wolf et al., 2015b).

1.8 Retrotransposons in Genome Regulation

Retrotransposon activation in permissive epigenetic reprogramming phases, along with retrotransposon de-repression in experimental contexts can pose a mutational threat to the genome with potentially deleterious consequences. Far from this negative aspect of retrotransposon biology, recent studies have shown that retrotransposons have played a prominent role in shaping the regulatory landscape of the mammalian genome (Thompson et al., 2016). The idea that retrotransposons could participate in gene regulation was proposed by the founder of transposable elements, Barbara McClintock over 60 years ago (McClintock, 1950). This was based on studies in maize whereby transposition resulted in gain of new traits for the recipient crop. Recently evidence has arisen that cements retrotransposons as drivers of genome regulation by their co-option as regulatory elements independent of their capacity for retrotransposition (Thompson et al., 2016). Many studies are now indicating mechanisms for this repetitive element mediated genome regulation. In normal mouse development, there is evidence of repeat elements with enhancer function (figure 1.8)(Chuong et al., 2013, 2016), or alternative promoter activities which regulates the expression of neighbouring genes (figure 1.8)(Peaston et al., 2004; Thompson et al., 2016). A study by Chuong et al in 2013 showed that LTRs are acting as enhancers for core factors necessary for defining the trophoblast stem cell (TSC) transcriptional network. Enhancer associated marks histone H3 lysine 4 monomethylation (H3K4me1) and histone H3 lysine 27 acetylation (H3K27ac) were enriched at

particular LTRs in TSCs as determined by ChIPseq and these elements were functionally capable of driving expression in a rat placental cell line (Chuong et al., 2013). LTRs can also function as alternative promoters in development (figure 1.8). An early version of a genome wide expression study from mouse oocytes used a cDNA library and analysed the amount of repetitive element expressed sequence tags (Peaston et al., 2004). They identified relatively high expression of the MaLR class of retrotransposons in full grown oocytes and determined that retrotransposon LTRs were functioning as primary or alternative promoters contributing to oocyte specific gene expression by analysis of chimeric transcripts (Peaston et al., 2004). Importantly, studies to date have indicated that co-opted LTRs are particularly prevalent in early embryonic development, germ cells and pluripotent stem cells (Thompson et al., 2016).

Alongside normal roles in development, aberrant de-repression of retrotransposons in genetic knockout contexts can result in non-developmentally programmed effects on neighbouring genes (Herquel et al., 2013; Rowe et al., 2013b). Conditional knockout of the transcriptional co-repressor *Kap1* in ES cells results in dramatic de-repression of LTR retrotransposons and their loss of repressive histone marks due to replacement with the marks of putative enhancers, H3K4me1 and H3K27ac (Rowe et al., 2010, 2013). Aberrantly de-repressed retrotransposon acting as alternative promoters has also been reported in ES cells lacking *Setdb1* or *Kap1* (Ecco et al., 2016; Karimi et al., 2011). Despite the wealth of evidence of de-repressed retrotransposons altering genic transcription in mouse knockout contexts (Ecco et al., 2016; Herquel et al., 2013; Rowe et al., 2013b), evidence of transcriptional effects on genes with functional relevance to the disease phenotype is yet to be shown. This presents an interesting possibility when determining the possible cause of disease phenotypes, correlating with aberrant retrotransposon de-repression.

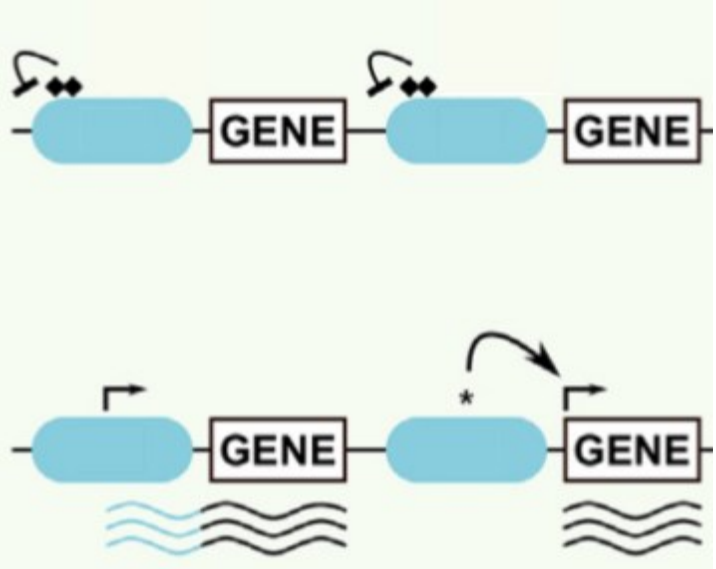


Figure 1.8: Effects of retrotransposons on neighbouring genes. Repression of retrotransposons (blue ovals) via DNA methylation or histone modifications (black diamonds) limits their activity (bar arrows). During normal pre-programmed development or when these epigenetic repressive mechanisms are perturbed, retrotransposons can become de-repressed, enabling them to gain alternative promoter activity (corner arrow) or ectopic enhancer function (asterisk) to stimulate transcription of nearby genes (wavy lines). Figure modified from Playfoot and Adams 2016.

1.9 Methylation Sensitive Genome Defence Genes

Numerous genes involved in supressing retrotransposons in the germline have been shown to be regulated by DNA methylation (Crichton et al., 2013; Hackett et al., 2012). Genes such as *Mael, Mvh, Zfp42* and *Dazl* require the *de novo* methyltransferase *Dnmt3b* dependent promoter DNA methylation to silence their expression in the early embryo (Borgel et al., 2010). Additional methylation sensitive genes were identified from a study to detect gene expression changes, whose expression is primarily and causally regulated by promoter DNA methylation in multiple hypomethylatated somatic cell models (Hackett et al., 2012). Mouse embryonic fibroblasts (MEFs) were treated with the DNA methyltransferase inhibitor, 5-Aza and allowed to recover of expression. Genes identified in this screen were then intersected with genes expressed in MEFs lacking the maintenance methyltransferase, *Dnmt1* (Hackett et al., 2012).

This largely identified methylation sensitive germline genes with roles in spermatogenesis and defence against retrotransposons in the germline such as *Tex19.1, Mili, Miwi2, Dazl, Asz1* and *Mov10L* among others (Hackett et al., 2012). The promoters of these genes are not enriched for repressive histone modifications when repressed in somatic cells indicating why they are highly responsive to DNA methylation.

In an *in vivo* developmental setting, the initial passive reduction in DNA methylation from E8.5 to E9.5 in PGCs correlates with increases in *Tex19.1* and *Mili* expression, the early responders to this epigenetic reprogramming (figure 1.9)(Hackett et al., 2012). *Dazl, Asz1* and *Mov10l* are induced during the proposed active DNA demethylation in PGCs from E10.5 to E11.5 (figure 1.9)(Hackett et al., 2012). Expression of both groups of methylation sensitive genes is maintained through germ cell development and adult oocytes, whilst expression declines in pachytene spermatocytes and round spermatids (figure 1.9)(Crichton et al., 2014). This induction of expression of methylation sensitive genome defence genes during the epigenetic reprogramming of PGCs and into spermatocyte development provides an essential level of retrotransposon repression when the genome is especially vulnerable to potentially deleterious retrotransposition events (Crichton et al., 2014; Goodier et al., 2016). Surprisingly, the hypomethylated state of the placenta results in the hypomethylation and expression of only one of the methylation sensitive genome defence genes, *Tex19.1* (Reichmann et al., 2013).

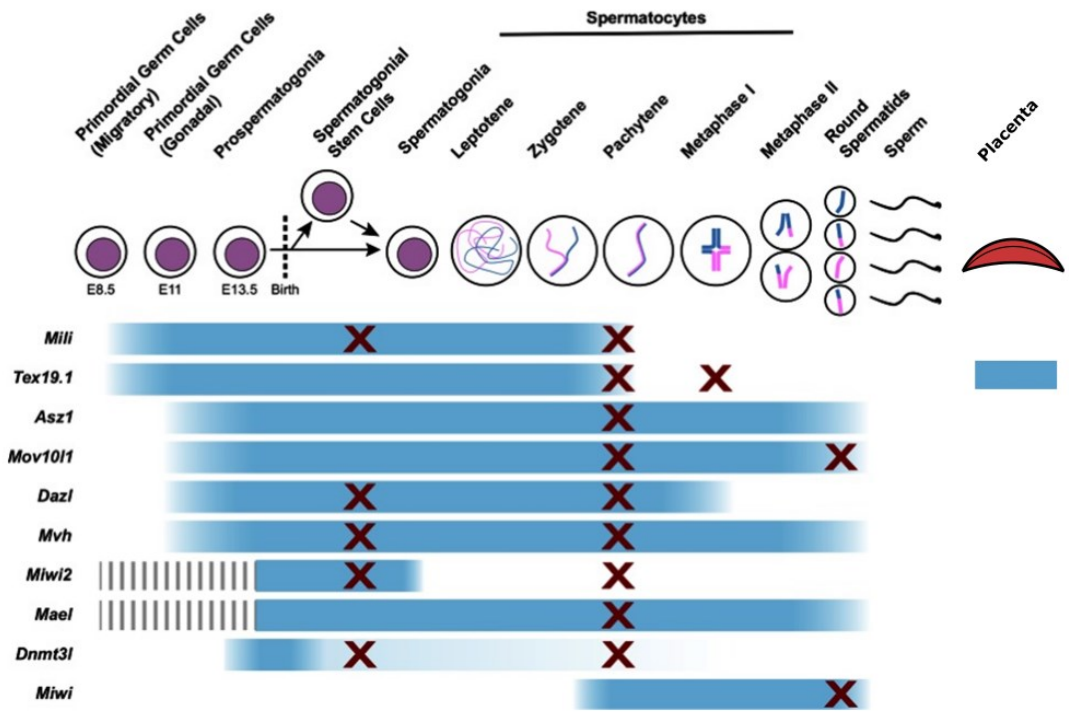


Figure 1.9 Germline genome defence gene expression throughout spermatogenesis and in the placenta and mutant phenotypes. The stages of spermatogenesis from PGCs to mature sperm are indicated along the top of the diagram. The placenta is at the farthest right and indicates that *Tex19.1* is the only methylation sensitive genome defence gene expressed in this tissue. Expression patterns of the indicated germline genome defence genes are indicated by the blue bars, with red crosses indicating the stages at which mutant mice have defects in progression through spermatogenesis. Modified from Crichton et al., 2014.

Experimental removal of each of these defence genes in mice results in de-repression of particular retrotransposons in germ cells and defects in progression through meiosis leading to sterility, alongside placental defects in *Tex19.1*$^{-/-}$ animals. *Tex19.1*$^{-/-}$ pachytene spermatocytes have increased *MMERVK10C* RNA, whereas *Mael, Mili, Miwi, Mov10l* and *Asz1* individual knockout mice have increases in LINE-1 expression and in some cases, IAP (Carmell et al., 2007; De Fazio et al., 2011; Ma et al., 2009; Ollinger et al., 2008; Soper et al., 2008; Zheng et al., 2010). The latter genes are all core components of the PIWI-piRNA mechanism which direct DNA methylation to retrotransposon loci for silencing in spermatocytes. MILI and MIWI physically interact with PIWI-interacting RNAs (piRNAs) from retrotransposons and piRNA clusters and facilitate retrotransposon silencing predominantly in the germline via

ping-pong amplification of piRNA. An antisense piRNA directs the cleavage of a complimentary mRNA from a piRNA cluster or retrotransposon sequence, generating sense piRNAs to further increase the amount of cleavage of complimentary mRNA. This then directs *de novo* DNA methylation to target loci for repression (figure 1.10)(Crichton et al., 2014). Male mice null for any one of these genes largely results in arrest at pachytene of meiosis I. Some mutant animals also exhibit defects earlier, in spermatogonial stem cell divisions (Crichton et al., 2014). Mutations in PIWI-piRNA genes result in defects largely restricted to spermatogenesis, with little effect on post-pachytene progression of oocytes (Crichton et al., 2014).

Together, these methylation sensitive germline defence genes are critical for germ cell development and the repression of retrotransposons at a time when active retrotransposition would have potentially deleterious consequences for future generations.

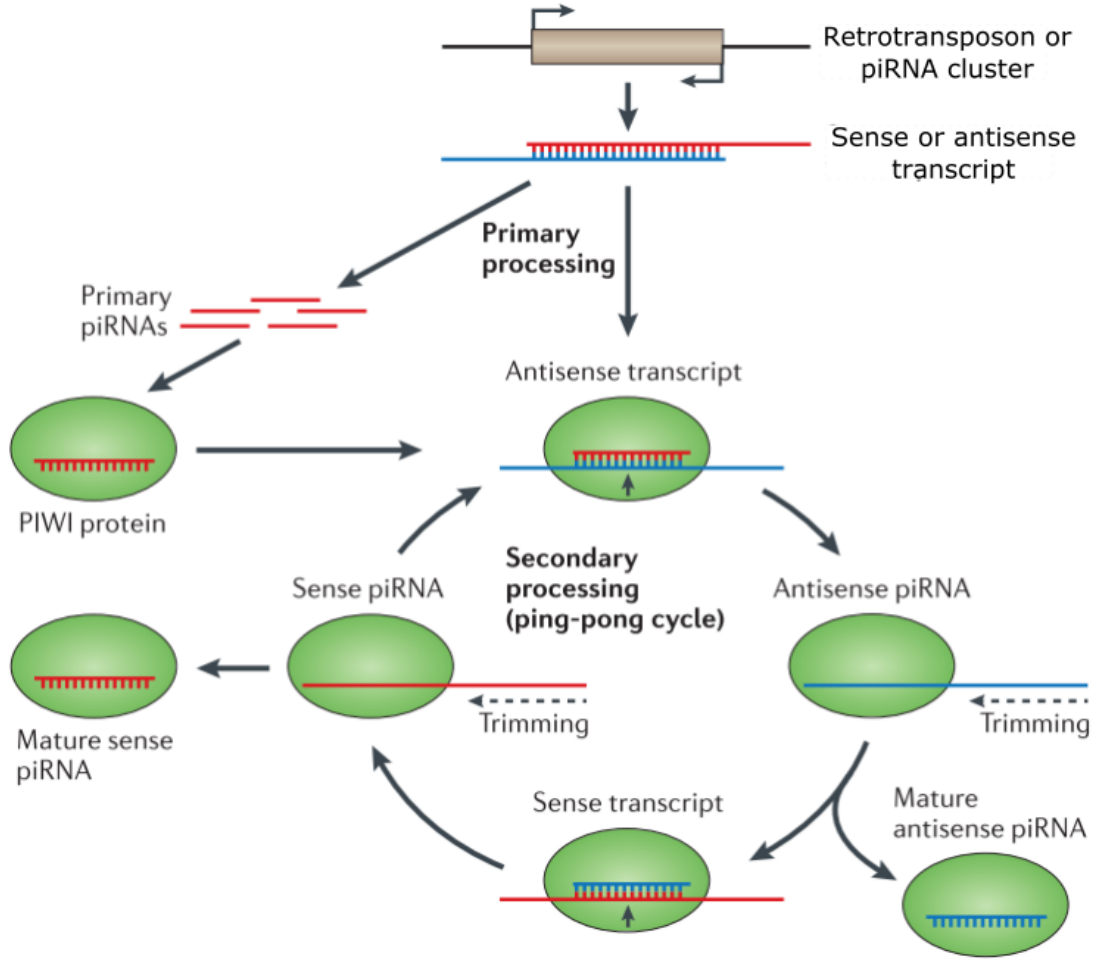


Figure 1.10: The piRNA ping-pong amplification loop. The primary processing pathway generates primary piRNAs from the long transposon transcript by nuclease cleavage in either sense or antisense directions (red and blue lines). piRNAs associate with PIWI proteins (green ovals) such as MILI and MIWI and guide PIWI proteins to complementary sequences on antisense transcripts from the same retrotransposon or piRNA cluster and use their slicer activity to cleave the transcript to generate new 5' ends ready for association with another PIWI protein. (Figure modified from Meister, 2013).

1.10 Spermatogenesis

1.9.1 The spermatogenic cycle

PGCs are one of the main globally identified hypomethylated cell types which eventually give rise to all subsequent germ cells throughout spermatogenesis. In this section I will describe the development of spermatocytes arising from these hypomethylated PGCs. Spermatogenic lineage development occurs in a complex but orderly, synchronised manner within seminiferous tubules and is referred to as the spermatogenic cycle (Clermont, 1963). Germ cells at all stages of development are closely supported by somatic sertoli cells, the 'nurse' cells of spermatogenesis (figure 1.11B). These have both structural and hormonal secretory roles to facilitate germ cell development (Griswold, 1998). In mouse there are 12 stages of spermatogenesis, characterised by a combination of spermatogonia, spermatocytes and spermatids in different stages of their respective development or cell divisions (figure 1.11B). The duration of each stage is precisely timed with a complete spermatogenic cycle lasting around 35 days in mice and 75 days in humans (Clermont 1963; Oakberg 1956).

Spermatogenic lineage development occurs when PGCs undergo sex specification at E11.5-E12.5 in response to signals that depend on Sertoli cells in the developing male gonad (Adams and McLaren, 2002). Sertoli cells are induced to differentiate along a male pathway by expression of the Y-encoded *Sry* gene in gonadal somatic cells (Palmer and Burgoyne, 1991) and by around E12.5 the male-committed prospermatogonia are surrounded by clusters of Sertoli cells which have organised into cord structures which later differentiate into seminiferous tubules. After a period of quiescence, the prospermatogonia resume proliferation a few days after birth and give rise to both mitotic spermatogonia and spermatogonial stem cells (SSCs). The process of self-renewal of SSCs provide the basis for all subsequent spermatogonia and germ cells necessary for spermatogenesis throughout the majority of a males lifespan (De Rooij and Russell, 2000; Yoshida et al., 2006).

1.9.2 Spermatogonial differentiation

Individual diploid SSCs called A_{single} (A_s) cells undergo mitotic division to form either two new diploid A_s spermatogonia or two clones called A_{pr} connected by an intercellular cytoplasmic bridge to facilitate exchange of gene products allowing for synchronised development (figure 1.11A)(Kato et al., 2004; Weber and Russell, 1987). Each A_{pr} spermatogonia will undergo mitosis to form spermatogonial clones of increasing length from 4, 8 and 16 $A_{aligned}$ cells connected by these bridges (figure 1.11A). Each of these cells are negative for the marker of differentiation of spermatogonia *C-Kit* and positive for undifferentiated spermatogonia specific genes such as *Plzf* and are therefore considered undifferentiated (Costoya et al., 2004; Nakagawa et al., 2006). *C-kit* begins to be expressed when $A_{aligned}$ spermatogonia undergo cellular transformation to A1 differentiated spermatogonia (figure 1.11A)(Schrans-Stassen et al., 2001). This is followed by rounds of mitosis and the formation of A_2, A_3 and A_4 spermatogonia (figure 1.11A)(De Rooij and Russell, 2000). A_4 spermatogonia number around 128 cells originating from the initial A_s undifferentiated spermatogonia and give rise to more mature intermediate spermatogonia (In) which further divide to form around 500 type B spermatogonia ready to undergo a final mitotic division, before entering meiotic prophase as pre-leptotene spermatocytes (figure 1.11A)(Oatley and Brinster, 2008).

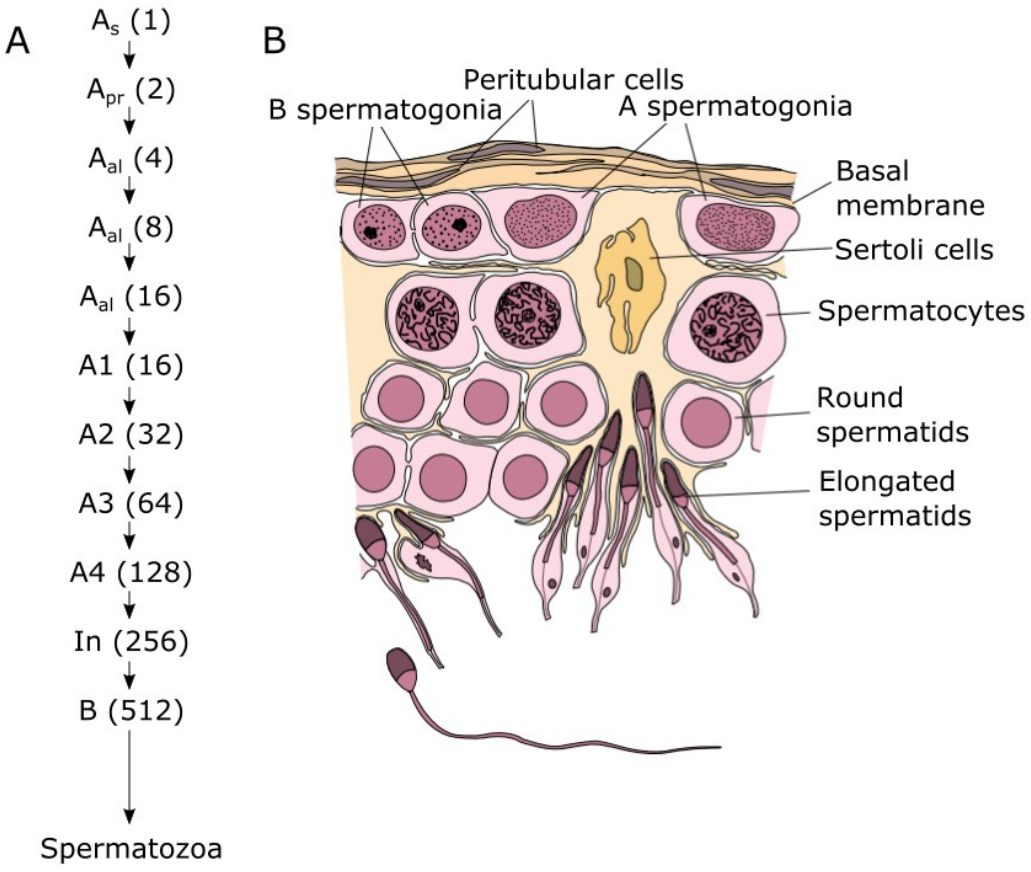


Figure 1.11 Spermatogenesis in murine testes. (A) A_s represents spermatogonial stem cells which give rise to all subsequent germ cells via A_{pr} spermatogonia which mitotically divide to give rise to $A_{al(4)}$ spermatogonia followed by $A_{al(8)}$ and $A_{al(16)}$ connected via cytoplasmic bridges. More mitotic division occur finally giving rise to type B spermatogonia. These enter meiosis and give rise to haploid spermatids which will undergo spermiogenesis to produce mature spermatozoa. Figure from Oatley and Brinster 2008. (B) Schematic cross section through a seminiferous tubule in the murine testis. Spermatogenesis occurs in a synchronous wave from the basal membrane to the lumen of the tubule supported by Sertoli cells. Type A spermatogonia are adjacent to the basal membrane. Cells undergoing the first round of meiosis are called primary spermatocytes whereas during meiosis II they are known as secondary spermatocytes. This final reductional division results in haploid elongated spermatids (figure was taken from de Rooij and Mizrak, 2008)

1.9.3 Meiosis and spermatogenesis

Meiosis is a specialised cell division involving a single round of DNA replication followed by two rounds of chromosome segregation. The first reductional division involves an extended prophase, whereby various meiosis specific events occur to ensure exchange of genetic material can occur between homologous chromosomes from the mother and father (Crichton et al., 2014). The initiation of meiosis begins with pre-leptotene where DNA replication occurs. During leptotene, individual chromosomes condense and become visible. Zygotene follows, with synapsis of homologous chomosomes and the formation of a proteinaceous scaffold between homologous pairs named the synapotenamal complex. At pachytene this synapsis is complete, with homologues joined by the synaptonemal complex (Crichton et al., 2014). The synaptonemal complex is then lost at diplotene with homologous chromosomes being held by physical crossovers (chiasmata). Diakinesis sees homologous chromosomes shorten and condense whilst preparing to align at the metaphase plate (Ollinger et al., 2010). Cells then undergo metaphase I, anaphase I and telophase I to produce two diploid daughter nuclei. A further round of meiosis called meiosis II results in four haploid nuclei which then undergo spermiogenesis and extensive chromatin remodelling to form elongated spermatozoa that are released into the lumen of the seminiferous tubule and released upon ejaculation (O'Donnell, 2015).

1.11 Placentation

1.10.1 Placental development

The cell divisions occurring shortly after fertilisation are arguably the most fundamental decisions in development. Fertilised gametes form a totipotent zygote which is able to differentiate into all of the tissues necessary for development of the fetus (Li et al., 2013). The zygote undergoes a series of cleavage divisions leading to the early preimplantation blastocyst and the first definitive differentiation event occurs around the 32 cell stage at E3 (Li et al., 2013). This leads to the establishment of the

hypomethylated, extraembryonic trophoblast cells, the specialised precursor to part of the placenta (figure 1.12). The remaining inner cells of the blastocyst eventually form the epiblast, which gives rise to the embryo proper, extraembryonic mesoderm, and the primitive endoderm, which contributes to the yolk sac (figure 1.12)(Cockburn and Rossant, 2010; Copp, 1978). Evidently, establishing a physical connection between embryo and the mother is one of the first priorities in development and maintenance of a healthy placenta is essential throughout gestation.

The outer trophectoderm monolayer of the blastocyst gives rise to many specialised hypomethylated trophoblast structures necessary for the complex molecular interactions occurring between the embryo and the uterus during implantation (Cockburn and Rossant, 2010). At E4.5 the blastocyst hatches from the surrounding protective zona pellucida matrix to implant into the uterine wall (Cockburn and Rossant 2010). Endoreduplication of outer trophoblast cells, distal from the ICM, form highly polyploid primary trophoblast giant cells (TGCs) which facilitate a process of active invasion into the maternal uterine stroma necessary for implantation (figure 1.12)(Ain et al., 2003; Hu and Cross, 2010). These cells are the first to achieve terminal differentiation during rodent embryogenesis and are the endocrine cells of the placenta responsible for hormone production. This is necessary to regulate the maternal immune and endocrine systems, whilst promoting maternal blood flow to the implantation site (Hu and Cross, 2010; Simmons et al., 2007). The polar trophectoderm cells nearby the ICM continue to proliferate forming both the extraembryonic ectoderm and the implanted ectoplacental cone surrounded by secondary TGCs at E6 (figure 1.12)(Rossant and Cross, 2001). The extraembryonic ectoderm then expands to form the chorion, a membrane which contacts with the allantois at the posterior end of the embryo at E8.5 to form the umbilical cord, a process called chorioallantoic fusion (Rossant and Cross, 2001). After chorioallantoic fusion feto-placental blood vessels develop inwards to the placenta from the allantois to form the fetal component of the torturously branched, complex vascular network of the placenta (figure 1.12). This area is called the labyrinth and is comprised of syncytial trophoblast cells called the syncytiotrophoblast, two layers of which line fetal blood vessels to facilitate feto-maternal nutrient and gas exchange (Rossant and Cross 2001). Overall there are five TGC subtypes in the mature placenta, Blood vessels are in apposition to sinusoidal-

TGCs formed around E10.5 (Simmons et al., 2007). In mid gestation maternal blood canal associated TGCs are found in the canal bringing maternal blood to the base of the labyrinth and spiral artery TGCs are associated with remodelling uterine spiral arteries to regulate blood flow to the placenta (Hu and Cross 2010). Additionally, channel-TGCs lining the channels that drain the deoxygenated blood from the labyrinth have been recently identified (Rai and Cross, 2014). Following implantation, parietal TGCs border the ectoplacental cone and invade the uterus via remodelling the uterine extracellular matrix and phagocytosis of uterine cells (Cross et al., 1994). The different TGC subtypes are defined based on localisation, cell morphologies and differential gene expression analyses (Simmons et al., 2007). In the mid-gestation placenta, secondary P-TGCs border the decidua and the junctional zone which is comprised of two distinct trophoblast populations; spongiotrophoblast cells and glycogen trophoblast cells (GlyT)(figure 1.12).

As with TGCs, the spongiotrophoblast is responsible for the secretion of hormones, angiogenic promoting factors, vascular endothelial growth factor and tissue remodelling factors (Rossant and Cross, 2001). GlyT cells harbour stores of glycogen thought to provide energy for placental processes (Coan et al., 2006). Spongiotrophoblast and glyT cells have distinct behaviours with spongiotrophoblast cells being nonmigratory and having a modest four-fold increase in number from E12.5 to E16.5 (Coan et al., 2006). Conversely, glyT cells with extensive glycogen stores are barely detectable at E12.5 but have an 80 fold increase in number to E14.5 and invade through the P-TGC barrier into the maternal decidua (Coan et al., 2006). From E16.5 to E18.5 there is a 50% drop in glyT numbers across the whole placenta which is thought to contribute to the decreased junctional zone volume at E18.5, possibly as a result of continued use of glycogen stores, alongside the migration of glyT cells into the maternal decidua (Coan et al., 2004, 2006). This is combined with continued expansion of the labyrinth (Coan et al., 2004, 2006). Maternal blood passes through the spongiotrophoblast via arterial sinuses towards the labyrinth where it bathes the fetal trophoblast lined blood sinusoids to facilitate nutrient and gas exchange.

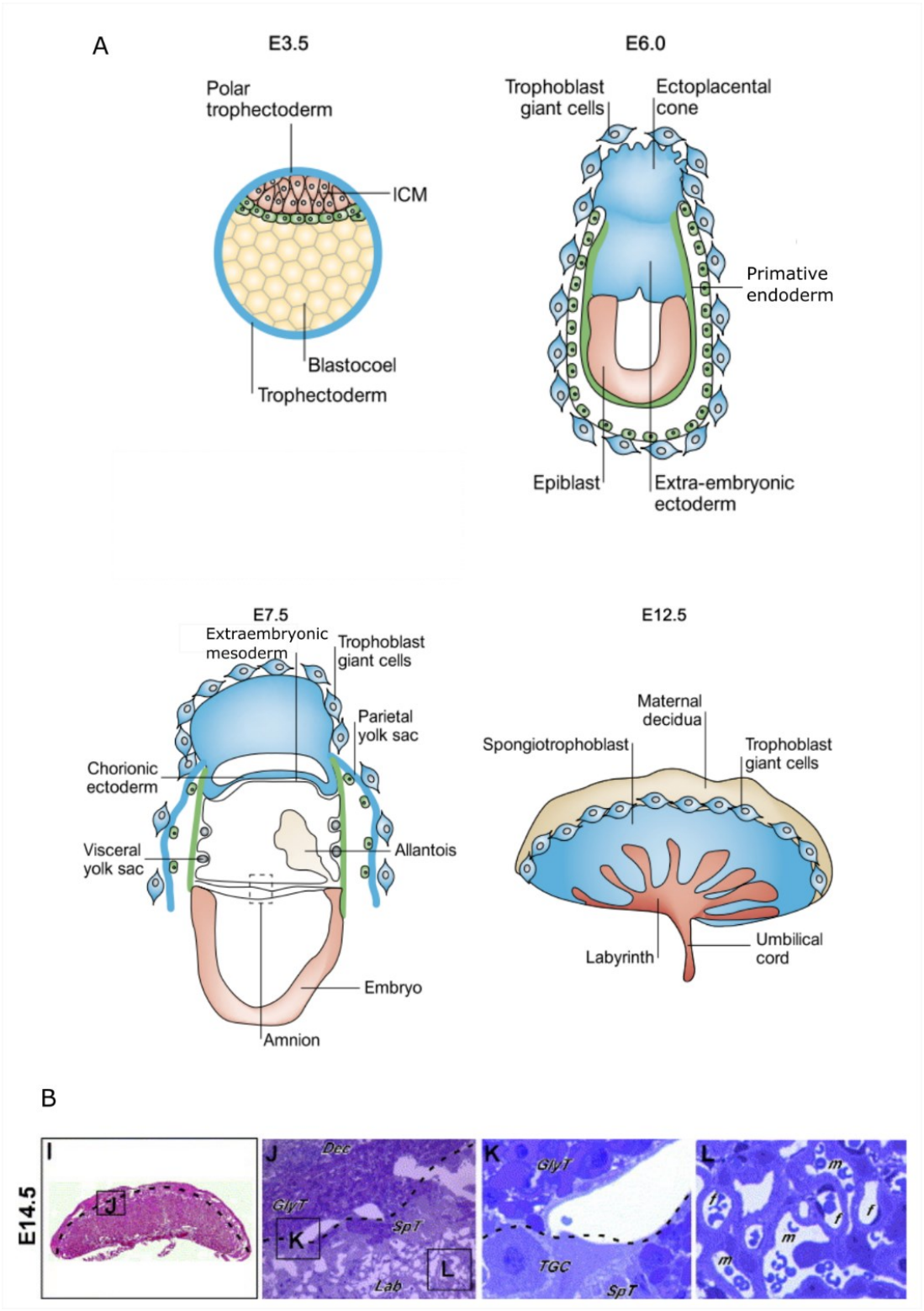


Figure 1.12: Murine placental development from the blastocyst to the placenta. (A) Diagram demonstrating the development of the murine placenta from the pre-implantation blastocyst to the ectoplacental cone and the placenta with all three established layers. Blue shading denotes the hypomethylated, trophoblast derived tissues whereas orange indicates normally methylated embryonic derived tissue

(Figure modified from Rossant and Cross 2001). (B) Histological haematoxylin and eosin (H+E) stained PFA sections of E14.5 placenta in I and semi-thin plastic, toluidine blue stained insets (J-L). The dashed line represents the boundary between the maternal decidua and the junctional zone. Dec, decidua; GlyT, glycogen trophoblast; SpT, spongiotrophoblast; Lab, labyrinth; TGC, trophoblast giant cell; m, maternal blood sinusoid space; f, fetal capillary space. (Figure taken from Simmons and Cross, 2005)

1.10.2 Differences between mouse and human placentas

It is important to note the similarities and differences between the mouse and human placenta when extrapolating findings from mice to humans. Both species have a haemochorial placenta, whereby the trophoblast layer is in direct contact with maternal blood (Schmidt et al., 2015). This is distinct from other species, such as dogs which possess endotheliochorial placentas, where the maternal blood vessels are enclosed in fetal epithelium, and pigs, which have epitheliochorial placentas where the chorion is next to the uterine lining with no uterine invasion (Schmidt et al., 2015). Both mouse and human placentas develop three distinct regions called the labyrinth, junctional zone and maternal decidua in mice which correlate to the fetal placenta, basal plate and maternal decidua in humans. The decidua of both species is largely analogous, based on their spatial location and the extent of trophoblast invasion into the maternal uterus (Georgiades et al., 2002). In contrast, the structure of the labyrinth and fetal placenta are relatively distinct between mice and humans. The chorionic projections within the murine labyrinth form a maze like, interconnected network, whereas the human fetal placenta has villous projections from the chorion, with extensive branching. In both, syncytiotrophoblasts are bathed in maternal and fetal blood allowing nutrient, gas and waste exchange, however this is performed by a monolayer of multinuclear syncytiotrophoblast in humans and two layers of multinuclear syncytiotrophoblast and a mononuclear layer in mice (Georgiades et al., 2002). In both species the junctional zone and basal plate is devoid of fetal blood and is thought to have an endocrine function, however the precise roles for this layer are unclear. Each is comprised of two types of cytotrophoblasts, one resembling glycogen cells and the other, endovascular trophoblasts, proposed to be analogous to mouse spongiotrophoblasts (Georgiades et al., 2002). The arrangement of this layer is distinct

from murine placenta as is loosely arranged proximal to column cytotrophoblasts associated with the villus. Despite the structural and cell composition differences between mice and human placentas, studies of the analogous roles of the labyrinth and related cell types throughout the placenta have led to many discoveries applicable to both normal and abnormal human placentation, leading to an understanding of the molecular basis of disorders such as intrauterine growth restriction (Rossant and Cross 2001).

1.10.3 Defects in placentation

Targeted genetic mutations in the mouse genome have provided insights into the role of different genes, structures, trophoblast cell types and essential signalling processes throughout development of the placenta. Interestingly, imprinted genes appear to play an important role in controlling placental and fetal growth (Tunster et al., 2013). Imprinted genes can be described as genes exhibiting monoallelic expression in a parent of origin dependent manner, with silencing of one parental allele brought about by DNA methylation (Reik and Walter, 2001). The evolution of genomic imprinting has been proposed to be due to "parental conflict", whereby asymmetry between parental contributions to the fetus during development is regulated by imprinted genes (Haig and Graham, 1991). This theory arose due to studies of the imprinted gene *Igf2* and its receptor *Igr2r*. *Igr2* is expressed from the paternal allele and imprinted on the maternal allele, with the converse being true for *Igf2r* (Barlow et al., 1991; DeChiara et al., 1991). Offspring receiving a disrupted *Igf2* allele from their father are smaller than if it is inherited from the mother, indicating the paternally expressed *Igf2* is responsible for enhanced growth of the embryo (DeChiara et al., 1991). By promoting embryonic growth the fathers genes gain fitness through the greater success of larger offspring, at the expense of the mother (Haig, 2000). Deletion of numerous imprinted genes can lead to a variety of placental defects in different trophoblast cell subtypes or placental layers, resulting in intrauterine growth restriction (IUGR) of mouse embryos or embryonic lethality (Tunster et al., 2016).

IUGR is defined as the failure of the fetus to reach its genetic growth potential and generally applies to babies born in the lowest 10th percentile of birthweights (Monk and Moore, 2004). IUGR in humans is associated with neonatal problems such as perinatal asphyxia, hypoglycaemia and premature death whereas problems into adulthood extend to growth retardation, neurodevelopmental defects and enhanced susceptibility to adult onset disease such as type II diabetes, ischemic heart disease and neurological disorders such as depression (Sharma et al., 2016). Mice have provided a good system to model human IUGR, whereby aberrant expression of evolutionary conserved genes in both species contributes to growth restriction. For example, deletion of the paternally expressed *Igf2* gene throughout placenta and embryo leads to reduction of the labyrinth layer and glycogen cell number, however deletion of the placental specific p0 *Igf2* promoter also led to a smaller placenta with compensatory increases in maternal amino acid transport and late onset IUGR (Constância et al., 2002; Sandovici et al., 2012).

Furthermore, when the maternally expressed, paternally imprinted gene, *Phlda2* is experimentally overexpressed in mice, it drives a reduction of the spongiotrophoblast compartment, with reduced amounts of placental glycogen and causes IUGR (Tunster et al., 2014, 2016). A potential role for IUGR in humans was suggested for the conserved human protein, *PHLDA2*. *PHLDA2* was found to be significantly elevated in expression in IUGR placentas with no alteration in DNA methylation at the locus (McMinn et al., 2006). Consequently, different studies in humans have found differing associations with *PHLDA2* expression and IUGR, such as overt differences in fetal weight, more subtle differences in placental weight, to milder associations with reduced fetal femur growth velocity and reduced infant bone mineral content (Apostolidou et al., 2007; Lewis et al., 2012). Regardless, the experiments on mice indicate the important role of imprinted genes in the regulation of placental endocrine function, and embryonic growth in mammals.

Interestingly, loss of imprinted genes can also result in an overgrowth of the placenta as observed for loss of function of *Phlda2,* which results in increased placental glycogen and increased fetal weight (Tunster et al., 2016). Additionally, one paternally imprinted gene encoding the cyclin dependent kinase inhibitor *Cdkn1c* negatively regulates cell proliferation (Takahashi and Nakayama, 2000). This was shown by

genetic knockout of *Cdkn1c* in C57BL/6 mice which resulted in a doubling of the spongiotrophoblast and labyrinth cell numbers, despite no alterations in TGC or glycogen cell number at E17.5 (Takahashi et al., 2000). Most *Cdkn1c* embryos died after birth, however those that survived were severely growth restricted (Takahashi and Nakayama, 2000). Interestingly, more recent analysis of *Cdkn1c*$^{-/-}$ placentas has been performed on 129S2/SvHsd background and again indicated large overgrowth of the placenta, accompanied by severe disruption of the labyrinth (Tunster et al., 2011). Sinusoidal trophoblast giant cell (S-TGC) number was reduced, along with impaired vascularisation and collagen deposits in the labyrinth and a reduction in spongiotrophoblast lineage (Tunster et al., 2011). *Cdkn1c* pups were born a similar weight to control animals despite having a growth advantage during late gestation which is not maintained (Tunster et al., 2011). This highlights the diverse defects which can lead to embryonic growth perturbations in mice and indicates strain specific defects arising in different genetic backgrounds.

Another example of placental overgrowth is observed upon disruption of the X-imprinted, homeobox transcription factor *Esx1* (Li and Behringer, 1998). When the *Esx1* null mutation is inherited by heterozygous females from their mother, this leads to a hypertrophic labyrinth compartment with a reduced vascular density (Li and Behringer, 1998). The overgrowth of labyrinth may be to compensate for the reduced blood flow and nutrient exchange due to the poor vascularisation of the labyrinth, however hemizygous female pups inheriting the maternal null allele are still born with IUGR (Li and Behringer, 1998).

The severity of placental defects upon disruptions in imprinted genes can have profound effects on embryonic development. For example, the paternally imprinted *Mash2 (Ascl2)* encodes a basic-helix-loop-helix transcription factor and null embryos die at E10.5 due to absence of the spongiotrophoblast layer, with little effect on secondary TGCs (Guillemot et al., 1994; Tanaka et al., 1997). Aggregation of *Mash2*$^{-/-}$ morulas with wild type tetraploid embryos, which contribute solely to extraembryonic tissues (Guillemot et al., 1994), results in viable embryos suggesting that the essential function of *Mash2* is in the placenta. Further studies on the chimeric mice indicated that *Mash2*$^{-/-}$ cells are unable to contribute to the spongiotrophoblast layer, however are abundant in the labyrinth of chimeric placenta (Tanaka et al., 1997).

This indicates a cell autonomous role of *Mash2* in the spongiotrophoblast layer in development of the placenta and embryo.

It is important to note that placental defects not only arise from disruptions of imprinted gene expression but also non-imprinted genes, many of which exhibit roles in different cell lineages or compartments of the placenta with differing consequences for embryonic development. *Hand1* is a non-imprinted basic-helix-loop-helix transcription factor and mice null for *Hand1* arrest at E7.5 of gestation due to defects in trophoblast giant cell differentiation (Riley et al., 1998). A similar tetraploid aggregation assay rescued this mutant, indicating an initial placental defect until E10.5 when embryos died with defects in ventricular myocardial differentiation (Riley et al., 1998).

Labyrinth development is another key process in placental development which when disrupted in a variety of mouse mutants of non-imprinted genes results in embryonic lethality, such as in mice null for *Gcm1* and *Gjb2*. *Gcm1* is a transcription factor expressed in the chorionic plate from E8.0 (Anson-Cartwright et al., 2000). Chorioallantoic branching fails to initiate in *Gcm1*$^{-/-}$ mice leading to absence of the labyrinth along with failure to form syncytiotrophoblast cells (Anson-Cartwright et al., 2000). Together, experiments on mice indicate the important role of imprinted and non-imprinted genes in the regulation of placental development, function and embryonic growth in mammals.

1.12 *Tex19.1*$^{-/-}$ phenotypes

TEX19.1 plays highly important roles in the development of both male germ cells and the placenta with de-repression of retrotransposons present in both *Tex19.1*$^{-/-}$ contexts. In the male gonad *Tex19.1* is expressed in the cytoplasm of spermatogonia and early spermatocytes, but not in late pachytene spermatocytes (Ollinger et al., 2008). *Tex19.1* is essential for spermatogenesis as *Tex19.1*$^{-/-}$ spermatocytes initiate meiotic recombination and homologous chromosomes assemble the axial elements of the synaptonemal complex but fail to fully synapse in around half of mutant nuclei (Ollinger et al., 2008). Of those *Tex19.1*$^{-/-}$ spermatocytes progressing through meiosis,

two thirds have univalent chromosomes at metaphase I, instead of pairing as bivalents (Ollinger et al., 2008). The spermatocyte asynapsis phenotype is similar with other germline genome defence genes such as *Miwi2, Mili* and *Mov10L* which correlates with de-repression of retrotransposon transcripts. *Tex19.1*$^{-/-}$ spermatocytes show increases of RNA from the *MMERVK10C* LTR retrotransposon as indicated by *in situ* hybridisation (Ollinger et al., 2008). The TEX19.1 mediated *MMERVK10C* repression is likely distinct from piRNA mediated mechanisms (figure 1.10) as LINE-1 and IAP element RNAs do not accumulate in *Tex19.1*$^{-/-}$ testes like in PIWI-piRNA mutants, and no differences in DNA methylation were detected at *MMERVK10C* loci in the absence of *Tex19.1* (Ollinger et al., 2008).

Recent evidence from our lab indicates that *Tex19.1* is operating to repress retrotransposons at multiple levels in their life cycle. *Tex19.1*$^{-/-}$ spermatocytes have increases in L1Orf1p protein levels determined by western blot and immunofluorescence but no change in the amount of LINE-1 RNA (MacLennan et al., submitted). This suggests TEX19.1 is post-translationally repressing LINE-1 to prevent its retrotransposition in the germline. In humans, TEX19 also regulates LINE-1 protein abundance at the post-translational level by stimulating its ubiquitination and subsequent turnover (MacLennan et al., submitted). This results in reduced retrotransposition of both mouse and human LINE-1 constructs (MacLennan et al., submitted). This post translational regulation of LINE-1 protein is in contrast to the transcriptional or post-transcriptional mechanisms leading to increases in RNA from *MMERVK10C* (Ollinger et al., 2008). The reasons for this specificity in repressing different retrotransposons at discrete time points in their lifecycle remains unclear, however it may be a consequence of the various interacting partners of TEX19.1 in the germline, such as the interaction with the E3 ubiquitin ligase, UBR2 (Yang et al., 2010).

As *Tex19.1* expression has been shown to be sensitive to DNA methylation (Hackett et al., 2012), its expression was investigated in the naturally hypomethylated tissue of the placenta (Reichmann et al., 2013). Of all the methylation sensitive germline genome defence genes identified in the previously discussed epigenetic screen (section 1.9)(Hackett et al., 2012), only *Tex19.1* was methylation sensitive and expressed in the placenta (Reichmann et al., 2013). TEX19.1 protein was restricted to the

hypomethylated trophectoderm derived cells detectable in the spongiotrophoblast layer and S-TGCs of the labyrinth. Surprisingly, TEX19.1 appeared to be partly nuclear in these cells via immunohistochemistry unlike the cytoplasmic localisation detected in spermatocytes (Ollinger et al., 2008; Reichmann et al., 2013). In the absence of *Tex19.1,* E18.5 placentas weigh significantly less than controls and have thinner junctional zones due to loss of trophectoderm derived spongiotrophoblast and glycogen trophoblast cells. S-TGCs of the labyrinth were also reduced in number and *Tex19.1*$^{-/-}$ pups exhibited IUGR. When the transcriptome of E18.5 control and *Tex19.1*$^{-/-}$ placentas was assessed by microarray and analysed for retrotransposon RNA, increased abundance of LINE-1 and VL30 RNA was detected, which was not detected for *MMERVK10C* (Reichmann et al., 2013). Therefore, *Tex19.1* not only facilitates repression at different stages in the retrotransposon life cycle in spermatogenesis, but also exhibits transcriptional or post-transcriptional repression of different retrotransposons in a tissue specific manner (Reichmann et al., 2013). It also remains unclear if chromatin marks are perturbed at these de-repressed retrotransposon loci and if *Tex19.1* is repressing MMERVK10C at a transcriptional or post-transcriptional level. Additionally, placental phenotypes have only been investigated at late stages of gestation at E18.5 and the role of interacting partners in the *Tex19.1*$^{-/-}$ defects in spermatogenesis have never been investigated. Therefore, the investigation of the mechanism of TEX19.1 mediated retrotransposon repression, the molecular etiology of the placental developmental defects and the possible contribution of TEX19.1 interacting partners in the defects in spermatogenesis of *Tex19.1*$^{-/-}$ mice are investigated in the following chapters of this thesis.

1.13 Thesis Objectives

The overall aims of this thesis are as follows:

1) To gain a greater insight into the possible mechanisms of *Tex19.1* mediated retrotransposon repression
2) To further characterise the phenotype of the *Tex19.1*$^{-/-}$ placenta, addressing the potential contribution of retrotransposon de-repression towards developmental defects.
3) To determine the contribution of the TEX19.1 interacting partners, UBR5 and UBR2 to the *Tex19.1*$^{-/-}$ spermatogenic phenotype.

Chapter 2: Materials and Methods

2.1 Mammalian Cell Culture

2.1.1 Routine cell culture and harvesting

To retrieve cells stored in liquid nitrogen, cells were quickly thawed at 37ºC in a water bath, added to 9ml of pre-warmed media and pelleted by centrifugation to remove DMSO before seeding into T25 or T75 culture flasks. For ES cell culture flasks were pre-coated with 0.2% gelatin (Sigma) in PBS which was aspirated prior to the addition of media and cells. Cell culture techniques have been described by (Hogan et al., 1994).

Feeder independent E14 ES cells (*Tex19.1*$^{+/+}$, *Tex19.1*$^{+/-}$, *Tex19.1*$^{-/-}$) were cultured at 37ºC 5% CO_2 in pre-gelatinised flasks in E14 media (GMEM, 10% fetal calf serum (FCS), 1% L-glutamine (L-glut), 1% penicillin/streptomycin (P/S),1% sodium pyruvate, 1% non-essential amino acids (NEAA), 0.1% β-mercaptoethanol and 1ml LIF-conditioned media per 500ml GMEM).

Feeder dependent J1 ES cells (*Dnmt1*$^{-/-}$*Dnmt3a*$^{-/-}$*Dnmt3b*$^{-/-}$ ES, J1 ES) were cultured on a fibroblast feeder layer in ES cell media (DMEM, 15% FCS, 1% L-glut, 1% P/S,1% sodium pyruvate, 1% NEAA, 0.1% β-mercaptoethanol and 1ml LIF per 500ml DMEM.

Hek293T cells, mouse embryonic fibroblasts (MEFs) (hypomorphic *Dnmt1*$^{n/n}$ *P53*$^{-/-}$ MEFs and *Dnmt*$^{+/+}$ *P53* MEFs) and primary embryonic fibroblast cells (for feeders) were cultured in STO media (DMEM, 10% FCS, 10% P/S, 10% L-glut). Primary embryonic fibroblast cells for use as feeders for J1 ES cells were cultured until 80% confluent and treated with 10ug/ml (final concentration) mitomycin C for 2.5 hours at 37ºC. Cells were then washed 2x with PBS, trypsinised and seeded at $2x10^5$ cells/ml and incubated at 37ºC for at least 1 hour before use as feeders for J1 ES culture.

All cell lines were passaged by 1x PBS wash then trypsinised for 5 minutes at 37ºC. To inactivate the trypsin 10 volumes of media containing FCS were added to the cells which were then spun down at 300 RCF for 5 minutes. Cells were then resuspended in fresh pre-warmed media and reseeded at $2x10^5$ cells/ml or frozen down for stocks.

Cells were frozen at $4x10^6$ cells/ml by adding 500µl of freezing media (20% DMSO, 20% FCS, 60% culture media) to 500µl cell suspension at $8x10^6$ cells/ml and placed at -70°C in a polystyrene box for at least 24 hours before storage in liquid nitrogen at -150°C.

Tex19.1$^{-/-}$ ES cells were generated previously by sequential targeting of E14 ES cells by Abigail Mann. The *Tex19.1* targeting vector was generated by inserting an IRES-GFP cassette into position chr11:121147942 (mm10 genome assembly) in the 3' untranslated region of *Tex19.1* in a bacterial artificial chromosome (BAC). A 13 kb region (chr11:121143511-121156687) containing *Tex19.1* was gap-repaired into PL253 vector, then a LoxP site was recombined upstream of the coding exon at position chr11:121146376. An Frt-flanked neomycin-resistance cassette and second LoxP site from PL451 was recombined downstream of the coding exon at chr11:121148877. Electroporation with the resulting targeting vector into E14 ES cells was followed by selection for neomycin resistance, and correct integrants identified by PCR. The *Tex19.1* coding exon in the targeted allele was removed by transfection with a Cre-expressing plasmid, and the resulting cells electroporated with the targeting vector again, selected for neomycin resistance, and correct integrants on the second *Tex19.1* allele identified by PCR. ES cells were then transiently transfected with a Flp-expressing plasmid to generate a conditional *Tex19.1*^{fl} allele. This was subsequently converted to a *Tex19.1*$^{-}$ allele by transient transfection with a Cre-expressing plasmid to remove the *Tex19.1* coding exon (MacLennan et al., 2017 Submitted).

2.1.2 2i culture

Where feeder independent *Tex19.1*$^{+/-}$ ES and *Tex19.1*$^{-/-}$ ES cells were cultured in 2i conditions, 500ml serum free ES medium (SFES) (45% Neurobasal media, 45% DMEM/F12, 0.5% N2-supplement, 1% B27+RA, 7.5% BSA, 1% P/S) was aliquoted into 10x 50ml falcons and stored at 4°C protected from light. When required, complete 2i media was made (50ml SFES, 1µM PD0325901, 3µM CHIR99021, 1% L-glut, 0.63µl monothioglycerol (11.9M), 50µl LIF). 50µl aliquots of 1mM PD0325901 (Stemgent), 3mM CHIR99021 (Stemgent) and LIF were stored at -20°C and added

daily. Where required Vitamin C (Sigma) was added daily at a concentration of 100µg/ml daily from frozen aliquots.

When serum/LIF cultured ES cells were 70-80% confluent E14 media was removed, the cells washed 1x with PBS and compete 2i media added. The following day, cells were washed 1x with PBS and trypsinised for 5 minutes until single cell suspension achieved. Trypsin was inactivated with addition of DMEM + 10% FCS, centrifuged and washed with DMEM only to remove residual trypsin and FCS. Cells were reseeded on pre-gelatinised flasks at $4x10^5$ cells/ml in fresh complete 2i media and split and until required.

2.1.3 Cell transfection

Hek293T cells were grown to 50% confluency in a 6 well plate. 1µg plasmid DNA was added to 50µl Optimem (Invitrogen). 3µl lipofectamine (Invitrogen) was added to 50µl optimem and the aliquots combined after 5 minutes with the 100µl added to the cells and incubated overnight at 37ºC.

2.1.4 Embryoid body formation

Differentiation of ES cells into embryoid bodies (EBs) was performed by hanging drop culture. 50 drops of E14 media without LIF were spotted onto a bacterial petri dish lid with each 20µl drop containing 600 ES cells. 10mls of PBS was placed in the tissue culture dish, the lid inverted and hanging drops cultured at 37ºC for 2 days. EBs were then washed from the lid with 10ml E14 media without LIF and cultured in suspension in new bacterial petri dishes, changing half of the media every two days. At day 7 the EBs were transferred to a gelatinised 10cm tissue culture dish to allow for outgrowths and further differentiation to occur. EBs were harvested for RNA at 0, 3, 6, 9 and 12 days of culture and qRT-PCR performed.

2.1.5 Sperm counts

For sperm counts the epididymis from each testes was dissected and homogenised in 1ml 1% sodium citrate to immobilise sperm and allowed to settle for 5 minutes to allow debris to settle. Sperm in the supernatant was then counted with a haemocytometer.

2.2 RNA extraction and manipulation

2.2.1 RNA isolation and purification

RNA was isolated from cultured cells and mouse tissues using TRIzol (Invitrogen) following the manufacturer instructions. Cultured cells were washed with PBS and lysed in 1ml TRIzol. After pipetting to homogenise cells the samples were either stored at -70^{o}C until required or incubated for 5 minutes at room temperature. 200µl of chloroform was added and shaken by hand for 15 seconds then incubated for 3 minutes at room temperature and centrifuged at full speed at 4^{o}C. The aqueous phase was collected, mixed with 500µl of 100% isopropanol, incubated for 10 minutes at room temperature and centrifuged full speed for 10 minutes at 4^{o}C to isolate the RNA pellet. If low numbers of cells were used, 1µl of glycoblue was added to aid RNA precipitation and identification. Pellets were washed with 1ml 75% ethanol, vortexed and centrifuged full speed at 4^{o}C for 5 minutes. RNA pellets were air dried for 10 minutes, resuspended in 30µl of RNAse-free milliQ dH_2O, incubated in a heat block at 55^{o}C for 10 minutes and concentration and purity measured by nanodrop (thermofisher). RNA was either diluted to necessary working concentrations with dH_2O or stored at -70^{o}C.

To digest and genomic DNA contamination, RNA was treated with either RQ1 DNAse (Promega) or Turbo DNA-free kit (ambion) as per manufacturer instructions and purified by an RNeasy mini kit (QIAgen).

2.2.2 RNAseq and Microarray

20µl of DNAse treated whole placental RNA at a concentration of 100ng/µl was sent to Edinburgh Genomics for RNAseq library preparation and sequencing. RNA had A260/280 of >1.9 as determined by Nanodrop (Thermo) and RNA integrity scores >7 determined by BioAnalyser (Agilent). TruSeq stranded mRNA-seq libraries (Illumina) were made by Edinburgh Genomics and sequenced on 1 lane of a HiSeq V4 high output (Illumina) yielding around 30 million 125 base paired-end reads per sample. The same quality control criteria and concentrations of *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ ES cell RNA were sent to Edinburgh Genomics for hybridisation on an Affymetrix Mouse Gene 2.1 array.

2.2.3 cDNA synthesis

1µg DNAse treated RNA was used to produce cDNA for qPCR by the Superscript III first strand synthesis system (Invitrogen) according to the manufacturer's protocol. Random primers were used to maximise transcript coverage and to ensure no genomic DNA contamination was present, reactions were performed with and without reverse transcriptase. cDNA reactions were prepared in 8 strip PCR tubes with 1µg of DNAse treated RNA, 0.2µl 250ng random primers, 1µl 10mM dNTP mix and dH_2O to 13µl. This was incubated in a PCR machine at 65^{o}C for 5 minutes then 2 minutes on ice. 4µl 5x first strand synthesis buffer 1µl 0.1M DTT, 1µM RNAse Inhibitor, 1µl Superscript III Reverse Transcriptase were added to reactions and incubated in a DNA Engine Tetrad PCR machine at 25^{o}C for 5 minutes, 50^{o}C for 60 minutes and 70^{o}C for 15 minutes and cDNA was stored at -20^{o}C.

2.2.4 qRT-PCR

To quantify levels of RNA transcripts, cDNA was used as a template for quantitative reverse transcription PCR (qRT-PCR). cDNA was diluted 1:50 in dH_2O and the standard curve method of quantification was used. 1:10, 1:50, 1:100 and 1:1000 fold serial dilutions of cDNA were incorporated on each plate to determine primer

efficiency and relative quantity of RNA transcript. Primers from the manufacturer (Sigma) were diluted to 100µM stocks and stored at -20°C. Primer sequences are given in table 2.1. These stocks were diluted to 1µM working concentration as required. For each 20µl reaction (96 well plate), 2.5µl of both forward and reverse primers, 10µl SYBR select master mix (LifeTech) and 5µl cDNA (1:50) was added per well. No template control reactions were set up with dH_2O and no reverse transcriptase reactions were set up from cDNA reactions without SuperScript III. qRT-PCR was performed on the LightCycler 480 (Roche) with pre-incubation of 50°C for 2 minutes and 95°C for 2 minutes. Amplification at 95°C for 15 seconds and 60°C for 30 seconds for 50 cycles was followed by melt curve generation to check primer binding specificity at 95°C for 5 seconds, 65°C for 1 minute with 0.5°C increment for 5 minutes to 95°C and cooled to 40°C. Only primers with an R^2 value of >0.98 were used. All reactions were run in triplicate, the Cts averaged and the quantity of transcript determined via standard curve. This quantity was normalised to a reference gene such as β-actin and the expression levels relative to wild type or heterozygous cultured cells or littermate controls are shown. A two tailed t-test was used to determine if gene expression changes were statistically significant.

2.3 DNA Extraction and Manipulation

2.3.1 DNA isolation from tissue and cells

DNA for bisulphite treatment was isolated from mammalian cells and E12.5 placentas using the Wizard Genomic DNA purification kit (Promega) as manufacturers protocol. Half placentas were mascerated with razorblades and added to 600µl of ice cold nuclei lysis solution and homogenised for 10 seconds with a hand homogeniser. The lysate was incubated at 65°C for 30 minutes, 3µl of RNase solution added, mixed, incubated for 30 minutes at 37°C and allowed to cool to room temperature. 200µl protein precipitation solution was added followed by vortexing for 20 seconds and placing on ice for 5 minutes to facilitate protein precipitation. The lysed samples were centrifuged for 4 minutes at 13000g and the supernatant containing the DNA removed to a fresh 1.5ml Eppendorf containing 600µl of 100% isopropanol and inverted until visible DNA precipitation. Precipitation was followed by centrifugation for 1 minute at

13,000g at room remperature and the supernatant discarded. To the white DNA pellet 600µl of room temperature 70% EtOH was added to wash the DNA and centrifuged as previously. The supernatant was removed and DNA allowed to air dry before addition of 100µl DNA rehydration solution and incubation at 65ºC for 1 hour. DNA was stored at 4ºC.

2.3.2 Agarose gel electrophoresis

To resolve DNA samples 1 – 2% agarose gels were prepared with 1x Tris/Borate/EDTA (TBE) buffer and 0.5µg/ml ethidium bromide was added to molten agarose after microwaving and mixed and performed as in (Sambrook and Russel, 2001). 6X Orange G loadng buffer was added to DNA samples before loading specific amounts of DNA dependent on well size. 500ng of 1kb or 100bp DNA ladder (NEB) was loaded on each gel and stained DNA was visualised on a transilluminator (Biorad Universal hood II) and printed with a thermal printer.

2.3.3 Gel purification of DNA

Different sized PCR products or cut plasmids were separated from undesired DNA fragments by gel electrophoresis, visualised on a transilluminator on low power to avoid crosslinking of DNA and excised with clean razorblades. The QIAQuick Gel Extraction Kit (QIAGen) was used to isolate DNA following the manufacturer's protocol and has been described by (Sambrook and Russel, 2001).

2.3.4 Quantifying quality and quantity of nucleic acids

Nucleic acid quantification was performed by measuring the absorbance of DNA and RNA at 260nm and 280nm using a spectrophotometer (Nanodrop ND-1000, Thermo). Purity of DNA and RNA was determined by A260/280 ratios.

2.3.5 Bisulphite treatment and sequencing

500ng genomic DNA prepared from the Wizard Genomic DNA kit was bisulphite treated using the EZ DNA Methylation kit (Zymo Research). 500ng DNA was diluted to 45µl, added to 5µl M-dilution buffer and mixed by pipetting followed by incubation at 37ºC for 15 minutes. To convert unmethylated cytosines to uracil for detection as thymidine after PCR, the DNA mixture was then incubated with 100µl of freshly made CT conversion reagent for 50ºC overnight on a rotor. 10 minutes on ice was followed by adding the 150µl of converted DNA to 400µl of M-binding buffer in a zymospin column and mixing. The column was then centrifuged at full speed for 30 seconds, the flow through discarded and the column washed with 100µl M-wash buffer and centrifuged. 200µl of M-desulphonation buffer was added and left to stand at room temperature for 15 minutes followed by centrifugation. For the final washes 200µl of M-wash buffer was added and centrifuged twice and the converted DNA was eluted in 10µl M-elution buffer and stored at -20ºC. Bisulphite primers were designed to amplify LINE-1 subtypes of interest (Table 2.1) and PCR amplification was performed with the following reaction on 1µl of converted genomic DNA: 1x buffer, 1.5mM $MgCl_2$, 200µM dNTPs, 1µM forward and reverse primers, 1.5U Platinum Taq DNA polymerase (Invitrogen). Reactions were incubated in a DNA Engine Tetrad PCR machine (MJResearch) at 95ºC for 5 minutes, 35 cycles of 94ºC for 45 seconds, 56ºC for 45 seconds, 72ºC for 90 seconds followed by 1 cycle at 72ºC for 10 minutes. PCR products were purified by agarose gel electrophoresis on a 1.5% gel. The 250bp band was excised for L1A and L1MdF2 bisulphite primers and gel purified as described above. No band was visible for the untreated DNA or dH_2O in the PCR reaction. Purified products were ligated into pGemTEasy Vector (promega) for 1 hour at room temperature following manufacturer's instructions. The plasmids were transformed into subcloning efficiency DH5α *E.coli* and transformants were determined by ampicillin selection and those colonies with inserts were determined by X-gal blue/white selection as previously described. White colonies were picked and sequenced using the M13 forward sequencing primer (performed by MRC HGU technical services). Resulting sequences were imported to BioEdit, aligned to the reference sequence via ClustalW and CT conversion for the repetitive sequences was

measured using the online programme BISMA (Rohde et al., 2010), discarding clonal sequences.

2.3.6 Cell and tissue preparation for ChIP

To prepare a single cell suspension from 16dpp testes, the tunica was removed and tubules were macerated with a razorblade in 100µl in PBS on ice, resuspended in 4ml PBS and left to settle for 10 minutes. The supernatant containing the single cell suspension was centrifuged at 2000rpm for 5 minutes and the pellet resuspended in 2ml of PBS. For preparation of placenta for ChIP, whole placentas were macerated with razor blades, added to 500µl ice cold PBS and homogenised with a hand homogeniser for 20 seconds. Crosslinked ChIP was performed as previously described (Mortazavi et al., 2006).

Cells were cross-linked by adding formaldehyde to a final concentration of 1% for 10 minutes and stopped by the addition of glycine to a final concentration of 0.125M. Cells were washed once with PBS, followed by resuspension in Farnham lysis buffer (5mM 1,4-piperazine-bis-[ethanesulphonic acid] pH8, 85mM KCl, 0.5% NP-40, protease inhibitor cocktail [Roche]). The crude nuclear preparation was collected by centrifugation and resuspended in 700µl RIPA buffer (1 X PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulphate [SDS], protease inhibitor cocktail). For placenta, cell lysis buffer (1% SDS, 10mM EDTA, 50mM Tris-HCL pH8.1, protease inhibitor cocktail) and IP dilution buffer (0.1% Triton X-100, 2mM EDTA, 150mM NaCl, 20mM Tris-HCL pH8.1) was used instead of Farnham and RIPA buffers. Sonication using the MSE SoniPrep 150 Probe Sonicator for three cycles of 30 seconds on and 30 seconds off was performed for testes DNA on ice, resulting in DNA fragments of size 300 to 500bp. For placental cells 10 cycles of 30 seconds on and 30 seconds off was required to generate appropriate sized DNA fragments. For ES cells 50x 30 seconds on 30 seconds off was used in a Diagenode bioruptor. Sonicated DNA was centrifuged at 14000 rpm for 15 minutes at 4oC. 5µl H3K4me3 antibody (Millipore 07-473), H3K9me3, Kap1, SetDB1 and IgG isotype control antibodies were coupled to 20µl Dynabeads Protein A or G (Life Technologies

10002D) in 1ml PBS/BSA for 2 hours at 4oC and resuspended in 100µl of PBS/BSA. 70µl of supernatant was retained as 10% input in qPCR analysis, whilst 630µl was added to the antibody coupled bead suspension, overnight at 4oC on a rotor. The beads and associated chromatin were washed four times with LiCl wash buffer (100mM Tris-HCL Ph7.5, 500mM LiCl, 1% NP-40, 1% Sodium deoxycholate), followed by one wash with TE buffer. Chromatin was eluted from beads in 50µl elution buffer (1% SDS, 0.1M NaHCO3), vortexing and heating to 65^{o}C for one hour in a shaking incubator, followed by collection of supernatant after centrifugation. Crosslinks were reversed by incubation of 10% input and eluted IP chromatin at 65oC overnight. 2µl of RNAse A (10mg/ml) was added for 30 minutes at 37oC, followed by 2µl of proteinase K (10mg/ml) for 1 hour at 42oC. DNA purification was performed with the MinElute PCR Purification Kit (28004) and eluted in 20µl of buffer EB.

2.3.7 Quantitative PCR

Real time PCR analysis was performed on three biological replicates of *Tex19.1*$^{-/-}$ and *Tex19.1*$^{+/-}$ 16dpp testes, in triplicate on the Roche LightCycler 480, using SYBR Green Master Mix (Roche). Oligonucleotides used for ChIP-PCR can be seen in table 2.1. 10% input and IP DNA was diluted 1:10 and 1µl, along with 2.5µM of forward and reverse primers (2.5µl of 1µM stock), 10µl Sybr Green Master Mix (Roche) and 4µl dH2O were used per qPCR reaction.

2.4 Protein Extraction and Manipulation

2.4.1 Co-Immunoprecipitation

Transfected cells were harvested and washed with PBS after overnight incubation at 37^{o}C. The cell pellet was resuspended in 200µl lysis buffer (10mM Tris Ph7.5, 150mM NaCl, 0.5mM EDTA, 0.5% NP40, 1mM PMSF (Sigma), 1x protease inhibitor cocktail EDTA-free (Roche)) and placed on ice for 30 minutes. Cell lystate was spun at full

speed for 10 minutes to pellet insoluble material and the supernatant was added to a new precooled tube. 20µl of lysate was taken as 10% input material and added to 20µl 2x loading buffer. 800µl of dilution buffer (10mM Tris pH7.5, 150mM NaCl, 0.5mM EDTA, 1mM PMSF (Sigma), 1x protease inhibitor cocktail EDTA-free (Roche)) was added for 1ml of cell lysate. Meanwhile 30µl flag-agarose bead slurry was resuspended in 500 µl of ice cold dilution buffer, spun at 2700g for 2 minutes at 4ºC and repeated 2 times to equilibrate the beads. The 1ml of soluble cell lyste was then added to the equilibrated beads and rotated at room temperature for 2 hours followed by centrifugation at 2700g for 2 minutes at 4ºC. 20µl of supernatant was added to 20µl 2x loading buffer. The flag-agarose beads and the immunoprecipitated material (IP) were washed twice in 500µl dilution buffer and resuspended in 20µl 2x loading buffer. Input, supernatant and IP samples were boiled for 2-3 minutes and 12µl of input and supernatant samples and 5µl of IP sample were run on 4-12% Bis-Tris gel and standard western blot protocol followed.

2.4.2 Western blotting

Protein samples prepared in 2x Laemmli buffer were run on 4-12% NuPAGE Novex Bis-Tris protein gels in an XCell SureLock gel tank. Broad range colour protein standard protein ladder (NEB) was used to determine approximate protein weights. The gel was washed in dH_2O and transferred to an iBlot nitrocellulose membrane by iBlot Mini transfer (ThermoFisher) for 7 minutes. Transfer was confirmed by rinsing membranes with Ponceau's stain (Sigma) then washed in 1x PBST (PBS, 0.1% Tween-20) and blocked with 5% (w/v) non-fat skimmed milk powder in PBST for 30 minutes at room temperature. Primary antibodies were diluted in milk and incubated with membranes overnight at 4ºC or for 2 hours at room temperature with rotation. The membrane was washed with three 5 minute washes with PBST and then incubated for 1 hour at room temperature with horse radish peroxidase conjugated second antibody

in block. Membranes were washed a further 3 times in PBST and proteins detected using ECL western blotting substrate (Pierce). Signals were detected by exposure onto X-ray film (Kodak). Western blotting procedures were performed as described (Sambrook and Russel, 2001). Antibodies used for western blotting are shown in table 2.2.

2.4.3 *In Vivo* ubiquitination assay

Hek293T cells were seeded at $2x10^6$ cells/ml in 10cm dishes the day before standard co-transfection with 1µg pCMV-*TEX19,* pCMV-His_6-myc-ubiquitin (Ward et al., 1995), pTRE-Gm6871-HA (Castro-Diaz et al., 2014; Ecco et al., 2016) or pLW-Ubi-KAP1-HA (Trono lab) using Lipofectamine 2000 (Invitrogen). Control plasmids for this experiment were p*Pgsk3β*-HA and empty vector (C-terminal FLAG tag). Transfected cells were incubated at 37ºC 5% CO_2 for 72 hours then harvested and lysed in 6M guanidinium-HCL, 0.1 M Na_2HPO_4, 0.1 M NaH_2PO_4, 0.01 M Tris–HCl pH 8.0, 5 mM imidazole and 10 mM β-mercaptoethanol. Lysates were sonicated for 6 times 10secs on/off with a probe sonicator and rotated with dH_2O and lysis buffer washed Ni-NTA agarose beads (Qiagen) at room temperature for 4 hours. Protein bound agarose beads were washed in lysis buffer without added imidazole, followed by washes in lysis buffer II and variants. First wash in 8M urea, 0.1M Na_2HPO_4, 0.1M NaH_2PO_4, 0.01M Tris-HCl pH 8.0 and 10mM β-mercaptoethanol, second wash in lysis buffer II pH6.3 in 0.2% triton X-100, third wash in lysis buffer II pH 6.3 and a fourth wash in lysis buffer II pH 6.3 plus 0.1% triton X-100. HIS-tagged proteins were eluted from washed beads in 200mM imidazole, 0.15M Tris–HCl pH 6.7, 30% glycerol, 0.72M β-mercaptoethanol and 5% SDS then analysed by Western blotting.

2.5 Bacterial Culture and Manipulation

2.5.1 Transfection of Dh5 E.Coli cells

Subcloning efficiency DH5α *Escherichia coli (E. coli)* (Invitrogen) were thawed on ice, gently mixed, aliquoted into 50µl and incubated on ice with 1-10ng of DNA for 30 minutes. The cells were heat shocked for 20 seconds at 42ºC, placed on ice for 2 minutes and 950µl of pre-warmed LB medium added. The cells were incubated at 37ºC for 1 hour at 225rpm followed by spreading 100µl on an LB plate with the appropriate antibiotic (Ampicillin at 50µg/L or Kanamycin at 50µg/L) for transformant selection. Where blue white selection was required, plates were spread with X-gal (50µg/ml) and IPTG to 1mM and left to dry before plating. Plates were inverted and incubated overnight at 37ºC. Growth of bacterial strains was carried out as described (Sambrook and Russel, 2001)

2.5.2 Bacterial growth

Plasmid DNA was transformed into chemically competent DH5-α *E. coli* (Invitrogen) followed by growth on either L-agar plates (10g/L NaCl, 10g/L Bacto-tryptone, 5g/L Yeast extract, 15g/L Difco Agar) and incubated overnight at 37ºC. For plasmid preparation for mini-preps or sequencing transformed cells were liquid cultured in LB medium (10g/L NaCl, 10g/L Bacto-tryptone, 5g/L Yeast extract). The plates or liquid culture were supplemented with the appropriate antibiotic to select for colonies transformed with the plasmid of interest (Ampicillin at 50µg/L or Kanamycin at 50µg/L).

2.5.3 Plasmid DNA isolation

Single transfected DH5α bacterial colonies were picked from agar plates and inoculated into 5ml LB medium with the appropriate selective antibiotic. These were cultured overnight at 37ºC at 225rpm and the following day DNA was isolated from pelleted cells using the MiniPrep spin column system (QIAGEN) following manufacturer's instructions.

2.6 Immunostaining and histology

2.6.1 Chromosome Spreads

Chromosome spreads were performed as previously described (Ollinger et al., 2008). Prior to preparation of a single cell suspension, dissected testes were trimmed of excess tissue and weighed using a fine balance. The tunica was removed from testes and the seminiferous tubules were added to 200µl RPMI media and mascerated with razor blades. The macerated material was added to 3mls of RPMI media in a 15ml falcon and debris left to settle for 10 minutes at room temperature. The supernatant containing the single cell suspension was transferred to a fresh tube and centrifuged at 1000rpm for 5 minutes. Cells were resuspended in fresh RPMI media and boiled, cooled Superfrost Plus microscope slides were prepared by 5 drops of 4.5% sucrose to the centre of each slide in a humid chamber to cause cell swelling. From a height of 15cm a single drop of mixed single cell suspension was added to the sucrose prepared slides, followed by a single drop of 0.05% Triton-X-100 to permeablise cells for fixation and leave for 10 minutes. 8 drops of fix were added to each slide and incubated at room temperature for 20 minutes. 4% PFA was prepared by adding 2g paraformaldehyde (PFA) to 25ml dH_2O in an inactive fume hood and dissolving at 60^oC. The solution was adjusted to pH8 with addition of HCl (1M) or NaOH (0.25M) and made to 50ml with dH_2O. The 4% PFA was aliquoted to 5mls and stored at -20^oC until required then made to 2% by the addition of 5ml 0.04% SDS (10µl 20% SDS + 5ml dH_2O). Fixed chromosome spreads were then washed by dipping slides 5 times and 1 time in two beakers of dH_2O and left to air dry. Spreads were stored at -80^oC or were directly used for immunostaining.

2.6.2 Immunocytochemistry and immunofluorescence

IF was performed as described in (Ollinger et al., 2008). Chromosome spread slides or dewaxed tissue sections were warmed in PBS in a coplin jar for 5 minutes then placed in a humid chamber and incubated with 50µl block (100µl 1.5% BSA, 10µl 10%

Tween-20, 50µl goat serum, 840µl PBS) and covered with a plastic cover slip for 1 hour. Block was drained off and 50µl primary antibody diluted to the required concentration in block was added with a new coverslip and incubated in a humid chamber for 2 hours at room temperature or overnight at 4°C. Slides were washed 3 times for 5 minutes in PBS and incubated with secondary antibody diluted in block for 1 hour at room temperature in the dark and washed further 3 times. Slides were mounted by applying two drops of mounting media (PD) (50mg p-phenylene diamine, 10% PBS to 90% with glycerol) and a coverslip sealed with nail varnish and imaged or stored at 4°C.

2.6.3 TUNEL assay

The fragmented DNA of apoptotic cells was measured using the DeadEnd Fluorometric TUNEL system (Promega) as the manufacturer's instructions. Paraffin embedded testes or placental tissue was dewaxed with 2 xylene washes followed by rehydration in 100%, 95%, 85%, 70% and 50% concentrations of EtOH. Incubation in 0.85% NaCl was followed by a 5 minute wash in PBS. Slides were fixed with 4% formaldehyde in PBS for 15 minutes then washed 2 times with PBS for 5 minutes each. Cells were permeabilised by adding proteinase K (20µg/ml) to each slide for 10 minutes at room temperature, washed twice for 5 minutes in PBS before further fixation with 4% formaldehyde for 5 minutes. Slides were equilibrated with 100µl of equilibration buffer for 10 minutes at room temperature. A positive control slide was treated with 10units/ml of DNAse I for 10 minutes, washed 4 times with deionised water and equilibrated with 100µl of equilibration buffer. Terminal deoxynucleotidyl transferase recombinant enzyme (rTDT) incubation mixture was prepared by 45µl equilibration buffer, 5µl nucleotide mix and 1µl rTdT enzyme.

2.6.4 Preparation of digoxigenin labelled probes for *in situ hybridisation*

A 460bp fragment of the MMERVK10C endogenous retrovirus was previously cloned into pBluescript II SK+ (Stratagene) (Ollinger et al., 2008) and cleaved with HindIII (NEB) or BamHI (NEB). 1µg linearised plasmid was used in a reaction with 0.1M DTT, 50U/µl RNase inhibitor, 10x transcription buffer, 10x DIG-rNTP mix, 20U/µl RNA polymerase (T3 for sense, T7 for antisense)(all Roche) and Rnase/Dnase free dH_2O for *in vitro* transcription for DIG-labelling of sense and anti-sense RNA. The reaction was incubated at 37ºC for 2 hours and followed by the addition of 10U/µl DNAse I (Roche) at 37ºC for 15 minutes. RNA was precipitated by adding 3M sodium acetate to 100% EtOH and incubating at -20ºC for 2 hours. The RNA pellet was rinsed with chilled 70% EtOH and resuspended in RNAse/DNAse free dH_2O and RNA integrity assessed by gel electrophoresis.

2.6.5 *In situ* hybridisation

6µm sections of 4% PFA fixed, embedded testes were dewaxed through three xylene washes, rehydrated through 100%, 95% and 70% EtOH washes and incubated in 0.2M HCL. Slides were washed in dH_2O and incubated in 2µg/ml proteinase K in 1M Tris-pH8, 0.5M EDTA and dH_2O. The reaction was stopped with chilled 0.2% glycine for 10 minutes 4ºC and slides washed twice in 0.1M Triethanolamine.HCL (TEA) buffer pH8 (Alfa Aesar) for 10 minutes with the addition of acetic anhydride at the second wash. A 15 minute wash in 5X SSC (Invitrogen) was followed by a 50ºC 2 hour incubation in a humidified chamber with prehybridisation buffer of formamide, 3X SSC, 1X Denhardts (Fisher) and dH_2O. Hybridisation buffer was comprised of formamide, 3X SSC, 1X Denhardts, 50% dextran suphate (Alfa Aesar), 100ng/ml yeast RNA (Ambion) and dH_2O, with 100µl of herring sperm DNA being denatured at 100ºC for 10 minutes and cooled on ice for 10 minutes and added to the hybridisation buffer. 5µl of DIG-labelled RNA sense and antisense probes (~100ng) were added to 200µl of hybridisation mix and denatured at 65ºC for 5 minutes. 200µl was added to each slide, covered with a plastic coverslip and incubated in a sealed humidified chamber at 50ºC annealing temperature overnight.

Slides were washed twice in 2X SSC for 15 minutes and any non-hybridised RNA was degraded by washing slides with 20µg/ml RNAseA and 2X SSC in dH_2O at 37°C for 30 minutes. Slides were washed in a graded series of 2X, 1X and 0.1X SSC solutions for 15 minutes each at 50°C, then washed in a maelic acid buffer solution of 0.1M maelic acid (Sigma) and 0.15M NaCl at pH 7.5 in dH_2O. Roche blocking reagent was diluted to 1X in maelic acid buffer pH7.5 with sheep serum and slides were blocked for 30 minutes at room temperature. To detect DIG labelled hybridised RNA, slides were incubated with 1:1000 alkaline phosphatase anti-DIG antibody (Roche 11093274910) diluted in 1X blocking solution for 1 hour at room temperature. Slides were washed twice for 15 minutes and equilibrated in 100mM Tris pH9.5, 100mM NaCl, 50mM $MgCl_2$, 0.1% tween, 0.5mg/ml levamisole (Sigma) and dH_2O for 5 minutes at room temperature then the conjugated DIG alkaline phosphatase antibody was detected by developing the signal in BCIP/NBT substrate solution (Vector Labs). Development was monitored and stopped by washing in tap water after an overnight incubation when signal was detectable. Slides were washed a final time in dH_2O and mounted with 90% glycerol in PBS and coverslips sealed with varnish. Stained sections were imaged by brightfield microscopy with phase contrast using Micromanager image capture software and ImageJ for visualisation.

2.6.6 Histology

Dissected tissue was either fixed with Bouins solution or 4% paraformaldehyde (PFA) in PBS. For Bouins and PFA fixation, whole testis and placenta were halved an fixed overnight at 4°C, followed by two 15 minute washes in 70% EtOH at room temperature. Tissue was stored in 70% EtOH until embedding with the Tissue-Tek VIP vacuum infiltration processor (Sakura) and hand embedding into wax moulds using the Tissue-Tek Embedding Console System (Sakura). For cryosections tissue was fixed in 4% PFA for 30 minutes, followed by 2 15 minute PBS washes and 30 minutes each in 5% sucrose and 10% sucrose followed by overnight in 20% sucrose. Ooptimal cutting temperature compound (OCT) and fixed tissue was placed in a plastic mould and frozen on dry ice.

6μm thick sections were cut using a microtome (Thermo), floated in a 47°C water bath and attached to superfrost non-static slides (Thermo). 12μm thick sections of OCT embedded cryofrozen tissue were cut using the Cryostat (Leica Biosystems) and attached to cold superfrost slides.

Haematoxylin and eosin staining of fixed tissue was carried out by dewaxing in 3 changes of xylene for 5 minutes each, followed by 3 changes of 100% EtOH and rehydrated in a series of 90%, 70%, 50% and 30% EtOH and then H_2O as described (Puchtler et al., 1986). Slides were stained in haematoxylin for 4 minutes, washed in tap water and differentiated in acid/alcohol (1% HCl in 70% EtOH) for a few seconds followed by washing in tap water. Slides were next immersed in saturated lithium carbonate solution for 10 seconds, washed in tap water and stained in eosin for 2 minues and washed again. Dehydration of slides was by rinsing in 100% EtOH for a few seconds, and a further 3 washes with 100% EtOH for 2 minutes each and cleared with 3 changes of xylene for 5 minutes each and mounted DPX mountant (Sigma).

For periodic acid schiffs (PAS) staining sections were deparaffinised as above, oxidised in 0.5% periodic acid solution for 5 minutes, rinsed in tap water and placed in Schiff reagent for 15 minutes. Slides were counterstained in Mayers haematoxylin for 1 minute, washed in tap water for 5 minutes and dehydrated and mounted as previously.

2.6.7 Measurement of placental layers

The area of placental layers was measured using at least three PAS stained selected midline sections per individual placentas with 20μm between each section. Sections were digitally imaged using the Olympus Dotslide system and areas calculated using the Dotslide software. Measurement of cell type abundance was performed on three rectangular segments spanning the width of the placenta on PAS stained sections. The number of spongiotrophoblast, glycogen trophoblast and sinusoidal trophoblast giant cells were counted and the area of each segment calculated using the Dotslide software.

2.6.8 Senescence assay

The senescence associated β galactosidase assay was performed using the Abcam Senescence Detection Kit (Abcam) to manufacturer's recommendations. 12µm frozen tissue sections were fixed with fixative solution (2% formaldehyde and 0.2% glutaraldehyde) for 3 minutes at room temperature and washed 3 times with PBS. 25µl 20mg/ml X-gal in N-N-dimethylformamide (DMF) was added to 470µl of Staining Solution and 5µl Staining Supplement at pH5.5 and slides were incubated overnight at 37^{o}C, then counterstained with eosin and imaged with a brightfield microscope using Micromanager image capture software.

2.7 Animals

2.7.1 Animal welfare and breeding strategy

Animals were housed and experiments were conducted according to UK Home Office regulations and local guidelines for animal welfare.

Tex19.1$^{-/-}$ mice used in this study were from colonies backcrossed to C57BL/6 for multiple generations which originated from the knockout line generated by Ollinger et al., 2008 in a 129/Ola x CD1 mixed genetic background. *Ubr5*^{flox} mice were from colonies generated on a C57/BL/6 background by Mark Ditzel and Bob Hill (ECRC and MRC HGU) (Kinsella et al., 2016). Stra8-Cre^{Tg} mice were obtained from the Jackson Laboratories and the line was maintained in house. The breeding strategy employed to produce the male specific conditional *Ubr5* knockout genotype *Stra8-Cre*^{tg} *Ubr5*$^{flox/\Delta}$ *(Ubr5*^{CKO}) at an expected frequency of 1:16 was used as it produced stable litter sizes of an average of 10 pups per mating (section 5.2.3 Table 5.2). These matings were of male *Stra8-Cre*^{tg} *Ubr5*$^{flox/wt}$ mice crossed with heterozygous female *Stra8-cre*$^{-}$ *Ubr5*$^{flox/wt}$ mice. Far fewer pups were born when the more optimal cross of male *Stra8-Cre*^{tg} *Ubr5*$^{flox/wt}$ and female *Stra8-Cre*^{tg} *Ubr5*$^{flox/\Delta}$ was used and very few *Ubr5*^{CKO} offspring were obtained. The reason for this is unclear. Male *Stra8-Cre*^{tg}

Ubr5$^{flox/flox}$ mice had smaller testes and a reduced sperm count and were therefore not used in the breeding strategy.

2.7.2 Embryo and placenta dissection

Females found plugged after mating were scored as positive for insemination and embryos were classed as E0.5 the morning the plug was observed. Pregnant dams were culled and recorded in accordance with home office regulations and the uterine sack containing embryos was removed and retained for embryo and placental dissection on ice. Individual embryos and their placenta were removed from the uterus and separated via cutting the umbilical artery, ensuring the yolk sac containing the embryo remained intact. Excess tissue was trimmed from the placenta and the embryo after removal from the yolk sac and weighed on a fine balance.

To dissect the placenta into its three constitutive layers of the decidua, junctional zone and labyrinth a Leica dissecting microscope was used. The placenta was placed with the dark red labyrinth facing upwards. The maternal decidua was removed from the placenta by inserting fine forecepts along the boundary between the pale decidua and darker junctional zone. This was repeated around the whole of the placenta getting progressively deeper into the tissue. Once sufficiently separated, the decidua was peeled from the darker junctional zone and labyrinth. This was repeated to separate the junctional zone from the labyrinth taking care to avoid contamination from the neighbouring layer.

2.7.3 Genotyping PCR

DNA was isolated from ear clips or embryo tail tips by using DNAreleasy (Anachem) and heating to 75^{o}C for 5 minutes and 96^{o}C for 2 minutes. Generally genotyping PCRs contained 0.5 units (U) Platinum Taq polymerase (Invitrogen), 0.5mM dNTPs (Invitrogen), 1X PCR buffer, 2.5mM Magnesium Chloride, 1mM each primer (Sigma) and 1µl DNA from the DNAreleasy reaction. Sequences of genotyping primers are

given in table 2.1. Typical PCR conditions were 95ºC for 5 minutes, 30 cycles of 95ºC for 30 secs, 65ºC for 30 secs, 72ºC for 1 minute and followed by 72ºC for 5 minutes on a DNA Engine Tetrad PCR machine (MJ Research). The 65ºC annealing temperature was optimal for *Tex19.1* genotyping but was altered to 58ºC for *Ubr5* primers, 66ºC for sex primers and 67ºC for Stra8-Cre transgene genotyping. PCR has been previously described (Sambrook and Russel, 2001).

2.8 Bioinformatics

2.8.1 RNAseq data

Paired end sequencing data was provided in the format of .sanfastq files from Edinburgh Genomics and quality of data generated was assessed using FastQC. If data passed quality criteria, remaining adapter sequences were removed using TrimGalore. Command line parameters: trim_galore --fastqc --stringency 5 --length 20 –paired. Sequencing reads were aligned to the mouse ensemble mm10 reference genome downloaded from UCSC using TopHat read aligner allowing mapping to the genome with 2 mismatches (Trapnell et al., 2012). Command line parameters: tophat -g 1 -p 8 --transcriptome-index --mate-inner-dist -62 --mate-std-dev 59 --no-coverage-search --b2-sensitive –o. –g = The number of alignments to the reference for a given read. –p = Number of threads to align reads (parallel processing). –transcriptome-index = Location of the .GTF mm10 known transcript file. –mate-inner-dist = Distance between paired end reads. –b2= sensitivity of Bowtie mapper. –o = Output file. The aligned reads from TopHat in a .bam file were converted to .bed format with genomic coordinates of mapped reads. To accurately quantify the expression level of transcripts with aligned reads, counts of aligned reads per transcript were consolidated and differential expression analysis was then performed on control and *Tex19.1*$^{-/-}$ RNAseq data with EdgeR (Robinson et al., 2009) with expression levels denoted as counts per million reads and a false discovery rate adjusted *P* value indicating significance values. The R programme was used for further analysis and plotting of gene expression levels using R base graphics. Distance to nearest repetitive element was determined utilising

BEDtools closest to report only the closest gene to each repetitive element in the RepeatMasker dataset (Smit, AFA, Hubley, R & Green, P. RepeatMasker Open-4.0.2013-2015 http://www.repeatmasker.org). Analysis of repeats was performed essentially as for genic transcripts. Aligned reads from the TopHat output file were intersected with genomic coordinates from the RepeatMasker dataset to generate repeat loci with mapped transcripts. Reads mapping to multiple repeat loci were distributed between all multimapping loci of that repeat and differential expression analysis was performed with EdgeR as for genic transcripts.

2.8.2 Microarray data

Affymetrix Mouse Gene 2.1 microarray data was imported to the *affy* Bioconductor package (Gautier et al., 2004) in R. Probe level expression was background corrected using the robust multi-array average algorithm (RMA) in *affy*. Expression from probes was quantile normalised and for gene level analysis summation of expression across gene probe sets was performed. Differential expression of summated gene probe sets in control and *Tex19.1*$^{-/-}$ ES cells was assessed using *limma* in R (Smyth, 2004) and data was graphed using R. For repeat probe analysis, defined repeat aligned probes were extracted from the affymetrix dataset and differential expression was analysed with *limma* as previously described (Reichmann et al., 2012).

Publically available microarray datasets from Affymetrix Mouse Genome 430 2.0 arrays were located on the Gene Expression Omnibus database (GEO). For gene expression in hypomethylated contexts datasets were obtained for fetus (GSE9954)(Thorrez et al., 2008), placenta (GSE9954)(Thorrez et al., 2008), 2i ES (GSE42580) (Yamaji et al., 2013) and primordial germ cell like cells (GSE45941)(Schemmer et al., 2013)analysed using the *affy* and *limma* in R to generate comparative gene expression values for placenta, 2i ES and primordial germ cell compared to fetal gene expression as described. For tissue specific analysis of *Ubr5* expression testis, placenta, ovary, muscle, heart, liver, lung, kidney, spleen and brain tissue microarray data was obtained from (GSE9954, Thorrez et al., 2008) and gene expression was analysed using the online GEO2R tool.

Gene / Retrotransposon	Forward	Reverse	Purpose of primer
Ubr5 wt	GTTTCTGGCAAGGTTCAGTGC	CACACATGCTGCACAAACACATG	Genotyping
Ubr5 flox	CGCGAAGAGTTTGTCCTCAC	GCCTCGATCCTCCCTTTATC	Genotyping
Ubr5 Δ	GACAATAGCAGGCATGCTGG	GCCTCGATCCTCCCTTTATC	Genotyping
Stra8-Cre +ve ctrl	CTAGGCCACAGAATTGAAAGATCT	GTAGGTGGAAATTCTAGCATCATCC	Genotyping
Stra8-Cre transgene	GTGCAAGCTGAACAACAGGA	AGGGACACAGCATTGGAGTC	Genotyping
Tex19.1	CTTCAGGAGGTCTGATGCCCTCT	GAGTGTTGTGTGGTGGGTGTTATGG	Genotyping
Tex19.1	CTTCAGGAGGTCTGATGCCCTCT	CACCGCCTGTGCTCTAGTAGCTT	Genotyping
XY	TGGTCTGGACCCAAACGCTGTCCACA	GGCAGCAGCCATCACATAATCCAGATG	Genotyping
β *Actin*	CCTCGATGCTGACCCTCATCC	GACACTGCCCCATTCAATGTCTC	ChIP
Pol2 Promoter	GACTCCGAACTGCACTCTCT	CAGCCTTTCCCTCCCTATCC	ChIP
Intergenic	CTGTCTCCAGGTCATTGAGAAGG	GTTGATCTTGGAACCCTTGGCTCC	ChIP
MMERVK10C	TTCGCCTCTGCAATCAAGCTCTC	TCGCTCRTGCCTGAAGATGTTTC	ChIP
LINE-1 5'UTR	AATCTGTCTCCCAGGTCTGC	CCTTTCGCCATCTGGTAATC	ChIP
L1MdF2	CGTAGTGTCTGTATAACCTGTCCA	GCAGGCCTATCAGAATTACAGC	ChIP, qRT-PCR
Peg3	GCCACTGCGGCAAAACA	GGTCTTCGCAATCTAGCCATCT	ChIP
KvDMR1	GGTCTATGATGGTGCATTTTGGT	AAGCCCACCGAAGTAATCCA	ChIP
IAP LTR	GATGGTGCTGACATCCTGTG	CTGACGTTCACGGGAAAAAC	ChIP
IAPEz	GCTCCTGAAGATGTAAGCAATAAAG	CTTCCTTGCGCCAGTCCCGAG	ChIP
L1Orf1	ATCCAGGAAATCCAGGACAC	TTTGCTGGACCTTTGAGTTG	ChIP
Tex19.1	AAAATGGGCCACCCACATCTC	CCACTGGCCCTTGGACCAGAC	qRT-PCR
VL30	CAGCCTTGGCCTGAGAGTTT	CTTTCTGGGCTGAAGTCCCT	qRT-PCR
MMERGLN	CCTGGAGGATTTGACCTAGACA	CCCAGGTATACCCTTTTGTCC	qRT-PCR
L1Orf2	GGAGGGACATTTCATTCTCATC	GCTGCTCTTGTATTTGGAGCATAGA	qRT-PCR
IAPEz	GCACCCTCAAAGCCTATCTTA	TCCCTTGGTCAGTCTGGATTT	qRT-PCR
Nanog	TCCTCGCCCTTCCTCTGAA	CAGGACTTGAGAGCTTTTGTTTGG	qRT-PCR
Gapdh	CGTCCCGTAGACAAAATGGT	TTGATGGGCAACAATCTCCAC	qRT-PCR
L1G	CCTGTGCCACAAACCTCTCA	GTGCCAGCCGTTTCTGGGACG	qRT-PCR
L1A	GCTTGTGCCTGCCCCAATCCAATC	GCAAGGTAGCCTGGGGCTCG	qRT-PCR
L1Mus1	GGAGTGCTGACAGAGGCTAA	GTCTTGCTTTTCGGGTGTGT	qRT-PCR
L1Mus2	CAGAGCTTCTTGGACAGGGT	GAGCTCAGATCCACCTCCTG	qRT-PCR
Lx	CAGAGCTGACCCTGTACCAC	ACTGGGTTCTGATGATGCCA	qRT-PCR
MMERVK10C	AACTGGTCGCAGGAGCTG	GGTAAAGTCTCCGAGGGTCA	qRT-PCR
β *Actin*	GGCTGTATTCCCCTCCATCG	ACATGGCATTGTTACCAACTGG	qRT-PCR
Cdx2	GGGTGGGGGTAGCAATACTT	CCCTTCCTGATTTGTGGAGA	qRT-PCR
Ctsq	AGAACAGCTGGGGTAGACG	GCCACATGCTTTCTTGTGAA	qRT-PCR
Prl8a8	AAATTATGTGGGTGCCTGGA	TCACGCAGAATTTGTCTGTTG	qRT-PCR
Gjb3	GGGGCTCTCCATCAGACATA	GTGCCAAACCTTCTCATGGT	qRT-PCR
Tmem173	CCTCAGTTGGATGTTTGGCC	AGATCAACCGCAAGTACCCA	qRT-PCR
Ifitm2	CTCTTCTTCAACGCCTGCTG	TGGGGTGTTCTTTGTGCAAA	qRT-PCR
Ifitm3	TATGAGGTGGCTGAGATGGG	TCACCCACCATCTTCCGATC	qRT-PCR
Oasl1	CTGCTCCAGGATCATGAGCT	TTCTGCCACATTGTTGGTGG	qRT-PCR
Oasl2	CCGGCAATCTACGAGACTCT	GCTCTCTGTACCCATCTCCC	qRT-PCR
Irf7	GTGTACTGGGAGGTAGGCAG	TGCCCAAAACCCAGGTAGAT	qRT-PCR
Ifi44	ATGGAGACCTGGTTCAGCAA	ACTGTCATCCTTGGCCTTGA	qRT-PCR
Psg18 L1 Antisense	CTATGGCTCTCAGGACACGT	ATCTAACACAGCGCTCCCTT	qRT-PCR
Psg28 L1 Antisense	CTATGGCTCTCAGGACACGT	GCATCTAACACAGCACTCCC	qRT-PCR
Psg21 L1 Antisense	CCAGACACTCAGAGCCAGAA	CCAGAAGGGCAGTATTGATGC	qRT-PCR
L1A Bis	TTATTTTGATAGTAGAGTT	CRAACCAAACTCCTAACAA	BisSeq
L1MdF2 Bis	TAGGTGAGAGGGTGTATTATAGAAT	AACAAATCCCAAAAACTATCAACTT	BisSeq

Table 2.1: Primer sequences are shown for each gene or genomic loci analysed Each primer pairs experimental usage is shown in the right most column (ChIP, chromatin immunoprecipitation; qRT-PCR, quantitative reverse transcription polymerase chain reaction, BisSeq, bisulphite sequencing)

Antibody	Species	Origin reference	Concentration	Use
Anti-SYCP1	Rabbit	Abcam, ab15090-100	1:200	ICC
Anti-SYCP3	Mouse	Santa Cruz, D-1, sc-74569	1:200	ICC
Anti-UBR5	Rabbit	Bethyl, A300-573	1:250	IF
Anti-PLZF (D-9)	Mouse	Santa Cruz, sc-28319	1:100	IF
Anti-H3K9me3	Rabbit	Abcam, ab8898	5µg	ChIP
Anti-H3K4me3	Rabbit	Millipore, 07-473	5µg	ChIP
Anti-KAP1	Mouse	Abcam, ab22553	1:2000/5µg	WB/ChIP
Anti-HA	Rabbit	Santa Cruz, sc-805	1:5000	WB
Anti-C-MYC	Mouse	Sigma, M4439	1:5000	WB
Anti-H3	Rabbit	Abcam, ab1791	1:100,000	WB
Anti-β-ACTIN	Mouse	Abcam, ab8226	1:1000	WB
Anti-P-STAT1 (Phos.Y7D1)	Mouse	Abcam, ab29045	1:500	WB
Anti-GFP	Rabbit	Abcam, ab290	1:5000	WB
Anti-LAMIN B1	Rabbit	Abcam, Ab16048	1:5000	WB
Anti-Mouse Alexa Fluor 488	Goat	Invitrogen, A-11001	1:500	IF
Anti-Mouse Alexa Fluor 594	Goat	Invitrogen, A-11005	1:500	IF
Anti-Rabbit Alexa Fluor 488	Goat	Invitrogen, A-11008	1:500	IF
Anti-Rabbit Alexa Fluor 594	Goat	Invitrogen, A-11012	1:500	IF
Anti-Mouse IgG HRP conjugate	Goat	BioRad, 1706516	1:1000	WB
Anti-Rabbit IgG HRP conjugate	Goat	Cell signalling, 7074S	1:1000	WB

Table 2.2: Antibodies used in experiments. Species, catalogue number, working concentration and experiments where the antibody was used are indicated. (ICC, immunocytochemistry; IF, immunofluorescence; ChIP, chromatin immunoprecipitation; WB, western blot)

Chapter 3: Investigating the mechanism of *Tex19.1* mediated transcriptional repression of retrotransposons

3.1 Introduction

The expression of the germline genome defence gene, *Tex19.1* is restricted to pluripotent cells, germ cells and the placenta (Kuntz et al., 2008; Ollinger et al., 2008; Reichmann et al., 2013). *Tex19.1,* along with other DNA methylation-sensitive germline genome defence genes such as the piRNA pathway associated *Mili, Mov10l1* and *Asz1,* are activated upon loss of methylation to repress retrotransposons (Crichton et al., 2013; Hackett et al., 2012). Genetic deletion of *Tex19.1* results in increased RNA of the MMERVK10C elements in the pachytene stage of meiosis in the germline, whereas deletion of *Mili, Mov10l1* and *Asz1* results in de-repression of the non-LTR class of retrotransposons, LINE-1 and the IAP LTR retrotransposons (Crichton et al., 2013; Ollinger et al., 2008). Taken together, the promoter hypomethylation and activation of these germline genome defence genes appear to be a compensatory response to loss of methylation which serves to prevent any consequent activation of retrotransposons. However, it is unclear if any of these methylation sensitive genes represent a generic global response in different hypomethylated contexts.

A hypomethylated tissue distinct from the germline is the trophectoderm derived murine placenta which is the first organ to form during embryogenesis. The hypomethylation occurs in the trophectoderm derived components of this tissue and is globally hypomethylated to around 40 to 50% methylation (Popp et al., 2010). Interestingly, the only DNA methylation-sensitive germline genome defence gene to be expressed in the placenta is *Tex19.1* (Reichmann et al., 2013). The increases in MMERVK10C RNA in *Tex19.1*$^{-/-}$ testis is not detected in the placenta. Instead, LINE-1 and VL30 RNA are increased in the absence of *Tex19.1* in this tissue. This indicates that the activation of *Tex19.1* may be an important generic compensatory mechanism for retrotransposon repression in response to hypomethylation. The main aim of this chapter is to investigate the potential compensatory mechanisms working through TEX19.1 to facilitate retrotransposon repression in vitro and in vivo.

It is not clear if the retrotransposon RNA increases are transcriptional or what gives *Tex19.1* its specificity for particular retrotransposons in particular tissues. Recent evidence suggests the transcriptional co-repressor KAP1 and the histone

methyltransferases SETDB1 and SUV39H1/2 form complexes acting to repress different retrotransposons via the deposition of H3K9me3 (Goodier, 2016; Matsui et al., 2010; Rowe et al., 2010). This KAP1 dependent mechanism represents another regulatory system acting to repress retrotransposons in the early embryo. KRAB-ZFPs, the largest family of transcription factors in the mouse, can target these repressive complexes to particular retrotransposon loci. This mechanism has evolved to target different classes and even specific subfamilies within different classes of retrotransposons (Castro Diaz et al., 2014). It is not clear if the KAP1 system interacts with the TEX19.1 pathway, however they do seem to have common retrotransposon targets in different cell types (Table 3.1).

In this chapter I aim to investigate if there is a generic response to repress retrotransposons in hypomethylated contexts, and if so, to characterise the potential mechanisms facilitating repression.

Tex19.1$^{-/-}$ Retrotransposons De-repressed	*Kap1*$^{-/-}$ Retrotransposons De-repressed
MMERVK10C (Testes)	MMERVK10C (ES)
LINE-1 (Placenta)	LINE-1 (ES)
VL30 (Placenta)	VL30 (ES)
-	IAP (ES)

Table 3.1: TEX19.1 represses the same retrotransposon families as in *Kap1*$^{-/-}$ ES cells in different in vivo *Tex19.1*$^{-/-}$ tissues. Table indicating overlap of retrotransposon families de-repressed in *Tex19.1*$^{-/-}$ animals (Ollinger et al., 2008; Reichmann et al., 2012, 2013) and in *Kap1* conditional knockout ES cells (Rowe et al., 2010).

3.2 Results

3.2.1 *Tex19.1* is the only gene activated in response to DNA hypomethylation in all in vitro and in vivo contexts analysed

Previous work from the Meehan and Adams labs has identified *Tex19.1* and a number of other 'methylation-sensitive genome defence genes' as being solely repressed by DNA methylation in vitro and in vivo (see appendix table 7.2 for list of genes from Hackett et al., 2012). NIH3T3 fibroblast cells were treated with the DNA demethylation inducing drug 5aza-dC and allowed to recover their expression (Hackett et al., 2012). Those genes which did not recover were deemed to be *bona fide* DNA methylation dependent genes. A further confirmation of this sensitivity to methylation was that these genes were also expressed in fibroblasts lacking the maintenance methyl transferase *Dnmt1*. *Tex19.1* and 24 other genes were activated in response to this DNA hypomethylation in vitro (Hackett et al., 2012).

To investigate if there is a generic response to hypomethylation, common to all hypomethylated contexts in vivo, I took an unbiased, genome wide approach using publicly available microarray datasets from 2i treated 'naive' ES cells, primordial germ cell like cells (PGCLCs) and E16.5 placenta which are all hypomethylated, and compared each to whole E16.5 embryo (Schemmer et al., 2013; Thorrez et al., 2008; Yamaji et al., 2013)(GEO ascensions in methods). Data was analysed with GEO2R to generate expression values and a stringent 6-fold expression increase threshold was used followed by intersection of datasets with R (Methods).

Similar numbers of genes were expressed in the placenta and naïve ES cells, but only 16% of these overlap in expression (figure 3.1A). Fewer genes were significantly expressed in PGCLCs, however 65% of PGCLC genes were also expressed in naïve ES cells compared to 16% overlap with placenta, possibly owing to the somatic nature of this tissue (figure 3.1A). 41 genes were common to all in vivo hypomethylated contexts analysed (figure 3.1A).

To address if these 41 genes were *bona fide* responses to loss of DNA methylation I intersected these with the 25 methylation sensitive genes identified previously in

experimentally induced hypomethylation contexts (Hackett et al., 2012). Surprisingly, of around ~20,000 genes, *Tex19.1* was the only gene activated in every hypomethylated context analysed (figure 3.1B).

This suggests that if there is a generic repressive mechanism to compensate for loss of DNA methylation in silencing retrotransposons then this mechanism must be working through TEX19.1.

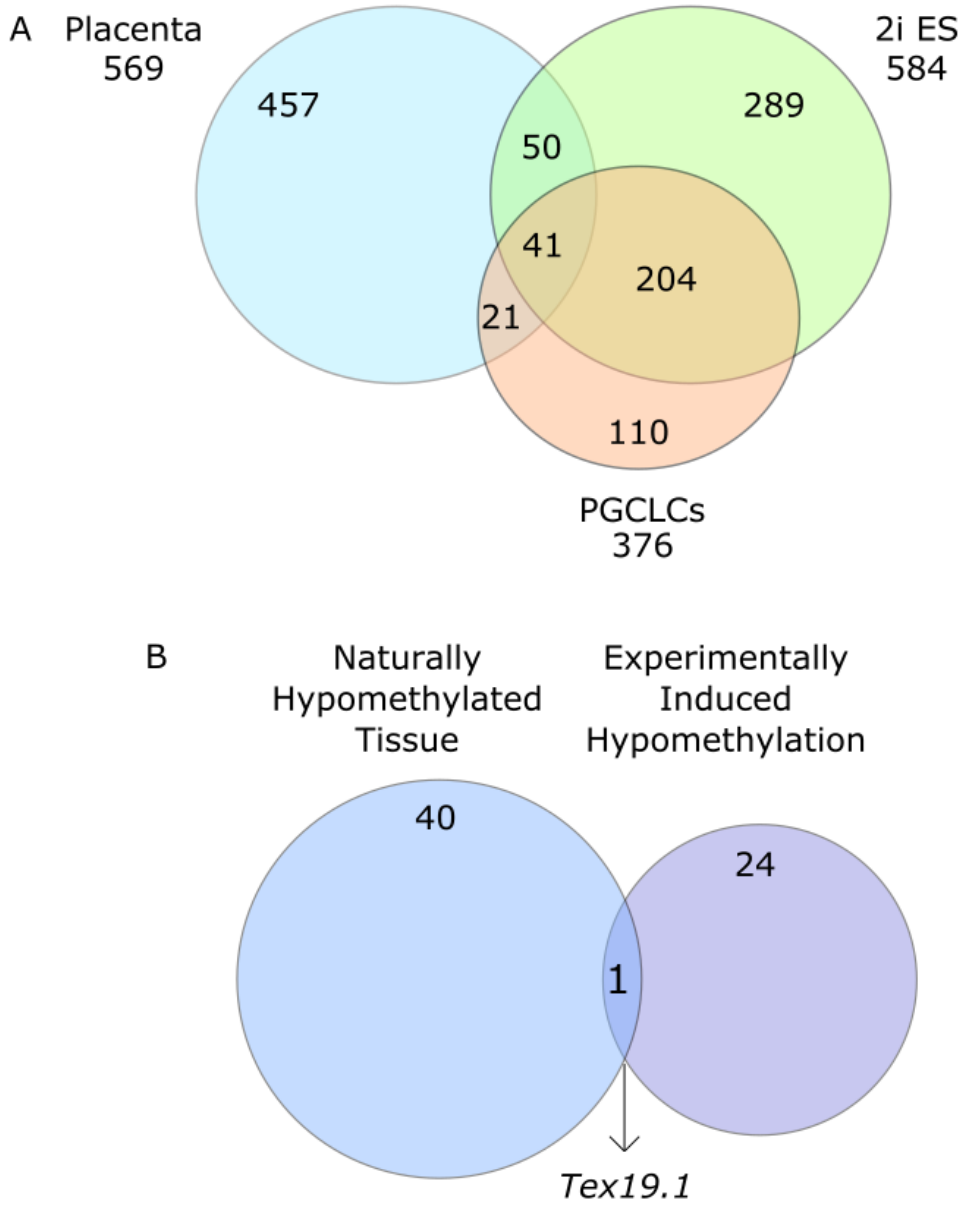


Figure 3.1: *Tex19.1* is the only gene responsive to hypomethylation in all in vivo and in vitro contexts analysed. (A) Venn diagram showing the number of differentially expressed genes in hypomethylated tissues relative to normally methylated embryo from publicly available microarray datasets. The number of genes common to different tissues are indicated in the overlaps. A stringent 6 fold differential expression threshold and FDR $P<0.05$ significance criteria were used to determine significantly different gene expression. (B) The 41 genes common to each hypomethylated dataset in A (see appendix table 7.1) were intersected with the 25 methylation sensitive genes determined in vivo (Hackett et al., 2012)(See appendix table 7.2).

3.2.2 TEX19.1 represses retrotransposons at the transcriptional level

Tex19.1$^{-/-}$ has already been shown to repress retrotransposon transcripts in testes and placenta as *Tex19.1*$^{-/-}$ tissue has increased MMERVK10C RNA and LINE-1 RNA respectively (Ollinger et al., 2008; Reichmann et al., 2012). However, the mechanism by which these retrotransposon RNAs increase in *Tex19.1*$^{-/-}$ tissues is not known. These increases in RNA could be due to an increase in RNA stability or de-repression of these retrotransposons at the transcriptional level. H3K4me3 is highly enriched at active promoters and can be used as a proxy for increased transcription (Heintzman et al., 2007). Therefore, to determine if TEX19.1 is functioning in the transcriptional repression of both MMERVK10C and LINE-1 in these different tissues I performed chromatin immunoprecipitation polymerase chain reaction (ChIP-PCR) for H3K4me3 on E12.5 placentas and P16 testes. Prior to these functional ChIP experiments on different tissues, the ChIP protocol required considerable optimisation. This was due to the limited cell numbers available from placenta and testes, the difficulty of creating a single cell suspension, especially from placental tissue, alongside differing formaldehyde fixation lengths and sonication lengths required. The optimisation of these ChIP experiments is discussed in appendix 7.1 (figure 7.1)

In the testes the *Pol2* promoter region, an active promoter, is shown as a positive control, and an intergenic region as a negative control (figure 3.2A). The positive control promoter region showed strong enrichment for H3K4me3 in the ChIP assay, whereas the negative control intergenic region did not. In addition, H3K4me3 was detected at MMERVK10C and LINE-1 5'UTR indicating active transcription of these retrotransposons in control and *Tex19.1*$^{-/-}$ testis. An increase in the abundance of H3K4me3 was detected at MMERVK10C promoters in *Tex19.1*$^{-/-}$ testes compared to control animals ($P<0.05$, Students T-test)(figure 3.2A). This increase was not observed for LINE-1 5'UTR. The H3K4me3 profiles correlate with previous reports of increased abundance of MMERVK10C RNA but not LINE-1 RNA in *Tex19.1*$^{-/-}$ testis (Ollinger et al., 2008; Reichmann et al., 2012). Therefore, TEX19.1 is responsible for transcriptional repression of MMERVK10C in the testes.

Similarly, in placentas from control mice, there is enrichment for H3K4me3 at MMERVK10C and LINE-1 indicating active transcription of these retrotransposons in this tissue in normal circumstances (Reichmann et al., 2013). This enrichment is significantly higher at LINE1 5' UTRs but not at MMERVK10C loci in *Tex19.1*$^{-/-}$ placentas compared to control placentas (figure 3.2B). As with *Tex19.1*$^{-/-}$ testis, these H3K4me3 profiles correlate with the increased abundance in LINE-1 RNA but not MMERVK10C RNA in *Tex19.1*$^{-/-}$ placentas. This indicates that TEX19.1 is responsible for transcriptional repression of LINE-1 in the placenta. Taken together, these data are consistent with TEX19.1 functioning to transcriptionally repress different retrotransposons in different tissues.

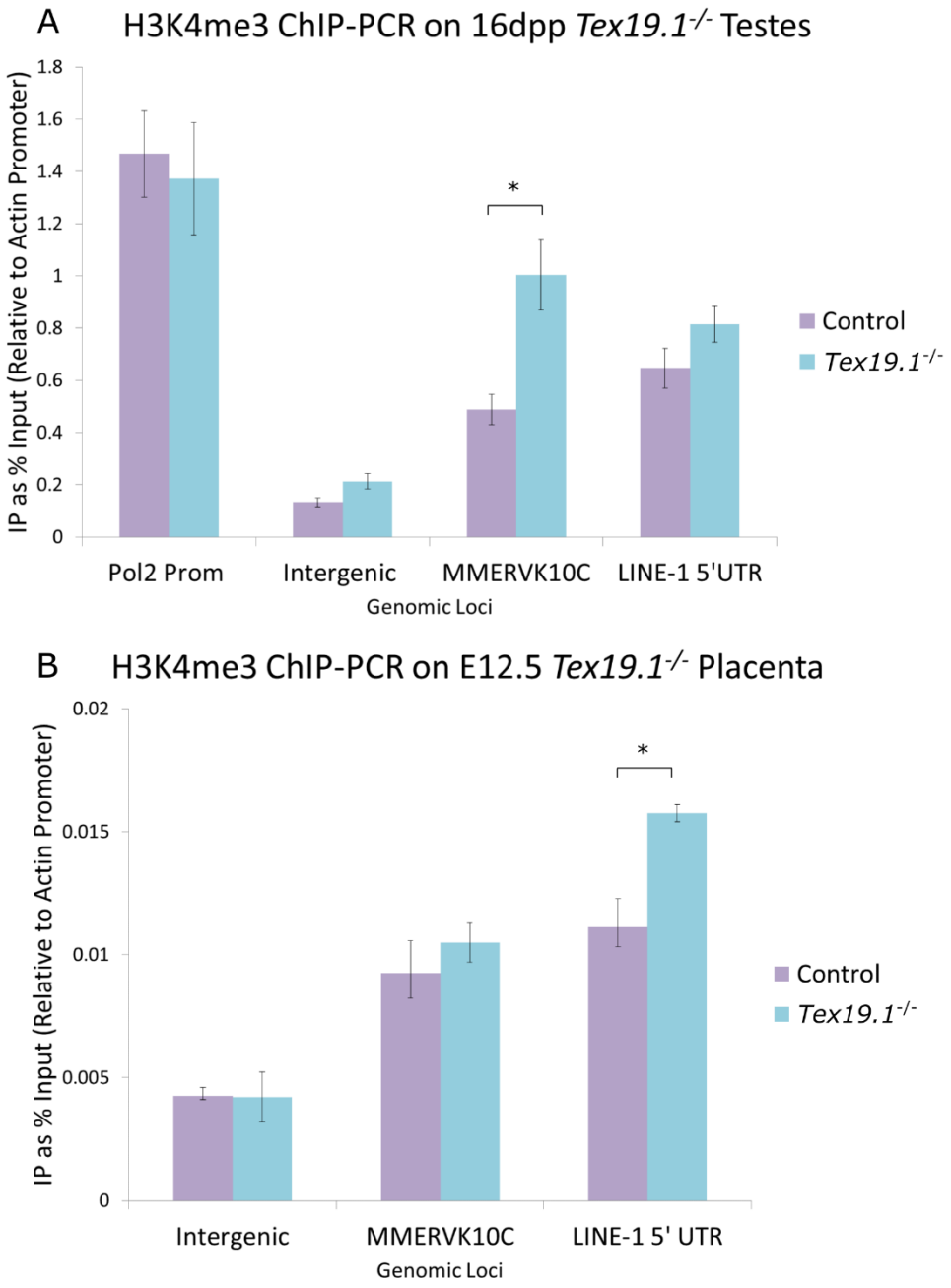


Figure 3.2: TEX19.1 represses MMERVK10C and LINE1 retrotransposons at the transcriptional level in different tissues. (A) H3K4me3 ChIP-PCR on 16dpp littermate control and *Tex19.1*$^{-/-}$ testes (biological replicates n=3, **P=0.025,* Students T-Test. (B) H3K4me3 ChIP-PCR on whole placentas from E12.5 littermate controls and *Tex19.1*$^{-/-}$ animals (n=5, **P=0.012,* Students T-Test). Graphs show mean ± standard error IP enrichment as a percentage of the input, relative to the actin promoter. qPCR was performed in triplicate for every sample and primer pair.

3.2.3 TEX19.1 interacts with the transcriptional co-repressor KAP1

I next aimed to determine the mechanism of transcriptional regulation of retrotransposons by TEX19.1. TEX19.1 is comprised of two structural domains of unknown function and the sequence gives no indication to the potential mechanism of the TEX19.1 mediated transcriptional regulation of retrotransposons. To address this, Marie MacLennan determined the interacting partners of TEX19.1 by co-immunoprecipitation mass spectrometry on the cytoplasmic fraction of TEX19.1-YFP ES cells. YFP is an enhanced yellow variant of GFP which can be detected using GFP reagents (Miyawaki et al., 1997). TEX19.1 was found to interact with KAP1; a transcriptional co-repressor of retrotransposons, which was the second most confident protein detected with 20 peptides matched and 18.7% coverage. The most confident interaction was the previously described interaction with UBR2 (Yang et al., 2010)(see appendix table 7.3 for full list of interacting peptides from mass spectrometry experiment). The interaction determined by mass spectrometry was confirmed by co-immunoprecipitation by Marie MacLennan (figure 3.3). Western blot for GFP confirmed successful immunoprecipitation of YFP or TEX19.1-YFP. In the supernatant and input material bands corresponding to endogenous KAP1 are detected in both YFP ES cells and TEX19.1-YFP ES cells. After immunoprecipitation with GFP-trap beads, endogenous KAP1 is only detected in the TEX19.1-YFP ES cells and absent from YFP ES cells. This confirms the mass spectrometry data and indicates that TEX19.1 interacts with the retrotransposon co-repressor, KAP1 in ES cells.

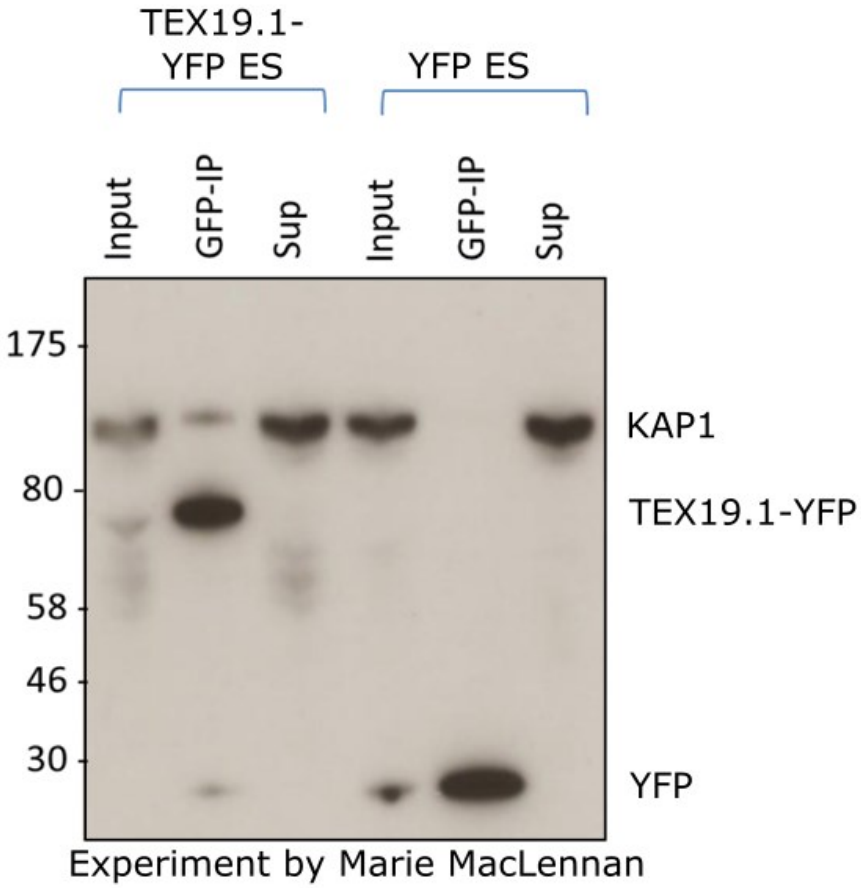


Figure 3.3: TEX19.1 interacts with KAP1 in ES cells – Experiment by Marie MacLennan. Western blot for KAP1 after co-immunoprecipitation with GFP antibodies in TEX19.1-YFP and YFP ES cells. Inputs and supernatants were not immunoprecipitated with GFP beads and a band for endogenous KAP1 is detected in each non precipitated sample. A band for KAP1 is detected in the IP sample in TEX19.1-YFP ES cells but not in YFP-ES cells (KAP1 ~100kDa, TEX19.1-YFP ~80kDa, YFP ~30kDa). Experiment was performed by Marie Maclennan.

3.2.4 TEX19.1 mediates deposition of the KAP1 dependent H3K9me3 repressive mark at LINE-1 loci in ES cells

KAP1 facilitates the deposition of tri methylation on H3K9 at retrotransposon loci (Rowe et al., 2010). I therefore aimed to determine if there were perturbations in this mechanism in the absence of *Tex19.1*. ChIP-PCR for the H3K9me3 repressive mark was performed on *Tex19.1*$^{-/-}$ ES cells for select retrotransposon candidates which are de-repressed in *in vivo Tex19.1*$^{-/-}$ contexts such as MMERVK10C, LINE-1 and VL30, alongside a known KAP1 bound subfamily of LINE-1, L1MdF2 (Castro-Diaz et al., 2014). As a control, I also analysed KAP1 dependent H3K9me3 levels at imprinted genes. H3K9me3 was enriched a similar amount at each genomic loci tested, but while there was no difference between the amount of H3K9me3 at MMERVK10C, VL30 or IAP elements in control and *Tex19.1*$^{-/-}$ ES cells, a specific KAP1-bound L1MdF2 element showed a significant 1.3 fold decrease in the amount of H3K9me3 bound (figure 3.4A)(n=6, P<0.01, Students T-test). Similar amounts of KAP1 dependent H3K9me3 were also present at the KAP1 imprinted genes, Peg3 and KvDMR in both control and *Tex19.1*$^{-/-}$ ES cells. These results indicate that in the absence of TEX19.1, deposition of H3K9me3 is perturbed at a particular LINE-1 subfamily with no effects at the other retrotransposon loci or KAP1 dependent imprinted genes tested.

I then aimed to investigate if this reduced amount of H3K9me3 was due to a reduction in the amount of KAP1 binding. ChIP-PCR for KAP1 in *Tex19.1*$^{-/-}$ ES cells showed no significant decrease in KAP1 abundance at MMERVK10C or IAP loci, however *Tex19.1*$^{-/-}$ ES cells have a significant 1.6 fold decrease in the amount of KAP1 bound at L1MdF2 elements in *Tex19.1*$^{-/-}$ (figure 3.4B)(n=5, P<0.05, Students T-test). Although the change in H3K9me3 and KAP1 binding is relatively small at this particular L1MdF2 element, it is possible that only a subset of the ~3000 full length L1MdF2 loci (Sookdeo et al., 2013) are being affected by TEX19.1. This data indicates that the physical interaction between TEX19.1 and KAP1 likely represents a functional relationship necessary for KAP1 binding and subsequent deposition of H3K9me3 at specific retrotransposon loci. Therefore, TEX19.1 acts to augment repressive histone modifications at specific retrotransposon loci as a compensatory mechanism for loss of DNA methylation.

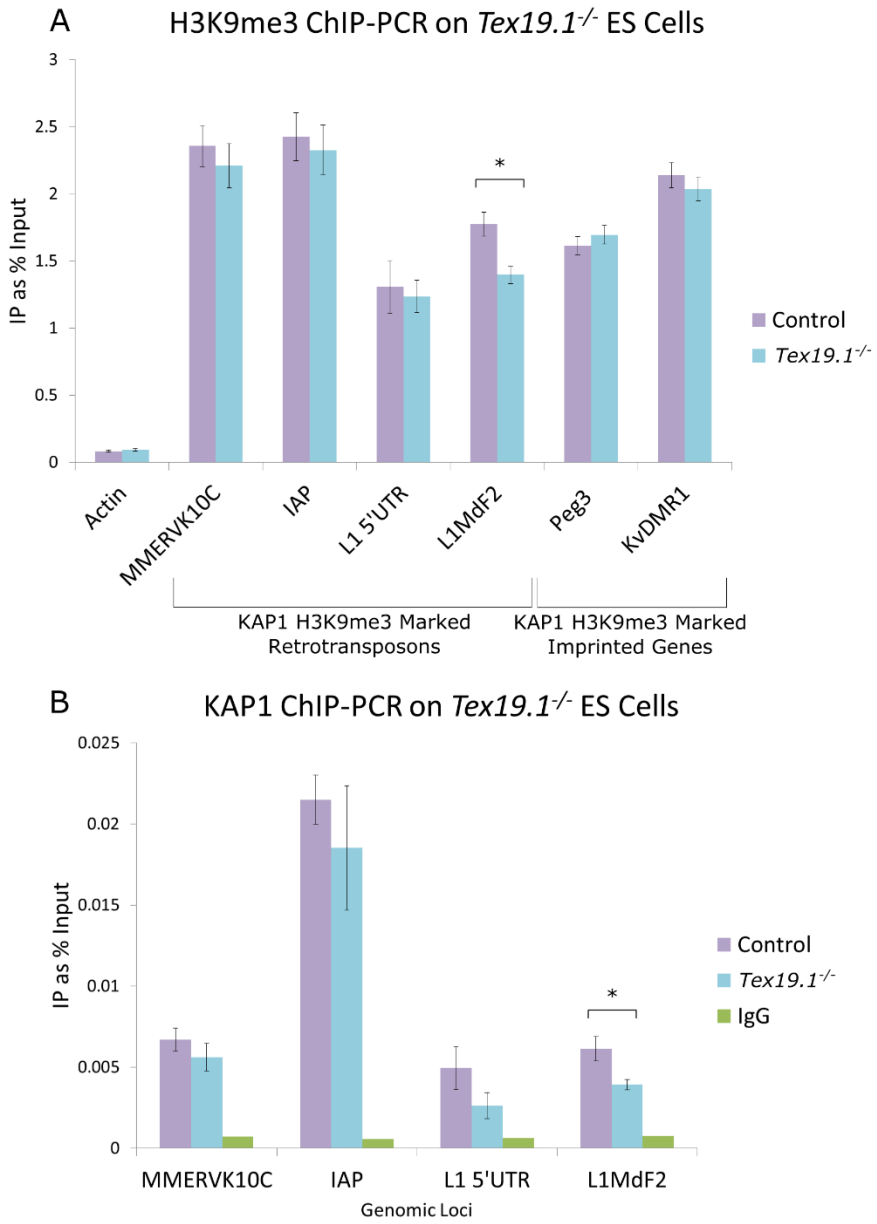


Figure 3.4: TEX19.1 is necessary for KAP1 binding and subsequent deposition of H3K9me3 at L1MdF2 elements in ES cells (A) H3K9me3 ChIP-PCR for selected retrotransposon loci in control *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ ES cells. *Actin* is shown as a negative control, H3K9me3 imprinted genes *Peg3* and *KvDMR1* as positive controls and candidate retrotransposon loci as experimental loci. (n=6, **P*=0.0066, Students T-test) (B) KAP1 ChIP-PCR for the retrotransposons assayed in A (n=5, **P*=0.033, Students T-test). IgG negative control is also shown to show the level of non-specific binding to beads during IP. Enrichment of H3K9me3 and KAP1 are shown as IP material as a percentage of the input material (mean ± standard error).

3.2.5 TEX19.1 is not bound to chromatin

I have shown that H3K9me3 and KAP1 are reduced at L1MdF2 loci in *Tex19.1*$^{-/-}$ ES cells, however it is unclear how TEX19.1 may be regulating KAP1 binding to retrotransposon loci. The KAP1-SETDB1 repressive complex is targeted to select genomic loci via particular KRAB-ZFPs which bind DNA in a sequence dependent manner. I aimed to investigate if TEX19.1 could be functioning as part of this complex by direct binding to retrotransposon chromatin. Therefore, to investigate if TEX19.1 is binding to chromatin at a global level, I prepared chromatin from ES cells stably expressing TEX19.1-YFP or YFP and performed a Western blot to detect YFP and Histone H3 as a loading control. In input and supernatant samples from TEX19.1-YFP and YFP ES cells, bands corresponding to the correct sizes of these fusion proteins and Histone H3 were detected (figure 3.5). In the TEX19.1-YFP ES samples there was a weaker band corresponding to YFP indicating that some cleavage of the TEX19.1-YFP fusion construct may have occurred (figure 3.5). The band for Histone H3 was overexposed, however, no band for TEX19.1-YFP was detected in the chromatin fraction, indicating that TEX19.1-YFP is not stably binding to chromatin in these conditions and is therefore unlikely to be acting as a KRAB-ZFP in this experimental context (figure 3.5).

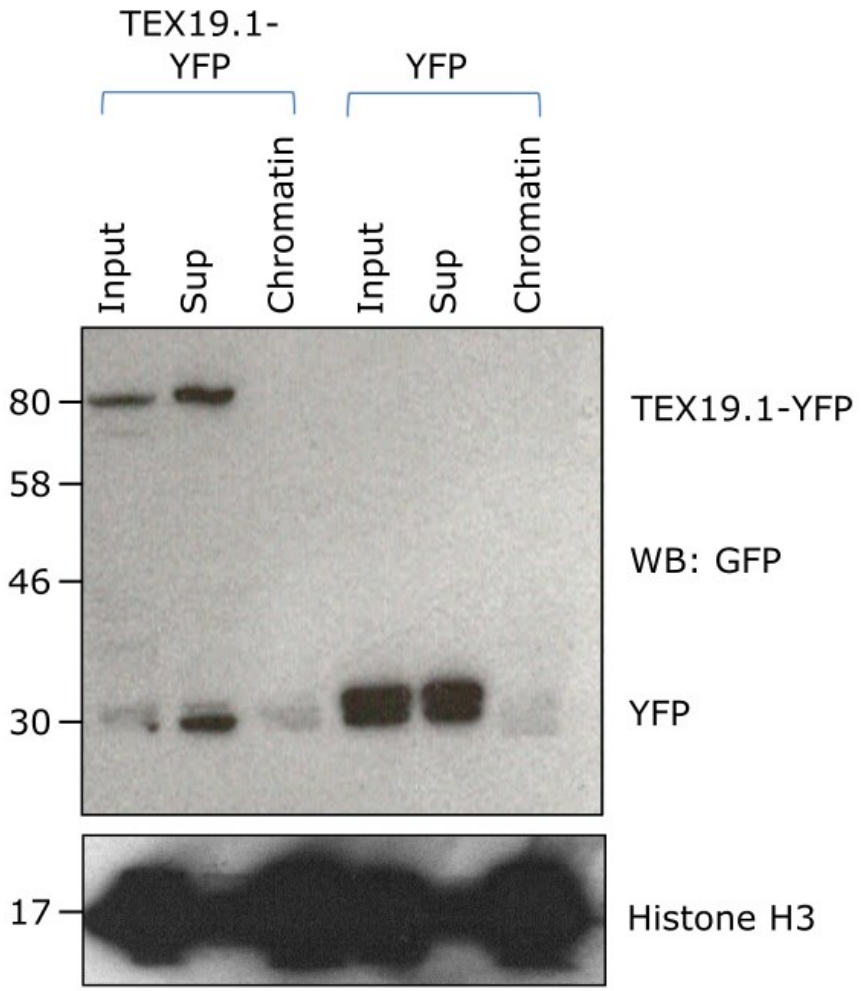


Figure 3.5: TEX19.1 is not bound to chromatin in ES cells. Western blot for GFP on chromatin preparations from Tex19.1-YFP ES cells and YFP ES cells. Input and supernatants (sup) were probed to indicate the abundance of non-chromatin bound Tex19.1-YFP and YFP. Histone H3 was used as a loading control. Pre-stained molecular weight markers in kDa are indicated.

3.2.6 KAP1 protein levels are not affected by *Tex19.1* in different cellular contexts

TEX19.1 interacts with RING domain E3 ubiquitin ligases (MacLennan et al., submitted; Yang et al., 2010) and as TEX19.1 is unlikely to be acting as a KRAB-ZFP, it is possible that the reduction in H3K9me3 and KAP1 at L1MdF2 loci is due to TEX19.1 affecting global KAP1 protein levels. To see if KAP1 levels are perturbed in an *in vivo* *Tex19.1*$^{-/-}$ context where LINE-1 is de-repressed, Western blots were performed for KAP1 on E12.5 control and *Tex19.1*$^{-/-}$ placentas (Reichmann et al., 2013). The highest enrichment in LINE-1 protein is in the junctional zone of *Tex19.1*$^{-/-}$ animals (Reichmann et al., 2013), I therefore dissected the junctional zone from the placenta to determine if specific reductions in KAP1 were occurring this layer. Control and *Tex19.1*$^{-/-}$ placentas have similar levels of KAP1 in the junctional zone indicating that TEX19.1 is likely not affecting global KAP1 stability in the placenta, however the band corresponding to KAP1 in the control junctional zone is moderately burnt in (figure 3.6). This may have implications for the quantification of KAP1 levels in control versus knockout junctional zones and, due to the small sample size more animals would need to be assessed.

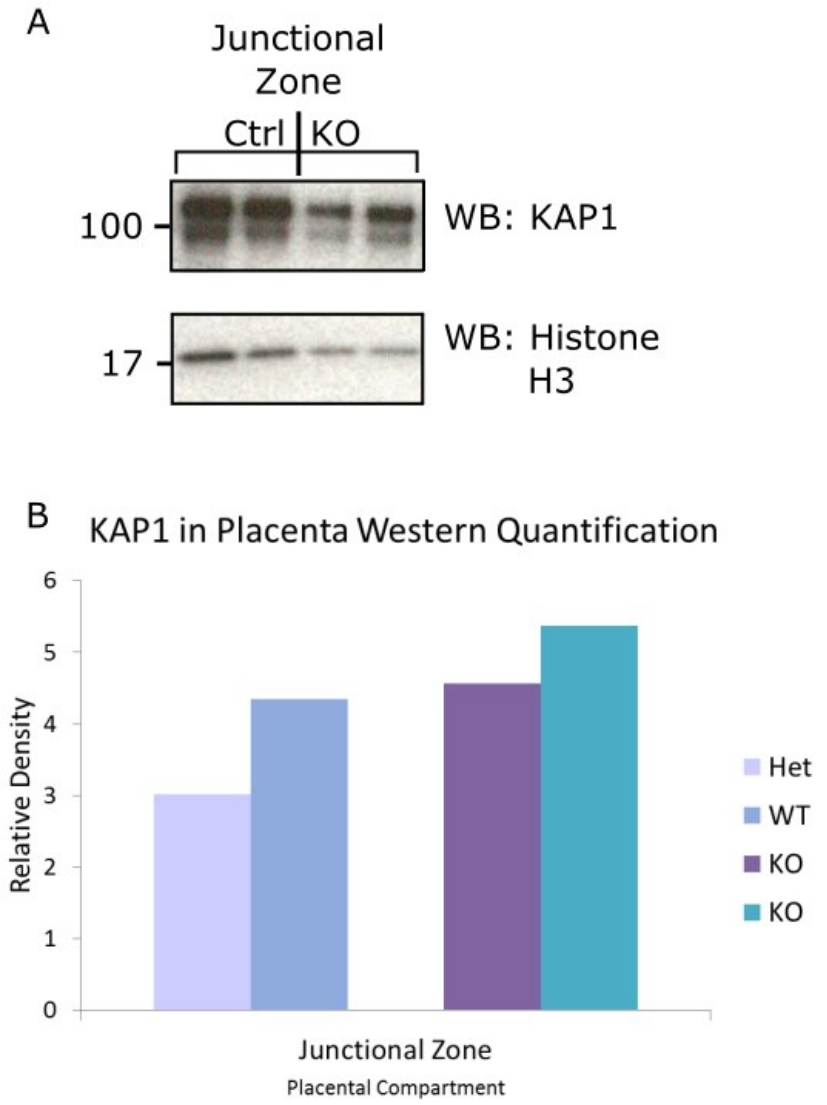


Figure 3.6: KAP1 levels are not affected by loss of TEX19.1 in the junctional zone of the placenta. (A) Western blot for KAP1 in the junctional zone of E12.5 control and *Tex19.1*$^{-/-}$ placentas. Histone H3 was used as a loading control. Vertical numbers on the left correspond to protein size (kDa). (B) Densitometry quantification of KAP1 western blot normalised to Histone H3.

I next performed an overexpression experiment to see if higher levels of TEX19 were sufficient to reduce KAP1 levels in a different cellular context. HEK293T cells were transfected with a vector expressing human *TEX19* or an empty vector as control. A vector expressing GFP acted as a transfection control. Western blotting for KAP1 showed that similar levels were present with and without TEX19 indicating that an increase in human TEX19 is not sufficient to reduce human KAP1 levels in this system. Even loading and transfection was determined by LAMIN B and GFP antibodies respectively, however the extent of TEX19 overexpression in each sample was not assessed (figure 3.7). Despite no detectable decrease in global KAP1 levels, KAP1 has many different functional sub-complexes (Iyengar and Farnham, 2011) some of which may be targeted by TEX19.1 and be responsible for the reduction in H3K9me3 in ES cells and the de-repression of LINE-1 in *Tex19.1*$^{-/-}$ placentas.

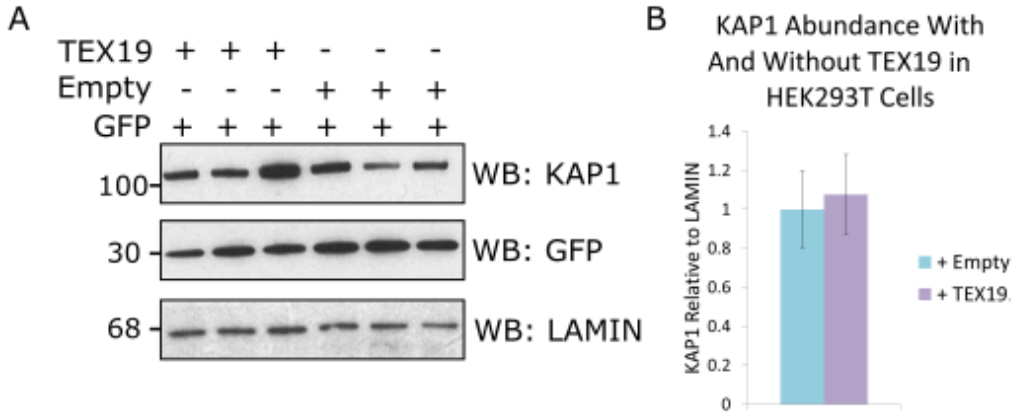


Figure 3.7: Human KAP1 levels are not dependent on TEX19 in HEK293T cells (A) Western blot and (B) quantification of KAP1 abundance in HEK293T cells after transfection with human TEX19 or an empty vector. Cells were co-transfected with GFP to control for transfection efficiency. LaminB1 was used as a loading control. KAP1 abundance was measured relative to LaminB1 then normalised to empty plasmid control cells. Pre-stained molecular weight markers in kDa are indicated.

3.2.7 TEX19.1 does not affect ubiquitination of KAP1

Previous work in the lab by Diana Best indicates that TEX19.1 inhibits the E3 ubiquitin ligase UBR2 from binding to its N-end rule substrates and stabilises N-end rule reporters (Diana Best, unpublished). I therefore aimed to determine if TEX19 could be regulating ubiquitination of endogenous human KAP1 by performing an *in vivo* ubiquitination assay in HEK293T cells transfected with tagged ubiquitin and human *TEX19*. Samples with and without his-myc-ub were pulled down with Ni^{2+} beads to isolate ubiquitinated KAP1. A band for KAP1 was present in the no his-myc-ub lane which indicates non-specific binding of KAP1 to the beads (figure 3.8). Despite this, the characteristic poly-ubiquitin ladder above the KAP1 band of the plus his-myc-ub samples indicates that KAP1 is ubiquitinated to a similar extent with and without *TEX19* (figure 3.8). Western blot for C-Myc is shown to indicate the amount of his-myc-ub present in each sample, however as with figure 3.7, the extent of over expression of *TEX19* is unknown. The input loading controls exhibited similar levels of ACTIN, however different levels of KAP1. Taken together these results indicate that global levels of KAP1 ubiquitination are not dependent on human TEX19 in this context.

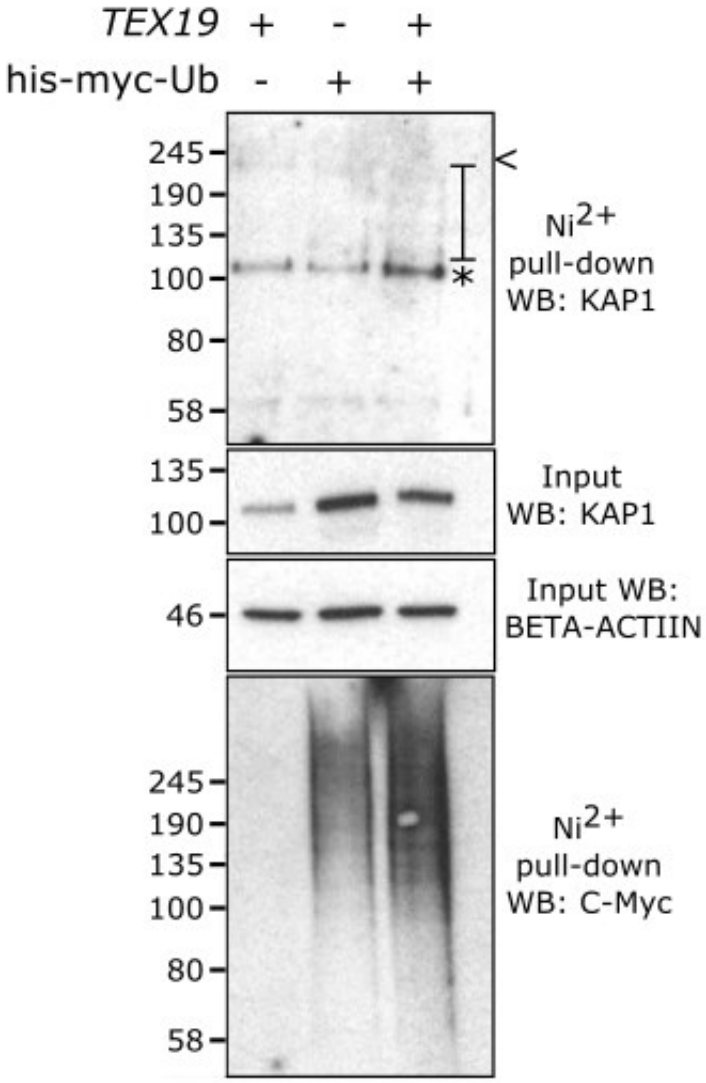


Figure 3.8: TEX19.1 does not regulate levels of endogenous KAP1 by ubiquitination
In vivo ubiquitination assay of endogenous human KAP1 in HEK293T cells in the presence and absence of human *TEX19.* Ni^{2+}-pull downs were Western blotted with anti-KAP1 antibody. Input samples were Western blotted with anti-KAP1, anti-β-ACTIN and anti-MYC antibodies. β-ACTIN serves as a loading control and C-Myc shows the amount of transfected his-myc-Ub. * indicates endogenous KAP1 whereas brackets indicates the faint ubiquitin ladder. A non-specific band at the top of the ubiquin ladder is marked by <. Representative Western blot of 3 independent experiments is shown. Pre-stained molecular weight markers in kDa are indicated.

3.2.8 Retrotransposons are not de-repressed in *Tex19.1*$^{-/-}$ ES cells

Retrotransposons are de-repressed in multiple *Tex19.1*$^{-/-}$ tissues in vivo and the reduction in the repressive H3K9me3 mark and KAP1 at L1MdF2 in *Tex19.1*$^{-/-}$ ES cells suggested that this element could also be de-repressed in ES cells. To address this globally, I used microarrays to analyse the transcriptome of *Tex19.1*$^{-/-}$ ES cells cultured in serum and LIF. To gain a general overview of any retrotransposon de-repression, I first analysed the broad classes of retrotransposon; LINE, SINE, LTR, DNA repeats and satellite repeats using probe repeat annotation (Reichmann et al., 2012). No de-repression of any class of transposon was detected (figure 3.9A). As the LTR retrotransposons MMERVK10C and VL30, members of the ERVK and ERVL families of elements, are de-repressed in *Tex19.1*$^{-/-}$ tissues I further extended this analysis to specific LTR families (ERV1, ERVK, ERVL and MalR). No upregulation of any LTR families was detected (figure 3.9B). This was confirmed by qRT-PCR for specific candidates de-repressed in *Tex19.1*$^{-/-}$ tissues (figure 3.9D). Additionally, as LINE-1 is de-repressed in *Tex19.1*$^{-/-}$ placentas I assessed levels of LINE-1 de-repression which indicated there was no effect on LINE-1 in ES cells (figure 3.9C & D). I further confirmed that there was no transcriptional de-repression of the LINE-1 class of retrotransposons by ChIP-PCR for H3K4me3 (figure 3.9E). Taken together, these data indicate that the decrease in H3K9me3 at LINE-1 loci in *Tex19.1*$^{-/-}$ ES cells is not sufficient to allow de-repression of these elements.

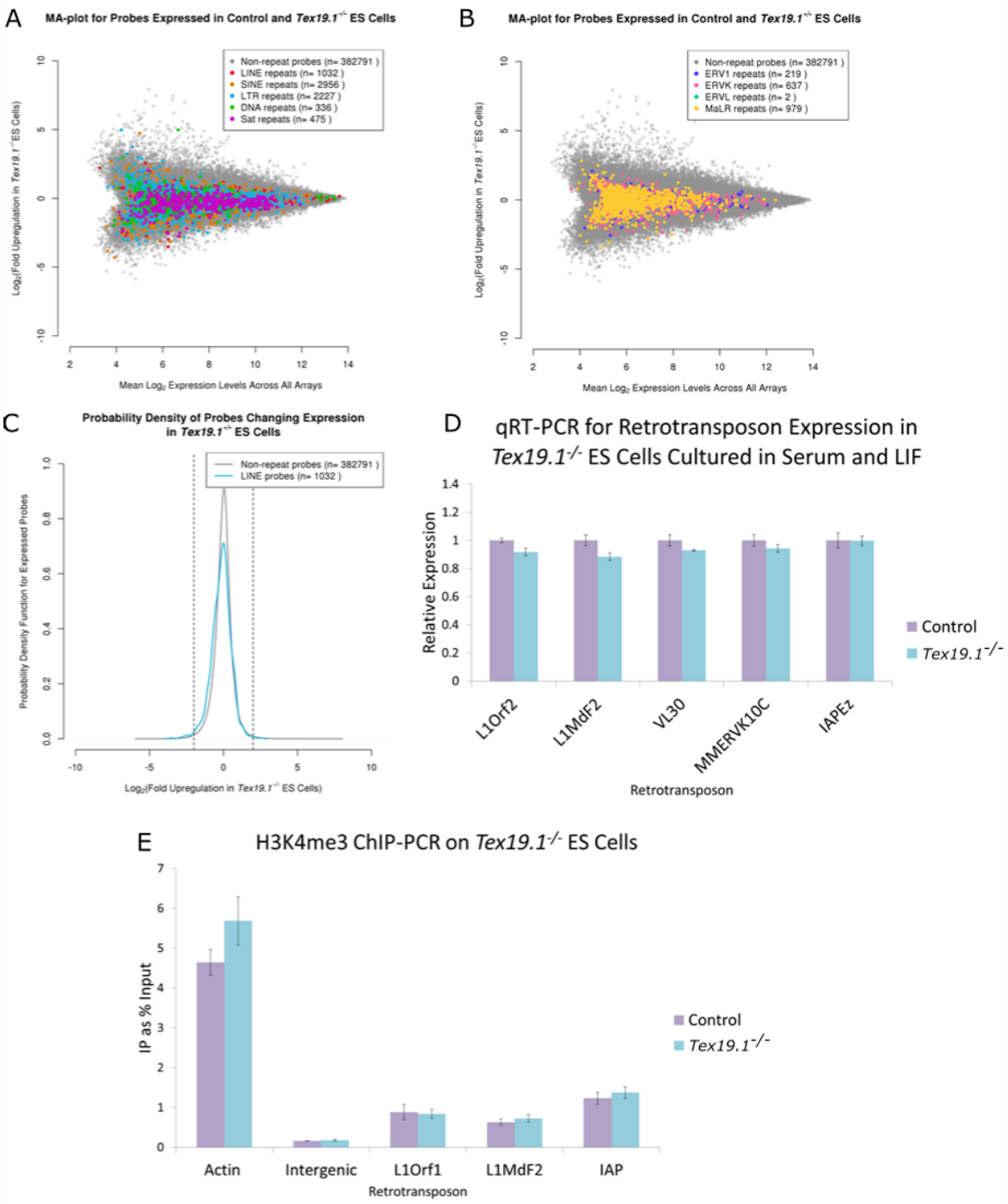


Figure 3.9: Genome wide analysis reveals retrotransposons are not de-repressed in *Tex19.1*$^{-/-}$ ES cells cultured in serum and LIF (A & B) Microarray plots showing the mean expression level for each expressed probe in *Tex19.1*$^{-/-}$ ES cell Affymetrics MoGene microarray data plotted against the fold upregulation of that probe in *Tex19.1*$^{-/-}$ ES cells. Probes corresponding to retrotransposon classes (A) and subsets of the LTR family (B) are coloured as in the legends. (C) Plot showing the expression of probes from the entire LINE-1 class of retrotransposons in *Tex19.1*$^{-/-}$ ES cells. (D) qRT-PCR verification of the lack of retrotransposon de-repression in *Tex19.1*$^{-/-}$ ES cells. Expression levels for each element (mean ± standard error) were normalised to

β-actin and then expressed relative to *Tex19.1*$^{-/-}$ ES cells (Control). (E) H3K4me3 crosslink-ChIP PCR in control in control and *Tex19.1*$^{-/-}$ ES cells (n=3). Primers to *β-actin* acted as a positive control for the ChIP with the intergenic region acting as a negative control region.

3.2.9 Differentiation of *Tex19.1*$^{-/-}$ ES cells does not induce retrotransposon de-repression

TEX19.1 represses LINE-1 and VL30 elements at the transcriptional level in the murine placenta (Reichmann et al., 2013). The placenta is a somatic tissue composed of many fully differentiated cell types. It is largely the differentiated cells of the junctional zone which have increased LINE-1 ORF1 protein in *Tex19.1*$^{-/-}$ placentas which made me question if LINE-1 de-repression was a consequence of the differentiated state of these *Tex19.1*$^{-/-}$ cells (Reichmann et al., 2013). To test if differentiation of *Tex19.1*$^{-/-}$ ES cells was sufficient to induce retrotransposon de-repression, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ ES cells were differentiated by hanging droplet culture and cells were harvested from day 0 to 9. Control and *Tex19.1*$^{-/-}$ ES cells formed spherical three dimensional aggregates over the duration of the time course (figure 3.10A) which was accompanied by a reduction in the expression of the pluripotency associated gene *Nanog*, indicating differentiation was occurring (figure 3.10A), and a reduction in expression of *Tex19.1*. To check if this differentiated cellular state was sufficient for retrotransposon de-repression in the absence of TEX19.1, qRT-PCR was performed for L1Orf2, L1MdF2, VL30 and MMERVK10C. *Tex19.1* expression is absent from *Tex19.1*$^{-/-}$ ES cells (figure 3.10B). No increase in RNA abundance for any retrotransposon was detected at any point relative to *Tex19.1*$^{+/-}$ cells suggesting that it is a property other than the differentiated state of *Tex19.1*$^{-/-}$ cells which is responsible for retrotransposon de-repression in the placenta (figure 3.10B). In contrast, VL30 and MMERVK10C expression was significantly decreased in differentiating day 6 *Tex19.1*$^{-/-}$ ES cells compared to *Tex19.1*$^{+/-}$ ES cells, implying that upon differentiation compensatory mechanisms may be increasing repression of these elements in the absence of TEX19.1. Alternatively, differentiation may be skewed in *Tex19.1*$^{-/-}$ ES cells resulting in cell composition differences, however the reason for this decreased expression is unclear (figure 3.10B).

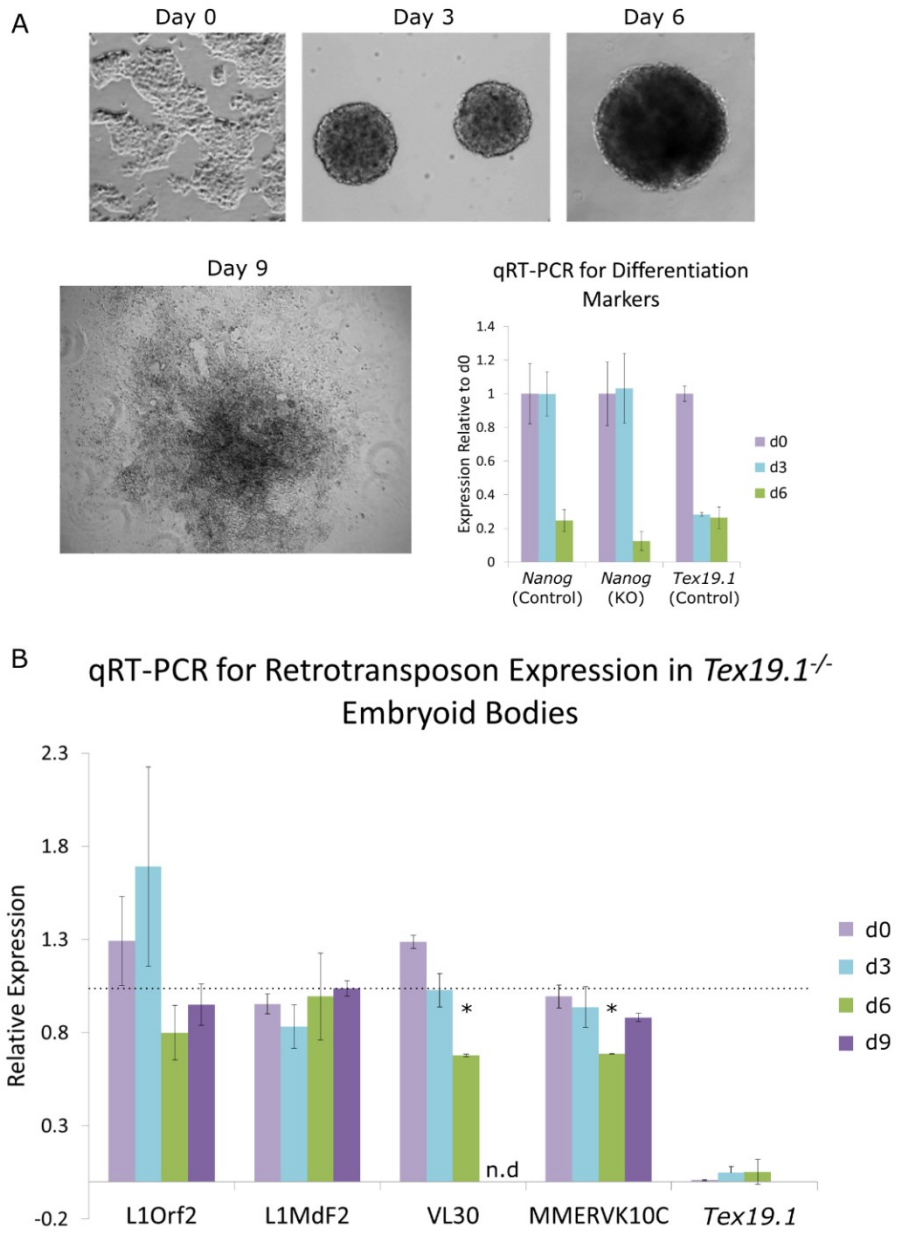
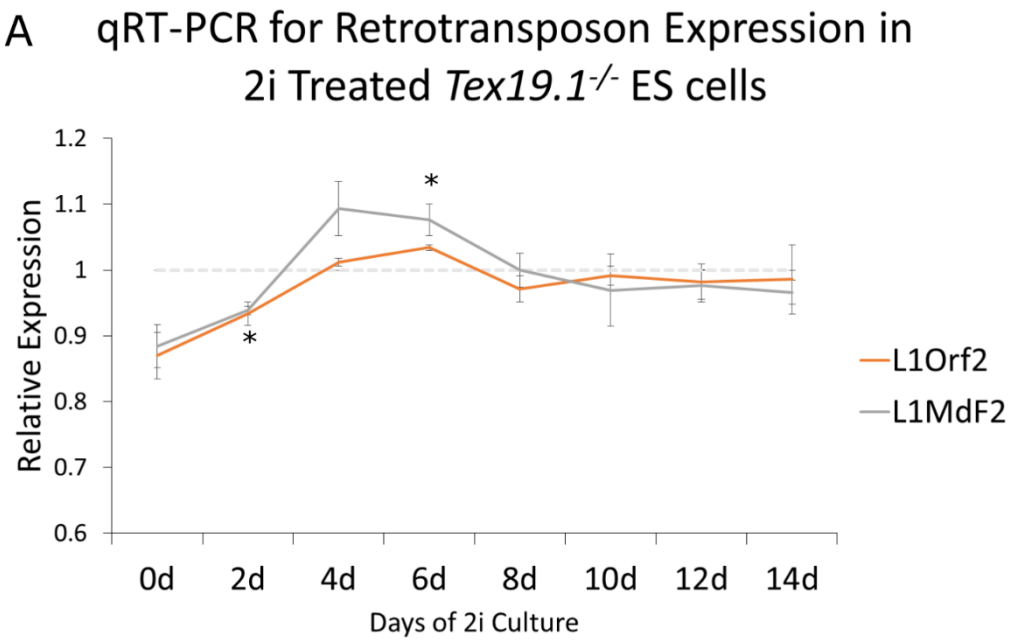


Figure 3.10: Retrotransposon de-repression does not occur upon differentiation of *Tex19.1*$^{-/-}$ ES cells. (A) ES cells were differentiated by hanging droplet culture in the absence of LIF. Embryoid body formation was assessed by phase contrast microscopy throughout the timecourse and loss of pluripotency is reflected by reduction in *Nanog* and *Tex19.1* expression after qRT-PCR relative to day 0. (d0 is 0 days of differentiation, d3 is the third day of differentiation etc.). (B**)** qRT-PCR for L1Orf2, L1MdF2, VL30 and MMERVK10C normalised to *β-actin* and expressed relative to *Tex19.1*$^{+/-}$ ES cells (mean ± standard error for biological replicates). *Tex19.1* is shown as confirmation of the absence of *Tex19.1* in the cells used. The dotted line indicates expression relative to control as 1 (n.d = not detected, **P<0.05* Students T-test)

3.2.10 Retrotransposon de-repression in 2i conditions

ES cells cultured in serum and LIF retain around 70-80% of global methylation, with around 40 – 60% methylation of LINE-1 elements (Walter et al., 2016). It is possible that these high levels of retrotransposon methylation prevent their de-repression in *Tex19.1*$^{-/-}$ ES cells, despite the reduction in H3K9me3 at L1MdF2. To determine if retrotransposons would be de-repressed in *Tex19.1*$^{-/-}$ ES cells upon reduction of DNA methylation, I cultured cells in the presence of two small molecule inhibitors of MEK and GSK3 (2i) for 14 days, which causes ES cells to enter a naïve pluripotent state and global hypomethylation to around 30% and 15- 20% at LINE-1 loci (Ficz et al., 2013; Habibi et al., 2013; Leitch et al., 2013; Seisenberger et al., 2012; Walter et al., 2016). Retrotransposon expression was then assessed by qRT-PCR every two days from entry into 2i conditions. From 0 days of culture to 4 days of culture no increase in LINE-1 expression was detected in *Tex19.1*$^{-/-}$ ES cells compared to *Tex19.1*$^{+/-}$ ES cells (figure 3.11A). At day 2, L1Orf2 was mildly but significantly decreased, however the reason for this is unclear (figure 3.11A). A mild but significant 1.08 fold increase in expression of L1MdF2 was detected at 6 days of 2i treatment (figure 3.11A&B)(n=3, **P*<0.05*,* Students T-test). These data indicate that treatment with 2i can induce mild LINE-1 de-repression in the absence of TEX19.1 at day 6 but that this de-repression is quenched from 8 days onwards.



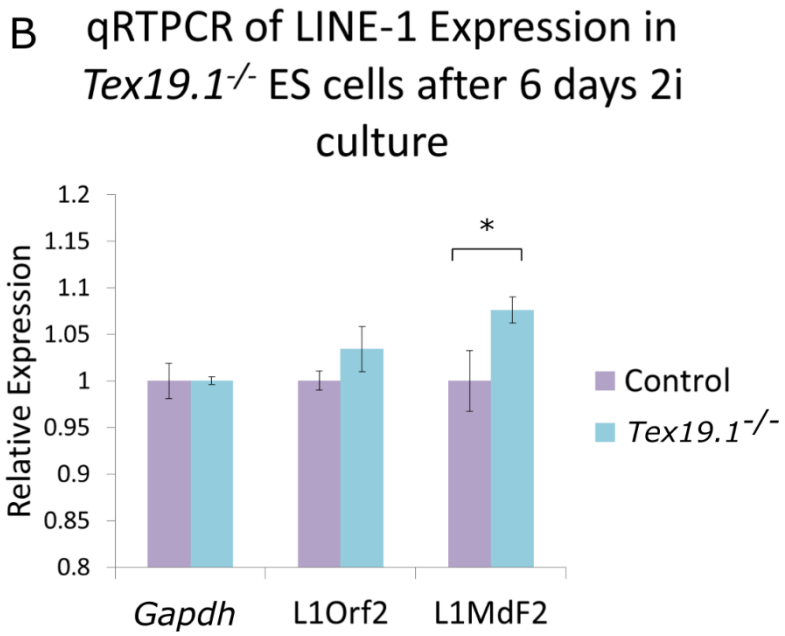


Figure 3.11: Mild de-repression of L1MdF2 occurs in *Tex19.1*$^{-/-}$ ES cells at 6 days of 2i treatment. (A) qRT-PCR for retrotransposon expression in *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ ES cells that have been cultured in 2i conditions for 14 days. RNA was harvested every other day. Expression levels for each retrotransposon (mean ± standard error) were normalised to *β-actin* and expressed relative to *Tex19.1*$^{+/-}$ ES cells (**P*=0.043 Students T-Test). (B) Bar chart representation of qRT-PCR results for 6 days culture of *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ ES cells in 2i conditions shown as line plot in (A).

3.2.11 2i Vitamin C treatment induces de-repression of many retrotransposons in *Tex19.1*$^{-/-}$ ES cells

Recent evidence from the Bourc'his lab indicates that supplementing 2i treatment with vitamin C dramatically reduces the DNA methylation levels of ES cells to approximately 5% global DNA methylation across all genomic regions, including retrotransposons (Walter et al., 2016). A transient spike in retrotransposon de-repression is observed at 6 days of 2i+vitamin C treatment which is then quenched by H3K9me3 or H3K27me3 polycomb based mechanism of repression depending on the retrotransposon family (Walter et al., 2016). Due to the global reduction in DNA methylation levels to 5% upon 2i + vitamin C treatment, I therefore asked if and when retrotransposons would become de-repressed in *Tex19.1*$^{-/-}$ ES cells relative to *Tex19.1*$^{+/-}$ ES cells in the presence of 2i and vitamin C by performing a time course followed by qRT-PCR expression analysis.

No de-repression of retrotransposons occurs from day 0 to day 4 and at day 6 only L1Orf2 is moderately de-repressed 1.09 fold which is lost at day 8 (figure 3.12, $P<0.05$, Students T-test). At 10 days of 2i vitamin C treatment MMERVK10C and L1MdF2 become significantly de-repressed in *Tex19.1*$^{-/-}$ compared to *Tex19.1*$^{+/-}$ ES cells (figure 3.12)(1.47 fold, $P<0.01$ and 1.42 fold, $P<0.05$, respectively, Students T-Test), whereas VL30 and L1Orf2 have more moderate increases in expression (1.16 fold; $P>0.05$, and 1.14 fold; $P<0.05$, respectively, Students T-Test). Interestingly, the de-repression of MMERVK10C is completely lost by 12 days, whereas, expression of L1MdF2 is moderately reduced compared to day 10 (1.42 fold to 1.22 fold; $P>0.059$) and maintained at this level to 14 days treatment (figure 3.12). IAP retrotransposons significantly increased their expression from 8 days to 14 days. It is important to note that the transient spike in retrotransposon de-repression that was observed by Walter et al was not detected in either the *Tex19.1*$^{+/-}$ or *Tex19.1*$^{-/-}$ ES cells. The reason for this is unclear, however vitamin C stability may have been a contributing factor. It is not known if the 2i vitamin C treated *Tex19.1* ES cells became hypomethylated to the same extent as in Walter et al. It was also surprising that upon 2i, vitamin C treatment de-

repression of all retrotransposons analysed occurred at some stage in the time course. This may indicate that in hypomethylated cellular states, TEX19.1 may be repressing more retrotransposons than previously appreciated.

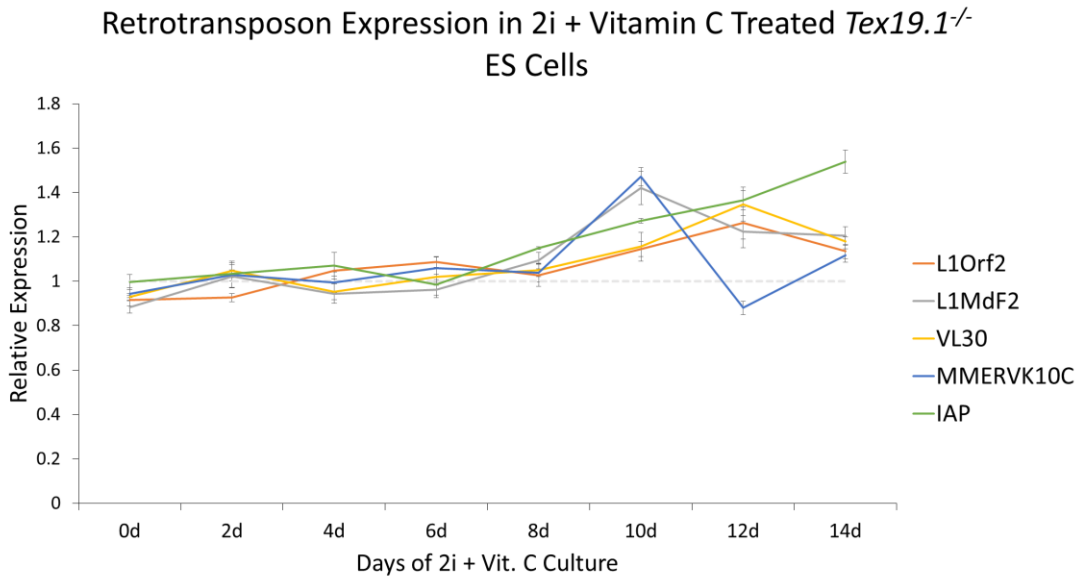


Figure 3.12: 2i Vitamin C treatment induces retrotransposon de-repression in *Tex19.1*$^{-/-}$ ES cells. qRT-PCR for retrotransposon expression in *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ ES cells that have been cultured in 2i plus vitamin C conditions for 14 days. RNA was harvested every other day. Expression levels for each retrotransposon (mean ± standard error) were normalised to *β-actin* and expressed relative to *Tex19.1*$^{+/-}$ ES cells. The dashed horizontal line indicates normalised expression in *Tex19.1*$^{+/-}$ ES cells. Significance *p* values (Students T-Test) are discussed in the main text.

3.2.12 Transcriptional de-repression of KAP1 bound, H3K9me3 repressed LINE-1 elements occurs in *Tex19.1*$^{-/-}$ placentas

I next aimed to extend my findings from hypomethylated cell culture models to *in vivo* hypomethylated placenta tissue. KAP1 ChIPseq on wild type ES cells indicates that KAP1 binds predominantly to L1MdF2 elements and at lower levels of the older (LX, L1Mus2 and L1Mus1) and younger LINE-1 elements (L1G and L1A) (figure 3.13B) (Castro Diaz et al., 2014). As the amount of H3K9me3 and KAP1 was decreased on the L1MdF2 subfamily element in *Tex19.1*$^{-/-}$ ES cells, I aimed to determine if this subfamily was de-repressed in *Tex19.1*$^{-/-}$ placentas to indicate a functional relationship between the reduction of H3K9me3 and KAP1 in vitro and LINE-1 de-repression *in vivo*.

No de-repression was detected for the LX, L1Mus2, L1Mus1, L1G and L1A subfamilies, each of which had a low level of KAP1 binding in ES cells (figure 3.13A&B)(Castro Diaz et al., 2014). The L1MdF2 element had a significant 1.7 fold de-repression (Figure 3.13A)(P<0.05, Students T-test,), indicating that the mechanism of KAP1 mediated transcriptional repression described in vitro may also be perturbed *in vivo*. This is also supported from RNAseq repeat analysis as discussed in chapter 4.

To test if H3K9me3 was reduced in the *Tex19.1*$^{-/-}$ placentas at the L1MdF2 site I performed H3K9me3 ChIP-PCR on whole placenta. The active *β-Actin* promoter acted as a negative control and showed no enrichment in H3K9me3 (figure 3.14). Enrichment of H3K9me3 was detected at repressed retrotransposon loci in both control and *Tex19.1*$^{-/-}$ placentas, however a reduction in H3K9me3 was not detected at the L1MdF2 loci or global LINE-1 loci (figure 3.14). This is possibly due to the inherent heterogeneity of the placenta, whereby subtle differences in H3K9me3 levels may be difficult to detect compared to a homogenous cell population like ES cells. As *Tex19.1*$^{-/-}$ is only expressed in the hypomethylated trophectoderm derived component of the placenta it would be expected that only this component would be affected upon loss of *Tex19.1*. Conversely, this result may suggest an alternative mechanism of TEX19.1 mediated retrotransposon repression is operating in the placenta.

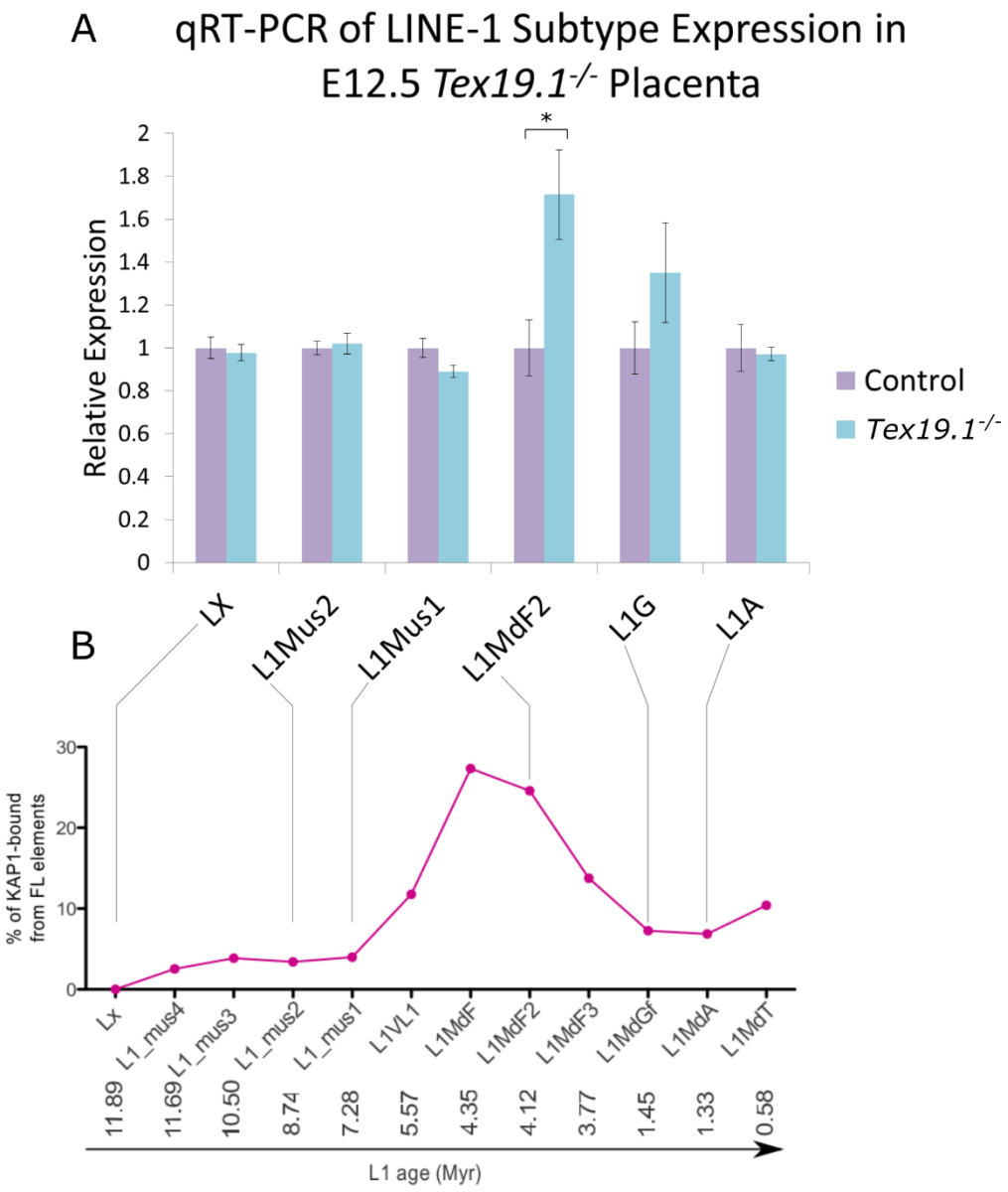


Figure 3.13: KAP1 bound LINE-1 subfamilies are de-repressed in *Tex19.1*$^{-/-}$ placentas
(A) qRT-PCR for retrotransposon expression in E12.5 *Tex19.1* control and *Tex19.1*$^{-/-}$ whole placentas. Expression levels for each LINE-1 subfamily (mean ± standard error) were normalised to *β-actin* and expressed relative to *Tex19.1* control placentas. LINE-1 subfamilies are ordered on evolutionary age *P<0.05, Students T-test. (B) Figure modified from Castro-Diaz et al., 2014. Percentage of full length KAP1 bound LINE-1 elements per LINE-1 subfamily in ES cells. Subfamilies are arranged from the oldest to the youngest subfamily (Castro-Diaz et al., 2014)

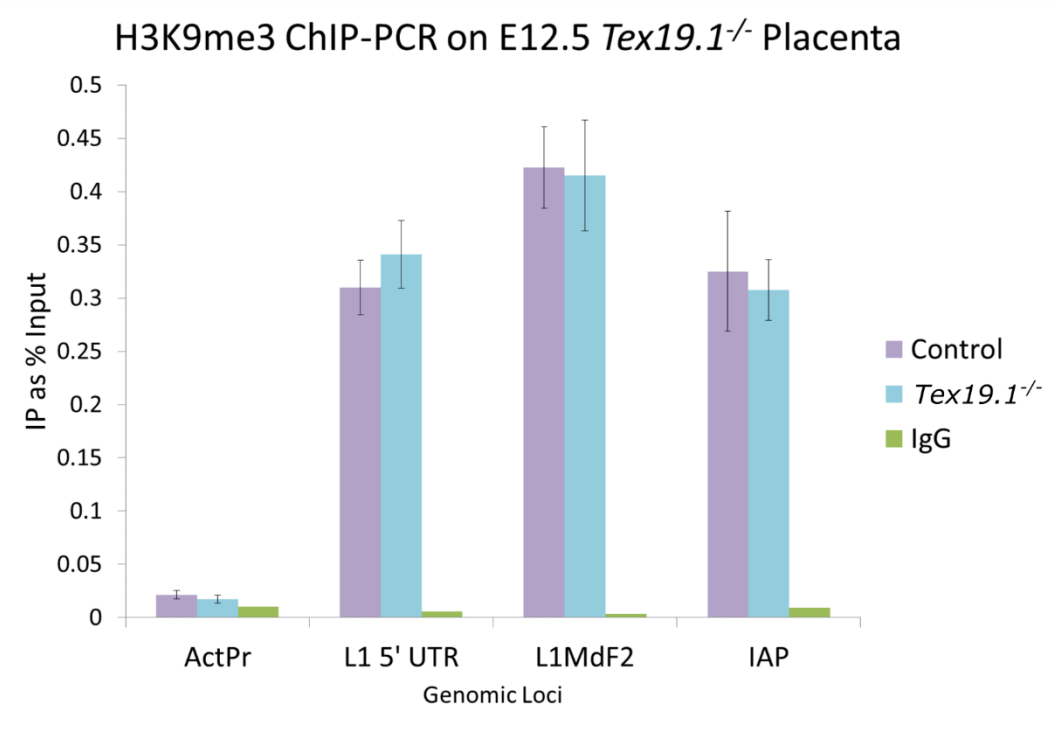


Figure 3.14: H3K9me3 is not reduced at LINE-1 loci in *Tex19.1*$^{-/-}$ placentas (A) H3K9me3 ChIP-PCR for selected retrotransposon loci in control *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ ES cells. *Actin* (ActPr) is shown as a negative control (n=3).

3.2.13 DNA methylation levels at retrotransposon loci are similar in *Tex19.1*$^{-/-}$ and control contexts

The placenta is globally hypomethylated when compared to other mouse tissues and retrotransposons are not resistant to this hypomethylation. Around 50% of cytosines in CpG dinucleotides are methylated within LINE-1 elements in the placenta compared to around 80% methylation in the embryo (Popp et al., 2010). In order to assess if TEX19.1 was affecting methylation of LINE-1 *in vivo*, I performed bisulphite sequencing on whole E12.5 placentas, alongside embryo controls. Two different subfamilies were analysed. The first was L1A which has only low level binding to KAP1 and is not de-repressed in the *Tex19.1*$^{-/-}$ placentas. L1A was 94% methylated in the embryo compared to 71% and 72% in control and *Tex19.1*$^{-/-}$ placentas respectively (figure 3.15A). I then designed bisulphite sequencing primers to detect CpG dinucleotides in the de-repressed L1MdF2 subfamily. Embryos had 82% methylation compared to 52% and 54% for control and knockout placentas respectively (figure 3.15B). This indicates that there does not appear to be a strong effect of loss of *Tex19.1* on the DNA methylation levels of these LINE-1 elements in the placenta.

L1MdF2 has 1.4 fold lower methylation than L1A which may contribute to the de-repression of this element in the absence of *Tex19.1*. As there is no de-repression of LINE-1 in *Tex19.1*$^{-/-}$ ES cells one hypothesis is that this could be due to high levels of DNA methylation. Therefore, the DNA methylation level of these particular LINE-1 subfamilies was assessed in control and *Tex19.1*$^{-/-}$ ES cells. L1A and L1MdF2 loci had approximately 70% and 75% of CpG methylation in both control and *Tex19.1*$^{-/-}$ ES cells respectively, similar to reported levels (figure 3.15C&D)(Popp et al., 2010). Therefore, TEX19.1 is not regulating the DNA methylation levels of different LINE-1 elements in ES cells of the placenta.

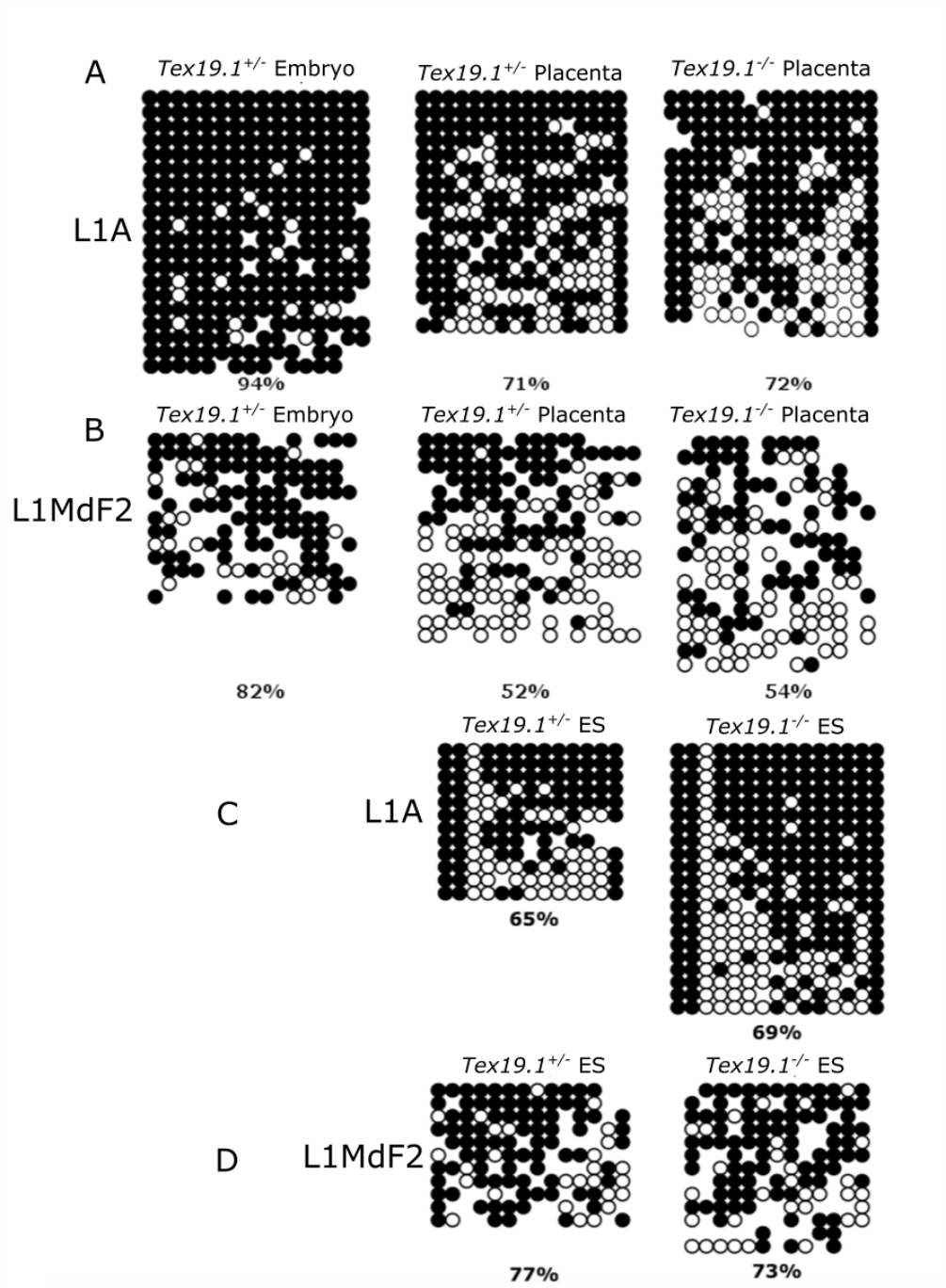


Figure 3.15: *Tex19.1* does not affect DNA methylation of different classes of LINE-1 element in the placenta or ES cells. Bisulphite sequencing showing the methylation status of L1A (A) and L1MdF2 (B) LINE-1 subfamilies in embryo, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ placentas and ES cells (C & D). Methylated and unmethylated CpGs are denoted by filled circles and empty circles respectively, with the percentage of CpG methylation shown below each. Empty spaces reflect ambiguous base calls in Sanger sequencing reads and were not included in percentage CpG calculations.

3.3 Discussion

In this chapter I investigated possible compensatory mechanisms of *Tex19.1* mediated retrotransposon repression in DNA hypomethylated conditions. Initially, I determined that *Tex19.1* is the only gene activated in *in vivo* and *in vitro* DNA hypomethylated contexts and that *Tex19.1* facilitates repression of retrotransposons at the transcriptional level in different tissues (table 3.2). I then showed that this transcriptional repression is likely due to a role of TEX19.1 in the KAP1 mediated repressive mechanism, as indicated by reduced KAP1 binding and H3K9me3 at a particular subfamily of retrotransposons *in vitro* (table 3.2). I demonstrate that de-repression of different retrotransposons occurs upon experimentally induced DNA hypomethylation in a temporal manner, in ES cells lacking *Tex19.1*. Finally, I extend these findings *in vivo* and indicate that the particular retrotransposon subfamily with reduced H3K9me3 in ES cells is de-repressed in *Tex19.1*$^{-/-}$ placentas (summarised in table 3.2).

Histone Modification / Histone Modifier Analysed	Retrotransposon Loci Studied	*Tex19.1*$^{-/-}$ context	Findings From This Study
H3K4me3 (Transcriptionally active mark)	LINE-1	ESCs	No change
		Placenta	Increased
		Testes	No change
	MMERVK10C	ESCs	No change
		Placenta	No change
		Testes	Increased
H3K9me3 (Transcriptionally repressive mark)	LINE-1	ESCs	Decreased
		Placenta	No change
		Testes	Not assessed
	MMERVK10C	ESCs	No change
		Placenta	No change
		Testes	Not assessed
KAP1	LINE-1	ESCs	Decreased
		Placenta	Not assessed
		Testes	Not assessed
	MMERVK10C	ESCs	No change
		Placenta	Not assessed
		Testes	Not assessed

Table 3.2: Summary table of histone modifications and the KAP1 repressive cofactor binding at LINE-1 and MMERVK10C loci in different *Tex19.1*$^{-/-}$ contexts. Summarised from the ChIP-PCR experiments discussed throughout chapter 3.

3.3.1 Compensatory mechanisms for retrotransposon repression upon DNA hypomethylation are working through *Tex19.1*

Previous work from the Meehan lab shows that upon experimentally induced hypomethylation *in vitro*, methylation sensitive genome defence genes are expressed, the majority of which silence retrotransposons (Crichton et al., 2014; Hackett et al., 2012). I extend these findings to *in vivo* hypomethylated conditions and show that of the *in vitro* methylation sensitive genes, only one gene, *Tex19.1,* is expressed in all hypomethylated tissues analysed (figure 3.1). This expression of *Tex19.1* in placenta,

ES cells and PGCs is consistent with previous reports (Hackett et al., 2012; Kuntz et al., 2008; Reichmann et al., 2013). In the germline, mutagenic consequences resulting from increased retrotransposon activity would be transmitted to the next generation, whereas in the placenta no transmission to offspring would occur. In normal circumstances, the placenta has co-opted numerous retrotransposon sequences, mainly of the LTR class, to provide dynamic regulation of developmental transcriptional networks in trophoblast stem cells, along with divergence of some LTRs into placental specific genes (Chuong et al., 2013; Emera et al., 2012; Mi et al., 2000; Thompson et al., 2016). This is contrary to aberrant retrotransposon de-repression in *Tex19.1*$^{-/-}$ placentas which are possibly associated with placental developmental defects as explored in Chapter 5 of this thesis. There may be other methylation sensitive genes expressed in the placenta with previously unappreciated genome defence roles, however these genes are unlikely to be methylation sensitive in all tissues, as these would have been detected in the Hackett et al. study. It is possible that some important methylation sensitive genome defence genes may have fallen below the stringent 6 fold change cut off compared to normally methylated conditions. Regardless, this suggests that utilising developmentally dynamic DNA methylation to switch on *Tex19.1* is an effective way to defend the genome in multiple developmental contexts vulnerable to increased retrotransposon activity (figure 3.16).

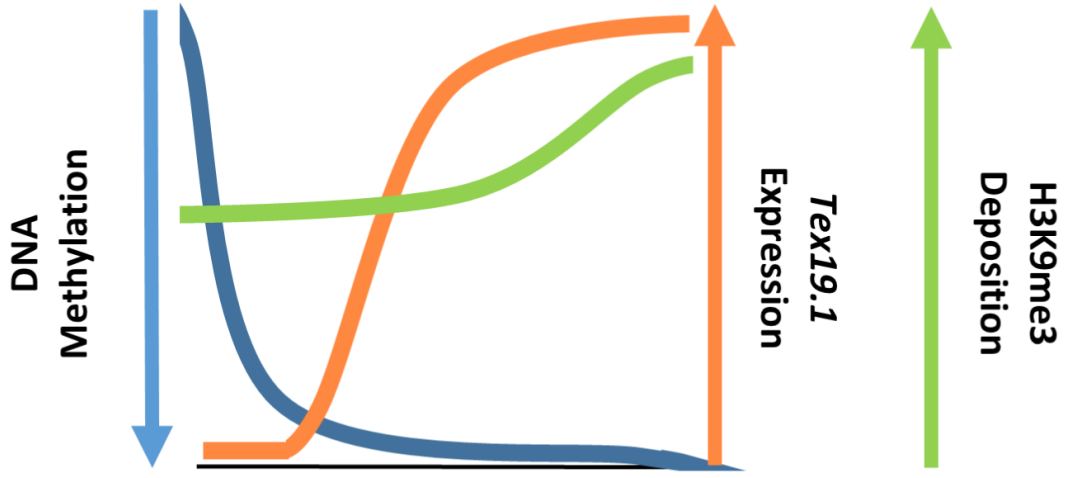


Figure 3.16: Tex19.1 augments repressive histone modifications to compensate for loss of DNA methylation. Schematic diagram indicating that as DNA methylation is

reduced, *Tex19.1* expression is induced which acts to augment trimethylation of H3K9 at particular loci.

3.3.2 TEX19.1 and the KAP1 mediated repressive mechanism

The potential mechanism of *Tex19.1* mediated retrotransposon repression at the transcriptional level has never been investigated, therefore one aspect of my PhD was to attempt to elucidate the molecular basis of this repression. I have shown that repression of retrotransposons by TEX19.1 can be transcriptional and data from Marie MacLennan indicates that TEX19.1 interacts with the transcriptional co-repressor KAP1. I determined that KAP1 and subsequent H3K9me3 levels were reduced at a KAP1 repressed LINE-1 subfamily in *Tex19.1*$^{-/-}$ which suggests that TEX19.1 was functioning in this KAP1 mediated mechanism. It is important to note that the small but significant 1.3 fold reduction in H3K9me3 binding at specific retrotransposon loci in *Tex19.1*$^{-/-}$ compared to *Tex19.1*$^{+/-}$ ES cells was in line with other studies where KAP1 is being perturbed (Castro-Diaz et al., 2014; Fasching et al., 2015). For example, retrotransposon de-repression occurs in neural progenitor cells conditionally deleted for *Kap1* (Fasching et al., 2015). This is one of the only somatic *Kap1* null cell types which has retrotransposon de-repression but when H3K9me3 ChIP-PCR was performed for the de-repressed IAP1 pol loci, only around a 1.5 fold decrease in H3K9me3 levels was detected (Fasching et al., 2015). Additionally, studies on ES cells knocked down for the LINE-1 targeting KRAB-ZFP, Gm6871 also revealed only a moderate decrease in H3K9me3 upon ChIP-PCR (Castro-Diaz et al., 2014). This suggests that additional KAP1 independent mechanisms or particular KAP1 sub-complexes may also be responsible for H3K9me3 mediated repression in different contexts. Furthermore, H3K9me3 ChIP sequencing of these *Tex19.1*$^{-/-}$ ES cells will highlight if any other retrotransposon subfamilies are targeted by TEX19.1 mediated H3K9me3.

3.3.3 Possible roles of TEX19.1 in the KAP1 repressive mechanism

The KAP1 repressive mechanism is dependent on the recruitment of the KAP1:HMT repressive complex to different loci in the genome by KRAB-ZFPs, often offering partial redundancy in the retrotransposon loci targeted in ES and somatic cells (Ecco et al., 2016; Najafabadi et al., 2015; Wolf et al., 2015b). KRAB-ZFPs have tissue specific expression patterns, therefore the differing retrotransposons de-repressed in different *Tex19.1*$^{-/-}$ contexts may reflect perturbations in tissue specific KRAB-ZFP proteins (Ecco et al., 2016; Rowe et al., 2013; Wolf et al., 2015). The co-immunoprecipitation mass spectrometry experiment performed by Marie MacLennan did not detect an interaction between TEX19.1 and either of the HMTs, SETDB1 or SUV39H1/2, however it did identify a KRAB-ZFP (ZNF295) as a potential binding partner of TEX19.1. In humans ZNF295 (ZBTB21) is a ubiquitously expressed transcriptional repressor, however it is not known if this KRAB-ZFP has roles for retrotransposon repression (Wang et al., 2005). It is possible that this ZFP could be responsible for the TEX19.1 mediated deposition of H3K9me3 at LINE-1 loci in ES cells and experiments determining if TEX19.1 is perturbing the interaction of KAP1 with a particular KRAB-ZFP or HMT would address this further.

Additionally, the specificity of the H3K9me3 reduction at a particular LINE-1 subfamily indicates that TEX19.1 may not be affecting global KAP1 function but could possibly be perturbing specific KRAB-ZFP function in some way. Indeed, global KAP1 levels are not perturbed in *Tex19.1*$^{-/-}$ placentas or in overexpression experiments with human TEX19. Also, it does not appear that TEX19.1 itself is acting as a KRAB-ZFP as I have not detected TEX19.1-YFP bound to chromatin in *in vitro* experiments on ES cells. A caveat of this experiment is that endogenous TEX19.1 is still present and may be preventing binding of TEX19.1-YFP. I is also unclear if this fusion protein is function. One future experiment to address this would be to transfect TEX19.1-YFP into *Tex19.1*$^{-/-}$ ES cells to detect any binding of TEX19.1-YFP to sites that otherwise would have been bound by endogenous TEX19.1.

One way in which levels of KRAB-ZFPs are regulated is by KAP1 dependent ubiquitylation. I have shown that the levels of ubiquitylation of KAP1 are similar with

and without TEX19, however it is unclear if KAP1 protein is regulated by ubiquitin dependent proteolysis. To address this, treatment of cells with a proteasome inhibitor such as MG132 followed by an *in vivo* ubiquitylation assay would determine if KAP1 turnover is dependent on the proteasome pathway. Regardless, KAP1 turnover itself is likely not being affected by TEX19, however KAP1 is also an E3 ubiquitin ligase and it is possible that TEX19 is regulating the ability of KAP1 to ubiquitylate a target protein. MAGE-I proteins have similar expression patterns to *Tex19.1* and have roles in binding to KAP1 which promotes the poly-ubiquitination and degradation of KRAB-ZFPs, such as ZNF382, leading to decreased binding to particular target loci (Xiao et al., 2011). Interestingly, previous work in the lab by Diana Best indicates that TEX19.1 inhibits the E3 ubiquitin ligase UBR2 from binding to its N-end rule substrates, stabilising N-end rule reporters (Diana Best, unpublished). The role of TEX19.1 in perturbing the ubiquitination and turnover of KRAB-ZFPs is yet to be investigated. If this is the case, this would provide a mechanism whereby reduced recruitment of KAP1 to retrotransposon loci occurs due to a reduced amount of KRAB-ZFP, resulting in less H3K9me3.

3.3.4 Combined repressive effects of DNA methylation and TEX19.1

In ES cells, silencing of LINE-1 transcription does not appear to be reliant on DNA methylation as ES cells null for the *de novo* methyltransferases, *Dnmt3a/b,* and maintenance methyltransferase, *Dnmt1,* have no change in LINE-1 expression (Karimi et al., 2011; Reichmann et al., 2012). *Tex19.1* becomes more highly expressed in these hypomethylated cells than wild type ES cells and may therefore be preventing de-repression of LINE-1 elements by augmenting H3K9me3 for wild type ES cells as I have shown in this chapter (Castro Diaz et al., 2014; Karimi et al., 2011). The levels of LINE-1 de-repression upon removal of *Kap1* or *Setdb1* are relatively small in ES cells but far greater in *Suv39h1/2* null ES cells indicating that there is redundancy in mechanisms of repression (Bulut-Karslioglu et al., 2014; Castro-Diaz et al., 2014; Karimi et al., 2011; Matsui et al., 2010). MMERVK10C are also only modestly expressed in ES cells null for *Dnmt3a, Dnmt3b* and *Dnmt1*, whereas are highly de-

repressed in *Kap1* and *SetDB1* null ES cells indicating that H3K9me3 silences these MMERVK10C *in vitro*. Conversely, the tissue specific expression of MMERVK10C, LINE-1 or VL30 in different *Tex19.1*$^{-/-}$ contexts may require specific transcription factors that are present in germ cells but absent from the somatic placenta and vice versa. This has been shown for MLV retrotransposon expression, a subset of which is dependent on the PAX5 transcription factor in *Setdb1* null B-lymphocytes (Fasching et al., 2015). Interestingly, loss of H3K9me3 did not correlate with retrotransposon de-repression in these cells, highlighting the importance of tissue specific transcription factors in controlling the expression of retrotransposons in different cellular contexts (Fasching et al., 2015).

I have shown that in *Tex19.1*$^{-/-}$ ES cells cultured in 2i + Vitamin C retrotransposon de-repression occurs in a temporal manner for different retrotransposons. One caveat of this experiment is that the transient spike in retrotransposon expression determined by Walter et al. (Walter et al., 2016), was not detected either in the control *Tex19.1*$^{+/-}$ or *Tex19.1*$^{-/-}$ ES cells. This may be a consequence of cell line differences or the extent of DNA hypomethylation achieved in our culture conditions. The levels of DNA methylation were not assessed in these cells and would need to be investigated further to support this conclusion. The dynamics of histone modifications in these *Tex19.1* null ES cells cultured in 2i and vitamin C was also not investigated in this study but would be intriguing to see if compensatory repressive histone modifications are being perturbed at different classes of retrotransposon loci than in serum / LIF cultured ES cells. It remains unclear why retrotransposons that have no detectable decrease in H3K9me3 in serum/LIF cultured ES cells are also de-repressed in the absence of *Tex19.1* in 2i / vitamin C conditions.

3.3.5 *In Vivo* de-repression of KAP1 bound LINE-1 elements

Previous work from the Trono lab has shown that particular subfamilies of LINE-1 are preferentially bound by KAP1 in ES cells compared to older and younger subfamilies (Castro Diaz et al., 2014). I show in this chapter that it is the most highly KAP1 bound LINE-1 subfamily, L1MdF2 which is de-repressed in the *Tex19.1*$^{-/-}$ placenta (figure

3.13). I demonstrate that this de-repression is transcriptional (figure 3.2), however changes in H3K9me3 at L1MdF2 loci could not be detected in *Tex19.1*$^{-/-}$ placentas (figure 3.14). This is possibly due to the heterogeneous nature of placental tissue masking any changes that may be occurring in subsets of hypomethylated trophectoderm-derived cells. Alternatively trophoblast cells are distinct from ES cells, therefore the mechanism of de-repression detected in ES cells may not be operating in different contexts. Indeed, KAP1 is known to regulate different retrotransposons in ES cells compared to relatively few examples in somatic cells and the roles for KAP1 or SETDB1 mediated transcriptional regulation have never been assessed in a placental context (Ecco et al., 2016; Fasching et al., 2015; Rowe et al., 2010). Furthermore, TEX19.1 is largely cytoplasmic in ES cells but appears more nuclear in the placenta (Ollinger et al., 2008; Reichmann et al., 2013). The effect of this differing intracellular localisation on TEX19.1 binding to KAP1 and molecular function are still unclear. To further investigate the role of *Tex19.1* in retrotransposon repression in a placental context, using trophoblast stem cells null for *Tex19.1* would allow ChIP assays on a more homogenous cell population. The epigenome of the placenta remains relatively unexplored and future conditional knockouts of key epigenetic modifiers such as KAP1 and SETDB1 in a variety of stages of trophectoderm development would enhance this understanding of retrotransposon regulation in this tissue.

In summary, I have shown that *Tex19.1* is the only gene to response to DNA hypomethylation in all different hypomethylated contexts assessed. This activation acts to augment repressive histone modifications at LINE-1 loci to limit their de-repression upon reductions in DNA methylation (figure 3.16).

Chapter 4: Investigation of the role of *Tex19.1* in development of the murine placenta

4.1 Introduction

The second aim of this thesis was to investigate why *Tex19.1*$^{-/-}$ placentas have developmental defects. As previously mentioned in chapter 3, DNA methylation is high in differentiated embryonic tissues, however is globally lower in placenta (Popp et al., 2010; Razin et al., 1984). *Tex19.1* is the only methylation-sensitive, germline genome defence gene expressed in the placenta and *Tex19.1*$^{-/-}$ animals have developmental defects with trophectoderm derived cells being perturbed at an undefined point in gestation (Hackett et al., 2012; Reichmann et al., 2013). *Tex19.1*$^{-/-}$ pups exhibit intrauterine growth restriction (IUGR), however it is unclear when and why these placental defects are occurring.

One interesting phenotype of the *Tex19.1*$^{-/-}$placentas is the de-repression of retrotransposons. In chapter 3 I show that TEX19.1 can influence the KAP1 dependent transcriptional repressive, H3K9me3 mark at retrotransposon loci to compensate for reduced DNA methylation (figure 3.4). I also show that the increases in retrotransposon RNA in *Tex19.1*$^{-/-}$ placentas and testes is transcriptional, as demonstrated by gains in the transcriptionally active associated mark, H3K4me3 at de-repressed retrotransposon loci (figure 3.2). This led me to hypothesise that the transcriptional de-repression of retrotransposons, due to changes in chromatin state, may be contributing to the defects in *Tex19.1*$^{-/-}$ placentas. One explanation may be the transcriptional de-repression of retrotransposons leading to an increase in retrotransposition to new genomic locations. Unique retrotransposition events may indeed be occurring in some cells, however the severity and uniformity of the placental phenotype between individual *Tex19.1*$^{-/-}$ mice suggests that this would be unlikely, requiring the same retrotransposition event to occur independently at early trophoblast stages of development of each *Tex19.1*$^{-/-}$ animal. A previously under-appreciated aspect of retrotransposons is their influence on transcriptional regulation of neighbouring genes, both in normal developmental contexts and in experimental contexts, where retrotransposon repression is defective (Chuong et al., 2013; Thompson et al., 2016). Chuong et al show that early transcriptional networks in trophoblast stem cells are regulated and driven by retrotransposon derived enhancers.

The gain of this regulatory capacity of retrotransposons was restricted to hypomethylated tissues, indicating that co-option of retrotransposons is generally restricted to tissues permissive for their activity (Chuong et al., 2013). An alternative hypothesis is that retrotransposon de-repression in *Tex19.1*$^{-/-}$ placentas could be causing increases in retrotransposon derived proteins or nucleic acids that interfere with host cell biology, or potentially induce innate immune responses leading to increased placental dysfunction and IUGR (Chiappinelli et al., 2015; Roulois et al., 2015). Aberrant removal of endogenous nucleic acids is associated with a variety of autoimmune diseases in humans such as Aicardi-Goutières Syndrome (Crow and Rehwinkel, 2009). Mouse models of AGS which are null for proteins mutated in humans, indicate increases of retrotransposon nucleic acids in the cytoplasm of cells and likely leads to activation of a Type I interferon response and a variety of associated defects (Gall et al., 2012; Morita et al., 2004; Stetson et al., 2008). Despite these interesting hypotheses, it is also possible that *Tex19.1* could be influencing placental development independently of any effect on retrotransposons and during this chapter I aimed to further characterise the *Tex19.1*$^{-/-}$ placental phenotype, whilst investigating the possible pathogenic causes of the placental defects observed.

Previous transcriptional profiling of *Tex19.1*$^{-/-}$ placentas at E18.5 indicated an increase in many retrotransposons from the LINE-1 class to the LTR class (Reichmann et al., 2013). This transcriptomic data, combined with histological analysis of *Tex19.1*$^{-/-}$ placental sections highlighted numerous compositional differences at E18.5. I therefore sought to investigate the developmental timing of defects in *Tex19.1*$^{-/-}$ placentas, with a view to performing transcriptome analysis at a developmental stage when *Tex19.1*$^{-/-}$ and control placentas have more similar cellular compositions. This would then allow me to determine genic and retrotransposon expression changes and gain potential insights into the pathology of the placental and IUGR phenotypes.

During this chapter, I identify that placental defects precede IUGR and establish that placental weight reduction is likely due to defects in early junctional zone development between E10.5 and E12.5. This is accompanied by reductions in particular trophoblast derived marker gene mRNA, which is associated with upregulated LINE-1 subfamilies, however the functional relevance of this association remains unclear. I also show that there is a type I interferon response in *Tex19.1*$^{-/-}$ placentas which may

be causative of the observed increased in apoptosis and reduced placental size. Thus, I have identified a possible role for *Tex19.1* in maintaining placental immune homeostasis possibly via transcriptional repression of retrotransposons, which is essential for correct development of the placenta and consequently the developing embryo.

4.2 Results

4.2.1 Reductions in Placenta Weight Precede Intrauterine Growth Restriction in *Tex19.1*$^{-/-}$ Animals

Previous analysis of the *Tex19.1*$^{-/-}$ embryonic and placental phenotypes has only been performed on E18.5 animals (Reichmann et al., 2013). Therefore, it was unclear at what stage of embryonic development the placental defects and intrauterine growth restriction was occurring. To determine this, I analysed placenta and embryo weights earlier in development at E14.5 and E12.5 as midpoints in placental development where the labyrinth, junctional zone and decidua are all established (Coan et al., 2004).

E14.5 *Tex19.1*$^{-/-}$ placentas had a significant 1.3 and 1.4 fold decrease in weight compared to *Tex19.1*$^{+/+}$ and *Tex19.1*$^{+/-}$ placentas respectively (figure 4.1A & C)(*Tex*$^{+/+}$ $P<0.01$, *Tex*$^{+/-}$ $P<0.01$, Mann Whitney U Test). To determine if there was a potential placental phenotype upon loss of one copy of *Tex19.1* I compared heterozygous *Tex19.1*$^{+/-}$ to *Tex19.1*$^{+/+}$ placentas. A small but significant 1.1 fold decrease in *Tex19.1*$^{+/-}$ placental weight was detected (figure 4.1A)($P<0.05$, Mann Whitney U Test) which was not observed at E18.5 implying that heterozygous placentas recover through development (Reichmann et al., 2013). These data indicate that loss of *Tex19.1* is affecting placental development by E14.5. Having one copy of *Tex19.1* also appears to be detrimental to early placental development, however to a far lesser extent than in *Tex19.1*$^{-/-}$ animals. This may imply that when comparing heterozygous *Tex19.1*$^{+/-}$ placental phenotypes to *Tex19.1*$^{-/-}$ placental phenotypes this is likely an underestimation of defects occurring in the *Tex19.1*$^{-/-}$ animals at E14.5. Furthermore, analysing the heterozygous phenotype in comparison to the *Tex19.1*$^{+/+}$ may give an indication of more subtle dosage dependent phenotypes as a result of loss of *Tex19.1.* To further statistically assess differences in placenta and embryo weights, power calculations would enable the determination of the minimum sample size of placenta and embryos required to reduce the possibility of false negative differences in weights due to small sample sizes.

To reach a time point closer to the onset of placental defects, I took litters at E12.5 and weighed the placentas. E12.5 *Tex19.1*$^{-/-}$ placentas were still significantly reduced in

weight 1.2 and 1.3 fold compared to *Tex19.1*$^{+/+}$ and *Tex19.1*$^{+/-}$ respectively (figure 4.1B)(P<0.05, P<0.01, Mann Whitney U Test), however to a lesser extent than at E14.5. Furthermore, preliminary data from E10.5 suggests that *Tex19.1*$^{-/-}$ placenta weights are similar to those of control placentas (*Tex19.1*$^{-/-}$ 34mg, n=2; *Tex19.1*$^{+/-}$ 27mg, n=7, data not shown). The placental defects in *Tex19.1*$^{-/-}$ animals therefore likely arise between E10.5 and E12.5, however additional embryos need to be analysed at E10.5 to allow a proper statistical comparison to be made. Together these data indicate *Tex19.1* is essential for correct placental development to E12.5 with reduced placental weight detectable at E12.5 and E14.5.

To investigate if placental defects preceded IUGR I weighed embryos at E14.5 (figure 4.1A & C) and E12.5 (figure 4.1B). At both E14.5 and E12.5 there was no reduction in *Tex19.1*$^{-/-}$ embryo weights compared to *Tex19.1*$^{+/+}$ and *Tex19.1*$^{+/-}$ indicating that growth restriction of the embryos was likely initiating after these developmental time points (figure 4.1)(E14.5; *Tex19.1*$^{+/+}$ P=0.06, *Tex*$^{+/-}$ P=0.08, E12.5; *Tex19.1*$^{+/+}$ P=0.4, *Tex19.1*$^{+/-}$ P=0.06, Mann Whitney U Test). No decrease in embryo weight was found between *Tex19.1*$^{+/-}$ and *Tex19.1*$^{+/+}$ animals at either time point. Taken together, these data show that the placental defects in *Tex19.1*$^{-/-}$ animals precede IUGR.

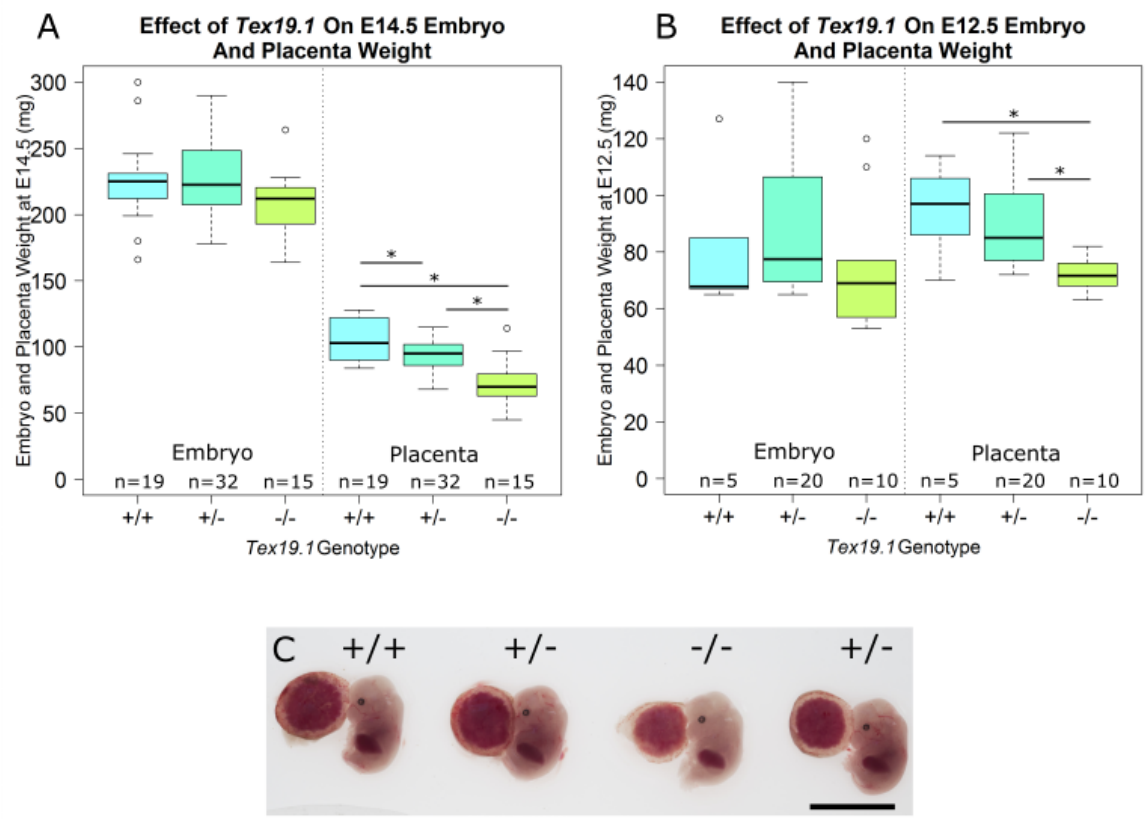


Figure 4.1: *Tex19.1*$^{-/-}$ placenta defects precede IUGR. (A) Boxplot showing weights of *Tex19.1*$^{+/+}$, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ embryos and placentas at E14.5. (B) Boxplot showing weights of *Tex19.1*$^{+/+}$, *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ embryos and placentas at E12.5. Numbers of pups is shown above the genotype. Note difference in scale between E14.5 and E12.5 boxplots (*P<0.05, Mann-Whitney U-test). (C) Photograph of E14.5 embryos and placentas from one litter. Genotype is indicated above the embryos. Scale bar = 1cm

4.2.2 E14.5 placentas have no increase in cellular senescence

From previous analysis at E18.5 reductions in placental weight were deemed to be due to loss of different trophoblast derived cell types, however the cause of the cell loss was not investigated (Reichmann et al., 2013). *Tex19.1* placentas have de-repression of LINE-1 elements but also the VL30 elements of the LTR class retrotransposons (Reichmann et al., 2013). Cellular senescence, a stable cell cycle arrest which can be sustained for long periods, has recently been associated with a human LTR class retrotransposon (Chuprin et al., 2013). This LTR derived gene encodes the protein Syncytin, which facilitates the fusion of the syncytiotrophoblast in the human placenta, which contributes to the feto-maternal barrier and facilitates gas and nutrient exchange (Mi et al., 2000). Expression of the LTR derived gene in human fibroblasts and cancer cells was sufficient to form multinuclear cells via fusion, and induce cellular senescence (Chuprin et al., 2013). Due to this induction of cellular senescence upon the activation of this LTR, I hypothesised that this may be occurring in the *Tex19.1*$^{-/-}$ placentas where LTR retrotransposons are de-repressed. To assess this, E14.5 *Tex19.1*$^{-/-}$ and control placentas were cryosectioned and stained for cellular senescence associated β-galactosidase (Campisi and d'Adda di Fagagna, 2007). To act as a positive control for the cellular senescence β-galactosidase assay I dissected the mesonephros from E14.5 embryos, as mesonephric tubules undergo p21 dependent, pre-programmed cellular senescence at this stage (Muñoz-Espín et al., 2013). The β-galactosidase staining of the mesonephros exhibited very similar patterns to those observed by Munoz-Espin et al., indicating that the assay was working as expected and was detecting previously described senescent cells in an *in vivo* context (figure 4.2A). Only a few sparse β-galactosidase positive cells were observed in control placentas (figure 4.2B). These were generally large polyploid trophoblast giant cells localised to the labyrinth. *Tex19.1*$^{-/-}$ placentas also had similar low numbers of β-galactosidase positive cells as with the controls, which appeared to be largely trophoblast giant cells (figure 4.2B). Little staining was observed in syncytiotrophoblast cells unlike in human placentas (Chuprin et al., 2013). Therefore, the levels of senescent, β-galactosidase positive cells are unlikely to be sufficient to cause the large decrease in *Tex19.1*$^{-/-}$ placental weight at this developmental time point.

To further confirm that cellular senescence was not playing a major role in the reduction in placenta weight, I performed qRT-PCR for classical markers of cellular senescence on whole E14.5 placentas. Alternate reading frame protein (*Arf/Cdkn2a*) positively regulates p53 through the inhibition of the E3 ubiquitin protein ligase HDM2 which facilitates p53 degradation. The expression of the cyclin dependent kinase inhibitor p21 is then induced by p53 to induce cellular senescence directly or via activation of the p16-retinoblastoma protein tumour suppressor pathways (Campisi and d'Adda di Fagagna, 2007). *Arf* and *p21* were expressed in control and knockout placentas but *p16* was not. No increased expression was observed for *Arf* or *p21* in *Tex19.1*$^{-/-}$ placentas (figure 4.2C). Furthermore, p21 was significantly down regulated 1.5 fold, which is in contrast to the upregulation of *p21* expected if the senescence pathway was activated (figure 4.2C)($P<0.05$, Students-T Test, n=3). P21 can also act as an inhibitor of apoptosis in a number of systems. The downregulation of *p21* may therefore suggest that an increase in apoptosis may be present in *Tex19.1*$^{-/-}$ placentas compared to controls (Gartel and Tyner, 2002). Together, the β galactosidase assay and the qRT-PCR results indicate that increased cellular senescence is unlikely to be a major factor in the reduction of placental weight of *Tex19.1*$^{-/-}$ animals at this time point.

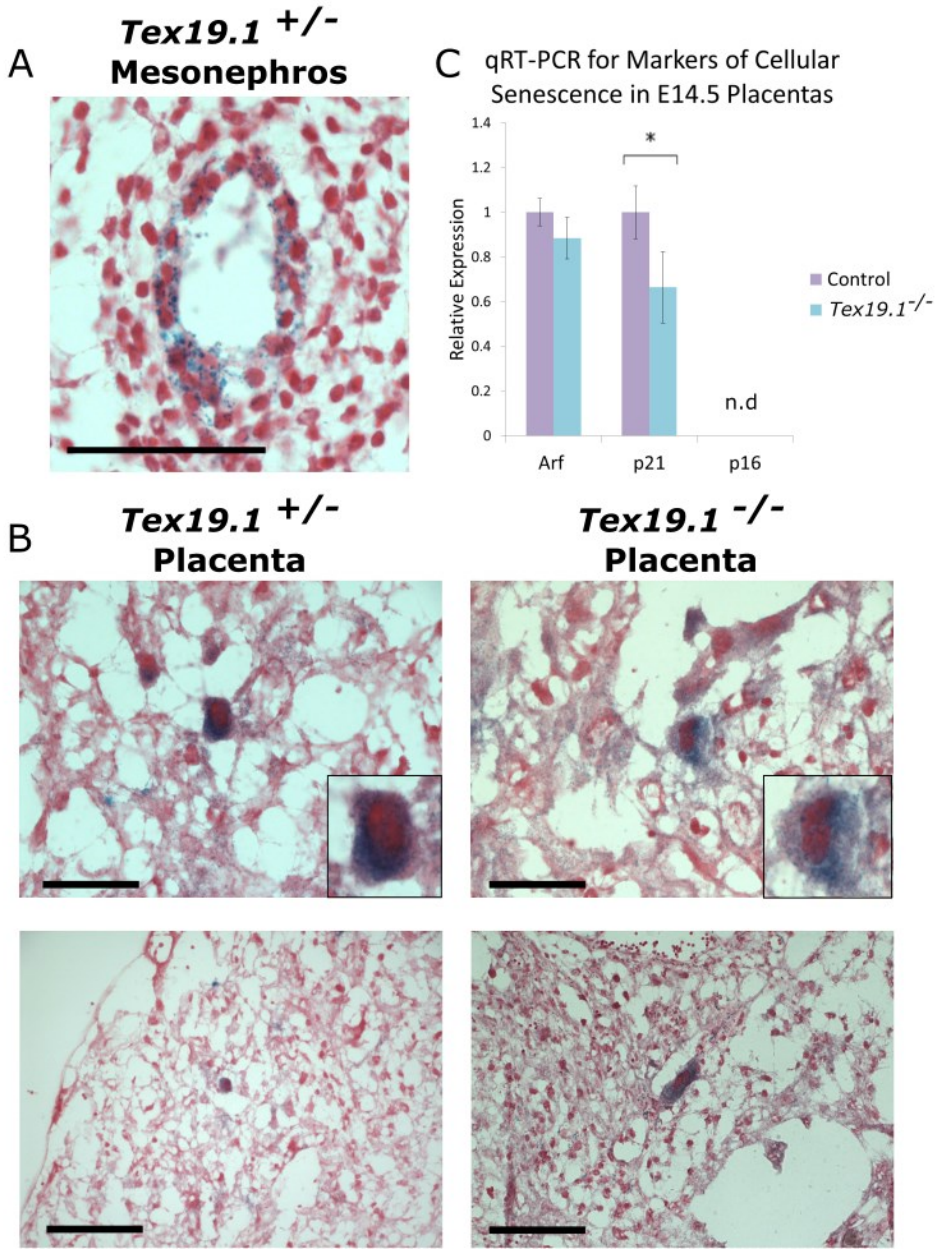


Figure 4.2: Senescence is unlikely to be a major cause of *Tex19.1*$^{-/-}$ placenta developmental defects at E14.5. (A) Cryosection of mesonephros from E14.5 embryo stained with senescence associated β galactosidase and counterstained with nuclear fast red. Blue colouring surrounding the mesonephric tubules indicates positive staining of senescent cells (Munoz-Espin et al, 2013). Scale bar=25µm. (B) Cryosections of littermate control and *Tex19.1*$^{-/-}$ placentas at E14.5 stained with senescence associated β galactosidase. Images from the labyrinth are shown with example positive trophoblast giant cells shown in the inset. Scale bar=50 µm, bottom scale bar=100µm. (C) qRT-PCR on whole E14.5 littermate control and knockout placentas for senescence associated cellular markers. Data is normalised to β actin and expressed relative to control (n=3, *P<0.05 Students T-test, n.d; not detected).

4.2.3 *Tex19.1*$^{-/-}$ placentas have increased cell death

As the decrease in *Tex19.1*$^{-/-}$ placental weights could not be attributed to cellular senescence, I therefore aimed to determine if there was an increase in cell death throughout the different layers of the placenta. E14.5 and E12.5 placental tissue sections were investigated by the TUNEL assay to detect fragmented DNA, indicative of cell death. *Tex19.1*$^{-/-}$ placentas had a similar number of TUNEL positive cells as control placentas in the labyrinth layer at both E14.5 and E12.5 (figure 4.3). In the junctional zone of E14.5 placentas there was a significant 1.9 fold increase in the amount of TUNEL positive cells detected in *Tex19.1*$^{-/-}$ compared to control animals (figure 4.3A & C)(n=3, $P<0.05$, Students T-Test one-tailed). At E12.5 the junctional zone of *Tex19.1*$^{-/-}$ placentas had a significant 1.8 fold increase in the number of TUNEL positive cells compared to control placentas (figure 4.3B & D)(n=4, $P<0.01$, Students T-Test one-tailed). This suggests that the reduction in placental weight in *Tex19.1*$^{-/-}$ animals is due to an increase in cell death in the junctional zone. The loss of cells at this mid-gestational developmental time point may be resulting in the decrease in junctional zone area at E18.5 (Reichmann et al., 2013).

Figure 4.3: *Tex19.1*$^{-/-}$ placentas have increased cell death in the junctional zone (figures overleaf). Images of PFA fixed sections of (A) E14.5 *Tex19.1*$^{+/+}$ and *Tex19.1*$^{-/-}$ placentas and (B) E12.5 *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$ placentas processed with the TUNEL assay. White dashed lines indicate the different layers of the placenta. Db, decidua basalis; Jz, junctional zone; Lb, labyrinth; Cp, chorionic plate. TUNEL positive cells are indicated by pointed white arrows, with auto-fluorescent background staining indicated by the square arrows. Negative control images are shown to indicate background fluorescence where no rTdT enzyme was added to the tissue sections. Auto-fluorescence was detected in the labyrinth, therefore only cells with a higher intensity than this background were scored as TUNEL positive**.** Nuclei were stained with DAPI. Scale bars: top=500µm, middle=100µm, bottom=50µm.

A
E14.5
Negative Control
Tex19.1+/+
Tex19.1-/-
Db
Jz
Lb
Cp
B
E12.5
Negative Control
Tex19.1+/-
Tex19.1-/-
Jz

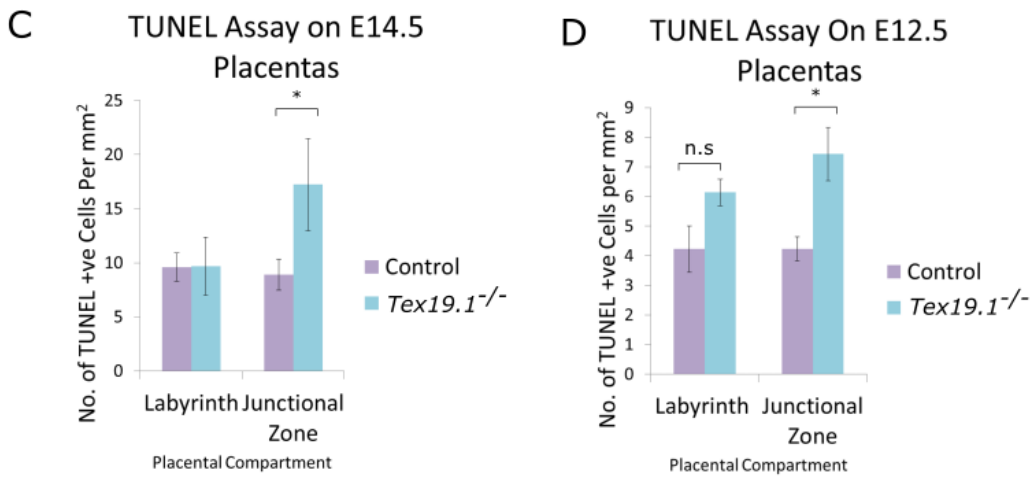


Figure 4.3 (continued): *Tex19.1*$^{-/-}$ placentas have increased cell death in the junctional zone (C) Graph showing the number of TUNEL cells per mm^2 in the labyrinth and junctional zone of control and *Tex19.1*$^{-/-}$ placentas at E14.5 (Control n=4, *Tex19.1*$^{-/-}$ n=3, Students 1 tailed T-Test) and (D) E12.5 (Control n=4, *Tex19.1*$^{-/-}$ n=4)(n.s = Non-significant, *P<0.05 Students 1 tailed T-test). Also see images in Figure 4.3A and 4.3B on previous page.

4.2.4 The junctional zone is significantly reduced in size in *Tex19.1*$^{-/-}$ placentas early in development

Analysis at E18.5 indicated that junctional zone size was reduced in the absence of *Tex19.1* (Reichmann et al., 2013). The modest but significant reduction in placental weight and increased cell death in the junctional zone at E12.5 indicated that placental developmental defects were still occurring at this time point. It was unclear if the increased cell death was resulting in a decrease in junctional zone size and if this could be contributing to the reduction in weight. I therefore measured the maternal decidua and the trophectoderm derived junctional zone and labyrinth areas of PAS stained sections of control and *Tex19.1*$^{-/-}$ placentas. E12.5 control placentas had similar proportions of decidua, labyrinth and junctional zone as has been previously reported (figure 4.4)(Coan et al., 2004). *Tex19.1*$^{-/-}$ placentas had a small but significant 1.3-fold increase in the contribution of the decidua (P=0.03, Students T-test) and a significant

1.5-fold decrease in the contribution of the junctional zone to the whole placenta (*P*=0.02, Students T-test). No change in the size of the labyrinth was detected. These results indicate that *Tex19.1* is necessary for correct development of the junctional zone at these early time points which therefore results in a proportional increase in the contribution of the maternally derived decidua. Together, these data suggest that the decreased junctional zone is the major contributor to the decrease in placental weight of *Tex19.1*$^{-/-}$ animals.

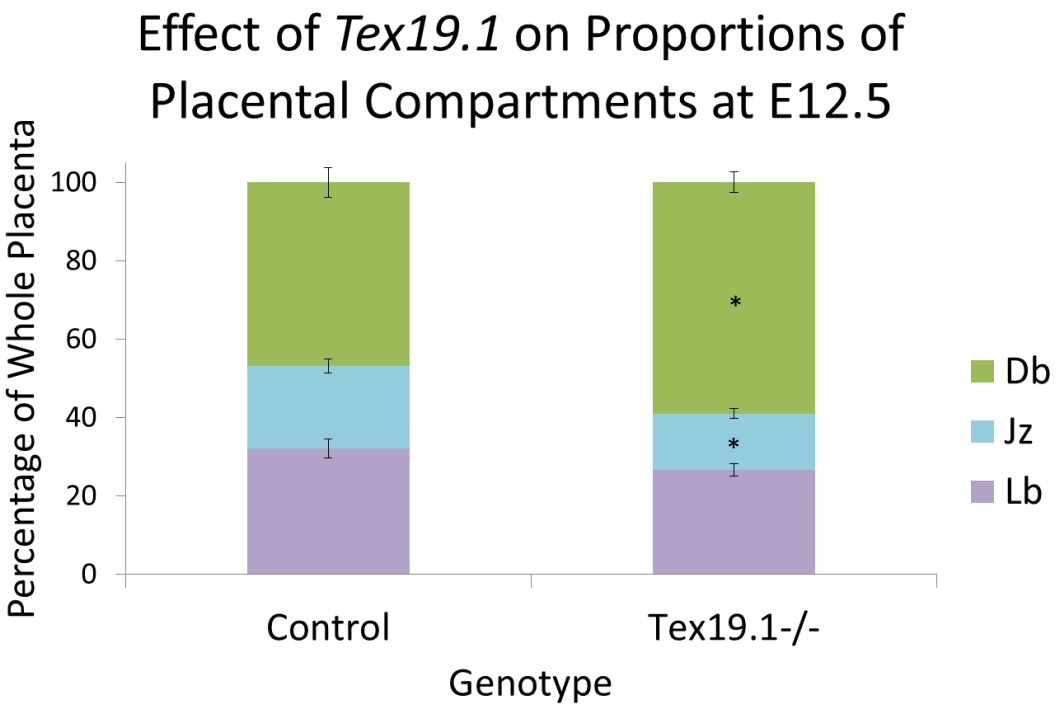


Figure 4.4: The junctional zone is reduced in *Tex19.1*$^{-/-}$ animals at E12.5. Graph showing the proportion contribution of different placental compartments to the whole placenta in E12.5 control and *Tex19.1*$^{-/-}$ animals. Db; Decidua basilis, Jz; junctional zone, Lb; Labyrinth (n=5, **P*<0.05 compared to control animals, Students-T test.)

4.2.5 Trophoblast derived cell types are not reduced in E14.5 *Tex19.1*$^{-/-}$ placentas

At E18.5 there was a reduction in the amount of trophoblast derived S-TGCs, spongiotrophoblast cells and glycogen trophoblast cells in the absence of *Tex19.1* (Reichmann et al., 2013). Despite only a smallincrease in TUNEL cells at E14.5 it was possible that decreased amounts of cells reflective of earlier defects could be contributing to the reduction in placental weight. To investigate if the number of trophoblast derived cells had already been perturbed by E14.5, I visually assessed the number of each cell type on H+E stained placental sections.

Trophoblast giant cells are large mononuclear polyploid cells identified by distinct expression profiles and their spatial location in the murine placenta (Simmons et al., 2007). S-TGCs line the maternal blood spaces in the labyrinth and have been shown to be reduced in the E18.5 *Tex19.1*$^{-/-}$ placenta. I therefore asked whether S-TGCs were reduced in number in E14.5 *Tex19.1*$^{-/-}$ placentas (Reichmann et al., 2013). Strict scoring criteria were used to determine S-TGCs where only cells larger than a defined area and adjacent to an obvious maternal blood space were counted as S-TGCs. *Tex19.1*$^{-/-}$ and control placentas had similar number of S-TGCs per area of labyrinth at E14.5 in contrast to the findings at E18.5 (figure 4.5). This indicates that the reduction in S-TGC number must be occurring after E14.5.

At E18.5, trophectoderm derived syncytiotrophoblast and glycogen trophoblast cells of the junctional zone were also reduced in number in *Tex19.1*$^{-/-}$ animals (Reichmann et al., 2013). The junctional zone is a layer thought to play a role in endocrine signalling via hormone secretion and also has large glycogen stores (Christie, 1967; Soares, 2004; Tunster et al., 2010). To see if there was a reduction in number of these cells at E14.5 I counted the syncytiotrophoblast and glycogen trophoblast cells present in the junctional zone layer of E14.5 control and *Tex19.1*$^{-/-}$ placentas. As with S-TGCs, the number of trophoblast derived cells per area was similar between control and *Tex19.1*$^{-/-}$ placentas at this time point, again suggesting that the reduction in cell number was occurring post E14.5 (figure 4.5).

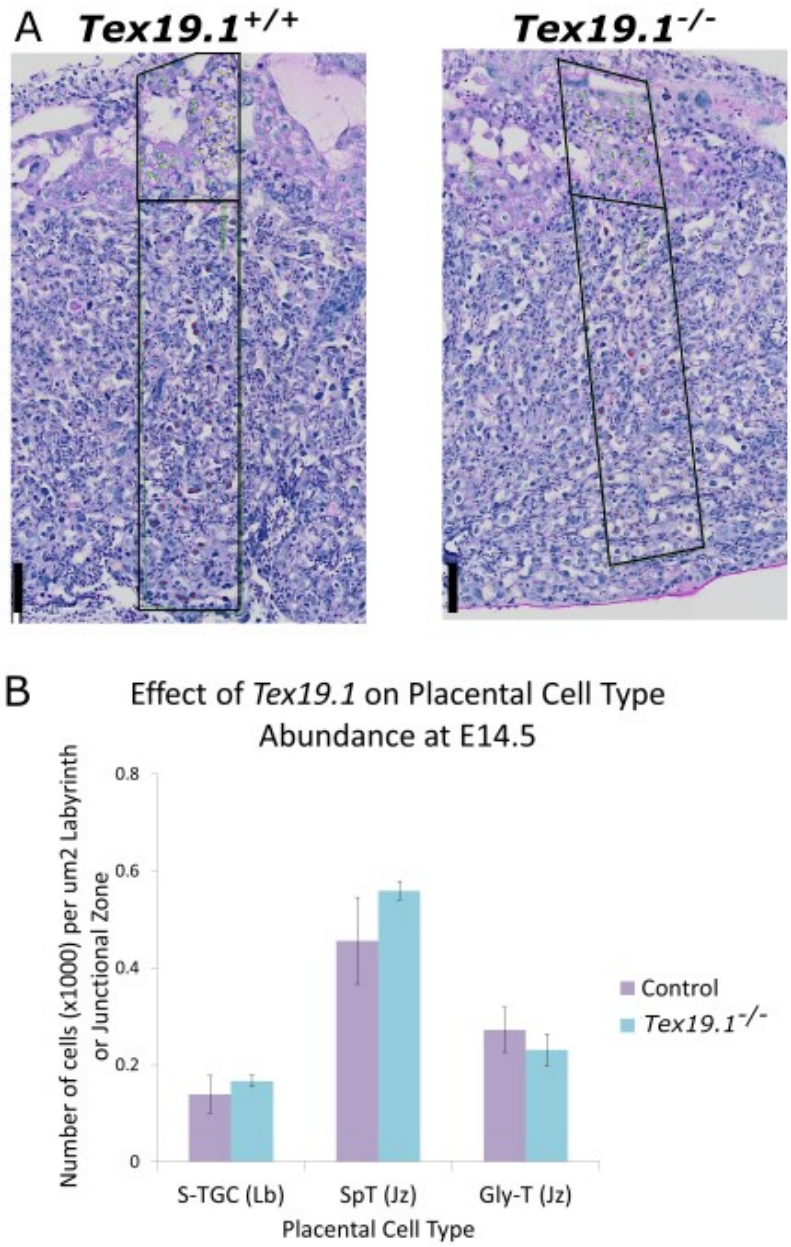


Figure 4.5: E14.5 *Tex19.1*$^{-/-}$ placentas have no reduction in numbers of trophoblast derived cells compared to control animals (A) PAS and haematoxylin stained sections of control and *Tex19.1*$^{-/-}$ E14.5 placentas. The black box represents the area of the placenta where cell types were scored, with the bottom rectangle corresponding to the labyrinth and the top corresponding to the junctional zone. An average cell count from three boxes per placenta was scored for 4 control and 4 *Tex19.1*$^{-/-}$ placentas (B) Graph showing the amount of sinusoidal-trophoblast giant cells (S-TGCs), spongiotrophoblast cells (SpT) and glycogen trophoblast cells (Gly-T) in the labyrinth (Lb) and junctional zone (Jz) per μm^2 placenta. Scale bar=100µm

4.2.6 Gene expression changes in E14.5 *Tex19.1*$^{-/-}$ placentas

Various changes in mRNA abundance were detected by microarray analysis in *Tex19.1*$^{-/-}$ placentas at E18.5 (Reichmann et al., 2013). Therefore, I next investigated various cellular biomarkers of expression at E14.5 that had lower mRNA abundance in E18.5 *Tex19.1*$^{-/-}$ placentas. To investigate the expression of marker genes for various cell types qRT-PCR was performed on whole E14.5 placentas. The general trophoblast derived cell marker *Cdx2*, S-TGC specific gene *Ctsq*, spongiotrophoblast specific gene *Prl8a8* and glycogen trophoblast cell marker *Gjb3* were expressed in control and *Tex19.1*$^{-/-}$ placentas. Significantly lower abundance of mRNA was detected for *Cdx2* and *Ctsq* in *Tex19.1*$^{-/-}$ animals (figure 4.6). As there is no decrease in the number of S-TGCs at E14.5 this suggests these changes may be transcriptional. *Prl8a8* and *Gjb3* showed no difference in abundance of mRNA in *Tex19.1*$^{-/-}$ placentas compared to controls supporting the cell counts from histology in the junctional zone (figure 4.5). Taken together, the large reduction in mRNA abundance for *Cdx2* and *Ctsq* and the fact that no differences in cell composition were detected in E14.5 *Tex19.1*$^{-/-}$ placentas raised the intriguing possibility of a *bona fide* role for *Tex19.1* in transcriptional control of genes in the placenta.

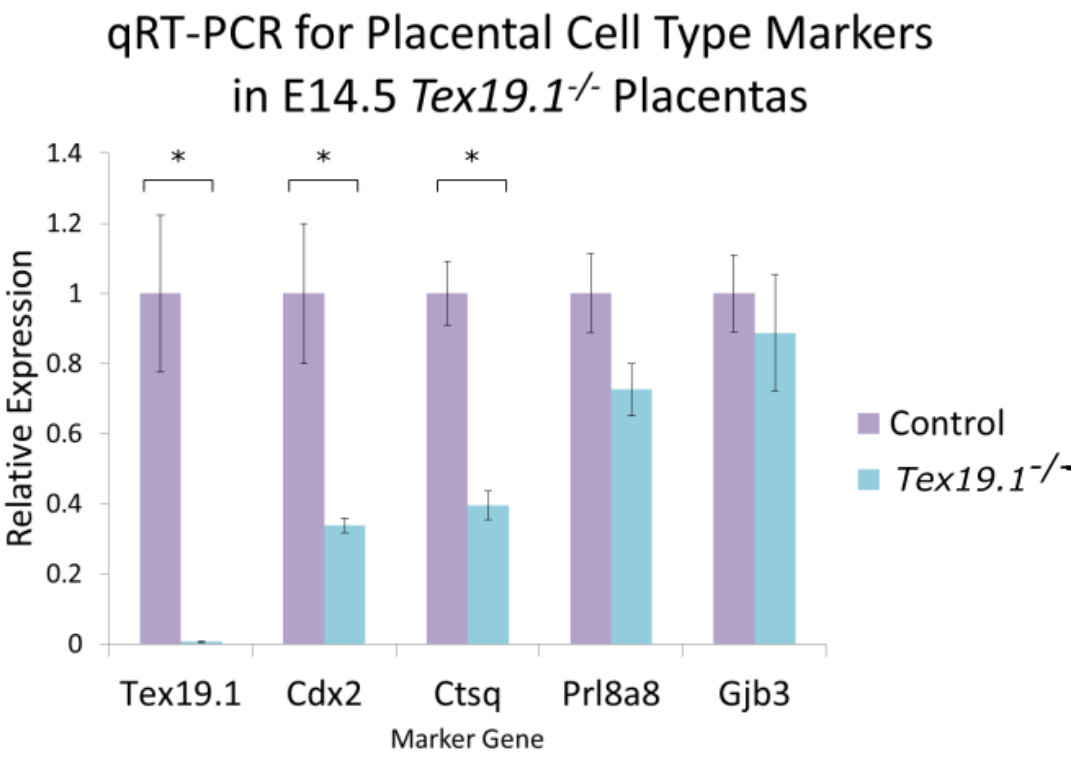


Figure 4.6: Gene expression changes are present in *Tex19.1*$^{-/-}$ placentas at E14.5
qRT-PCR for trophoblast specific cell type marker genes from whole placentas at E14.5. Data is normalised to β actin and expressed relative to control (n=3, **P*<0.05 Students T-test)

4.2.7 LINE-1 retrotransposons are de-repressed at E12.5

One aspect of the E18.5 *Tex19.1*$^{-/-}$ placenta phenotype is the de-repression of retrotransposons (Reichmann et al., 2013). Recent evidence suggests that retrotransposons contribute to the control of placental gene expression and that de-repression of retrotransposons can having consequences on the transcription of neighbouring genes (Chuong et al, 2013)(Goke and Ng, 2016). Other findings implicate retrotransposon nucleic acids in activating the type I interferon response leading to cell death or disease (Chiapinelli et al., 2015)(Roulois et al., 2015)(Stetson et al., 2008). It is therefore possible that retrotransposon de-repression in *Tex19.1*$^{-/-}$ placentas is contributing to the developmental defects in this tissue. To address this, I focussed on the LINE-1 class of retrotransposons as these had the highest level of de-repression in E18.5 *Tex19.1*$^{-/-}$ placentas (Reichmann et al., 2013). It was unclear when this LINE-1 de-repression was occurring in the development of *Tex19.1*$^{-/-}$ placentas and if this correlated with the onset of placental developmental defects. I therefore assessed expression of LINE-1 in control and *Tex19.1*$^{-/-}$ placentas at E10.5, E12.5 and E14.5. Interestingly, no difference in the expression of LINE-1 Orf2 was detected at E10.5 between control and *Tex19.1*$^{-/-}$ placentas (figure 4.7). In contrast, LINE-1 is significantly de-repressed 1.6 fold in E12.5 *Tex19.1*$^{-/-}$ placentas compared to controls (figure 4.7). This correlates with the onset of the placental defects. De-repression was further detected as a 1.7 fold increase at E14.5. The 1.6-fold increase in L1Orf2 expression from E18.5 *Tex19.1*$^{-/-}$ from Reichmann et al., 2013 is shown as a comparison (figure 4.7). Therefore, once these retrotransposons become de-repressed in the absence of *Tex19.1* at around E12.5, re-silencing does not occur through development.

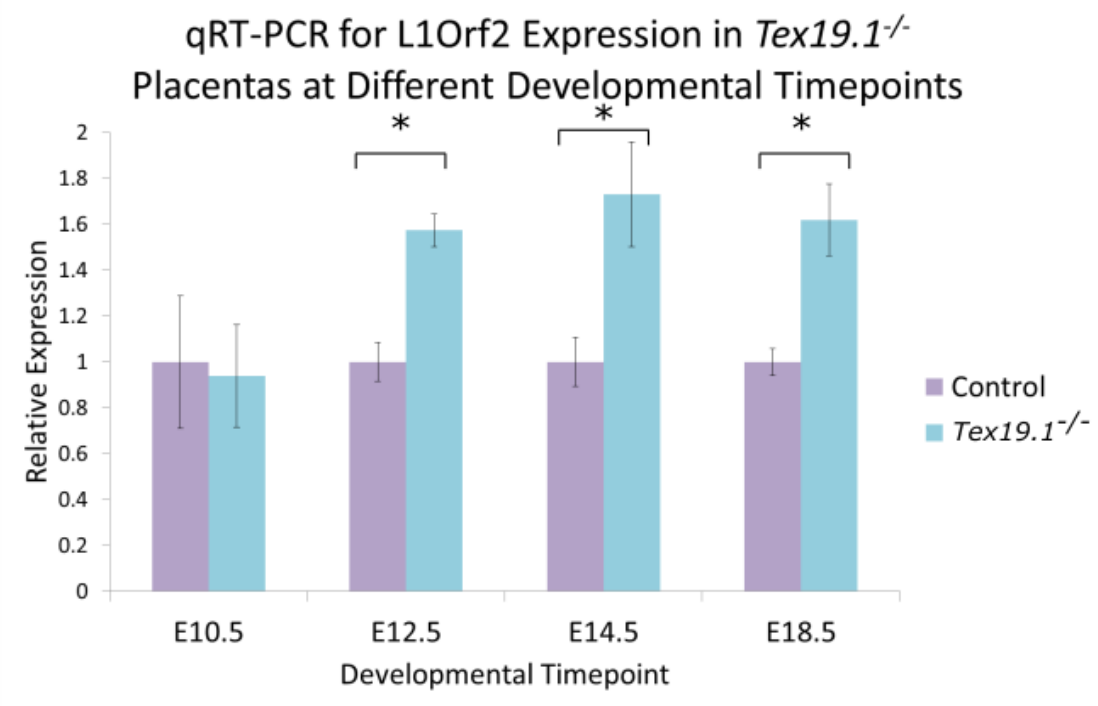


Figure 4.7: LINE-1 is de-repressed in the placenta from E12.5 onwards. qRT-PCR of L1Orf2 expression in placentas from E10.5, E12.5, E14.5 and E18.5. Data is normalised to β actin and expressed relative to control values (n=2, n=5, n=6, n=6 respectively, **P*<0.05 Students T-test)

4.2.8 Different classes of retrotransposons are de-repressed in *Tex19.1*$^{-/-}$ placentas

As I hypothesised that the retrotransposon de-repression at E12.5 may be contributing to placental defects I aimed to determine if retrotransposons in addition to LINE-1 and VL30 were being de-repressed in *Tex19.1*$^{-/-}$ placentas. RNA was isolated from three whole *Tex19.1*$^{-/-}$ placentas and three control littermate placentas (2 *Tex19.1*$^{+/+}$ and 1 *Tex19.1*$^{-/-}$) and sent for paired end, 125bp read length, RNAseq. The RNAseq data was mapped to the mouse genome using TopHat allowing two mismatches, and reads mapping to loci belonging to each class of repeat as defined by Repeatmasker annotation were counted. Multiple mapping reads were randomly assigned to a matching locus. Around 4% of control and *Tex19.1*$^{-/-}$ RNAseq reads mapped to repetitive elements with the highest proportion of reads being from the SINE class of elements (1.8%)(figure 4.8A). LTR elements comprised 1.2% of all mapped reads, with LINE-1 and DNA elements comprising 0.5% and 0.2% of all reads respectively (figure 4.8A). The number of different subfamilies within each class is shown on the pie chart. The proportion of reads for each class of retrotransposon was similar for control and *Tex19.1*$^{-/-}$ placentas.

The placenta is a highly permissive environment for repetitive element de-repression presumably due to it containing hypomethylated components (Reichmann et al., 2013). To determine if any additional repetitive elements were de-repressed in the absence of Tex19.1, I performed differential expression analysis on the repeat mapped RNAseq data from control and Tex19.1 whole placentas. A 0.01 false discovery rate (FDR) cut off was utilised to determine those repetitive elements whose changes in mRNA abundance were likely to be *bona fide* expression alterations. Of the 1099 repetitive element subfamilies with mapped reads, a total of 23 different repetitive elements were significantly upregulated in *Tex19.1*$^{-/-}$ placentas versus control placentas (figure 4.8B & figure 4.9). Two of these corresponded to the LINE-1 family, 18 to the LTR class and 3 to the DNA elements. 25 repetitive elements were significantly downregulated in Tex19.1-/- placentas compared to controls with 4 LINE-1 elements, 20 LTR elements and 1 DNA repeat being repressed (figure 4.8B & figure 4.9).

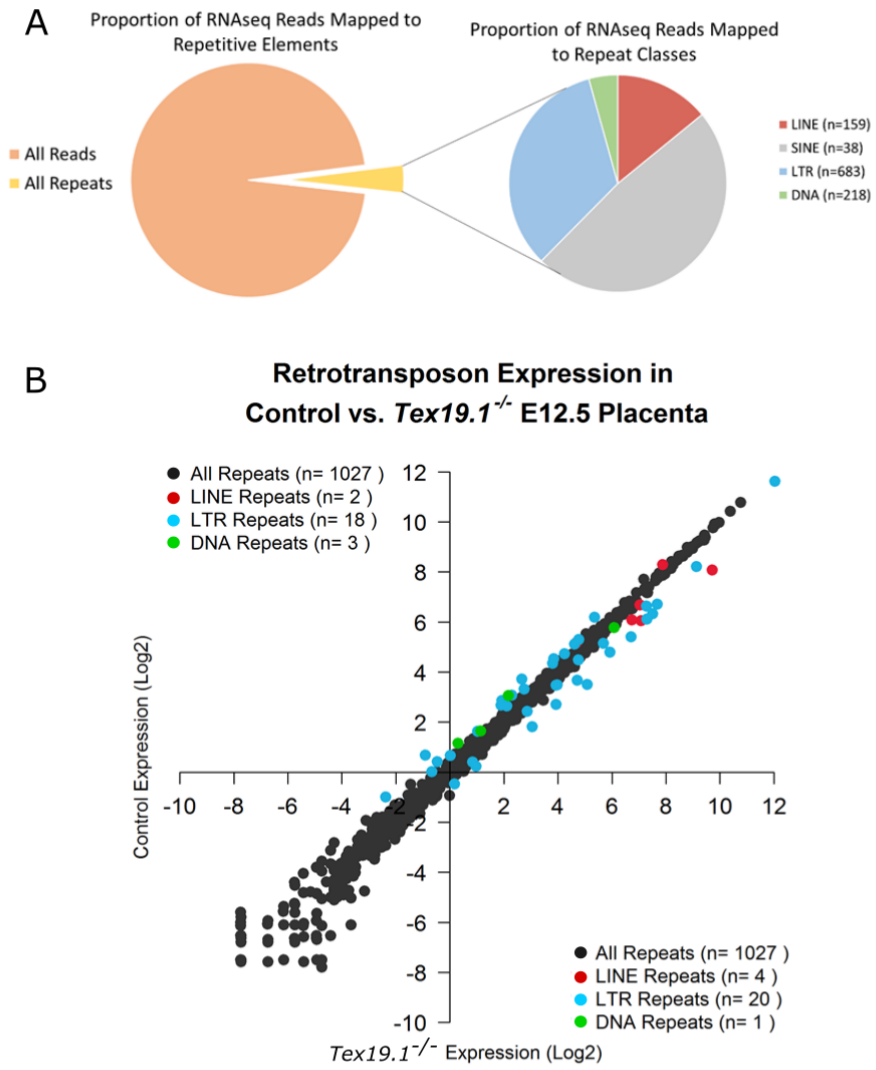


Figure 4.8: Different repetitive elements are de-repressed in E12.5 *Tex19.1*$^{-/-}$ placentas in addition to LINE-1 and VL30. (A) Pie chart illustrating the proportion of all RNAseq reads mapping to repeat and non-repeat regions of the genome. The pie chart on the right represents the proportion of each repetitive class that reads are mapped to with the number of different subfamilies in brackets. (B) Graph showing the average Log_2 expression values (fragments per kilobase of transcript per million mapped reads, FPKM) for repetitive elements in *Tex19.1* littermate control (n=3) and *Tex19.1*$^{-/-}$ placentas (n=3). Each point represents a different repetitive element subfamily and are coloured dependent on their class if their expression is changing significantly (FDR<0.01). The number of repetitive element subclasses are indicated in the legend by triangles signifying elements changing up or down. Black points correspond to repetitive elements with no significant changes in expression between control and *Tex19.1*$^{-/-}$ placentas.

4.2.9 Distinct subfamilies of repeat elements are differentially controlled in *Tex19.1*$^{-/-}$ placentas

I next aimed to determine which subfamilies of elements were being de-repressed in *Tex19.1*$^{-/-}$ placentas. Of the 23 upregulated elements two corresponded to LINE-1 (figure 4.9). Earlier qRT-PCR results from E12.5, indicated that the KAP1 bound, L1MdF2 is significantly de-repressed in *Tex19.1*$^{-/-}$ placentas compared to control placentas (Chapter 3 - figure 3.13). This was confirmed in the placental RNAseq data. The other LINE-1 element corresponds to L1MC and represents a different LINE-1 subfamily potentially controlled by TEX19.1. Another confirmation that our RNAseq repeat analysis approach was working as expected was the detection of the RLTR6_Mm repeat which is the LTR driving expression of the VL30 retrotransposon, which has previously been shown to be de-repressed in *Tex19.1*$^{-/-}$ placentas (figure 4.9)(Reichmann et al., 2013).

Interestingly, the most upregulated element was MLT1A (figure 4.9). This ancient element predates the human-mouse speciation and is present in both species (Smit 1993; Waterston et al. 2002). Recent evidence indicates that human MLT1A elements are associated with placental specific genes and have been proposed to have enhancer functions in a tissue specific manner (Pavlicev et al., 2015). Similar observations have also been made for MER52 (Pavlicev et al., 2015) which is also upregulated in *Tex19.1*$^{-/-}$ placentas. A member of the RLTR13 subfamily is also de-repressed in *Tex19.1*$^{-/-}$ placentas. In mouse trophoblast stem cells, the mouse superfamily of RLTR13 LTR elements play important roles as placental specific enhancers (Chuong et al., 2013). RLTR13D5 acts as enhancers for the core genes of the TSC transcriptional network *Cdx2, Eomes* and *Elf5* (Chuong et al., 2013). The de-repression of the closely related element RLTR13D3 in *Tex19.1*$^{-/-}$ placentas raises the possibility that enhancer function may be being perturbed though development in the absence of *Tex19.1*. Recent studies also indicate LTR elements as alternative promoters for a variety of different genes (Thompson et al., 2016). The permissive environment of the placenta likely allows detection of the de-repression of many different retrotransposons due to reduced DNA methylation and the absence of other genome defence genes (Reichmann et al., 2013).

25 repetitive element subfamilies had reduced mRNA abundance in *Tex19.1*$^{-/-}$ placentas. The reason for this is not clear, however this may either indicate additional defence mechanisms acting in a compensatory way to repress these repetitive elements in the absence of *Tex19.1* or that at least a subset of these elements are controlled by DNA methylation and may reflect alterations in the contribution of the hypomethylated trophectoderm derived cells in *Tex19.1*$^{-/-}$ placentas. The repetitive element with the most reduced abundance of mRNA was Lx4A. In a global screen for KRAB-ZFP repressors of repetitive elements Ecco et al. included an Lx4A sequence in their analysis. No KRAB-ZFP repression of this sequence was detected in mESCs suggesting that this element is unlikely to be repressed by KRAB-ZFPs and may therefore by more susceptible to DNA methylation mediated repression. Also mRNA abundance from IAP elements was significantly reduced in *Tex19.1*$^{-/-}$ animals. IAP is highly expressed in the trophectoderm derived, hypomethylated cells of the junctional zone and labyrinth (Reichmann et al., 2013) which suggests that the reductions in the junctional zone size could be contributing to the reduced IAP mRNA abundance.

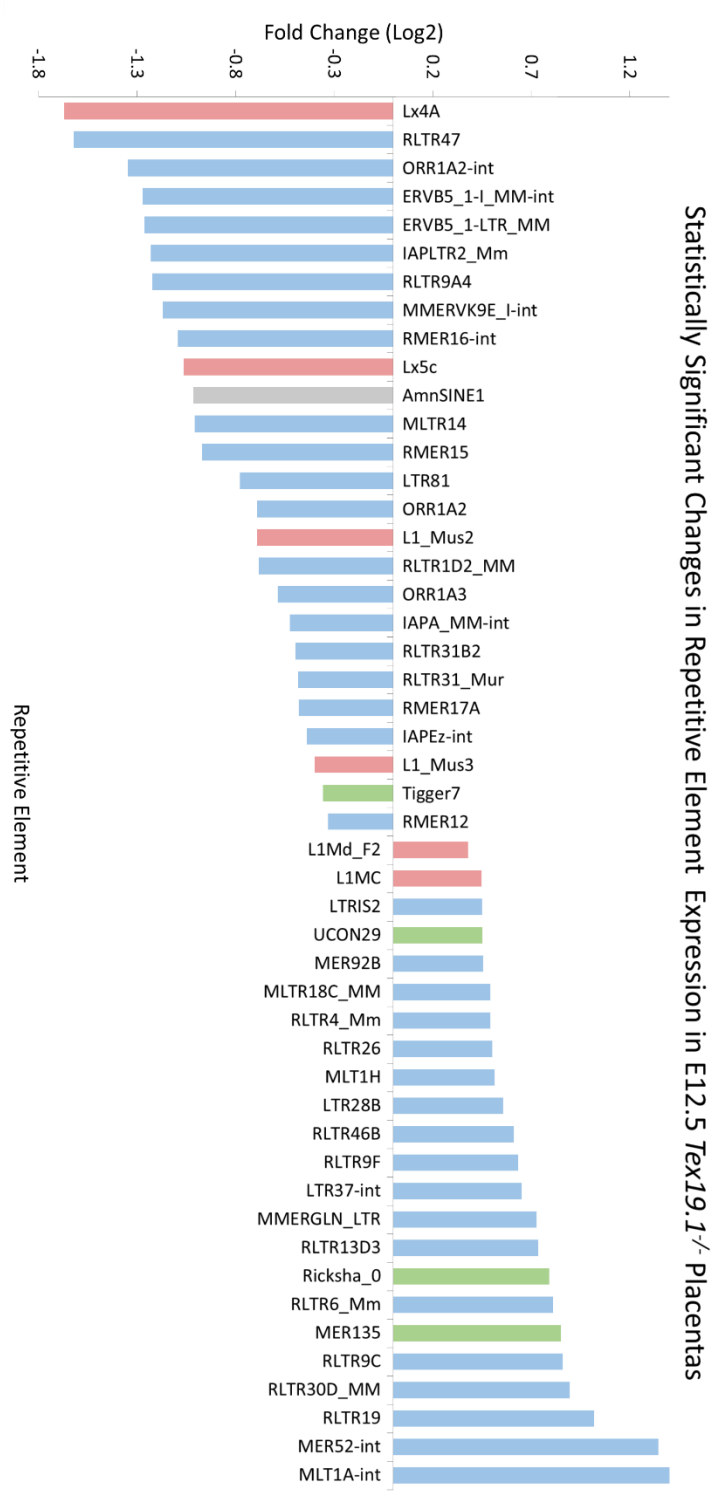


Figure 4.9: Many subtypes of repetitive element are differentially regulated in *Tex19.1*$^{-/-}$ placentas. Plot indicating repetitive elements significantly down and upregulated with FDR<0.01 in *Tex19.1*$^{-/-}$ placentas. Elements are ordered by their Log_2 FC from most downregulated to most upregulated. Classes of repetitive elements are coloured as follows: LINE-1 = red, SINE=grey, LTR=blue, DNA=green.

4.2.10 E12.5 *Tex19.1*$^{-/-}$ placentas have numerous gene expression changes

No retrotransposon de-repression or decrease in placental weight was detected at E10.5. At E12.5 LINE-1 has increased mRNA abundance (figure 4.7, 4.8 & 4.9) but the extent of alterations of mRNA abundance for genic loci was unclear. I therefore focussed specifically on addressing if there were any changes in mRNA abundance in E12.5 *Tex19.1*$^{-/-}$ placentas where retrotransposon de-repression is first detectable.

Paired end reads from RNA sequencing of 3 matched littermate control and *Tex19.1*$^{-/-}$ whole placentas were mapped to the genome using the TopHat read mapper, with mRNA detected for 14118 genes across all biological replicates. A false discovery rate cut off of 0.05 was applied to the data to determine genes with significantly different transcript abundance between control and knockout placentas. An initial arbitrary threshold of 2-fold change in mRNA abundance was applied to the significantly differentially expressed genes. This resulted in the identification of 77 genes with significantly increased RNA abundance and 65 genes with significantly reduced RNA abundance (figure 4.10A). These will be referred to as 'strongly differentially expressed genes'. The subsets of genes being differentially expressed is discussed further in from sections 4.2.17. It is possible that some of these changes may reflect changes in cell composition in the *Tex19.1*$^{-/-}$ placenta compared to control placentas, but for at least some genes such as the S-TGC marker gene *Ctsq,* these changes are likely to be transcriptional.

4.2.11 Analysis of all significantly up and downregulated genes

The strict 2-fold change in expression was an arbitrary cut off to determine those subsets of genes which were changing the most, meaning that some of the genes changing expression to a lesser extent would have been missed. Because of this, I determined how many genes were significantly up and down regulated, regardless of the extent of the changes. This resulted in 731 genes significantly upregulated and 599

genes significantly downregulated (figure 4.10B) and are discussed further from section 4.2.17.

At E18.5 it was shown that mRNAs of imprinted genes are not specifically altered in *Tex19.1*$^{-/-}$ placentas. To investigate if this was the case at E12.5, I determined how many of the 151 imprinted genes in the mouse genome (imprinted gene list from: http://www.har.mrc.ac.uk/research/genomic_imprinting/) were significantly changing in *Tex19.1*$^{-/-}$ compared to control placentas. Similar proportions of imprinted genes (11/151) and non-imprinted genes (1319/12648) showed significantly different levels of mRNA in *Tex19.1*$^{-/-}$ placentas compared to controls (χ^2-test, P=0.366)(Data not shown). This indicates that imprinted genes are not preferentially affected in *Tex19.1*$^{-/-}$ placentas.

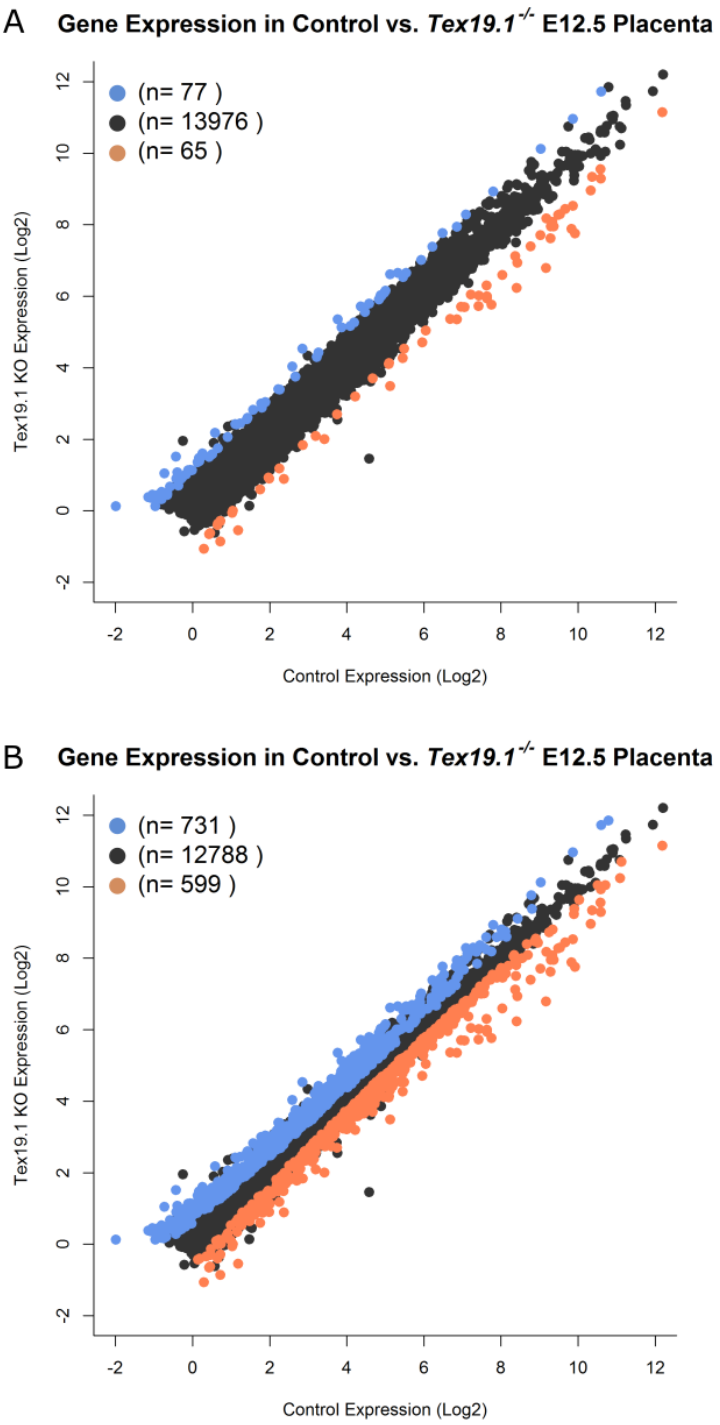


Figure 4.10: RNAseq gene expression analysis from control and *Tex19.1*$^{-/-}$ placentas
Graphs showing the average Log_2 expression values for *Tex19.1* littermate control (n=3) and *Tex19.1*$^{-/-}$ placentas (n=3). Points coloured blue or orange correspond to upregulated or downregulated genes changing greater than (A) 2 fold compared to control and (B) no fold change cut off for genes with a false discovery rate of less than 0.05. Black points correspond to genes with no significant changes between control and *Tex19.1*$^{-/-}$ placentas.

4.2.12 Full length LINE-1 elements preferentially flank strongly downregulated genes

As the *Tex19.1*$^{-/-}$ placentas have numerous gene expression changes, along with retrotransposon de-repression I aimed to explore the possibility that de-repression of retrotransposons may be contributing to placental gene expression changes. When utilising the UCSC genome browser to view genes with significantly increased and decreased mRNA in *Tex19.1*$^{-/-}$ placentas, it appeared that placental genes significantly reduced in expression such as *Ctsq* (figure 4.6 & 4.10) were flanked by large regions of LINE-1 elements (figure 4.11A). I therefore initially focussed analysis on LINE-1 elements.

As many LINE-1 elements are 5' truncated and likely transcriptionally incapable (Waterston et al., 2002), I generated a dataset containing coordinates of LINE-1 elements greater than 5kb in length, to enrich for those most likely to be full length (~6kb). This resulted in a dataset containing 15,141 'full length' LINE-1 elements. 3951 L1MdF2, 5059 L1MdT and 3304 L1MdA elements were identified in the total full length LINE-1 dataset compared with the older more degenerate LINE-1 families such as L1Mus2 (n=350) and L1Mus3 (n=451). Because of this, further analysis was focussed on L1MdF2, L1MdT and L1MdA. I have also shown previously that L1MdF2 elements are significantly de-repressed in *Tex19.1*$^{-/-}$ placentas (Chapter 3 - figure 3.13 and figure 4.9). It is also of note that significant 1.3-fold de-repression of L1MdT (FDR=0.04) elements also occurs which was not reported in figure 4.9 as it was above the 0.01 FDR cut off.

In order to determine if the 13976 non-changing or 77 upregulated and 65 strongly differentially expressed downregulated genes (2 fold change, FDR<0.05)(figure 4.10A) showed any enrichment for either of these full length LINE-1 elements in the regions flanking their transcriptional start sites (TSS), TSS coordinates from placenta RNAseq and full length LINE-1 datasets were intersected with BedTools. Some genes had multiple full length LINE-1 elements either upstream, downstream or both, however for the initial analysis, only the presence or absence of a full length LINE-1 element was scored. No cut off for distance from the neighbouring gene was used with the largest distance from a gene being 1.1Mb.

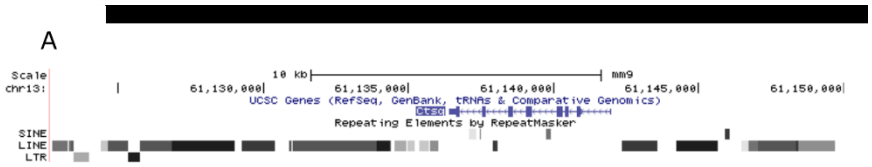
I initially focussed on the L1MdF2 element. 3.9% of upregulated genes and 2.9% of non-changing genes had an L1MdF2 element either upstream or downstream of their TSS. Surprisingly, 15.4% of downregulated genes were flanked by L1MdF2 (P=0.00002, Fishers Exact Test) (figure 4.11B). This suggests that there is an inverse correlation between the presence of a full length L1MdF2 element and the downregulation of its neighbouring gene. I then aimed to determine if there was an association of the different gene groups with the L1MdT elements. A similar pattern was observed, whereby 4%, 9% (P=0.03, Fishers Exact Test) and (P=0.0003, Fishers Exact Test) of non-changing, upregulated and downregulated genes were flanked by L1MdT elements respectively (figure 4.11B). It is important to note that the 15.4% of downregulated genes flanked by L1MdT were different to the 15.4% of genes flanked by L1MdF2. This indicated that there is a significant association of both upregulated and downregulated genes with full length L1MdT elements. L1MdA, which has no significant de-repression in *Tex19.1*$^{-/-}$ placentas, flanks 2.7%, 2.6% and 7.7% (P=0.03, Fishers Exact Test) of non-changing, upregulated and downregulated genes respectively (figure 4.11B). This suggests that there is a weaker but still significant enrichment of full length L1MdA elements in downregulated gene flanks. *Psg29* was the only gene to be proximal to each of the three subtypes. Therefore, there is a significant association of downregulated gene with full length LINE-1 elements of different subtypes.

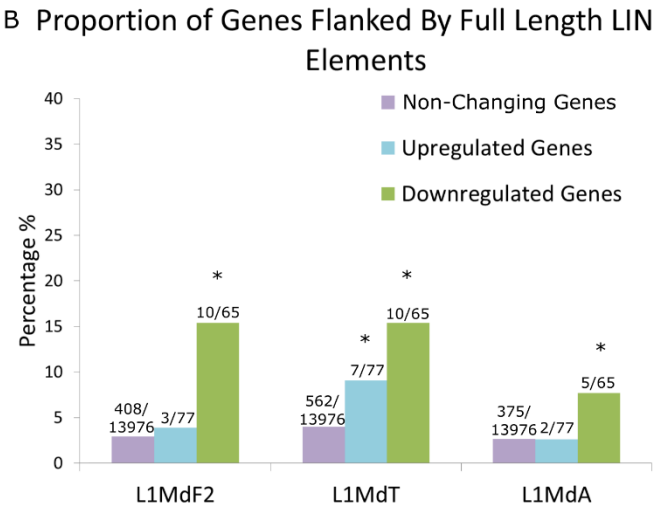
The spatial relationship of this association was unclear, I therefore next aimed to determine the position of the full length LINE elements up or downstream relative to its neighbouring gene. Around half of non-changing, strongly differentially upregulated and downregulated genes flanked by full length L1MdF2, L1MdT or L1MdA have the element upstream of the TSS and half have the element downstream (data not shown). Together these data indicate an association of gene downregulation and full length LINE-1 families which are de-repressed in *Tex19.1*$^{-/-}$ placentas, both up and downstream of the downregulated TSS.

4.2.13 L1MdF2 elements are in close proximity to their neighbouring downregulated genes

I next aimed to determine if the transcriptional start sites of strongly differentially expressed genes in *Tex19.1*$^{-/-}$ placentas are closer to full length LINE-1 elements than non-changing genes expressed in the placenta. Only genes associated with full length LINE-1 elements determined in the previous analysis (figure 4.11A) were used for this analysis. The average distance from TSS was calculated regardless of LINE-1 location up or downstream. The genes not changing in expression were an average distance of approximately 50kb from each subfamily of LINE-1 (figure 4.11C). Both upregulated and downregulated genes were an average distance of around 35kb from the L1MdF2 elements, however it is important to note that the number of genes included in this analysis is relatively small for the upregulated, LINE-1 flanked gene (upregulated n=3, downregulated n=10). This was the shortest distance for any subfamily tested, as up and downregulated genes were an average of 40kb and 60kb from L1MdT elements respectively. L1MdA was an average of 78kb and 30kb from nearby upregulated and downregulated genes respectively (figure 4.11C). The small number of genes present in these categories precluded statistical analysis. This indicates that downregulated genes are not only significantly enriched in L1MdF2, but they are also in close proximity. This association possibly suggests a detrimental role for L1MdF2 de-repression perturbing placental gene expression in the absence of *Tex19.1,* which in turn may be contributing to placental developmental defects.

I next aimed to determine if there was a relationship between the extent of gene expression alterations and the distance from full length L1MdF2 elements. The fold change of each significantly differentially expressed gene was plotted against distance to its neighbouring L1MdF2 element. This indicated a very small correlation with R^2 value of 0.18 for strongly differentially expressed downregulated genes (n=12) compared to all non-significantly changing genes with an R^2 of 0.0057 (n=519)(Data not shown). The correlation for upregulated genes flanked by L1MdF2 could not be performed due to only 3 upregulated genes being flanked by L1MdF2. Together these data suggest that any contribution of full length L1MdF2 de-repression on nearby gene expression is likely only minor.


A
Scale
chr13:
10 kb
mm9
61,130,000
61,135,000
61,140,000
61,145,000
61,150,000
UCSC Genes (RefSeq, GenBank, tRNAs & Comparative Genomics)
Repeating Elements by RepeatMasker
SINE
LINE
LTR


B Proportion of Genes Flanked By Full Length LINE-1 Elements
Non-Changing Genes
Upregulated Genes
Downregulated Genes
Percentage %
40
35
30
25
20
15
10
5
0
*
*
*
10/65
10/65
5/65
7/77
408/ 13976
3/77
562/ 13976
375/ 13976
2/77
L1MdF2
L1MdT
L1MdA

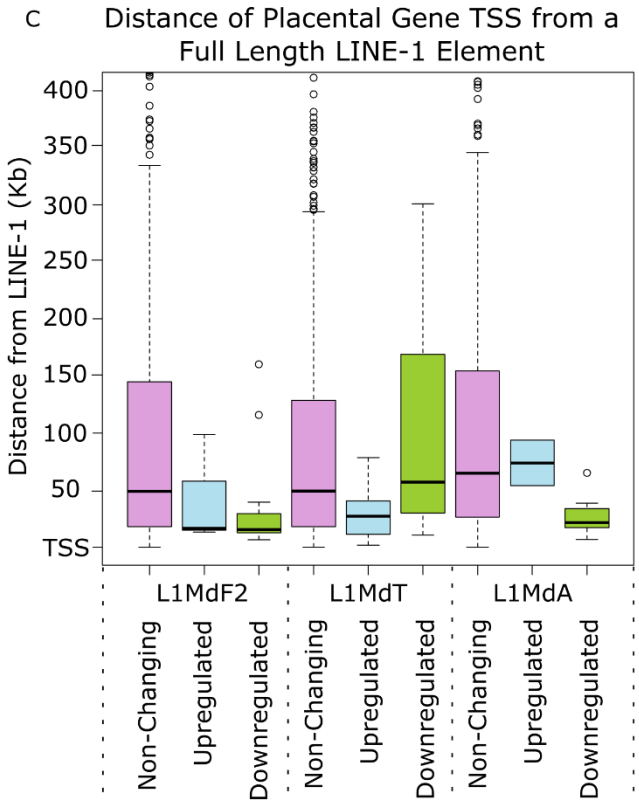

C Distance of Placental Gene TSS from a Full Length LINE-1 Element
Distance from LINE-1 (Kb)
400
350
300
250
200
150
100
50
TSS
L1MdF2
L1MdT
L1MdA
Non-Changing
Upregulated
Downregulated
Non-Changing
Upregulated
Downregulated
Non-Changing
Upregulated
Downregulated

Figure 4.11: Downregulated genes are flanked by increased numbers of full length LINE-1 elements compared to non-changing genes in *Tex19.1*$^{-/-}$ placentas, with no up or downstream bias (A) UCSC browser track indicating *Ctsq* (blue line) and the large amount of flanking LINE-1 sequences (black/grey bars). The lighter shading reflects an increase in of base mismatch, base deletion, and base insertion associated with a repeat element compared to consensus sequence. (B) Bar chart showing the percentage of all placental genes (purple, FDR >0.05), significantly upregulated genes (blue, Log2FC>1, FDR<0.05) and significantly downregulated genes (green, Log2FC<-1, FDR<0.05) which have at least one neighbouring full length LINE-1 element. The number of LINE-1 flanked genes is shown above each bar. The subfamily of LINE-1 element is shown below the bars. (**P*<0.05, Fishers exact test) (C) Boxplots showing the distance of non-changing, upregulated and downregulated genes in *Tex19.1*$^{-/-}$ placentas from full length LINE-1 elements (statistical testing was not performed on these data due to insufficient numbers for a Mann Whitney U-Test).

4.2.14 Antisense transcripts are detectable from L1 promoters but are not enriched in *Tex19.1*$^{-/-}$ placentas

I have shown in the previous analyses that highly differentially downregulated genes are enriched and closer to full length L1MdF2 elements, however the functional relevance of this association is unclear. In other mouse mutants, retrotransposon de-repression can have effects on neighbouring genes. For example, mouse hepatocytes null for *Trim24* exhibit retrotransposon de-repression which results in long non-coding RNAs likely altering nearby gene expression (Göke and Ng, 2016; Herquel et al., 2013). Interestingly, the *Psg* gene cluster has 5 highly significantly down regulated genes, 3 of which have full length L1MdF2 elements in the antisense orientation less than 5kb from the neighbouring TSS (figure 4.12A&B). The mouse LINE-1 element initiates an antisense transcript from the L1ORF1 protein (Li et al., 2014) which could potentially interfere with transcription of nearby genes. I therefore asked if an antisense transcript was increased in *Tex19.1*$^{-/-}$ knockout placentas from L1MdF2 elements flanking neighbouring *Psg* genes.

qRT-PCR for L1 antisense transcripts was performed on cDNA from control and *Tex19.1*$^{-/-}$ placentas. Robust amplification was detected for L1MdF2 antisense transcripts upstream of *Psg18* and *Psg28* in control animals, indicating that these elements are transcriptionally active in normal circumstances (figure 4.12C). These antisense transcripts were detected to the same extent in *Tex19.1*$^{-/-}$ placentas and are therefore unlikely to be affecting *Psg18* and *Psg28* gene expression in *Tex19.1*$^{-/-}$ animals (figure 4.12C). No transcript was detected for the L1MdF2 element upstream of *Psg21* in either control or *Tex19.1*$^{-/-}$ placentas. Together these results show that flanking LINE-1 elements are transcriptionally active in this gene cluster, however antisense transcripts are unlikely to be directly interfering with neighbouring gene expression at the loci analysed. The mechanism of disruption of gene expression by L1MdF2 elements, if any, remains to be elucidated and warrants further investigation.

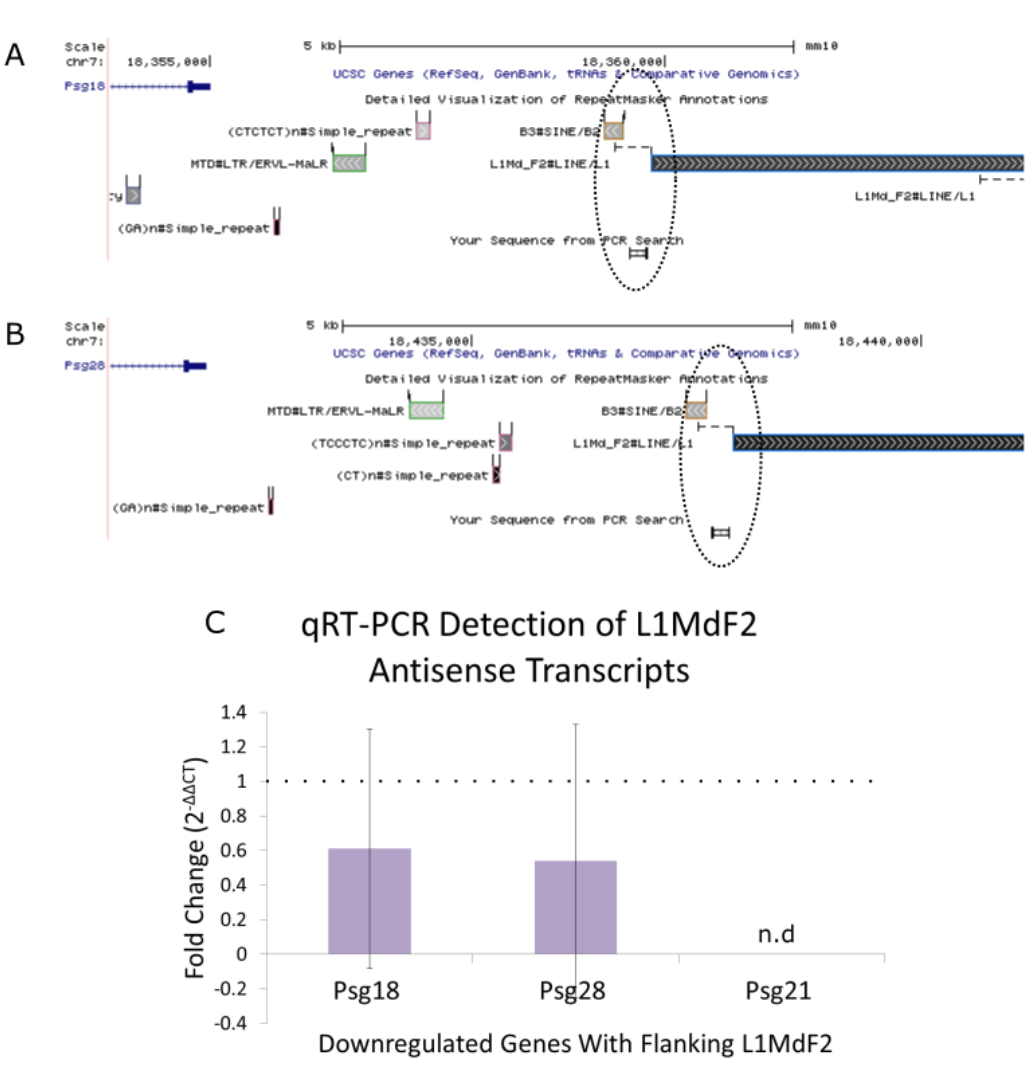


Figure 4.12: L1MdF2 antisense transcripts are present in control and *Tex19.1*$^{-/-}$ placentas to a similar extent. (A & B) UCSC genome browser image of the repeat genomic environment upstream of *Psg18* and *Psg28* (blue bars). The large black rectangle on the right indicates a full length L1MDF2 element on the sense strand with arrows indicating its transcriptional direction. The dotted oval indicates the location of primer pairs to detect antisense transcripts for the qRT-PCR reaction. These lie upstream of the L1MdF2 TSS and between a SINE B3 element in both examples. Chromosome and genomic coordinates are shown at the top of each chart below a 5kb scale bar. Other small repeat regions are indicated by additional rectangles. (C) qRT-PCR for L1MdF2 antisense transcripts using primers shown in A and B. Fold change relative to wild type is shown following the ΔΔCT method of quantification with control levels indicated by a horizontal dotted line (n=3 control and n=3 *Tex19.1*$^{-/-}$ placentas)(n.d=not detected).

4.2.15 LTR elements are not enriched in flanking placental gene regions

LTR elements have been observed to be controlling gene expression in a variety of trophoblast and placental contexts and have recently established roles in controlling transcriptional networks beyond pluripotency (Chuong et al., 2013; Ecco et al., 2016; Göke and Ng, 2016). 18 different subfamilies of LTR repeats are de-repressed in *Tex19.1*$^{-/-}$ placentas, I therefore hypothesised that this de-repression may be having consequences for placental gene expression, as with LINE-1 de-repression. I therefore aimed to determine if the LTR retrotransposons with the highest expression level and significant de-repression in *Tex19.1*$^{-/-}$ placentas are associated with up or downregulation of neighbouring genes.

RLTR6, the LTR flanking VL30 elements, has a significant 1.8-fold increase in expression and the highest expression level for any LTR de-repressed in *Tex19.1*$^{-/-}$ placentas compared to control (figure 4.9). RLTR4, RLTR13D3 and RLTR19 were all also analysed for this association with nearby differentially expressed genes. All individual elements from these subfamilies were extracted from the RepeatMasker dataset and identified 237 RLTR6, 239 RLTR4, 339 RLTR13D3 and 944 RLTR19 elements throughout the genome. These coordinates were intersected with genome TSS coordinates for all genes expressed in our placental RNAseq dataset to determine if non-changing, all significantly upregulated or all significantly downregulated genes were associated with de-repressed LTR retrotransposons. The presence or absence of a flanking LTR element was scored regardless of whether multiple LTRs were flanking a gene. For most assessed LTR elements, there was no significant association of significantly differentially expressed genes with flanking LTR loci (figure 4.13A). Interestingly, downregulated genes were significantly enriched for RLTR4 in the regions flanking the TSS suggesting possible regulatory roles for this element in gene downregulation (P=0.02, Fishers Exact Test)(figure 4.13A).

There was generally no up or downstream bias for LTR positioning relative to the TSS. Another de-repressed retrotransposon in *Tex19.1*$^{-/-}$ placentas, RLTR13D3, was also not significantly enriched at either upregulated or downregulated genes, therefore the functional relevance of these RLTR associations remain to be seen. As these were

detected at only a small number of loci, they are unlikely to represent a major mechanism contributing to the changes in RNA abundance in *Tex19.1*$^{-/-}$ placentas. Overall, these data suggest that there is little association of these de-repressed LTR elements with increases or decreases in genic transcription and these are unlikely to be driving large proportions of the changes in RNA abundance in the RNAseq from *Tex19.1*$^{-/-}$ placentas.

4.2.16 RLTR6 elements are close to upregulated genes

I next aimed to investigate if there were any differences in the distance of each subfamily of element from each flanked gene. This was performed regardless of up or downstream orientation. Non-changing gene TSS's were an average distance of 80kb from an RLTR6 element (4.13B). Interestingly this was highly reduced to only 19kb for the 10 upregulated genes with associated RLTR6 elements, whilst for downregulated genes RLTR6 elements could be found an average of 175kb away for the 3 genes flanked (4.13B). This indicates that upregulated genes are flanked by RLTR6 far closer than would be expected from the non-changing gene category. The non-changing, upregulated and downregulated genes associated with RLTR13D3 had the element similar distances away. In contrast, upregulated genes flanked by RLTR4 had the element less than 2kb away, however this was only two RLTR4 elements which represented the 5' and 3' LTR of a single LTR element. Interestingly, where non-changing genes are flanked by RLTR19 these are an average of 100kb away. In contrast, upregulated genes with RLTR19 associations are far closer to the element at 45kb away (figure 4.13B). As with the LINE-1 analysis the small amount of data has precluded statistical analysis. Functional analysis of these associations will need to be performed to elucidate the possible relevance of these findings. The lack of up or downregulated genes being significantly enriched in LTR elements up or downstream of their TSS suggests that major gene expression changes are likely not occurring as a result of the de-repression of the RLTR elements analysed, despite reduced distances to TSS for some LTRs.

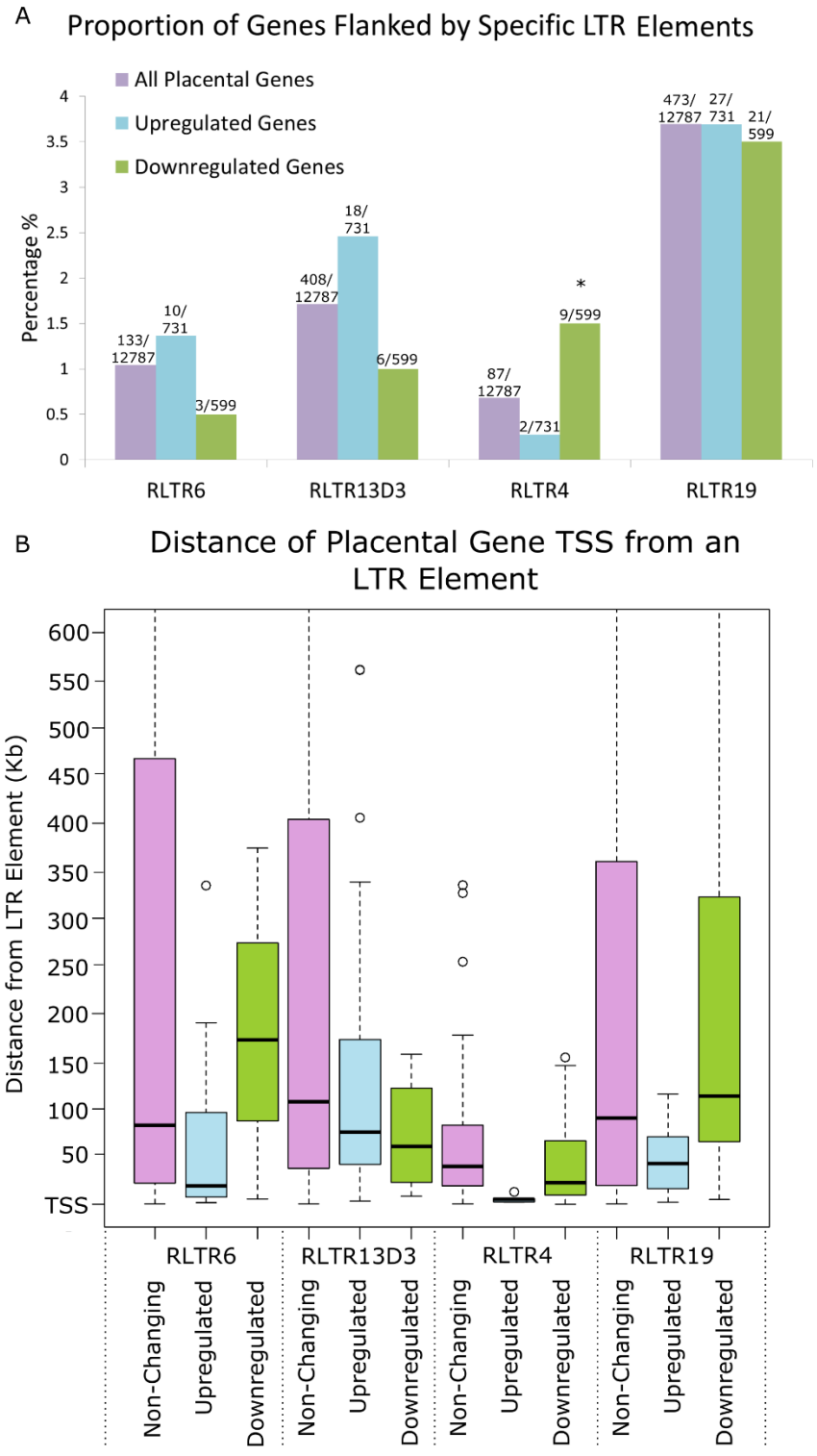


Figure 4.13: No enrichment is detected for LTR elements flanking up or downregulated compared to non-changing genes in *Tex19.1*$^{-/-}$ placentas (A) Bar chart showing the percentage of non-changing genes (purple, FDR >0.05), significantly upregulated genes (blue, Log2FC >0, FDR<0.05) and significantly

downregulated genes (green, Log2FC>0, FDR<0.05) which have at least one neighbouring LTR element. The number of LTR flanked genes is shown above each bar. The subfamily of LTR element is shown below the bars. *P<0.05 Fishers Exact Test. (B) Barchart indicating the distance of a genic TSS from a significantly de-repressed LTR family. If a gene is flanked by multiple full length LTR elements these are included in the data. The number of flanked genes and the number of LTR elements in the downstream region are indicated above the bars in (A).

4.2.17 Gene ontology analysis of genes changing mRNA abundance in *Tex19.1*$^{-/-}$ placental RNAseq

To gain a better appreciation of the general roles of the genes altered in the E12.5 *Tex19.1*$^{-/-}$ placentas, PANTHER gene ontology analysis was performed on the 77 upregulated and the 65 downregulated strongly differentially expressed genes. For the 65 genes with decreased mRNA abundance, the only GO terms significantly enriched in *Tex19.1*$^{-/-}$ placentas were hormone activity and receptor binding (40-fold enrichment, P=4.83x10^{-14}; 4-fold enrichment, P=0.00412 respectively) indicating that a decrease in hormonal gene expression and signalling in the *Tex19.1*$^{-/-}$ placentas may be contributing to the defects in placental development. No GO terms were significantly enriched in the 77 significantly 2-fold increased genes and unlike the genes with reduced mRNA abundance, no obvious gene clusters were being perturbed. The genes significantly reduced over 2 fold largely corresponded to the prolactin (*Prl*), cathepsin and pregnancy specific glycogen (*Psg*) families. The 23 *Prl* genes clustered on chromosome 13 each have distinct temporal and spatial expression patterns largely restricted to spongiotrophoblast and glycogen trophoblast cells of the junctional zone (Simmons et al., 2008). The cathepsin gene cluster comprises 8 genes expressed in a variety of distinct and overlapping cell types. *Ctsm* and *Ctsq* are downregulated and expressed in the spongiotrophoblast, and sinusoidal trophoblast giant cells respectively (Bode et al., 2005). *Psg* genes were also reduced and are also found in the spongiotrophoblast layer. Many of the significantly down regulated genes had been identified previously at E18.5 and possibly reflect compositional changes present at this earlier time point (Reichmann et al., 2013). However, as *Ctsq* RNA abundance is reduced in *Tex19.1*$^{-/-}$ placentas at E14.5 in the absence of any change in the number of

S-TGCs (figure 4.5 & 4.6) it is possible that some of these changes in RNA abundance reflect changes in gene transcription.

To determine any enrichment in GO categories missed by using only the strongly differentially expressed genes, I re-ran all 731 significantly upregulated genes through the PANTHER software. This revealed 23 GO categories significantly enriched $P<0.01$ (figure 4.14A). The top significantly enriched GO category was 'positive regulation of apoptotic process' (3.16 fold enrichment, $P=2.01\text{x}10^{-8}$) which supported the increase in TUNEL positive cells in *Tex19.1*$^{-/-}$ placentas at E12.5, confirming that cells are likely dying via the apoptotic pathway (figure 4.14A & 4.3). A variety of other GO terms related to placental function were significantly enriched in the upregulated gene category, giving insights into the possible reasons for placental dysfunction in *Tex19.1*$^{-/-}$ animals. Taken together, 'regulation of angiogenesis' (3.8 fold enrichment, $P=0.000379$), 'negative regulation of vasculature development' (5.47 fold enrichment, $P=0.00388$) and 'angiogenesis' (2.13 fold enrichment, $P=0.0044$) all indicate that there may be problems with the development of blood vessels and blood supply within the placenta (figure 4.14A). A study analysing gene expression in the maternal decidua indicated that GO terms associated with 'angiogenesis and blood vessel development' are enriched in the decidua which would suggest that these results may partially reflect the compositional differences in *Tex19.1*$^{-/-}$ placentas (Knox and Baker 2008). However, defects in correct vasculature of the placenta has been shown before to contribute to IUGR due to hypoxia as a result of reduced blood supply and nutrient transport to the fetus (Cetin and Alvino, 2009).

Additional significantly enriched categories could be largely grouped into immune response categories. 'Innate immune response' (2.7 fold enrichment, $P=0.000768$), 'positive regulation of immune system process' (2.3 fold enrichment, $P=0.00224$) and 'regulation of immune effector process' (2.78 fold enrichment, $P=0.00458$) all indicated that an immune response may be occurring in the *Tex19.1*$^{-/-}$ placentas, however these categories have again been indicated as a property of the maternal decidua. Interestingly however, one of the most highly enriched and significant categories was 'cellular response to interferon beta' which had a 12.61 fold enrichment ($P=1.81\text{x}10^{-5}$)(figure 4.14A). 'Cellular response to interferon beta' was not reported as enriched in the maternal decidua of E12.5 placentas (Knox and Baker, 2008),

indicating that a *bona fide* immune response may be occurring in *Tex19.1*$^{-/-}$ placentas, possibly leading to the increased apoptosis and placental defects.

Far fewer terms were enriched in the significantly reduced mRNA category from 'anion binding' (1.8 fold enrichment, P=5.2x10^{-5}), 'hormone activity' (5.5 fold enrichment, P=0.00024) and 'Ras guanyl-nucleotide exchange factor activity' (5.3 fold enrichment, P=0.00024). These categories likely indicate a major cause of placental dysfunction in *Tex19.1*$^{-/-}$ animals as disruptions in the levels of hormones of the placenta are known to have consequences for placental and fetal development (John, 2013)(figure 4.14B).

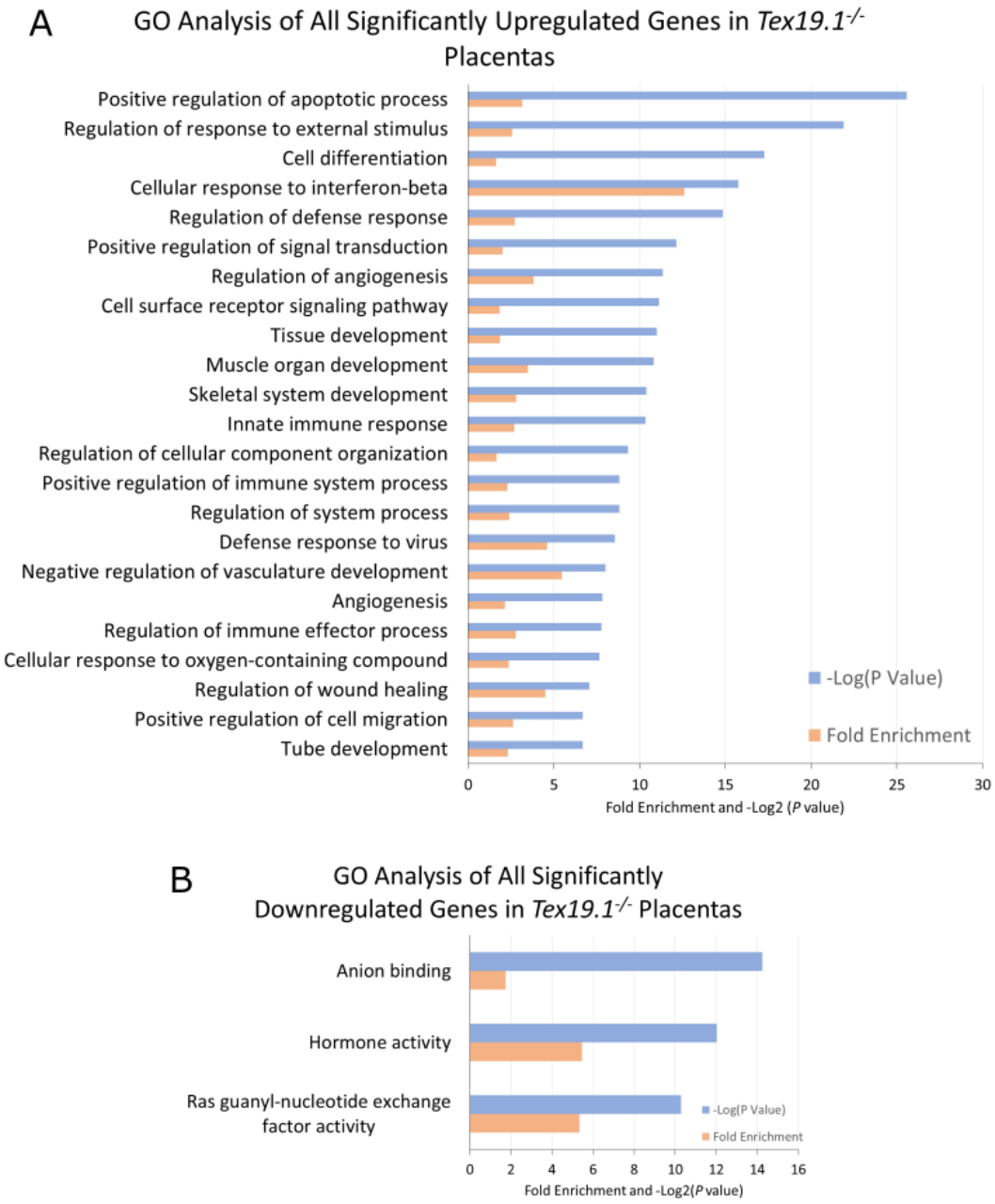


Figure 4.14: Gene ontology categories significantly enriched in *Tex19.1*$^{-/-}$ RNAseq. (A) GO analysis of all significantly upregulated genes with FDR<0.05. (B) GO analysis of all significantly downregulated genes with FDR <0.05. For both plots the enriched GO category is shown on the left of the barplot ordered by significance of enrichment (-Log *P*-value, blue bars). Fold enrichment is shown by the orange bars. Only GO categories with a significance of *P*<0.01 are shown.

4.2.18 The interferon pathway is activated in *Tex19.1*$^{-/-}$ placentas

As the GO category of 'Cellular response to interferon β' had the highest enrichment level for the upregulated genes I aimed to confirm the mRNA changes observed in the interferon and immune response categories from the RNAseq (figure 4.15A). I performed qRT-PCR in control and *Tex19.1*$^{-/-}$ placentas and confirmed the small but significant increase in expression of genes in the 'cellular response to interferon-beta' category (figure 4.15B). Importantly, the extent of upregulation of these genes is modest but in line with other mouse knockout models where an interferon β response is induced such as in *Samhd1*$^{-/-}$ mice (Maelfait et al., 2016). When significantly upregulated genes in the GO category 'Cellular response to interferon beta' were overlapped with genes in 'cellular response to interferon gamma' only, *Gbp2, Gbp3, ifitm2, ifitm3* were common to both. When intersected with 'cellular response to interferon alpha' only the upregulated genes *Ifit1, Tgtp1, Ifitm2* were common suggesting this interferon response is specific for interferon β. Interferon β is known to be involved in nucleic acid sensing and the immune response to retrotransposons in other mouse knockouts and experimental models of retrotransposon activation, further supporting activation of the interferon β response (Chiappinelli et al., 2015; Herquel et al., 2013; Maelfait et al., 2016; Roulois et al., 2015). Interestingly, recent evidence from *Mov10L* null testes which have LINE-1 de-repression, suggest that this activation of interferon β can occur not only in response to LTRs but also to LINE-1 elements (Yu et al., 2015). Together, these results indicate that an increase in the interferon β immune response may be present in the absence of *Tex19.1*.

Type I interferons function via signalling through activation of the Janus activated kinases (JAKs) which results in tyrosine phosphorylation of STAT1, leading to formation of interferon stimulated gene complexes (ISGs) which activate interferon stimulated response elements (ISREs)(Chapter 1 – figure 1.3)(Platanias, 2005). I therefore, aimed to determine if phosphorylation of STAT-1 is occurring in *Tex19.1*$^{-/-}$ placentas, as an indicator that the increased expression of the ISREs detected in the RNAseq is as a response to activation of the interferon β pathway.

Control and *Tex19.1*$^{-/-}$ placentas were dissected into junctional zone and labyrinth compartments at E12.5 to see if activation of interferon was specific to a particular layer of the placentas. Western blot for phosphorylated STAT1 indicated that STAT1 was phosphorylated in both layers of the control and *Tex19.1*$^{-/-}$ placentas. Interestingly, there was increased abundance of p-STAT1 specifically in the labyrinth of *Tex19.1*$^{-/-}$ placentas compared to control (n=2)(figure 4.16). No difference was observed in the junctional zone between control and knockout placentas. Although more *Tex19.1*$^{-/-}$ placentas need to be analysed to confirm this effect, this indicates that the upregulation of interferon β associated genes represents a response to activation of the interferon pathway in the labyrinth. Taken together, these data suggest that loss of *Tex19.1* induces aberrant innate immune responses in the developing placenta.

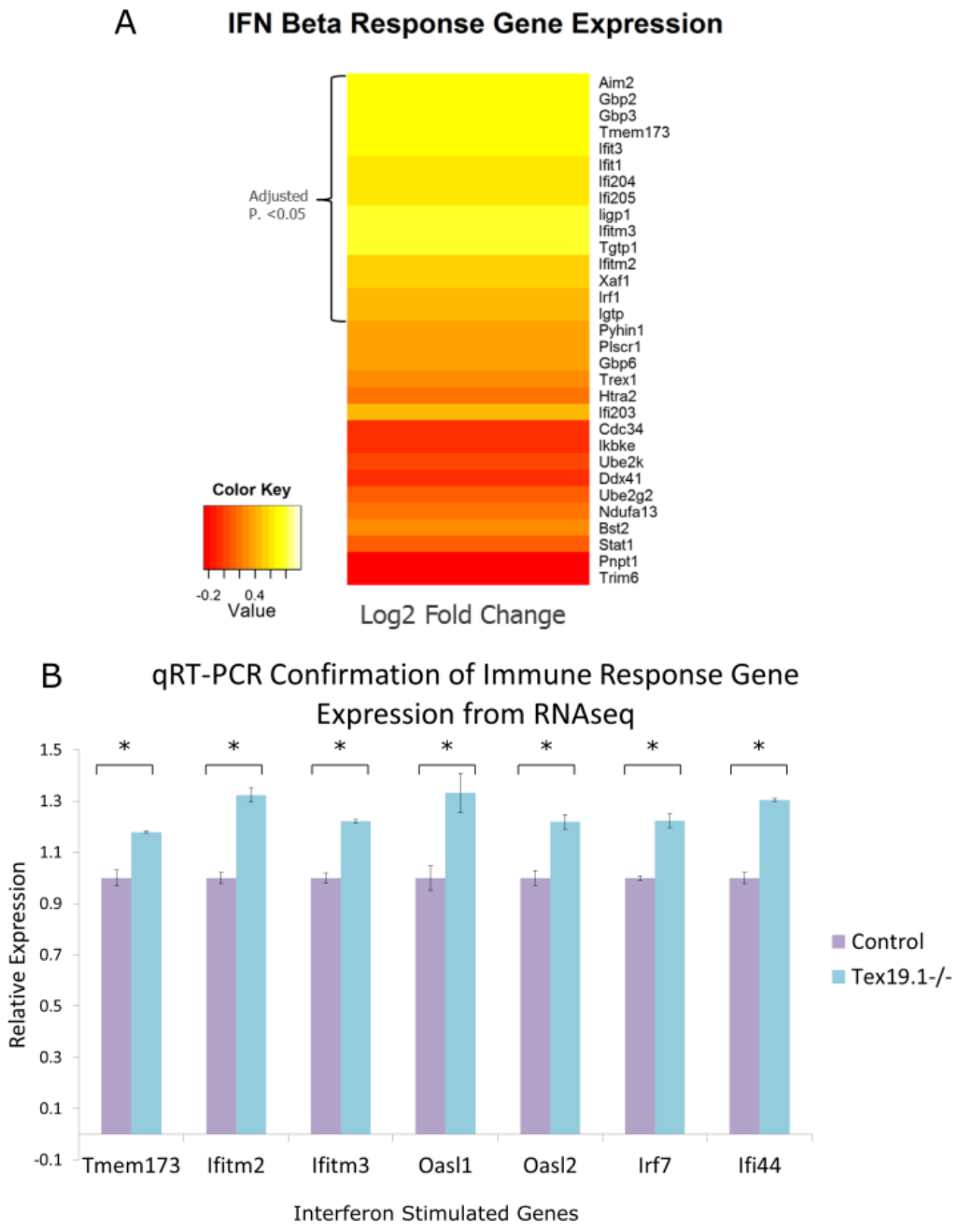


Figure 4.15: Immune genes are activated in *Tex19.1*$^{-/-}$ placentas (A) Heatplot of RNAseq genes corresponding to the 'Cellular response to interferon β' GO category ordered by significance of fold change. Yellow indicates upregulated genes whilst red indicates genes not changing expression. Those genes with an FDR of <0.05 are indicated by the bracket to the left of the plot. (B) qRT-PCR confirmation of classical interferon stimulated genes and immune response genes. Data is normalised to β-actin and expressed relative to gene expression in control placentas (n=3, **P*<0.05 Students T-test).

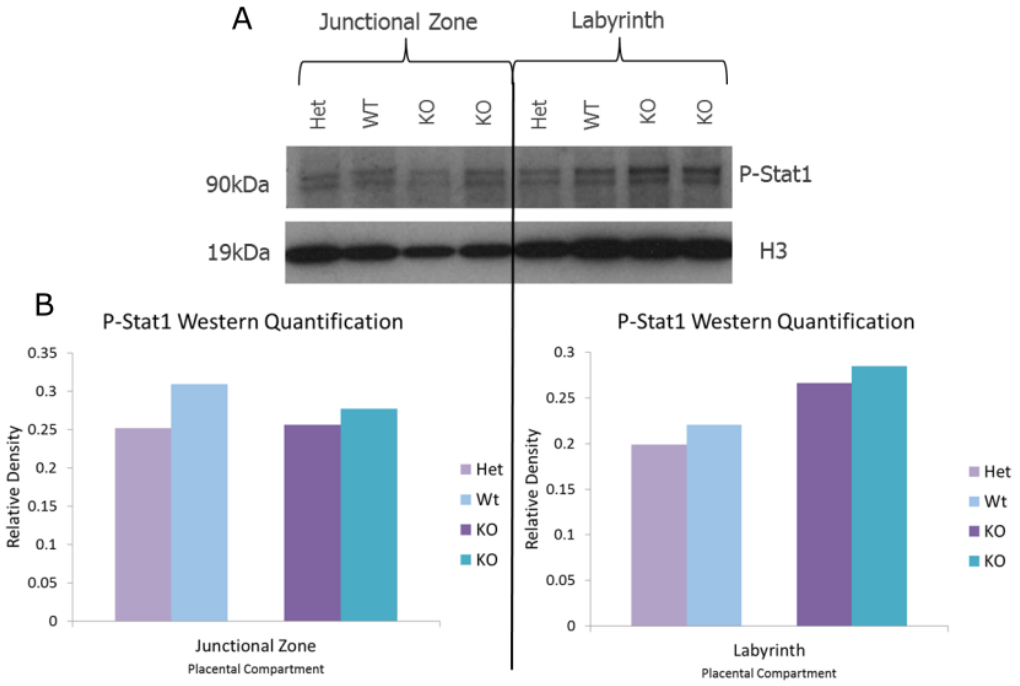


Figure 4.16: The interferon response is active predominantly in the labyrinth of *Tex19.1*$^{-/-}$ placentas. (A) Western blot for phosphorylated STAT1 as an indicator of activation of an interferon response. E12.5 control and *Tex19.1*$^{-/-}$ placentas were dissected into junctional zone and labyrinth and probed with antibody to phosphorylated STAT1 and histone H3 as a loading control. Protein size is indicated in kDa. (B) Densitometry quantification of western blot relative to histone H3.

4.3 Discussion

In this chapter I aimed to gain a better understanding of the etiology of the *Tex19.1*$^{-/-}$ placental phenotypes previously investigated at E18.5 by Reichmann et al. I identified that placental defects precede IUGR of the embryo in the absence of *Tex19.1* and that *Tex19.1* has a novel role in the regulation of placental specific genes (summarised in table 4.1). I next sought to further characterise the retrotransposon de-repression in the placentas of *Tex19.1*$^{-/-}$ animals and provided evidence for intriguing potential links between retrotransposon de-repression and alterations in gene expression in the placenta. I finished the chapter with an analysis of differentially expressed genes in *Tex19.1*$^{-/-}$ placentas and have discovered that an aberrant innate immune response is occurring (table 4.1). I propose that this may be a major contributor to defects in placental development, such as the increased cell death, which is then leading to the consequent failure of *Tex19.1*$^{-/-}$ embryos to reach their maximal growth potential during gestation.

Phenotype Assessed	*Tex19.1*$^{-/-}$ Phenotype	Figure
Placenta weight	Decreased	4.1
Embryo weight	Decreased	4.1
Cell death	Increased in junctional zone	4.3
Placental layer composition	Junctional zone reduced	4.4
Gene expression	Increases (immune response) and decreases (hormonal genes, enriched next to LINE-1 elements)	4.6, 4.10, 4.11, 4.15 & 4.16
Retrotransposon expression	Increases in LINE-1 and some LTR elements. Some decreases	4.8 & 4.9

Table 4.1: Summary of *Tex19.1*$^{-/-}$ placental phenotypes identified in this study

4.3.1 *Tex19.1*$^{-/-}$ placental defects precede IUGR

Embryos lacking *Tex19.1* are born significantly smaller than their littermate counterparts but it was not clear if placental defects were contributing to this reduction in embryo size (Reichmann et al., 2013). Going back in development enabled me to show that there is increased cell death at E12.5, which likely leads to the reduction in junctional zone size (chapter results summarised in table 4.1). These defects precede any reduction in embryo size. This is consistent with other mouse mutants with IUGR, whereby a reduction in the junctional zone results in IUGR (Hitz et al., 2005; Tunster et al., 2010), possibly due to reduced placental hormones influencing fetal maternal blood supply, which in turn reduces the exchange of nutrients to the developing embryo (John 2013; Simmons et al., 2008; Tunster et al., 2016). The fact that *Tex19.1*$^{-/-}$ embryos do not exhibit early embryonic lethality, such as *Mash2* null embryos with defective spongiotrophoblast development, indicates that *Tex19.1* is not absolutely

necessary for placental development (Guillemot et al., 1994; Reichmann et al., 2013; Rossant and Cross 2001).

Placental development occurs from the polar trophectoderm layer of the blastocyst at E4.5, with the development of the ectoplacental cone following implantation into the maternal uterine wall from E6.5 to E9.5 (Watson and Cross, 2005). The hypomethylated trophoblast derived cells of the ectoplacental cone gives rise to hypomethylated cells of the junctional zone and the different types of TGCs (Hu and Cross, 2010). Previous analysis indicates that *Tex19.1* is highly expressed in the ectoplacental cone in both post-implantation extra-embryonic and epiblast tissues at E6.5 and is maintained in the extra-embryonic tissue at E7.5, whereas expression is downregulated in the epiblast (Hackett et al., 2012). This raises the possibility that perturbing *Tex19.1* may have effects earlier in placental development than is investigated in this study. However, the preliminary data suggesting that there is no difference in weights of E10.5 *Tex19.1*$^{-/-}$ and control placentas possibly suggests that any defects in developing hypomethylated cell types are either more subtle at these earlier stages or are occurring after this time point, however more samples will be needed to investigate this further. Another possibility is that differentiation of some common precursors of spongiotrophoblast and glycogen trophoblast cells of the junctional zone may be being perturbed, leading to the reduced junctional zone size. Careful analysis of the ectoplacental cone earlier in development would need to be performed to assess this possibility.

4.3.2 Effects of retrotransposons and their transcriptional control of placental genes

The association of de-repressed L1MdF2 retrotransposons with downregulated genes in *Tex19.1*$^{-/-}$ placentas is the inverse of what has generally been reported previously (figure 4.11)(Goodier 2016). Largely, evidence exists for the LTR family of retrotransposons acting to enhance transcription via behaving as alternative promoters producing long non coding RNAs, or acting as enhancers upon their de-repression (Ecco et al., 2016; Herquel et al., 2013; Rowe et al., 2013b). The SINE class of

elements has also been shown to have a regulatory role in genome regulation with some SINE B2 elements functioning as insulators or enhancer blockers to facilitate gene expression in humans (Lunyak et al., 2007; Raab & Kamakaka 2010). This data suggested that transcription of repetitive elements may represent a strategy to establish functionally distinct domains within the mammalian genome to control gene activation (Lunyak et al., 2007). The importance of this association of retrotransposons with the regulation of these placental genes, along with potential mechanisms of genic regulation is unclear at present. Indeed, antisense transcripts from two full length L1MdF2 loci which are flanking significantly downregulated *Psg* genes was detected to the same extent in control and *Tex19.1*$^{-/-}$ animals. This likely points to a different mechanism of retrotransposon mediated repression distinct from antisense transcript interference, at least at the loci tested. Furthermore, the mechanism of retrotransposon repression could possibly be due to the perturbed function of KAP1 in the absence of TEX19.1 at some retrotransposon loci as indicated in chapter 3 of this thesis.

Also, the RNAseq analysis discussed in this chapter incorporated multiple mapping repeats. Repeating the analysis to include only uniquely mapping repeats may shed more light on the specific repeat loci which are being de-repressed and address if all elements of a retrotransposon subfamily are de-repressed or if it is a subset which are sensitive to repression by *Tex19.1*. The differential sensitivity of different elements of the same subfamily to histone based repressive mechanisms has been previously established (Brunmeir et al., 2010).

4.3.3 Gene expression changes in *Tex19.1*$^{-/-}$ placentas

When genic loci were analysed from the mapped RNAseq data, it is important to note that many of the strongly differentially downregulated genes in *Tex19.1*$^{-/-}$ placenta are normally expressed in the junctional zone, which I indicate is reduced in E12.5 *Tex19.1*$^{-/-}$ placentas. Therefore, mRNAs expressed in the junctional zone may be under represented in *Tex19.1*$^{-/-}$ placentas in the RNAseq data, due to this compositional difference. Despite this, transcriptional regulation by *Tex19.1* is likely occurring for at least some genes as indicated by the significant reduction in the expression of the S-

TGC marker gene, *Ctsq* despite similar amounts of S-TGCs. The caveat of the possible contribution of compositional differences was further addressed for highly differentially expressed transcripts, when GO terms enriched in *Tex19.1*$^{-/-}$ placentas were compared alongside published GO analysis from the trophectoderm derived placenta compartments and the decidua (Bode et al., 2005; Knox and Baker, 2008; Simmons and Cross, 2005; Simmons et al., 2007, 2008). This allowed identification of enriched pathways unlikely to be a consequence of compositional differences.

The fact that many genes encoding placenta derived hormones were reduced in expression, such as the *Psg* genes and placental lactogens, likely indicates reduced hormone secretion into maternal blood. This may therefore be negatively influencing maternal-fetal processes and reducing communication across the fetal-maternal interface (John, 2013, Simmons et al., 2008). PSG proteins have roles in protecting the fetus from the maternal immune system, along with roles in remodelling both placental and maternal vasculature (Wu et al., 2008). Interestingly, reduced levels of these hormones in maternal blood in humans is associated with placental dysfunction and pregnancy complications such as pre-eclampsia and IUGR (Bersinger and Ødegård, 2004).

In addition to placental specific gene clusters, some interesting genes were detected as highly down regulated out-with these genes families. For example, Keratin intermediate filament 5 (*Krt5*) was downregulated 16 fold (P=2.14x10^{-33}). This gene is hypermethylated in low birth weight human females compared to normal birthweight (Rumbajan et al., 2016). The expression of this gene as a consequence of this increased DNA methylation is unclear but indicates an association of this gene with low birth weight humans and mice (Rumbajan et al., 2016). Interestingly, mice null for *Krt5* die shortly after birth due to defective development of the basal epidermis, indicating an essential role for this gene in development, however the role of this gene in the mouse placenta remains unclear (Peters et al., 2001b).

Some individual highly expressed genes provided interesting links to placental dysfunction. The most highly upregulated gene *Adamdec1* (4-fold upregulation, P=3.80x10^{-8}) is localised to the spongiotrophoblast layer at E12.5 and is negatively regulated by prolactin when expressed in the uterus (Baran, 2002). This suggests that

the reduction in the expression of prolactin genes from the junctional zone leads to upregulation of *Adamdec1* which is thought to have functional consequences on the immune environment and angiogenesis of placentas (Baran, 2002). The upregulation of this gene in response to the reduction in prolactins indicates that perturbations in the hormonal environment of the placenta can have dramatic regulatory effects on placental genes.

4.3.4 Response to Interferon β

To investigate further mechanism which may be contributing to impaired placental development GO analysis was performed. Interestingly, the most highly enriched and significantly changing GO category was 'cellular response to interferon β'. This is likely not explained by the overrepresentation of the maternal decidua as this GO category has not been reported to be enriched in genes expressed in the decidua (Knox and Baker 2008). To further address this I confirmed that the GO category represented a *bona fide* activation of the interferon β pathway by demonstrating an increase in the phosphorylation of STAT1, a major effector protein responsible for activation of interferon stimulated genes (Kassiotis and Stoye 2016). This was specifically detected in the labyrinth compartment and suggests that the increased cell death that is occurring in the junctional zone may be likely due to paracrine effects of secreted interferon β, if this is indeed causative of the cell death. More animals will be needed to confirm this finding and additional analysis of proteins such as ISG15, an interferon inducible, ubiquitin-like protein modifier with antiviral functions, will provide further evidence for the activation of a *bona fide* type I interferon response in the labyrinth of *Tex19.1*$^{-/-}$ placentas (Maelfait et al., 2016; Morales and Lenschow, 2013). The response to interferon β in *Tex19.1*$^{-/-}$ placentas and its potential contribution to IUGR of the embryo is also intriguing due to evidence from humans. Interferon β is used as a therapy for multiple sclerosis and in multiple studies has been associated with IUGR, however the placental pathology in these patients has never been assessed (Amato et al. 2010; Boskovic 2015).

At present the cause of the increased interferon β response in *Tex19.1*$^{-/-}$is unclear, however a few intriguing possibilities have arisen from this study and other recent research. The first thing to hint at a possible cause of the interferon response was the enrichment in GO categories 'regulation of defence response' and 'defence response to virus' (P<0.01, 2.73 fold and 4.6 fold respectively). The de-repression of retrotransposons in *Tex19.1*$^{-/-}$ placentas raises the possibility that the interferon response is activated to sense these retrotransposons. Indeed, mice null for the methylation sensitive genome defence gene *Mov10L* have de-repression of LINE-1 in spermatocytes accompanied activation of interferon β (Yu et al., 2015; Zheng et al., 2010). The activation of interferon β was shown to be due to LINE-1 increases by *in vitro* overexpression experiments (Yu et al., 2015). Numerous studies in humans and mice indicate activation of interferon β responses are induced via the sensing of de-repressed retrotransposon derived nucleic acids in the cytosol of affected cells (Chiappinelli et al., 2015; Herquel et al., 2013; Roulois et al., 2015). This response can be mediated by the cytosolic DNA sensor *Tmem173 (Sting)* or the cytosolic dsRNA sensors *Rig1* or *Mda5* (Paludan and Bowie, 2013). Interestingly, *Tmem173* was significantly upregulated in *Tex19.1*$^{-/-}$ placentas. To determine if the interferon β response and placental developmental defects are dependent on the sensing of increased cytosolic retrotransposon DNA, inter-crosses of *Tex19.1*$^{+/-}$ and *Tmem173*$^{-/-}$ mice could be performed, to generate double *Tex19.1*$^{-/-}$*Tmem173*$^{-/-}$ mice. If mutating *Tmem173* abrogates the IUGR phenotype of *Tex19.1*$^{-/-}$ pups, this would provide evidence that inappropriate activation of an innate immune response is contributing to the placental defects and IUGR seen in *Tex19.1*$^{-/-}$ mice.

From my findings in this chapter, I propose that the response to interferon β, along with reduced hormonal gene expression, are likely exerting combined effects on development of the placenta, which is exacerbated throughout gestation and contributes to IUGR of *Tex19.1*$^{-/-}$ embryos.

Chapter 5: Investigation of the contribution of the TEX19.1 interacting proteins, the E3 ubiquitin ligases, to the *Tex19.1*$^{-/-}$ testicular phenotype

5.1 Introduction

TEX19.1 is a protein comprised of two unique domains, the sequence of which gives no indication of its biochemical function. Therefore, to impart potential mechanistic functions of TEX19.1 it is useful to consider its interacting partners. The phenotype of mice null for an interacting protein can then give an indication of the possible functional contribution of that protein interaction to the *Tex19.1*$^{-/-}$ phenotype. TEX19.1 interacts with the E3 ubiquitin ligase UBR2 (Yang et al., 2010). *Ubr2*$^{-/-}$ testis have severe atrophy owing to the arrest of meiotic spermatocytes early in prophase I, similar to *Tex19.1*$^{-/-}$ animals (Kwon et al., 2003). The de-repression of MMERVK10C retrotransposons that are observed in *Tex19.1*$^{-/-}$ animals (Ollinger et al., 2008) has never been investigated in *Ubr2*$^{-/-}$ mice and it is unclear if other interacting partners may also contribute to the phenotypes observed in *Tex19.1*$^{-/-}$ testes. Therefore, the final aim of my PhD is to investigate the hypothesis that the defective spermatogenesis and transcriptional retrotransposon de-repression phenotypes observed in *Tex19.1*$^{-/-}$ testes may be, at least in part, due to loss of TEX19.1 interacting partners.

To determine other interacting partners of TEX19.1 in addition to UBR2, Marie MacLennan performed co-immunoprecipitation (Co-IP) mass spectrometry on the cytoplasmic fraction of TEX19.1-YFP ES cells. This indicated that TEX19.1 can also be co-immunoprecipitated with the E3 ubiquitin ligases UBR5 and KAP1 (figure 5.1 & figure 3.3). Each interaction determined by mass spectrometry was confirmed by Co-IP in TEX19.1-YFP ES cells and western blot for GFP confirmed successful immunoprecipitation of YFP or TEX19.1-YFP (figure 5.1 & figure 3.3 Marie MacLennan). Conditional germ line deletion of *Kap1* has no gross effect on meiosis (Weber et al., 2002), has well characterised roles in retrotransposon repression (Rowe et al., 2010) and has been previously investigated, however potential roles for *Ubr5* in spermatogenesis and retrotransposon regulation have never been studied.

UBR5 is a HECT domain E3 ligase implicated in a broad array of processes such as N-end rule proteolysis and ubiquitin mediated degradation (Varshavsky, 2011), cell cycle checkpoint control (Munoz et al., 2007) and regulation of histone ubiquitylation at sites of DNA damage (Gudjonsson et al., 2012). It is unknown if *Ubr5* is necessary

for spermatogenesis or retrotransposon repression, as embryonic lethality upon loss of *Ubr5* has precluded previous analysis (Saunders et al., 2004). To address this, I generated *Stra8-Cre*^{tg} *Ubr5*$^{flox/\Delta}$ (*Ubr5*^{CKO}) mice where *Ubr5* expression is conditionally ablated in undifferentiated spermatogonia due to the expression of the *Stra8-Cre* transgene from 3 days post-partum onwards (Sadate-Ngatchou et al., 2008).The main objective of this chapter is to determine the extent to which the TEX19.1 interacting partners UBR5 and UBR2 may be contributing to the *Tex19.1*$^{-/-}$ phenotype.

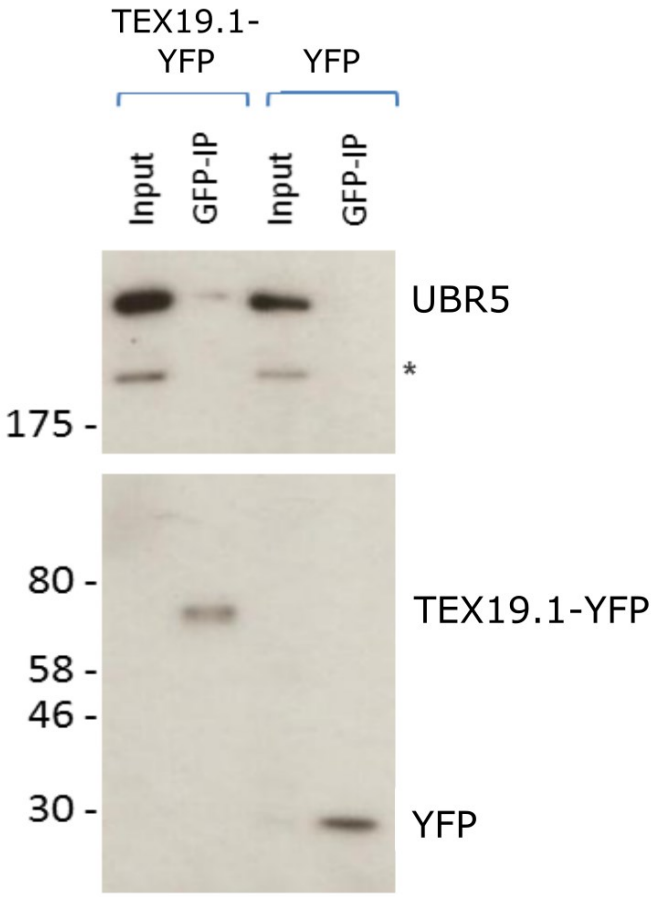


Figure 5.1: TEX19.1 interacts with UBR5 (experiment performed by Marie Maclennan). Western blot of co-immunoprecipitation experiment confirming the interaction of TEX19.1 and UBR5 identified by IP-mass spectrometry. Bands corresponding to UBR5 are identified in the input lanes of both Tex19.1-YFP and YFP ES cells. Immunoprecipitation was performed using GFP beads and a band corresponding to UBR5 is detected in Tex19.1-YFP ES IP but is absent from YFP ES IP samples (Marie Maclennan). (* denotes an unidentified non-specific band present in input samples)

5.2 Results

5.2.1 *Ubr5* is expressed throughout spermatogenesis

*Ubr*5 mRNA and protein is present in rat gonocytes and spermatogonia and in humans *Ubr5* has the highest levels of expression in testis compared to other tissues (Callaghan et al., 1998; Manku et al., 2012). The extent of *Ubr5* expression in mouse testis has never been investigated. Therefore, to address the expression levels of the TEX19.1 interacting partners *Ubr5* and *Ubr2* in mouse testis in relation to a range of different tissues, I analysed publicly available mouse gene expression microarray datasets (Thorrez et al., 2008; GSE9954). *Ubr5* expression is highest in ovaries, placenta and testis, with slightly lower expression in muscle and heart (figure 5.2A). The lowest levels of *Ubr5* expression are present in liver, lung, kidney, spleen and brain. *Ubr2,* has highest expression in the ovaries, placenta and muscle, however lowest expression in the testis (figure 5.2A). The germline specific gene *Dazl* is shown as a positive control for testis expression and the trophoblast specific gene, *Cdx2* is shown as a positive control for placental expression (figure 5.2A).

The high levels of *Ubr5* expression in the testis prompted me to more specifically determine at what stage of spermatogenic development *Ubr5* is expressed. Spermatogenic development in seminiferous tubules can be divided into twelve stages in the mouse. Each stage has characteristic combinations of spermatogonia, spermatocytes and spermatozoa synchronously proceeding through the spermatogenic cycle (figure 1.11)(Russell et al., 1990**)**. I therefore analysed gene expression microarray data from different microdissected seminiferous tubule stages (Hasegawa and Saga, 2012; GSE34758). This indicated that *Ubr5* has consistent expression in each stage of tubule, similar to *Ubr2* (figure 5.2B). It is not clear why there is disparity between *Ubr2* and *Ubr5* expression data in figure 5.2A compared to figure 5.2B. Expression of the germline specific gene *Dazl* is indicated as a marker of germline expression and *Cdx2,* a placental specific gene which exhibits no germline expression, acted as a negative control (figure 5.2B). Interestingly, *Ubr5* and *Ubr2* are significantly more highly expressed in seminiferous tubules than FACs sorted sertoli cells, indicating that the majority of expression is likely due to developing germ cells.

In order to determine if this was the case, I investigated the expression dynamics of *Ubr5* and *Ubr2* throughout spermatogenesis using publicly available RNAseq datasets of purified germ cells (Gan et al., 2013; GSE35005 Supplementary data 2). *Ubr5* is expressed throughout spermatogenesis from primary spermatogonia-A to elongating spermatids and has a 1.8 fold increase in expression from pre-leptotene spermatocytes to pachytene spermatocytes (figure 5.2C). *Ubr2* is also expressed at all stages of spermatogenesis, however undergoes a 2.1 fold reduction in expression from preleptotene spermatocytes to pachytene spermatocytes (figure 5.2C). *Ubr2* null mice are infertile and exhibit numerous defects in progression through spermatogenesis as previously discussed (An et al., 2012; Kwon et al., 2003). These data indicate that *Ubr5* may be functioning at all stages of spermatogenesis, with the increase in expression during meiosis possibly indicative of a meiosis specific function.

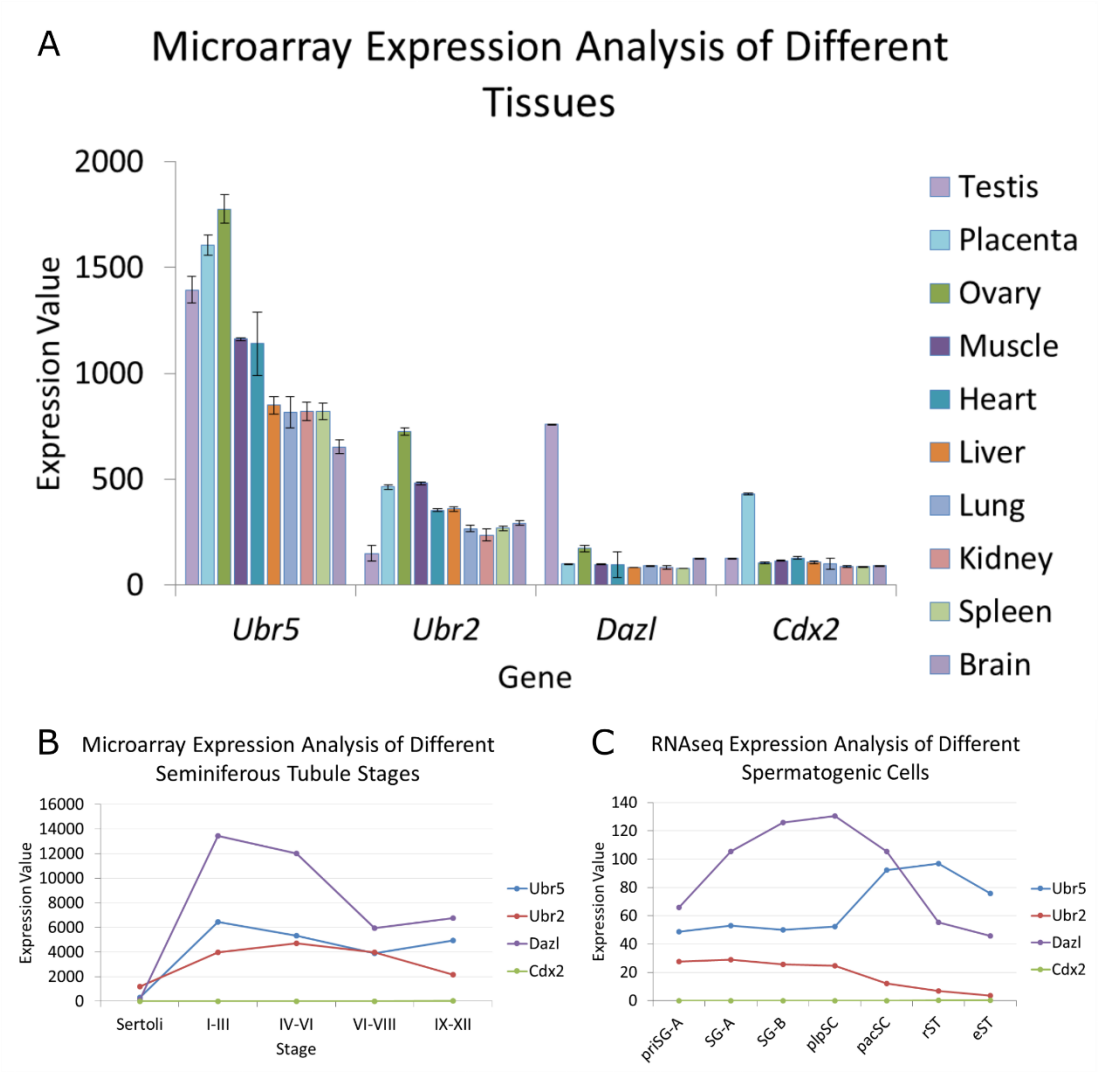


Figure 5.2: *Ubr*5 is expressed throughout spermatogenesis with peak expression in pachytene spermatocytes and round spermatids. (A) Microarray expression levels (mean ± standard error for three replicates) for the TEX19.1 interacting partners UBR5 and UBR2 in a range of different tissues. *Dazl* and *Cdx2* are included as germline and placenta specific genes respectively (Data from Thorrez et al., 2008, GSE9954). (B) Microarray expression levels from microdissected seminiferous tubule stages (Mean shown. No standard error bars due to only two replicates; data from Hasegawa & Saga 2012, GSE34758). (C) RNAseq expression levels from sorted spermatogenic cell types. Primary spermatogonia-A (priSG-A), spermatogonia-A (SG-A), spermatogonia-B (SG-B), pre-leptotene spermatocyte (plp-SC) and pachytene spermatocyte (pacSC) are from prepubertal mice undergoing the first wave of spermatogenesis and round spermatids (rST) and elongating spermatids (eST) are from 60 day postpartum adult mice (Gan et al., 2013, GSE35005, Supplementary data 2). (Mean shown. No standard error bars due to only two replicates; data from Gan et al., 2013, GSE35005, Supplementary data 2).

5.2.2 Generation of conditional *Ubr5^{CKO}* animals

Loss of UBR5 is embryonic lethal due to failure of yolk sac vascularisation and chorioallantoic fusion (Saunders et al., 2004), therefore to assess the potential role of UBR5 during spermatogenesis I generated conditional *Stra8-Cre^{tg} Ubr5$^{flox/\Delta}$* (*Ubr5^{CKO}*) mice. To interfere with mRNA expression and UBR5 protein function C57BL/6J mice carrying the *EUCOMM Ubr5* conditional gene trap (EUCE0171f01) in the antisense orientation between exons 20-21 were obtained from Mark Ditzel (figure 5.3A)(Kinsella et al., 2016; Schnütgen et al., 2005).

This conditional *Ubr5* gene trap allele contains a splice acceptor site, a β-galactasidase:aminoglycoside 3' phosphotransferase (βGEO) coding sequence conferring X-gal staining and Neomycin resistance followed by a poly-adenylation signal (figure 5.3A). The cassette is flanked by two site-specific recombination loci (FLPe/frt and Cre/loxP) which allow for gene trap cassette inversions from the sense, coding strand of *Ubr5* to the antisense, noncoding strand and back (figure 5.3A). In the antisense orientation the cassette does not disrupt *Ubr5* gene expression and the UBR5 protein retains all domains (*Ubr5^{flox}*)(figure 5.3A). After CRE-mediated recombination, the cassette recombines in the sense, mutagenic orientation where the predicted fusion protein consists of βGEO fused to UBR5's UBA domain (*Ubr5$^{\Delta}$*) (figure 5.3B). This UBR5 fusion protein is predicted to be functionally impaired due to lack the UBR domain responsible for N-end rule function (Kozlov et al., 2007), the HECT domain E3 ubiquitin ligase activity (Matta-Camacho et al., 2012) and miRNA regulatory function via the Poly(A)-binding protein C-terminal (PABC) domain (figure 5.3B) (Lim et al., 2006).

Ubr5 expression was conditionally ablated in undifferentiated spermatogonia by crossing with mice transgenic for *Stra8-Cre* (Sadate-Ngatchou et al., 2008). This transgene drives expression of Cre recombinase, which recombines the gene trap cassette in the mutagenic, sense orientation specifically in male germ cells from 3 days post-partum onwards generating the *Ubr5$^{\Delta}$* allele (figure 5.3B).

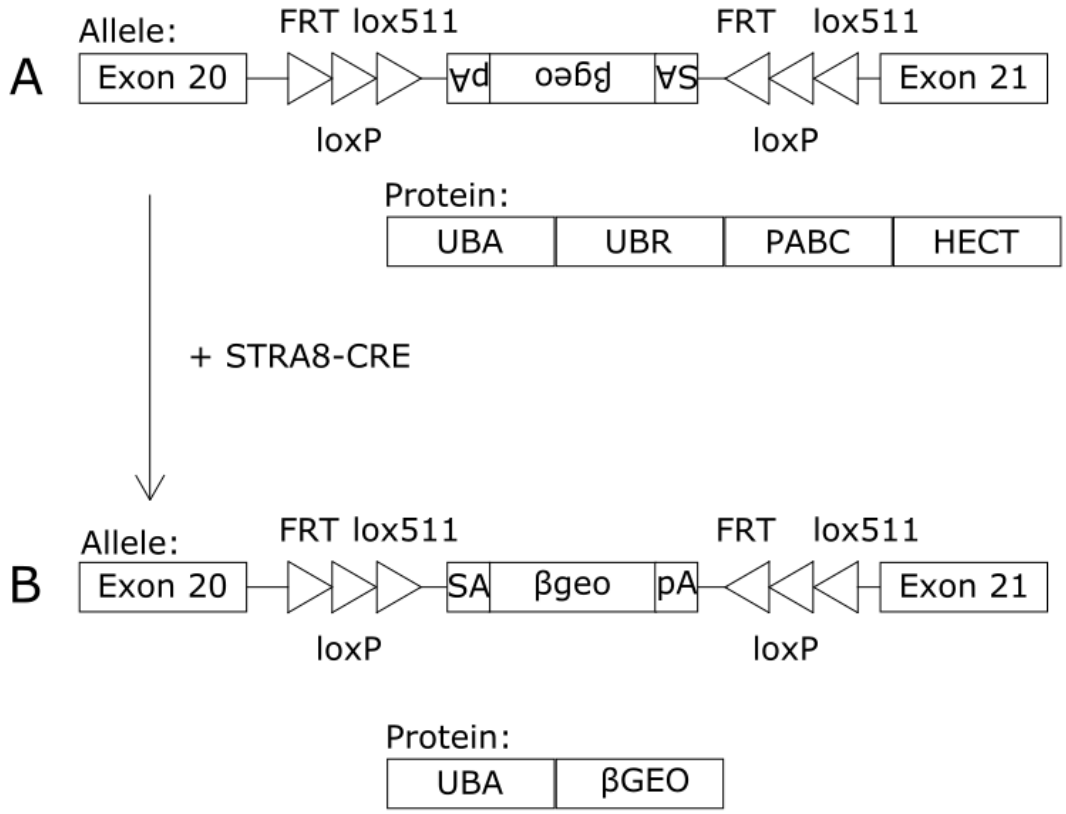


Figure 5.3: Schematic of the *Ubr5* gene trap allele and resultant protein domains before and after STRA8-CRE mediated recombination (Modified from Kinsella et al., 2016). (A) Schematic of the EUCOMM EUCE0171f01 gene trap introduced between exon 20 and 21 of *Ubr5.* The orientation of FRT, loxP and lox511 recombination sites are shown as triangles flanking the gene trap which contains a splice acceptor site (SA), a LacZ and neomycin coding sequence (βgeo) and a poly-adenylation signal (pA). This allele codes for whole UBR5 protein as indicated below the schematic. (B) Representation of the effect of Cre mediated recombination on the gene trap allele and the resultant non-functional fusion protein with UBR, PABC and HECT domains replaced by the functional βgeo cassette.

5.2.3 *Ubr5^{CKO}* mice are born at normal Mendelian ratios

Generally, heterozygous male *Stra8-Cre^{tg} Ubr5$^{flox/wt}$* mice were crossed with heterozygous female *Stra8-cre^{-} Ubr5$^{flox/wt}$* mice to generate the male specific conditional *Ubr5* knockout genotype *Stra8-Cre^{tg} Ubr5$^{flox/\Delta}$ (Ubr5^{CKO})* at an expected frequency of 1:16 (see section 2.7.1 for further information about breeding strategy). To ensure the *Stra8-cre* transgene was present and the animals were heterozygous for the *Ubr5$^{flox/del}$* allele genotyping PCR was performed on ear clips from adult animals (figure 5.4).

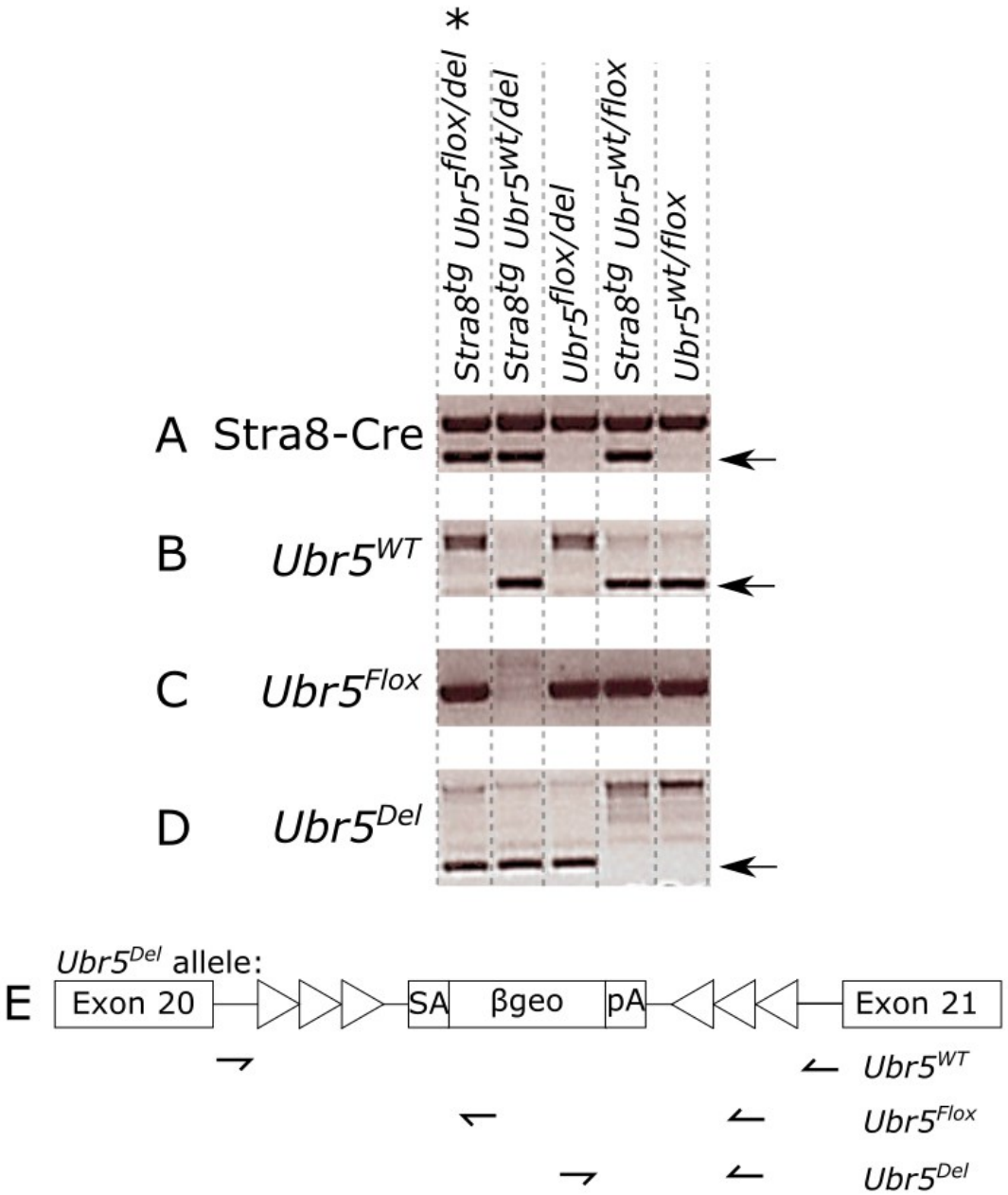


Figure 5.4: Genotyping of *Ubr5* transgenic animals. Genotyping PCR results from tail tips of *Ubr5* transgenic mice. (A) The presence of the Stra8-Cre transgene is determined by the detection of a lower band indicated by the arrow. The upper band is a control PCR in a non-transgenic region of the genome. (B) The wild type allele is detected by primers flanking the βGEO gene trap cassette determined by the lower band indicated by the arrow. The upper band is detected due to the presence of the βGEO cassette. (C) Presence of a band is used to detect the βGEO cassette in the antisense orientation (flox). (D) The sense orientation βGEO cassette, resulting in non-functional UBR5 fusion protein, is determined by the presence of a lower band indicated by the arrow. * indicates the $Ubr5^{CKO}$ genotype. (E) Schematic indicating the locations of primers to detect each *Ubr5* allele.

To ensure there was no embryonic lethality associated with any particular genotype, Chi-squared tests were performed. *Ubr5*^{wt} homozygous, heterozygous and *Ubr5*^{flox} homozygous offspring were obtained at the expected Mendelian ratio of 1:2:1 respectively (χ^2-test p=0.3, n=142)(table 5.1). Males and females were born at the expected 1:1 ratio (χ^2-test p=0.4, n=142) and *Stra8-Cre*^{tg} and *Stra8-Cre*$^{-}$ alleles were transmitted at the expected ratio of 1:1 (χ^2-test p=0.09, n=142) with no sex specific preference.

All Pups	Number of Pups Observed	Number of Pups Expected
Stra8- wt/wt	20	17.75
Stra8- wt/flox(del)	29	35.5
Stra8- flox/flox(del)	12	17.75
Stra8+ wt/wt	22	17.75
Stra8+ wt/flox(del)	43	35.5
Stra8+ fox/flox(del)	16	17.75
Total	142	142
Chi test P value	0.295408967	

Table 5.1: Observed and expected distribution of all *Ubr5*^{CKO} genotypes. Table indicating the observed and expected ratios of pup genotypes from crosses of male *Stra8-Cre*^{tg} *Ubr5*$^{wt/flox}$ with female *Stra8-Cre*$^{-}$ *Ubr5*$^{wt/flox}$.

Recombination by Cre transgenes can have varying efficiencies (Sadate-Ngatchou et al., 2008). Therefore, to determine the excision efficiency of *Stra8-Cre* for the $Ubr5^{flox}$ allele, male $Stra8\text{-}cre^{tg}$ $Ubr5^{flox/wt}$ and female $Stra8\text{-}cre^{-}$ $Ubr5^{flox/wt}$ animals were mated and the frequency of the transmission of the $Ubr5^{del}$ allele was determined in the offspring. *Stra8-cre* is male specific and has no activity in the female germline, therefore the $Ubr5^{del}$ allele must be generated in the paternal germline (Sadate-Ngatchou et al., 2008). If *Stra8-cre* was 100% efficient, no $Ubr5^{flox/flox}$ pups would be expected. The ratio of $Ubr5^{flox/flox}$ and $Ubr5^{flox/del}$ pups therefore indicates the excision efficiency of *Stra8-Cre.* A total of 3 pups were homozygous $Ubr5^{flox/flox}$, whereas 25 pups were heterozygous $Ubr5^{flox/del}$. This indicates an 89% $Ubr5^{flox}$ excision efficiency, suggesting incomplete excision of the$Ubr5^{flox}$ allele in 11% of spermatocytes.

5.2.4 *Ubr5* protein levels are reduced in $Ubr5^{CKO}$ testes

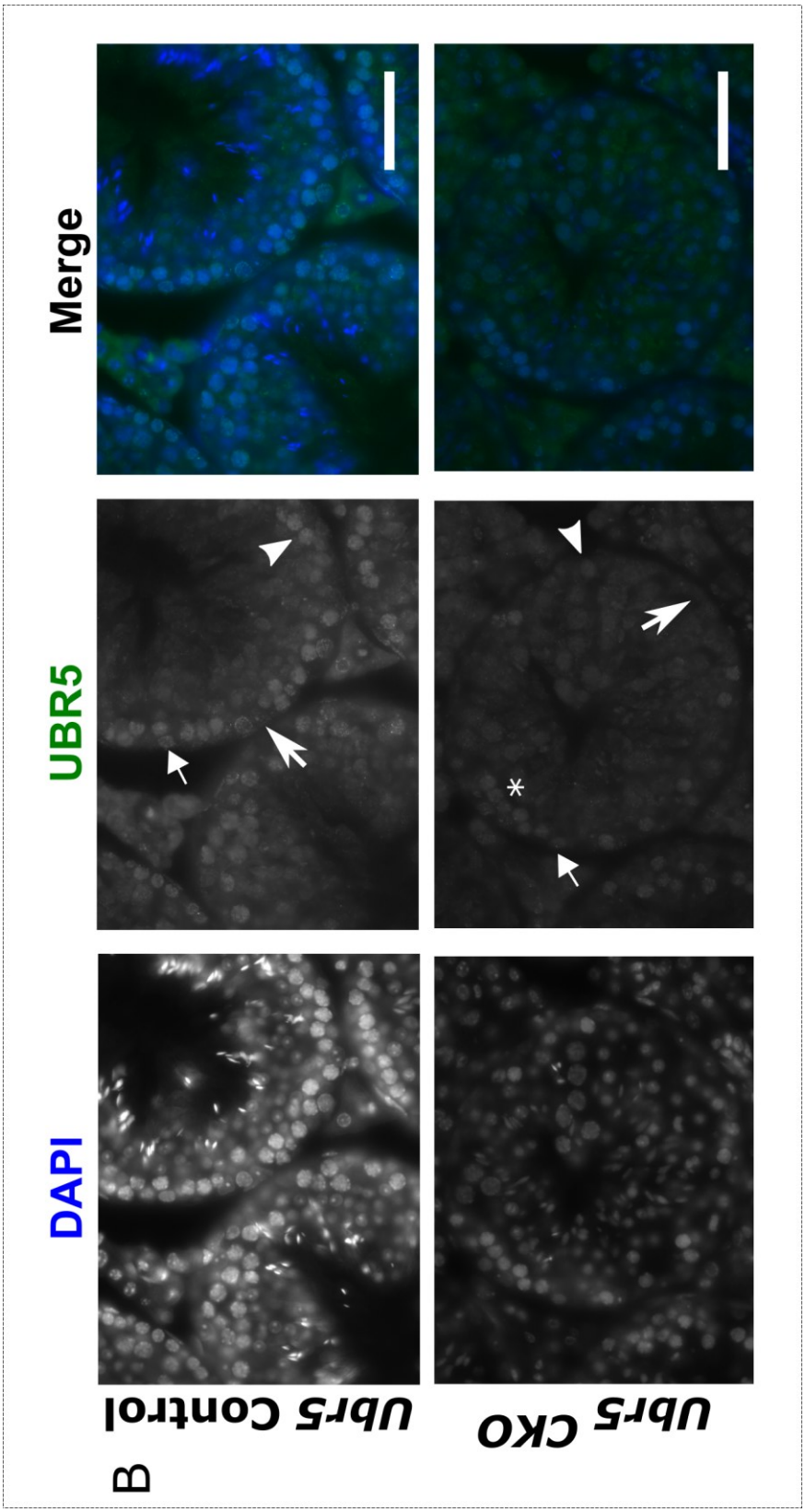
I next aimed to determine if there were lower levels of UBR5 protein in seminiferous tubules of $Stra8\text{-}Cre^{tg}$ $Ubr5^{flox/\Delta}$ ($Ubr5^{CKO}$) animals, indicative of successful knockout of *Ubr5*. Immunostaining with an anti-UBR5 antibody was performed on cross sections of seminiferous tubules from control and $Ubr5^{CKO}$ animals (figure 5.5).

Previous studies have identified UBR5 protein in rat spermatogonia and my microarray and RNAseq data analysis identified *Ubr5* expression in cells in all stages of spermatogenesis from spermatogonia to elongating spermatids (Manku et al., 2012)(figure 5.2B&C). In line with this expression analysis, UBR5 protein was detected in adult control testis, in the nuclei of spermatogonia (triangle arrow), with strongest staining detected in meiotic spermatocyte nuclei (arrowhead)(figure 5.5B). Sertoli cells were not stained for UBR5, in line with expression analysis (pointed arrow (figure 5.5B & 5.2B). Moderately weaker nuclear staining was detected in round spermatids and UBR5 staining was largely absent from elongating spermatids (figure 5.5). Modest staining in interstitial cells possibly reflects background staining as was also detected in sections assessed with a negative isotype control (figure5.5A

Secondary only). In contrast, reduced amounts of staining was observed in *Ubr5*^{CKO} testes in all spermatogenic cell types, indicating that conditional *Ubr5* deletion results in reduced protein levels as expected (figure 5.5). Interestingly, some spermatocytes still stain for UBR5 despite neighbouring cells being largely devoid of staining (* in figure 5.5B) suggesting that these are retaining *Ubr5* expression and have escaped excision by Stra8-Cre.

Figure 5.5: UBR5 staining is present in all cell types through spermatogenesis and is reduced in *Ubr5*^{CKO} testes (figures 5.5A and 5.5B on subsequent full pages) (A) Immunofluorescent staining of cross sections of adult control and *Ubr5*^{CKO} testis with antibody to UBR5 protein. Scale bar represents 100µm. (B) Higher magnification images of control and *Ubr5*^{CKO} testis. Sertoli cells negative for UBR5 are indicated by the pointed arrow, UBR5 positive spermatogonia by the triangle arrow and meiotic spermatocytes by the arrowhead. Spermatocytes retaining UBR5 expression in *Ubr5*^{CKO} testis are indicated by *. A secondary antibody only, negative isotype control is also shown to highlight background non-specific staining. Scale bar=50µm

A
Ubr5 Control
Ubr5 CKO
Secondary only
DAPI
UBR5
Merge


B
Ubr5 Control
Ubr5 CKO
DAPI
UBR5
Merge

5.2.5 *Ubr5*^{CKO} animals have significantly reduced testes weight and sperm count

To determine if there were any overt defects in testis development after *Ubr5* deletion, I weighed the testes and performed epididymal sperm counts. Adult control mice had an average testis weight of 90mg which was reduced 1.7 fold to 52mg in *Ubr5*^{CKO} mice (figure 5.6A)(n=8, P<0.01, Mann Whitney U test). This is a less severe phenotype than the 2.6 fold reduction in testis weights from adult control (111mg) to *Tex19.1*$^{-/-}$ animals (42.5mg) (Ollinger et al., 2008). Testis were visibly reduced in size, however there were no apparent perturbations in general shape compared to control animals (figure 5.6B). To determine if this reduction in weight was in part due to a reduction in mature spermatozoa I counted sperm from the epididymis. There was a 7 fold reduction in sperm count from $1.9x10^7$ cells/epididymis in control animals to $2.9x10^6$ cells/epididymis in *Ubr5*^{CKO} animals, indicating that *Ubr5* is contributing to germ cell development (figure 5.6C&D)(n=6, P<0.01 Mann Whitney U test). This reduction in sperm is less severe than the 100 fold reduction in *Tex19.1*$^{-/-}$ animals where control testes have $1.3x10^7$ sperm/ml versus $1.3x10^5$ sperm/ml in *Tex19.1*$^{-/-}$ mice (Ollinger et al, 2008). This is likely partially due to the *Stra8-Cre* being 86% efficient and leaving around 11% of spermatocytes with *Ubr5*^{flox} alleles, but also indicates that the *Ubr5*^{CKO} spermatogenic defect may be distinct from the meiotic defects in *Tex19.1*$^{-/-}$ testes.

Different previously studied mutant mice with defects in spermatogenesis exhibit different dynamics of testicular degeneration over time (Crichton et al., 2013). For example, *Tex19.1*$^{-/-}$ mice showed no correlation between the severity of testis degeneration and the age at which the mice were examined from 6 weeks to 9 months, whereas *Ubr2*$^{-/-}$ mice have severe testicular defects at 2 weeks post-partum with progressive decline in testis weight to 9 (Kwon et al., 2003; Ollinger et al., 2008). Mice lacking the *Zfp145* gene which encodes the undifferentiated spermatogonia restricted protein, PLZF have a progressive, age dependent decline (Costoya et al., 2004). To determine if there was a progressive decline in *Ubr5*^{CKO} testis weight upon aging I analysed 6-7 week and 5 month testis weights. 5 month old control mice had a 1.1 fold increase in testis weight compared to 6 to 7 week old control mice. In contrast,

Ubr5^{CKO} mice has a 1.3 fold decrease in testes weights from 6-7 week old to 5 month *Ubr5*^{CKO} animals, suggesting that there is only moderate progressive testes degeneration with age (See appendix figure 7.2 for comparison of 6 week and adult testes weights and sperm counts).

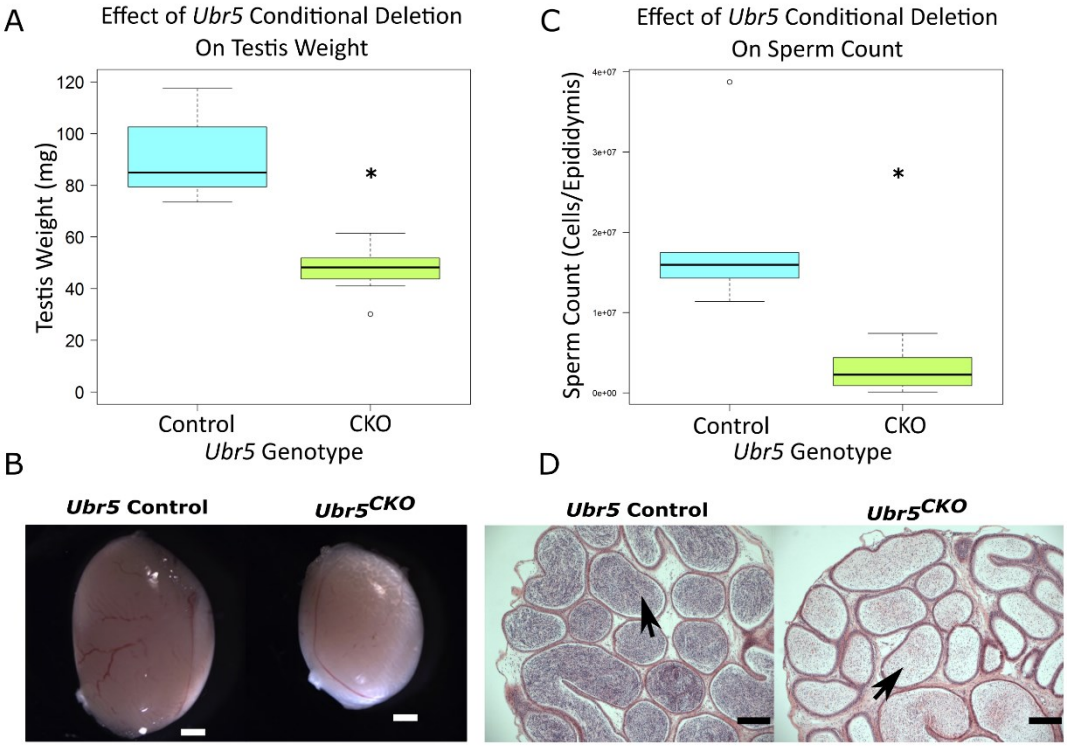


Figure 5.6: Testes weight and sperm count of adult *Ubr5*^{CKO} animals is significantly reduced. (A) Boxplot showing the median testis weights of adult control and *Ubr5*^{CKO} mice (n=8 for both genotypes, **P*<0.01, Mann Whitney U-test). (B) Photograph showing that testes from *Ubr5*^{CKO} animals are smaller in size than control animals. Scale bar=500µm. (C) Boxplot showing that median epididymal sperm count is significantly reduced in *Ubr5*^{CKO} mice (n=6, **P*<0.01, Mann Whitney U-test). (D) Image of a cross section of adult control and *Ubr5*^{CKO} epididymis after haematoxylin and eosin staining. Tubules contain visibly less mature spermatozoa and confirm the sperm counts as indicated by the arrows. Scale bar=200µm.

5.2.6 *Ubr5^{CKO}* animals are fertile and produce offspring at Mendelian ratios

To investigate if the severe testicular atrophy and 5.2 fold reduction in sperm count in adult *Ubr5^{CKO}* animals rendered the mice infertile, three 5 month *Ubr5^{CKO}* males were mated with wild type C57BL/6 females. Each mating pair was allowed to produce two litters. Offspring were born at the expected 1:1 ratio of *Ubr5$^{wt/flox}$* and *Ubr5$^{wt/del}$* (table 5.2) and the transmission of the *Stra8-Cre* transgene was also 1:1 as expected as transmission was from the father. There was an average litter size of 10 pups, indicating that the *Ubr5^{CKO}* males are fertile despite the reduction in sperm count. This also suggests that *Ubr5* is essential for spermatogenesis and that the sperm contributing to offspring represent germ cells where excision of the *Ubr5^{flox}* allele has not occurred.

All Pups	Number of Pups Observed	Number of Pups Expected
Stra8- wt/flox	13	15.25
Stra8- wt/del	20	15.25
Stra8+ wt/flox	16	15.25
Stra8+ wt/del	12	15.25
Total	61	61
Chi test P value	0.467929938	

Table 5.2: Observed and expected ratios of pups from *Ubr5^{CKO}* males and WT females. Table indicating the number of pups observed from mating experiment and their possible genotypes. Stra8- or + indicates the absence or presence of the Stra8-Cre transgene in offspring. wt/flox or wt/del indicates the combination of *Ubr5* alleles each offspring may inherit. The Chi^2 test represents the ratios of genotypes in the table compared to their expected ratios of 1:1. The numbers of genotypes do not differ from the expected ratios.

5.2.7 *Ubr5*^{CKO} animals have severe testicular atrophy and increased apoptosis

Spermatogenesis occurs in a synchronised wave of development from mitotic spermatogonia lining the edge of the tubule, followed by meiotic progression via reductional chromosome divisions of primary spermatocytes, equational division of secondary spermatocytes followed by mature spermatozoa being released into the lumen (figure 1.11)(Russell et al., 1990). To investigate whether the reduced testis weight and sperm count reflected defective spermatogenesis, testes were fixed for histology and stained with haematoxylin and eosin. In control animals the testis histology is normal with the seminiferous tubules containing sertoli cells, spermatogonia, spermatocytes in meiotic prophase, metaphase nuclei and round and elongating spermatids (figure 5.7A&B).

In contrast, *Ubr5*^{CKO} animals have a severe phenotype featuring many distinct vacuoles generally visible in the meiotic layer of the seminiferous tubules, indicating severe testicular atrophy in the absence of UBR5 (figure 5.7A&B). Despite these vacuoles, all cells types could be identified in many tubules (figure 5.7B) however, as in the *Tex19.1*$^{-/-}$ testes some tubules were more severely degenerated than others, at both 6 week and 5 month time points. To quantify the severity of the testicular atrophy, I visually assessed 140 tubules from 3 control and 3 *Ubr5*^{CKO} 5 month old testes, scoring tubules with the criteria of no vacuoles, one or more vacuoles or sertoli cell only tubules. Control animals had 99% intact tubules with very few vacuoles (figure 5.7C). The tubule degeneration in *Ubr5*^{CKO} animals had varying degrees of severity with the most severely affected testis having 27% sertoli cell only tubules compared to only 1% for the least severely affected animal (figure 5.7C). Vacuoles were also detected in 21dpp *Ubr5*^{CKO} testes indicating that defects are present even at early stages and are no more severe at adult time points (data not shown). This indicates that *Ubr5* deletion induces testicular atrophy with varying severity.

The large quantity of vacuoles suggested that there could be increases in germ cell death in the absence of UBR5. I therefore performed the TUNEL assay to label dying

cells in 6 week old testes. Control animals had an average of 0.64 TUNEL positive cells per tubule compared to 1.75 TUNEL positive cells per tubule in *Ubr5*^{CKO} testes (figure 5.7D)(n=3, Students T-test $P<0.01$). This indicates that the absence of UBR5 results in increased cell death during progression through spermatogenesis.

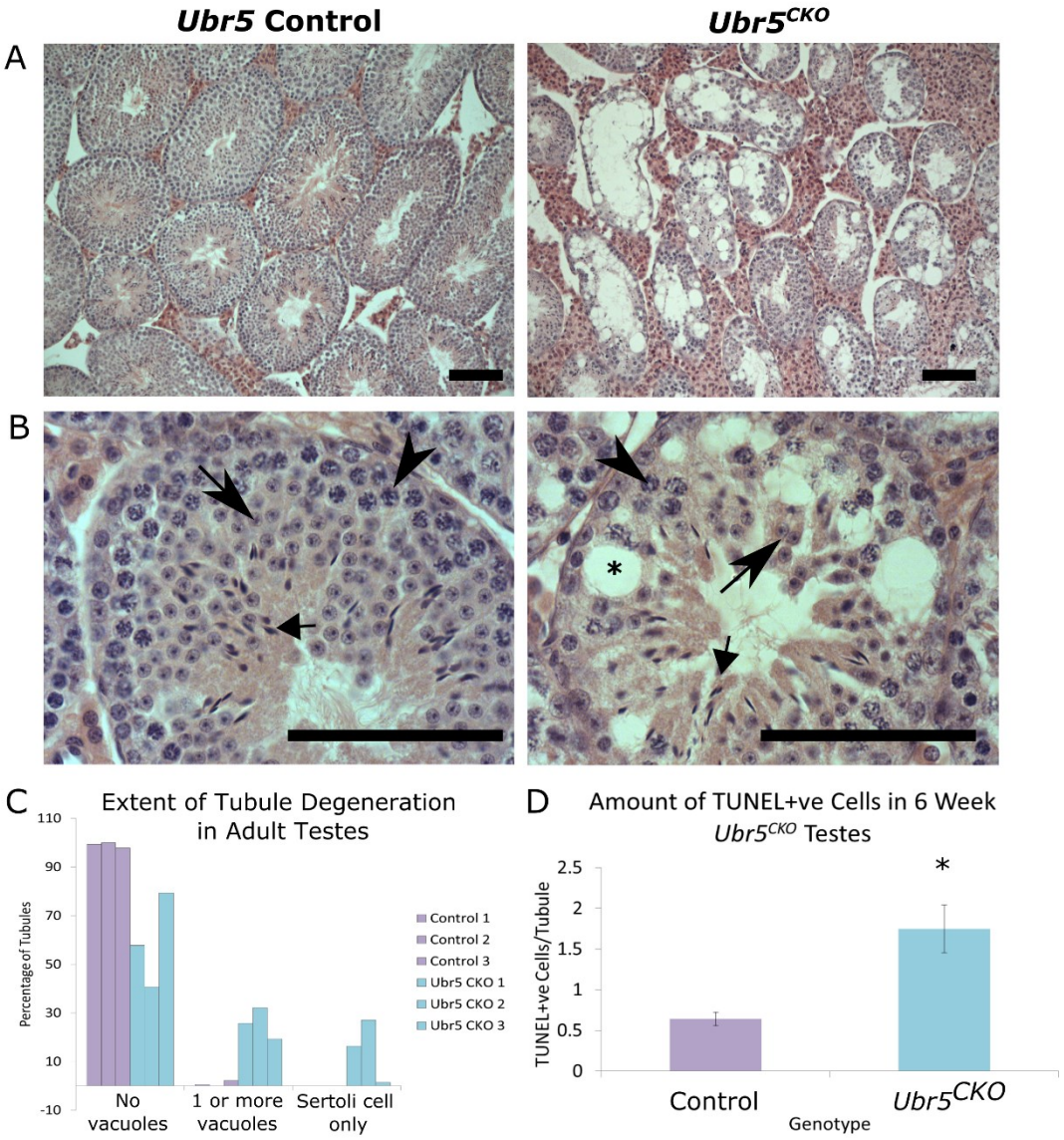


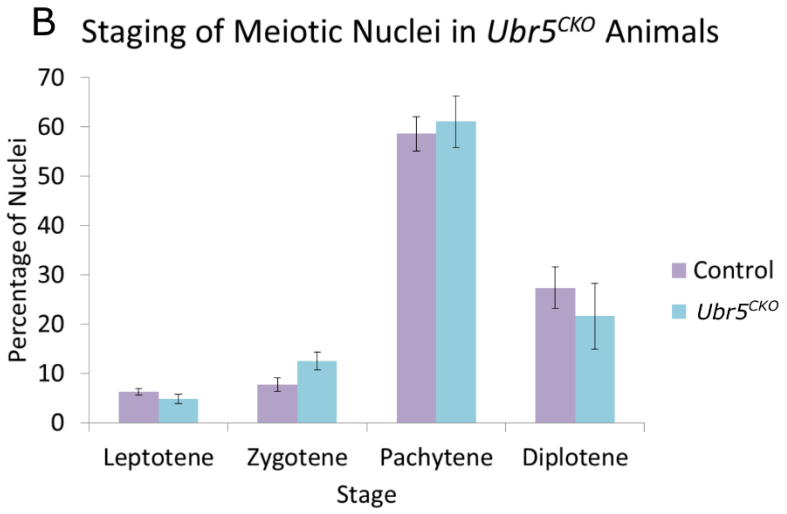
Figure 5.7: *Ubr5*^{CKO} testes have severe seminiferous tubule degeneration of varying severities. (A) H+E stained cross section of adult *Ubr5* control and *Ubr5*^{CKO} testis. Distinct voids can be seen in many tubules of the *Ubr5*^{CKO} testis which are absent from the seminiferous tubules of control animals. (B) Higher magnification image of H+E stained testis. Pointed arrow head indicates meiotic spermatocytes in prophase I showing distinct staining of meiotic chromosomes, pointed arrow indicates round spermatids with a distinct round morphology and the absence of dense nuclear staining and triangle arrowheads indicate elongating spermatids maturing proximal to the lumen, prior to release. The * in the *Ubr5*^{CKO} cross section indicates a void in the seminiferous tubule. Scale bars for all images represent 100µm. (C) Graph showing the proportion of tubules with or without voids or being sertoli cell only (n=3). (D) Graph showing the number of TUNEL positive cells per seminiferous tubule (n=3, *P<0.05 Students T-test)

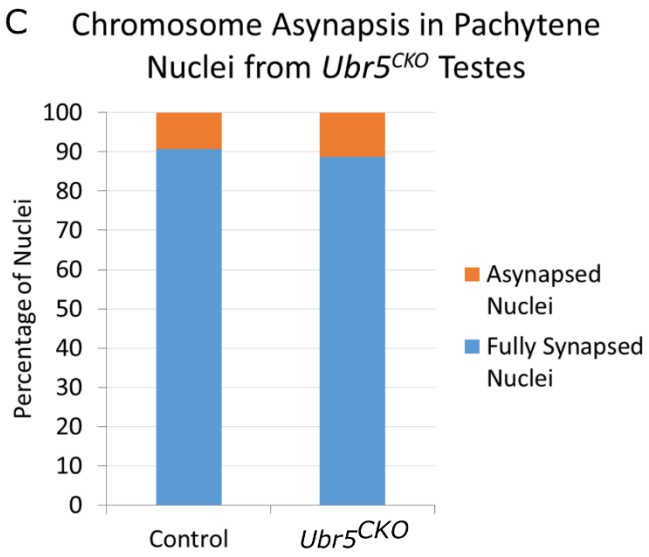
5.2.8 No detectable pachytene arrest occurs in the absence of UBR5

Nuclear morphology and location of the TUNEL positive cells was insufficient to determine their developmental stage, therefore I aimed to determine if the increases in cell death could be due to defects in the progression through meiotic prophase, similar to the meiotic pachytene arrest in *Tex19.1*$^{-/-}$ testes. To investigate this, meiotic chromosome spreads from three 6-7 week old *Ubr5*^{CKO} and control testes were stained with antibodies for SCP1 and SCP3 to mark transverse filaments and axial/lateral elements of the synaptonemal complex respectively, to determine meiotic stages (figure 5.8A). Around 100 nuclei were staged per animal and no significant differences in the proportion of nuclei in specific substages of meiotic prophase were detected (figure 5.8A&B). This suggests that the defects leading to the severe testicular atrophy must be occurring before or after meiotic prophase. No increase in chromosome asynapsis was observed for *Ubr5*^{CKO} pachytene nuclei, further indicating that the *Ubr5*^{CKO} and *Tex19.1*$^{-/-}$ phenotypes are distinct (figure 5.8A&C). A caveat with this immunocytochemistry analysis is that I have been unable to co-stain chromosome spreads with UBR5 as the antibodies I used for immunocytochemistry on testis sections did not work on chromosome spreads. It is therefore possible that the meiotic nuclei analysed had a functioning copy of *Ubr5*^{flox} that was incompletely excised, however at least some UBR5 negative spermatocytes were detected on testis sections (figure 5.5B). This would mean *Ubr5* null spermatocytes may have been underrepresented in the data, therefore a role for *Ubr5* in meiosis cannot be fully excluded.

Figure 5.8: Absence of *Ubr5* does not cause meiotic pachytene arrest or any increases in chromosome asynapsis (figures on subsequent pages) (A) Pachytene and diplotene spermatocytes from *Ubr5* control or *Ubr5*^{CKO} mice stained with axial element marker SYCP3 (red) and the central element marker (SYCP1). Genotype and antibody is represented at the left and above the images respectively. Co-localisation of SYCP3 and SYCP1 indicates chromosome synapsis in pachytene images (arrow heads). Pachytene was determined by the detection of at least one fully synapsed chromosome (arrow heads). Sex chromosomes synapsed at their pseudoautosomal region are indicated by triangle arrows. The white dashed line separates two nuclei in close proximity. Representative images of diplotene spermatocytes are shown to indicate that *Ubr5*^{CKO} spermatocytes can progress beyond pachytene. Pointed arrows indicate autosomal chromosomes undergoing de-synapsis in both *Ubr5* control and *Ubr5*^{CKO} spermatocytes (Scale bar = 50µm) (B) Graph showing the percentage of spermatocytes in each particular stage in meiotic prophase in control and *Ubr5*^{CKO} animals. Legend is shown on the left. Control and *Ubr5*^{CKO} animals n=3, spermatocytes scored=~100). (C) Graph indicating the percentage of asynapsed spermatocytes from control and *Ubr5*^{CKO} animals. (Number of control and *Ubr5*^{CKO} animals n=3, spermatocytes scored = 50)

A
Pachytene
Diplotene
Ubr5 Control
Ubr5CKO
Ubr5 Control
Ubr5CKO
SYCP1
SYCP3
Merge


B
Staging of Meiotic Nuclei in Ubr5CKO Animals
Percentage of Nuclei
70
60
50
40
30
20
10
0
Leptotene
Zygotene
Pachytene
Diplotene
Stage
Control
Ubr5CKO


C
Chromosome Asynapsis in Pachytene Nuclei from Ubr5CKO Testes
Percentage of Nuclei
100
90
80
70
60
50
40
30
20
10
0
Control
Ubr5CKO
Asynapsed Nuclei
Fully Synapsed Nuclei

5.2.9 *Ubr5*^{CKO} animals have reduced PLZF positive spermatogonia

As no defects were observed in meiotic spermatocytes I explored the possibility that loss of *Ubr5* may affect the number of spermatogonia in the testis. I therefore performed immunostaining of testis sections with antibodies to the transcription factor PLZF. PLZF is reported to be expressed in undifferentiated spermatogonia and in the absence of PLZF male mice exhibit a progressive testicular degeneration and age-dependent loss of germ cells, due to an inability to maintain the spermatogonial stem cell population (Costoya et al., 2004). Consistent with previous reports, cells with strong anti-PLZF staining in their nuclei were present in the basal layer of the seminiferous tubules where undifferentiated spermatogonia reside (figure 5.9C white arrows). There was no clear enrichment of these PLZF-positive cells in specific seminiferous tubule sub-stages. PLZF staining was confined to the nuclei of cells at the basal layer in both control and *Ubr5*^{CKO} testis. Quantification of the number of these positive cells indicated control animals had an average of 7.2 PLZF positive spermatogonia per tubule whereas *Ubr5*^{CKO} tubules had a significant 1.5 fold reduction to 4.7 PLZF positive spermatogonia per tubule (figure 5.9 & figure 5.10A)(n=3, P=0.034 Students T-Test). The reduction in PLZF-positive cells presumably reflects a loss of undifferentiated spermatogonia in *Ubr5*^{CKO} testes. However, it is also possible that the reduction in PLZF staining reflects a loss of *Plzf* expression rather than a loss of undifferentiated spermatogonia. Immunostaining with additional markers of undifferentiated spermatogonia would help to distinguish between these possibilities. Together, this analysis suggests that UBR5 has a role in maintaining undifferentiated spermatogonia populations in mouse testes. This is in contrast to *Tex19.1*$^{-/-}$ testes where there is change in the number of PLZF-positive cells (Ollinger et al., 2008).

I previously noted that the severity of tubule degeneration was variable between *Ubr5*^{CKO} animals. PLZF counts were also variable between *Ubr5*^{CKO} animals (figure 5.10A), I therefore aimed to investigate if the severity of the tubule degeneration and number of PLZF positive spermatogonia were related. Interestingly, the average number of PLZF positive spermatogonia per tubule is positively correlated with the differing levels of tubule degeneration in figure 5.7C (figure 5.10B, R^2=0.8969).

Taken together, these data suggest that tubule degeneration in *Ubr5*^{CKO} testes may be due to the reduced number of PLZF spermatogonia.

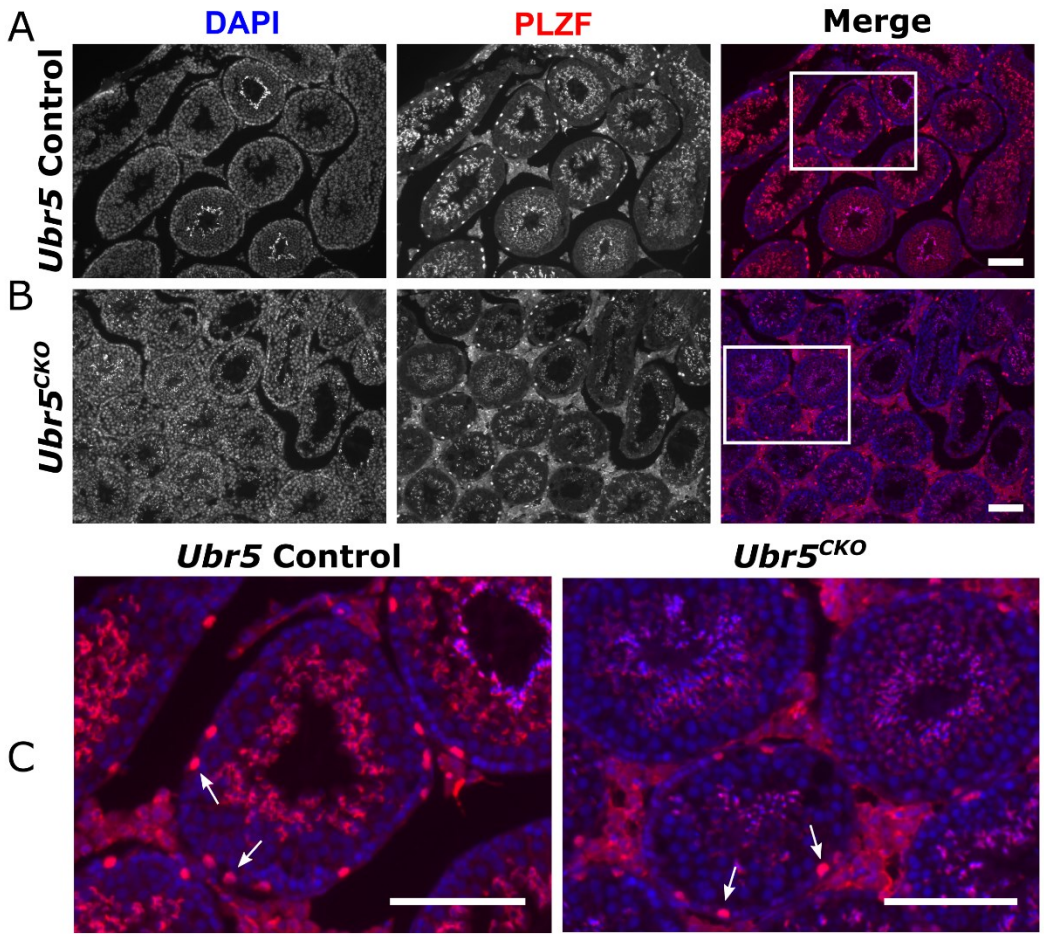


Figure 5.9: The number of PLZF stained spermatogonia are reduced in adult *Ubr5*^{CKO} testes (A) Cross section of seminiferous tubules of adult *Stra8-Cre Ubr5*$^{wt/flox}$ (*Ubr5* control) testis and (B) *Ubr5*^{CKO} testis stained for DAPI and PLZF. (Scale bar=100µm). (C) Higher magnification of white boxes in merged images from A and B. (Scale bar=100µm)

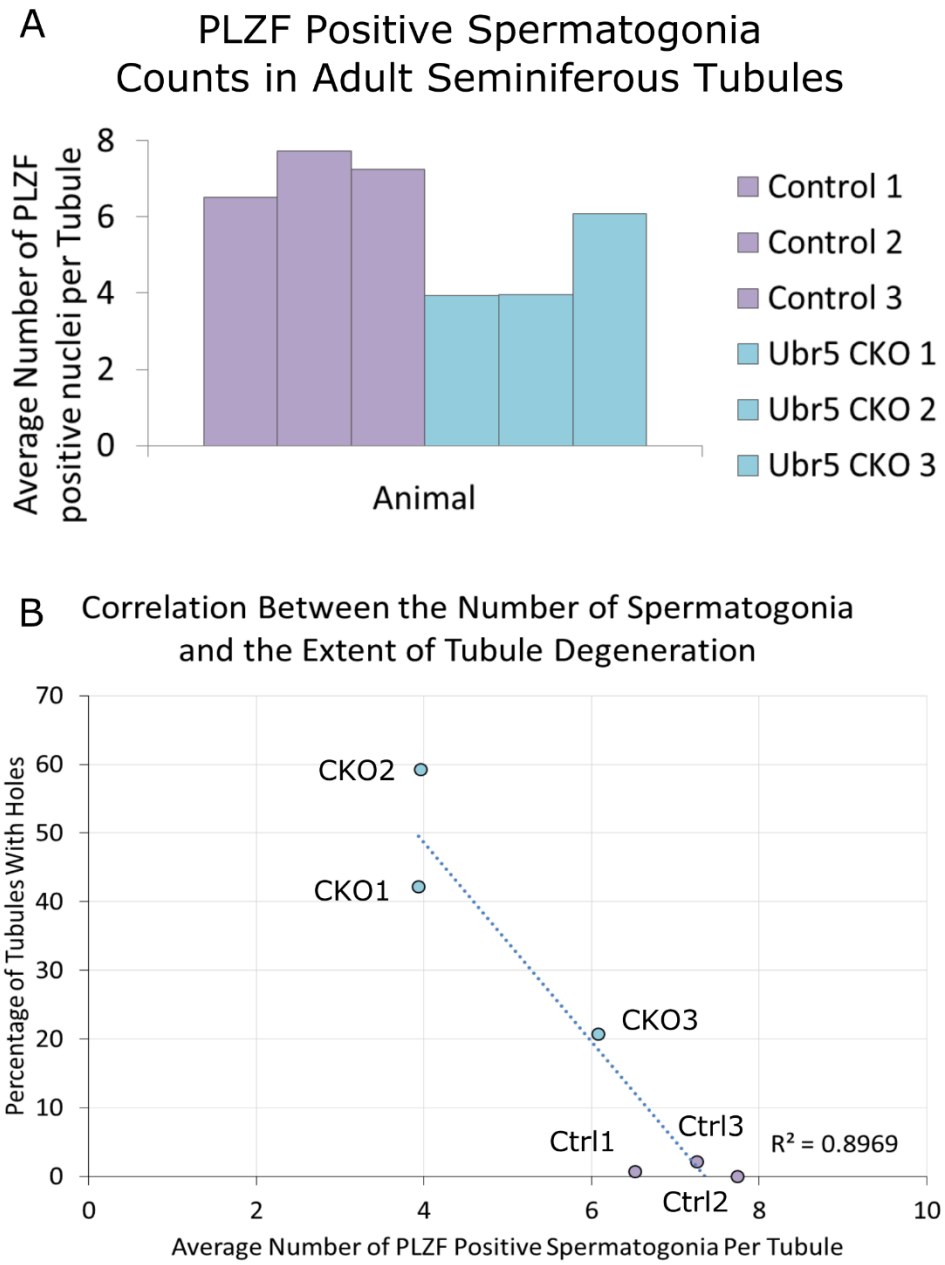


Figure 5.10: The decrease in spermatogonia is correlated to the extent of tubule degeneration in adult *Ubr5^{CKO}* testes (A) Graph indicating the average number of PLZF positive spermatogonia per tubule of adult *Stra8-Cre Ubr5$^{wt/flox}$* (Control) and *Ubr5^{CKO}* testis (n=3). 30 tubules were analysed per animal. (B) Scatterplot indicating correlation between the average number of PLZF positive spermatogonia per tubule from figure 5.10A and the extent of tubule degeneration from figure 5.7C for control (Ctrl) and *Ubr5^{CKO}* (CKO) testis. Animal number on the graph corresponds to animals in A. R^2=0.8969.

5.2.10 MMERVK10C is transcriptionally regulated in the testes by UBR2 and not UBR5

Many mouse mutants with defective spermatogenesis also have de-repression of retrotransposons (Crichton et al., 2014). Previous work from Ollinger et al (Ollinger et al., 2008) has shown that RNA for the LTR retrotransposon, MMERVK10C is increased in *Tex19.1*$^{-/-}$ testes and experiments in this thesis indicate this is likely at the transcriptional level (chapter 3 - figure 3.2A). To investigate if this transcriptional repression could be mediated by one of the TEX19.1 interacting partners from the UBR family I performed *in situ* hybridisation for MMERVK10C *env* transcripts on *Ubr2*$^{-/-}$ and *Ubr5*^{CKO} adult testis and the corresponding control animals. *Tex19.1*$^{-/-}$ testes were used as a positive control.

In control animals of each genotype, low levels of MMERVK10C *env* transcripts were present in a subset of seminiferous tubules as has been reported previously (figure 5.11 & figure 5.12)(Ollinger et al., 2008). Adult *Tex19.1*$^{-/-}$ testes acted as a positive control for the detection of MMERVK10C RNA and increased levels of transcript were detected in some *Tex19.1*$^{-/-}$ seminiferous tubules (figure 5.11A & 5.12A)(Ollinger et al., 2008). Interestingly, *Ubr2*$^{-/-}$ testes also have increased MMERVK10C RNA in some tubules compared to control testes (figure 5.11B & 5.12B). In contrast, *Ubr5*^{CKO} testes had little increase in MMERVK10C RNA compared to control testes (figure 5.11C). Importantly, all samples probed with a negative control MMERVK10C sense probe showed no staining (figure 5.11). .

MMERVK10C is de-repressed in meiotic spermatocytes of *Tex19.1*$^{-/-}$ animals (Ollinger et al., 2008). To determine if this was also the case in *Ubr2*$^{-/-}$ testes, higher magnification images were taken. *Tex19.1*$^{+/-}$ control testes indicated very modest staining in spermatocytes of a low number of tubules (figure 5.12A). The increased levels of MMERVK10C in *Tex19.1*$^{-/-}$ testes appear to be due to strongly expressing cells towards the lumen region of the tubules, likely corresponding to meiotic spermatocytes as previously determined (figure 5.12A – black arrows)(Ollinger et al., 2008). In *Ubr2*$^{+/-}$ sections as with *Tex19.1*$^{+/-}$ testes, faint staining could be detected in meiotic spermatocytes of some tubules (figure 5.12B). Interestingly *Ubr2*$^{-/-}$ testes have

a defined increase in MMERVK10C expression in the layer corresponding to meiotic spermatocytes when compared to *Ubr2*$^{+/-}$ testes (figure 5.12B black arrows). For *Ubr5*^{Ctrl} tubules weak staining was detected in two different locations in differing seminiferous tubules. Staining for MMERVK10C RNA was detected in basal regions, likely corresponding to meiotic spermatocytes of some tubules, but was also present closer to the lumen in other tubules, possibly corresponding to staining in round spermatids (figure 5.12C). No increase in the amount of MMERVK10C transcript was detected in *Ubr5*^{CKO} testes compared to control testes (figure 5.12C), however, some staining was detected possibly in round spermatids near to the lumen of the tubule. MMERVK10C signal was also detected in the spermatocytes of some tubules in *Ubr5*^{CKO} testes, but to a similar extent as *Ubr5*^{Ctrl} tubules (figure 5.11C & 5.12C). Together, these data identify UBR2, but not UBR5, as a novel transcriptional repressor of retrotransposons in the germline. This also further indicates a distinct overlap between the *Tex19.1*$^{-/-}$ and the *Ubr2*$^{-/-}$ phenotypes, but differences with the *Ubr5*^{CKO} phenotype.

To determine if UBR2 is functioning in transcriptional repression of retrotransposons specifically in the testes or if this extends to other tissues, I performed qRT-PCR expression analysis for retrotransposons on RNA extracted from P16 *Ubr2*$^{-/-}$ and control testes and cerebellum. *Ubr2*$^{-/-}$ testes had a 1.5-fold increase in MMERVK10C expression compared to *Ubr2*$^{+/-}$ testes, which was absent from the cerebellum (P=0.003, Students T-test)(figure 5.12D). This indicates the transcriptional repressive capacity of UBR2 may be specific to the germ line however, the mechanism for this tissue specific repression is still unclear. No de-repression was observed for LINE-1, however a small but significant 1.2 fold de-repression of VL30 (P=0.02, Students T-test) was detected in *Ubr2*$^{-/-}$ testes compared to *Ubr2*$^{+/-}$ testes indicating that in the testes, UBR2 specifically represses not only MMERVK10C, but also VL30 (figure 5.12D). This increase in expression of VL30 was not detected in *Ubr2*$^{-/-}$ cerebellum and has not been reported in *Tex19.1*$^{-/-}$ testes (figure 5.12D)(Ollinger et al., 2008).

In *Tex19.1*$^{-/-}$ testes, the increase in MMERVK10C expression is around 4 fold, compared to the modest 1.5 fold expression increase observed in *Ubr2*$^{-/-}$ animals (Ollinger et al., 2008). It is possible that this may reflect differences in genetic background as repetitive element expression is known to be variable between mouse

strains (Nellåker et al., 2012). Previously MMERVK10C expression was assessed in *Tex19.1*$^{-/-}$ mice on a 129/Ola X CD1 mixed genetic background (Ollinger et al., 2008), whereas *Ubr2*$^{-/-}$ mice are of the C57/BL/6 background. In line with this hypothesis, qRT-PCR for MMERVK10C in *Tex19.1*$^{-/-}$ testes on a C57/BL/6 background show a 2 fold increase in RNA (Reichmann et al., 2012), similar to the 1.5 fold increase observed in *Ubr2*$^{-/-}$ testes (figure 5.12D). Together, this provides the first evidence of the transcriptional repressive capability of UBR2 for regulating retrotransposons in the germline.

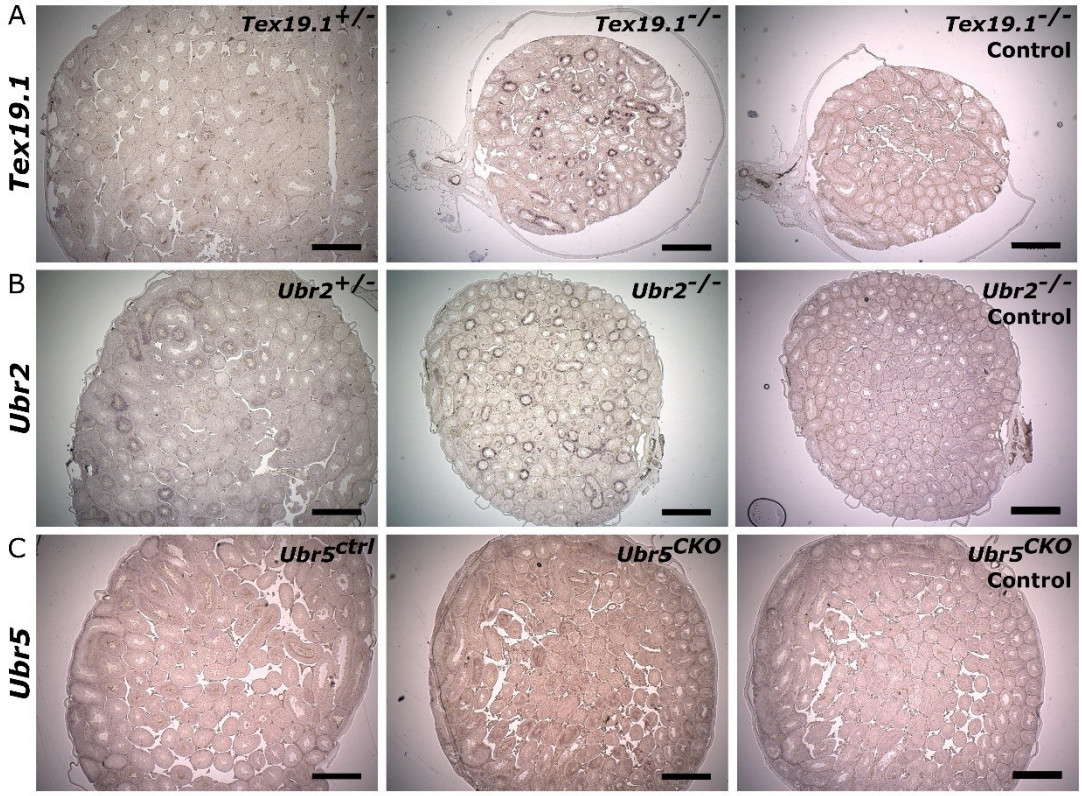


Figure 5.11: Expression of MMERVK10C is upregulated in *Ubr2*$^{-/-}$ testes. (A) *In situ* hybridisation with an antisense MMERVK10C probe (brown precipitate) in *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$, (B) *Ubr2*$^{+/-}$ and *Ubr2*$^{-/-}$, (C) *Ubr5*^{Ctrl} (Stra8-Cre *Ubr5*$^{wt/flox}$) and *Ubr5*^{CKO} testes from littermates. Images show the whole cross section of one testis to indicate the extent of MMERVK10C staining. The column on the right (*Tex19.1*$^{-/-}$ Control, *Ubr2*$^{-/-}$ Control and *Ubr5*^{CKO} Control) shows representative images from control ISH with a sense probe and indicates no staining in any tubules. Scale bar=500µm.

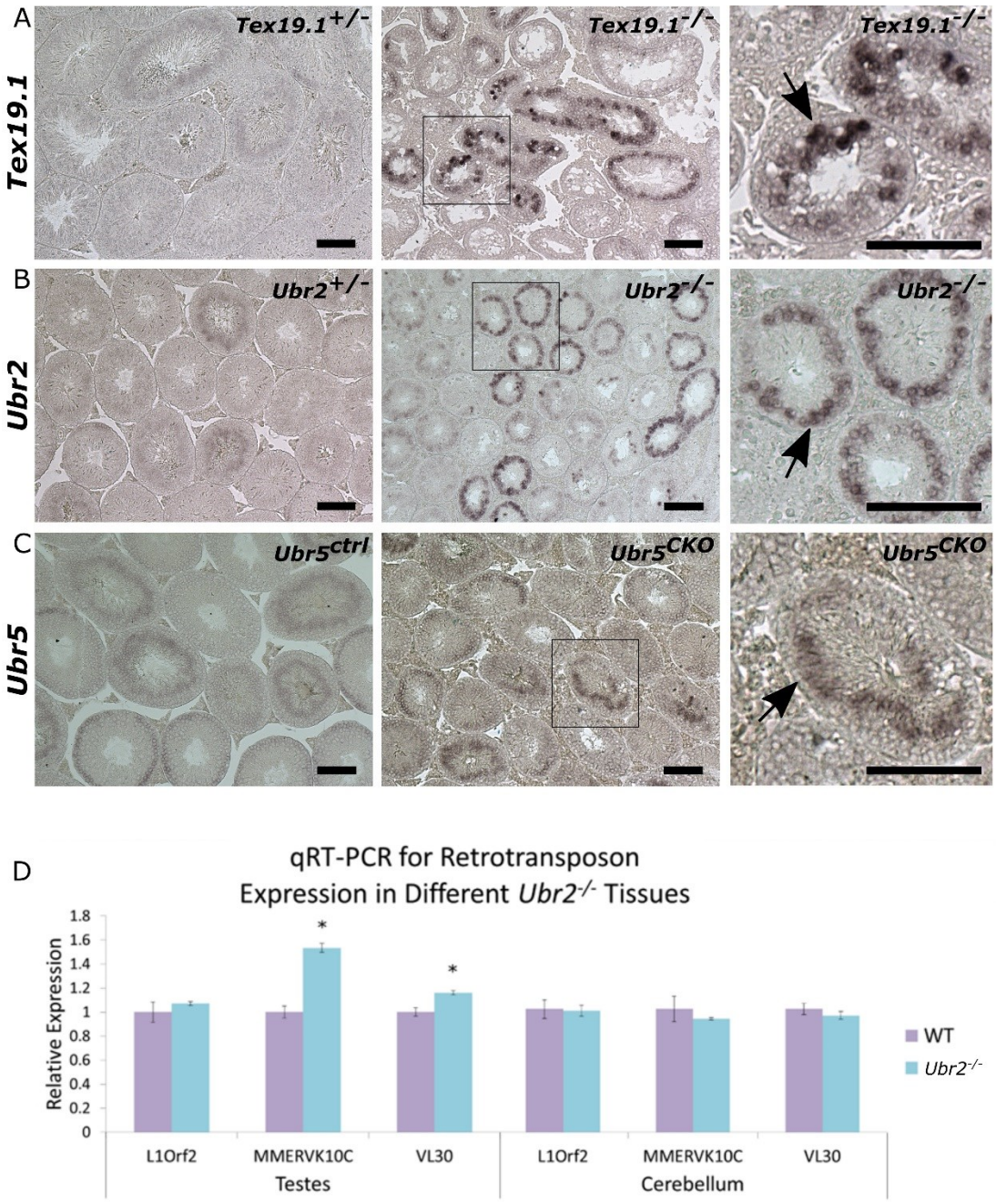


Figure 5.12: UBR2 represses MMERVK10C specifically in meiotic spermatocytes and not the cerebellum (A) A) Higher magnification images of *In situ* hybridisation with an antisense MMERVK10C probe (brown precipitate) in *Tex19.1*$^{+/-}$ and *Tex19.1*$^{-/-}$, (B) *Ubr2*$^{+/-}$ and *Ubr2*$^{-/-}$, (C) *Ubr5*^{Ctrl} (Stra8-Cre *Ubr5*$^{wt/flox}$) and *Ubr5*^{CKO} testes from littermates. Boxes in middle column indicate the images shown on the right hand column at higher magnification. Black arrows indicate positively stained spermatocytes. Scale bar=100µm. (D) qRT-PCR on RNA extracted from whole P16 *Ubr2*$^{-/-}$ testes and cerebellum for retrotransposon transcripts. Data is normalised to *β-actin* and expressed relative to *Ubr2*$^{+/+}$ (*Ubr2*$^{+/+}$ n=2, *Ubr2*$^{+/-}$ n=3, **P*<0.05 Students T-test)

5.3 Discussion

In this chapter I aimed to gain an understanding of the potential contribution of two TEX19.1 interacting partners, UBR5 and UBR2, to the *Tex19.1*$^{-/-}$ testicular phenotype (Ollinger et al., 2008). I show for the first time that UBR5 is essential for spermatogenesis, as conditional deletion of *Ubr5* early in spermatogenesis results in reduced testicular weight, reduced sperm count and large vacuoles in some seminiferous tubules of *Ubr5*^{CKO} animals, accompanied by an increase in cell death (summarised in table 5.3). I have also identified a role for UBR5 in maintaining the undifferentiated spermatogonia population which is likely contributing to the severe testicular degeneration in the absence of UBR5 (table 5.3). Together, these data suggest distinct differences between *Ubr5*^{CKO} and *Tex19.1*$^{-/-}$ testes phenotypes which was extended to the lack of MMERVK10C retrotransposon de-repression in *Ubr5*^{CKO} animals (table 5.3). Conversely, I have shown that repression of MMERVK10C in the testes is dependent on UBR2, suggesting overlapping roles of TEX19.1 and UBR2 in retrotransposon repression in this context.

Phenotype Assessed	*Ubr5*^{CKO} Phenotype	*Tex19.1*$^{-/-}$ Phenotype	Figure
Testis weight	Decreased	Decreased	5.6
Sperm count	Decreased	Decreased	5.6
Testis morphology	Severe atrophy	Severe atrophy	5.7
Fertility	Normal	Sterile	Table 5.2
Cell death	Increased	Increased	5.7
Proportion of meiotic substages	Normal	Arrested at pachytene	5.8
Chromosome synapsis	Normal	Asynapsis	5.8
Spermatogonia numbers	Reduced	Normal	5.9
MMERVK10C expression	Normal (Increased in *Ubr2*$^{-/-}$ testes)	Increased	5.11 & 5.12

Table 5.3: Summary table of *Ubr5*^{CKO} phenotypes compared to *Tex19.1*$^{-/-}$ phenotypes

5.3.1 UBR5 is essential for spermatogenesis and maintains spermatogonial stem cells in mouse testes

Quantifying the number of PLZF positive undifferentiated spermatogonia in *Ubr5*^{CKO} testes identified a role of UBR5 in maintenance of undifferentiated spermatogonia. This is distinct from *Tex19.1*$^{-/-}$ mice which have similar PLZF positive spermatogonia numbers as control testes (Ollinger et al., 2008). Mice null for *Zfp145,* which encodes PLZF, undergo progressive testicular degeneration in an age dependent manner due to an inability to maintain the spermatogonial stem cell population (Costoya et al., 2004). This phenotype is severe with many adult tubules being devoid of germ cells (Costoya et al., 2004). Indeed the number of PLZF positive spermatogonia per tubule in *Ubr5*^{CKO}

animals correlated with severity of tubule degeneration, however it is difficult to determine if this correlation is a direct effect of loss of *Ubr5* on undifferentiated spermatogonia, or if spermatogonia are being lost in a non-cell autonomous manner. This may occur as a direct result of disrupted paracrine signalling, such as the hedgehog pathway or due to composition differences (Makela et al., 2011). Analysing $Ubr5^{CKO}$ testes earlier in development before global composition alterations would address if this decrease in PLZF positive spermatogonia is directly contributing to the defects in progression through spermatogenesis.

Paracrine hedgehog signalling has been shown to play a major role in spermatogenesis (Bitgood et al., 1996). Desert hedgehog (*Dhh)* expression is initiated and maintained in Sertoli cells of the testes and is a small signalling molecule essential for spermatogenesis (Bitgood et al., 1996). Mice null for *Dhh* exhibit loss of germ cells and a severe 90% decrease in testicular mass by six weeks of age, however the molecular pathology is unclear (Bitgood et al., 1996). Generally when hedgehog molecules reach a target cell it binds to its receptor, PTCH1, which then inhibits Smoothened (SMO*)* and consequently GLI transcription factors which regulate normal gene expression (Briscoe and Therond, 2013).

Interestingly, UBR5 has roles in the maintenance of other stem cell populations such as chondrocytes in bone development (Mark Ditzel, personal communication). Conditional removal of *Ubr5* from chondrocyte precursors caused a decreased number of chondrocytes, along with perturbed expression of members of the hedgehog signalling pathway, *Ihh, Ptch* and *Gli* (Mark Ditzel, personal communication). Also, *Ubr5* has been determined as a negative regulator of hedgehog expression in *Drosophila* (Lee et al., 2002; Moncrieff et al., 2015). Therefore, misregulation of UBR5 mediated hedgehog paracrine signalling may be an additional mechanism leading to defective spermatogenesis in $Ubr5^{CKO}$ testes, which requires further investigation with antibody staining or gene expression analysis.

Alternatively, UBR5 has been demonstrated to have key roles in mitosis (Scialpi et al., 2015). The UBR5 protein physically interacts with components of the spindle assembly checkpoint and depletion of *Ubr5* by treatment with siRNA dramatically reduced mitotic cell viability (Scialpi et al., 2015). Undifferentiated spermatogonia are

proliferative cells, therefore if mitosis is perturbed in the absence of *Ubr5* it may be contributing to the loss of undifferentiated spermatogonia.

5.3.2 UBR2 represses MMERVK10C in spermatocytes

UBR2 has previously been shown to have roles in spermatogenesis and *Ubr2*$^{-/-}$ testes have been proposed to phenocopy *Tex19.1*$^{-/-}$ mice (Kwon et al., 2003). Furthermore, *Tex19.1* is transcribed in *Ubr2* null germ cells but TEX19.1 protein is absent, indicating that UBR2 stabilises TEX19.1 protein in the germline (Yang et al., 2010). Data from Marie MacLennan indicates that TEX19.1 is not required for the stability of UBR2 in the testes, as similar UBR2 protein levels to wild type are detected in *Tex19.1*$^{-/-}$ testes (Marie MacLennan personal communication). This suggests that the phenotypes occurring in both *Ubr2*$^{-/-}$ and *Tex19.1*$^{-/-}$ testes are possibly due to specific roles of TEX19.1, which is absent in both mutants, or both proteins acting together in a complex. Subunits of many protein complexes are unstable if differing binding partners are absent (Eskeland et al., 2010). I indicate for the first time that UBR2 is essential for repression of MMERVK10C in the testes, however as with other spermatogenic phenotypes of *Ubr2*$^{-/-}$ mice, this may be a consequence of loss of TEX19.1 (Yang et al., 2010). The differences in the extent of de-repression of MMERVK10C in each mutant may be a result of the different genetic background of the strains however, direct roles for UBR2 in transcriptional repression of MMERVK10C cannot be ruled out.

TEX19.1 has been shown to enhance poly-ubiquitination of LINE-1 ORF1 protein via UBR2, however effects on ubiquitination of other potential substrates such as KRAB-ZFPs, many of which have retrotransposon targets, is yet to be investigated (Ecco et al., 2016; Wolf et al., 2015a). UBR2 has also been suggested to regulate transcription of meiotic genes by histone ubiquitination in meiosis, however only representative ICC images were shown indicating reductions in H2A ubiquitination (Young An et al., 2010). It is unclear how much differences in cell composition, caused by the meiotic arrest of *Ubr2*$^{-/-}$ spermatocytes, are contributing to the loss of H2A ubiquitination on transcriptional repression reported from this study (Young An et al.,

2010). The mechanism of TEX19.1/UBR2 mediated transcriptional repression of MMERVK10C in the germline is unclear and will require further study, however it is unlikely to involve UBR5 as no increase in MMERVK10C RNA was detected in *Ubr5*^{CKO} testes compared to control testes.

In summary, I have determined that UBR5 is essential for spermatogenesis and has a phenotype distinct from that of *Tex19.1*$^{-/-}$ testes. This suggests that the contribution of UBR5 to the *Tex19.1*$^{-/-}$ phenotype is minimal, with the retrotransposon de-repression in *Tex19.1*$^{-/-}$ spermatocytes likely being in part due to UBR2.

Chapter 6: Discussion

6.1 Mechanism of TEX19.1 mediated genome defence

The main aim of this thesis was to gain a better understanding of TEX19.1 mediated genome defence in hypomethylated contexts, from the mechanism of retrotransposon repression, to the possible developmental defects associated with their activation. Furthermore, I aimed to investigate the possible contribution of the TEX19.1 interacting partners to these *Tex19.1*$^{-/-}$ phenotypes.

Tex19.1$^{-/-}$ mice have numerous developmental defects, from meiotic pachytene arrest and male infertility, to defective placental development, with de-repression of different retrotransposons in both contexts (Ollinger et al., 2008; Reichmann et al., 2013). Modification of histones is a major mechanism contributing to retrotransposon repression in ES cells, however is thought to play a lesser role in differentiated cells, where repression is mainly facilitated by DNA methylation (Karimi et al., 2011; Matsui et al., 2010; Rowe et al., 2010). In the germline, the piRNA pathway is an additional mechanism which ensures retrotransposon repression via endonuclease 'slicer' activity and *de novo* methylation of retrotransposons (Bortvin, 2013). Despite these differing mechanisms, a certain amount of redundancy in their target loci ensures continued repression throughout development and into adult tissues (Goodier 2016).

One outstanding question arising when considering the *Tex19.1*$^{-/-}$ phenotype is why different retrotransposons are de-repressed in different *Tex19.1*$^{-/-}$ tissues. Is the *Tex19.1* mediated repressive mechanism the same in ES cells, testes and the somatic placenta? Are the transcriptional repressive interacting partners of TEX19.1 different in each tissue, possibly based on the different cytoplasmic and nuclear localisations of TEX19.1? Is redundancy in repressive mechanisms masking retrotransposon de-repression in the absence of *Tex19.1*?

In this thesis I have identified that TEX19.1 functions, at least in part, through the KAP1 repressive mechanism in ES cells to augment repressive histone modifications to compensate for loss of DNA methylation in development (figure 3.4). This only targets a subset of LINE-1 elements, L1MdF2, which are known KAP1 targets in ES cells (Castro-Diaz et al., 2014). These are de-repressed in the *Tex19.1*$^{-/-}$ placenta,

suggesting that the KAP1 repressive mechanism may be perturbed *in vivo* in the absence of *Tex19.1* (figure 3.13)(Castro-Diaz et al., 2014).

One inherent property of the KAP1 repressive mechanism is its targeting to specific genomic loci by KRAB-ZFPs (Wolf et al., 2015b). KRAB-ZFPs constitute the largest transcription factor family in mice and humans with around 200 and 300 genes respectively (Emerson and Thomas, 2009). This provides redundancy in the KRAB-ZFPs targeting specific retrotransposons loci and regulating their expression (Castro-Diaz et al., 2014; Ecco et al., 2016; Najafabadi et al., 2015). Previously, the KRAB-ZFP mechanism was thought to establish transcriptional repression in embryonic development, with the majority of repression in somatic differentiated tissues being due to DNA methylation (Wolf et al., 2015), however at least some KRAB-ZFP:KAP1 complexes can continually facilitate retrotransposon repression into adult cells, suggesting a more dynamic transcriptional regulation than previously appreciated (Ecco et al., 2016; Playfoot and Adams 2016). If TEX19.1 is affecting KRAB-ZFPs in the placenta is currently unclear but if so, it would provide another somatic context for their transcriptional repressive function. Indeed, many KRAB-ZFPs are highly expressed in evolutionarily young adult somatic tissues such as the mammalian placenta, but to a lesser extent in other cell types (Liu et al., 2014a). Furthermore, co-IP mass spectrometry indicates that TEX19.1 interacts with a known KRAB-ZFP, ZNF295, however any role for this ZFP in retrotransposon repression is yet to be established (Marie MacLennan, unpublished data).

Tex19.1$^{-/-}$ placenta has de-repression of a subfamily of LINE-1, L1MdF2 and the LTR element VL30, whereas testes have de-repression of a subfamily of the MMERVK10C elements (Ollinger et al., 2008; Reichmann et al., 2013). L1MdF2, VL30 and MMERVK10C are all known targets of different KRAB-ZFPs and KAP1 in ES and somatic cells (Castro-Diaz et al., 2014; Ecco et al., 2016; Wolf et al., 2015). In ES cells the specific KRAB-ZFP targeting KAP1 to L1MdF2 is Gm6871, which is distinct from the KRAB-ZFPs targeting VL30, ZFP809 and MMERVK10C, Gm15446 (Castro-Diaz et al., 2014; Ecco et al., 2016; Wolf et al., 2015). Each KRAB-ZFP is expressed to a similar extent in control and *Tex19.1*$^{-/-}$ placentas (data not shown). Interestingly, when I assessed the *Tex19.1*$^{-/-}$ placentas for additional retrotransposons repressed by TEX19.1, I identified a number of significantly upregulated retrotransposons which

are known targets of KRAB-ZFP and the H3K9me3 machinery. For example, the most upregulated element in *Tex19.1*$^{-/-}$ placenta RNAseq (figure 4.9) was MLT1A, the human homolog of which is bound by ZNF264 and ZNF519 in HEK293T cells (Najafabadi et al., 2015). In addition, the second most highly de-repressed element in *Tex19.1*$^{-/-}$ placentas, MER52, is also bound by KRAB-ZFPs in humans, specifically ZNF273 and is marked by H3K9me3 (Najafabadi et al., 2015). If these elements are regulated by KRAB-ZFPs in mouse remains to be seen, however this gives further support to the hypothesis that TEX19.1 may be playing an important role in this repressive mechanism depositing H3K9me3, to compensate for the low levels of DNA methylation in this tissue. The specific targeting of H3K9me3 mediated by TEX19.1 could therefore possibly be due to interactions with different KRAB-ZFPs, alongside KAP1, however it is unclear if TEX19.1 is binding with KRAB-ZFPs previously associated with specific de-repressed retrotransposons. Generally the majority of KRAB-ZFPs were not detected in the CoIP-mass spectrometry experiment by Marie MacLennan, however this was performed on the cytoplasmic fraction from ES cells. Therefore, this may not necessarily represent the physiological interacting partners detected in the placenta where TEX19.1 localisation appears more nuclear (Reichmann et al., 2013). Additionally, the permissive environment of the placenta due to reduced DNA methylation and the absence of other genome defence genes likely allows detection of retrotransposon repressed by TEX19.1, whereas any transcriptional de-repression of LINE-1 in the testis would presumably be prevented by the action of the PIWI proteins (Crichton et al., 2013).

A study by Collins et al., also highlighted that reduced levels of H3K9me3 upon loss of *Setdb1* in somatic cells does not necessarily result in de-repression of the affected retrotransposon loci (Collins et al., 2015). For example, loss of KAP1:SETDB1 dependent H3K9me3 at L1MdF2 loci does not result in their transcriptional activation in SETDB1 deficient B-lymphocyte cells (Fig. 2A). In these cells they determined that tissue specific transcription factors are necessary for the de-repression of at least some retrotransposons with reduced H3K9me3 but not others (Collins et al., 2015). This may provide another explanation as to the highly tissue specific de-repression of particular retrotransposons in different *Tex19.1*$^{-/-}$ tissues.

Kap1 is expressed in the placenta and its precursors, however any roles for KAP1 in retrotransposon repression or molecular aspects of development in this tissue are yet to be fully determined (Cammas et al., 2000; Shibata et al., 2011). ChIPseq experiments to assess KAP1 binding to the genome have never been performed on wild type placental tissue. Investigation of the retrotransposons de-repressed in the absence of *Kap1* in the placenta has also never been assessed as the embryonic lethality of global $Kap1^{-/-}$ embryos around E5.5, post-implantation has precluded any analysis (Cammas et al., 2000). It would be informative to perform a conditional *Kap1* knockout in placental lineages by crossing a *Tpbpa-Cre* mouse with a $Kap1^{flox/flox}$ mouse to delete *Kap1* in progenitor cells of differentiated trophoblast cells at around E8.5 (Simmons et al., 2007). This would allow analysis of the extent of KAP1 mediated retrotransposon repression, if any, in this tissue, but also of the roles of KAP1 in placental development. So far, in somatic contexts, KAP1 has roles repressing retrotransposons in neural progenitor cells (Fasching et al., 2015) and examples of retrotransposon de-repression in somatic tissues has only been detected in a few mouse genetic knockouts or knockdowns of KRAB-ZFPs with offspring generally being fertile and healthy, however the eaerly trophoblast or mature placentas have only been assessed for ZNF568 (Ecco et al., 2016; Wolf et al., 2015; Wolf et al., 2015b). A mutation resulting in hypomorphic expression of *Znf568* results in defective placental development correlated with defects in migration of the embryonic derived extraembryonic mesoderm leading to defective yolk sac development (Shibata and Garcia-Garcia, 2011). A direct role for ZNF568 in trophoblast development was not detected, however the vast number of KRAB-ZFPs and their generally high expression in the placenta could possibly indicate as yet undiscovered roles for these transcription factors in placental development (Emerson and Thomas, 2009; Liu et al., 2014a).

Furthermore, deletion of *Tex19.1, Kap1* or particular KRAB-ZFPs in trophoblast stem cells would provide an alternative, more homogenous cellular model to address their roles in both retrotransposon and gene regulation, along with placental development at earlier time points. Specific LTR retrotransposons have been shown to act as enhancer elements facilitating the control of key TSC lineage regulators such as *Cdx2*, *Eomes* and *Elf5,* therefore disruption of the expression of these elements early in development may have serious consequences in the progression of the trophoblast

lineages (Chuong et al., 2013). This would also address the caveat of the differences in composition detected in *Tex19.1*$^{-/-}$ placentas compared to controls which, for at least some genes, will be causative of their significant increases or decreases in RNA detected in the RNAseq experiment (figure 4.10). Undirected differentiation of *Tex19.1*$^{-/-}$ ES cells did not result in de-repression of any retrotransposons (figure3.10), however differentiation of *Tex19.1*$^{-/-}$ TS cells to TGCs may provide an *in vitro* model of retrotransposon de-repression in a hypomethylated somatic context. *Tex19.1*$^{-/-}$ ES cells do not de-repress retrotransposons in normally methylated conditions (figure 3.9) and upon hypomethylation, de-repression appears to be transient, with different dynamics for different retrotransposons, however the reason for this remains unclear (figure 3.12). Additionally, it is unclear if any of the genes changing in RNA abundance in *Tex19.1*$^{-/-}$ placentas are targets of KAP1. KAP1 ChIPseq on TS cells would provide an informative dataset to determine which genes and retrotransposons are targeted by KAP1, in a more relevant context to placenta than existing genome wide KAP1 datasets from ES cells and normally methylated somatic tissues.

At present, the molecular mechanism of TEX19.1 binding to KAP1 and promoting deposition of H3K9me3 is unclear. A number of differing mechanisms are possible. For example, TEX19.1 may be stabilising the interaction of KAP1 and KRAB-ZFPs to facilitate targeting of the complex to specific genomic loci. Alternatively, TEX19.1 binding to KAP1 may be disrupting its E3 ubiquitin ligase capacity as has been determined for the interaction of TEX19.1 and UBR2, whereby TEX19.1 inhibits the binding of UBR2 to its N-end rule substrates (Diana Best unpublished). Interactions between KAP1 and SETDB1 are sumoylation-dependent (Zeng et al., 2008), with sumolyated KAP1 being highly transcriptionally repressive, due to stimulation of the HMT activity of SETDB1 (Ivanov et al., 2007). If this sumoylation mechanism of KAP1 is disrupted in *Tex19.1*$^{-/-}$ animals leading to reduced activity of SETDB1 and therefore a reduced repressive capacity, is unclear. A future experiment to assess if TEX19.1 is affecting the sumoylation of KAP1 would be an *in vitro* sumoylation assay similar to the *in vivo* ubiquitination assays used in this thesis (Li et al., 2010). By co-transfecting cells with tagged KAP1 and SUMO expression constructs, with and without *Tex19.1,* any effect of TEX19.1 on KAP1 sumoylation levels would be detectable by Western blotting.

Interestingly, I have determined that UBR2 is a novel repressor of MMERVK10C transcription in mouse testes but has no role in transcriptional repression of LINE-1 in this tissue, similar to TEX19.1 (figure 5.11 & 5.12). Analysis of placentas in *Ubr2*$^{-/-}$ mice may be informative to test if there is any similarity with *Tex19.1*$^{-/-}$ phenotypes. Although IUGR has not been reported in *Ubr2*$^{-/-}$ mice, loss of *Ubr2* is associated with female embryonic lethality (Kwon et al., 2003) and it will be of interest to determine if there are placenta defects and IUGR in these embryos.

In testes UBR2 is necessary to stabilise TEX19.1 as TEX19.1 protein is undetectable in *Ubr2*$^{-/-}$ testes (Yang et al., 2010). In contrast, UBR2 is present at similar amounts in *Tex19.1*$^{-/-}$ testes as control testes (Marie MacLennan unpublished results). Despite this, it is unclear if UBR2 is required to stabilise TEX19.1 in different contexts such as the placenta. In the placenta TEX19.1 protein is largely detectable in the nucleus, compared to mainly the cytoplasm in germ cells (Ollinger et al., 2008; Reichmann et al., 2013). The significance of this difference in localisation is yet to be determined, however if some TEX19.1 protein is detectable in *Ubr2*$^{-/-}$ placentas it may suggest a different mechanism of retrotransposon repression is occurring in the placenta, distinct from the testes.

Another unexplored possibility may be that in both tissues, the association of TEX19.1 and UBR2 is having an effect on KAP1. This may extend to effects on the molecular function of KAP1 or perturbations of the binding to other proteins in the KAP1 repressive complex (Iyengar and Farnham, 2011). Indeed, both KAP1 and UBR2 are RING domain E3 ubiquitin ligases and TEX19.1 has a role in UBR2 dependent poly-ubiquitination of LINE-1 ORF1p, leading to its degradation (MacLennan et al., submitted). If KAP1 is a substrate for UBR2:TEX19.1 mediated ubiquitination has not been investigated, however there appears to be no reduction in global KAP1 protein levels or ubiquitination levels when TEX19 is overexpressed in human cells and no difference in KAP1 levels in the junctional zone of *Tex19.1*$^{-/-}$ placenta compared to control animals (figure 3.6, 3.7 & 3.8).

An alternative possibility is that the E3 ubiquitin ligase activity of KAP1 itself may be being perturbed in the absence of *Tex19.1*. KAP1 E3 ubiquitin ligase activity has been determined to be promoted by a subclass of proteins with similar expression patterns

as *Tex19.1* called the MAGE proteins (Doyle et al., 2010). MAGE proteins can promote KAP1 dependent ubiquitination of KRAB-ZFPs thereby increasing their turnover, similar to the role of TEX19.1 in promoting UBR2 dependent poly-ubiquitination and turnover of LINE-1 ORF1p (MacLennan et al., submitted; Xiao et al., 2011). If this is extending to KAP1 and the KRAB-ZFPs is unclear. Alternatively it is possible that TEX19.1 may be preventing MAGE proteins interacting with KAP1, thus leading to reduced ubiquitination of KRAB-ZFPs. If TEX19.1 was regulating KAP1 in this manner this would result in an increase in KRAB-ZFP proteins and therefore presumably an increased binding of KRAB-ZFP proteins to their target loci. The effects of having increased KRAB-ZFP and KAP1 repressive complexes binding would possibly result in increased retrotransposon repression, the converse of *Tex19.1*$^{-/-}$ tissues (Jacobs et al., 2014; Castro-Diaz et al., 2014; Ecco et al., 2016), however if TEX19.1 dependent ubiquitination of KRAB-ZFPs is actually occurring remains to be investigated. Furthermore, no MAGE proteins have yet been associated with retrotransposon de-repression (Feng et al., 2011).

Together, an arsenal of defence mechanisms from DNA methylation, transcription factor availability, tissue specific KRAB-ZFPs and KAP1 dependent histone modifications are co-ordinately regulating retrotransposon expression to varying extents in different contexts (Goodier 2016).

6.2 Relationship between retrotransposon de-repression and developmental defects

In a somatic context, any retrotransposition events are not heritable and are unlikely to be occurring in the same progenitor cells, at the same loci in each mouse, causing the same phenotype. Expanding evidence exists, distinct from classical *de novo* insertion events, for the role of LTR and non-LTR retrotransposon expression itself. The de-repression of some retrotransposon loci can control expression of neighbouring genes, both in normal development and in genetic knockouts experiments (Elbarbary et al., 2016; Göke and Ng, 2016; Thompson et al., 2016). Retrotransposons regulate the genome in a variety of ways, from behaving as alternative promoters producing long

non coding RNAs, to acting as enhancers for nearby genes (Ecco et al., 2016; Herquel et al., 2013; Rowe et al., 2013). Therefore, the number of retrotransposon subfamilies de-repressed and the amount of gene expression changes in *Tex19.1*$^{-/-}$ placentas led me to investigate if the two could be correlated. Generally, I observed an association of full length elements of the de-repressed L1MdF2 subfamily with downregulated genes, however if this association is causative remains unclear (figure 4.11). Generally, LINE-1 dependent genic repression can occur via different mechanisms such as antisense interference which alters chromatin state of nearby regions (Cruickshanks et al., 2013). Antisense transcripts of full length LINE-1 flanking downregulated genes were detected at similar levels in control and *Tex19.1*$^{-/-}$ placentas (figure 4.12). Therefore this is unlikely causative of neighbouring gene expression changes, at least at the loci assessed. The fact that little association was detected for the LTR subfamilies analysed, further suggests that this is possibly not a major cause of the RNA abundance changes, however de-repression of a retrotransposon which is altering the expression of a single gene critical for placental development could be causative of the downstream *Tex19.1*$^{-/-}$ placental phenotype.

Other genome defence mutant mice, along with *Tex19.1*$^{-/-}$ mice, have de-repression of numerous retrotransposons in the germline along with defective spermatogenesis (Crichton et al., 2013). In this thesis I have expanded the repertoire of genes involved in germline genome defence to include *Ubr2* (figure 5.11 & 5.12). In each of these knockout mice it is unclear if retrotransposon de-repression is the cause of developmental defects in the germline (Crichton et al., 2013). No developmental defect has yet been directly associated with perturbed expression of a gene with a known function in the defective pathway due to nearby retrotransposon de-repression, but examples will inevitably emerge with the use of high throughput sequencing experiments.

6.3 The innate immune response

Far from any intricate effects on gene expression, recent evidence from a multitude of different cell types and developmental contexts has identified a substantial role for retrotransposon de-repression in the activation of the immune system (Kassiotis and Stoye, 2016). For example, germ cells of mice lacking the methylation sensitive genome defence gene *Mov10l* have increased LINE-1 expression, alongside an interferon β response (Yu et al., 2015). Yu et al demonstrated in an *in vitro* experiment, transfecting MEFs with a LINE-1 expressing plasmid, that endonuclease dependent LINE-1 retrotransposition stimulates interferon β expression in somatic cell lines (Yu et al., 2015). Furthermore, interferon β and associated interferon stimulated genes restrict LINE-1 propagation, implying a role for the interferon pathway in genome defence and stability (Yu et al., 2015). Interestingly, I have detected a response to interferon β in the placentas of *Tex19.1*$^{-/-}$ animals (figure 4.14, 4.15 & 4.16) and it is tempting to hypothesise that this is due to increases in retrotransposon de-repression (see chapter 1, section 1.5 for details).

Despite this alternate possibilities the innate immune response appears to provide an additional mechanism of defence against increased retrotransposons in developmental contexts or in the absence of genome defence genes. For example, mice null for the transcriptional repressor *Trim24* exhibit increases in VL30 and MMERVK10C cDNA in the cytoplasm of null hepatocytes, the same retrotransposon subfamilies de-repressed in *Tex19.1*$^{-/-}$ contexts (Herquel et al., 2013; Ollinger et al., 2008; Reichmann et al., 2013). This correlates with the activation of the viral defence IFNβ response and an increase in apoptosis (Herquel et al., 2013).

Furthermore, human somatic tumour cell lines treated with the DNA demethylating drug 5-Aza induce a potent IFNβ response due to recognition of retrotransposon dsRNA and cytosolic DNA by the MDA5-MAVs proteins and TMEM173 respectively (Chiappinelli et al., 2015; Roulois et al., 2015). *Tmem173* was significantly upregulated in *Tex19.1* suggesting possible activation of this cytosolic DNA sensing pathway. Crossing the *Tex19.1*$^{+/-}$ and *Tmem173*$^{-/-}$ mice would determine if the interferon response was operating via TMEM173 and presumably sensing

retrotransposon DNA in the cytoplasm of trophoblast derived cells. If this abrogated the immune response, placental defects and IUGR this would suggest that cytosolic DNA sensing is leading to the immune response which may be causative of developmental defects in *Tex19.1*$^{-/-}$ animals. Additionally, determining if this same response is occurring in *Tex19.1*$^{-/-}$ and *Ubr2*$^{-/-}$ testes in response to MMERVK10C de-repression would lend further support to the hypothesis that de-repression of retrotransposons is causing an interferon response leading to apoptosis and developmental defects in multiple cellular contexts. The investigation of interferon autoimmune activation in a number of human diseases also provides evidence of significant associations with retrotransposon de-repression (Kassiotis and Stoye, 2016). One such disease is Aicardi-Goutières syndrome which is discussed in section 1.5.

It is important to note that mouse ESCs are thought to have limited IFNα and IFNβ activity (Cross et al., 1990; Wang et al., 2013), however somatic cells have a robust interferon response (Wang et al., 2013). The general roles for interferons in mouse and human somatic contexts is well characterised, however the role of interferons in placental development is still to be fully elucidated (Platinas 2005). The experimental evidence for a role of interferons in mouse placentation is very sparse, especially for IFNβ. Most evidence of IFN in normal placental development supports roles for IFNγ which, in the post-implantation placenta, is produced by NK cells of the uterine decidua and is important for spiral artery dilation (Ashkar et al., 2000). TGCs can secrete hormones to prevent natural killer (NK) cells producing IFNγ to regulate this process (Hu and Cross, 2010; Muller et al., 1999). Both rodent and human trophoblast cells show dampened responses to INFγ. IFNα is released by trophoblast giant cells in the mouse placenta which, as with IFNγ, stimulates the expression of interferon stimulated genes in the decidua (Bany and Cross, 2006; Hu and Cross, 2010; Platt and Hunt, 1998).

As regards placental pathogenesis and the interferon response, in humans interferon β is used as a therapy for multiple sclerosis (Patti et al., 2008). Numerous studies have assessed the effects on fetal outcome from pregnant women taking interferon β in the first trimesters of pregnancy. In four separate studies, exposure to formulations of interferon β was associated with low birth weight, shorter birth length or shorter

gestation periods, but no increased risk of pregnancy loss or health problems later in life (Amato et al., 2010; Boskovic, 2015; Hellwig et al., 2012; Patti et al., 2008; Weber-Schoendorfer and Schaefer, 2009). Together, these data suggest that increased interferon β during pregnancy has negative effects on gestation, however the reason for this remains unclear, and effects on placenta are uncharacterised.

One interesting line of research is starting to provide insight into the defences against viral invasion of the placenta. The number of human diseases with maternal-fetal transmission has stimulated intense research into the mechanisms of defence and viral replication occurring in trophoblast cells. For example, Zika virus infection in pregnant women causes IUGR, spontaneous abortion and microcephaly of the offspring (Miner et al., 2016). Viral transmission from the mother to the offspring occurs by the virus crossing the maternal-fetal blood barrier of the syncytiotrophoblasts in infected mice (Miner et al., 2016). In humans an innate immune response is activated upon infection of the placenta with Zika virus, with increased IFNλ (Bayer et al., 2016). This IFNλ is sufficient to protect primary human syncytiotrophoblasts from infection in an autocrine and paracrine manner, providing an efficient immune response to exogenous viral infection (Bayer et al., 2016). Whether an immune response occurs due to de-repression of retrotransposons, leading to IUGR in humans is unclear and few studies have investigated the consequences of aberrant retrotransposon de-repression in the human placenta. This highlights the value of the *Tex19.1*$^{-/-}$ mouse when considering the possible pathogenic consequences of loss of genome defence mechanisms in human pregnancies.

6.4 Conclusion

Overall, I have identified that *Tex19.1* is the only methylation sensitive genome defence gene activated in response to reduced levels of DNA methylation in all hypomethylated tissues in the mouse. This activation acts to augment repressive histone modifications at specific retrotransposon loci to compensate for reduced DNA methylation, thereby limiting retrotransposon de-repression and possibly a retrotransposon sensing innate immune response which may be detrimental to

development. I have identified a role for the interacting partner of TEX19.1, UBR2, in defending the genome from activation of MMERVK10C in the mouse germline and have determined that UBR5 is likely not making a large contribution to the *Tex19.1*$^{-/-}$ phenotype. Despite this, I establish UBR5 as a novel factor for the maintenance of spermatogonia and progression of spermatogenesis in the mouse germline.

Taken together, this thesis highlights the inherent importance of defending the genome from retrotransposons and provides insights into the potential far reaching physiological consequences when these mechanisms fail. Indeed, we cannot live with them, nor without them and it is becoming increasingly apparent that striking the balance between the benefits of genome innovation provided by these ancient viruses, alongside the possible severe consequences when things go wrong, are of critical importance.

Bibliography

Aasland, R., and Stewart, a F. (1995). The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. Nucleic Acids Res. *23*, 3168–3173.

Adams, I.R., and McLaren, A. (2002). Sexually dimorphic development of mouse primordial germ cells: switching from oogenesis to spermatogenesis. Development *129*, 1155–1164.

Ain, R., Canham, L.N., and Soares, M.J. (2003). Gestation stage-dependent intrauterine trophoblast cell invasion in the rat and mouse: Novel endocrine phenotype and regulation. Dev. Biol. *260*, 176–190.

Amato, M.P., Portaccio, E., Ghezzi, A., Hakiki, B., Zipoli, V., Martinelli, V., Moiola, L., Patti, F., La Mantia, L., Mancardi, G.L., et al. (2010). Pregnancy and fetal outcomes after interferon-beta exposure in multiple sclerosis. Neurology *75*, 1794–1802.

An, J.Y., Kim, E., Zakrzewska, A., Yoo, Y.D., Jang, J.M., Han, D.H., Lee, M.J., Seo, J.W., Lee, Y.J., Kim, T.-Y., et al. (2012). UBR2 of the N-end rule pathway is required for chromosome stability via histone ubiquitylation in spermatocytes and somatic cells. PLoS One *7*, e37414.

Anson-Cartwright, L., Dawson, K., Holmyard, D., Fisher, S.J., Lazzarini, R.A., and Cross, J.C. (2000). The glial cells missing-1 protein is essential for branching morphogenesis in the chorioallantoic placenta. Nat. Genet. *25*, 311–314.

Apostolidou, S., Abu-Amero, S., O'Donoghue, K., Frost, J., Olafsdottir, O., Chavele, K.M., Whittaker, J.C., Loughna, P., Stanier, P., and Moore, G.E. (2007). Elevated placental expression of the imprinted PHLDA2 gene is associated with low birth weight. J. Mol. Med. *85*, 379–387.

Arand, J., Wossidlo, M., Lepikhov, K., Peat, J.R., Reik, W., and Walter, J. (2015). Selective impairment of methylation maintenance is the major cause of DNA methylation reprogramming in the early embryo. Epigenetics Chromatin *8*, 1.

Ashkar, A.A., Di Santo, J.P., and Croy, B.A. (2000). Interferon gamma contributes to initiation of uterine vascular modification, decidual integrity, and uterine natural killer cell maturation during normal murine pregnancy. J. Exp. Med. *192*, 259–270.

Bannert, N., and Kurth, R. (2006). The evolutionary dynamics of human endogenous retroviral families. Annu. Rev. Genomics Hum. Genet. *7*, 149–173.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Bany, B.M., and Cross, J.C. (2006). Post-implantation mouse conceptuses produce paracrine signals that regulate the uterine endometrium undergoing decidualization. Dev. Biol. *294*, 445–456.

Baran, N. (2002). Decysin, a New Member of the Metalloproteinase Family, Is Regulated by Prolactin and Steroids During Mouse Pregnancy. Biol. Reprod. *68*, 1787–1792.

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

Bayer, A., Lennemann, N.J., Ouyang, Y., Bramley, J.C., Morosky, S., Marques, E.T.D.A.,

Cherry, S., Sadovsky, Y., and Coyne, C.B. (2016). Type III Interferons Produced by Human Placental Trophoblasts Confer Protection against Zika Virus Infection. Cell Host Microbe *19*, 1–8.

Beck, C.R., Collier, P., Macfarlane, C., Malig, M., Kidd, J.M., Eichler, E.E., Badge, R.M., and Moran, J. V. (2010). LINE-1 retrotransposition activity in human genomes. Cell *141*, 1159–1170.

Behrendt, R., and Roers, A. (2014). Mouse models for Aicardi-Gouti??res syndrome provide clues to the molecular pathogenesis of systemic autoimmunity. Clin. Exp. Immunol. *175*, 9–16.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482–485.

Bersinger, N.A., and Ødegård, R.A. (2004). Second- and third-trimester serum levels of placental proteins in preeclampsia and small-for-gestational age pregnancies. Acta Obstet. Gynecol. Scand. *83*, 37–45.

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R. a. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes Dev. *23*, 2484–2489.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21.

Bitgood, M.J., Shen, L., and McMahon, a P. (1996). Sertoli cell signaling by Desert hedgehog regulates the male germline. Curr. Biol. *6*, 298–304.

Blaschke, K., Ebata, K.T., Karimi, M.M., Zepeda-Martínez, J. a, Goyal, P., Mahapatra, S., Tam, A., Laird, D.J., Hirst, M., Rao, A., et al. (2013). Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature *500*, 222–226.

Bode, S., Peters, C., and Deussing, J.M. (2005). Placental cathepsin M is alternatively spliced and exclusively expressed in the spongiotrophoblast layer. Biochim. Biophys. Acta - Gene Struct. Expr. *1731*, 160–167.

Bojkowska, K., Aloisio, F., Cassano, M., Kapopoulou, A., Santoni de Sio, F., Zangger, N., Offner, S., Cartoni, C., Thomas, C., Quenneville, S., et al. (2012). Liver-specific ablation of Kr??ppel-associated box-associated protein 1 in mice leads to male-predominant hepatosteatosis and development of liver adenoma. Hepatology *56*, 1279–1290.

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. Nat. Genet. *42*, 1093–1100.

Bortvin, A (2013). PIWI-interacting RNAs (piRNAs) - a mouse testis perspective. Biochem. Biokhimiia *78*, 592–602.

Boskovic (2015). The reproductive effects of beta interferon therapy in pregnancy - A longitudinal cohort The reproductive effects of beta.

Boyes, J., and Bird, A. (1991). DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell *64*, 1123–1134.

Branco, M.R., King, M., Perez-garcia, V., Dean, W., Hemberger, M., Reik, W., Development, E.T., Branco, M.R., King, M., Perez-garcia, V., et al. (2016). Maternal DNA Methylation Regulates Early Trophoblast Development Article Maternal DNA Methylation Regulates. Dev. Cell *36*, 152–163.

Briscoe1, J., Thérond, P.P., Briscoe, J., and Thérond, P.P. (2013). The mechanisms of Hedgehog signalling and its roles in development and disease. Nat. Rev. Mol. Cell Biol. *14*, 416–429.

Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J. V, and Kazazian, H.H. (2003). Hot L1s account for the bulk of retrotransposition in the human population. Proc. Natl. Acad. Sci. U. S. A. *100*, 5280–5285.

Brunmeir, R., Lagger, S., Simboeck, E., Sawicka, A., Egger, G., Hagelkruys, A., Zhang, Y., Matthias, P., Miller, W.J., and Seiser, C. (2010). Epigenetic regulation of a murine retrotransposon by a dual histone modification mark. PLoS Genet. *6*, e1000927.

Bulut-Karslioglu, A., DeLaRosa-Velázquez, I. a., Ramirez, F., Barenboim, M., Onishi-Seebacher, M., Arand, J., Galán, C., Winter, G.E., Engist, B., Gerle, B., et al. (2014). Suv39h-Dependent H3K9me3 Marks Intact Retrotransposons and Silences LINE Elements in Mouse Embryonic Stem Cells. Mol. Cell *55*, 277–290.

Callaghan, M.J., Russell, A.J., Woollatt, E., Sutherland, G.R., Sutherland, R.L., and Watts, C.K. (1998). Identification of a human HECT family protein with homology to the Drosophila tumor suppressor gene hyperplastic discs. Oncogene.

Cammas, F., Mark, M., Dollé, P., Dierich, a, Chambon, P., and Losson, R. (2000). Mice lacking the transcriptional corepressor TIF1beta are defective in early postimplantation development. Development *127*, 2955–2963.

Campisi, J., and d'Adda di Fagagna, F. (2007). Cellular senescence: when bad things happen to good cells. Nat. Rev. Mol. Cell Biol. *8*, 729–740.

Cantone, I., and Fisher, A.G. (2013). Epigenetic programming and reprogramming during development. Nat. Struct. Mol. Biol. *20*, 282–289.

Cardoso, M.C., and Leonhardt, H. (1999). DNA methyltransferase is actively retained in the cytoplasm during early development. J. Cell Biol. *147*, 25–32.

Carmell, M. a, Girard, A., van de Kant, H.J.G., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev. Cell *12*, 503–514.

Castro-Diaz, N., Ecco, G., Coluccio, A., Kapopoulou, A., Yazdanpanah, B., Friedli, M., Duc, J., Jang, S.M., Turelli, P., and Trono, D. (2014). Evolutionally dynamic L1 regulation in embryonic stem cells. Genes Dev. *28*, 1397–1409.

Cedar, H., and Bergman, Y. (2012). Programming of DNA methylation patterns. Annu Rev Biochem *81*, 97–117.

Cetin, I., and Alvino, G. (2009). Intrauterine Growth Restriction: Implications for Placental Metabolism and Transport. A Review. Placenta *30*, 77–82.

Cheng, C.-T. (2014). KAPtain in charge of multiple missions: Emerging roles of KAP1. World J. Biol. Chem. *5*, 308.

Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A., et al. (2015). Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell *162*, 974–986.

Cho, S., Park, J.S., and Kang, Y.K. (2011). Dual functions of histone-lysine N-methyltransferase Setdb1 protein at promyelocytic leukemia-nuclear body (PML-NB):

Maintaining PML-NB structure and regulating the expression of its associated genes. J. Biol. Chem. *286*, 41115–41124.

Christie, G.A. (1967). Comparative histochemical distribution of glycogen and alkaline phosphatases in the placenta. Histochemie *9*, 93–116.

Chuong, E.B., Rumi, M. a K., Soares, M.J., and Baker, J.C. (2013). Endogenous retroviruses function as species-specific enhancer elements in the placenta. Nat. Genet. *45*, 325–329.

Chuong, E.B., Elde, N.C., and Feschotte, C. (2016). Regulatory evolution of innate immunity through co-option of endogenous retroviruses. Science *in press*, 1–28.

Chuprin, A., Gal, H., Biron-Shental, T., Biran, A., Amiel, A., Rozenblatt, S., and Krizhanovsky, V. (2013). Cell fusion induced by ERVWE1 or measles virus causes cellular senescence. Genes Dev. *27*, 2356–2366.

Clermont, Y. (1963). The cycle of the seminiferous epithelium in man. Am. J. Anat. *112*, 35–51.

Coan, P.M., Ferguson-Smith, A.C., and Burton, G.J. (2004). Developmental dynamics of the definitive mouse placenta assessed by stereology. Biol. Reprod. *70*, 1806–1813.

Coan, P.M., Conroy, N., Burton, G.J., and Ferguson-Smith, a C. (2006). Origin and characteristics of glycogen cells in the developing murine placenta. Dev. Dyn. *235*, 3280–3294.

Cockburn, K., and Rossant, J. (2010). Making the blastocyst: Lessons from the mouse. J. Clin. Invest. *120*, 995–1003.

Collins, P.L., Kyle, K.E., Egawa, T., Shinkai, Y., and Oltz, E.M. (2015). The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. Proc. Natl. Acad. Sci. *112*, 8367–8372.

Constância, M., Hemberger, M., Hughes, J., Dean, W., Ferguson-Smith, A., Fundele, R., Stewart, F., Kelsey, G., Fowden, A., Sibley, C., et al. (2002). Placental-specific IGF-II is a major modulator of placental and fetal growth. Nature *417*, 945–948.

Copp, A.J. (1978). Interaction between inner cell mass and trophectoderm of the mouse blastocyst. I. A study of cellular proliferation. J. Embryol. Exp. Morphol. *48*, 109–125.

Cordaux, R., and Batzer, M. a (2009). The impact of retrotransposons on human genome evolution. Nat. Rev. Genet. *10*, 691–703.

Cortázar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A.L., Siegrist, F., et al. (2011). Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. Nature *470*, 419–423.

Costoya, J. a, Hobbs, R.M., Barna, M., Cattoretti, G., Manova, K., Sukhwani, M., Orwig, K.E., Wolgemuth, D.J., and Pandolfi, P.P. (2004). Essential role of Plzf in maintenance of spermatogonial stem cells. Nat. Genet. *36*, 653–659.

Crichton, J.H., Dunican, D.S., Maclennan, M., Meehan, R.R., and Adams, I.R. (2013). Defending the genome from the enemy within: mechanisms of retrotransposon suppression in the mouse germline. Cell. Mol. Life Sci.

Crichton, J.H., Playfoot, C.J., and Adams, I.R. (2014). The Role of Chromatin Modifications in Progression Through Mouse Meiotic Prophase. J. Genet. Genomics.

Cross, J.C., Farin, C.E., Sharif, S.F., and Roberts, R.M. (1990). Characterization of the

Antiviral Activity Constitutively Produced by Murine Conceptuses : Absence of Placental mRNAs for Interferon Alpha and Beta. *128*.

Cross, J.C., Werb, Z., and Fisher, S.J. (1994). Implantation and the placenta: key pieces of the development puzzle. Science *266*, 1508–1518.

Crow, Y.J., and Rehwinkel, J. (2009). Aicardi-Goutie'res syndrome and related phenotypes: Linking nucleic acid metabolism with autoimmunity. Hum. Mol. Genet. *18*, 130–136.

Cruickshanks, H. a, Vafadar-Isfahani, N., Dunican, D.S., Lee, A., Sproul, D., Lund, J.N., Meehan, R.R., and Tufarelli, C. (2013). Expression of a large LINE-1-driven antisense RNA is linked to epigenetic silencing of the metastasis suppressor gene TFPI-2 in cancer. Nucleic Acids Res. *41*, 6857–6869.

Davis, C.M., Constantinides, P.G., van der Riet, F., van Schalkwyk, L., Gevers, W., and Parker, M.I. (1989). Activation and demethylation of the intracisternal A particle genes by 5-azacytidine. Cell Differ. Dev. *27*, 83–93.

Dawlaty, M.M., Breiling, A., Le, T., Raddatz, G., Barrasa, M.I., Cheng, A.W., Gao, Q., Powell, B.E., Li, Z., Xu, M., et al. (2013). Combined Deficiency of Tet1 and Tet2 Causes Epigenetic Abnormalities but Is Compatible with Postnatal Development. Dev. Cell *24*, 310–323.

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

Denli, A.M., Narvaiza, I., Kerman, B.E., Pena, M., Benner, C., Marchetto, M.C.N., Diedrich, J.K., Aslanian, A., Ma, J., Moresco, J.J., et al. (2015). Primate-specific ORF0 contributes to retrotransposon-mediated diversity. Cell *163*, 583–593.

Dewannieux, M., Esnault, C., and Heidmann, T. (2003). LINE-mediated retrotransposition of marked Alu sequences. Nat. Genet. *35*, 41–48.

Dewannieux, M., Dupressoir, A., Harper, F., Pierron, G., and Heidmann, T. (2004). Identification of autonomous IAP LTR retrotransposons mobile in mammalian cells. Nat. Genet. *36*, 534–539.

Dewannieux, M., Heidmann, T., and Yaniv, M. (2005). L1-mediated retrotransposition of murine B1 and B2 SINEs recapitulated in cultured cells. J. Mol. Biol. *349*, 241–247.

Dodge, J.E., Kang, Y., Beppu, H., and Lei, H. (2004). Histone H3-K9 Methyltransferase ESET Is Essential for Early Development Histone H3-K9 Methyltransferase ESET Is Essential for Early Development. Mol. Cell. Biol. *24*, 2478–2486.

Doyle, J.M., Gao, J., Wang, J., Yang, M., and Potts, P.R. (2010). MAGE-RING protein complexes comprise a family of E3 ubiquitin ligases. Mol. Cell *39*, 963–974.

Ecco, G., Cassano, M., Kauzlaric, A., Duc, J., Coluccio, A., Offner, S., Imbeault, M., Rowe, H.M., Turelli, P., and Trono, D. (2016). Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. Dev. Cell *36*, 611–623.

Elbarbary, R.A., Lucas, B.A., and Maquat, L.E. (2016). Retrotransposons as regulators of gene expression. Science *351*, aac7247.

Emera, D., Casola, C., Lynch, V.J., Wildman, D.E., Agnew, D., and Wagner, G.P. (2012). Convergent evolution of endometrial prolactin expression in primates, mice, and elephants through the independent recruitment of transposable elements. Mol. Biol. Evol. *29*, 239–247.

Emerson, R.O., and Thomas, J.H. (2009). Adaptive evolution in zinc finger transcription factors. PLoS Genet. *5*.

Eskeland, R., Leeb, M., Grimes, G.R., Kress, C., Boyle, S., Sproul, D., Gilbert, N., Fan, Y., Skoultchi, A.I., Wutz, A., et al. (2010). Ring1B Compacts Chromatin Structure and Represses Gene Expression Independent of Histone Ubiquitination. Mol. Cell *38*, 452–464.

Fasching, L., Kapopoulou, A., Sachdeva, R., Petri, R., Jönsson, M.E., Männe, C., Turelli, P., Jern, P., Cammas, F., Trono, D., et al. (2015). TRIM28 Represses Transcription of Endogenous Retroviruses in Neural Progenitor Cells. Cell Rep. *10*, 20–28.

De Fazio, S., Bartonicek, N., Di Giacomo, M., Abreu-Goodger, C., Sankar, A., Funaya, C., Antony, C., Moreira, P.N., Enright, A.J., and O'Carroll, D. (2011). The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature *480*, 259–263.

Feng, Y., Gao, J., and Yang, M. (2011). When MAGE meets RING: Insights into biological functions of MAGE proteins. Protein Cell *2*, 7–12.

Ficz, G., Hore, T. a, Santos, F., Lee, H.J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y.-L., Walter, J., et al. (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. Cell Stem Cell *13*, 351–359.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G., and Rauscher, F.J. (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev. *10*, 2067–2078.

Gall, A., Treuting, P., Elkon, K.B., Loo, Y.M., Gale, M., Barber, G.N., and Stetson, D.B. (2012). Autoimmunity Initiates in Nonhematopoietic Cells and Progresses via Lymphocytes in an Interferon-Dependent Autoimmune Disease. Immunity *36*, 120–131.

Gan, H., Wen, L., Liao, S., Lin, X., Ma, T., Liu, J., Song, C.-X., Wang, M., He, C., Han, C., et al. (2013). Dynamics of 5-hydroxymethylcytosine during mouse spermatogenesis. Nat. Commun. *4*, 1995.

Gartel, A.L., and Tyner, A.L. (2002). The role of the cyclin-dependent kinase inhibitor p21 in apoptosis. Mol Cancer Ther *1*, 639–649.

Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). Affy - Analysis of Affymetrix GeneChip data at the probe level. Bioinformatics *20*, 307–315.

Georgiades, P., Ferguson-smith, a C., and Burton, G.J. (2002). CURRENT TOPIC Comparative Developmental Anatomy of the Murine and Human De nitive Placentae. Placenta 3–19.

Globisch, D., M??nzel, M., M??ller, M., Michalakis, S., Wagner, M., Koch, S., Br??ckl, T., Biel, M., and Carell, T. (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. PLoS One *5*, 1–9.

Göke, J., and Ng, H.H. (2016). CTRL+INSERT: retrotransposons and their contribution to regulation and innovation of the transcriptome. EMBO Rep. *17*, e201642743.

Goodier, J.L. (2016). Restricting retrotransposons: a review. Mob. DNA *7*, 16.

Griswold, M.D. (1998). The central role of Sertoli cells in spermatogenesis. Semin. Cell Dev. Biol. *9*, 411–416.

Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-F., Liu, W., Xie, Z.-G., Shi, L., He, X., Jin,

S., et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *477*, 606–610.

Gudjonsson, T., Altmeyer, M., Savic, V., Toledo, L., Dinant, C., Grøfte, M., Bartkova, J., Poulsen, M., Oka, Y., Bekker-Jensen, S., et al. (2012). TRIP12 and UBR5 suppress spreading of chromatin ubiquitylation at damaged chromosomes. Cell *150*, 697–709.

Guillemot, F., Nagy, a, Auerbach, a, Rossant, J., and Joyner, a L. (1994). Essential role of Mash-2 in extraembryonic development. Nature *371*, 333–336.

Guo, F., Li, X., Liang, D., Li, T., Zhu, P., Guo, H., Wu, X., Wen, L., Gu, T.P., Hu, B., et al. (2014). Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote. Cell Stem Cell *15*, 447–458.

Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H.D., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C., et al. (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell *13*, 360–369.

Hackett, J. a, Reddington, J.P., Nestor, C.E., Dunican, D.S., Branco, M.R., Reichmann, J., Reik, W., Surani, M.A., Adams, I.R., and Meehan, R.R. (2012). Promoter DNA methylation couples genome-defence mechanisms to epigenetic reprogramming in the mouse germline. Development *139*, 3623–3632.

Hackett, J. a, Dietmann, S., Murakami, K., Down, T. a, Leitch, H.G., and Surani, M.A. (2013). Synergistic Mechanisms of DNA Demethylation during Transition to Ground-State Pluripotency. Stem Cell Reports *1*, 518–531.

Haig, D. (2000). T k t g i. 9–32.

Haig, D., and Graham, C. (1991). Genomic Imprinting and the Strange Case of the Insulin-like Growth Factor II Receptor. Cell *64*, 1045–1046.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. Mech. Dev. *117*, 15–23.

Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U.C., Cesari, F., Lee, C., Almouzni, G., Schneider, R., and Surani, M.A. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature *452*, 877–881.

Hancks, D.C., and Kazazian, H.H. (2016). Roles for retrotransposon insertions in human disease. Mob. DNA *7*, 9.

Hasegawa, K., and Saga, Y. (2012). Retinoic acid signaling in Sertoli cells regulates organization of the blood-testis barrier through cyclical changes in gene expression. Development *4355*, 4347–4355.

Havecker, E.R., Gao, X., and Voytas, D.F. (2004). The diversity of LTR retrotransposons. Genome Biol. *5*, 225.

Hellwig, K., Haghikia, A., Rockhoff, M., and Gold, R. (2012). Multiple sclerosis and pregnancy: experience from a nationwide database in Germany. Ther. Adv. Neurol. Disord. *5*, 247–253.

Herquel, B., Ouararhni, K., Martianov, I., Le Gras, S., Ye, T., Keime, C., Lerouge, T., Jost, B., Cammas, F., Losson, R., et al. (2013). Trim24-repressed VL30 retrotransposons regulate gene expression by producing noncoding RNA. Nat. Struct. Mol. Biol. *20*, 339–346.

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

Hitz, C., Vogt-Weisenhorn, D., Ruiz, P., Wurst, W., and Floss, T. (2005). Progressive loss of the spongiotrophoblast layer of Birc6/Bruce mutants results in embryonic lethality. Genesis *42*, 91–103.

Hu, D., and Cross, J.C. (2010). Development and function of trophoblast giant cells in the rodent placenta. Int. J. Dev. Biol. *54*, 341–354.

Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., Gong, W., et al. (2015). Structural insight into substrate preference for TET-mediated oxidation. Nature *527*, 118–122.

Hu, X., Zhang, L., Mao, S.Q., Li, Z., Chen, J., Zhang, R.R., Wu, H.P., Gao, J., Guo, F., Liu, W., et al. (2014). Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. Cell Stem Cell *14*, 512–522.

Hutnick, L.K., Huang, X., Loo, T.-C., Ma, Z., and Fan, G. (2010). Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation-independent mechanism. J. Biol. Chem. *285*, 21082–21091.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., Zhang, Y., Goll, M.G., Bestor, T.H., et al. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300–1303.

Ivanov, A. V., Peng, H., Yurchenko, V., Yap, K.L., Negorev, D.G., Schultz, D.C., Psulkowski, E., Fredericks, W.J., White, D.E., Maul, G.G., et al. (2007). PHD Domain-Mediated E3 Ligase Activity Directs Intramolecular Sumoylation of an Adjacent Bromodomain Required for Gene Silencing. Mol. Cell *28*, 823–837.

Iyengar, S., and Farnham, P.J. (2011). KAP1 protein: an enigmatic master regulator of the genome. J. Biol. Chem. *286*, 26267–26276.

Iyengar, S., Ivanov, A. V, Jin, V.X., Rauscher, F.J., and Farnham, P.J. (2011). Functional analysis of KAP1 genomic recruitment. Mol. Cell. Biol. *31*, 1833–1847.

Jacobs, F.M.J., Greenberg, D., Nguyen, N., Haeussler, M., Ewing, A.D., Katzman, S., Paten, B., Salama, S.R., and Haussler, D. (2014). An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. Nature *516*, 242–245.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. *33 Suppl*, 245–254.

John, R.M. (2013). Epigenetic regulation of placental endocrine lineages and complications of pregnancy. Biochem. Soc. Trans. *41*, 701–709.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900–903.

Kang, Y.-K. (2014). SETDB1 in Early Embryos and Embryonic Stem Cells. Curr. Issues Mol. Biol. *17*, 1–10.

Karimi, M.M., Goyal, P., Maksakova, I. a, Bilenky, M., Leung, D., Tang, J.X., Shinkai, Y., Mager, D.L., Jones, S., Hirst, M., et al. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in

mESCs. Cell Stem Cell *8*, 676–687.

Kassiotis, G., and Stoye, J.P. (2016). Immune responses to endogenous retroelements: taking the bad with the good. Nat. Rev. Immunol. *16*, 207–219.

Kato, A., Nagata, Y., and Todokoro, K. (2004). ??-Tubulin Is a Component of Intercellular Bridges and Both the Early and Mature Perinuclear Rings During Spermatogenesis. Dev. Biol. *269*, 196–205.

Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y., Okano, M., Li, E., Nozaki, M., and 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*, 2272–2280.

Khan, H., Smit, A., and Boissinot, S. (2006). Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. Genome Res. *16*, 78–87.

Kim, U., Garner, T.L., Sanford, T., Speicher, D., Murray, J.M., and Nishikura, K. (1994). Purification and characterization of double-stranded RNA adenosine deaminase from bovine nuclear extracts. J. Biol. Chem. *269*, 13480–13489.

Kinsella, E., Dora, N., Mellis, D., Lettice, L., Deveney, P., Hill, R., and Ditzel, M. (2016). Use of a Conditional Ubr5 Mutant Allele to Investigate the Role of an N-End Rule Ubiquitin-Protein Ligase in Hedgehog Signalling and Embryonic Limb Development. PLoS One *11*, e0157079.

Knox, K., and Baker, J.C. (2008). Genomic evolution of the placenta using co-option and duplication and divergence. Genome Res. *18*, 695–705.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell *128*, 693–705.

Kozlov, G., Nguyen, L., Lin, T., De Crescenzo, G., Park, M., and Gehring, K. (2007). Structural basis of ubiquitin recognition by the ubiquitin-associated (UBA) domain of the ubiquitin ligase EDD. J. Biol. Chem. *282*, 35787–35795.

Kramerov, D.A., and Vassetzky, N.S. (2011). SINEs. Wiley Interdiscip. Rev. RNA *2*, 772–786.

Kuntz, S., Kieffer, E., Bianchetti, L., Lamoureux, N., Fuhrmann, G., and Viville, S. (2008). Tex19, a mammalian-specific protein with a restricted expression in pluripotent stem cells and germ line. Stem Cells *26*, 734–744.

Kwon, Y.T., Xia, Z., An, J.Y., Davydov, I. V, Seo, J.W., Sheng, J., Varshavsky, A., Tasaki, T., and Xie, Y. (2003). Female Lethality and Apoptosis of Spermatocytes in Mice Lacking the UBR2 Ubiquitin Ligase of the N-End Rule Pathway Female Lethality and Apoptosis of Spermatocytes in Mice Lacking the UBR2 Ubiquitin Ligase of the N-End Rule Pathway.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

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

Lee, J.D., Amanai, K., Shearn, A., and Treisman, J.E. (2002). The ubiquitin ligase Hyperplastic discs negatively regulates hedgehog and decapentaplegic expression by

independent mechanisms. Development *129*, 5697–5706.

Lei, H., Oh, S.P., Okano, M., Jüttermann, R., Goss, K. a, Jaenisch, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195–3205.

Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A., et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. Nat. Struct. Mol. Biol. *20*, 311–316.

Leung, D.C., and Lorincz, M.C. (2012). Silencing of endogenous retroviruses: when and why do histone marks predominate? Trends Biochem. Sci. *37*, 127–133.

Levin, H.L., and Moran, J. V (2011). Dynamic interactions between transposable elements and their hosts. Nat. Rev. Genet. *12*, 615–627.

Lewis, R.M., Cleal, J.K., Ntani, G., Crozier, S.R., Mahon, P.A., Robinson, S.M., Harvey, N.C., Cooper, C., Inskip, H.M., Godfrey, K.M., et al. (2012). Relationship between placental expression of the imprinted PHLDA2 gene, intrauterine skeletal growth and childhood bone mass. Bone *50*, 337–342.

Li, Y., and Behringer, R.R. (1998). Esx1 is an X-chromosome-imprinted regulator of placental development and fetal growth. Nat. Genet. *20*, 309–311.

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

Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362–365.

Li, J., Kannan, M., Trivett, A.L., Liao, H., Wu, X., Akagi, K., and Symer, D.E. (2014). An antisense promoter in mouse L1 retrotransposon open reading frame-1 initiates expression of diverse fusion transcripts and limits retrotransposition. Nucleic Acids Res. *42*, 4546–4562.

Li, L., Lu, X., and Dean, J. (2013). The maternal to zygotic transition in mammals. Mol. Aspects Med. *34*, 919–938.

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., and Ferguson-Smith, A.C. (2008). A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and Paternal Imprints. Dev. Cell *15*, 547–557.

Li, Y., Li, Y., Kao, G.D., Kao, G.D., Garcia, B. a, Garcia, B. a, Shabanowitz, J., Shabanowitz, J., Hunt, D.F., Hunt, D.F., et al. (2006). A novel histone deacetylase pathway regulates mitosis by modulating Aurora B kinase activity. Genes Dev. 2566–2579.

Lim, N.S., Kozlov, G., Chang, T.C., Groover, O., Siddiqui, N., Volpon, L., De Crescenzo, G., Shyu, A. Bin, and Gehring, K. (2006). Comparative peptide binding studies of the PABC domains from the ubiquitin-protein isopeptide ligase HYD and Poly(A)-binding Protein: Implications for HYD function. J. Biol. Chem. *281*, 14376–14382.

Lindahl, T., and Wood, R.D. (1999). Quality control by DNA repair. Science *286*, 1897–1905.

Liu, H., Chang, L.H., Sun, Y., Lu, X., and Stubbs, L. (2014a). Deep vertebrate roots for mammalian zinc finger transcription factor subfamilies. Genome Biol. Evol. *6*, 510–525.

Liu, S., Amour, J.B., Karimi, M.M., Shirane, K., Bogutz, A., Lefebvre, L., Sasaki, H., Shinkai, Y., and Lorincz, M.C. (2014b). Setdb1 is required for germline development and

silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. 2041–2055.

Ma, L., Buchold, G.M., Greenbaum, M.P., Roy, A., Burns, K.H., Zhu, H., Han, D.Y., Harris, R.A., Coarfa, C., Gunaratne, P.H., et al. (2009). GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline. PLoS Genet. *5*.

Maelfait, J., Bridgeman, A., Benlahrech, A., Cursi, C., and Rehwinkel, J. (2016). Restriction by SAMHD1 Limits cGAS/STING-Dependent Innate and Adaptive Immune Responses to HIV-1. Cell Rep. *16*, 1492–1501.

Makela, J. -a., Saario, V., Bourguiba-Hachemi, S., Nurmio, M., Jahnukainen, K., Parvinen, M., and Toppari, J. (2011). Hedgehog signalling promotes germ cell survival in the rat testis. Reproduction *142*, 711–721.

Maksakova, I. a, Romanish, M.T., Gagnier, L., Dunn, C. a, van de Lagemaat, L.N., and Mager, D.L. (2006). Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line. PLoS Genet. *2*, e2.

Manku, G., Wing, S.S., and Culty, M. (2012). Expression of the Ubiquitin Proteasome System in Neonatal Rat Gonocytes and Spermatogonia: Role in Gonocyte Differentiation. Biol. Reprod. *87*, 44–44.

Mannion, N.M., Greenwood, S.M., Young, R., Cox, S., Brindle, J., Read, D., Nellåker, C., Vesely, C., Ponting, C.P., McLaughlin, P.J., et al. (2014). The RNA-Editing Enzyme ADAR1 Controls Innate Immune Responses to RNA. Cell Rep. *9*, 1482–1494.

Mätlik, K., Redik, K., and Speek, M. (2006). L1 antisense promoter drives tissue-specific transcription of human genes. J. Biomed. Biotechnol. *2006*, 1–16.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I. a, Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M.C., and Shinkai, Y. (2010). Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature *464*, 927–931.

Matta-Camacho, E., Kozlov, G., Menade, M., and Gehring, K. (2012). Structure of the HECT C-lobe of the UBR5 E3 ubiquitin ligase. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *68*, 1158–1163.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. Nature *403*, 501–502.

McClintock, B. (1950). the Origin and Behavior of Mutable Loci in Maize. These Proc. Am. J. Bot. J. Genet. Nat. Nat. Proc. *36*, 344–355.

McClintock, B. (1956). Controlling Elements and the Gene. Cold Spring Harb. Symp. Quant. Biol. *21*, 197–216.

McMinn, J., Wei, M., Schupf, N., Cusmai, J., Johnson, E.B., Smith, A.C., Weksberg, R., Thaker, H.M., and Tycko, B. (2006). Unbalanced Placental Expression of Imprinted Genes in Human Intrauterine Growth Restriction. Placenta *27*, 540–549.

McNab, F., Mayer-Barber, K., Sher, A., Wack, A., and O'Garra, A. (2015). Type I interferons in infectious disease. TL - 15. Nat. Rev. Immunol. *15 VN-r*, 87–103.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766–770.

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet. *14*, 447–459.

Meyenn, F. Von, Iurlaro, M., Habibi, E., He, C., Reik, W., and Stunnenberg, H.G. (2016). Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells Article Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. 1–14.

Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., Lavallie, E., Tang, X., Edouard, P., Howes, S., et al. (2000). Syncytin is a captive retroviral envelope protein involved. *403*, 785–789.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

Miner, J.J., Cao, B., Govero, J., Smith, A.M., Fernandez, E., Cabrera, O.H., Garber, C., Noll, M., Klein, R.S., Noguchi, K.K., et al. (2016). Zika Virus Infection during Pregnancy in Mice Causes Placental Damage and Fetal Demise. Cell *165*, 1–11.

Miyawaki, a, Llopis, J., Heim, R., McCaffery, J.M., Adams, J. a, Ikura, M., and Tsien, R.Y. (1997). Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. Nature *388*, 882–887.

Moncrieff, S., Moncan, M., Scialpi, F., and Ditzel, M. (2015). Regulation of hedgehog ligand expression by the N-end rule ubiquitin-protein ligase hyperplastic discs and the drosophila GSK3β homologue, Shaggy. PLoS One *10*, 1–24.

Monk, D., and Moore, G.E. (2004). Intrauterine growth restriction--genetic causes and consequences. Semin. Fetal Neonatal Med. *9*, 371–378.

Morales, D.J., and Lenschow, D.J. (2013). The antiviral activities of ISG15. J. Mol. Biol. *425*, 4995–5008.

Moran, J., Holmes, S., and Naas, T. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87*, 917–927.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: Implications for epigenetic reprogramming. J. Biol. Chem. *279*, 52353–52360.

Morita, M., Stamp, G., Robins, P., Dulic, A., Rosewell, I., Hrivnak, G., Daly, G., Lindahl, T., and Barnes, D.E. (2004). Gene-targeted mice lacking the Trex1 (DNase III) 3'-->5' DNA exonuclease develop inflammatory myocarditis. Mol. Cell. Biol. *24*, 6719–6727.

Mortazavi, A., Chen, E., Thompson, L., Garcia, S.T., Myers, R.M., and Wold, B. (2006). Comparative genomics modeling of the NRSF / REST repressor network : From single conserved sites to genome-wide repertoire Comparative genomics modeling of the NRSF / REST repressor network : From single conserved sites to genome-wide repertoire. Genome Res. 1208–1221.

Muller, Molecular, D., and M, I.P.H. (1999). Uterine Natural Killer Cells Are Targets for a Protein A *. Endocrinology *140*, 2711–2720.

Munoz, M. a, Saunders, D.N., Henderson, M.J., Clancy, J.L., Russell, A.J., Lehrbach, G., Musgrove, E. a, Watts, C.K.W., and Sutherland, R.L. (2007). The E3 ubiquitin ligase EDD regulates S-phase and G(2)/M DNA damage checkpoints. Cell Cycle *6*, 3070–3077.

Muñoz-Espín, D., Cañamero, M., Maraver, A., Gómez-López, G., Contreras, J., Murillo-Cuesta, S., Rodríguez-Baeza, A., Varela-Nieto, I., Ruberte, J., Collado, M., et al. (2013). Programmed cell senescence during mammalian embryonic development. Cell *155*, 1104–1118.

Nabel, C.S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H.L., Stivers, J.T., Zhang, Y., and Kohli, R.M. (2012). AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. Nat. Chem. Biol. *8*, 751–758.

Najafabadi, H.S., Mnaimneh, S., Schmitges, F.W., Garton, M., Lam, K.N., Yang, A., Albu, M., Weirauch, M.T., Radovani, E., Kim, P.M., et al. (2015). C2H2 zinc finger proteins greatly expand the human regulatory lexicon. Nat Biotechnol *33*, 555–562.

Nakagawa, T., Nabeshima, Y., and Yoshida, S. (2006). Supplemental Data Functional Identification of the Actual and Potential Stem Cell Compartments in Mouse Spermatogenesis. Dev. Cell *12*, 1–11.

Nellåker, C., Keane, T.M., Yalcin, B., Wong, K., Agam, A., Belgard, T.G., Flint, J., Adams, D.J., Frankel, W.N., and Ponting, C.P. (2012). The genomic landscape shaped by selection on transposable elements across 18 mouse strains. Genome Biol. *13*, R45.

O'Donnell, L. (2015). Mechanisms of spermiogenesis and spermiation and how they are disturbed. Spermatogenesis *4*, e979623.

O'Geen, H., Squazzo, S.L., Iyengar, S., Blahnik, K., Rinn, J.L., Chang, H.Y., Green, R., and Farnham, P.J. (2007). Genome-wide analysis of KAP1 binding suggests autoregulation of KRAB-ZNFs. PLoS Genet. *3*, 0916–0926.

Oatley, J.M., and Brinster, R.L. (2008). Regulation of spermatogonial stem cell self-renewal in mammals. Annu. Rev. Cell Dev. Biol. *24*, 263–286.

Oda, M., Oxley, D., Dean, W., and Reik, W. (2013). Regulation of Lineage Specific DNA Hypomethylation in Mouse Trophectoderm. PLoS One *8*, 1–12.

Okano, M., Bell, D.W., Haber, D. a, and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247–257.

Ollinger, R., Childs, A.J., Burgess, H.M., Speed, R.M., Lundegaard, P.R., Reynolds, N., Gray, N.K., Cooke, H.J., and Adams, I.R. (2008). Deletion of the pluripotency-associated Tex19.1 gene causes activation of endogenous retroviruses and defective spermatogenesis in mice. PLoS Genet. *4*, e1000199.

Ollinger, R., Reichmann, J., and Adams, I.R. (2010). Meiosis and retrotransposon silencing during germ cell development in mice. Differentiation. *79*, 147–158.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. *10*, 475–478.

Palmer, S.J., and Burgoyne, P.S. (1991). In situ analysis of fetal, prepuberal and adult XX----XY chimaeric mouse testes: Sertoli cells are predominantly, but not exclusively, XY. Development *112*, 265–268.

Paludan, S., and Bowie, A. (2013). Immune Sensing of DNA. Immunity *38*, 870–880.

Patti, F., Cavallaro, T., Lo Fermo, S., Nicoletti, A., Cimino, V., Vecchio, R., Laisa, P., Zarbo, R., and Zappia, M. (2008). Is in utero early-exposure to interferon beta a risk factor

for pregnancy outcomes in multiple sclerosis? J. Neurol. *255*, 1250–1253.

Pavlicev, M., Hiratsuka, K., Swaggart, K.A., Dunn, C., and Muglia, L. (2015). Detecting endogenous retrovirus-driven tissue-specific gene transcription. Genome Biol. Evol. *7*, 1082–1097.

Peaston, A.E., Evsikov, A. V., Graber, J.H., de Vries, W.N., Holbrook, A.E., Solter, D., and Knowles, B.B. (2004). Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. Dev. Cell *7*, 597–606.

Peng, H., Begg, G.E., Schultz, D.C., Friedman, J.R., Jensen, D.E., Speicher, D.W., and Rauscher, F.J. (2000). Reconstitution of the KRAB-KAP-1 repressor complex: a model system for defining the molecular anatomy of RING-B box-coiled-coil domain-mediated protein-protein interactions. J. Mol. Biol. *295*, 1139–1162.

Peng, H., Feldman, I., and Rauscher, F.J. (2002). Hetero-oligomerization among the TIF family of RBCC/TRIM domain-containing nuclear cofactors: A potential mechanism for regulating the switch between coactivation and corepression. J. Mol. Biol. *320*, 629–644.

Peters, A.H.F.M., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schöfer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., et al. (2001a). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323–337.

Peters, B., Kirfel, J., Büssow, H., Vidal, M., and Magin, T.M. (2001b). Complete cytolysis and neonatal lethality in keratin 5 knockout mice reveal its fundamental role in skin integrity and in epidermolysis bullosa simplex. Mol. Biol. Cell *12*, 1775–1789.

Peterson, C.L., and Laniel, M.-A. (2004). Histones and histone modifications. Curr. Biol. *14*, R546–R551.

Phalke, S., Nickel, O., Walluscheck, D., Hortig, F., Onorati, M.C., and Reuter, G. (2009). Retrotransposon silencing and telomere integrity in somatic cells of Drosophila depends on the cytosine-5 methyltransferase DNMT2. Nat Genet *41*, 696–702.

Piccolo, F.M., and Fisher, A.G. (2014). Getting rid of DNA methylation. Trends Cell Biol. *24*, 136–143.

Platanias, L.C. (2005). Mechanisms of type-I- and type-II-interferon-mediated signalling. Nat. Rev. Immunol. *5*, 375–386.

Platt, J.S., and Hunt, J.S. (1998). Interferon- □ gene expression in cycling and pregnant mouse uterus : temporal aspects and cellular localization pregnant mouse uteri . In cycling uteri , IFN- □ was during the estrus phase of the cycle . In pregnant. J. Leukoc. Biol. *64*, 393–400.

Playfoot, C.J., and Adams, I.R. (2016). Previews KRABs RegulaTE Gene Expression beyond the Embryo Embracing the Enemy : Cell-to-Cell Force Transmission Enhances Cytotoxicity. Dev. Cell *36*, 591–592.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature *463*, 1101–1105.

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

Raab, J.R., and Kamakaka, R.T. (2010). Insulators and promoters: closer than we think. Nat. Rev. Genet. *11*, 439–446.

Rai, A., and Cross, J.C. (2014). Development of the hemochorial maternal vascular spaces in the placenta through endothelial and vasculogenic mimicry. Dev. Biol. *387*, 131–141.

Rasmussen, K.D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. Genes Dev. *30*, 733–750.

Razin, A., Szyf, M., Yisraeli, J., and Cedar, H. (1984). during differentiation. *81*, 2275–2279.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593–599.

Reichmann, J., Crichton, J.H., Madej, M.J., Taggart, M., Gautier, P., Garcia-Perez, J.L., Meehan, R.R., and Adams, I.R. (2012). Microarray analysis of LTR retrotransposon silencing identifies Hdac1 as a regulator of retrotransposon expression in mouse embryonic stem cells. PLoS Comput. Biol. *8*, e1002486.

Reichmann, J., Reddington, J.P., Best, D., Read, D., Ollinger, R., Meehan, R.R., and Adams, I.R. (2013). The genome-defence gene Tex19.1 suppresses LINE-1 retrotransposons in the placenta and prevents intra-uterine growth retardation in mice. Hum. Mol. Genet. *22*, 1791–1806.

Reik, W., and Walter, J. (2001). Genomic imprinting: parental influence on the genome. Nat. Rev. Genet. *2*, 21–32.

Reynolds, N., Collier, B., Maratou, K., Bingham, V., Speed, R.M., Taggart, M., Semple, C.A., Gray, N.K., and Cooke, H.J. (2005). Dazl binds in vivo to specific transcripts and can regulate the pre-meiotic translation of Mvh in germ cells. Hum. Mol. Genet. *14*, 3899–3909.

Riley, P., Anson-Cartwright, L., and Cross, J.C. (1998). The Hand1 bHLH transcription factor is essential for placentation and cardiac morphogenesis. Nat. Genet. *18*, 271–275.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Rohde, C., Zhang, Y., Reinhardt, R., and Jeltsch, A. (2010). BISMA--fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. BMC Bioinformatics *11*, 230.

de Rooij, D.G., and Mizrak, S.C. (2008). Deriving multipotent stem cells from mouse spermatogonial stem cells: a new tool for developmental and clinical research. Development *135*, 2207–2213.

De Rooij, D.G., and Russell, L.D. (2000). All you wanted to know about spermatogonia but were afraid to ask. J. Androl. *6512*, 776–798.

Rossant, J., and Cross, J.C. (2001). Placental development: lessons from mouse mutants. Nat. Rev. Genet. *2*, 538–548.

Rothbart, S.B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A.I., Barsyte-lovejoy, D., Martinez, J.Y., Bedford, M.T., Fuchs, S.M., et al. (2012). Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. Nat. Struct. Mol. Biol. *19*, 1155–1160.

Roulois, D., Loo Yau, H., Singhania, R., Wang, Y., Danesh, A., Shen, S.Y., Han, H., Liang, G., Jones, P.A., Pugh, T.J., et al. (2015). DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. Cell *162*, 961–973.

Rowe, H.M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., Maillard, P. V, Layard-Liesching, H., Verp, S., Marquis, J., et al. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. Nature *463*, 237–240.

Rowe, H.M., Friedli, M., Offner, S., Verp, S., Mesnard, D., Marquis, J., Aktas, T., and Trono, D. (2013a). De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. Development *140*, 519–529.

Rowe, H.M., Kapopoulou, A., Corsinotti, A., Fasching, L., Macfarlan, T.S., Tarabay, Y., Viville, S., Jakobsson, J., Pfaff, S.L., and Trono, D. (2013b). TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. Genome Res. *23*, 452–461.

Rumbajan, J.M., Yamaguchi, Y., Nakabayashi, K., Higashimoto, K., Yatsuki, H., Nishioka, K., Matsuoka, K., Aoki, S., Toda, S., Takeda, S., et al. (2016). The HUS1B promoter is hypomethylated in the placentas of low-birth-weight infants. Gene *583*, 141–146.

Sadate-Ngatchou, P.I., Payne, C.J., Dearth, A.T., and Braun, R.E. (2008). Cre recombinase activity specific to postnatal, premeiotic male germ cells in transgenic mice. Genesis *46*, 738–742.

Sandovici, I., Hoelle, K., Angiolini, E., and Constância, M. (2012). Placental adaptations to the maternal-fetal environment: Implications for fetal growth and developmental programming. Reprod. Biomed. Online *25*, 68–89.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev. Biol. *241*, 172–182.

Saunders, D.N., Hird, S.L., Withington, S.L., Dunwoodie, S.L., Henderson, M.J., Biben, C., Sutherland, R.L., Ormandy, C.J., Colin, K., Watts, W., et al. (2004). Edd , the Murine Hyperplastic Disc Gene , Is Essential for Yolk Sac Vascularization and Chorioallantoic Fusion Edd , the Murine Hyperplastic Disc Gene , Is Essential for Yolk Sac Vascularization and Chorioallantoic Fusion.

Saurin, A.J., Borden, K.L.B., Boddy, M.N., and Freemont, P.S. (1996). Does this have a familiar RING? Trends Biochem. Sci. *21*, 208–214.

Schemmer, J., Araúzo-Bravo, M.J., Haas, N., Schäfer, S., Weber, S.N., Becker, A., Eckert, D., Zimmer, A., Nettersheim, D., and Schorle, H. (2013). Transcription Factor TFAP2C Regulates Major Programs Required for Murine Fetal Germ Cell Maintenance and Haploinsufficiency Predisposes to Teratomas in Male Mice. PLoS One *8*, 1–15.

Schmidt, A., Morales-Prieto, D.M., Pastuschek, J., Fruhlich, K., and Markert, U.R. (2015). Only humans have human placentas: Molecular differences between mice and humans. J. Reprod. Immunol. *108*, 65–71.

Schneider, W.M., Chevillotte, M.D., and Rice, C.M. (2014). Interferon-stimulated genes: a complex web of host defenses. Annu. Rev. Immunol. *32*, 513–545.

Schnütgen, F., De-Zolt, S., Van Sloun, P., Hollatz, M., Floss, T., Hansen, J., Altschmied, J., Seisenberger, C., Ghyselinck, N.B., Ruiz, P., et al. (2005). Genomewide production of multipurpose alleles for the functional analysis of the mouse genome. Proc. Natl. Acad. Sci. U. S. A. *102*, 7221–7226.

Schrans-Stassen, B.H., Saunders, P.T., Cooke, H.J., and de Rooij, D.G. (2001). Nature of the spermatogenic arrest in Dazl -/- mice. Biol. Reprod. *65*, 771–776.

Schultz, D.C., Friedman, J.R., and Rauscher, F.J. (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: The PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2?? subunit of NuRD. Genes Dev. *15*, 428–443.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J. (2002). SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. *16*, 919–932.

Scialpi, F., Mellis, D., and Ditzel, M. (2015). EDD, a ubiquitin-protein ligase of the N-end rule pathway, associates with spindle assembly checkpoint components and regulates the mitotic response to nocodazole. J. Biol. Chem. *290*, 12585–12594.

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W., and Reik, W. (2012). The Dynamics of Genome-wide DNA Methylation Reprogramming in Mouse Primordial Germ Cells. Mol. Cell *48*, 849–862.

Seisenberger, S., Peat, J.R., Hore, T.A., Santos, F., Dean, W., Reik, W., Hemberger, M., Dean, W., Reik, W., Bostick, M., et al. (2013). Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *368*, 20110330.

Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M., and 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*, 440–458.

Sharma, D., Shastri, S., and Sharma, P. (2016). Intrauterine Growth Restriction: Antenatal and Postnatal Aspects. Clin. Med. Insights Pediatr. 67.

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*, 459–470.

Shibata, M., and Garc??a-Garc??a, M.J. (2011). The mouse KRAB zinc-finger protein CHATO is required in embryonic-derived tissues to control yolk sac and placenta morphogenesis. Dev. Biol. *349*, 331–341.

Shibata, M., Blauvelt, K.E., Liem, K.F., and García-García, M.J. (2011). TRIM28 is required by the mouse KRAB domain protein ZFP568 to control convergent extension and morphogenesis of extra-embryonic tissues. Development *138*, 5333–5343.

Simmons, D.G., and Cross, J.C. (2005). Determinants of trophoblast lineage and cell subtype specification in the mouse placenta. Dev. Biol. *284*, 12–24.

Simmons, D.G., Fortier, A.L., and Cross, J.C. (2007). Diverse subtypes and developmental origins of trophoblast giant cells in the mouse placenta. Dev. Biol. *304*, 567–578.

Simmons, D.G., Rawn, S., Davies, A., Hughes, M., and Cross, J.C. (2008). Spatial and temporal expression of the 23 murine Prolactin/Placental Lactogen-related genes is not associated with their position in the locus. BMC Genomics *9*, 352.

Sio, F.R.S. De, Massacand, J., Barde, I., Offner, S., Corsinotti, A., Kapopoulou, A., Bojkowska, K., Fernandez, M., Ghia, P., Thomas, J.H., et al. (2013). Europe PMC Funders Group KAP1 regulates gene networks controlling mouse B lymphoid cell differentiation and

function. *119*, 4675–4685.

Smallwood, S.A., and Kelsey, G. (2012). De novo DNA methylation: A germ cell perspective. Trends Genet. *28*, 33–42.

Smit, A.F.A. (1993). Identification of a new, abundant superfamily of mammalian LTR-transposons. Nucleic Acids Res. *21*, 1863–1872.

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

Smyth, G.K. (2004). Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. Stat. Appl. Genet. Mol. Biol. *3*, 1–26.

Soares, M.J. (2004). The prolactin and growth hormone families: pregnancy-specific hormones/cytokines at the maternal-fetal interface. Reprod. Biol. Endocrinol. *2*, 51.

Solyom, S., Ewing, A.D., Rahrmann, E.P., Doucet, T., Nelson, H.H., Burns, M.B., Harris, R.S., Sigmon, D.F., Casella, A., Erlanger, B., et al. (2012). Extensive somatic L1 retrotransposition in colorectal tumors. Genome Res. *22*, 2328–2338.

Sookdeo, A., Hepp, C.M., McClure, M. a, and Boissinot, S. (2013). Revisiting the evolution of mouse LINE-1 in the genomic era. Mob. DNA *4*, 3.

Soper, S.F.C., van der Heijden, G.W., Hardiman, T.C., Goodheart, M., Martin, S.L., de Boer, P., and Bortvin, A. (2008). Mouse Maelstrom, a Component of Nuage, Is Essential for Spermatogenesis and Transposon Repression in Meiosis. Dev. Cell *15*, 285–297.

Speek, M. (2001). Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Mol. Cell. Biol. *21*, 1973–1985.

Sripathy, S.P., Stevens, J., and Schultz, D.C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. Mol. Cell. Biol. *26*, 8623–8638.

Stetson, D.B., Ko, J.S., Heidmann, T., and Medzhitov, R. (2008). Trex1 Prevents Cell-Intrinsic Initiation of Autoimmunity. Cell *134*, 587–598.

Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S., and Jacobsen, S.E. (2011). 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. Genome Biol. *12*, R54.

Szwagierczak, A., Bultmann, S., Schmidt, C.S., Spada, F., and Leonhardt, H. (2010). Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Res. *38*.

Tachibana, M., Sugimoto, K., Nozaki, M., Ueda, J., Ohta, T., Ohki, M., Fukuda, M., Takeda, N., Niida, H., Kato, H., et al. (2002). G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. Genes Dev. *16*, 1779–1791.

Tachibana, M., Ueda, J., Fukuda, M., Takeda, N., Ohta, T., Iwanari, H., Sakihama, T., Kodama, T., Hamakubo, T., and Shinkai, Y. (2005). Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev. *19*, 815–826.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930–935.

Takahashi, K., and Nakayama, K. (2000). Mice lacking a CDK inhibitor, p57Kip2, exhibit skeletal abnormalities and growth retardation. J. Biochem. *127*, 73–83.

Takebayashi, S., Tamura, T., Matsuoka, C., and Okano, M. (2007). Major and essential role for the DNA methylation mark in mouse embryogenesis and stable association of DNMT1 with newly replicated regions. Mol. Cell. Biol. *27*, 8243–8258.

Tanaka, M., Gertsenstein, M., Rossant, J., and Nagy, A. (1997). Mash2 acts cell autonomously in mouse spongiotrophoblast development. Dev. Biol. *190*, 55–65.

Thompson, P.J., Macfarlan, T.S., and Lorincz, M.C. (2016). Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire. Mol. Cell *62*, 766–776.

Thorrez, L., Van Deun, K., Tranchevent, L.C., Van Lommel, L., Engelen, K., Marchal, K., Moreau, Y., Van Mechelen, I., and Schuit, F. (2008). Using ribosomal protein genes as reference: A tale of caution. PLoS One *3*.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. *7*, 562–578.

Tunster, S.J., Tycko, B., and John, R.M. (2010). The imprinted Phlda2 gene regulates extraembryonic energy stores. Mol. Cell. Biol. *30*, 295–306.

Tunster, S.J., Van de Pette, M., and John, R.M. (2011). Fetal overgrowth in the Cdkn1c mouse model of Beckwith-Wiedemann syndrome. Dis. Model. Mech. *4*, 814–821.

Tunster, S.J., Jensen, A.B., and John, R.M. (2013). Imprinted genes in mouse placental development and the regulation of fetal energy stores. Reproduction *145*.

Tunster, S.J., Van De Pette, M., and John, R.M. (2014). Isolating the role of elevated Phlda2 in asymmetric late fetal growth restriction in mice. Dis. Model. Mech. *7*, 1185–1191.

Tunster, S.J., Creeth, H.D.J., and John, R.M. (2016). The imprinted Phlda2 gene modulates a major endocrine compartment of the placenta to regulate placental demands for maternal resources. Dev. Biol. *409*, 251–260.

Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. Protein Sci. *20*, 1298–1345.

Walsh, C.P., Chaillet, J.R., and Bestor, T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat. Genet. *20*, 116–117.

Walter, M., Teissandier, A., Pérez-Palacios, R., and Bourc'his, D. (2016). An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. Elife 1–30.

Wang, P.J., McCarrey, J.R., Yang, F., and Page, D.C. (2001). An abundance of X-linked genes expressed in spermatogonia. Nat. Genet. *27*, 422–426.

Wang, R., Wang, J., Paul, A.M., Acharya, D., Bai, F., Huang, F., and Guo, Y.L. (2013). Mouse embryonic stem cells are deficient in type i interferon expression in response to viral infections and double-stranded RNA. J. Biol. Chem. *288*, 15926–15936.

Ward, C.L., Omura, S., and Kopito, R.R. (1995). Degradation of CFTR by the ubiquitin-proteasome pathway. Cell *83*, 121–127.

Watanabe, D., Suetake, I., Tada, T., and Tajima, S. (2002). Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. Mech. Dev. *118*, 187–190.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., et al. (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520–562.

Watson, E., and Cross, J. (2005). Development of Structures and Transport Functions in the Mouse Placenta. Physiology *20*, 180–193.

Weber, J.E., and Russell, L.D. (1987). A study of intercellular bridges during spermatogenesis in the rat. Am. J. Anat. *180*, 1–24.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. *39*, 457–466.

Weber, P., Cammas, F., Gerard, C., Metzger, D., Chambon, P., Losson, R., and Mark, M. (2002). Germ cell expression of the transcriptional co-repressor TIF1β is required for the maintenance of spermatogenesis in the mouse. Development *129*, 2329–2337.

Weber-Schoendorfer, C., and Schaefer, C. (2009). Multiple sclerosis, immunomodulators, and pregnancy outcome: a prospective observational study. Mult.Scler. *15*, 1037–1042.

Williams, K., Christensen, J., and Helin, K. (2012). DNA methylation: TET proteins-guardians of CpG islands? EMBO Rep. *13*, 28–35.

Wiznerowicz, M., Jakobsson, J., Szulc, J., Liao, S., Quazzola, A., Beermann, F., Aebischer, P., and Trono, D. (2007). The Kruppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. J. Biol. Chem. *282*, 34535–34541.

Wolf, G., Yang, P., Fuchtbauer, A.C., Fuchtbauer, E.M., Silva, A.M., Park, C., Wu, W., Nielsen, A.L., Pedersen, F.S., and Macfarlan, T.S. (2015a). The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. Genes Dev. *29*, 538–554.

Wolf, G., Greenberg, D., and Macfarlan, T.S. (2015b). Spotting the enemy within: Targeted silencing of foreign DNA in mammalian genomes by the Krüppel-associated box zinc finger protein family. Mob. DNA *6*, 17.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat. Commun. *2*, 241.

Wu, J.A., Johnson, B.L., Chen, Y., Ha, C.T., and Dveksler, G.S. (2008). Murine Pregnancy-Specific Glycoprotein 23 Induces the Proangiogenic Factors Transforming-Growth Factor Beta 1 and Vascular Endothelial Growth Factor A in Cell Types Involved in Vascular Remodeling in Pregnancy. Biol. Reprod. *79*, 1054–1061.

Xiao, T.Z., Bhatia, N., Urrutia, R., Lomberk, G.A., Simpson, A., and Longley, B.J. (2011). MAGE I transcription factors regulate KAP1 and KRAB domain zinc finger transcription factor mediated gene repression. PLoS One *6*.

Yamaji, M., Ueda, J., Hayashi, K., Ohta, H., Yabuta, Y., Kurimoto, K., Nakato, R., Yamada, Y., Shirahige, K., and Saitou, M. (2013). PRDM14 ensures naive pluripotency through dual

regulation of signaling and epigenetic pathways in mouse embryonic stem cells. Cell Stem Cell *12*, 368–382.

Yang, F., Cheng, Y., An, J.Y., Kwon, Y.T., Eckardt, S., Leu, N.A., McLaughlin, K.J., and Wang, P.J. (2010). The ubiquitin ligase Ubr2, a recognition E3 component of the N-end rule pathway, stabilizes Tex19.1 during spermatogenesis. PLoS One *5*, e14017.

Yoder, J. a, Walsh, C.P., and Bestor, T.H. (1997). Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. *13*, 335–340.

Yoshida, S., Sukeno, M., Nakagawa, T., Ohbo, K., Nagamatsu, G., Suda, T., and Nabeshima, Y. (2006). The first round of mouse spermatogenesis is a distinctive program that lacks the self-renewing spermatogonia stage. Development *133*, 1495–1505.

Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B., et al. (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell *149*, 1368–1380.

Yu, Q., Carbone, C.J., Katlinskaya, Y. V., Zheng, H., Zheng, K., Luo, M., Wang, P.J., Greenberg, R.A., and Fuchs, S.Y. (2015). Type I interferon controls propagation of long interspersed element-1. J. Biol. Chem. *290*, 10191–10199.

Zamudio, N., and Bourc'his, D. (2010). Transposable elements in the mammalian germline: a comfortable niche or a deadly trap? Heredity (Edinb). *105*, 92–104.

Zemach, A., and Zilberman, D. (2010). Evolution of eukaryotic DNA methylation and the pursuit of safer sex. Curr. Biol. *20*, R780–R785.

Zeng, L., Yap, K.L., Ivanov, A. V, Wang, X., Mujtaba, S., Plotnikova, O., Rauscher 3rd, F.J., and Zhou, M.M. (2008). Structural insights into human KAP1 PHD finger-bromodomain and its role in gene silencing. Nat Struct Mol Biol *15*, 626–633.

Zhang, L., Lu, X., Lu, J., Liang, H., Dai, Q., Xu, G.-L., Luo, C., Jiang, H., and He, C. (2012). Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. Nat. Chem. Biol. *8*, 1–27.

Zhao, K., Du, J., Han, X., Goodier, J.L., Li, P., Zhou, X., Wei, W., Evans, S.L., Li, L., Zhang, W., et al. (2013). Modulation of LINE-1 and Alu/SVA Retrotransposition by Aicardi-Gouti??res Syndrome-Related SAMHD1. Cell Rep. *4*, 1108–1115.

Zheng, K., Xiol, J., Reuter, M., Eckardt, S., Leu, N.A., McLaughlin, K.J., Stark, A., Sachidanandam, R., Pillai, R.S., and Wang, P.J. (2010). Mouse MOV10L1 associates with Piwi proteins and is an essential component of the Piwi-interacting RNA (piRNA) pathway. Proc. Natl. Acad. Sci. U. S. A. *107*, 11841–11846.

Zuo, X., Sheng, J., Lau, H.T., McDonald, C.M., Andrade, M., Cullen, D.E., Bell, F.T., Iacovino, M., Kyba, M., Xu, G., et al. (2012). Zinc finger protein ZFP57 requires its co-factor to recruit DNA methyltransferases and maintains DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. J. Biol. Chem. *287*, 2107–2118.

Zhu, J., He, F., Hu, S., and Yu, J. (2008). On the nature of human housekeeping genes. Trends Genet. *24,* 481-4.

Appendix

Over the course of my PhD I have participated in research outwith my own projects which has contributed to the submission of research papers to different journals. In addition I have contributed to two review articles and have a first author 'Preview' article, reviewing a novel research article in the wider context.

Publications:

Patricia Yeyati, P., Schiller, R., Mali, G., Kasioulis, I., Kawamura, A., Adams, I. R., **Playfoot**, **C.J.,** Gilbert, N., van Heyningen, V., Wills, J., von Kriegsheim, A., Finch, A., Sakai, J., Jackson, I., Schofield, C. and Mill., P. (2017) The JmjC Protein KDM3A Coordinates Actin Dynamics with Intraflagellar Transport to Modulate Cilia Number and Length. J. Cell. Biol. *doi: 10.1083/jcb.201607032*

Playfoot, C.J., and Adams, I.R. (2016) KRABs RegulaTE Gene Expression Beyond the Embryo. Dev. Cell *36,* 591-2.

MacLennan, M., Crichton, J.H., **Playfoot, C.J.**, and Adams, I.R. (2015) Oocyte development, meiosis and aneuploidy. Semin. Cell. Dev. Biol. *45,* 68-76.

Crichton, J.H., **Playfoot, C.J.**, and Adams, I.R. (2014). The Role of Chromatin Modifications in Progression Through Mouse Meiotic Prophase. J. Genet. Genomics 41, 97-106.

Publications in submission:

The Ubiquitin-Proteasome System Protects the Mammalian Germline from L1 Retrotransposons

Marie MacLennan, Marta García-Cañadas, Judith Reichmann, Carmen Salvador-Palomeques, Abigail R. Mann, Paula Peressini, Laura Sanchez, **Christopher J. Playfoot**, David Read, Chao Chun Hung, Ragnhild Eskeland, Richard R. Meehan, Jose Luis García-Pérez, Ian R. Adams

Publications in review:

Meiotic DNA Double Strand Break Frequency in Mammalian Spermatocytes is Regulated by *TEX19.1*

James H. Crichton, Marie MacLennan, **Christopher J. Playfoot**, David Read, Howard J. Cooke, and Ian R. Adams

Symbol	logFC_Placenta	logFC_PGCLCs	logFC_2iES
4933402E13Rik	7.983185158	3.961511196	4.179866353
A4galt	3.737236305	2.930979413	4.309420306
AA467197	6.417371007	3.736588066	5.402206519
Abcb1a	3.814330111	3.57831586	4.91421334
Abcb1b	4.369522631	3.676160665	5.352394988
AU018091	5.857654313	5.500459258	6.32521332
Cbx7	3.203157974	3.680920776	3.383435617
Cul4b	3.036131778	2.914766316	2.593616571
Dpp7	2.729644104	3.049419123	3.091902914
Fmr1nb	6.433691764	4.637685407	6.011252617
Folr1	5.586053317	4.231227271	4.784011553
Gcnt2	3.024394769	2.759626933	3.313339825
Gjb3	3.600998384	4.758097435	5.776291172
Hic2	2.968619883	2.718687479	3.250494199
Ildr1	2.849969542	3.052900878	4.580187712
Itpka	2.679080264	3.120155307	3.940295059
Jade1	2.922938861	2.992270133	3.434462249
Klf9	3.150905644	3.549836128	4.985934759
L1td1	3.486774498	7.247493591	6.960350571
Liph	4.633272039	3.721535787	3.701884199
Mageb16	3.307315339	4.603314251	5.181832026
Manba	2.762748999	3.791015622	3.873169169
Nabp1	3.954014382	3.411020412	3.384798759
Nmrk1	2.92345741	3.160769684	3.157196729
Pfkp	3.026999762	2.597023375	3.285298645
Plekhh1	4.431253702	3.417931103	3.104544519
Pramef12	5.877864928	3.647035173	3.70088934
Rhox5	7.94329767	6.834147562	5.337591616
Slc30a2	4.421529677	3.979874367	4.322193751
Slc35f2	2.623168212	4.082909855	4.329657991
Slc4a5	2.85062507	2.692356882	2.914986113
Stk31	5.492047732	5.08420851	3.341806772
Stx3	3.440116423	3.671175366	3.782642486
Taf7	2.916138833	2.991875457	4.500128901
Tdrp	5.33352406	3.812611925	4.534097239
Tex19.1	4.64023022	3.989158083	5.945526078
Tfap2c	4.682119987	4.851598927	3.081758011
Trap1a	8.222338447	6.924496437	8.641808957
Trim6	4.26142354	6.371426027	6.868824669
Upp1	3.972513128	2.813811106	6.35300458
Zfp42	5.065454448	8.589615418	7.995127119

Table 7.1: Genes commonly expressed in hypomethylated placenta, PGCLCs and 2i cultured ES cells compared to whole embryo (see Figure 3.1B). A 274tringent 6 fold cut off in expression difference was applied and only genes with a false discovery rate $P<0.05$ were included.

Symbol
Asz1
1700034E13Rik
4930550L24Rik
Akr1c12
Aldh3a1
Casp1
Dazl
Gpr97
Gstp2
IAP
Iigp2
Mili
Mov10l1
Nckap1l
Prelp
Rhox4d
Rps4y2
Slc15a3
Slc47a1
Tex13
Tex19.1
Tex19.2
Tnnt1
Xist
Xlr4a

Table 7.2: Genes determined as methylation sensitive in Hackett et al., 2012. (See figure 3.1B)

7.1: Optimisation of ChIP on placental tissue

The initial ChIP protocol was performed and optimised on around 10 million ES cells and ChIP was confirmed to be working as expected by use of positive and negative control loci such as the actin promoter for the mark of active transcription H3K4me3 and an intergenic region respectively. For H3K9me3 positive control loci were retrotransposons and imprinted genes whereas actin promoter was used as a negative control as it is not marked by H3K9me3. ChIP on 16dpp testes material was performed essentially as with ES cells due to the ease of generating a single cell suspension and efficient lysis in Farnham buffer. Volumes were reduced due to the low cell number

to limit cell loss. The major optimisation step in performing ChIP on whole placenta was the generation of a single cell suspension. In initial experiments placentas were enzymatically digested with collagenase or trypsin whilst being agitated at 37ºC for differing incubation lengths, followed by filtering for single cells. For all incubation lengths, a low number of cells were recovered. Maceration of the tissue with razor blades, followed by homogenisation with a hand homogeniser provided a greater yield of single cells recovered and was therefore used in all future experiments. Next, it was necessary to optimise sonication of extracted chromatin as different cells require different lengths of sonication to achieve the desired 200- 400bp DNA fragments. For ES cells, lysing in Farnham lysis buffer (detergent 0.5% Np-40) and sonicating for 9x 30 seconds on/off in RIPA buffer (detergent 1% Np-40, 0.1% SDS) was sufficient to shear chromatin to a size of 100 - 400bp (figure 7.1A). This was not sufficient for shearing E12.5 placental chromatin which remained a high molecular weight indicating inefficient sonication (figure 7.1A). Use of a different lysis buffer containing more concentrated detergent at 1% SDS and 0.1% Triton X-100 resulted in effective sonication even when sonicating for 3x 30 seconds on/off (figure 7.1B). These conditions were used in all future ChIP experiments on placental material. Positive and negative control loci were used in qPCR reactions on ChIP IP and input material and indicated differing enrichment in different tissues, likely reflecting both technical differences in the preparation of the tissues and real biological differences in the enrichment of histone modifications at certain control loci (for H3K4me3 compare figure 3.2 (testes and placenta) and figure 3.9E (ESCs), for H3K9me3 compare figure 3.4A (ESCs) with figure 3.14 (placenta)).

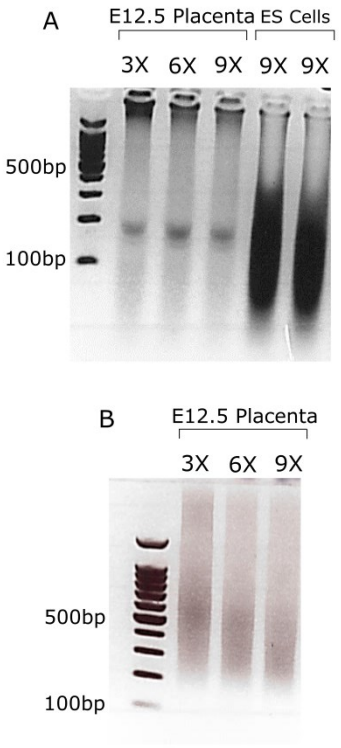


Figure 7.1: Optimisation of ChIP on E12.5 placenta. (A) Gel electrophoresis showing DNA fragment sizes after sonication of E12.5 placenta and ES cells for 3x, 6x or 9x 30 seconds on/off after lysis in Farnham lysis buffer and RIPA buffer. (B) DNA fragment sizes of E12.5 placenta after lysis in 1% SDS lysis buffer and sonication for 3x, 6x or 9x 30 seconds on/off. DNA ladder is along the left side of both gels with 500bp and 100bp bands indicated.

LASU1 [Mus musculus]
pericentriolar material-1 [Mus musculus]
E3 ubiquitin-protein ligase UBR4 [Mus musculus]
alpha-fetoprotein [Mus musculus]
midasin homolog (yeast) [Mus musculus]
leucine zipper protein 1 [Mus musculus]
mKIAA1227 protein [Mus musculus] (INVESTIGATE FURTHER) L1Mdf3 flanking it at 5' end. Associated w
TI-225 [Mus musculus]
pol [Mus musculus]
mCG13235 [Mus musculus]
leucine zipper protein 1 [Mus musculus]
KRAB-A interacting protein [Mus musculus]
mKIAA1227 protein [Mus musculus]
TI-225 [Mus musculus]
mCG13235 [Mus musculus]
nuclear receptor subfamily 0 group B member 1 [Mus musculus]
matricin [Mus musculus]
ruvB-like 2 [Mus musculus]
TI-225 [Mus musculus]
ATP-dependent zinc metalloprotease YME1L1 [Mus musculus]
prolyl-tRNA synthetase (mitochondrial)(putative), isoform CRA_a [Mus musculus]
L-threonine 3-dehydrogenase, mitochondrial precursor [Mus musculus]
translocase of inner mitochondrial membrane 50 homolog (yeast), isoform CRA_a [Mus musculus]
fructose-bisphosphate aldolase A isoform 2 [Mus musculus]
TI-225 [Mus musculus]
Wdr18 protein [Mus musculus]
Otx2 [Mus musculus]
coatomer subunit epsilon [Mus musculus]
transforming growth factor-beta homolog [Mus sp.]
aldolase 1 A retrogene 1 [Mus musculus]
26S protease regulatory subunit 7 [Mus musculus]
Rfc5 protein [Mus musculus]
mCG13052 [Mus musculus]
histone H2A [Mus musculus domesticus]
ubiquitin-conjugating enzyme E2 A [Mus musculus]
tumor metastatic process-associated protein NM23 [Mus musculus]
exportin-2 [Mus musculus]

Table 7.3: Proteins identified as binding to TEX19.1-YFP in mass spectrometry experiment by Marie Maclennan

Significantly Upregulated Genes				Significantly Downregulated Genes		
GeneSymbol	logFC	FDR		GeneSymbol	logFC	FDR
3300005D01Rik	1.029278	0.000534		1600012P17Rik	-1.29707	5.99E-05
4932415M13Rik	1.276428	0.005945		1600015I10Rik	-1.02898	0.039436
A2m	1.031734	0.033304		Adra2c	-1.13998	1.37E-05
Adamdec1	2.044209	3.8E-08		Aldh1a3	-1.07004	3.03E-05
Adcy2	1.187963	0.044569		Apc2	-1.21629	1.86E-07
Adrb2	1.01024	0.007003		Brsk1	-1.10599	0.0048
Ankrd1	1.708689	0.014032		Ceacam11	-1.27393	3.77E-05
Anxa8	1.00943	0.00025		Ceacam12	-2.05259	1.69E-10
Aqp1	1.214705	0.031963		Ceacam13	-1.22688	1.41E-08
Arl4d	1.045138	0.00199		Ceacam14	-1.40467	0.003365
Bok	1.04439	0.001141		Ceacam3	-1.3127	8.61E-09
C1qtnf9	1.267154	2.2E-05		Ceacam5	-1.40655	1.35E-07
Calm4	1.382947	0.002826		Ceacam9	-1.37514	8.52E-05
Cck	1.276482	0.027942		Creg1	-1.47472	9.57E-05
Cldn1	1.145211	0.008903		Cts3	-1.73347	1.86E-07
Cldn10	1.402986	0.007085		Ctsm	-2.24019	1.42E-17
Ctla2a	1.061052	0.000691		Ctsq	-1.71	1.69E-07
Cxcl14	1.123491	9.44E-05		Cxcl12	-1.1133	0.024546
Cyp11b1	1.082185	0.016763		Cyp4f18	-1.12622	5.14E-06
Cyp2ab1	1.151176	0.014514		Fa2h	-1.40411	0.036957
Des	1.065278	0.002138		Fbln7	-1.24311	0.013239
Dppa4	1.381569	1.12E-05		Flt1	-1.31618	0.000144
Dppa5a	1.065411	0.000746		Fndc5	-1.13143	0.000725
Ear1	1.616574	9.43E-05		Gm16565	-1.6193	6.61E-06
Ear10	1.41865	0.000255		Gm5155	-1.37879	3.53E-06
Ear2	1.523278	2E-07		Gm5506	-1.01909	0.031963
Edn1	1.088938	0.000832		Gpr156	-1.07276	0.000146
Eno2	1.135123	0.006025		Gpr44	-1.14329	0.002991
Epdr1	1.015268	0.01455		Hsd11b2	-1.01211	7.59E-09
Eva1c	1.018623	0.00475		Krt5	-4.04683	2.14E-33
Fabp4	1.017015	0.000651		Mmp1a	-1.48207	0.000267
Fam13c	1.141139	0.031829		Mn1	-1.08153	9.99E-05
Fam189a2	1.19341	0.003003		Pappa2	-1.29173	0.000534
Fbln2	1.043691	0.021354		Prl2b1	-1.36814	0.000522
Gdf10	1.522726	0.005528		Prl3b1	-1.09815	4.07E-06
Gml	1.157103	0.001679		Prl3c1	-1.5419	9.94E-09
Golga7b	1.096617	0.000675		Prl4a1	-1.46978	1.06E-09
Gpihbp1	1.069785	1.31E-05		Prl6a1	-1.30845	6.86E-15

Gpr87	1.12969	0.027023		Prl7a1	-1.05781	0.037819
Havcr1	1.211973	0.035784		Prl7a2	-1.4327	3.01E-09
Hemgn	1.112949	0.036095		Prl7c1	-1.16321	1.44E-11
Hemt1	1.030674	0.040951		Prl8a6	-1.04913	0.041918
Hist1h2bp	1.008618	9.89E-05		Prl8a8	-1.42854	0.038621
HOXA11-AS1_6	1.017666	0.000671		Prl8a9	-1.09554	2.67E-05
Hoxd10	1.116209	5.93E-05		Psg16	-1.56743	0.005213
Hsd11b1	1.006912	0.001468		Psg17	-1.75492	7.89E-05
Ifne	1.189526	9.09E-05		Psg18	-1.55582	1.32E-05
Igfbp6	1.081969	0.000384		Psg19	-1.39033	5.57E-07
Klra7	1.078566	0.028572		Psg21	-2.22963	1.1E-05
Krt14	1.246553	0.000376		Psg23	-2.4419	4.7E-05
Mfap5	1.023264	0.000396		Psg25	-1.76078	3.01E-09
Mir295	1.458335	3.8E-08		Psg26	-1.7005	8.06E-06
Mir322	1.015705	4.29E-05		Psg27	-2.01002	4.97E-07
Mt3	1.8738	0.00162		Psg28	-1.51336	7.76E-05
Mustn1	1.057059	0.000248		Psg29	-1.44549	4.1E-18
Mybpc3	1.117705	0.010629		Ptch1	-1.10514	0.009094
Prkg2	1.145144	0.014032		Ptpru	-1.08068	9.39E-07
Prss23	1.190848	0.021679		Rgs5	-1.3205	1.54E-22
Qpct	1.010915	0.018972		Slc26a4	-1.07783	0.003716
Ramp1	1.080112	5.8E-05		Slc39a5	-1.05929	0.013246
S100a4	1.020215	2.2E-05		Stc1	-1.09586	0.00025
Sfrp4	1.426423	0.003152		Stox2	-1.14883	0.000255
Sfrp5	1.061702	0.000972		Tet1	-1.78423	1.69E-10
Shisa3	1.009005	0.033951		Tex19.1	-12.4887	3.96E-84
Siglecg	1.161228	0.003152		Tpbpb	-1.09519	0.005897
Smoc2	1.104101	0.001411				
Sox5	1.025403	0.045749				
Sphk1	1.08044	0.004662				
Srgn	1.020796	0.020438				
Steap4	1.192673	0.014877				
Tac2	1.234373	0.008903				
Thbs2	1.270068	0.006599				
Tmem141	1.079668	0.000163				
Trem2	1.284251	9.2E-06				
Trpc4	1.210299	0.008012				
Twist2	1.251686	1.52E-05				
Wfdc3	1.131974	9.95E-05				

Table 7.4: Genes significantly up and downregulated in *Tex19.1*$^{-/-}$ placentas compared to control placentas. Genes have LogFC >= 1 and FDR <= 0.05 and correspond to points shown in figure 4.10.

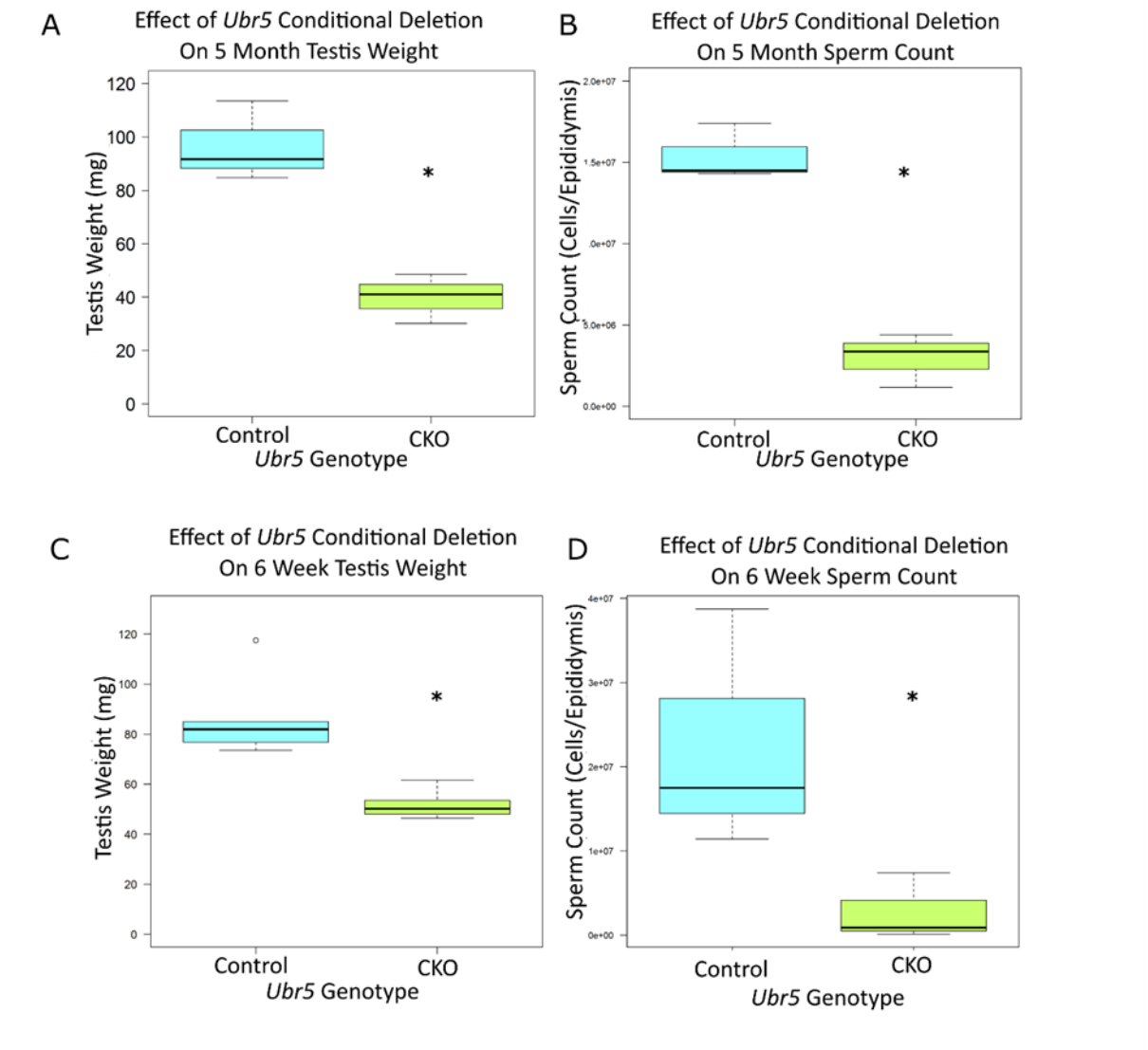


Figure 7.2: Comparison of testis weights and sperm counts from 6 week and 5 month old *Ubr5^{CKO}* animals (A) Boxplot showing the median testis weights of 5 month control and *Ubr5^{CKO}* mice and (C) 6 week old mice (5 month: n=3 for both genotypes, 6 week: n=5 for both genotypes *P<0.05, Mann Whitney U-test). (B) Boxplot showing the median sperm count from 5 month old epididymis from control and *Ubr5^{CKO}* mice and (D) 6 week old epididymis (5 month: n=3 for both genotypes, 6 week: n=5 for both genotypes. *P<0.05, Mann Whitney U-test)