The role of the X chromosome in embryonic and postnatal growth

Daniel Mark Snell

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Francis Crick Institute/Medical Research Council National Institute for Medical
Research
University College London

January 28, 2018
I, Daniel Mark Snell, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.
Abstract

Women born with only a single X chromosome (XO) have Turner syndrome (TS); and they are invariably of short stature. XO female mice are also small: during embryogenesis, female mice with a paternally-inherited X chromosome (X^P_O) are smaller than XX littermates; whereas during early postnatal life, both X^P_O and X^M_O (maternal) mice are smaller than their XX siblings. Here I look to further understand the genetic bases of these phenotypes, and potentially inform areas of future investigation into TS.

Mouse pre-implantation embryos preferentially silence the X^P via the non-coding RNA Xist. X^P_O embryos are smaller than XX littermates at embryonic day (E) 10.5, whereas X^M_O embryos are not. Two possible hypotheses explain this observation. Inappropriate expression of Xist in X^P_O embryos may cause transcriptional silencing of the single X chromosome and result in embryos nullizygous for X gene products. Alternatively, there could be imprinted genes on the X chromosome that impact on growth and manifest in growth retarded X^P_O embryos. In contrast, during the first three weeks of postnatal development, both X^P_O and X^M_O mice show a growth deficit when compared with XX littermates. This deficit is not observed in the presence of a second sex chromosome - i.e. in normal XX female mice, or in females with a Y chromosome that lacks Sry - suggesting haploinsufficiency of genes with homologues present on, and expressed from, both sex chromosomes as a cause.
In this thesis I have investigated the role of \textit{Xist} in \textit{XPO} embryonic growth retardation; and utilised mouse stem cells to perform an \textit{in vitro} screen to identify X-linked imprinted genes. To characterise postnatal haploinsufficiency, I identify four candidate genes and, utilising CRISPR-Cas genome editing, delineate the role of each in the growth deficit phenotype. I further use these X-linked mutants to investigate the functional divergence of the X and Y chromosomes in the context of postnatal survival.
"There is nothing like looking, if you want to find something. You certainly usually find something, if you look, but it is not always quite the something you were after."

Whilst this thesis bears my name, undoubtedly I couldn’t have finished it without the generous support, advice and encouragement of many colleagues and friends. Firstly, thanks to my supervisor, James, for agreeing to take a punt on me, probably somewhat unexpectedly, after our initial meeting. I am particularly appreciative of your input during the writing process, during which my ability to construct short, logical, unambiguous sentences has improved immeasurably...

Thanks to the whole Turner lab, past and present, for collectively making it a pleasure to come to work: Andrew, Alex, Ben, Bryony, Charlotte, Elias, Shantha, Taka, and Valdone. I pay particular thanks to Fanny, without whose patient teaching and timely words of encouragement I would not be half the molecular biologist (and person) I am today. I am also grateful to Mahesh for analysing the RNA-seq data, and for helpful discussions regarding data interpretation, and likewise to Jasmin. Many thanks to Oana, who is likely the most conscientious and hard working BSc and MSc project student anyone could ever wish for.

The writing of this thesis has coincided with the lab migration from an old, characterful asylum in Mill Hill to "Sir Paul’s Cathedral" in central London. I’d like to thank Kathy and the Niakan lab for providing me space to work during the migration
Acknowledgements

and beyond. I’m particularly indebted to Sissy for many fascinating, wide-ranging
communications, both scientific and philosophical.

Thanks to the former Developmental Genetics Division for their scientific advise
and weekly cake; and in particular to Christophe for motivational words at key
times (and for having very slightly less hair than me at all times).

None of the work presented here would have been possible without the support of
the Biological Research Facility. In chronological order, thanks to Hannah, Sue,
Pete, and Pat for providing excellent husbandry over many years. I’m also majorly
indebted to the former Procedural Services Section from Mill Hill: almost with-
out exception, Sophie expertly carried out all of the microinjection work for my
projects, as well as tolerating my occasional moans when things didn’t go to plan.
Thanks must also be extended to Sarah, Marta, Nicolle, and Katharine.

From UCL, I will be forever grateful to (now) Emeritus Professor Gordon Stewart
for giving me the opportunity to pursue the slightly different MBPhD academic
track during medicine. I appreciate the input of my Thesis Committee, Alex Gould,
Vivian Li, and Robin Lovell-Badge; and thanks to both Rosetrees Trust and Medical
Research Council for funding my work.

Thanks to my friends - Jack, Lara, Charlie, Ben, and Saba - for showing great
understanding when I’ve either been late or not made it to things because I was in
the lab. I can’t promise I’ve quite got the balance right just yet, but I’m making
progress. This progress is indisputably the result of the influence of Daria. Thank
you for helping me to realise the need to live outside of work, and for encouraging
me to do things I never imagined I would a year ago. Thank you for being there,
unquestioning; and thank you for tolerating me.

Saving the final, biggest acknowledgment until last, I thank my parents, Liz and
Mark, for their unconditional love, support, and understanding; not only during the
past four years, but the past 32 years. Here is, undoubtedly, the most expensive
book you’ll never read. This thesis is dedicated to you both.
# Contents

1 Introduction ........................................ 22
   1.1 Sex .................................................. 22
      1.1.1 Sex determination and the evolution of the mammalian sex chromosomes .................. 23
   1.2 The mammalian Y chromosome ....................... 25
      1.2.1 The Testis Determining Factor: Sry .................. 25
      1.2.2 Transposed and ampliconic genes .................... 26
      1.2.3 X degenerate genes .................................. 26
   1.3 The mammalian X chromosome ......................... 27
      1.3.1 Sexual antagonism drives X chromosome evolution .......... 28
      1.3.2 Pseudoautosomal region (PAR) ....................... 29
   1.4 Dosage compensation .................................. 30
      1.4.1 X upregulation(XUR) .............................. 30
      1.4.2 X chromosome inactivation (XCI) ................... 32
      1.4.3 Escape from X chromosome inactivation ............. 42
   1.5 Aneuploidy .......................................... 43
      1.5.1 Sex chromosome aneuploidies ....................... 43
      1.5.2 Turner syndrome: XO ............................. 46
      1.5.3 XO aneuploid mice .............................. 50
   1.6 X-Y gene pairs with candidate roles in postnatal growth .......... 53
      1.6.1 Kdm5c ........................................... 53
      1.6.2 Kdm5d ........................................... 55
2 Materials and Methods 68

2.1 Mouse work 68
  2.1.1 Mouse strains 68
  2.1.2 Timed matings 70
  2.1.3 Blastocyst isolation 70
  2.1.4 Isolation of post implantation embryos 71
  2.1.5 Weighing of post implantation embryos 71
  2.1.6 Somite counting of post implantation embryos 71
  2.1.7 Generation of CRISPR targeted lines 71
  2.1.8 Weighing postnatal animals 72
  2.1.9 Identification of individuals 73

2.2 Molecular biology 73
  2.2.1 Genomic DNA (gDNA) extraction 73
  2.2.2 RNA extraction 74
  2.2.3 cDNA synthesis 74
  2.2.4 Polymerase Chain Reaction (PCR) 74
  2.2.5 Quantitative real time PCR 75
  2.2.6 TA cloning 76
  2.2.7 RFLP digest 76

2.3 Cytology 77
  2.3.1 Metaphase chromosome spreads for karyotyping 77
  2.3.2 DNA FISH 78
2.4 Protein extraction and western blotting ........................................ 80
  2.4.1 Tissue harvest and protein extraction .................................. 80
  2.4.2 Protein quantification ..................................................... 80
  2.4.3 Western blotting ........................................................... 80
2.5 Cell culture ............................................................................. 81
  2.5.1 Mouse trophoblast stem cells (mTSCs) ................................. 81
  2.5.2 Mouse embryonic stem cells (mESCs) ................................. 83
  2.5.3 Derivation of MEFs ........................................................... 84
2.6 Next generation sequencing ....................................................... 85
  2.6.1 Generation of data to screen CRISPR targeted loci .............. 85
  2.6.2 Data analysis ..................................................................... 86
  2.6.3 RNA-seq ......................................................................... 86
2.7 Statistical analysis .................................................................... 88
  2.7.1 Weight data analysis .......................................................... 88
  2.7.2 Postnatal survival analysis .................................................. 89

3 Examining the role of Xist in the post-implantation growth deficit observed in XPO embryos 90
  3.1 Introduction ........................................................................... 90
  3.2 Results .................................................................................. 92
  3.3 Discussion ............................................................................. 96
    3.3.1 Xist coating may not directly result in imprinted XCI ....... 98
    3.3.2 The X chromosome may harbour imprinted genes that are primarily expressed in the extraembryonic tissue ........... 100
    3.3.3 Mouse trophoblast stem cells can be utilised to identify imprinted genes in vitro ........................................... 102

4 Using in vitro models to identify candidates for X linked imprinted genes in mouse extraembryonic tissue 104
  4.1 Introduction ........................................................................... 104
    4.1.1 Genomic imprinting ......................................................... 104
4.1.2 The placenta .................................................. 106
4.2 Results .............................................................. 110
  4.2.1 Derivation of stem cell lines .......................... 110
  4.2.2 $X^P_O$, $X^M_O$, and XX mTSCs show equivalent differentiation potential \textit{in vitro} ..................... 112
  4.2.3 Derivation of mESC and mEpiLC lines and qRT-PCR verification of identity .......................... 119
  4.2.4 RNA-seq analysis of gene expression in mTSC, mESC and mEpiLC lines .................................. 121
  4.2.5 Identification of candidate imprinted genes by differential gene expression analysis (DGE) ............ 126
  4.2.6 A significant number of genes were differentially expressed between XO and XX mTSCs ..................... 133
4.3 Discussion .......................................................... 137
  4.3.1 No XY mTSC lines were recovered during the derivation process ........................................ 137
  4.3.2 Significantly more genes are differentially expressed between XX and XO cell lines in mTSCs than mESCs or mEpiLCs .......................................................... 138
  4.3.3 Non genetic phenotypic heterogeneity in aneuploid cells .................................................. 140
  4.3.4 Confounding factors in analysis of RNA-seq data from aneuploid samples .............................. 141
  4.3.5 Validation of cell lines for the study of imprinted genes .................................................. 143
  4.3.6 Alternative models for investigating the $X^P_O$ embryonic growth deficit .......................... 145
  4.3.7 Potential use of proteogenomics to further investigate the $X^P_O$ growth deficit phenotype ........ 146

5 Assessing the contribution of X-linked dosage sensitive genes to the postnatal growth deficit in XO female mice 148
  5.1 Introduction ....................................................... 148
5.2 Results ................................................................. 149
  5.2.1 Recapitulation of previous results ......................... 149
  5.2.2 Verification of potential postnatal growth retardation candidates ........................................... 155
  5.2.3 The generation of mutant alleles by CRISPR Cas9 genome engineering ........................................ 158
  5.2.4 Kdm5c .......................................................... 160
  5.2.5 Kdm6a .......................................................... 163
  5.2.6 Ddx3x .......................................................... 166
  5.2.7 Eif2s3x ........................................................ 168
  5.2.8 Haploinsufficiency for Kdm5c, Kdm6a, or Ddx3x does not fully recapitulate the postnatal XO growth deficit ..................................................... 171
5.3 Discussion ......................................................... 176
  5.3.1 Comparison of birth weight data ............................. 176
  5.3.2 XY females at birth and post weaning ........................ 177
  5.3.3 Increasing the XO:XX weight difference ..................... 178
  5.3.4 The postnatal growth phenotype of heterozygous females mutant for Kdm5c, Kdm6a, or Ddx3x ............... 179
  5.3.5 Targeting and re targeting Eif2s3x .......................... 183

6 The contribution of X-Y homologous genes to postnatal survival in the mouse ........................................ 185
  6.1 Introduction ..................................................... 185
  6.2 Results .......................................................... 187
    6.2.1 Kdm6a and Kdm5c are required for postnatal survival in female mice ........................................... 187
    6.2.2 Y linked genes partially compensate for loss of their X linked homologues ................................ 191
    6.2.3 Some Y linked genes can compensate for loss of their X linked homologue independent of the hormonal milieu ... 197
  6.3 Discussion ...................................................... 202
## Contents

6.3.1 The *Ddx3x* mutant allele is a hypomorph, not a knockout . . 202
6.3.2 Conservation of protein coding sequence does not equate to
conservation of gene function . . . . . . . . . . . . . . . . . . 203
6.3.3 Differences in gene regulation likely underlie the apparent
functional differences in X-Y gene pairs . . . . . . . . . . . . . 204

7 Summary 207

Appendices 211

A Cell culture media and solutions 211

B Oligonucleotide sequences 214

C Antibodies and BACs 222

D Colophon 223

Bibliography 224
List of Figures

1.1 Evolution of the mammalian sex chromosomes . . . . . . . . . . . . 24
1.2 The theory of dosage compensation in eutherian mammals . . . . . 31
1.3 X chromosome inactivation in the mouse . . . . . . . . . . . . . . 34
1.4 Cartoon depicting relative positions of X-Y gene pairs on mouse X
    and Y chromosomes. . . . . . . . . . . . . . . . . . . . . . . . . 65
3.1 Xist KO does not rescue the developmental deficit observed in X^{P}O
    embryos at E10.5 . . . . . . . . . . . . . . . . . . . . . . . . . 96
4.1 Mouse placenta development . . . . . . . . . . . . . . . . . . . . 108
4.2 Derivation of mouse trophoblast stem cells (mTSCs) . . . . . . . . 112
4.3 Differentiation of mTSCs . . . . . . . . . . . . . . . . . . . . . . 117
4.4 qRT-PCR analysis of gene expression during mTSC differentiation . 119
4.5 Derivation of mouse embryonic stem cells (mESCs) and mouse epi-
    blast like stem cells (mEpiLCs) . . . . . . . . . . . . . . . . . . . 121
4.6 Initial analyses of RNA-seq on mTSC, mESC and mEpiLC . . . . . 124
4.7 Genotyping confirmation from RNA-seq data in all cell types . . . . 126
4.8 Differential gene expression analyses . . . . . . . . . . . . . . . . 129
4.9 Differential gene expression analysis summary . . . . . . . . . . . 130
4.10 Genes differentially expressed between XO and XX mTSCs . . . . 136
5.1 Recapitulation of postnatal growth retardation phenotype in XO fe-
    male mice . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 153
5.2 Expression analysis of genes on the mouse Y chromosome by RT-PCR 156
List of Figures

5.3  Restriction fragment length polymorphism analysis to confirm escape from XCI in XX female mouse brain and liver .......................... 158
5.4  Demonstrating sgRNA cutting of target genomic DNA ...................... 159
5.5  Generation of mutant Kdm5c alleles by CRISPR ............................. 163
5.6  Generation of mutant Kdm6a allele by CRISPR .............................. 166
5.7  Generation of mutant Ddx3x allele by CRISPR .............................. 167
5.8  Generation of mutant Eif2s3x alleles by CRISPR ............................ 171
5.9  Assessment of postnatal growth in heterozygous mutant females compared to wildtype littermates .................................................. 174
5.10 Model depicting XCI escapee expression in XX, XO, and XX- females 182

6.1  Postnatal survival of XO female mice carrying a single mutated allele, relative to XO littermates ................................................. 190
6.2  Postnatal survival of female mice carrying two mutated alleles, relative to males with one mutated allele and an intact Y chromosome 194
6.3  Postnatal survival of male and female mice carrying a single mutated allele, relative to wildtype littermates ................................. 196
6.4  Postnatal survival of XX and XY male, and XX and XY female mice carrying a single mutated allele, relative to wildtype littermates 201
## List of Tables

1.1 Key features of human sex chromosome aneuploidies .......................... 44
1.2 Extant X-Y gene pairs in mouse and human ................................. 66
3.1 Weight and somite comparison of E10.5 embryos ......................... 97
4.1 RNA-seq mapping percentages .............................................. 122
4.2 X linked genes significantly differentially expressed in X\textsuperscript{M}O vs X\textsuperscript{P}O mTSCs .................................................. 131
4.3 X linked genes significantly differentially expressed in X\textsuperscript{M}O vs X\textsuperscript{P}O mESCs .................................................. 132
4.4 X linked genes significantly differentially expressed in X\textsuperscript{M}O vs X\textsuperscript{P}O mEpiL Cs ............................................... 132
4.5 Gene groups over represented in both DGE analyses comparing X\textsuperscript{P}O with XX mTSCs and X\textsuperscript{M}O with XX mTSCs ....................... 134
4.6 Chromosomal locations over represented in DGE analyses comparing X\textsuperscript{P}O with XX mTSCs and X\textsuperscript{M}O with XX mTSCs ....................... 134
5.1 Birth weight data compared by mean weighted difference (MWD) .... 154
5.2 Means and mean weighted differences (MWD) for phasic weight gain between genotype comparisons shown in Figure 5.1 ............... 154
5.3 Guide efficacy as tested in XY mESCs ....................................... 160
5.4 Comparing numbers of wildtype and \textit{Kdm6a} hemizygous males recorded across five litters ............................................. 164
5.5 Targeting \textit{Eif2s3x} in mouse embryos .................................... 170
5.6 Birth weight compared by mean weighted difference (MWD) for
XX vs. XX': ................................................................. 175
5.7 Means and mean weighted differences (MWD) for phasic weight
gain between genotype comparisons shown in Figure 5.8 ............ 175
5.8 Birth weight compared by mean weighted difference (MWD) for
XX vs. XY ................................................................. 178
6.1 Postnatal survival for XO vs X'O females. .......................... 188
6.2 Postnatal survival for X'X' vs X'Y animals ......................... 191
6.3 Postnatal survival for pooled XY vs X'Y males .................... 192
6.4 Postnatal survival for pooled XY^{Tdy} vs X'Y^{Tdy} females ........ 198
6.5 X-Y pair copy number summary for all comparisons ............ 199
6.6 Phenotype: amino acid similarity correlation for X linked mutants 206
A.1 Components of embryo culture media used in this thesis ....... 211
A.2 Components of cell culture media used in this thesis .......... 212
A.3 Components of solutions used in this thesis ...................... 213
C.1 Antibodies used in this thesis ..................................... 222
C.2 BAC clones used in this thesis .................................... 222
Nomenclature

3C  Chromosome conformation capture
ALL  Acute lymphocytic leukaemia
ARID  AT rich interacting domain
ASF  Advanced sequencing facility
BAC  Bacterial artificial chromosome
BCA  Bicinchoninic acid
bHLH  Basic helix loop helix
BSA  Bovine serum albumin
ccRCC  Clear cell renal cell carcinoma
cDNA  Complimentary DNA
ChIP-seq  Chromatin immunoprecipitation sequencing
CML  Chronic myeloid leukaemia
CpG  Cytosine phosphate guanine
CRISPR  Clustered regularly interspaced short palindromic repeats
CTA  Cancer testis antigen
DAPI  4’,6-diamidino-2-phenylindole
DE  Differentially expressed
DGE  Differential gene expression
DMSO  Dimethyl sulphoxide
DTT  Dithiothreitol
dUTP  Deyoxyuridine triphosphate
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECL</td>
<td>Enhanced chemiluminescence</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGFP</td>
<td>Enhanced green fluorescent protein</td>
</tr>
<tr>
<td>EPC</td>
<td>Ectoplacental cone</td>
</tr>
<tr>
<td>EtOH</td>
<td>Ethanol</td>
</tr>
<tr>
<td>ExE</td>
<td>Extraembryonic ectoderm</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence activated cell sorting</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>FHM</td>
<td>Follicle holding medium</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescence <em>in situ</em> hybridisation</td>
</tr>
<tr>
<td>gDMR</td>
<td>Germline differentially methylated region</td>
</tr>
<tr>
<td>gDNA</td>
<td>Genomic DNA</td>
</tr>
<tr>
<td>GO</td>
<td>Gene ontology</td>
</tr>
<tr>
<td>Gy</td>
<td>Gray</td>
</tr>
<tr>
<td>HCC</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td>HCV</td>
<td>Hepatitis C virus</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HLA</td>
<td>Human leukocyte antigen</td>
</tr>
<tr>
<td>ICM</td>
<td>Inner cell mass</td>
</tr>
<tr>
<td>IF</td>
<td>Immunofluorescence</td>
</tr>
<tr>
<td>iPSC</td>
<td>Induced pluripotent stem cell</td>
</tr>
<tr>
<td>IUGR</td>
<td>Intrauterine growth restriction</td>
</tr>
<tr>
<td>iXCI</td>
<td>Imprinted X chromosome inactivation</td>
</tr>
<tr>
<td>KEGG</td>
<td>Kyoto encyclopaedia of genes and genomes</td>
</tr>
<tr>
<td>KS</td>
<td>Kabuki syndrome</td>
</tr>
<tr>
<td>KSOM</td>
<td>Potassium supplemented simplex optimised medium</td>
</tr>
<tr>
<td>LB</td>
<td>Luria broth</td>
</tr>
<tr>
<td>LIF</td>
<td>Leukaemia inhibitory factor</td>
</tr>
<tr>
<td>LINE</td>
<td>Long interspersed nuclear element</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>IncRNA</td>
<td>Long non coding RNA</td>
</tr>
<tr>
<td>MA</td>
<td>Methanol:acetic acid</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen activated protein kinase</td>
</tr>
<tr>
<td>MDS</td>
<td>Multidimensional scaling</td>
</tr>
<tr>
<td>MEF</td>
<td>Mouse embryonic fibroblast</td>
</tr>
<tr>
<td>Meg</td>
<td>Maternally expressed gene</td>
</tr>
<tr>
<td>mEpiLC</td>
<td>Mouse epiblast like cell</td>
</tr>
<tr>
<td>mESC</td>
<td>Mouse embryonic stem cell</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MSCI</td>
<td>Meiotic sex chromosome inactivation</td>
</tr>
<tr>
<td>mTSC</td>
<td>Mouse trophoblast stem cell</td>
</tr>
<tr>
<td>MWD</td>
<td>Mean weighted difference</td>
</tr>
<tr>
<td>mXEN</td>
<td>Mouse extraembryonic endoderm</td>
</tr>
<tr>
<td>MYA</td>
<td>Million years ago</td>
</tr>
<tr>
<td>ORA</td>
<td>Over representation analysis</td>
</tr>
<tr>
<td>PAM</td>
<td>Protospacer adjacent motif</td>
</tr>
<tr>
<td>PAR</td>
<td>Pseudoautosomal region</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PE</td>
<td>Primitive endoderm</td>
</tr>
<tr>
<td>Peg</td>
<td>Paternally expressed gene</td>
</tr>
<tr>
<td>PGC</td>
<td>Primordial germ cell</td>
</tr>
<tr>
<td>PRC</td>
<td>Polycomb repressor complex</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>Quantitative real time polymerase chain reaction</td>
</tr>
<tr>
<td>RCC</td>
<td>Renal cell carcinoma</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction fragment length polymorphism</td>
</tr>
<tr>
<td>RIN</td>
<td>RNA integrity number</td>
</tr>
<tr>
<td>RIPA</td>
<td>Radioimmunoprecipitation assay</td>
</tr>
</tbody>
</table>
RNA-seq  RNA sequencing
RNAi    RNA interference
RT-PCR  Reverse transcription polymerase chain reaction
rXCI    Random X chromosome inactivation
s-TGC   Sinusoidal trophoblast giant cell
SDS PAGE Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sgRNA   Small guide RNA
siRNA   Small interfering RNA
SNP     Single nucleotide polymorphism
SPRI    Solid phase reversible immobilisation
SSC     Saline sodium citrate
SWI/SNF SWItch/sucrose non-fermentable
Syn     Syncytiotrophoblast
t-SNE   T-distributed stochastic neighbour embedding
TAD     Topologically associating domain
TBS-T   Tris-buffered saline Tween
TDF     Testis determining factor
TE      Trophectoderm
TGC     Trophoblast giant cell
TPM     Transcripts per million
TPR     Tetratricopeptide repeat
TS      Turner syndrome
UHC     Unsupervised hierarchical clustering
X^M     Maternally inherited X chromosome
X^P     Paternally inherited X chromosome
XAR     X added region
XCI     X chromosome inactivation
XLID    X linked intellectual disability
XUR     X upregulation
<table>
<thead>
<tr>
<th>YAR</th>
<th>Y added region</th>
</tr>
</thead>
</table>

*List of Tables*
Chapter 1

Introduction

Fundamental differences exist between males and females, encompassing areas including anatomy, physiology, and behaviour. Such differences undoubtedly play a part in the well documented, yet poorly understood, disparity in disease susceptibility between the sexes. Although traditionally attributed to differences in sex hormone levels, recent work has begun to shed more light on the contribution of genetics - and in particular the sex chromosomes - to this sexual dimorphism. This chapter begins with an overview of sex chromosome evolution, before focusing on topics most pertinent to this thesis, including dosage compensation, aneuploidy, and growth.

1.1 Sex

Sexual reproduction - meiotic recombination and the subsequent production and fusion of gametes - is widespread throughout the eukaryotic kingdom. This encompasses all mammals and birds, most reptiles, fish and amphibia, and about 99.9% of flowering plant species (Charlesworth, 2006). Whilst many plant species combine male and female sex organs within individuals as hermaphrodites, most animals have, through the evolution of anisogamy, segregated these organs into different individuals, thus resulting in two sexes (Bachtrog et al., 2014). During development, these sexes must be accurately differentiated. Evolution has resulted in multiple
solutions to this sex determination problem, which can be broadly grouped into environmental and genotypic. Environmental sex determination is exemplified by external stimuli - such as social factors or temperature - determining the sex of the offspring. Such a mechanism has been shown to be favoured over a genetic mechanism when specific environments benefit one sex. Conversely, genetic sex determination is found when the environment is either unpredictable or not varied enough, and is mediated by elements ranging in size from genes through chromosomes to entire genomes (Bachtrog et al., 2014). Mammalian sex determination is a primary example of this mechanism.

1.1.1 Sex determination and the evolution of the mammalian sex chromosomes

Hermann Muller posited the first theory of sex chromosome evolution following observations that at least one gene could be linked to each of the *Drosophila* autosomes but not the Y chromosome (Muller, 1914). He hypothesised that the X and Y chromosomes evolved from an ordinary pair of autosomes. The Y chromosome subsequently accumulated deleterious mutations that, due to a lack of recombination, could not be repaired, leading to the dearth of reported Y linked genes. This was consistent with observations from human genetics: whilst multiple traits were mapped to the X chromosome, everything attributed to the Y was publicly discredited (Stern, 1957). Following the development of karyotyping techniques in the 1950s, a role for the Y chromosome in sex determination was firmly established. It was shown that individuals with only a single X chromosome are phenotypically female (XO; Turner syndrome(TS)), whereas those with an XXY karyotype are male (Klinefelter syndrome, Ford et al., 1959a; Jacobs and Strong, 1959). Ohno built on this work, and that of Muller, in arguing that a spontaneous mutation producing a proto sex determining gene (see section 1.2.1) had arisen on one of the ancestral vertebrate autosomes to create a pair of heteromorphic sex chromosomes (Ohno, 1967). He went on to suggest that a pericentric inversion containing this proto sex determining gene would suppress crossing over between the nascent sex chromo-
1.1. Sex

The resulting pair of heteromorphic sex chromosomes would in turn cause isolation of the sex specific chromosome, thereby facilitating divergence through either innovation or degeneration.

**Figure 1.1:** Evolution of the mammalian sex chromosomes from autosomes. A sex determining locus (white) was acquired around 180-200 million years ago (mya). Sexually antagonistic alleles (purple) then evolved at nearby loci, selecting for recombination suppression between the X and Y chromosomes, and resulting in the first X chromosome evolutionary stratum (I), and the so called male specific Y. Following the end of recombination in a given region, genes without sex specific benefits often become pseudogenes (dark blue). The appearance and selection for sexually antagonistic alleles, followed by further recombination suppression, then becomes iterative and, as the non recombining region expands, further strata appear (II-V). Between the evolution of strata I and II/III, a fusion event took place between the sex chromosomes and a pair of autosomes, resulting in the expansion of both sex chromosomes in the X added region (XAR) and Y added region (YAR). Over evolutionary time, the lack of recombination leads to the accrual of repetitive DNA sequences and a short term increase in the size of the chromosomes; though this eventually results in large deletions and explains the relatively diminutive size of the Y chromosome in most mammals.

Innovation in sequencing technology provided the first sequences of the human sex chromosomes. These sequences added weight to the theory of a common evolutionary origin, which was based on nucleotide divergence patterns (Lahn and Page, 1999). By comparing nucleotide divergence between Y linked and X linked gene copies, Lahn and Page highlighted a strong correlation with the position of the X linked gene. X-Y gene pairs were shown to be present in at least four groups of increasing divergence, from the short to the long arm of the X chromosome. Thus, the authors posited the existence of distinct evolutionary strata (Lahn and Page,
1.2. The mammalian Y chromosome

As previously highlighted, the Y mammalian chromosome contains the sex determining gene, and genes that have survived the process of Y degeneration (X degenerate, Bellott and Page, 2009) in the strata identified by X-Y divergence. Additionally, the Y chromosome also harbours genes that have been transposed from the X chromosome later in history, and genes that have become ampliconic (Skaletsky et al., 2003; Bellott et al., 2014; Cortez et al., 2014).

1.2.1 The Testis Determining Factor: Sry

Following on from the confirmation that presence of a Y chromosome genetically determines maleness in humans (Ford et al., 1959a; Jacobs and Strong, 1959), the identity of a specific locus was sought. This search for a testis determining factor (TDF) utilised sex reversed patients, i.e. XX males and XY females, hoping to isolate the smallest fragment of Y chromosome required for a male phenotype. Sex determining region on the Y chromosome, or SRY, was identified first in human and mouse, then shown to be present in a "Noah’s Ark" (Gubbay et al., 1990; Sinclair et al., 1990) of mammals. This conservation extends across eutherians, and the sequence is present in therians and prototherians (Cortez et al., 2014), though whether it retains a role in sex determination in these clades remains an open question.

Sry is expressed for a brief period at E11.5 in XY mouse embryonic gonads, preceding the morphological differentiation of the testes (Gubbay et al., 1990). A 14kb transgene containing the Sry open reading frame (ORF), along with 8kb of upstream
and 5 kb of downstream sequence, has been shown sufficient to induce testis formation (though not spermatogenesis) in XX mice (Koopman et al., 1991).

### 1.2.2 Transposed and ampliconic genes

The first resolved examples of ampliconic genes were found on the human Y chromosome, where there are nine multicopy gene families. Two of these families originated on the pre-degenerate Y, and seven have been transposed from autosomes (Skaletsky et al., 2003). More recently, sequencing of the mouse Y chromosome has identified massively ampliconic sequence comprising >95% of the euchromatin and containing three protein coding gene families with significant homology to autosomal genes (Soh et al., 2014). Ampliconic genes on the Y chromosome were originally thought to be important in intrachromosomal recombination and gene conversion to slow degeneration of the Y (Rozen et al., 2003). It was later hypothesised that the high copy number inherent to these ampliconic genes might also facilitate high levels of protein expression, specifically in the testis (Mueller et al., 2013). An alternative explanation cites the presence of the same gene families amplified on the mouse X chromosome (Mueller et al., 2013), and suggests that genomic conflict between the two sex chromosomes during spermatogenesis to increase transmission is the driving force behind amplification (Bachtrog, 2014). The latter hypothesis serves to contradict the hypothesis often popularised in the press that the Y chromosome will lose all protein coding genes within the next 10 million years (Aitken and Graves, 2002).

### 1.2.3 X degenerate genes

Following the publication of the first complete sequence of the human Y chromosome in 2003 (Skaletsky et al., 2003), it became clear that only around 3% of ancestral genes survived 200 million years of sex chromosome evolution on the human Y, in stark contrast to 98% on the X chromosome (Mueller et al., 2013). This decay has been observed in the sex specific chromosomes across therian mammals.
It is especially true of the mouse Y chromosome, which has fewer ancestral genes than most other mammals studied. Given this extensive loss of genes from mammalian Y chromosomes, are the survivors simply lucky, or do they carry biological importance?

The evolutionary significance of the X degenerate genes lies in their convergent survival and longevity across not only primates (Hughes et al., 2005, 2010, 2012), but also eutherians and the wider therian subclass (Bellott et al., 2014; Cortez et al., 2014). This convergent survival is exemplified by \textit{KDM5D} and \textit{UBE1Y}. The stratum containing both genes evolved independently in eutherian and metatherian lineages, containing a combined 184 shared genes. The former lineage has retained three Y linked genes, and in the latter nine survived on the Y chromosome. Remarkably, \textit{KDM5D} and \textit{UBE1Y} are still present in both lineages, an event unlikely to happen under the assumption of random genetic decay (Bellott et al., 2014).

**1.3 The mammalian X chromosome**

Ohno (1967) hypothesised that dosage compensation keeps the X chromosome constant in terms of gene content. This prediction was initially borne out by sequencing data, as high conservation was noted within eutherians at the gene level (Carver and Stubbs, 1997; Ross et al., 2005), and fewer interchromosomal translocations were uncovered than on the mammalian autosomes (Carver and Stubbs, 1997). Changes have, however, undoubtedly taken place. The rate of global X chromosome expression adjusted quickly during early sex chromosome evolution, perhaps in order to enact faster functional adaptation (Brawand et al., 2011). Moreover, events such as translocations from the pseudoautosomal region (PAR, see section 1.3.2) to the autosomes in mouse have served to reshape the gene content of the X chromosome towards enrichment for specific physiological functions (Palmer et al., 1997; Carver and Stubbs, 1997; Rugarli et al., 1995).
1.3.1 Sexual antagonism drives X chromosome evolution

The X chromosome spends two thirds of its time in females, and one third in males (Fisher, 1931; Rice, 1984). Therefore if a mutation arises with significant fitness effect on heterozygous carriers that benefits females and is deleterious in males, it will be subject to positive selection more often than negative selection, and will increase in frequency. A recessive, or partially recessive, mutation that is beneficial to males and deleterious to females can spread by virtue of positive selection in male carriers, and masking in female carriers (Vicoso and Charlesworth, 2006). These theoretical examples of sexual antagonism explain the accumulation of either male or female biased genes on the X chromosome relative to the autosomes, via processes such as retrotransposition and gene duplication.

Reverse transcription and integration via a mature mRNA intermediate - retrotransposition - has generated and recruited a disproportionately high number of functional genes from and to the X chromosome relative to the autosomes (Emerson et al., 2004). A significant number of X to autosome transposition events involved genes with testis biased expression (Emerson et al., 2004). This bias has been explained by both mutational bias and natural selection for expression in germ cells, possibly resulting from need to maintain gene expression during meiotic sex chromosome inactivation (MSCI, Turner, 2007).

A further driver of enrichment for testis biased genes has been noted in the form of gene duplication events on the X chromosome. Duplication of existing genes has created vast swathes of ampliconic sequence which, in human, often take the form of cancer testis antigens (CTAs, Warburton et al., 2004). CTA genes exist in families that have expanded independently in both primate and rodent lineages, in a somewhat similar manner to the expansion of ampliconic sequence on the mammalian Y chromosome (Chomez et al., 2001; Birtle et al., 2005; Ross et al., 2005). They are usually expressed only in male germ cells, but also aberrantly in certain cancers (Simpson et al., 2005).
In addition to enrichment for genes expressed in testis, the human and mouse X chromosomes show over representation of genes expressed in brain; and a correlation has been reported between patterns of X linked gene expression in these two tissues (Guo et al., 2003, 2005). It has been suggested that this relates to females selecting more intelligent males (Zechner et al., 2001). A more likely explanation puts forward that genes involved in cognition and fertility are especially important in the evolution of mammals, and are therefore more likely to experience positive selection (Vicoso and Charlesworth, 2006). Finally, the X chromosome also shows specific enrichment for genes expressed in the ovary (Khil et al., 2004), and the placenta (Hemberger, 2002).

1.3.2 Pseudoautosomal region (PAR)

The PAR is the only region of continuous sequence homology between the two heteromorphic mammalian sex chromosomes, where evolutionary divergence has not taken hold. This conservation enables pairing, synopsis, and recombination during meiosis (Koller and Darlington, 1934; Moses et al., 1975; Solari, 1974), as also takes place on the autosomes, hence the derivation of the name PAR (Burgoyne, 1982). The state of diploidy with reference to loci located within the PAR also means that there is no requirement for dosage compensation by XCI (Ellis and Goodfellow, 1989; Ross et al., 2005). The PAR is, therefore, also subject to the effects of aneuploidy (see section 1.5).

As all eutherian pseudoautosomal genes are autosomal in metatherians, it is thought that the PAR evolved following a large autosome to sex chromosome translocation between 80 and 130 MYA (Park et al., 2005; Waters et al., 2007). PAR evolution has since proceeded at significantly different rates in mammalian species. Humans are the only known eutherian species to have two PARs, with 25 genes annotated in PAR1 and four genes in PAR2 (Hughes et al., 2012). In the mouse PAR, which shares very little homology with either human PAR, there are five known protein coding genes (Mueller et al., 2013; Bellott et al., 2014).
1.4 Dosage compensation

Both Muller (1964) and Ohno (1967) predicted that the evolution of a heterogametic sex would require a system of dosage compensation to avoid the "...great peril..." of hemizygous gene expression: effectively a form of natural aneuploidy, the negative impact of disrupted chromosome balance has long been noted (see 1.5). The appearance across multiple species of efficient and stable mechanisms to relieve this situation attests to its significance (reviewed by Disteche, 2012). Two distinct aims of dosage compensation have been identified: i) to balance sex chromosome gene expression with that of the autosomes; and ii) to balance sex linked gene expression between the homogametic and heterogametic sexes. Ohno hypothesised that this could be achieved by two fold transcriptional upregulation of the X chromosome in males, X chromosome upregulation (XUR). However, this would leave females expressing X linked genes at twice the level of autosomal genes. Therefore a further step is the inactivation of one X chromosome in the homogametic sex - XCI (Figure 1.2).

1.4.1 X upregulation(XUR)

Initial evidence for XUR in mouse was limited to one gene, Clcn4-2 (Adler et al., 1997), before the appearance of more data produced by microarray in the mid 2000s (Nguyen and Disteche, 2006; Gupta et al., 2006). Assays across several tissues in multiple species supported increased expression of X linked genes relative to autosomal genes, thus resulting in an X:autosome expression ratio of 0.8-1.0. RNA sequencing (RNA-seq) analyses then appeared to contradict the previous data, with X:autosome expression ratios of 0.5 (Xiong et al., 2010), and subsequent work has seen the balance of evidence constantly shift. What has become clear is that the parameters surrounding the analysis, particularly normalisation, play a significant part in the outcome. Even expression levels between different autosomes can vary substantially (Castagné et al., 2011). A number of ways around this problem have been suggested. X:autosome expression ratios were shown to increase following...
1.4. Dosage compensation

Figure 1.2: The theory of dosage compensation in eutherian mammals. (i) To balance gene dosage between the autosomes (A) and sex chromosomes, eutherian mammals upregulate gene expression from one X chromosome (X in males and X\textsuperscript{a} in females). This is depicted in the figure as increased size of the X chromosome relative to the autosomes. X upregulation in females leads to excess X chromosome gene expression relative to males. To balance X chromosome gene expression between males and females, one X chromosome is transcriptionally silenced in females by X chromosome inactivation (X\textsuperscript{i}). The inactivated chromosome is depicted as a white chromosome. (ii) Dosage compensation in numbers. The gene expression from a single autosome is arbitrarily represented as 1.

the exclusion of low/non expressed testis specific genes (Deng et al., 2011). Furthermore, when certain idiosyncrasies of the X chromosome are accounted for - such as treating multi copy gene families differently to single copy genes - evidence for XUR in multiple tissues mounts (Sangrithi et al., 2017; Jue et al., 2013). Alternative approaches have also been trialled. Two recent studies compared expression of conserved ancestral genes on the mammalian X chromosome to the expression of autosomal orthologues in species where the sex chromosomes diverged before the origin of the X, i.e. chicken (Julien et al., 2012; Lin et al., 2012a). Although inter species comparisons are complicated by different levels of reference sequence annotation, neither of these studies found evidence for chromosome wide XUR. However, Lin and colleagues did find evidence to support a refinement to Ohno’s hypothesis (Lin et al., 2012a). Along with a second study (Pessia et al., 2012), each provides transcriptomic and proteomic data showing upregulation of X linked
1.4. Dosage compensation

genes encoding the members of multi protein complexes. Additionally, comparing X linked genes to autosomal orthologues in divergent species has revealed a down-regulation of autosomal genes whose protein products interact with those from X linked genes (Julien et al., 2012). Taken together, these data suggest that XUR, and possibly autosomal downregulation, are important for specific dosage sensitive genes where the stoichiometry of protein products involved in macromolecular complexes affects function, as reviewed by Veitia (Veitia and Birchler, 2010). Whether or not this equalising of X:autosome gene products was sufficient to drive the evolution of equalisation between XX females and XY males through XCI remains to be shown.

1.4.2 X chromosome inactivation (XCI)

XCI balances sex chromosome dosage between XX females and XY males in mammals via transcriptional inactivation of one X chromosome in the homogametic sex. The first cytological evidence was presented by Barr and Bertram (1949), whereby they observed nucleolar satellites in female but not male cat motor neurons. Ohno & Hauschka suggested that these satellites were heteropyknotic X chromosomes (Ohno and Hauschka, 1960), and Lyon was able to show that such a heteropyknotic X chromosome could be inherited from either parent, and was transcriptionally silent (Lyon, 1961). She went on to posit that the random inactivation of one of two different X linked coat colour alleles in each cell would result in the tortoishell coat of the cat, and thus began a field of research that has matured to become a fundamental paradigm in epigenetics.

XCI is mediated by the long non-coding RNA (IncRNA) Xist in eutherian mammals (Brown et al., 1991; Borsani et al., 1991; Brockdorff et al., 1991), and Rxx in metatherians (Grant et al., 2012). Following transcriptional upregulation of Xist from the future inactive X chromosome, a series of events takes place to ensure that the subsequent transcriptional silencing is locked in and stably transmitted to future cell progeny. Xist RNA coats the chromosome in cis, leading to the creation of an
inactive domain containing silenced genes, and excluding RNA polymerase II and transcription factors (Okamoto et al., 2005). Following the removal of active histone marks, PRC 1 and 2 are recruited (Silva et al., 2003; Schoeftner et al., 2006; Tavares et al., 2012), possibly via the repA domain in \textit{Xist} (Wutz et al., 2002; Zhao et al., 2008), to add repressive marks such as H3K27me3 to histone tails (Plath et al., 2003; Okamoto et al., 2004). Finally, replication timing of the entire chromosome during mitosis changes relative to the rest of the nucleus (Takagi et al., 1982; Sugawara et al., 1983), and methylation of CpG islands situated at the 5' end of genes on the inactive X chromosome functions to stably maintain the inactive state (Norris et al., 1991; Pfeifer et al., 1990a,b).

In the mouse, \textit{Xist} is first expressed from the paternally inherited X chromosome (X\textsubscript{P}) at approximately the four cell stage (Okamoto et al., 2004, 2005). This imprinted XCI (iXCI) is maintained across the whole embryo until the first lineage specifications take place. It continues in the trophectoderm (Takagi and Sasaki, 1975; Mak et al., 2004), which goes on to contribute to the placenta, and in the primitive endoderm (PE) (Takagi and Sasaki, 1975; West et al., 1977). In the epiblast of the inner cell mass (ICM), \textit{Xist} is downregulated. This downregulation results in a transient two active X state, followed by \textit{Xist} mediated inactivation of either X\textsuperscript{M} or X\textsuperscript{P} (Takagi et al., 1982; Rastan, 1982): random XCI (rXCI). Both of these are now described in more detail and summarised in Figure 1.3.

1.4.2.1 Imprinted XCI

Imprinted inactivation, i.e. inactivation in a parent of origin dependent manner, of the X\textsuperscript{P} is found in all tissues of metatherian mammals (Sharman, 1971; Cooper et al., 1971; Richardson et al., 1971), and the extraembryonic tissue of some eutherians, i.e. mice (Takagi and Sasaki, 1975).

iXCI is necessary for embryonic survival in mouse. X\textsuperscript{M}X\textsuperscript{P}\textit{Xist}- female embryos show transcriptional activity from both X chromosomes in the early embryo (Borensztein et al., 2017) and in trophoblast, become severely growth retarded from E6.5,
1.4. Dosage compensation

Figure 1.3: X chromosome inactivation in the mouse. From top, XCI is initiated in the XX female embryo at the 4 cell stage by expression of \( Xist \) (in red) from the paternally inherited X chromosome (\( X^p \), purple). \( Xist \) RNA spreads to coat the \( X^p \) by the 8-16 cell stage which, with subsequent modifications, establishes XCI. The \( X^p \) remains inactivated until lineage specification takes place, concurrent with blastocyst formation. In the part of the inner cell mass (ICM) that goes on to form the embryo proper, the epiblast, \( Xist \) is downregulated, and this is followed by the process of random XCI, which is passed through to the adult mouse. In the primitive endoderm (PE) and the trophectoderm (TE), which contribute largely to the extraembryonic tissue, XCI is maintained on the \( X^p \). Each of these lineages gives rise to a population of stem cells that can be used as \textit{in vitro} models with different XCI states: mouse embryonic stem cells (mESCs) represent the preimplantation ICM, and have two active X chromosomes; and mouse trophoblast stem cells (mTSCs) and extraembryonic endoderm stem cells (mXENs) are derived from the TE and and PE respectively, and both show imprinted XCI of the \( X^p \). mESCs can be used to derive cells representative of the post implantation epiblast, epiblast like cells (mEpiLCs), which have randomly inactivated one X chromosome. This process of differentiation can be used to model rXCI \textit{in vitro}.
and die before E12.5 (Marahrens et al., 1997). This phenotype is similar to that shown by PRC protein EED deficient embryos (Wang et al., 2001a), implicating a role for histone modifications in iXCI in the extraembryonic tissue. Extraembryonic tissue is known to show reduced methylation (Manes and Menzel, 1981; Rossant et al., 1986). This has been correlated with incomplete silencing of the inactive X in vivo, as assayed via either endogenous (Patrat et al., 2009) or transgene expression (Hadjantonakis et al., 2001). Such a state has also been observed in vitro in trophoblast stem cells (Dubois et al., 2014), in addition to a transient two active X chromosome state (Prudhomme et al., 2015). These instances of plasticity could result from the coexistence of both heterochromatic and euchromatic marks, as observed in trophoblast giant cells (Corbel et al., 2013); and may serve to compensate for occurrences of perturbation of X linked monoallelic gene expression (Prudhomme et al., 2015). Other work has shown the presence of low level methylation of CpG islands at promoter regions in trophoblast tissues when compared to epiblast (Senner et al., 2012). When an enzyme responsible for the maintenance of this mark is removed (Dnmt1), reactivation of a reporter transgene is detected in the embryo but not in the extraembryonic tissue (Sado et al., 2000). Interpreted with evidence showing that neither Dnmt1, Dnmt3a, or Dnmt3b are required for survival and proliferation of mTSCs (Sakaue et al., 2010), it can be concluded that DNA methylation is not the primary epigenetic mark "locking in" the inactive state in extraembryonic iXCI. Much work remains still to be done on the role of histone modifications in this context.

**Mechanism of iXCI**

Although it has previously been thought that the X<sup>p</sup> may be pre-inactivated in the preimplantation embryo as a result of MSCI (Huynh and Lee, 2003), subsequent work has shown active X<sup>p</sup> transcription at the two cell stage (Okamoto et al., 2005). Multiple alternative mechanisms have been proposed to explain the data, a number of which are now reviewed.
1.4. Dosage compensation

A priori, an imprint could protect the $X^M$ from silencing, or target the $X^P$ to be inactivated. Early evidence from experiments utilising parthenogenetic embryos ($X^M X^M$) strongly supports the former hypothesis, as development is stalled in the early postimplantation stages following the failed growth of extraembryonic tissue (Shao and Takagi, 1990; Goto and Takagi, 1998, 2000). Furthermore, it seems that such an imprint is imposed during oogenesis, as combining maternal genomes from non-growing (ng) and fully-grown (fg) oocytes always resulted in inactivation of the ng derived genome in extraembryonic tissue. This observation is consistent with a mark imposed on the fg oocyte during maturation (Tada et al., 2000), though unlike classical genomic imprinting, this is independent of DNA methylation (McDonald et al., 1998; Liu et al., 2010). Recent work has implicated the H3K9me3 histone mark in $X^M$-$Xist$ silencing, showing that reduction of H3K9me3 levels via over-expression of its specific demethylase $Kdm4b$, in addition to a histone deacetylase inhibitor trichostatin, results in detectable $X^M$-$Xist$ de-repression (Fukuda et al., 2014). However, no difference in H3K9me3 levels was found when comparing ng and fg oocytes at the $X^M$-$Xist$ locus (Fukuda et al., 2015). Most recently, the same authors showed that both transient histone alterations and chromatin decondensation are able to partially rescue the embryonic lethality that results from $X^M X^P Xist$- (Fukuda et al., 2015, 2016).

Also utilising histone alterations, Federici and colleagues reported partial rescue of $X^M X^P Xist$- female embryonic lethality (Federici et al., 2016). Using round spermatid injection, they were able to achieve fertilisation without the protamine associated condensation and decondensation of the paternal genome. Taken together, these results implicate a role for H3K9me3 in iXCI, though it seems likely this is within a wider context of genome organisation in the regulation of gene expression.

Namekawa and colleagues suggest that iXCI in mouse can be separated into two steps, with silencing of repetitive sequences preceding that of genic sequences (Namekawa et al., 2010). Using FISH and immunofluorescence (IF), they identified
1.4. Dosage compensation

a Cot1-, Pol-II-, Pol-III- silent compartment within which the repeat sequences of the X<sup>p</sup> are localised at the two cell stage. The formation of this silent compartment is Xist-independent. In contrast, the inactivation of genic regions takes place a number of cell divisions later (reasonably consistent with other data: Okamoto et al., 2004, 2005, 2011), and this requires Xist expression. They favour a model whereby repeats arrive in the embryo pre-inactivated following MSCI. These repeats could comprise an imprint on the X<sup>p</sup> predisposing to inactivation, independent of any X<sup>M</sup> imprint laid down in the oocyte.

Experiments utilising Xist transgenes suggest a role for hemizygosity in iXCI. A 200-kb Xist locus transgene present in a multicopy array on an autosome is sufficient for ectopic Xist expression and chromatin silencing in the hemizygous state, but not the homozygous state (Sun et al., 2015). It is possible that chromatin modifications (like in MSCI) resulting from the unpaired state of the Xist transgenes are directly responsible for Xist imprinting in next generation (Sun et al., 2015). However, this is in contrast to previous work (Okamoto et al., 2005). Using a single copy 460-kb transgene, no difference was observed between homozygous and hemizygous males in terms of ectopic Xist expression. Moreover, there was also no evidence of the Xist transgene associating with the sex body, as would be expected if passage through the XY body is required for iXCI (Okamoto et al., 2005). Payer suggests that the differences observed may be due to the multicopy nature of the transgene in the former study, which might more completely reflect the endogenous situation (Payer, 2016). This has previously been reported when using Xist transgenes to recapitulate rXCI, whereby multicopy transgenes succeed whilst single copies failed (Heard et al., 1996, 1999b; Lee et al., 1999).

1.4.2.2 Random XCI

Early biochemical and cytogenetic evidence provided the first link between cellular differentiation and XCI in vivo (Rastan, 1982; Monk and Harper, 1979). It showed that both X chromosomes are active in the peri implantation mouse embryo until
rXCI takes place in the epiblast around E4.5. This event is re captured in the most well used model of XCI in vitro, mouse embryonic stem cells (mESCs). XX female mESCs have two active X chromosomes, and undergo XCI upon differentiation (Martin et al., 1978; Takagi and Martin, 1984), thereby facilitating investigation of the mechanisms underlying rXCI.

**Regulation of Xist expression**

XCI is controlled by the X inactivation centre (Xic), a region of around 1Mb shown to be both necessary and sufficient to trigger XCI when present in at least two copies (Heard et al., 1996; Lee et al., 1996; Heard et al., 1999a). Within the Xic locus are situated a number of genes - protein coding and non coding - that are involved in the process, many of which have been characterised using the mESC model.

*Tsix* is a negative regulator of *Xist*: it is a non-coding RNA that overlaps with and is transcribed antisense to *Xist* (Lee et al., 1999). Current evidence suggests that the act of *Tsix* transcription, as opposed to the RNA itself, represses *Xist* expression (Stavropoulos et al., 2001; Shibata and Lee, 2004; Sado et al., 2006). Heterozygous deletion of *Tsix* results in skewed XCI of the mutated chromosome in XX female cells, but deletion does not cause inactivation in XY male cells. This suggests that *Tsix* is involved in the choice of which chromosome to inactivate, but not the counting that ensures one X chromosome is left active (Lee and Lu, 1999; Lee, 2002; Stavropoulos et al., 2001). Mechanistically, it has been proposed that pairing observed between the *Tsix* and *Xpr* loci might mediate this choice process (Xu et al., 2006, 2007; Bacher et al., 2006). However, recent work using heterokaryons has provided evidence that pairing is not essential for XCI initiation, as XCI can take place in different nuclei within the same cytoplasm via *trans* acting factors (Barakat et al., 2014).

*Tsix* has also been implicated in iXCI. It is downregulated in the postimplantation embryo, but expression from the X_M persists in the extraembryonic tissue. Disruption of this allele results in ectopic *Xist* expression and early embryonic lethality.
following \textit{Xist} mediated inactivation of the single X in males and both Xs in females (Sado et al., 2001). Recent work has cast doubt on this interpretation, suggesting \textit{Tsix} is not required for initiation of iXCI, but instead is involved in the differentiation of mouse trophoblast stem cells (mTSCs) towards a giant cell fate (Maclary et al., 2014). A further role for \textit{Tsix} has been highlighted in mouse epiblast stem cells, where it is required to silence \textit{Xist} on the active X chromosome (Xa) \textit{in vivo} and \textit{in vitro} as they differentiate (Gayen et al., 2015).

Loci upstream of \textit{Xist} have also been proposed to play roles in the regulation of its expression, either in \textit{cis} or in \textit{trans}. \textit{Jpx} was initially thought unlikely to regulate XCI (Chureau et al., 2002), however, later chromosome conformation capture (3C) analysis suggested a relationship between \textit{Jpx} and \textit{Xist} (Tian et al., 2008). The \textit{Jpx} locus produces a long non coding RNA (lncRNA) that may act in \textit{trans} to regulate \textit{Xist} expression from the Xi (Tian et al., 2010) via dose dependent eviction of CTCF at the promoter (Sun et al., 2013a). Alternatively, the lncRNA could act in \textit{cis} to coactivate \textit{Xist} through promoter contacts (Barakat et al., 2014), though Sun and colleagues (2013) suggest that such contacts are not significant to XCI. Barakat et al. (2014) reached similar conclusions for the lncRNA \textit{Ftx}. Whilst initial work in male mESCs suggested that \textit{Ftx} deletion reduced expression of \textit{Xist} and resulted in dysregulation of other loci across the Xic (Chureau et al., 2011), subsequent evidence from female mESCs and \textit{in vivo} showed that \textit{Ftx} deletion does not affect \textit{Xist} expression or iXCI (Soma et al., 2014). Taken together, these data suggest that both \textit{Jpx} and \textit{Ftx} may act in \textit{cis} to play a part in \textit{Xist} regulation, lowering the threshold for activation by \textit{trans} acting factors such as \textit{Rnf12} (Barakat et al., 2014). It is, however, clear that neither acts alone, and whether the mechanism of regulation is driven by the act of transcription, the transcript itself, or at the DNA level, remains to be investigated. Finally, existence of two microRNAs within the Xic has recently been identified, though work is ongoing to explore their function (Kobayashi et al., 2013).
1.4. Dosage compensation

Pluripotency and XCI

The pluripotent state inhibits XCI (Minkovsky et al., 2012) via a number of pluripotency associated transcription factors that can be loosely grouped as either repressing *Xist* expression or promoting *Tsix*. *Xist* intron 1 can be bound by OCT4, SOX2, NANOG, and PRDM14 (Payer et al., 2013), and can also spatially associate with the *Xist* promoter (Tsai et al., 2008), repressing *Xist* transcription and therefore XCI (Navarro et al., 2008). However, whilst deletion of intron 1 in male and female mESC lines carrying a *Xist* transgene led to slightly increased *Xist* expression (Nesterova et al., 2011), recent work has removed endogenous intron 1 in both mESCs and mice without obvious effect on *Xist* expression or phenotype (Minkovsky et al., 2013). OCT4 has also been shown to bind to the promoter of *Tsix* and stimulate transcription, in addition to KLF4, c-MYC and REX1 (Navarro et al., 2008; Donohoe et al., 2009; Nesterova et al., 2011).

*Rex1* likely sits in both groups as an *Xist* repressor and an activator of *Tsix* (Gontan et al., 2012). During differentiation, REX1 is targeted for degradation by the E3 ubiquitin ligase Rnf12(RLIM), which has been proposed to behave as a *trans* acting factor, working dosage dependently to trigger XCI (Gontan et al., 2012). Male mESCs usually have a single copy of *Rnf12* and therefore do not inactivate the single X chromosome; however, when extra copies are added, XCI is aberrantly initiated (Jonkers et al., 2009). The picture in female mESCs is not so clear. Homozygous inheritance of an allele producing a 340 amino acid mutant protein (Jonkers et al., 2009) abolished XCI in differentiating cells; whereas production of a 83 amino acid mutant peptide (Shin et al., 2010) had no effect *in vitro*. The latter allele also had no effect on rXCI (Shin et al., 2014), but did perturb iXCI when inherited maternally (Shin et al., 2010).

Exit from the pluripotent state through differentiation is affected by X linked gene dosage. In female mESCs, when both X chromosomes are active, X linked genes show on average 2-fold higher expression than in male cells (Lin et al., 2007). This seems to stabilise the pluripotent state by inhibiting MAPK/Erk signaling and stim-
1.4. Dosage compensation

ulating Akt (Schulz et al., 2014). As differentiation requires MAPK signaling, the only way to exit pluripotency is for the cell to undergo XCI (Schulz et al., 2014). Utilising multiple mESC lines in serum and leukemia inhibitory factor (LIF) culture conditions, Schulz and colleagues showed that XX female cells are delayed in exiting the pluripotent state, following the induction of differentiation, when compared to XY and XO cells. They observed reduced activity of MAPK signaling, correlated with DNA hypomethylation, and suggest that these factors result in a population of cells shifted towards naive ground state pluripotency. Following induced expression of $Xist$, female cells were able to differentiate with similar kinetics to XY and XO cells. These observations support the hypothesis that two active X chromosomes inhibit downregulation of pluripotency factors, and only following feedback from XCI - indicative of X chromosome dosage compensation - is this inhibition released.

Further support for the link between exit from pluripotency and XCI can be identified from mouse development data, where it was first noted that XY male embryos show slightly accelerated growth relative to XX female embryos (Tsunoda et al., 1985). Paul Burgoyne utilised a plethora of different mutant mouse crosses to show the Y chromosome of most strains accelerates preimplantation development (Burgoyne, 1993); and furthermore that embryos with two X chromosomes show slightly delayed post implantation development (Thornhill and Burgoyne, 1993; Burgoyne et al., 1995). Intriguingly, this observation was true when comparing XX embryos to XY and $X^M$O embryos (with a single X chromosome inherited maternally). However, $X^P$O embryos (with a single X chromosome inherited paternally) showed delayed development relative to XX, XY and $X^M$O embryos at E10.5 (Thornhill and Burgoyne, 1993). This phenotypic difference suggests that $X^M$ and $X^P$ are not equal in terms of their effects on embryonic development, and warrants further attention. The aneuploid XO model system isolates maternal and paternal X chromosomes efficiently in order to address this (see section 1.5).
1.4. Dosage compensation

1.4.3 Escape from X chromosome inactivation

A number of X-linked genes are not silenced during the process of XCI. In humans, it has been estimated that around 12% show consistent escape, and a further 8% escape variably in different individuals and different tissues (Schultz et al., 2015; Carrel and Willard, 1999); whereas in mouse, the numbers are 3% and 4% respectively (Berletch et al., 2015). These genes have been found to lack the Xist coating and repressive histone marks usually characteristic of the inactive X (Simon et al., 2014; Goto and Takagi, 2000; Yang et al., 2010; Marks et al., 2015), in addition to displaying unique methylation marks (reviewed in Balaton and Brown, 2016).

Whether these differences are the cause of escape or simply a result of escape has yet to be determined. Interestingly, large clones incorporating the locus of escapee gene Kdm5c can be integrated at different X linked locations and still show expression, suggesting that escape is a locus intrinsic property (Li and Carrel, 2008). Further work deleting a non-coding region within the cloned locus resulted in expression from genes outside the integrated clone, showing that the escape property can spread (Horvath et al., 2013).

It has previously been put forward that specific DNA sequences, such as long interspersed nuclear elements (LINEs) or CTCF binding domains, might be enriched around those genes subject to or escaping from XCI (Bailey et al., 2000; Chow et al., 2010; Dixon et al., 2012). Most recently, work on nuclear ultrastructure has described the existence of topologically associating domains (TADs), intra chromosomal compartments thought to represent functional domains of gene regulation (Dekker and Heard, 2015). Groups of human and mouse genes escaping XCI have been shown to exist within the same TADs (Marks et al., 2015), and in naturally occurring autosomal translocations onto the inactive X chromosome, the translocated gene(s) followed the general pattern of the TAD, i.e. subject to or escape from XCI (Cotton et al., 2014). The significant question of why these genes escape chromosome wide silencing remains unanswered.
1.5 Aneuploidy

Aneuploidy - a deviation from euploidy, i.e. not having a normal number of chromosomes - is a term that encompasses loss and gain of chromosomes, both of which are usually problematic to the organism. One of the first reports was in 1920, when it was shown that the flowering plant *Datura stramonium* exhibited fewer negative phenotypic effects from duplication of the whole genome than duplication of a single chromosome (Blakeslee et al., 1920). This result suggested that the balance between chromosomes is more important that the overall number of copies of the genome. Further examples can be found in human disease. The only monosomy compatible with perinatal life is monosomy X (TS), whereas three autosomal trisomies are compatible with perinatal life - trisomy 13 (Patau syndrome), 18 (Edwards syndrome), and 21 (Down’s syndrome, Tybulewicz, 2006). Those afflicted with the latter present with severe phenotypes including developmental abnormalities, mental retardation and shortened life expectancy (weeks, months and around 50 years, respectively, Tybulewicz, 2006). Sex chromosome trisomies are significantly less severe, possible due to the transcriptional inactivation of all but one X chromosomes (with the exception of XCI escapees: see section 1.4.3), and the relative paucity of widely expressed genes from the Y chromosome.

1.5.1 Sex chromosome aneuploidies

A brief description of the salient features of the most prevalent sex chromosome aneuploidies are summarised in Table 1.1

1.5.1.1 Klinefelter syndrome: XXY

First described in 1942 as male patients presenting primarily with small testes and gynaecomastia, Klinefelter syndrome has a prevalence of around 0.1-0.2% of males (Klinefelter et al., 1942; Lanfranco et al., 2004). It is the most common cause of genetic infertility (Lanfranco et al., 2004). There is some evidence to suggest that XXY embryos are selected against *in utero*, as one study showed that only 55%
<table>
<thead>
<tr>
<th>Aneuploidy</th>
<th>Prevalence (birth)</th>
<th>Infertility</th>
<th>Intellectual disability</th>
<th>Congenital abnormalities</th>
<th>Growth phenotype</th>
<th>Karyotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klinefelter syndrome</td>
<td>0.1-0.2% of males</td>
<td>Yes</td>
<td>Mild</td>
<td>No</td>
<td>Tall</td>
<td>XXY</td>
</tr>
<tr>
<td>Triple X syndrome</td>
<td>0.1% of females</td>
<td>No</td>
<td>Mild</td>
<td>No</td>
<td>Tall</td>
<td>XXX</td>
</tr>
<tr>
<td>XYY syndrome</td>
<td>0.1-0.2% of males</td>
<td>Variable</td>
<td>Mild</td>
<td>No</td>
<td>Tall</td>
<td>XYY</td>
</tr>
<tr>
<td>Turner syndrome</td>
<td>0.04% of females</td>
<td>Yes</td>
<td>Mild-moderate</td>
<td>Heart, kidney</td>
<td>Short</td>
<td>XO</td>
</tr>
</tbody>
</table>

**Table 1.1:** Key features of human sex chromosome aneuploidies.
survive to term (Jacobs, 1990). Karyotypic analysis revealed the aetiology as an abnormal sex chromosome complement of XXY (Jacobs et al., 1959; Ford et al., 1959b) in around 80% of cases, with higher grade aneuploidies and structurally abnormal chromosomes making up the further 20% (Bonomi et al., 2017). Subsequent characterisation has added a number of other features to the syndrome, including osteoporosis, sexual dysfunction, and tall stature (Bonomi et al., 2017). It is likely that at least part of the spectrum of phenotypes attributed to Klinefelter syndrome result from abnormal dosage of genes escaping XCI, or those present in the PAR (see below, section 1.5.2).

1.5.1.2 Triple X syndrome

Originally described as a human "super female" (Jacobs et al., 1959), Triple X syndrome was reported only after the development of karyotyping techniques, as many cases, then and now, go undiagnosed due to the lack of clinical signs and symptoms (Otter et al., 2009). With a prevalence of around 0.1% of female births, Jacobs (1990) estimates that this represents around 70% of XXX fertilisation events; i.e. some negative in utero selection takes place. The relatively mild reported phenotype of slight developmental delay during childhood, possible correlation with increased psychiatric diagnoses in adulthood, and tall stature (Otter et al., 2009), suggest that this aneuploidy is at least partly compensated for by XCI.

1.5.1.3 Y chromosome aneuploidies

Whilst other Y chromosome aneuploidies exist as case reports, by far the most prevalent is XYY, which is detected at a rate of 0.1% of live male births (Geerts et al., 2003), with no detected selection in utero (Jacobs, 1990). Some individuals show delayed language acquisition and motor development, along with above average height, weight and head circumference from birth onwards. There is also an increased risk of child psychiatric disorders, including autism (Geerts et al., 2003), and variable rates of infertility (Kim et al., 2013).
1.5.2 Turner syndrome: XO

The clinical entity of Turner syndrome (TS) was first reported by Ullrich (Ullrich, 1930), before later being made eponymous by Henry Turner in his 1938 address to The Association for the Study of Internal Secretions (Turner, 1938). TS encompasses a core phenotype of short stature, streak ovaries, webbing of the neck, and congenital cardiovascular and renal defects (Ranke and Saenger, 2001). Notably, the severity of each of these shows significant inter individual variation (Ranke and Saenger, 2001). In 1959, Ford and colleagues linked karyotype with phenotype and thus formally identified TS as monosomy X (Ford et al., 1959a). TS is the most commonly diagnosed sex chromosome abnormality in humans, occurring in 1-2% of clinically recognised pregnancies (Hassold et al., 1992); though survival to term only occurs in about 1% of XO foetuses (Hook and Warburton, 1983). Such high in utero lethality supports the case for the importance of X linked genes in embryonic development. There are a number of mechanistic hypothesis that account for this, and the subsequent variable postnatal phenotype in survivors, including genomic imprinting and haploinsufficiency. These are now explored in further detail.

1.5.2.1 Genomic imprinting

In therian mammals, embryos derived solely from paternal genomes (androgenotes) or from maternal genomes (gynogenotes) do not survive in utero development. Androgenotes are characterised by poor progression in the embryo proper, but significant development of the extraembryonic tissue (Barton et al., 1984). Gynogenetic embryos progress further with some growth defects, though show pronounced abnormalities in the extraembryonic derivatives (Kaufman et al., 1977; Surani and Barton, 1983; Surani et al., 1984; McGrath and Solter, 1984). This requirement for one copy each of the paternal and maternal genome to initiate successful embryonic development is explained by genomic imprinting, the process whereby a group of genes is epigenetically regulated to show monoallelic expression in a parent of origin specific manner (Cleaton et al., 2014). In mouse, around 100 autosomal genes
are regulated in this way: they are usually found in small groups clustered around an imprinting control region (ICR) that is differentially methylated in either the male or female germline during development (Cleaton et al., 2014). Autosomal genomic imprinting is also found in humans - attested to by a number of severe disease phenotypes related to erroneous imprinting (see Soejima and Higashimoto, 2013) - and across the wider mammalian class, with the exception of monotremes. This observation has led to the hypothesis that the process evolved alongside evolution of the placenta (Smith et al., 2011). Interestingly, of the small number of X linked imprinted genes reported in mouse, most are expressed in placenta (Li and Behringer, 1998; Rodriguez et al., 2004; Shi et al., 2004), with brain being the other tissue (Davies et al., 2005). No related X linked ICR has been reported.

**Genomic imprinting in TS**

The parental origin of the X chromosome in XO spontaneous abortuses, late foetal deaths, and liveborn females is relatively constant, with approximately 70% maternal and 30% paternal (Hassold et al., 1985, 1988; Cockwell et al., 1991; Sanger et al., 1977; Jacobs et al., 1990). Hook and Warburton (2014) hypothesise that this deviation from 50:50 can be partly explained by a post conception mitotic sex chromosome loss. An XX embryo can lose either XM or XP, whereas an XY embryo must lose the Y and retain the XM because YO embryos are not viable (Morris, 1968a; Burgoyne and Biggers, 1976). Jacobs further suggests that the paternal X chromosome from XX embryos may also be lost before pronuclear fusion, and that the Y chromosome may be lost due to its small size (Jacobs et al., 1997).

As the available data show a bias towards a maternally inherited X chromosome in individuals living with TS, does this manifest as a difference in phenotype? Evidence has been presented that skills mediating social interactions are better developed in XP0 females, and it was hypothesised this could be explained by the existence of an X linked imprinted locus (Skuse et al., 1997; Bishop et al., 2000). However, no specific loci were found. Further suggestion of X linked imprinting in humans was provided by studying the transmission of X linked maternal alleles in
families free of detectable genetic disease (Naumova et al., 1998). A significant deviation from the Mendelian ratio resulted in distorted inheritance at a specific locus among male offspring, though no follow up has since been published.

1.5.2.2 Haploinsufficiency

Ferguson-Smith (Ferguson-Smith, 1965) suggested that the phenotype observed in patients with an XO karyotype could be explained by the presence of homologous genes on the X and Y chromosomes that escape XCI, and thus are haploinsufficient in patients lacking a second sex chromosome. Following their identification, the PARs became an obvious location for such genes (Canki et al., 1988). In human, the short stature homeobox containing gene (SHOX) certainly influences the growth deficit observed in TS patients (see next section); however, it does not seem to be involved in any of the other phenotypic abnormalities (Ellison et al., 1997).

More recently, Urbach and Benvenisty used XX and XO human embryonic stem cells (hESCs) differentiated towards various lineages to propose that CSF2RA, located in PAR1, is haploinsufficient when present in only one copy in the placenta; and is therefore responsible for the embryonic lethality (Urbach and Benvenisty, 2009). This is in keeping with the concept of the existence of a karyotypically normal rescue line in all liveborn XO individuals. In 1983, Hook and Warburton suggested that the difference between viable and non viable XO conceptuses could be the existence of undetected mosaicism for a karyotypically viable cell line, e.g. XX, suggesting haploinsufficiency at key sex linked loci as the underlying reason for death in utero (Hook and Warburton, 1983, 2014). This position was supported by extensive abortion datasets from New York and New Jersey, highlighting the ratio of apparent non mosaic XO to XO/XX mosaics was higher in embryonic and foetal deaths than in living individuals, 13.5:1 and 3.6:1, respectively (Hook and Warburton, 1983). Subsequent work has specifically searched for such hidden mosaicism in living patients with a diagnosis of TS, with varying success. Jacobs initially counted 100 cells from peripheral blood (greater than the usual 30 used for clinical diagnosis), and found 2/84 mosaics (Jacobs et al., 1997). Cells from buccal smears
1.5. Aneuploidy

were then counted, and no further mosaicism was found. More recent work used FISH on buccal smears (subsequent to the initial diagnostic 30 cells from peripheral blood) and reported a detection rate of 30% (Freriks et al., 2013). These data are largely consistent with the cryptic mosaicism hypothesis: if the patients with "complete monosomy" (Hook and Warburton, 2014) have a more severe phenotype than those with rescue lines, they would be more likely to present with symptoms and signs for medical investigation, and thus the rate of undetected mosaicism is relatively low. Furthermore, screening for mosaicism in these patients using peripheral blood and buccal swabs will not detect placental mosaicism, and it is the latter that could explain their survival (Hook and Warburton, 2014).

**Turner syndrome as an experimental paradigm**

In order to more fully understand the role of X linked genes in development and, more specifically, differences between \(X^M\) and \(X^P\), a measurable characteristic is required: the most quantifiable aspect of the TS phenotype is the growth deficit.

Short stature has long been recognised as an ever present feature of TS (Brook et al., 1974), starting with intrauterine growth retardation, as identified by sono- graphic studies. Using data resulting from prenatal screening, it has been shown that alongside the TS diagnostic anomalies of cystic hygroma and fetal hydrops, foetuses are generally small for gestational age (Wladimiroff et al., 1995). This continues postnatally, as evidenced by a retrospective study of an Italian cohort of patients with reduced birth weight (Larizza et al., 2002; Wisniewski et al., 2007). Children show a decline in growth velocity that is not rescued by a pubertal growth spurt, and adults are typically \(~2.5-3.5\) standard deviations below normal female height (Brook et al., 1974; Robles Valdés et al., 2003).

A number of other sex chromosome abnormalities have been associated with short stature, primarily including deletions of the short arms of the X or Y chromosome (Curry et al., 1984; Zuffardi et al., 1982). Further work localised the locus responsible to the PAR1 (Ballabio et al., 1989), and the gene was subsequently identified
1.5. Aneuploidy

as SHOX. Deficiency in this gene could explain the short stature observed in individuals with TS (Rao et al., 1997; Ellison et al., 1997). These individuals carry only one copy of SHOX, whereas individuals with a normal XX or XY karyotype carry two copies. Furthermore, some patients with idiopathic short stature (Rao et al., 1997; Fukami et al., 2016) and others with limb bone abnormalities (Belin et al., 1998; Chen et al., 2009; Benito-Sanz et al., 2005; Barca-Tierno et al., 2011) have mutations in the SHOX gene, independent of specific sex chromosome abnormalities. SHOX expression has since been observed in the developing limb, mainly in perichondrial layer, and in the pharyngeal arches (Clement-Jones et al., 2000; Rao et al., 2001). In vitro work suggests SHOX is an activator of the aggrecan enhancer in chondrogenesis and skeletal development (Aza-Carmona et al., 2011), thus supporting significant involvement in the TS phenotype. There is, however, some evidence to suggest that individuals with TS are generally shorter than those with SHOX mutations, and that other loci may be involved in the short stature phenotype (Ross et al., 2001; Fukami et al., 2016; Zinn and Ross, 1998).

1.5.3 XO aneuploid mice

The deduction that the female mouse can also harbour an XO sex chromosome complement came contemporaneous with the first reported human karyotype in 1959 (Welshons and Russell, 1959). A line was subsequently developed by Cattanach and colleagues in Edinburgh, UK (Cattanach, 1962). Animals of this karyotype showed no gross difference in growth from XX littermates, and were used for breeding, with apparently fewer XO daughters from XO mothers than expected. This observation was explained by the hypothesis that many oocytes segregated the chromosome set lacking an X chromosome to the polar body (Cattanach, 1962). Further work has gone on to show that XO female mice have reduced fertility when compared to XX females, both in terms of litter size and reproductive lifespan (Morris, 1968b; Burgoyne and Biggers, 1976; Burgoyne and Baker, 1981). This reduction is largely due to the presence of an unpaired X chromosome during meiosis (Burgoyne and Baker, 1985).
1.5. Aneuploidy

Originally reported as phenotypically normal (Welshons and Russell, 1959; Cat-
tanach, 1962), XO female mice were subsequently found to recapitulate at least one aspect of the TS phenotype, with growth deficits occurring during embryonic and postnatal development. X<sup>P</sup>O embryos have a developmental lag relative to XX littermates at preimplantation (Banzai et al., 1995a) and egg cylinder (E7.25) stages (Burgoyne et al., 1983b; Ishikawa et al., 1999). This lag could be explained, at least in part, by the reduced volume of the ectoplacental cone (Jamieson et al., 1998), part of the early placenta (Gardner et al., 1973). In contrast, the X<sup>P</sup>O placenta during late development is increased in volume (Burgoyne et al., 1983b). Interestingly, placentas from X<sup>M</sup>O and XY conceptuses also showed this hyperplastic phenotype (Ishikawa et al., 2003; Zechner et al., 1996, 2002). These apparently contradictory data can be reconciled by two separate explanations. The reduced ectoplacental cone volume in X<sup>P</sup>O embryos was suggested to be a result of genomic imprinting, and late development placenta hyperplasticity was attributed to X chromosome dosage (Ishikawa et al., 2003).

Imprinting in XO mice

Further, more detailed work has identified additional evidence for a potential X chromosome imprinting effect, showing that the embryonic developmental lag is present only in X<sup>P</sup>O embryos at E10.5. X<sup>M</sup>O embryos and XY embryos are significantly larger than XX littermates at the same stage, and all are significantly larger than X<sup>P</sup>O embryos (Thornhill and Burgoyne, 1993). Alternatively, this could be explained by inappropriate expression of Xist in the preimplantation embryo. Xist expression is imprinted in the early mouse embryo, such that it is always expressed from the X<sup>P</sup> at 4-8 cell stage (Okamoto et al., 2005). Notably, Xist expression also occurs in X<sup>P</sup>O embryos, which only have a single X chromosome (Matsui et al., 2001). Xist is then downregulated in the epiblast of all embryos at the blastocyst stage prior to random XCI, whereas imprinted XCI continues in the trophectoderm. Downregulation of Xist was found to take place in all cell types of X<sup>P</sup>O embryos at the blastocyst stage. It is therefore conceivable that inappropriate Xist expres-
sion results in XPO embryos functionally nullizygous for embryo derived X linked gene products for a number of cell cycles during this key developmental time window. This functional nullizygosity for an entire chromosome would likely have a significant impact on embryonic development.

**Haploinsufficiency in XO mice**

Whilst there is no growth deficit observed in XO mice during late embryonic development, and reports contrast about birth weight data (Burgoyne et al., 1983b; Deckers and Van der Kroon, 1981; Burgoyne et al., 1983a, 2002), it is clear that a postnatal growth deficit exists (Burgoyne et al., 1983a, 2002). In the earlier study, Burgoyne and colleagues compared XPO female mice to XX littermates, and showed that the rate of growth differed significantly during the first three weeks of postnatal life. This difference disappeared during subsequent phases of growth (Burgoyne et al., 1983a). The follow up study sought to determine whether this effect could be attributable to imprinting, i.e. is it present in the XmO, or X linked gene dosage deficiency (Burgoyne et al., 2002). By comparing the weights of XPO and XmO female mice with their XX littermates during the first five weeks of postnatal life, it was shown that a small but significant growth deficit exists regardless of the parental origin of the X chromosome. Reasoning that this could be due to haploinsufficiency of PAR localised genes, they utilised sex chromosome variant mice to show that two copies of the PAR did not rescue the growth deficit. Interestingly, a growth deficit was not observed in female mice carrying a Y chromosome deficient in Sry (an XYTdy female, Lovell-Badge and Robertson, 1990; Gubbay et al., 1992). The authors concluded that the postnatal growth deficit observed in XO female mice is not the result of genomic imprinting on the X chromosome, or of haploinsufficiency for genes localised to the PAR. The growth deficit was likely caused by haploinsufficiency of X linked genes that escape XCI, and that have homologues on the Y chromosome (Burgoyne et al., 2002). The identification of these genes - Ddx3x/y, Kdm5c/d, Kdm6a/Uty, and Eif2s3x/y - as likely candidates for a haploinsufficiency hypothesis is consistent with the conclusions made from the recent Y chromosome
1.6. X-Y gene pairs with candidate roles in postnatal growth

1.6.1 Kdm5c

*KDM5C* (lysine demethylase 5C) was the first gene described to escape XCI, initially in human, and very shortly after in mouse, showing ubiquitous expression in both organisms (Wu et al., 1994; Agulnik et al., 1994b; Sheardown et al., 1996; Carrel et al., 1996). Also known as *Smcx* and *Jarid1c*, the protein has a number of significant domains suggestive of histone lysine demethylase function: JmjC, a histone lysine demethylase domain (Chen et al., 2006; Klose et al., 2006), supported by a JumjN domain (Chen et al., 2006); an AT rich domain interacting domain (ARID) associates with DNA in both sequence specific and sequence independent manners (Kortschak et al., 2000; Wilsker et al., 2002); and C5HC2 and PHD zinc finger domains provide histone methyl lysine binding motifs (Shi et al., 2006; Wysocka et al., 2006; Horton et al., 2009).

Confirmation of histone lysine demethylase function was provided by a number of reports, showing activity at both H3K4 me3 and me2, and highlighting a key role in repression of transcription (Iwase et al., 2007; Christensen et al., 2007; Tahiliani et al., 2007). Work in *Drosophila* on the orthologue *Little imaginal discs* (*LID*, Eisenberg et al., 2007; Lee et al., 2007a) and in *C. elegans* on *rbr-2* has indicated that this function is highly conserved across species (Christensen et al., 2007). It is at least partly mediated by association with the transcriptional repressor REST in multi...
1.6. **X-Y gene pairs with candidate roles in postnatal growth**

protein complexes (Tahiliani et al., 2007). KDM5C and REST were shown to cooccupy neuron restrictive silencing elements, and RNA-interference (RNAi) mediated knockdown of Kdm5c de-repressed several downstream targets (Tahiliani et al., 2007). Independently, a role for Kdm5c in neural development was shown utilising zebrafish morpholinos, and isolated primary granule neurons from rat pups manifest impaired dendritic growth following Kdm5c RNAi (Iwase et al., 2007).

The link between reduced KDM5C activity and impaired neuronal function in humans was first noted in a subset of patients with X linked intellectual disability (XLID) syndrome carrying mutations in KDM5C (Jensen et al., 2005; Tzschach et al., 2006; Santos et al., 2006; Abidi et al., 2008; Rujirabanjerd et al., 2010; Ōunap et al., 2012; Peng et al., 2015; Fieremans et al., 2015). In XLID syndrome, affected individuals have an IQ<70 (Roeleveld et al., 1997). Some patients with mutations in KDM5C also have diagnoses of autism spectrum disorders, and many are of short stature (Adegbola et al., 2008). Patients reported in these papers are almost always male, and females carrying one copy of the mutation are usually described as phenotypically normal. There are, however, some reports of KDM5C mutant females with mild intellectual disability and XCI skewed to silence the mutation harbouring X chromosome. Where parental origin has been reported, the mutant allele is always maternally inherited (Abidi et al., 2008; Santos-Reboucas et al., 2011; Ōunap et al., 2012). This case based evidence supports the hypothesis that dosage of KDM5C is important. Although the gene escapes XCI, expression from the inactive X allele has been estimated at between 20%-70% of that on the active X (Sheardown et al., 1996). Moreover, expression analysis in mouse brain indicates that Kdm5c is expressed more highly in adult females than adult males, and expression of the Y linked homologue Kdm5d does not fully compensate (Xu et al., 2002). This is consistent with the significantly higher incidence of XLID in males than females, suggesting that either the Y linked homologue does not provide complete functional compensation for loss of KDM5C, or it is not expressed at equivalent levels in correct subset of cells required for normal development.
1.6. **X-Y gene pairs with candidate roles in postnatal growth**

Interestingly, and supporting a fundamental housekeeping role for *KDM5C* across multiple cell types, mutations are associated with a wide spectrum of cancers. In renal cell carcinoma (RCC), *KDM5C* could act as a tumour suppressor (Dalgliesh et al., 2010; Niu et al., 2012). *KDM5C* is abundantly expressed in hepatocellular carcinoma (HCC) cell lines, and depletion reduced migration and epithelial mesenchymal transition (Ji et al., 2015). *KDM5C* is significantly upregulated in breast cancer, and levels positively correlate with metastasis (Wang et al., 2015). Finally, DNA methylation is significantly reduced in the leukocytes of these individuals carrying these mutations (Grafodatskaya et al., 2013).

### 1.6.2 *Kdm5d*

First described as an H Y antigen responsible for the rejection of male skin grafts by female mice of the same strain (Koo et al., 1977; Simpson, 1986; Simpson et al., 1987), *KDM5D* was then named *SMCY* following mapping in mouse and human, where it was shown to be expressed widely (Agulnik et al., 1994a; Scott et al., 1995). Work utilising both *Drosophila* and human cell lines has shown that in addition to functional conservation with *KDM5C* with regards to H3K4 demethylase activity, *KDM5D* associates with Ring6a/MBLR, which enhances demethylation activity (Lee et al., 2007a). Furthermore, *KDM5D* is implicated in the regulation of *Engrailed 2*. Depletion by RNAi led to increased H3K4 di- and tri-methylation levels at *Engrailed 2* promoter, increased transcription of this gene, and enhanced recruitment of the chromatin remodelling complex NURF (Lee et al., 2007a). This evidence suggests a more general role for *KDM5D* in transcriptional regulation. A similar role has also been implicated in mouse, where KDM5D co localises with MSH5, a meiosis regulator, during spermatogenesis (Akimoto et al., 2008). Finally, mutations in *KDM5D* have been found in a significant proportion of prostate cancer cases; and knockdown in prostate cancer cell lines resulted in an increased rate of growth alongside reduced apoptosis (Perinchery et al., 2000).
1.6. X-Y gene pairs with candidate roles in postnatal growth

1.6.3 Kdm6a

*Kdm6a* (lysine demethylase 6a), also referred to by its former name *Ubiquitously transcribed Tetratricopeptide repeat on chromosome X (Utx)*, was first described in 1998 in both mouse and human tissue as showing ubiquitous expression and escaping from XCI (Greenfield et al., 1998). A decade later, histone lysine demethylase function was attributed to this previously under annotated protein, and subsequent work has shown it essential for normal development.

Both UTX and the closely related Jumonji D3 (JMJD3) have been shown to catalyse the removal of di- and trimethyl moieties on histone H3 lysine 27, thus providing a counter balance for the activity of *Polycomb Repressor Complex 2* (PRC2), a H3K27 methyltransferase (Agger et al., 2007; Lan et al., 2007; Lee et al., 2007b). Whether the Y linked homologue *Uty* also has *in vivo* demethylase remains an open question (see section 1.6.4). The H3K27 methylation mark is a highly significant mediator of transcriptional repression, involved in a plethora of biological processes including genomic imprinting (Greenberg et al., 2017), XCI (Plath et al., 2003), and cancer (Conway et al., 2015). UTX demethylase activity is effected by a JmjC catalytic domain (Lan et al., 2007), and specificity towards H3K27me3 over the highly similar H3K9me3 results from the presence of a zinc binding domain (Sengoku and Yokoyama, 2011). Protein protein interactions are thought to be mediated by the tetratricopeptide repeat domains (TPRs), though these are not present in JMJD3, suggesting a slight functional divergence (Lan et al., 2007). Despite subtle changes between family members, the H3K27 demethylases are generally well conserved across multiple model organisms, including mouse (*Utx, Uty, Jmjd3*), *Drosophila* (*dUTX*), and *C. elegans* (*UTX1* - closest to mouse *Utx*, and *UTX2*, 3 and 4 - closest to *Jmjd3*, Greenfield et al., 1998; Herz et al., 2010; Jin et al., 2011).

KDM6A has been shown to contribute to multiple protein complexes involved in the remodelling of chromatin via histone modifications. Complexes involving mixed lineage leukaemia protein 2 (MLL2) and KDM6A exhibit H3K4 methyltransferase activity, thus promoting transcription in humans (Issaeva et al., 2007;
1.6. X-Y gene pairs with candidate roles in postnatal growth

Cho et al., 2007), Drosophila (Mohan et al., 2012), and C. elegans (Vandamme et al., 2012). This suggests a dynamic relationship between transcriptional de-repression via H3K27 demethylation and transcriptional activation via H3K4 methylation. KDM6A, along with JMJD3, also interacts with the ATP dependent chromatin remodelling complex SWI/SNF (Miller et al., 2010). The authors showed that either KDM6A or JMJD3 demethylase is necessary for regulating inducible T-box factor mediated transcription independent of the demethylase activity. Finally, an interaction has also been demonstrated between dUTX and the histone acetylase Creb binding protein (CBP) (Tie et al., 2012). As histone acetylation functions to promote active chromatin, the co precipitation of dUTX, CBP and the ATPase Brm at Polycomb target genes enriched for H3K27ac might be interpreted as collective antagonism of PRC activity.

In C. elegans UTX-1 is required for progression to adulthood, as mutants lacking the protein die in late embryogenesis or early larval stages (Vandamme et al., 2012). Similarly, in Drosophila, dUTX is essential for survival: <5% of homozygous mutants pass the pupal stage but then die immediately after eclosion (Herz et al., 2010). dUTX homozygous mutant males have fewer teeth within the sex combs than wild-type males, along with rough eyes and wing defects (Herz et al., 2010). Such phenotypic features are reminiscent of Trithorax mutants, suggesting that dUTX is a Trithorax group gene, counteracting repression by Polycomb genes in the regulation of Hox genes. Unsurprisingly, H3K27me3 levels are increased in dUTX mutants and, consistent with previous work in vitro (Mohan et al., 2012), H3K4me1 levels are also affected, in a JmjC independent manner. Mutant tissue was shown to have an H3K27me3 dependent growth advantage when compared to wildtype in the developing eye. This increased proliferation was caused by increased Notch activity, and inactivation of the Drosophila orthologue of tumour suppressor Retinoblastoma, Rbf, contributes to this phenotype (Herz et al., 2010).

In order to address the function of Kdm6a in mammalian development, a number of models have been created, first in mESCs and hESCs, and then subsequently in
mice. \textit{KDM6A} is not required for ESC maintenance (Lee et al., 2012; Jiang et al., 2013), though there is evidence for a role in the differentiation of all germ cell layers. For example, \textit{KDM6A} knockdown results in impaired endoderm differentiation via WNT signaling (Jiang et al., 2013). In the mesoderm, \textit{Kdm6a} deficient cells show defects in cardiac differentiation (Wang et al., 2012; Welstead et al., 2012; Morales Torres et al., 2013), independent of demethylase activity (Lee et al., 2012). Moreover, \textit{Kdm6a} is involved in the specification of ectoderm, which is also independent of demethylase activity (Morales Torres et al., 2013). Recent work has implicated \textit{Kdm6a} in maintenance of DNA integrity, as use of a KDM6A inhibitor, along with knockdown of the gene, activated the DNA damage response in differentiating but not undifferentiated mESCs (Hofstetter et al., 2016). Interestingly, two independent groups have reported that in mESCs, but not in embryos, KDM6A regulates expression of \textit{Uty} (Lee et al., 2012; Wang et al., 2012).

Several different \textit{Kdm6a} knockout mouse lines have been created, all exhibiting relatively consistent phenotypes. Most prominently, female homozygous mutant embryos have severe cardiac malformations, neural tube closure defects, and reduced somite counts. A reduction in number of somites is indicative of retarded development (Tam, 1981). These embryos also show anaemia, myelodysplasia, and aberrant splenic haematopoiesis (Thieme et al., 2013), and invariably die between E9.5 and E13.5 (Lee et al., 2012; Wang et al., 2012; Shpargel et al., 2012; Welstead et al., 2012; Mansour et al., 2012; Thieme et al., 2013). Finally, there is some suggestion of disrupted primordial germ cell (PGC) development in \textit{Kdm6a} knockout female embryos at E12.5, leading to reduced transmission of \textit{Kdm6a} deficient germ cells in chimeras produced from targeted mESCs (Mansour et al., 2012).

Mutations in \textit{KDM6A} have generally been detected in the context of either XLID or cancer. Many patients with Kabuki syndrome (KS) have complete deletion or loss of function mutations in \textit{KDM6A}, or the H3K4 methyltransferase MLL2 (Miyake et al., 2012; Lederer et al., 2012; Priolo et al., 2012; Banka et al., 2013). KS is a developmental disorder, signs of which include craniofacial abnormalities, skele-
1.6. X-Y gene pairs with candidate roles in postnatal growth

Males with KS and a mutation in \textit{KDM6A} show a more severe phenotype than females, suggesting \textit{UTY} has functionally diverged from KDM6A (Miyake et al., 2012). Whilst the pathogenesis of KS is still not understood, a number of case reports have described predisposition to multiple different malignancies; and a number patients with malignancy have a \textit{KDM6A} mutation but do not have KS. This was first shown in a study involving multiple tumour types (van Haaften et al., 2009), and later work has provided more detail on specific cancers including clear cell renal cell carcinoma (ccRCC, Dalgliesh et al., 2010), chronic myeloid leukaemia (CML, Jankowska et al., 2011), acute lymphoblastic leukaemia (ALL, Van der Meulen et al., 2015), and medulloblastoma (Jones et al., 2012; Robinson et al., 2012; Dubuc et al., 2012; Bunt et al., 2012). Initial mechanistic work suggests that \textit{KDM6A} is putatively a tumour suppressor gene (Arcipowski et al., 2016), and the broad range of cell types affected by mutations here attest to the importance of regulated chromatin remodelling in health (Plass et al., 2013).

1.6.4 \textbf{Uty}

\textit{Uty} (\textit{Ubiquitously transcribed tetratricopeptide repeat gene on the Y chromosome}) is the Y linked homologue of \textit{Kdm6a}, with expression reported across multiple tissues and developmental stages in male mice (Greenfield et al., 1996) and humans (Lahn and Page, 1997). Although there is a high degree of sequence conservation between the two genes, conflicting \textit{in vitro} evidence exists regarding whether or not \textit{Uty} has histone demethylase activity. Walport and colleagues support common functionality, whereas separate reports by Lan and Hong do not (Walport et al., 2014; Lan et al., 2007; Hong et al., 2007).

\textit{UTY} is required for normal spermatogenesis in human (Foresta et al., 2000), and the protein also has a role in self recognition as part of the innate immune response. An epitope of UTY forms an MHC peptide (Warren et al., 2000) that contributes to graft rejection after HLA identical sex mismatched stem cell transplantation (Vogt
et al., 2000). Furthermore, in a male donor/female recipient situation, UTY specific peptides show gender specific antitumour effect in vitro (Ivanov et al., 2005; Riddell et al., 2002; Warren et al., 2000).

In the mouse mutant, UTY partially compensates for the absence of UTX in male embryos, suggesting some degree of functional conservation between the homologues. Whether this is demethylase dependent remains an open question (Shpargel et al., 2012; Walport et al., 2014). $Utx$ mutant embryos are born at a sub Mendelian ratio and show life long growth retardation, though have normal fertility (Shpargel et al., 2012). The $Utx/Uty$ compound mutant phenocopies the homozygous mutant female and is embryonic lethal (Shpargel et al., 2012).

### 1.6.5 $Eif2s3x$

Eukaryotic initiation factor 2 (eIF-2) is made up of three subunits, $\alpha, \beta,$ and $\gamma$. In the presence of guanosine triphosphate (GTP), the three subunits form a ternary complex with Met-tRNA and recruit it to the 40S ribosomal complex for translation initiation (Merrick, 1992). Subunit $2\gamma$ was initially described by Gaspar and colleagues (Gaspar et al., 1994), and subsequent work recognised that whilst this is a single X linked gene in human, there are two highly homologous genes on the X and Y chromosomes in mouse, $Eif2s3x$ and $Eif2s3y$, respectively (Ehrmann et al., 1998).

Both mouse and human X linked orthologues are expressed ubiquitously and escape XCI (Ehrmann et al., 1998). Interestingly, higher concentrations of $Eif2s3x$ transcript were found in XX females in both developing (Armoskus et al., 2014) and adult brain, and adult liver, when compared to XY males (Xu et al., 2006). However, this did not translate into differences at the protein level in adult brain or kidney (Xu et al., 2006). It is certainly possible that the low number of tissues and timepoints assayed in the latter experiment could have failed to detect relevant biological differences.
Similar to a number of other ubiquitously expressed X linked genes, EIF2S3X functions in multi protein complexes, and mutations are implicated in causation of XLID (Borck et al., 2012). Three male patients of Moroccan Jewish ancestry showed microcephaly, facial dysmorphism, and short stature, along with more specific neurological signs (Borck et al., 2012). Using human cells overexpressing one of these patient derived γ mutations, and yeast with an equivalent mutation, a defect in binding between the β and γ subunits was uncovered (Borck et al., 2012). In the yeast model, this defect impaired translation start codon selection and eIF2 function; and was rescued by overexpression of the β subunit. Although there is no Y linked homologue in human, an autosomally transposed copy of the gene exists (Ehrmann et al., 1998; Hughes et al., 2015), and it would be interesting to sequence this gene in the context of these patients to assess the degree of functional conservation. Eif2s3x is able to compensate for the loss of Eif2s3y in mouse spermatogenesis when present as a transgene in high copy number (Yamauchi et al., 2016), supporting the hypothesis that dosage of the X-Y gene pairs is critical in certain stages in certain tissues.

### Eif2s3y

Eif2s3y was mapped to the mouse Y chromosome concurrent with the identification of Eif2s3x (Ehrmann et al., 1998). It is the only Y linked gene required to complete the first meiotic division in male mouse gametogenesis (Mazeyrat and Mitchell, 1998; Mazeyrat et al., 2001). Like Eif2s3x, Eif2s3y also shows ubiquitous expression (Xu et al., 2002). In mouse brain, three isoforms of the gene have been detected, compared to one isoform of the X linked homologue; with the former showing a higher ratio of embryonic:adult expression than the latter (Xue et al., 2002). There is no human orthologue of Eif2s3y. It is hypothesised that the loss is partly compensated for by retrotransposition to an autosome, which has occurred on three independent occasions in primates (Hughes et al., 2015); and that DDX3Y has an equivalent role in human spermatogenesis (Ditton et al., 2004).
1.6.7 **Ddx3x**

The first *Ddx3x* sequence was identified in mouse erythroid cells as mDEAD3 during a screen for RNA helicase genes (Gee and Conboy, 1994). This family had been previously been characterised as ATP dependent remodellers of nucleic acid secondary structure, all having in common the amino acid sequence D-E-A-D within the ATP binding domain (Linder et al., 1989), and showing high degrees of DNA sequence conservation from yeast to mammals (Gee and Conboy, 1994; Linder, 2006; Chu et al., 2006). The human orthologue, originally known as *DBX*, was mapped to the X chromosome (Park et al., 1998), shortly after its Y homologue was identified in a somatic cell hybrid search (Lahn, 1997). Both human and mouse genes have been shown to escape XCI, and are expressed ubiquitously (Lahn, 1997; Distech et al., 2002).

As a member of the DEAD box protein family of RNA helicases, DDX3X has nine conserved structural motifs that facilitate its involvement in a wide variety of cellular processes (Cordin et al., 2006), including: pre mRNA splicing (Deckert et al., 2006), nuclear export and transport of RNA (Yedavalli et al., 2004; Kanai et al., 2004), translation initiation and regulation (Beckham et al., 2008; Shih et al., 2008), cell cycle regulation (Chang et al., 2005; Chao et al., 2006; Huang et al., 2004; Sekiguchi et al., 2007), and apoptosis (Sun et al., 2008). Likely resulting from this key role in RNA homeostasis, DDX3X is a host factor required for the replication of human immunodeficiency virus (HIV, Yedavalli et al., 2004; Sharma and Bhattacharya, 2010; Liu et al., 2011; Garbelli et al., 2011) and hepatitis C virus (HCV, Owsianka and Patel, 1999; Mamiya and Worman, 1999; Ariumi et al., 2007; Li et al., 2009b; Sun et al., 2010), in addition to its interaction with tank binding kinase 1 (TBK1) that results in the production of antiviral type I interferons (Soulat et al., 2008).

The ubiquitous expression and wide ranging involvement of DDX3X in multiple intracellular processes raises the question of what happens when a mutation arises, both experimentally and in the context of human health. The first reported work
in mouse embryos suggested an essential role for Ddx3x in early development (Li et al., 2014). Using short interfering RNA (siRNA) microinjection into mouse zygotes, a reduction in cell numbers was observed, which correlated with an increase in apoptosis and decreased progression to blastocyst when compared to scramble controls. The authors also noted that this phenotype was associated with p53 accumulation (Li et al., 2014), confirmed by later work (Chen et al., 2016a). Chen and colleagues created a conditional allele by inserting loxP sites 5’ to exon 3 and 3’ to exon 17 of Ddx3x, and initially showed that F1 animals inheriting the floxed allele on the Xp had a normal phenotype. When this same floxed allele was inherited maternally, females showed growth retardation from E12.5, then perished between E14.5 and E16.5, concurrent with evidence of oedema and peripheral haemorrhage. On further analysis of the placenta in these mutants, the labyrinth was vascularised abnormally, the spongiotrophoblast was thinner, and all placentas were smaller than wildtype littermates. A Sox2-Cre was then used to target the deletion specifically to the embryo and not the extraembryonic tissue, and these embryos were grossly normal with no significant deviation from Mendelian ratio at birth. Male conceptuses carrying the deleted allele died around E9.5, and those with the Sox2-Cre floxed allele in just the embryo were found to have abnormalities in neural tube closure, cardiac development, and chorioallantoic fusion, along with increased apoptosis in the embryo proper. Taken together, these results suggest that Ddx3x either shows tissue and developmental time specific variable escape from XCI; or that although the paternal allele escapes, expression is not sufficient to compensate for the lack of the maternal allele. Moreover, at least one copy of Ddx3x is required for normal embryonic development in the mouse, and the presence of Ddx3y alone is not sufficient.

Mutations in DDX3X have been associated with a variety of cancers, including HCC (Chang et al., 2005), head and neck squamous cell carcinoma (Stransky et al., 2011), oral squamous cell carcinoma (Koshio et al., 2013), lung cancer (Bol et al., 2015a), chronic lymphocytic leukaemia (CLL, Wang et al., 2012), colorectal cancer (Su et al., 2015), and medulloblastoma (Jones et al., 2012; Pugh et al., 2012; Robinson et al.,
1.6. **X-Y gene pairs with candidate roles in postnatal growth**

2012). Whilst there seems to be no consensus over the exact role of *DDX3X* in cancer pathogenesis, a number of small molecule inhibitors are being developed (Samal et al., 2015; Bol et al., 2015a,b).

Finally, *DDX3X* mutations have also been correlated with incidence of XLID and autism spectrum disorders, as part of a syndromic disorder with significant neurological involvement (Blok et al., 2015). The authors suggest functional mosaicism results in haploinsufficiency and a dose dependent phenotype (Blok et al., 2015).

### 1.6.8 *Ddx3y*

*DDX3Y*, also known as *DBY*, is ubiquitously transcribed in both human and mouse (Lahn and Page, 1997; Ditton et al., 2004; Foresta et al., 2000). Similar to *UTY*, *DDX3Y* was first recognised as a Y linked antigen, responsible for the rejection of human leucocyte antigen (HLA) matched allografts (Chen et al., 2004b). *In vitro* data suggest that it is able to functionally compensate for the loss of its X linked homologue *DDX3X* (Sekiguchi et al., 2004), thus implicating the protein in RNA folding and metabolism (for full description see section 1.6.7). *DDX3Y* is almost exclusively detected in the cytoplasm of spermatogonia, and individuals carrying mutations have severe germ cell loss (Ditton et al., 2004). In order to model this phenotype *in vitro*, human induced pluripotent stem cells (iPSCs) can be directed to differentiate into germ cell like cells. This ability is compromised in cells carrying *DDX3Y* mutant allele, but can be rescued by the addition of *DDX3Y* (Ramathal et al., 2015). Furthermore, during a neuronal differentiation programme, siRNA knockdown of *DDX3Y* reduced cell growth, proliferation and viability, and increased apoptosis (Vakilian et al., 2015), though humans with reduced levels of/mutations in *DDX3Y* show no apparent somatic phenotype (Ditton et al., 2004): this discrepancy remains to be explained.
1.6.9 X-Y gene pairs: summary

As these X-Y gene pairs have persisted across evolutionary space and through time, it can be deduced that they might have adopted a common survival strategy, implying a collective functional significance. A number of lines of evidence support this conclusion (summarised in Table 1.2). Following a gene expression analysis in eight mammalian species across multiple tissues and through developmental time, it was shown that X-Y pair genes are more broadly expressed than other ancestral survivors on the Y chromosome (Bellott et al., 2014). The authors reasoned that wide expression could constrain selection to prevent both loss of the Y linked gene and the appearance of a dosage compensated X linked replacement. Most X linked genes of the X-Y pairs escape X chromosome inactivation (see section 1.4.2) in human, mouse and opossum, thus preserving biallelic expression (Carrel and Willard, 2005; Yang et al., 2010; Wang et al., 2014). These genes are subject to stronger purifying selection than genes subject to XCI (Bellott et al., 2014; Park et al., 2010; Slavney et al., 2016). When added to the observation that all have functions pertaining to transcription, translation, splicing, chromatin modification, and ubiquitination (see above, and Bellott et al., 2014; Cortez et al., 2014), it becomes clear that the common factor is preservation of dosage.

Figure 1.4: Cartoon depicting relative positions of X-Y gene pairs on mouse X and Y chromosomes. Banding is representative and not to scale.
<table>
<thead>
<tr>
<th>Mouse gene (alternative names)</th>
<th>XCI status</th>
<th>Y homologue (alternative names)</th>
<th>%AA homology</th>
<th>Expression pattern</th>
<th>Mutant phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ddx3x (Dbx)</td>
<td>Escapes</td>
<td>Ddx3y (Dby)</td>
<td>88</td>
<td>Ubiquitous</td>
<td>Heterozygous female: No phenotype when inherited paternally; growth retardation and embryonic lethality when inherited paternally; specifically required for ExE development. Homozygous female: Not generated; likely lethal. Hemizygous male: Early embryonic lethality.</td>
</tr>
<tr>
<td>Eif2s3x</td>
<td>Escapes</td>
<td>Eif2s3y</td>
<td>98</td>
<td>Ubiquitous</td>
<td>None reported</td>
</tr>
<tr>
<td>Kdm5c (Jarid1c, Smcx)</td>
<td>Escapes</td>
<td>Kdm5d (HY, Jarid1d, Smcy)</td>
<td>80</td>
<td>Ubiquitous</td>
<td>None reported</td>
</tr>
<tr>
<td>Kdm6a (Utx)</td>
<td>Escapes</td>
<td>Uty</td>
<td>76</td>
<td>Ubiquitous</td>
<td>Cardiac malformation, neural tube closure defects, embryonic lethality. Reduced perinatal viability, life long growth deficit.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human gene (alternative names)</th>
<th>XCI status</th>
<th>Y homologue (alternative names)</th>
<th>%AA homology</th>
<th>Expression pattern</th>
<th>Mutant phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDX3X (DBX)</td>
<td>Escapes</td>
<td>DDX3Y (DBY)</td>
<td>92</td>
<td>Ubiquitous</td>
<td>Female: XLID. Male: XLID.</td>
</tr>
<tr>
<td>EIF2S1X</td>
<td>Escapes</td>
<td>n/a</td>
<td>n/a</td>
<td>Ubiquitous</td>
<td>Mutations reported, no obvious phenotype. Male: XLID, short stature.</td>
</tr>
<tr>
<td>KDM5C (Jarid1C, Smcx)</td>
<td>Escapes</td>
<td>KDM5D (HY, Jarid1D, Smcy)</td>
<td>85</td>
<td>Ubiquitous</td>
<td>Mild XLID, skewed XCI to inactivate mutation harbouring X. Male: XLID.</td>
</tr>
<tr>
<td>KDM6A (Utx)</td>
<td>Escapes</td>
<td>UTY</td>
<td>83</td>
<td>Ubiquitous</td>
<td>Female: XLID, short stature, Kabuki syndrome. Male: As female, more severe.</td>
</tr>
</tbody>
</table>

**Table 1.2:** Extant X-Y gene pairs in mouse and human.
1.7 Aims

In conclusion, there is abundant evidence that sex chromosome genes control growth at two stages of mammalian development. *In utero*, the growth deficit observed in X\(^{PO}\) embryos is likely explained either by genomic imprinting or inappropriate imprinted *Xist* expression. Postnatally, haploinsufficiency underlies the growth deficit seen in all XO mice relative to XX littermates. However, the genetic bases of these effects is not known.

The overall aim of my thesis is to identify possible mechanisms for both phenotypes, which could subsequently open new avenues of exploration to further understand the phenotypic effects of Turner syndrome for the three million women with the condition worldwide.

- In Chapter 3, I will test whether inappropriate *Xist* expression is responsible for the embryonic growth deficit in X\(^{PO}\) embryos.

- In Chapter 4, I will describe experiments carried out to identify X linked imprinted genes expressed in mouse trophoblast stem cells as a model for mouse placenta.

- In Chapter 5, I will test whether haploinsufficiency for *Ddx3x*, *Kdm5c*, *Kdm6a*, or *Eif2s3x* are responsible for the postnatal growth deficit in XO female mice.

- Chapter 6 describes initial experiments to address the functional divergence of X encoded genes *Ddx3x*, *Kdm5c*, *Kdm6a*, and *Eif2s3x* from their Y homologues *Ddx3y*, *Kdm5d*, *Uty*, and *Eif2s3y*. 
Chapter 2

Materials and Methods

2.1 Mouse work

2.1.1 Mouse strains

All mice used were generated and maintained on an outbred albino MF1 background (originating from MRC National Institute for Medical Research and Harlan (Blackthorn, UK)), unless otherwise stated. Specific crosses are detailed within the relevant chapters, and a summary of the origin of their constituent mutants is provided below.

2.1.1.1 In(X)1H: a large inversion in the X chromosome

The In(X)1H mutation was first reported in the offspring of a male that had received a fractionated dose of X rays to the spermatogonia (Russell, 1961; Evans and Phillips, 1975). It was noted that In(X)1H heterozygous females produced greater than expected numbers of X\textsuperscript{P}O progeny (Evans and Phillips, 1975). Furthermore, the presence of an In(X)1H allele suppressed crossing over between previously reported X-linked mutants, including \textit{Bpa}, \textit{Ta}, and \textit{Blo} (Phillips and Kaufman, 1974), suggesting a structural mutation of the X chromosome (Evans and Phillips, 1975). Cytogenetic and genetic evidence showed that the inversion spans almost the entire
length of the chromosome, and further work revealed that the formation of a dicentric chromatid during meiosis leads to the production of nulli-X oocytes (Koehler et al., 2002). Around 15% of offspring from a XY x In(X)/X cross are XpO (Burgoyne et al., 2002).

In this thesis, the In(X)1H line was used to generate XpO embryos for experiments described in Chapters 3 and 4, and XpO females for experiments described in Chapter 5.

2.1.1.2 Y*: a rearrangement in the Y chromosome

The Y* chromosome is a Y chromosome with a non-Y centromere that is attached to an inverted region of the Y PAR (Eicher et al., 1991). During male meiosis, recombination between the Y* and X chromosomes results in the generation of a large X*Y chromosome and a small X*Y chromosome, and additionally a number of gametes lacking a sex chromosome (Burgoyne and Evans, 2000). The large chromosome consists of almost the entire X chromosome, in addition to the Y* chromosome without the Y* centromere. The small X*Y is essentially the centromeric portion of the Y* chromosome in addition to a PAR (Burgoyne and Evans, 2000).

Offspring inheriting the X*Y chromosome develop as males but are sterile, likely due to the presence of two X chromosomes (Cattanach, 1961). In the absence of a second X chromosome, i.e. X*YO, a proportion of males carrying the X*Y chromosome are fertile (Trent et al., 2012). Offspring inheriting the X*Y chromosome are female and fertile. Offspring inheriting gametes lacking a sex chromosome will contribute to forming XM0 embryos.

In this thesis, the Y* line was used to generate XM0 and XY*X females for experiments described in Chapter 5. It was also used to generate X*Y0 males for Chapter 4.
2.1.1.3 $Y^{T_{dym}1}TgN(Sry-129)2Ei$: a deletion in $Sry$

The $Y^{T_{dym}1}TgN(Sry-129)2Ei$ mutation consists of a 14kb deletion in $Sry$, and was generated in mESCs before injection into blastocysts to produce chimeric founders (Lovell-Badge and Robertson, 1990; Gubbay et al., 1992). It can be used to generate XY female offspring and, when complemented by an autosomally located $Sry$ transgene, XX male offspring (Mahadevaiah et al., 1998).

In this thesis, the $Y^{T_{dym}1}TgN(Sry-129)2Ei$ line was used to generate XY females (denoted XY$^-$) in Chapter 4. It was complemented by an $Sry$ transgene and used to generate XX males (denoted XX$Sry$) in Chapter 6.

2.1.2 Timed matings

All animal work was carried out under UK Home Office license. Mice were kept on a 12:12 light:dark cycle, with food and water available ad libitum. Female and male mice were housed separately, with stud males alone and females in groups according to the cage size. Matings were set up as follows: at around 17:00 female mice would be placed into the male mouse cage, removed early the next morning and checked for the presence of a vaginal plug, which usually indicates a mating has taken place. Mating was assumed to take place during the hours of 00:00-02:00.

2.1.3 Blastocyst isolation

For the derivation of cell lines, blastocysts were retrieved at embryonic day (E) 3.5, as follows. Previously plugged females were culled via cervical dislocation, and death confirmed by severing the carotid artery. The uterine horns were removed into follicle holding medium (FHM, Appendix A), and flushed with a further 0.5-1ml of FHM utilising a 27G steel needle. Blastocysts were visualised on a Leica MC80 dissecting microscope (Leica: Wetzlar, Germany), collected using a Stripper® pipette (Origio: Malov, Denmark), and washed through a series of drops
of FHM, before a final wash and plating in the required derivation media (see below: 2.5.1, 2.5.2).

### 2.1.4 Isolation of post implantation embryos

Pregnant dams were culled as above at the appropriate time *post coitum*, and the uterine horns removed and placed in ice cold phosphate buffered saline (PBS, Thermo Scientific: Waltham, USA). Individual embryo sacs were dissected free of maternal tissue, and the embryo was freed from the membranes. After weighing and somite counting (when applicable: see section 2.1.6), representative images were taken with the Leica IC80 camera attached to Leica MC80 dissecting microscope, and acquired using LeicaAcquire software version 3.1 (all Leica: Wetzlar, Germany). Embryos were then frozen in liquid nitrogen before storage at 80°C awaiting genotyping.

### 2.1.5 Weighing of post implantation embryos

Each embryo was placed on a 30mm Petri dish and fluid removed with the corner of a lint free tissue, before weighing on a Sartorius Cubis analytical balance MSE225P (Sartorius: Gottingen, Germany).

### 2.1.6 Somite counting of post implantation embryos

Following imaging, phase contrast was used to obtain the best lighting, and somites were counted manually.

### 2.1.7 Generation of CRISPR targeted lines

#### 2.1.7.1 Guide design and synthesis

Guide sequences were designed using MIT CRISPR Designer (Cong et al., 2013); *BbsI* restriction sequences were added to the 5’ end; and oligonucleotides were
ordered (Appendix B, Eurofins: Ebersberg, Germany). Oligonucleotides were then cloned into plasmid PX330 (Cong et al., 2013), and this was transformed into chemically competent *E. coli*. Clones were screened for positive ligation by Sanger sequencing. A T7 promoter was then added to the guide sequence via PCR amplification, and sgRNA was produced by *in vitro* transcription using HiScribe T7 High Yield RNA Synthesis Kit (NEB: Ipswich, USA), and purified by MEGAclear Transcription Clean-Up (Thermo Scientific: Waltham, USA). Cas9 mRNA was synthesised using the same method with alternative primers.

### 2.1.7.2 Microinjection of pronuclear mouse embryos

Cas9 mRNA and sgRNA were mixed at 100ng/ul and 50ng/ul respectively, centrifuged at 13,000rpm for 15 minutes at 4°C, and the supernatant moved to a new tube to remove any contamination from cleaning column fibres. Microinjection was performed by Genetic Manipulation Service Science Technology Platform (STP) into pronuclear stage mouse embryos following on from a standard superovulation protocol and *in vitro* fertilisation. Injected embryos were cultured overnight before transfer to pseudopregnant females.

### 2.1.7.3 Analysis of founder animals

Pups were ear biopsied at three weeks of age and, following DNA extraction, PCR (see 2.2.4) was carried out to amplify the locus surrounding the CRISPR target site. During early stages of the project PCR products were TA cloned (see 2.2.6) and Sanger sequenced; subsequently this was replaced by the Illumina MiSeq platform (see 2.6.1).

### 2.1.8 Weighing postnatal animals

Following successful mating, pregnant dams were housed alone and, on the day of birth, pups were weighed following tattooing/toe clipping (see below), and weekly until 35 days old, using a Scout Pro balance (Ohaus: New Jersey, USA). Pups were
housed with the dam until weaning at three weeks of age, and subsequently male and female pups were housed separately.

2.1.9 Identification of individuals

In initial experiments, new born pups were marked using a subcutaneous injection of black pigment (VWR: Lutterworth, UK) as follows: pups were incubated in a 10cm plastic dish on ice water for three minutes to provide light anaesthesia/analgesia; skin was cleaned with 70% EtOH and sterile gauze, and a 33G needle and glass syringe (Hamilton Company: Nevada, USA) were used to inject 2ul pigment subcutaneously into the paw pad; pups were kept on a heat pad until righting reflex was regained.

In subsequent experiments, toe clipping was performed in order to identify individual animals: this method was faster and provided tissue for genotyping where required, removing the requirement for a subsequent ear biopsy procedure later. Local anaesthesia was applied to the paw. After an appropriate time, the distal phalanx was removed from one digit per animal, and any bleeding cauterised chemically with silver nitrate. Pups were again kept on a heat pad until return to the dam.

2.2 Molecular biology

For PCR, primer sequences were either obtained from previously published work or designed using Primer3 (Untergasser et al., 2007). All PCR primer sequences are provided in Appendix B; where commercial sources have been used, this is indicated.

2.2.1 Genomic DNA (gDNA) extraction

Generally, gDNA was extracted from mouse tissue samples lysed in KT buffer (Appendix A), by overnight incubation at 55°C, followed by 15 minutes at 95°C and five minutes centrifugation at 13,000rpm, room temperature (RT). For qRT-PCR (sec-
2.2. Molecular biology

2.2.5.2), gDNA was extracted using PureLink (Thermo Scientific, Waltham, USA), as per the manufacturer’s protocol.

2.2.2 RNA extraction

Total RNA was extracted from frozen tissue samples and cell cultures using TRIzol (Thermo Fisher: Waltham, USA) or TRI reagent (Sigma-Aldrich: St Louis, USA) as per the manufacturers’ protocols. When amount of available tissue was limited (i.e. E10.5 embryo) and both RNA and DNA were required, DNA was extracted from the lower phase, also per manufacturer’s instructions. Extracted RNA was dissolved in ddH2O and quantified using a Nanodrop spectrophotometer (Thermo Fisher: Waltham, USA).

2.2.3 cDNA synthesis

Synthesis of cDNA for qRT-PCR was described in section 2.2.5.2. For general cDNA synthesis, total RNA was added to oligo(dT) (Thermo Fisher: Waltham, USA) and 10mM dNTPs and incubated at 65°C for five minutes. First strand buffer, DTT and RNAse Out (all Thermo Fisher) were then added and, following a further two minute incubation at 42°C, Superscript II (Thermo Fisher: Waltham, USA) was added, and the mixture incubated at 42°C for another 60 minutes. The enzyme was heat inactivated at 70°C for 15 minutes.

2.2.4 Polymerase Chain Reaction (PCR)

For standard genotyping, 0.5ul of tissue lysis supernatant was added to 2X MyTaq Redmix (Bioline: London, UK). Primers (Eurofins MWG: Ebersberg, Germany) were used at a final concentration of 0.2nM. For PCR to be followed by sequencing, the high fidelity q5 Taq polymerase was used, as per the manufacturer’s protocol (NEB: Ipswich, USA). Product was loaded onto an agarose (Sigma-Aldrich: St Louis, USA) gel with 1kb ladder (Thermo Scientific: Waltham, USA) and run at
2.2. Molecular biology

around 85V for an appropriate time. Images were then acquired using AlphaImager 2200 (ProteinSimple: California, USA) with AlphaImager software.

2.2.5 Quantitative real time PCR

2.2.5.1 qRT-PCR on gDNA

To assess X chromosome copy number, gDNA was quantified using qRT-PCR, as previously described (Yamauchi et al., 2015), but with minor variation. gDNA extracted using PureLink was amplified with primers targeting *Amelx* and *Prdx4* on the X chromosome, and *Myogenin* as an autosomal control. Both X linked gene products were normalised to the autosomal control (to generate \( \Delta \text{Ct} \) values), before comparison with a known XX sample (to generate \( \Delta \Delta \text{Ct} \) values, as per Livak method (Livak and Schmittgen, 2001)). The reaction was performed using SensiMix LoRox polymerase (Bioline: London, UK), on a ABI 7500 thermal cycler (Thermo Scientific: Waltham, USA).

2.2.5.2 qRT-PCR on cDNA

RNA was extracted as described (section 2.2.2). For relative quantification of gene expression, RNA was converted to cDNA using the Maxima First Strand cDNA Synthesis Kit following the manufacturer’s protocol (Thermo Scientific: Waltham, USA). Briefly, between 500ng and 5ug total RNA was first treated with double stranded DNase, before conversion to cDNA using reverse transcriptase with a mix of random hexamer and oligo(dT) 18 primers. Minus reverse transcription controls were utilised to detect the presence of contaminating genomic DNA. Primers were first tested for specificity and efficiency via melt curve analysis and standard curve methods, respectively: the pairs that showed a single peak during the melt curve analysis and an efficiency of 2 +/- 0.1 were deemed suitable for use. On occasion, multiple primer pairs designed to amplify a given locus failed testing, so a commer-
cial alternative was sought and used following manufacturer’s protocol (Taqman, Thermo Scientific: Waltham, USA).

384 plates were prepared with Lightcycler 480 Sybr Green I mastermix (Roche: Basel, Switzerland) as per the manufacturer’s protocol; reactions were performed in triplicate and run on the Lightcycler 480 II thermal cycler (Roche: Basel, Switzerland). Data were exported from the proprietary software to Excel (Microsoft: Washington, USA) for statistical testing based on the Livak method (previous section).

### 2.2.6 TA cloning

PCR products were TA cloned (TOPO, Thermo Fisher: Waltham, USA). Cultures were grown from colonies in Luria Broth (LB), DNA was extracted by mini prep according to the manufacturer’s instructions (Qiagen: Manchester, UK), and clones were Sanger sequenced by Beckman-Coulter genomics (Takley: Essex, UK) using M13 primers.

### 2.2.7 RFLP digest

SNPs were selected using the Mouse Genomes Project online browser (Wellcome Trust Sanger Institute), DNA sequences were acquired from UCSC Genome Browser (build GRCh38/mm10). Restriction enzymes were selected using NEB-cutter (NEB: Ipswich, USA). Total RNA was extracted from mouse brain and liver and transcribed into cDNA as described (see 2.2.2, 2.2.5.2). Following PCR, products were run on a 2% agarose gel, isolated using a scalpel blade and UV light box, and purified using NucleoSpin PCR Gel and PCR Clean-up (Machery Nagel: Duren, Germany) according to the manufacturer’s instructions. Products were then digested with restriction enzymes (all NEB: Ipswich, USA) as per manufacturer’s guidelines, run on a 2% agarose gel, and visualised as described (2.2.4).
2.3 Cytology

2.3.1 Metaphase chromosome spreads for karyotyping

Mice were culled as described above. Using a 1ml syringe filled with HEPES buffered RPMI medium (Thermo Fisher: Waltham, USA) supplemented with colcemid (0.1ug/ul; Sigma-Aldrich: St Louis, USA), a 27G needle was inserted into the distal end of the femur, and bone marrow was flushed into a 15ml round bottomed polypropylene tube (Greiner Bio-One: Stonehouse, UK) with 0.5ml solution per femur. The tube was incubated horizontally for 10 minutes at 31°C, vortexed briefly and centrifuged at 180 x g for five minutes in a swinging rotor centrifuge at RT. The supernatant was drained, the pellet re suspended in 1-2ml 0.56% w/v potassium chloride (Sigma-Aldrich: St Louis, USA) dissolved in ddH2O, and the tube incubated at RT for at least 20 minutes. Following a further brief vortex and spin as previously, the supernatant was carefully drained, methanol:glacial acetic acid (3:1; MA) was added down the side of the tube, drop wise, until the pellet was coated, and the tube rotated to ensure even coverage. 1ml of MA was then added, and the tube vortexed briefly to re suspend the pellet, quickly followed by further centrifugation as before. The supernatant was drained carefully, 1-2ml of MA was added, and the tube incubated for 15 minutes at RT. Two further consecutive washes were carried out: first, the tubes were centrifuged as previously, followed by draining of the supernatant, fresh 1ml MA added, and the solution vortexed. After the second drainage step, a suitable amount of MA was added to result in a good final density of cells on the slides (judged based on experience, around 1ml); the tube was vortexed, left to settle for 30 seconds, and then three drops were spread onto a cleaned (98% EtOH/2% acetic acid) glass slide (VWR: Lutterworth, UK), from a height of about 30cm. After air drying, the slides were stained in 4% Giemsa buffered in PBS (pH 6.8, Sigma-Aldrich: St Louis, USA) for eight minutes, then viewed under an Olympus BH2 microscope (Olympus: Tokyo, Japan), and imaged using a Leica camera (Leica: Wetzlar, Germany). When used for DNA FISH, one
drop of the final MA fixed cell suspension was spread onto multi well glass slides (Thermo Scientific: Waltham, USA) and left to air dry.

2.3.2 DNA FISH

2.3.2.1 BAC extraction

BAC clones (Appendix C) were selected using Ensembl genome browser (Mouse, GRCm38) and obtained as stabs from CHORI BACPAC (California, USA). Cultures were set up using 20ml LB inoculated with chloramphenicol (25ug/ml, dissolved in 100% EtOH) and incubated at 37°C, 200rpm overnight. Glycerol stock was made using 500ul culture and 500ul 100% glycerol (autoclaved; Sigma-Aldrich: St Louis, USA), and stored in a 1ml Nunc tube (Thermo Fisher: Waltham, USA) at -80°C. The cultures were then centrifuged at 4500rpm, 4°C for 15 minutes. After discarding the supernatant, the pellet was resuspended in P1 (Appendix A) by vortexing; P2 (Appendix A) was then added, and the solution was inverted 20 times to mix, followed by an incubation of five minutes at RT. P3 (Appendix A) was added, and once again the solution was inverted 20 times to mix and incubated on ice for five minutes. The solution was then centrifuged at RT for 10 minutes at 13,000rpm, the supernatant transferred to fresh tubes, and 0.6 volumes of RT isopropanol was added, with inversion to mix; followed by immediate centrifugation at RT for 15 minutes, 13,000rpm. The supernatant was discarded, the pellet washed with 75% EtOH, and subsequently centrifuged down at 12,000rpm for 5 minutes at RT. The EtOH was then discarded and the pellet air dried, followed by resuspension in ddH2O. DNA was incubated overnight at 37°C with RNAse A (Sigma Aldrich: St Louis, USA). The next day, RNAse A was removed using subsequent washes of equal volume phenol, phenol:chloroform, and chloroform:isoamyl alcohol (all Sigma-Aldrich: St Louis, USA). 1:10 3M NaAc at pH5.2 and 2.5X 100% EtOH were added to the final supernatant, which was centrifuged at 13,000rpm for 15 minutes at 4°C. The supernatant was discarded, and the pellet washed with 75% EtOH; finally, the DNA was dissolved in ddH2O.
2.3. Cytology

2.3.2.2 Probe labelling

Approximately 1ug BAC DNA was incubated with Abbott nick translation enzyme and either Spectrum Green dUTP, Spectrum Red dUTP (all Abbott Molecular: Illinois, USA), or Cy5 dUTP (Jena Bioscience: Jena, Germany) at 15°C for 12 hours. 1/10 of this reaction was run on a 2% agarose gel at 85V for 45 minutes to check for probe size of around 200bp. If the probe was found to be greater than 300-400bp, the reaction was incubated for varying times (minutes) at 37°C and the electrophoresis step repeated.

In order to prepare one probe for one slide, 1/10 of this reaction was added to 3ul mouse Cot1 DNA and 1.5ul salmon sperm (both Thermo Fisher: Waltham, USA); the mix was then precipitated using 2.5X 100% EtOH, mixed, and then centrifuged at 4°C for 15 minutes, 13,000rpm. The supernatant was removed, the pellet washed with 70% EtOH, and after centrifugation for five minutes at 13,000rpm, left to air dry. The probe was then dissolved in deionised formamide (Sigma-Aldrich: St Louis, USA) and either used immediately or stored at -20°C.

2.3.2.3 Hybridisation

Slides were washed in 2X saturated sodium citrate (SSC, Appendix A), and denatured in 70% formamide/30% 2XSSC at 75°C, before dehydration through a series of EtOH. Concurrently, the probe was denatured at 75°C and pre hybridised in hybridisation buffer (Appendix A). Probes were added to slides, followed by coverslips, and hybridisation took place in a formamide humid chamber overnight at 37°C. The following day, slides were washed four times in 2XSSC, then 0.1XSSC, before transfer to 4XSSC/0.1% Tween (Sigma-Aldrich: St Louis, USA). 4’,6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich: St Louis, USA) was then used to stain the slides for 10 minutes, before a final brief wash in 4XSSC/0.1% Tween, followed by mounting in Vectashield (Vector Laboratories: Burlingame, USA). Slides were visualised using an Olympus IX70 inverted microscope (Olympus: Tokyo,
2.4. Protein extraction and western blotting

2.4.1 Tissue harvest and protein extraction

Tissue was collected by dissection, snap frozen in liquid nitrogen, and stored at -80°C until required. Upon thawing, tissue was placed into a 2.0ml reinforced plastic tube (Bertin Instruments: Montigny-le-Bretonneux, France) containing RIPA buffer (Appendix A) and two-three metal beads (Bertin Instruments: Montigny-le-Bretonneux, France), and stored on ice for 15 minutes. Subsequently the samples were run on the Precellys 24 (Bertin Instruments: Montigny-le-Bretonneux, France) and for 45 seconds to break up the tissues. Following centrifugation at 13,000rpm, 4°C for 30 minutes, the supernatant was removed to a new tube and stored on ice.

2.4.2 Protein quantification

Protein extract supernatant was diluted 1:100 and subjected to a bicinchoninic acid assay (BCA, Thermo Scientific: Waltham, USA) as per the manufacturer’s protocol. A standard curve was prepared using bovine serum albumin (BSA, Sigma-Aldrich: St Louis, USA). Colour change was quantified with Infinite M1000 Pro (Tecan: Mannedorf, Switzerland), and analysed using proprietary Tecan software. Following quantification, protein extract was stored at -80°C until required.

2.4.3 Western blotting

20-30mg protein was prepared for sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) by dilution with Laemmli loading buffer (Appendix A). Samples were run on a polyacrylamide gel (generally 4-15%, Bio-Rad: Her-
2.5. Cell culture

Unless otherwise stated, all cell culture plasticware was obtained from Nunc (Thermo Fisher: Waltham, USA), and media components from Invitrogen (Thermo Fisher: Waltham, USA).

2.5.1 Mouse trophoblast stem cells (mTSCs)

mTSCs were derived and cultured as described previously (Himeno et al., 2008), and summarised below.

2.5.1.1 Derivation and maintenance

Gamma irradiated MEFs (section 2.5.3.1) were plated in 4 well plates at $4 \times 10^4$ cells per well, using mTSC medium (Appendix A). The next day, human recom-
binant FGF4 (25μg/ml; Peprotech: London, UK) and heparin (1mg/ml; Sigma-Aldrich: St Louis, USA) were freshly added to mTSC medium (mTSCM), and this was applied to the MEFs. Blastocysts were harvested at E3.5 as described above (section 2.1.3), and one embryo was added per well of MEFs. When embryos were observed to attach to the MEFs, medium was changed, and then subsequently every second day. Outgrowths appeared over the next three-five days, and upon reaching a size of 1mm, were dissociated in 0.1% trypsin/1mM EDTA back into the same well. Cultures were subsequently fed with a mix of MEF conditioned medium (CM, section 2.5.3.1) and fresh mTSC medium at a ratio of 70:30, in addition to 1.5X F4H. Tight epithelial mTSC like colonies were observed to appear over the next two weeks, and often differentiation was also noted. In the case of the latter, colonies were picked, dissociated, and re plated to obtain morphologically consistent mTSC lines. Upon reaching 50% confluence, colonies were expanded and passed onto new MEF feeders, and fed mTSCM + 1.5X F4H every second day. This process continued, with 1:20 passages approximately once per week, until reaching passage 10. At this point, lines became more stable, and could be passed without feeders using 70CM + 1X F4H medium. Cells could then be genotyped and characterised via gDNA and RNA extraction and DNA FISH.

2.5.1.2 Differentiation

To induce differentiation of mTSCs, cells were plated feeder free in mTSCM without F4H. RNA was isolated on days two, four, and six using TRI reagent as described (section 2.2.2).

2.5.1.3 Freezing

Prior to freezing, supernatant was removed from confluent cultures for Mycoplasma testing in-house (Cell Services Science Technology Platform, Francis Crick Institute). mTSCs were frozen in mTSCM with 10% dimethyl sulfoxide (DMSO, Sigma-Aldrich: St Louis, USA): approximately three x 1ml cryovials were ob-
tained from one confluent well of a 6 well plate. Vials were placed in a cell freezing container at -80°C overnight, before transfer to permanent storage in liquid nitrogen.

2.5.2 Mouse embryonic stem cells (mESCs)

mESCs were derived and maintained as described previously (Ying et al., 2008), and briefly outlined below.

2.5.2.1 Derivation

Blastocysts were harvested at E3.5 as described above (section 2.1.3), and one embryo was added per well of a 24 well plate containing 2i+LIF (Appendix A). After approximately seven-ten days, spherical outgrowths were observed, picked, and dissociated using TrypLE reagent. This was re plated into one well of a 4 well plate previously coated with poly-L-ornithine and laminin (as described previously Hayashi and Saitou, 2013), whilst a small quantity of the dissociated outgrowth was reserved for PCR genotyping. Upon reaching confluence, cells were expanded and frozen. X chromosome copy number was assessed via gDNA qRT-PCR and/or DNA FISH.

2.5.2.2 Freezing

One well of a 6 well plate was frozen in a single vial using 2i+LIF + 20% DMSO. Supernatant was reserved for Mycoplasma testing as described above.

2.5.2.3 Maintenance

mESCs were passed every two-three days at approximately 1:5. Passage number was kept low in order to reduce the occurrence of karyotypic abnormalities, as previously reported (Eggan et al., 2002; Robertson et al., 1983).
2.5.2.4 Electroporation

Electroporation was used to deliver a vector containing Cas9 and guide sequences, along with a drug selection cassette (pPGKpuro, Tucker et al., 1996). mESCs in the early stage of growth were disaggregated into solution and counted using Scepter 2.0 (Merck Millipore: Billerica, USA). 2 x 10^6 cells were then electroporated using Mouse ES Cell Nucleofector Kit (Lonza: Basel, Switzerland), as per the manufacturer’s protocol. Drug selection was carried out using 1.5ug/ml puromycin (Sigma-Aldrich: St Louis, USA) to identify clones that had taken up plasmid and, following a period of growth, colonies were picked, and genotyped by PCR followed by restriction digest.

2.5.2.5 Mouse epiblast like cell (mEpiLC) differentiation

mESCs were differentiated towards an epiblast-like cell fate according to a previously published protocol (Hayashi et al., 2011a; Hayashi and Saitou, 2013). In brief, mESCs were grown to confluence, disaggregated with TrypLE, and plated at 1 x 10^5 cells per well of a 12 well plate coated with human plasma fibronectin (Merck Millipore: Billerica, USA). N2B27 medium (Takara Bio: Tokyo, Japan) was supplemented with activin A at 20ng/ml (Peprotech: London, UK), FGF2 at 12ng/ml (Peprotech: London, UK), and 1% knockout serum replacement (KSR, Thermo Fisher: Waltham, USA); and cells were cultured for 48 hours prior to RNA extraction.

2.5.3 Derivation of MEFs

Gamma irradiated MEFs were used as feeder cells for mTSCs: originating primary cells were generated according to a published protocol (Xu, 2005). Briefly: 14 days following timed mating, pregnant dams were culled, and uterine were horns dissected out using sterile instruments followed by a thorough wash in sterile PBS (Thermo Fisher: Waltham, USA). In a laminar flow hood, the embryos were cleanly dissected away from the uterus and the viscera and brain were removed. Individual
2.6. Next generation sequencing

Embryos were then teased apart using fine forceps in ice cold trypsin EDTA, and incubated overnight at 4°C. The following morning, most of the trypsin EDTA was removed without disturbing the settled tissue, and the remainder was incubated at 37°C. Tissues were then dissociated into cells, plated in MEF medium (Appendix A) as one embryo per plate, and allowed to grow to confluence before freezing in vials with MEF medium/20% DMSO. Vials were placed in a cell freezing container at -80°C overnight, before transfer to permanent storage in liquid nitrogen.

2.5.3.1 Gamma irradiation of MEFs and production of conditioned medium

One vial of primary MEFs was thawed and expanded for two passages, generating between 5 and 10 x 10^7 cells. Following dissociation, cells were exposed to a radiation dose of approximately 35 Gy via a caesium source, counted, and frozen as described above. For the production of MEF conditioned medium (CM), irradiated cells were plated at a density of 6 x 10^6 per 150mm plate. mTSCM was added for five x three day periods: subsequent to collection of the final batch, all batches were pooled, centrifuged and filtered to remove cell debris, and stored at -20°C until use.

2.6 Next generation sequencing

2.6.1 Generation of data to screen CRISPR targeted loci

Illumina MiSeq platform was used to generate sequencing data for genotyping of CRISPR targeted cells and animals. Primer sequences were designed around the region of interest using Primer3 (Untergasser et al., 2007), with a maximum ampli-con size of 500bp and synthesised by Eurofins MWG (Eurofins MWG (Ebersberg, Germany). Adapter sequence was added to the 5’ and 3’ ends of the forward and reverse primers, respectively, in order to enable further amplification with Illumina TruSeq indexing primers (Illumina: San Diego, USA). PCR was performed us-


2.6. Next generation sequencing

Using q5 high fidelity polymerase (NEB: Ipswich, USA), according to manufacturer’s protocol. Unused polymerase and enzyme was removed from the reaction using solid phase reverse immobilisation (SPRI) beads, to the specification of AMPure XP (Beckman-Coulter: Brea, USA), but made in house. Indexing was performed as per the Illumina protocol, and a further clean up step utilised SPRI beads. Sample quantification utilised the dsDNA dye Quanti-fluor (Promega: Fitchburg, USA); signal was recorded with Infinite M1000 Pro (Tecan: Mannedorf, Switzerland), and analysed using proprietary Tecan software. Samples were then normalised and pooled before submission to the Advanced Sequencing Facility (ASF) at the Francis Crick institute. Following quality control - library quantitation via Eco Real-Time PCR (Illumina: San Diego, USA) and sizing via 2100 Bioanalyzer (Agilent: Santa Clara, USA) - the library was loaded onto the MiSeq at 2nM, using MiSeq Reagent Kit v2 (Illumina: San Diego, USA) for 2 x 250bp sequencing.

2.6.2 Data analysis

Fastq files were collapsed using FastX Toolkit (v0.0.13) and aligned to the reference genome (mm10) using blastn.

2.6.3 RNA-seq

2.6.3.1 Library preparation

Total RNA was extracted as described above (section 2.2.2). 500ng was submitted to ASF and assessed for RNA integrity (RIN) using Caliper LabChip GX (Perkin Elmer: Waltham, USA). Samples that passed quality control were prepared for sequencing by ASF using TruSeq Stranded mRNA Library Prep Kit as per manufacturer’s protocol (Illumina: San Diego, USA). Libraries were sequenced across two runs on Illumina HiSeq 4500.
2.6.3.2 Data quality control

Quality control was performed on FASTQ output files using the FastQC package (Andrews, 2010). Paired end reads from each library were each trimmed of 13 base pairs from the 5’-end and 3 base pairs from the 3’ end using Trim-Galore (Krueger, 2012), based on output from the Per Base Sequence Quality metric.

2.6.3.3 Mapping and differential expression analysis

Trimmed reads were then aligned to the mm10 genome using HISAT2 (v2.0.6: Kim et al., 2015). Uniquely mapped reads, (ie. fragments that mapped once only to the reference genome) and with a MAPQ score > 40 were retained for further analysis. These were then reverted to FASTQ format using PicardTools SamToFastq prior to input for further steps.

Transcript abundances were then determined using Salmon (v.0.7.2, Patro et al., 2017). Mouse Gencode (VM12) transcriptome annotation was used to create a quasi Salmon index.

Differential gene expression analysis was performed in R (v.3.2.2, R Core Team, 2013) utilising two independent methods. Firstly output from Salmon were passed to Voom, limma package (Ritchie et al., 2015) using tximport (Soneson et al., 2015), and gene level differential expression (DE) testing was performed. Secondly, outputs from Salmon were also passed onto Sleuth (Pimentel et al., 2017) via kallisto (Bray et al., 2016), and again gene level DE testing was carried out. A comparison of XY and XX mESCs was included as a positive control for detection of DE genes, which guided the setting of filtering thresholds for the exclusion of lowly expressed genes.
2.7. Statistical analysis

2.6.3.4 Over representation analysis (ORA)

Over representation analyses were carried out as using the Web based gene set analysis toolkit (WebGestalt, Zhang et al., 2005; Wang et al., 2017) with standard settings.

2.7 Statistical analysis

2.7.1 Weight data analysis

There are major litter effects on growth during pregnancy and early postnatal life independent of genotype; therefore comparisons between genotypes were made within litters (Burgoyne et al., 1983). Mean weighted differences (MWD) were used in order to compare weight data, thus taking into account varying litter sizes and numbers of animals per genotype within each litter. The difference between the required two genotypes was weighted according to the number (n) of animals of those genotypes (i.e. \( w = n_1 \times n_2 / (n_1 + n_2) \)). The MWD was then calculated by dividing the sum of the weighted differences over the sum of the weights (w), and standard error calculated from estimates of the variances in each genotype in each litter. MWD over standard error is distributed as Student’s t (Burgoyne et al., 2002).

At E10.5, embryos are growing exponentially, therefore a log transformation was performed before the MWD calculation to bring the data into a normal distribution. Additionally, it has previously been noted that some XO embryos are more runted than expected, and can thus seriously affect XY-XO comparisons (Burgoyne et al., 1983b). To test for this effect, the differences between XY and XO, and XY and \( X^{\text{Xist}} \)-O embryos were collated across litters and the distribution checked for outliers (Grubbs’ test): none were found.

Postnatal growth rates were calculated by plotting each individual animal’s weight data and fitting a regression line; the gradients from 0-3 weeks and 3-5 weeks were then compared using MWD as described in (Burgoyne et al., 2002). Data were
analysed using Excel (Microsoft: Washington, USA) and Prism (Graphpad: La Jolla, USA).

### 2.7.2 Postnatal survival analysis

The statistical significance of the differences between expected and observed numbers of each genotype was calculated using Fisher’s Exact Test of Independence, using the package rcompanion in R (R Core Team, 2013) I corrected for multiple comparisons, where applicable, using the Holm-Bonferroni method (Holm, 1979). Mosaic plots were created using the package ggplot2 in R.
Chapter 3

Examining the role of \textit{Xist} in the post-implantation growth deficit observed in \textit{X}^{P}\textit{O} embryos

3.1 Introduction

TS is invariably associated with low birth weight, reduced growth during childhood, and adult short stature. These effects likely manifest from intrauterine growth retardation (IUGR) during embryogenesis. XO mouse embryos are similarly growth retarded during development. There is limited evidence to suggest that this manifests at the pre-implantation stage, though it is certainly present as a weight deficit by the egg cylinder stage, at around E7.25 (Banzai et al., 1995b; Burgoyne et al., 1983b). The weight deficit persists until around E12.25, with a further lag period between E14.5 and E16.5 (Burgoyne et al., 1983b). Interestingly, subsequent work focusing on the parent of origin of the single X chromosome in XO mice showed that this effect is only observed in \textit{X}^{P}\textit{O} embryos (Thornhill and Burgoyne, 1993). In contrast, at E10.5 \textit{X}^{M}\textit{O} embryos were shown to be heavier than XX littermates, and of equivalent weight to normal \textit{X}^{M}\textit{Y} males (Thornhill and Burgoyne, 1993). Bringing these data together, it was hypothesised that the \textit{X}^{P} carries an imprint that reduces
3.1. Introduction

X linked gene expression across the whole chromosome, thus delaying growth; and that the X<sup>M</sup> is more efficient at supporting early development. This could be explained by imprinted Xist expression from the X<sup>P</sup>, thus silencing X linked genes in early X<sup>P</sup>O embryos.

These observations were confirmed, and the reduced X linked gene expression hypothesis tested, using a lacZ transgene to assay gene activity from the single X chromosome in X<sup>P</sup>O embryos (Jamieson et al., 1998). In XY males and homozygous carrier females, X-gal staining was positive in the embryo proper and in cells of both extraembryonic ectoderm (ExE) and ectoplacental cone (EPC). Heterozygous female embryos carrying the transgene on the X<sup>M</sup> showed a similar pattern, whereas those carrying the transgene on the X<sup>P</sup> displayed no expression in the ExE or EPC, in line with imprinted XCI of the X<sup>P</sup> in these tissues. If the single X chromosome in the X<sup>P</sup>O showed significantly reduced or absent staining, this would support a case for chromosome wide downregulated gene expression as a cause of the growth deficit in these embryos. Intriguingly, X<sup>P</sup>O embryos stained positively for X-gal in all tissues, suggesting that the single X<sup>P</sup> can behave as per the X<sup>M</sup> in the absence of a second sex chromosome. The authors also observed that the EPC of X<sup>P</sup>O embryos is significantly smaller than XX littermates, likely due to a reduction in the number of cells, as the packing density was normal (Jamieson et al., 1998).

Taken together, whilst these data show that the X<sup>P</sup> is transcriptionally active in the X<sup>P</sup>O context at E7.5, it is still possible that global transcription across the chromosome is reduced, or indeed absent, at an earlier stage, possibly as a result of paternal Xist expression. Alternatively, specific genes may be required from the X<sup>M</sup> that are not expressed from the X<sup>P</sup> due to epigenetic differences independent of iXCI, and their absence compromises development.

XCI controls transcription across the X chromosome. In the pre-implantation embryo this process is mediated by Xist in an imprinted manner, such that X<sup>P</sup> is always inactivated. Xist expression was assayed by RNA FISH in 8 cell and blastocyst stage mouse embryos (Matsui et al., 2001). Whereas in XY male embryos no expression
was observed at either stage - as would be expected in the presence of a single $X^M$ - a distinct signal was present in every cell of the 8 cell $X^P$O embryos, with a decreased frequency of 20% of cells in fully expanded blastocysts (Matsui et al., 2001). Based on these data, it can be hypothesised that inappropriate $Xist$ expression in the pre-implantation $X^P$O embryo results in temporary chromosome wide silencing, and an embryo nullizygous for $X$ linked gene products. Functional nullisomy for $X$ linked gene products has previously been shown to be extremely harmful to early embryonic development (Morris, 1968a; Burgoyne and Biggers, 1976), therefore it is possible that this causes the growth retardation phenotype observed from E7.25 onwards. I sought to test this hypothesis by asking whether deletion of $Xist$ rescues the developmental retardation of $X^P$O embryos.

3.2 Results

I set out first to reproduce previously published data confirming that $X^P$O embryos show significant growth retardation when compared with XX littermates at E10.5. $X^P$O embryos resulted from cross 1 (Figure 3.1 A(i)). Here, I utilised female mice heterozygous for a large pericentric inversion on the $X$ chromosome, In(X), as a source of ‘O’ gametes (Evans and Phillips, 1975). When mated to an XY male, there are six possible offspring (Figure 3.1 A(i)), five of which survive to E10.5. Four of these genotypes can be identified by PCR on genomic DNA (Figure 3.1B), whereas RT-PCR to detect $Xist$ was used to distinguish between XX and XO embryos (see section 2.2.3).

Preliminary analyses showed no significant difference in weight or somite counts between XX and InX/X female embryos or between XY and InX/Y male embryos. These were therefore pooled as "XX" and "XY" for all future analyses. $X^P$O embryos were significantly smaller than XX littermates: this was clear from morphology (Figure 3.1D), weight data (Figure 3.1E), and somite counts (Table 3.1). I then compared $X^P$O embryos with XY littermates to confirm whether a single $X^M$ is able to support development more effectively than a single $X^P$, as shown previ-
3.2. Results

Previously (Thornhill and Burgoynes, 1993). XPO embryos are also significantly smaller than XY littermates (Figure 3.1D, E), and have fewer somites (Table 3.1). Taken together, these data confirm previous results showing XPO growth retardation at E10.5, relative to XX littermates. Furthermore, I observed that a single XM is more capable of supporting normal development than a single XP. This result does not, however, enable us to determine whether inappropriate $Xist$ expression or X linked gene imprinting is responsible for the reported embryonic growth deficit.

In order to test the $Xist$ hypothesis, I utilised a male carrying the $Xist^{tm1Jae}$ allele (Marahrens et al., 1997) to produce XPO mice lacking $Xist$ (hereafter denoted $XXist^{-}$). This allele is lethal to all XX females when inherited paternally because imprinted XCI is ablated, resulting in embryonic death and degradation by E10.5 (Marahrens et al., 1997). Accordingly, there are no XX females available as controls, and only three offspring genotypes are produced in this cross at E10.5 (Figure 3.1A(ii)). As an alternative, I used XYRIII male embryos as controls. Whilst most Y chromosomes accelerate pre-implantation development, the RIII strain Y chromosome has previously been shown to lack this effect (Burgoyne, 1993). In order to confirm there is no acceleration of development in XYRIII embryos, I compared to XX littermates in cross 1. No significant difference was observed in either weight (Figure 3.1E(i)) or somite data (Table 3.1). I therefore utilised XYRIII male embryos as controls throughout this experiment.

I then assessed the contribution of inappropriate $Xist$ expression to the growth deficit observed in XPO embryos at E10.5. If the cause of the growth deficit was $Xist$ expression, no difference in weight or somite count would be expected between XYRIII and $XXist^{-}$ embryos (Figure 3.1C(ii)). Most $XXist^{-}$ embryos showed growth retardation when compared to XYRIII littermates, which was clear both from embryo morphology (Figure 3.1D, right hand side), weights (Figure 3.1E(ii)), and somite counts (Table 3.1). Based on these data, I conclude that inappropriate expression of $Xist$ is not responsible for the growth retardation phenotype observed in XPO embryos at E10.5.
Embryonic growth controls the kinetics of somitogenesis, both in wildtype and growth impaired embryos (Tam, 1981; Power and Tam, 1993). When embryo size is halved prior to the eight cell stage, morphogenesis and pattern formation are appropriate for tissue volume as opposed to chronological age (Power and Tam, 1993). I was interested to see if a similar relationship holds for X\textsuperscript{P}O embryos, thereby gaining further insight into the nature of the growth deficit. I used log weight and somite count metrics to validate this relationship in the current model (Figure 3.1E(iii)): all four genotypes showed highly significant correlations (Figure 3.1E(iii)). Utilising the same data, I then created a linear regression model of developmental progression (Figure 3.1E(iv)). The gradients of the regression lines were not significantly different between any of the genotypes ($p = 0.3310$; Figure 3.1E(iv)). I conclude that the growth retardation and reduced somite counts represent a global developmental delay in X\textsuperscript{P}O embryos at E10.5, i.e. the somite count is appropriate for the log weight. This was also true for X\textsuperscript{$\text{Xist}$-O} embryos, further supporting my conclusion that $\text{Xist}$ does not play a role in the retardation phenotype.
3.2. Results

A
(i) 

(ii) 

B

C
(i) 

(ii) 

D

E
(i) 

(ii) 

(iii) 

(iv)
Figure 3.1: Xist KO does not rescue the developmental deficit observed in XPO embryos at E10.5. (A)(i) Cross 1 utilised a XY^{RIII} male and a In(X)1HY female, with five possible offspring: Y^{RIII}O embryos are not viable. (ii) Cross 2 utilised a X^{Xist-}Y^{RIII} male and a In(X)1HY female, with three possible offspring: females with two X chromosomes are not viable due to a lack of imprinted XCI. (B) Representative genomic DNA PCR results for the inversion X chromosome (top row); the Y chromosome (middle row); and the mutated Xist allele (bottom row). (C) (i) and (ii) illustrate experimental hypotheses. (D) Representative images of all genotypes at E10.5 (scale bar = 100 μM). (E) (i) The log weights of XX and XPO, XX and XY^{RIII}, XY^{RIII} and XPO embryos, and (ii) XY^{RIII} and X^{Xist-}O embryos were compared at E10.5, with P values given by Student’s t; bars represent mean±SEM; (iii) Somite count was significantly correlated with log weight for all genotypes (r^2 for XY^{RIII}=0.8173, XX=0.9245, XPO=0.8814, X^{Xist-}O=0.8248); (iv) linear regression analysis included somite count and log weight as factors, with no significant difference between the gradients for different genotypes (p=3.31x10^{-1}).

3.3 Discussion

Whilst it has been known for a number of years that XPO embryos are significantly smaller than XX littermates at E10.5 (Burgoyne et al., 1983b; Thornhill and Burgoyne, 1993; Jamieson et al., 1998), here I show that this is not as a result of inappropriate Xist expression. The Xist mutant allele has been shown to abolish Xist expression from the X^{P} (Marahrens et al., 1997). Nevertheless, XPO mice carrying Xist- still exhibit a significant growth deficit at E10.5.

There are a number of possible explanations that could underlie this result, from both technical and biological perspectives. Firstly, loss of Xist expression should be confirmed by transcriptional analysis of X^{Xist-} embryos. Secondly, significant variation exists between mouse litters, largely resulting from the uterine environment (Hughes, 1979; Holson & Pearce 1992; Lazic and Essioux, 2013). I reduced this effect by keeping the maternal genotype constant across experiments, and utilising the MWD statistical technique to compare embryos within litters. However, it was not possible to generate normal X^{P} females and X^{P} females lacking Xist within the same litter. I cannot therefore fully exclude the effect of inter litter differences. In the XY^{RIII} male I believe I opted for the best proxy comparison; however, due to the genetically outbred nature of the MF1 strain, it is formally possible that this is an
<table>
<thead>
<tr>
<th>Cross</th>
<th>Genotypes compared</th>
<th>Number of foetuses</th>
<th>Mean±SEM log weight (weight in mg)</th>
<th>Mean±SEM weighted difference</th>
<th>Significance of difference (p)</th>
<th>Mean somites±SEM (n)</th>
<th>Regression line gradient±SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>XX</td>
<td>vs. $X^P O$</td>
<td>14</td>
<td>0.644±0.118 (4.30)</td>
<td>0.243±0.066</td>
<td>6.30×10^{-5}</td>
<td>26.1±2.331 (10)</td>
<td>13.17±1.708</td>
</tr>
<tr>
<td>$XY^{R III}$</td>
<td>vs. $X^P O$</td>
<td>14</td>
<td>0.644±0.118 (4.30)</td>
<td></td>
<td></td>
<td>26.1±2.331 (10)</td>
<td>13.17±1.708</td>
</tr>
<tr>
<td>$XY^{R III}$</td>
<td>vs. XX</td>
<td>26</td>
<td>0.880±0.062 (7.08)</td>
<td></td>
<td></td>
<td>31±1.190 (9)</td>
<td>19.60±2.117</td>
</tr>
<tr>
<td>$XY^{R III}$</td>
<td>vs. $XX$</td>
<td>26</td>
<td>0.880±0.062 (7.08)</td>
<td>0.070±0.049</td>
<td>4.05×10^{-1}</td>
<td>31±1.190 (9)</td>
<td>19.60±2.117</td>
</tr>
<tr>
<td>$XY^{R III}$</td>
<td>vs. $X^{Xist-} Y^{R III}$</td>
<td>31</td>
<td>0.920±0.058 (8.31)</td>
<td></td>
<td></td>
<td>29.24±0.8948 (30)</td>
<td>16.38±1.130</td>
</tr>
<tr>
<td>$XY^{R III}$</td>
<td>vs. $XX$</td>
<td>26</td>
<td>0.880±0.062 (7.08)</td>
<td></td>
<td></td>
<td>31±1.190 (9)</td>
<td>19.60±2.117</td>
</tr>
<tr>
<td>$XY^{R III}$</td>
<td>vs. $X^{Xist-} Y^{R III}$</td>
<td>36</td>
<td>0.925±0.053 (8.41)</td>
<td></td>
<td></td>
<td>29.37±1.421 (19)</td>
<td>16.38±1.130</td>
</tr>
<tr>
<td>$XY^{R III}$</td>
<td>vs. $X^{Xist-} Y^{R III}$</td>
<td>14</td>
<td>0.678±0.0898</td>
<td></td>
<td></td>
<td>26.20±2.154 (10)</td>
<td>16.89±2.753</td>
</tr>
</tbody>
</table>

Table 3.1: Weight and somite comparison of E10.5 embryos
uncontrolled source of variation. A further source of variation lies within developmental timing between litters (Figure 3.1E); i.e. some litters were likely conceived many hours earlier than others, despite maintaining constant timing by embryonic day. Again, the mean weighted difference technique helped to ameliorate this effect in the data by making only relative comparisons within litters. Furthermore, the logistic regression analysis utilised both somite number and log weight, thus controlling for developmental timing.

3.3.1 Xist coating may not directly result in imprinted XCI

Previous work has shown by RNA FISH almost complete Xist coating of the paternally derived X chromosome(s) in X\textsuperscript{PO} and X\textsuperscript{PX}\textsuperscript{P} embryos at the 8-16 cell stage of pre-implantation development (Matsui et al., 2001; Okamoto et al., 2000). This expression was subsequently reduced at the blastocyst stage such that only around 20% of cells showed inappropriate numbers of Xist foci. Neither of these experiments, however, assayed whether this Xist expression causes X linked gene silencing.

Utilising a similar model system, but assaying instead by qRT-PCR, Latham and Rambhatla screened five X linked genes, including Xist, for differential expression between androgenetic and gynogenetic embryos at both 8 cell and late morula stages (Latham and Rambhatla, 1995). Androgenetic embryos showed reduced expression in all genes except Xist, which was more highly expressed. These results should, however, be explicated in the context of the experiment, which utilised pooled embryos and early qRT-PCR techniques. For example, a population of androgenetic embryos may include those with two X chromosomes, two Y chromosomes, or one of each, thus making interpretation of differential expression challenging. Later work from the same group rectified this ambiguity by using single embryos for experiments, though focusing on mid to late blastocyst stage embryos (Latham et al., 2000). There was a significant difference in Xist expression between XX and XY blastocysts in the control group. The same observation was made comparing X\textsuperscript{PX}\textsuperscript{P}
with X\textsuperscript{P}Y embryos, though the difference was not statistically significant. Finally, they reported no significant difference in \textit{Xist} expression between XX control and X\textsuperscript{P}X\textsuperscript{P} androgenetic blastocysts, despite the latter carrying two paternally derived X chromosomes (Latham et al., 2000). This is consistent with the RNA FISH data from androgenetic blastocysts (Okamoto et al., 2000). The embryo becomes competent to count at this stage, which effects a reduction in the expression of \textit{Xist} appropriately in both androgenetic and control blastocysts (Latham et al., 2000; Okamoto et al., 2000).

Taken together, these data leave open two possibilities. Firstly, imprinted XCI may take place in X\textsuperscript{P}O embryos and reduce transcription from the X\textsuperscript{P}. This would be in keeping with the fact that X linked nullizygosity, as seen in the YO embryo, is detrimental to pre-implantation embryonic development (Morris, 1968a; Burgoyne and Biggers, 1976). Alternatively, \textit{Xist} is expressed, but the downstream pathway that results in transcriptional inactivation is not complete. This implies that whilst imprinted expression of \textit{Xist} from the X\textsuperscript{P} is absolute, i.e. it is blind to the number of X chromosomes present, a downstream counting or regulatory mechanism must exist to ensure continued X linked gene expression. Hypothetically, such a mechanism could act on X\textsuperscript{M}, so that \textit{Xist} expression is only converted to transcriptional inactivation in the presence of at least one X\textsuperscript{P} and one X\textsuperscript{M}. A candidate for the effector of this mechanism is \textit{Rnf12} (also known as \textit{Rlim}). A maternally expressed copy of \textit{Rnf12} is absolutely required for imprinted XCI in XX female embryos, as shown by the early post implantation female lethality of an \textit{Rnf12} mutant allele (Shin et al., 2010). These \textit{Rnf12} mutant embryos do, however, also display defective \textit{Xist} expression at the blastocyst stage, and this defect has not been observed in X\textsuperscript{P}O embryos (Shin et al., 2010).

In order to distinguish between these two hypotheses, an analysis of X linked transcription in the early embryo is required. RNA FISH would enable a number of genes to be screened, but RNA-seq would provide a more comprehensive answer. Recently, single cell RNA-seq has facilitated the study of gene transcription in the
3.3. Discussion

early embryo in even greater depth, thus providing a transcriptome wide view of iXCI. Transcription from XP in wildtype XX embryos has been shown equivalent to that from XM at the 4 cell stage, with a paternal to maternal ratio of 1.0. (Petropoulos et al., 2016). The ratio then progressively decreased to 0.4 at the early blastocyst stage (Petropoulos et al., 2016). Xist expression peaked at the 16 cell stage and was maintained until E4.5 (Deng et al., 2014; Wang et al., 2016). Borensztein and colleagues further classified the genes undergoing inactivation into early, intermediate and late silenced, at 16, 32 and blastocyst stages, respectively (Borensztein et al., 2017). As these data are entirely consistent with previously reported RNA FISH datasets, a complementary approach of RNA-seq validated by RNA FISH is ideally suited to begin to address the role of Xist in imprinted XCI in XP/O pre-implantation embryos.

3.3.2 The X chromosome may harbour imprinted genes that are primarily expressed in the extraembryonic tissue

Inappropriate Xist expression is not the cause of the growth retardation observed in XP/O embryos at E10.5, therefore I hypothesise the existence of imprinted genes on the X chromosome that are differentially expressed during embryonic development. This could also serve to explain the observation that XM/O female embryos show accelerated growth relative to XX littermates at E10.5 (Thornhill and Burgoyne, 1993). Such imprinted genes could function either as a normally silent growth suppressor(s) on the single XP, or as a growth enhancer(s) from the XM. There are a number of reported examples of X linked genes proposed to be silenced on XP due to imprinting, e.g. Esx1 (Li and Behringer, 1998) and Plac1 (Jackman et al., 2012). However, whether the silencing observed in these reports results from true imprinting or imprinted XCI remains to be shown. In order to separate the conceptually similar occurrences of imprinted XCI and genomic imprinting, the presently described XO model is required. Imprinted XCI does not occur in this model, thus any differences observed between XP and XM can be attributed to genomic imprint-
ing. This approach led to the discovery of the Xlr locus, a group of imprinted genes expressed in the developing mouse brain (Davies et al., 2005; Raefski and O’Neill, 2005).

Whilst imprinting could affect genes expressed in the embryonic or in the extraembryonic tissue, a number of lines of evidence make the latter more likely. Firstly, embryos that do not show growth retardation, i.e. XMXp, XMpY, and XMpO have an active XM in extraembryonic tissue, whereas XPpO has an active XP. Secondly, one of the recurring themes in the genomic imprinting literature is the correlation between the evolution of imprinting and the evolution of an invasive, gestation spanning placenta (Cleaton et al., 2014; Renfree et al., 2013). A significant number of the 100 or so currently recognised imprinted genes affect prenatal development through placental expression (Coan et al., 2005; Keverne, 2015). Thirdly, as previously described here, the placentas of XO embryos show morphological abnormalities during development. XPpO embryos show reduced EPC volume when compared to XX littermates, while XMpO embryos do not (Jamieson et al., 1998). Earlier work utilising the XO mouse model showed that XPpO placenta are larger than placenta from XX littermates, and further work described this hyperplasia in both XPpO and XMpO placenta at E17.5, with the latter larger than the former (Burgoyne et al., 1983a; Ishikawa et al., 2003). The authors reconcile this apparent contradiction by suggesting that the growth trajectories are simply maintained from EPC onwards, and that placenta with one X chromosome are generally larger than those with two X chromosomes (Ishikawa et al., 2003). Finally, XO mESCs have been shown to exit the pluripotent state and differentiate faster than XX mESCs, which must trigger XCI before exit from pluripotency (Schulz et al., 2014). If this also applies to the situation in vivo, it might be expected that the XO embryo proper is actually at an advantage relative to the XX embryo proper, at least initially.

In order to determine whether an embryonic or extraembryonic defect underlies the growth deficit in XPpO embryos, the definitive experiment would require an analysis of embryos in which sex chromosome complement differs between these two
3.3. Discussion

By introducing an $X^P O$ embryo into an XX/XY extraembryonic environment, and *vice versa*, the effects on development could be assessed. This can be achieved by either mechanical or chemical means. The cells of a two cell stage embryo can be electrically fused, thus making the cell tetraploid (4n). These cells can then be complemented at the four cell stage with donor mESCs, resulting in the creation of high grade chimeras. The recipient 4n cells form the extraembryonic compartment, and the mESCs differentiate into the embryo proper (Nagy et al., 1990, 1993). Alternatively, recipient embryos can be pre-treated with FGF4 and heparin in order to promote formation of PE and inhibit the development of the epiblast (Yamanaka et al., 2010). Very recently it has been shown that donor mESCs can be injected into FGF/heparin pre-treated embryos at the blastocyst stage, and that these cells make a significant contribution to the formation of the embryo proper (Dupont et al., 2017). There are compromises with both techniques. Firstly, tetraploid complementation has been correlated with a growth deficit in all wild-type embryos (Tam, 2003). More importantly, the karyotype of the extraembryonic tissue would be 4n, thus making it no longer possible to assess the contribution of a single X chromosome. FGF treatment does not currently achieve equivalent rates of chimerism to the tetraploid method. However, with appropriate controls, it should still be possible to detect relative changes to embryonic development. These experiments are in progress.

3.3.3 Mouse trophoblast stem cells can be utilised to identify imprinted genes *in vitro*

A tractable model is required to facilitate the search for any X linked imprinted genes expressed in the extraembryonic compartment. Undoubtedly the gold standard would be to show differential gene expression (DGE) between $X^P$ and $X^M$ chromosomes *in vivo*, followed by functional experiments to link cause and effect. However, the placenta is a dynamic tissue that changes significantly during development, therefore identifying the correct time and location for DGE may be chal-
lenging without an initial screen. This could be carried out in mouse trophoblast stem cells (mTSCs): an *ex vivo* derived population of multipotent cells originating from the extraembryonic compartment, amenable to propagation *in vitro*, manipulations such as overexpression or genome editing, and shown to maintain imprints originally observed in the embryo (Tanaka et al., 1998; Calabrese et al., 2015). The derivation and use of mTSCs to identify X linked imprinted genes will be the focus of the next chapter.
Chapter 4

Using *in vitro* models to identify candidates for X linked imprinted genes in mouse extraembryonic tissue

4.1 Introduction

4.1.1 Genomic imprinting

Parent of origin biased gene expression - genomic imprinting - has been observed across fauna and flora, though among vertebrates it is most well characterised in eutherian mammals. Early studies utilising androgenetic and gynogenetic mouse embryos established that one copy of the genome from each parent was required for successful embryonic development (Barton et al., 1984; Kaufman et al., 1977; Surani and Barton, 1983; Surani et al., 1984). Subsequent work showed that this uniparental embryonic lethality phenotype resulted from a necessity for expression of a small subset of genes from a specific parental allele, denoted *paternally expressed genes, Pegs* and *maternally expressed genes, Megs*, respectively (Cleaton
et al., 2014). These subsets are most often found in groups known as imprinted domains, which themselves are coordinately regulated by a germline differentially methylated region (gDMR) or imprinting control region (ICR). gDMRs are CpG rich, and the two parental alleles are differentially methylated during gametogenesis (Iurlaro et al., 2017). The majority of gDMRs are maternally methylated, though a small number are paternally methylated, including the \textit{Igf2-H19} cluster. Whilst it is clear that such differential methylation drives the imprinting status of the cluster, there are also a number of other mechanisms by which allele specific gene expression can be achieved, including lncRNA mediated recruitment of epigenetic modifiers (Nagano et al., 2008).

The biological purpose of imprinting is unknown; though the timing of its appearance in both eutherian mammals and, to a limited extent, metatherians has been correlated with the evolution of viviparity through placentation (Smith et al., 2011). It has previously been suggested that imprinting regulates the balance between the demands of the maternal and paternal genomes via the \textit{in utero} foetus; the kinship, or parental conflict hypothesis (Haig and Westoby, 1989). This posits that paternally derived genes generally enhance foetal growth because they look to increase resource grabbing from the mother, thus maximising the probability of further genetic propagation. Maternally derived genes, in contrast, often suppress foetal growth in order to balance resources between current and future litters (Haig and Westoby, 1989).

Alternatively, imprinting may serve to regulate the balance of resources \textit{in utero} between mother and foetus, with maternal interests represented by the placenta (Reik et al., 2003). By this hypothesis, nutrient supply is genetically controlled by expression of imprinted genes in the placenta, and demand is controlled by expression of imprinted genes in the foetus. Such a hypothesis also accounts for the evolution of clusters, suggesting that the existence of supply and demand genes in close proximity (e.g. the region around \textit{Igf2}) facilitates the fine balance required in expression during development (Reik et al., 2003).
A third theory suggests that imprinting evolved because it increases the genetic integration of co adapted maternal and offspring traits, thereby enhancing the fitness of offspring (Wolf and Hager, 2006). No single theory fully explains the existence of all imprinted genes; though it is clear that all theories take into account the evolution of the placenta and the significant number of imprinted genes expressed by the shortest lived of all organs.

### 4.1.2 The placenta

The placenta is the physiological interface between mother and foetus (Smith et al., 2011). It mediates the exchange of nutrients and waste products between foetal and maternal blood, as well as acting as a centre for the production of hormones and growth factors during pregnancy (Rossant and Cross, 2001). Whilst the placenta is most commonly associated with eutherian mammals, it is also present in metatherians, and has been found in diverse vertebrate classes including amniotes such as Amphibia, Reptilia, and the anamniotic Chondrichthyes (sharks, Blackburn and Flemming, 2011; Hamlett, 1989). In mammals, and amniotes more generally, placental development includes the fusion of pairs of the four foetal membranes, the amnion, the yolk sac, the allantois, and the chorion (Smith et al., 2011). The amnion directly surrounds the foetus and serves to protect it from shock. The yolk sac encapsulates the egg yolk in egg bearing species, and carries out many duties of the placenta before the allantoic placenta is established, including blood synthesis and blood vessel formation and nutrient and gas exchange. The allantois is continuous with the embryonic bladder; and the chorion is formed by the extraembryonic mesoderm and a layer of outer trophoblast - it joins with the yolk sac to form the chorio vitelline placenta or the allantois to form the chorio allantoic placenta (Smith et al., 2011).

The mature mouse placenta is a chorio allantoic placenta, composed of multiple cell types across three layers: i) the decidua, formed of maternal tissue; ii) the junctional zone, and; iii) the labyrinth (John and Hemberger, 2012). Within the
labyrinth, exchange of nutrients, waste, and gases takes place between maternal and foetal circulations. The exchange interface is made up of two layers of multinucleated syncytiotrophoblast cells, in addition to sinusoidal trophoblast giant cells (TGCs, Simmons et al., 2008). TGCs also exist within the decidua, in the form of spiral artery associated TGCs and canal associated TGCs, and between the layers as parietal TGCs. They produce a number of endocrine molecules, including prolactin, that serve to mediate the continuation of pregnancy, maternal vascular remodelling, and possibly parturition (Hu and Cross, 2010). Placental prolactins are also produced by spongiotrophoblast cells from within the junctional zone, in addition to other pregnancy specific glycoproteins thought to have a role in the immune privileged status of the foetus (Wu et al., 2008). The function of the spongiotrophoblast within the junctional zone was previously thought to be primarily structural (Rossant and Cross, 2001), however, in addition to the production of hormones and glycoproteins, it also produces molecules that affect the maternal brain to mediate maternal instinct (Glynn and Sandman, 2011). Glycogen cells are also found within the junctional zone and, by migration, also within the maternal decidua, though the function of these aptly named glycogen heavy cells remains to be found (Prudhomme and Morey, 2015). A summary of mouse placenta development is provided in Figure 4.1.
4.1. Introduction

Figure 4.1: Mouse placenta development. Around the time of implantation (E3.5), the trophoectodermal cells overlying the ICM, the polar TE, proliferate and give rise to diploid extraembryonic ectoderm and the ectoplacental cone (EPC). The trophoectodermal cells in contact with the blastocoele cavity, the mural TE, endoreduplicate their DNA to become polyploid primary trophoblast giant cells (TGCs). Secondary TGCs are formed from the EPC, which migrate to surround the conceptus. Between E6.0 and E7.5, the extraembryonic ectoderm expands and joins with extraembryonic mesoderm to form the chorion, which itself then makes contact with the mesoderm derived allantois. This attachment is followed by folding of the chorion, thus marking the sites of foeto placental blood vessel growth from the allantois. By E12.5, the foetal blood vessels become a densely packed labyrinth with extensive villous branching, and chorionic trophoblast cells differentiate to become syncytiotrophoblast cells. These multinucleated cells form two distinct layers which, in addition to sinusoidal TGCs, form the interface between the maternal and foetal blood within the placental labyrinth. In order to make contact with the foetal blood supply via the trophoblast villi, the maternal blood passes through the junctional zone, comprised mainly of spongiotrophoblast cells. This is formed by differentiation from the EPC. After Rossant and Cross, 2001.

Many imprinted genes are expressed in the placenta (Cleaton et al., 2014; Tunster et al., 2013). A number of these imprinted genes have been functionally characterised by targeted deletion, often resulting in growth phenotypes affecting either the placenta and/or the foetus. Placental growth restriction has been reported following individual deletion of the paternally expressed genes Igf2, Peg1, and Plag1; in contrast to deletion of the maternally expressed Igf2r, Cdkn1c, Phlda2 and Grb10, which resulted in enhanced placenta growth (Cleaton et al., 2014). Foetal overgrowth was observed following targeting of the maternally expressed Igf2r or H19 (Wang et al., 1994; Leighton et al., 1995). An embryonic growth retardation phenotype was reported after deletion of Igf2, Peg1/Mest, and Peg3, all normally ex-
pressed from the paternal allele (DeChiara et al., 1991; Lefebvre et al., 1998; Li et al., 1999).

\(^{X^P}\)O female embryos are smaller than XX littermates at E10.5 (Thornhill and Burgoyne, 1993). In this thesis, I previously showed that this growth deficit is not caused by \(Xist\) expression (Chapter 3). Furthermore, \(^{X^P}\)O embryos also have a smaller ectoplacental cone (EPC) than XX littermates at E7.5 (Jamieson et al., 1998). Neither of these defects are present in \(^{X^M}\)O embryos, which are heavier than XX littermates at E10.5 (Thornhill and Burgoyne, 1993; Jamieson et al., 1998). This \(^{X^P}\)O growth deficit is likely to result from genomic imprinting affecting the X chromosome.

\(XCI\) always silences the \(^X^P\) in mouse extraembryonic tissue, whereas in the embryo, \(XCI\) randomly affects either \(^X^M\) or \(^X^P\) (Okamoto et al., 2004). Genes from the \(^X^P\) are therefore expressed in 50% of cells in the normal XX embryo proper, but are never expressed in the XX, XY, or \(^X^M\)O placenta. I hypothesise that the absence of the \(^X^M\) in \(^{X^P}\)O placentas could result in loss of expression of a growth promoting gene(s) expressed only from \(^X^M\) in placenta. Alternatively, there may be a gene(s) present on \(^X^P\) that is usually silenced by \(XCI\) in the XX placenta, and whose expression in the \(^{X^P}\)O is disadvantageous to embryonic growth. It is also formally possible, though less likely, that the phenotype could be explained by genes expressed in the embryo proper. In order to search for candidate genes to support these hypotheses, I performed RNA sequencing (RNA-seq) on \(^{X^P}\)O and \(^{X^M}\)O mouse stem cell lines to model the different embryonic compartments: pluripotent - mouse embryonic stem cells (mESCs); epiblast - mouse epiblast like cells (mEpiLCs); and the trophectoderm - mouse trophoblast stem cells (mTSCs).
4.2 Results

4.2.1 Derivation of stem cell lines

First I set out to derive mTSC lines from blastocyst stage embryos, as described previously (Tanaka et al., 1998; Himeno et al., 2008) and summarised in Figure 4.2A. In order to generate X^P^O embryos, I used the cross described in Chapter 3 between an In(X)/X female and a XY^R^III male (Figure 4.2B(i)), here designated Cross I. X^M^O embryos were generated by crossing an X^Y^*O male (Burgoyne et al., 1998) with a wildtype XX female (Figure 4.2B(ii)), hereafter Cross II. After ten passages supported by mouse embryonic fibroblast (MEF) feeders, mTSC lines stabilised. Feeders were removed, cell lines were maintained in MEF conditioned media, and genotyping was carried out. PCR revealed the presence of only female mTSC lines from Cross I. No XY or InX/Y lines were derived (Figure 4.2C(i)). Three of these female mTSC lines were determined to be X^P^O, with the presence of a single X chromosome shown by DNA FISH (Figure 4.2D) and qRT-PCR. Four mTSC lines were derived from Cross II outgrowths (Figure 4.2C(ii)), and two of these lines were determined to be X^M^O, based on PCR and DNA FISH (Figure 4.2D).

Transcription factors including Cdx2, Eomes, Esrrb, Elf5, and Tfap2c, and the growth factor Bmp4, have previously been shown to be highly expressed in mTSCs (Tanaka et al., 1998; Kubaczka et al., 2014). In order to confirm mTSC identity, expression levels of these key genes were assayed by qRT-PCR, and compared with expression levels from the EGFP mTSC line, originally derived in the Rossant laboratory and here designated Rossant (Tanaka et al., 1998). All newly derived cell lines showed equivalent expression to the Rossant line (Figure 4.2E), as measured by log2 fold change, with the exception of lines X^M^O 2 and X^P^O 1. These two XO lines had increased expression for most of the genes assayed. I conclude that I had successfully derived X^P^O and X^M^O mTSC lines, in addition to female control lines.
4.2. Results

A

Plate blastocyst in FGF4/heparin

Day 1

Dissociate outgrowth

Day 4-6

P1

Day 14-28

Plate feeder-free DNA FISH/PCR

Day 75-100

tRT-PCR

RNA-seq

DNA FISH

Remove FGF4/heparin

Pick and passage (−P10)

Differentiation assay

B

(i)

X

InX

O

(ii)

X

X

XYIII

YIII

XX

InX/X

XPO

XYIII

InX/YIII

YYIII

XX

XX

XX

XMO

XMO

C

(i)

5

3

3

0

0

0

(ii)

2

2

D

XX

XMO

XPO

E

Log2-fold change
4.2. Results

Figure 4.2: Derivation of mouse trophoblast stem cells (mTSCs). (A) Schematic depicting derivation process. (B) (i) Mouse cross used to generate XX and X<sup>M</sup>O blastocysts, and (ii) X<sup>M</sup>O blastocysts. (C) (i) Number of cell lines derived, corresponding to genotypes in (B) (i), and (ii) corresponding to (B) (ii). (D) (i) Brightfield microscopy showing feeder free mTSCs of different sex chromosome genotypes (scale bar = 100μM); (ii) Representative images of DNA FISH probing for sex chromosomes (scale bar = 10μM). 50 cells were counted per cell line, with 80% agreement required to denote a sex chromosome genotype. (E) Heatmap summarising qRT-PCR results following screening of all cell lines for mTSC marker genes. Results were first normalised to Gapdh and Srp72, before log-fold change was determined based on a reference cell line (Rossant).

4.2.2 X<sup>P</sup>O, X<sup>M</sup>O, and XX mTSCs show equivalent differentiation potential in vitro

The EPC originates from the polar trophectoderm, and later gives rise to the spongiotrophoblast in the mature placenta (Cross et al., 1994). Previous work has shown that the EPC in X<sup>P</sup>O conceptuses is deficient in size at E7.5 relative to XX littermates (Jamieson et al., 1998). In contrast, X<sup>M</sup>O EPC was morphologically normal (Jamieson et al., 1998). As the EPC represents a relatively early, transient progenitor population of cells in the placenta, I was interested to see if the in vivo phenotype manifests as a differentiation bias in X<sup>P</sup>O mTSCs in vitro. I hypothesised that, in an undirected differentiation protocol, X<sup>P</sup>O mTSCs might be less likely to form the differentiation products of EPC, i.e. spongiotrophoblast cells, when compared with X<sup>M</sup>O and XX mTSCs.

In order to test this hypothesis, three X<sup>P</sup>O, two X<sup>M</sup>O, three XX, and the Rossant EGFP mTSC lines were differentiated by removal of FGF4 and heparin. I first noted that there were no obvious morphological differences between mTSC genotypes during the differentiation (Figure 4.3). To assay for the presence of cells from different placental lineages within the differentiated cultures, I performed qRT-PCR. If X<sup>P</sup>O differentiated mTSC cultures are deficient in EPC cells and other cell types downstream, I would expect to detect reduced expression of markers for these cell types, relative to XX differentiated mTSC cultures or the Rossant EGFP control. I might also observe increased expression of markers for other placenta lineages, as
when Ascl2 function is ablated *in vivo*, trophoblast giant cells predominate (Guillemot et al., 1994, 1995).

I first looked to confirm the effectiveness of the differentiation protocol by assaying expression of the key mouse trophoblast marker gene Cdx2 (Tanaka et al., 1998; Niwa et al., 2005). All mTSC genotypes (X_P, X_M, and XX) showed downregulation of Cdx2 expression at day two of differentiation when compared to the Rossant XX EGFP line at day 0 (Figure 4.4) and this was maintained throughout the six day time course. For each individual timepoint, there were no significant differences in Cdx2 expression between mTSC genotypes. I conclude that mTSC genotype does not affect the ability of the cell to exit the multipotent state, as tested by Cdx2 expression.

Next I tested the ability of the three mTSC genotypes to differentiate into spongiotrophoblast via EPC, compared with the Rossant XX EGFP line. EPC is characterised by expression of the basic helix loop helix (bHLH) transcription factor Ascl2/Mash2 (Guillemot et al., 1994, 1995), and one of the key markers of the spongiotrophoblast is Tpbpa/4311 (Carney et al., 1993). Notably, I observed increased expression of Ascl2 across all mTSC genotypes at day two (Figure 4.4), consistent with the downregulation of Cdx2 and exit from the stem cell state. There was, however, no significant difference in Ascl2 between genotypes. I also observed upregulation of the spongiotrophoblast marker Tpbpa, which continued throughout the differentiation timecourse, though again there was no significant difference between genotypes. Finally, I found downregulation of the glycogen trophoblast marker Pcdh12 across all cell lines at day two, and this level of expression was maintained through to day six. There was no significant difference between the mTSC genotypes. From these data I conclude that X_P or X_M mTSCs do not show a differentiation bias away from EPC, or its derivative spongiotrophoblast, when compared to Rossant XX EGFP mTSCs.
Although there was no difference in differentiation potential between X<sup>P</sup>O and XX mTSCs towards the spongiotrophoblast lineage, I continued to investigate differential potential towards other placenta lineages.

In order to assay trophoblast giant cell (TGC) differentiation potential, I first measured *Hand1* expression. *Hand1* is a bHLH transcription factor co-expressed with *Ascl2* in the EPC (Scott et al., 2000), though these two genes have mutually antagonistic functions in differentiation. As described above, *Ascl2* maintains spongiotrophoblast, whereas *Hand1* is required for giant cell differentiation. Unsurprisingly, in mTSCs *Hand1* showed a similar expression pattern to *Ascl2*, with upregulation at day two, which then reduced almost back to baseline at day six (Figure 4.4). There was no significant difference in *Hand1* expression between X<sup>P</sup>O, X<sup>M</sup>O and XX mTSCs, though all three genotypes showed lower expression than the Rossant XX EGFP line at all time points.

I next looked for the presence of markers for the terminally differentiated TGC cell types, including sinusoidal TGCs (s-TGCs) found lining the maternal sinusoids (Simmons et al., 2007). *Pl1, Prl2c2*, and *Prl3d3* are expressed in TGCs (Simmons et al., 2007), and were upregulated from day four of differentiation in all cell lines (Figure 4.4). These three markers were generally less expressed in XX mTSCs than both X<sup>P</sup>O and X<sup>M</sup>O mTSCs, though the differences were not statistically significant. *Ctsq*, a marker of s-TGCs, showed slight upregulation at day two, and became highly upregulated on days four and six. There was no significant difference in *Ctsq* expression between mTSC genotypes. I conclude that, whilst the data suggest lower expression of TGC markers in XX mTSCs when compared to XO mTSCs and Rossant XX EGFP mTSCs, these differences were not significant. Therefore I found no difference in the expression of giant cell lineage markers between XX, X<sup>P</sup>O, X<sup>M</sup>O, and Rossant XX EGFP mTSCs. Notably, both X<sup>P</sup>O and X<sup>M</sup>O mTSC lines generally showed larger confidence intervals than those for XX mTSCs.

Finally I assayed for syncytiotrophoblast (Syn) differentiation. As a marker for Syn progenitors, I used *Rhox4b/Ehox* (Watson et al., 2011). Expression relative
to Rossant XX EGFP was reduced in all cell lines at day zero, and this difference persisted throughout differentiation (Figure 4.4). Rhox4b was downregulated in all mTSC lines and, notwithstanding the difference when compared to Rossant XX EGFP, there was no difference in expression between XX, X^pO and X^M_O lines.

Coincident with the reduction in Syn progenitors, expression of markers for terminally differentiated syncytiotrophoblast-I (SynT-I) increased at day two, and this increase persisted through the differentiation timecourse (Figure 4.4). Syna (Simmons et al., 2008) showed marked upregulation at day two, which was maintained, whereas Snai1 (Watson et al., 2011) showed a small upregulation at day two, which continued to day six. Expression of Snai1 in Rossant XX EGFP mTSCs was slightly higher than XX, X^pO and X^M_O mTSCs, though there was no difference between the latter three genotypes. Expression of Synb, a marker for syncytiotrophoblast-II (SynT-II, Simmons et al., 2008), was decreased in all cell lines relative to Rossant XX EGFP at day 0. Synb transcript levels then increased incrementally in all cell lines to day six. Interestingly, expression of a second marker of SynT-II, Gcm1, decreased in all mTSC lines on day two, and continued to decrease in XX and Rossant XX EGFP mTSC lines through to day six. Expression of Gcm1 in X^pO and X^M_O cell lines troughed at day two this level was maintained relative to Rossant XX EGFP day 0. From these data I was able to conclude that there was no difference between X^pO, X^M_O, and XX mTSCs in their ability to differentiate to all placenta cell lineages.
4.2. Results
Figure 4.3: Differentiation of mTSCs. Representative brightfield images of mTSC lines with different sex chromosome genotypes following a six day differentiation timecourse (scale bars = 100μM).
Figure 4.4: qRT-PCR analysis of gene expression during mTSC differentiation. Data are presented in the form of a differentiation tree, whereby Cdx2, a marker for the stem cell lineage, is at the base; with the three main lineages of the mouse placenta branching out. Each graph shows the mean logFC and 95% CI at days 0, 2, 4 and 6 for each sex chromosome genotype per gene, relative to day 0 Rossant EGFP. Data represent values obtained from three independent XX mTSC lines, three independent X<sup>P</sup>O mTSC lines, and two independent X<sup>M</sup>O mTSC lines.

4.2.3 Derivation of mESC and mEpiLC lines and qRT-PCR verification of identity

I considered it most likely that X linked imprinting candidates would be expressed in the extraembryonic tissue (as described in 4.1), making mTSC lines the most relevant for my screening experiment. However, I also wanted to model the embryo proper as an alternative hypothesis. To this end, I derived mESC lines to capture the pluripotent compartment, and differentiated these lines into mEpiLCs to model the epiblast. X<sup>P</sup>O and X<sup>M</sup>O mESC lines were derived from blastocysts from cross 1 and cross 2, respectively. Each embryo was plated individually in 2i+LIF, as previously described (Ying et al., 2008). Outgrowths were picked after 10-14 days and, concurrent with genotyping, cell lines were expanded to passage five (Figure 4.5A). DNA FISH was used to confirm sex chromosome genotype, and three mESC lines were maintained for each genotype, including XX and XY control lines (Figure 4.5B). A corresponding, isogenic mEpiLC line was differentiated from each mESC line using bFGF and Activin A (Figure 4.5A), as described (Hayashi et al., 2011b; Hayashi and Saitou, 2013). EpiLCs have been shown to represent the epiblast of the mouse blastocyst stage embryo, thus downregulating pluripotency markers including Sox2 and Nanog, and upregulating genes associated with exit from pluripotency, including Dnmt3b and Fgf5 (Hayashi et al., 2011b). I compared expression of these genes in mEpiLCs with the mESCs from which they were derived by qRT-PCR, and noted downregulation of pluripotency associated genes and upregulation of epiblast associated genes (Figure 4.5C). These results were entirely consistent with the literature, and were taken to confirm the lineage identity of the derived cell lines.
4.2. Results

A

Plate blastocyst in 2i/LIF

Dissociate outgrowth (P1)

DNA FISH, qRT-PCR, RNA-seq

Remove 2i, add bFGF, gsk3

RNA-seq

Day 1

Days 10-14

Days 19-29

Days 20-31

PCR, expand and freeze

EpEpLC differentiation

B

mESC

mEpiLC

DNA FISH

XX

XY

X'Y

X'O

Nanog
Sox2
Esrb
Pou5f1
Fgf5
Wnt3
Dmnt3b

Log2-fold change

-5

0

5
4.2. Results

Figure 4.5: Derivation of mouse embryonic stem cells (mESCs) and mouse epiblast like stem cells (mEpiLCs). (A) Schematic depicting derivation process. (B) Representative brightfield images of mESCs and the resultant mEpiLCs following differentiation (scale bars = 50uM), alongside DNA FISH used for genotyping. 50 cells were counted per cell line, with 80% agreement required to denote a sex chromosome genotype. (C) Heatmap summarising qRT-PCR data comparing mEpiLC lines with their originating mESC line for mEpiLC and mESC marker gene expression. Results were first normalised to Gapdh and Srp72, before log fold change was determined based on the originating mESC line for each mEpiLC line.

4.2.4 RNA-seq analysis of gene expression in mTSC, mESC and mEpiLC lines

In order to look for genes differentially expressed between X^PO and X^MO cell lines, and thus identify candidates that may be imprinted, RNA-seq was performed on RNA isolated from all mTSC, mESCs and mEpiLCs. Sequencing was performed by the Advanced Sequencing Facility STP at Francis Crick Institute, and returned between 39-54x10^6 paired end, strand specific reads per sample (Table 4.1). Data were analysed and plotted by a Turner lab post doc, Mahesh Sangrithi, as per the pipeline detailed in Figure 4.6A.

To gain an initial, graphical overview of how the samples related to one another, and specifically look at the relationship between biological replicates, multi dimensional scaling (MDS: Figure 4.6B) and t-stochastic neighbour embedding (t-SNE: Figure 4.6C) analyses were performed. Both techniques showed tight clustering of samples by cell type and, pleasingly, within cell type clustering by genotype. Further analysis of the mTSC data by unsupervised hierarchical clustering (UHC) showed with greater resolution that XX mTSCs clustered with the XX Rossant reference line, and separately from all XO mTSC lines (Figure 4.6D). Within XO mTSCs, all X^PO lines were shown to be more closely related to one another than to the X^MO mTSCs lines.
# 4.2. Results

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Sample</th>
<th>Raw reads</th>
<th>Percentage mapped</th>
</tr>
</thead>
<tbody>
<tr>
<td>mESC</td>
<td>XX 2</td>
<td>41552922</td>
<td>87.23</td>
</tr>
<tr>
<td></td>
<td>XX 1</td>
<td>43290415</td>
<td>88.83</td>
</tr>
<tr>
<td></td>
<td>XX 0</td>
<td>48235209</td>
<td>88.81</td>
</tr>
<tr>
<td></td>
<td>XY 2</td>
<td>33649772</td>
<td>91.26</td>
</tr>
<tr>
<td></td>
<td>XY 1</td>
<td>39843527</td>
<td>86.18</td>
</tr>
<tr>
<td></td>
<td>XY 0</td>
<td>52491568</td>
<td>85.55</td>
</tr>
<tr>
<td></td>
<td>XMO 2</td>
<td>50469556</td>
<td>89.3</td>
</tr>
<tr>
<td></td>
<td>XMO 1</td>
<td>40778312</td>
<td>78.15</td>
</tr>
<tr>
<td></td>
<td>XMO 0</td>
<td>46829273</td>
<td>87.38</td>
</tr>
<tr>
<td></td>
<td>XPO 1</td>
<td>46481635</td>
<td>87.34</td>
</tr>
<tr>
<td></td>
<td>XPO 0</td>
<td>40710940</td>
<td>93.3</td>
</tr>
<tr>
<td>mEpiLC</td>
<td>XX 2</td>
<td>55837430</td>
<td>90.82</td>
</tr>
<tr>
<td></td>
<td>XX 1</td>
<td>45576928</td>
<td>92.35</td>
</tr>
<tr>
<td></td>
<td>XX 0</td>
<td>50715628</td>
<td>90.91</td>
</tr>
<tr>
<td></td>
<td>XY 2</td>
<td>51420726</td>
<td>91.85</td>
</tr>
<tr>
<td></td>
<td>XY 0</td>
<td>50108108</td>
<td>91.91</td>
</tr>
<tr>
<td></td>
<td>XMO 3</td>
<td>33436541</td>
<td>90.62</td>
</tr>
<tr>
<td></td>
<td>XMO 2</td>
<td>46689812</td>
<td>91.04</td>
</tr>
<tr>
<td></td>
<td>XMO 1</td>
<td>53982073</td>
<td>91.63</td>
</tr>
<tr>
<td></td>
<td>XMO 0</td>
<td>48211831</td>
<td>91.39</td>
</tr>
<tr>
<td></td>
<td>XPO 2</td>
<td>47362417</td>
<td>91.36</td>
</tr>
<tr>
<td></td>
<td>XPO 0</td>
<td>60231527</td>
<td>90.94</td>
</tr>
<tr>
<td>mTSC</td>
<td>XX 7</td>
<td>41863904</td>
<td>91.89</td>
</tr>
<tr>
<td></td>
<td>XX 6</td>
<td>44532910</td>
<td>92.42</td>
</tr>
<tr>
<td></td>
<td>XX 5</td>
<td>39112596</td>
<td>93.35</td>
</tr>
<tr>
<td></td>
<td>XX 4</td>
<td>41064513</td>
<td>91.62</td>
</tr>
<tr>
<td></td>
<td>XX 3</td>
<td>31578074</td>
<td>94.77</td>
</tr>
<tr>
<td></td>
<td>XX 2</td>
<td>37020197</td>
<td>92.7</td>
</tr>
<tr>
<td></td>
<td>XX 1</td>
<td>41419287</td>
<td>92.11</td>
</tr>
<tr>
<td></td>
<td>XX 0</td>
<td>39396116</td>
<td>93.07</td>
</tr>
<tr>
<td></td>
<td>Rossant</td>
<td>42482645</td>
<td>93.08</td>
</tr>
<tr>
<td></td>
<td>XMO 2</td>
<td>52206246</td>
<td>91.37</td>
</tr>
<tr>
<td></td>
<td>XMO 0</td>
<td>54152268</td>
<td>90.27</td>
</tr>
<tr>
<td></td>
<td>XPO 3</td>
<td>63225432</td>
<td>91.23</td>
</tr>
<tr>
<td></td>
<td>XPO 1</td>
<td>54999428</td>
<td>91.46</td>
</tr>
<tr>
<td></td>
<td>XPO 0</td>
<td>43603825</td>
<td>90.46</td>
</tr>
</tbody>
</table>

**Table 4.1:** RNA-seq mapping percentages
4.2. Results

A. Workflow diagram showing the steps from raw fastq files to downstream validation.

- Quality control and trimming
- Align to gcm38
- Convert uniquely-mapped reads back to fastq
- Determine transcript abundances
- DGE analysis

B. Heatmap of leading logFC dimension 1 vs. leading logFC dimension 2.

C. t-SNE plot with dimension 1 vs. dimension 2.

D. Hierarchical clustering dendrogram.
4.2. Results

Figure 4.6: Initial analyses of RNA-seq on mTSC, mESC and mEpiLC. (A) Schematic of bioinformatic analysis pipeline: tasks are shown in orange and software utilised is shown in purple. (B) Multi-dimensional scaling plot of all RNA-seq datasets. (C) T-distributed stochastic neighbour embedding plot of RNA-seq datasets. (D) Unsupervised hierarchical clustering plot of mTSC RNA-seq data.

Expression of specific marker genes was analysed next. Y linked gene expression was first used to confirm the XY cell line genotyping data (Figure 4.7A). XY cell lines showed relatively higher expression of Y linked genes when compared with either XX or XO cell lines. Notably, Kdm5d appears to be expressed equally in cell lines of all sex chromosome genotypes, in particular mESCs and mEpiLCs (Figure 4.7A). This artifact is likely the result of mismapping of Kdm5c reads to Kdm5d, an effect documented for highly expressed genes, such as Kdm5c in mESCs (Conesa et al., 2016).

As there were no XY mTSC lines available, expression of Xist, the IncRNA that mediates XCI, was analysed. XCI takes place in XX mTSCs, such that Xist expression was expected from the Xp (Mak et al., 2002). As XO mTSCs carry only a single X chromosome, XCI is not required, and Xist expression was not expected. These predictions were met (Figure 4.7A). XCI is imprinted in extraembryonic tissue, such that Xist is always expressed from the Xp in XX mTSCs (Kunath et al., 2005). Notably, this imprinted expression was not maintained in XpO mTSCs, consistent with the requirement for X linked gene expression in extraembryonic tissue. Furthermore, I observed relatively low level Xist expression in mESCs, likely because they have not undergone XCI. This process has been initiated in mEpiLCs, hence the slight upregulation of Xist when compared to mESCs.
4.2. Results

**A**

<table>
<thead>
<tr>
<th>Gene</th>
<th>XX</th>
<th>XY</th>
<th>XX</th>
<th>XY</th>
</tr>
</thead>
<tbody>
<tr>
<td>mESC</td>
<td><img src="image1.png" alt="Heatmap" /></td>
<td></td>
<td><img src="image2.png" alt="Heatmap" /></td>
<td></td>
</tr>
<tr>
<td>mEpil.C</td>
<td><img src="image3.png" alt="Heatmap" /></td>
<td></td>
<td><img src="image4.png" alt="Heatmap" /></td>
<td></td>
</tr>
<tr>
<td>mTSC</td>
<td><img src="image5.png" alt="Heatmap" /></td>
<td></td>
<td><img src="image6.png" alt="Heatmap" /></td>
<td></td>
</tr>
</tbody>
</table>

**B**

XX vs. XY mESC

- **XX vs. XY relative to XY**
- **Upregulated in XX**
- **Downregulated in XX**

- **Red**: DE X-linked
- **Blue**: DE Y-linked
- **Black**: DE autosomal
- **Gray**: Not significant

**C**

XX vs. XY mEpilLC

- **XX vs. XY relative to XY**

- **Red**: DE X-linked
- **Blue**: DE Y-linked
- **Black**: DE autosomal
- **Gray**: Not significant
4.2. Results

Figure 4.7: Genotyping confirmation from RNA-seq data in all cell types. (A) Heatmap showing relative expression of Y linked genes and Xist in all RNA-seq datasets. (B) Volcano plot representing genes differentially expressed in XX mESCs compared to XY mESCs. (C) Volcano plot representing genes differentially expressed in XX mEpiLCs compared to XY mEpiLCs.

4.2.5 Identification of candidate imprinted genes by differential gene expression analysis (DGE)

Before comparing gene expression between X\textsuperscript{P}O and X\textsuperscript{M}O cell lines to identify potential candidates for the imprinting hypothesis, the differential gene expression (DGE) analysis pipeline was validated. Expression of genes from the Y chromosome sets apart the transcriptomes of XY from XX and XO cell lines. XX and XY mESC and XX and XY mEpiLC datasets were analysed for differentially expressed genes (Figure 4.7B,C). These two plots compare XX mESC and mEpiLC lines to XY mESC and mEpiLC baselines, respectively, therefore relative downregulation of Y linked genes was predicted. As expected, most highly significantly downregulated genes were Y linked (Figure 4.7B,C, circled). This provided confirmation that the bioinformatic pipeline was able to detect biologically relevant differential expression.

In order to identify all differentially expressed genes within cell type, pairwise comparisons were made between all sex chromosome genotypes (Figure 4.9A(i)). Focusing on the comparison between X\textsuperscript{P}O and X\textsuperscript{M}O lines, I used volcano plots to visualise the degree of differential expression for each gene and the significance of the hit (Figure 4.8).

For mESC, the total number of genes differentially expressed between X\textsuperscript{M}O and X\textsuperscript{P}O was 123, which were split evenly between upregulation and downregulation (Figure 4.8A). Interestingly, of these 123, 8 were X linked, again distributed evenly between upregulation and downregulation.
For mEpiLCs, 244 genes were differentially expressed in $X^M_O$ relative to $X^P_O$, and more of these were upregulated (Figure 4.8B). Of the 244 total differentially expressed genes, 10 were X linked genes, mostly upregulated in $X^M_O$ relative to $X^P_O$ EpiLCs.

For mTSCs, 201 genes were differentially expressed in $X^M_O$ relative to $X^P_O$, and significant number of these were downregulated (Figure 4.8C). Of 201 total differentially expressed genes, 24 were X linked, and these X linked genes were mostly downregulated in $X^M_O$ relative to $X^P_O$ mTSCs.
4.2. Results

- **A**
  - $X^{\text{mO}}$ vs. $X^{\text{oO}}$
  - mESC

- **B**
  - $X^{\text{mO}}$ vs. $X^{\text{oO}}$
  - mEpiLC

- **C**
  - $X^{\text{mO}}$ vs. $X^{\text{oO}}$
  - mTSC
I next compared the lists of differentially expressed genes between cells types to look for commonality. As mEpiLCs are differentiated from mESCs, and both contribute to the embryo proper, I expected to observe a degree of conservation between the differentially expressed genes. This proved to be correct, as 20 genes were significantly differentially expressed between X\textsuperscript{P}O and X\textsuperscript{M}O mESC and mEpiLC lines (Figure 4.9A(ii)). In contrast, only two genes were differentially expressed between X\textsuperscript{P}O and X\textsuperscript{M}O mTSCs and mESCs. Two other genes were differentially expressed between X\textsuperscript{P}O and X\textsuperscript{M}O in mTSCs and mEpiLCs. There were no genes differentially expressed between X\textsuperscript{P}O and X\textsuperscript{M}O cells common to all three cell types (Figure 4.9A(ii)).
4.2. Results

Figure 4.9: Differential gene expression analyses. (A) (i) Summary data of all genes differentially expressed between sex chromosome genotypes, within cell type; (ii) Venn diagram showing cross over between genes differentially expressed in XMO relative to XP0 cell lines; (iii) Summary data of all X linked genes differentially expressed between sex chromosome genotypes, within cell type; (iv) Venn diagram showing cross over between X linked genes differentially expressed in XMO relative to XP0 cell lines.

The fundamental difference between XP0 and XMO cell lines is the parental origin of the X chromosome, therefore one would expect any candidate genes for the imprinting hypothesis to be localised to the X chromosome. Autosomal genes may also be differentially expressed, though these genes would theoretically be downstream of any X linked candidates. I therefore chose to focus on the X chromosome, and looked to see how many of the genes identified in the DGE analysis localised to the X (Figure 4.9A(iii)). Once again focusing on the XP0 and XMO comparison, I compared the relatively short lists to look for any conservation. Unsurprisingly, there was only one gene differentially expressed between XP0 and XMO in more than one cell type. Mageb10-ps appeared in both mESC and mEpiLC lists (Table 4.3, 4.4), otherwise all candidates were unique to the cell type.
4.2. Results

Table 4.2: X linked genes significantly differentially expressed in X\textsuperscript{MO} vs X\textsuperscript{PO} mTSCs.
* indicates genes also significantly differentially expressed in X\textsuperscript{PO} relative to XX mTSC analysis. $ indicates genes also significantly differentially expressed in X\textsuperscript{MO} relative to XX mTSC analysis.

Most of the candidate genes across all three cell types were annotated as either predicted genes or pseudogenes, and were therefore not analysed further at this point.

Focusing on the mTSCs, some of the annotated candidates with ascribed function have potential based on what is already known. For example, \textit{Hs6st2} was upregulated in X\textsuperscript{MO} mTSCs (Table 4.2). It is a heparan sulphate 6-O-sulphotransferase enzyme involved in the metabolism of the extracellular matrix (Habuchi et al., 2000). Its autosomal homologue \textit{Hs6st1} is homozygous lethal during development, and conceptuses showed vascular abnormalities in the labyrinth of the placenta (Habuchi et al., 2007). Multiple genes from the reproductive homeobox (\textit{Rhox}) cluster were highlighted as downregulated in X\textsuperscript{MO} mTSCs. The \textit{Rhox} genes all have a conserved homeodomain, and are expressed in placenta, testis, and ovary (Maclean et al., 2006; Kobayashi et al., 2006). Interestingly, it has been suggested that a number of genes within this cluster are imprinted, though \textit{Rhox3} is not among these (Maclean et al., 2011).
The X<sup>P</sup>O embryonic growth deficit phenotype has been quantified relative to either XX or XY littermates (Chapter 3). Any candidate genes that might have an affect on this X<sup>P</sup>O phenotype might therefore be expected to also be differentially expressed in X<sup>P</sup>O mTSCs relative to XX mTSCs. It would also be interesting to compare XO to XY mTSCs, however, XY mTSCs were not recovered during the derivation process. I therefore looked in more detail at the comparisons between X<sup>P</sup>O and XX mTSCs, and then X<sup>M</sup>O and XX mTSCs.

First considering genes downregulated in X<sup>M</sup>O relative to X<sup>P</sup>O mTSCs; this result could either be interpreted as genes downregulated in X<sup>M</sup>O relative to other genotypes, or upregulated in X<sup>P</sup>O relative to other genotypes. Ten genes were found in common between those downregulated in X<sup>M</sup>O relative to X<sup>P</sup>O mTSCs and those upregulated in X<sup>P</sup>O relative to XX mTSCs (*, Table 4.2). I conclude that these ten genes are upregulated in X<sup>P</sup>O mTSCs. These genes could potentially contribute to the X<sup>P</sup>O embryonic growth deficit phenotype, and their expression would merit further investigation by qRT-PCR.
I then considered genes upregulated in $X^M O$ relative to $X^P O$ mTSCs. This result could be interpreted as either genes downregulated in $X^P O$ relative to other genotypes, or upregulated in $X^M O$ relative to other genotypes. Three genes were found in common to the lists of differentially expressed genes comparing $X^M O$ relative to $X^P O$ mTSCs, and $X^M O$ relative to $XX$ mTSCs (S, Table 4.2). I conclude that these three genes are upregulated in $X^M O$ mTSCs. As no difference in expression was observed for these genes between $X^P O$ and $XX$ mTSCs, I considered it unlikely that they would contribute to the $X^P O$ embryonic growth deficit phenotype.

### 4.2.6 A significant number of genes were differentially expressed between XO and XX mTSCs

The comparisons utilised in the previous section also highlighted the significant number of genes differentially expressed between $X^P O$ and $XX$ mTSCs, and $X^M O$ mTSCs and $XX$ mTSCs: 8124 and 5175, respectively (Figure 4.10A). Following this unexpected result, I analysed the overlap between the two lists. Given that the $X^P$ is always inactivated in $XX$ mTSCs, $XX$ mTSCs are theoretically more similar to $X^M O$ mTSCs than $X^P O$ mTSCs. I therefore expected to observe a greater number of genes differentially expressed in $X^P O$ relative to $XX$ mTSCs than $X^M O$ relative to $XX$ mTSCs, taking into account commonality between the two lists. I first excluded the 3817 genes commonly differentially expressed between both $X^P O$ and $X^M O$ relative to $XX$ mTSCs (Figure 4.9B). There were significantly more genes specifically differentially expressed in $X^P O$ relative to $XX$ mTSCs, than in $X^M O$ relative to $XX$ mTSCs.

The numbers of genes differentially expressed between XO and XX mTSCs represent a difference of more than an order of magnitude when making the same comparisons for the mESCs and mEpiLCs (Figure 4.9A(i)). As there is no obvious biological explanation for the observation, I looked for order or pattern within the differentially expressed genes by performing over representation enrichment analysis (ORA) using WebGestalt (Zhang et al., 2005). As the observation related to all
4.2. Results

XO mTSCs, I first considered those genes significantly differentially expressed in both $X^P_O$ relative to XX and $X^M_O$ relative to XX:

<table>
<thead>
<tr>
<th>Analysis type</th>
<th>over representation category</th>
<th>Expected genes</th>
<th>Observed genes</th>
<th>$p$ value</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO(biological process)</td>
<td>morphogenesis of an epithelium</td>
<td>82.27</td>
<td>115</td>
<td>$7.04 \times 10^{-5}$</td>
<td>$4.50 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>tube morphogenesis</td>
<td>63.20</td>
<td>91</td>
<td>$1.31 \times 10^{-4}$</td>
<td>$4.50 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>cell junction organization</td>
<td>31.35</td>
<td>51</td>
<td>$1.91 \times 10^{-4}$</td>
<td>$4.50 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>gland development</td>
<td>75.64</td>
<td>104</td>
<td>$2.95 \times 10^{-4}$</td>
<td>$4.50 \times 10^{-2}$</td>
</tr>
<tr>
<td></td>
<td>cell-substrate adhesion</td>
<td>44.78</td>
<td>67</td>
<td>$3.15 \times 10^{-4}$</td>
<td>$4.50 \times 10^{-2}$</td>
</tr>
<tr>
<td>GO(cellular process)</td>
<td>cell-substrate junction</td>
<td>65.08</td>
<td>108</td>
<td>$1.59 \times 10^{-8}$</td>
<td>$2.31 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>cell leading edge</td>
<td>55.93</td>
<td>79</td>
<td>$5.96 \times 10^{-4}$</td>
<td>$4.32 \times 10^{-2}$</td>
</tr>
<tr>
<td>GO(molecular function)</td>
<td>rRNA binding</td>
<td>9.08</td>
<td>208</td>
<td>$3.13 \times 10^{-4}$</td>
<td>$7.51 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

| Pathway (KEGG)         | Metabolic pathways                                  | 189.32         | 248            | $5.95 \times 10^{-7}$ | $1.78 \times 10^{-4}$ |
|                        | Lysosome                                             | 17.93          | 36             | $1.92 \times 10^{-5}$ | $2.86 \times 10^{-3}$ |
|                        | Influenza A                                          | 24.59          | 42             | $2.50 \times 10^{-4}$ | $2.49 \times 10^{-2}$ |

Table 4.5: Gene groups over represented in both DGE analyses comparing $X^P_O$ with XX mTSCs and $X^M_O$ with XX mTSCs. FDR cut off = 0.05.

<table>
<thead>
<tr>
<th>Chromosome band</th>
<th>Expected genes</th>
<th>Observed genes</th>
<th>$p$ value</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>11E2</td>
<td>29.54</td>
<td>80</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11B1.3</td>
<td>24.52</td>
<td>45</td>
<td>$2.75 \times 10^{-5}$</td>
<td>$5.27 \times 10^{-3}$</td>
</tr>
<tr>
<td>4D2.2</td>
<td>20.64</td>
<td>39</td>
<td>$4.60 \times 10^{-5}$</td>
<td>$5.89 \times 10^{-3}$</td>
</tr>
<tr>
<td>11C</td>
<td>16.49</td>
<td>32</td>
<td>$1.23 \times 10^{-4}$</td>
<td>$1.18 \times 10^{-2}$</td>
</tr>
<tr>
<td>11D</td>
<td>43.02</td>
<td>65</td>
<td>$3.61 \times 10^{-4}$</td>
<td>$2.59 \times 10^{-2}$</td>
</tr>
<tr>
<td>12C3</td>
<td>10.47</td>
<td>22</td>
<td>$4.05 \times 10^{-4}$</td>
<td>$2.59 \times 10^{-2}$</td>
</tr>
<tr>
<td>11B2</td>
<td>7.46</td>
<td>17</td>
<td>$6.40 \times 10^{-4}$</td>
<td>$3.02 \times 10^{-2}$</td>
</tr>
<tr>
<td>14D2</td>
<td>5.59</td>
<td>14</td>
<td>$6.61 \times 10^{-4}$</td>
<td>$3.02 \times 10^{-2}$</td>
</tr>
<tr>
<td>4D2.3</td>
<td>6.88</td>
<td>16</td>
<td>$7.21 \times 10^{-4}$</td>
<td>$3.02 \times 10^{-2}$</td>
</tr>
<tr>
<td>14B</td>
<td>15.20</td>
<td>28</td>
<td>$7.87 \times 10^{-4}$</td>
<td>$3.02 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

Table 4.6: Chromosomal locations over represented in DGE analyses comparing $X^P_O$ with XX mTSCs and $X^M_O$ with XX mTSCs. FDR cut off = 0.05.
4.2. Results

Figure 4.10: Genes differentially expressed between XO and XX mTSCs. (A) Venn diagram showing overlap between genes differentially expressed in \( X^P O \) relative to XX mTSCs and genes differentially expressed in \( X^M O \) relative to XX mTSCs. (B) (i) Ideogram of the mouse genome showing upregulation or downregulation based on relative log fold change of genes differentially expressed in \( X^P O \) relative to XX mTSCs. All genes are depicted as a single coloured bar in their relative chromosomal location; (ii) Ideogram of the mouse genome showing upregulation or downregulation based on relative log fold change of genes differentially expressed in \( X^M O \) relative to XX mTSCs. All genes are depicted as a single coloured bar in their relative chromosomal location; (iii, iv) Ideograms of the mouse genome showing upregulation and downregulation of genes differentially expressed in both \( X^P O \) and \( X^M O \) relative to XX mTSCs. (iii) represents logFC in \( X^P O \) mTSCs relative to XX mTSCs; and (iv) represents logFC in \( X^M O \) relative to XX mTSCs. All genes are depicted as a single coloured bar in their relative chromosomal location. (B) plotted by Jasmin Zohren.

A limited number of over representation categories were identified per analysis type, using a relatively conservative false discovery rate of 0.05 (Zhang et al., 2005), as detailed in Tables 4.5 & 4.6. Most notably, the analysis with the highest number of significant hits was Chromosomal location (Table 4.6). The results suggest significant over representation of genes from chromosome 11, in addition to some bands within chromosomes 4, 12 and 14. In order to determine whether there was a prevailing general trend of upregulation or downregulation, idograms were plotted based on logFC (Jasmin Zohren: Figure 4.10). First, direction of change, i.e. upregulation or downregulation, was plotted for all genes differentially expressed in \( X^P O \) relative to XX mTSCs (Figure 4.10B(i)); and for all genes differentially expressed in \( X^M O \) relative to XX mTSCs (Figure 4.10B(ii)). For both comparisons, there was a mix of upregulation and downregulation. The upregulation and downregulation of genes common to both comparisons (overlap in Venn, Figure 4.10A) was then plotted for \( X^P O \) relative to XX mTSCs (Figure 4.10B(iii)), and for \( X^M O \) relative to XX mTSCs (Figure 4.10B(iv)). This revealed that the 3817 genes in common to both comparisons were differentially expressed in the same direction, i.e. upregulated or downregulated, for all XO mTSCs. Graphical representation of the common genes also aided visualisation of the regions highlighted in the ORA. For example, 11E2, the most significant chromosomal location identified in the ORA can be seen as upregulated in both \( X^P O \) and \( X^M O \) relative to XX mTSCs. In contrast, 4D2.2 and 4D2.3 were downregulated in both \( X^P O \) and \( X^M O \) relative to XX mTSCs. I
conclude that, based on these results, further work is required in order to further elucidate the mechanisms underlying this significant differential gene expression between XO and XX mTSCs.

### 4.3 Discussion

In mouse extraembryonic tissue, X linked genes are expressed from the $X^M$, and the $X^P$ is always inactivated (Takagi and Sasaki, 1975; Okamoto et al., 2004). This is true for XX, XY and $X^M O$ conceptuses. Embryos with an $X^P O$ sex chromosome genotype, in contrast, only have a single $X^P$ and no $X^M$, therefore the $X^P$ is active in the extraembryonic tissue. When combined with the observations that $X^P O$ embryos are growth retarded relative to XX (Thornhill and Burgoyne, 1993) and $XY^{RIII}$ (Chapter 3) littermates at E10.5; and that $X^P O$ embryos have smaller EPCs at E7.5 (Jamieson et al., 1996), a causative link is plausible. In this Chapter I set out to identify genes differentially expressed between $X^P$ and $X^M$ chromosomes, thus serving as candidates to potentially explain the $X^P O$ embryonic growth deficit phenotype. I utilised *in vitro* models for the extraembryonic, pluripotent and epiblast compartments of the mouse embryo and, following RNA-seq analysis, highlighted a number of interesting differentially expressed genes that now require further testing. Prior to addressing the next steps for this experiment, a number of notable observations will be discussed.

#### 4.3.1 No XY mTSC lines were recovered during the derivation process

In order to generate $X^P O$ mTSCs, I crossed an $XY^{RIII}$ male with an InX/X female. I expected to recover roughly equal numbers of female (XX;InX/X;$X^P O$) and male ($XY^{RIII}$; InX/Y$^{RIII}$) blastocysts, and therefore derive cell lines in approximately equal proportions. It was therefore surprising to find no male mTSC lines following genotyping. This could be explained by a number of factors. Firstly, it is possible that there we no male blastocysts on day 1 of the derivation. However, as the same
cross was utilised to derive mESCs, and male lines were recovered, this reasoning is unlikely to be true. Secondly, the Y chromosome may have a previously uncharacterised deleterious effect on the survival of mTSCs. This explanation is also unlikely, as I successfully derived XX\textsuperscript{Y} mTSC lines as an unused product of the cross to generate X\textsuperscript{MO} mTSCs (Figure 4.2B(ii)). In these XX\textsuperscript{Y} cells, the Y chromosome is physically linked to one of the X chromosomes, with no overt adverse phenotypic effects reported (Burgoyne et al., 1998; Trent et al., 2012). Moreover, XY mTSC lines have previously been reported in the literature (Tanaka et al., 1998; Motomura et al., 2016), though most papers continue to fail to report the sex chromosome karyotype of cell lines used.

A third theoretical explanation for the failure to derive XY mTSCs could be X chromosome number. However, the successful derivation to XO mTSCs effectively rules out any single X effect, and derivation specifically of X\textsuperscript{MO} and X\textsuperscript{PO} lines eliminates the possibility of an X chromosome imprinting effect on XY mTSC survival. The final explanation is experimental artifact, which asserts that if the experiment were repeated, male mTSCs would be obtained. This could be addressed by a repeat derivation utilising the same mouse cross, but instead of using feeders of mixed sex chromosome karyotype, use XX MEFs. Genotyping by PCR could then be performed around day 28.

4.3.2 Significantly more genes are differentially expressed between XX and XO cell lines in mTSCs than mESCs or mEpiLCs

Differential expression analyses comparing X\textsuperscript{MO} with X\textsuperscript{PO} RNA-seq datasets for mTSCs, mESCs and mEpiLCs identified of a small number of genes for each cell type. In contrast, the analyses of X\textsuperscript{PO} relative to XX mTSCs, and X\textsuperscript{MO} relative to XX mTSCs, revealed surprisingly high numbers, at 8124 and 5175, respectively. 3817 genes were common to both lists, and over representation analysis revealed relative enrichment for genes located on chromosome 11, among others. Several
possible explanations exist for these observations, including aneuploidy independent of the sex chromosomes, and a previously unrecognised effect of sex chromosome genotype on mTSC gene expression.

There are very little published data on mTSC karyotypes. A number of studies found euploid and aneuploid mTSC lines following routine karyotyping, but none characterised the nature of the aneuploidies (Tanaka et al., 1998; Kubaczka et al., 2015; Benchetrit et al., 2015). A contributory factor to this apparent oversight may stem from studies of tetraploidy. Whilst tetraploid embryos do not survive gestation (Kaufman and Webb, 1990; Henery et al., 1992), diploid embryos with tetraploid extraembryonic tissue do (Nagy et al., 1990). This could be taken to imply that the extraembryonic tissue is more tolerant than the epiblast to aneuploidy. However, duplication of the whole genome has significantly fewer transcriptional, and phenotypic, consequences in comparison to loss or gain of single chromosomes (Birchler, 2014).

In contrast, it is well documented that mESC lines have a tendency to develop karyotypic abnormalities following prolonged culture. Trisomies 8 and 11 are both common in mESCs and, when present together, result in a recognisable growth advantage (Sugawara et al., 2006; Rebuzzini et al., 2008; Gaztelumendi and Nogués, 2014; Rebuzzini et al., 2016). The X chromosome is also often affected in XX mESC lines, with loss resulting in the creation of a pseudo XO line (Robertson et al., 1983).

Based on the data presented here, if autosomal aneuploidy is present, it is possible that either all eight independently derived XX mTSC lines carry similar chromosomal abnormalities, or all five independently derived XO mTSC lines carry similar chromosomal abnormalities. Importantly, I confirmed that the sex chromosome complements were as reported for all cell lines, by DNA FISH. Autosomal aneuploidy could be investigated first by karyotyping of each cell line, thus revealing whether the overall ploidy is normal. If there is any doubt, spectral karyotyping, multiple (M) FISH, or comparative genomic hybridisation would provide more de-
4.3. Discussion

Detailed information on specific chromosomal rearrangements (Liyanage et al., 1996; Anderson and Brown, 2005; Saldarriaga et al., 2015).

Alternatively, chromosome territories in mTSCs may be disturbed by the sex chromosome aneuploid state. In the nucleus, chromosomes are positioned within defined areas, with gene rich chromosomes located towards the interior and gene poor chromosomes closer to the periphery (Cremer et al., 2006). This organisation is altered in disease states such as cancer (Meaburn and Misteli, 2008; Barutcu et al., 2015), which frequently involves structural chromosomal rearrangements and aneuploidy. Such chromosome territory alterations can have a significant impact on gene expression profiles (Meaburn and Misteli, 2008; Barutcu et al., 2015). Little is known about the nuclear organisation in mTSCs, as the only study to date focused on escape from XCI and, as such, only utilised female cell lines. It would therefore be highly informative to investigate the nuclear organisation of XX, XY, X^P^O and X^M^O mTSCs using techniques such as 3-D FISH and chromosome conformation capture (Fudenberg and Imakaev, 2017).

A third explanation for the high numbers of differentially expressed genes in XO mTSCs relative to XX mTSCs is non genetic phenotypic heterogeneity.

4.3.3 Non genetic phenotypic heterogeneity in aneuploid cells

Recent work has highlighted that aneuploidy may cause phenotypic variability independent of genetic variation. Working primarily in yeast, Beach and colleagues created karyotypically homogeneous populations of aneuploid cells that showed significant phenotypic changes within the population when compared with euploid cells (Beach et al., 2017). They showed that cells with chronic aneuploidities - defined as a chromosomal gain/loss propagated for more than two cell divisions - exhibited heterogeneity in cell cycle timing, variability in gene expression in response to cellular stress, and altered protein expression levels relative to wildtype. Data
are also provided to show phenotypic variability in inbred mice carrying identical Robertsonian translocations (Beach et al., 2017).

The heterogeneity in gene expression reported in Beach et al. study is likely to apply equally to sex chromosome aneuploidy as it applies to autosomal aneuploidy. For example, this phenomenon could underlie the increased variation in gene expression observed for XO but not XX mTSC lines during the differentiation analysis (section 4.2.2). It is, therefore, conceivable that such heterogeneity could also impact on differential gene expression analysis. In the present study I have relied upon reproducibility between biological replicates to identify genes as differentially expressed. However, if the findings of Beach and colleagues apply to these cell lines, there would theoretically be very few conserved differentially expressed genes found across all X^P^O biological replicates or across all X^M^O biological replicates. Importantly, the number of differentially expressed genes identified here in euploid cell line comparisons, i.e. XX and XY mESCs (322 genes), is within the same order of magnitude as the number identified in aneuploid cell line comparisons, i.e. X^P^O and XX (109 genes). It therefore seems likely that, despite the reported phenotypic variability in aneuploid conditions, a core set of genes underlies conserved aspects of the phenotype. Undoubtedly further work is required to determine the effects of non genetic phenotypic variability in human aneuploidies, though this could potentially provide an explanation for the highly variable phenotype observed in women with Turner syndrome (Hook and Warburton, 2014).

### 4.3.4 Confounding factors in analysis of RNA-seq data from aneuploid samples

It has previously been assumed that a change in chromosomal complement inside a cell will only impact on gene expression from the chromosome(s) directly affected, and thus gene expression from the unaffected chromosomes can be used for normalisation. This approach was taken in the analysis of early work on X chromosome upregulation (XUR), as previously described (section 1.4.1). However, based on
4.3. Discussion

In the preparation of libraries for sequencing, the total RNA input was equalised across all samples, and the concentrations of all libraries were equal on the flow cell. These normalisations assume that total RNA per cell is approximately equal. Samples were then normalised for sequencing depth by calculating transcripts per million (tpm). Tpm is defined as the number of transcripts of one specific type (e.g. GAPDH-001) that would be detected if one sequenced one million full length transcripts, given the abundance of the other transcripts in the sample (Li et al., 2009a). It is therefore important that relative transcript abundances for the whole transcriptome are comparable between samples. In the case of aneuploid cell lines, the potential global changes in gene expression mean that this criterion may not be met and, as such, downstream differential gene expression analyses may not be optimal.

In order to mitigate these difficulties, a number of approaches could be utilised; the first two as an extension to current work, and the second two in future work. Using existing data, a ratio of read counts for each transcript in the X<sup>p</sup>O versus XX could be calculated, and subsequently plotted in a distribution. This would serve to highlight if there are any abnormal patterns of potential interest (Sun et al., 2013a, 2010, 2013b). Although relatively low throughput, quantitative RNA FISH allows direct visualisation of the nascent transcript, facilitating comparison of per cell expression in two different sex chromosome genotypes (Pal Bhadra, 2005). The addition of a pre determined concentration of exogenous spike in RNA to every sample prior to cDNA preparation would provide an absolute reference for across sample RNA abundance comparisons (Jiang et al., 2011; Lovén et al., 2012). Alternatively, the

the Beach et al. (2017) study, in addition to previous work detailing the genomic instability of aneuploid cells (Nicholson et al., 2015; Blank et al., 2015; Passerini et al., 2016), and the literature on XUR, this assumption is not supported by the recent data. It is becomingly increasingly clear that there are global transcriptional changes in aneuploidy, and these must be taken into account during the data analysis to enable accurate differential expression analyses (Birchler, 2014).
mRNA to genomic DNA relationship could be preserved during sample preparation such that genomic DNA acts as a control for the calculation of transcriptome size. This has been developed primarily for polyploid cells, but could effectively be used in the study of aneuploidies, too (Coate and Doyle, 2010).

4.3.5 Validation of cell lines for the study of imprinted genes

The primary aim of the derivation and transcriptome analysis of X_P0 and X_M0 stem cell lines was to identify X linked differentially expressed genes. A subset of these candidate genes could be imprinted, and therefore underlie the X_P0 embryonic growth deficit phenotype. It has been shown that culture conditions can influence the maintenance of genomic imprinting in mESCs (Habibi et al., 2013), and the potential impact of such an effect should have been assessed in my cell lines. Previously reported imprinted loci, such as Peg3 and Kcnq1ot in both mESCs and mTSCs (Leitch et al., 2013; Calabrese et al., 2015), could have been used as positive controls for parent of origin expression. However, in order to determine the parental origin of a given allele, single nucleotide polymorphisms (SNPs) are required. I used MF1 strain sex chromosome variant mice to generate both X_P0 and X_M0 blastocysts for cell line derivation. Whilst the MF1 strain is outbred, it does not have enough SNPs to perform accurate and effective variant calling. As a consequence, it was not possible to confirm that autosomal imprinting is reliably maintained within the model. The ability to assess potential X chromosome imprinting was not affected, as two X chromosome monosomic states were used whereby parent of origin was known a priori.

A number of studies in mESCs have suggested that culture conditions impact the DNA methylation status of imprinting control regions, creating a potential risk of culture artifacts (Ficz et al., 2013; Habibi et al., 2013; Blaschke et al., 2013). Cells derived and maintained in 2iLIF conditions were reported to show faithful maintenance of parental imprints, as per the ICM of the blastocyst stage embryos, at three
loci examined (Leitch et al., 2013). However, two other reports found sexually di-

morphic ICR methylation maintenance at the *Peg3* and *Peg10* loci in both serum and
LIF and 2iLIF conditions (Ooi et al., 2010; Zvetkova et al., 2005). Recently it has
been shown that histone modifications, such as H3K27me3, also play a role in the
regulation of imprinted gene expression in the early mouse embryo independently
of DNA methylation (Inoue et al., 2017). It will therefore be important to deter-
mine whether culture conditions impact on the genes regulated by this mechanism,
too.

There are no reports of culture induced changes in DNA methylation or imprinting
in mTSCs, however, there are also only a small number of reports showing im-
printing in mTSCs. The most comprehensive study to date utilised eight F1 mTSC
lines produced by reciprocal cross between C57BL/6J and *Mus musculus castaneus*
mouse strains (Calabrese et al., 2015). The authors found evidence for imprinted ex-
pression of 31 previously described imprinted genes and 17 new candidates, based
on RNA-seq and pyrosequencing data. They also noted that several of the known
imprinted genes showed variability in parent of origin expression across their lines,
and suggest this could either be an example of developmental plasticity at imprinted
loci (Radford et al., 2011), or alternatively an artifact of culture.

It would be useful to further understand the mechanisms of imprint maintenance in
mTSCs, as this was not addressed in the Calabrese et al. study. Recently it was
shown that DNA methylation marks originating in the oocyte are important for tro-
phoblast development, and play a role in trophoblast cell adhesion both in vivo and
in vitro (Branco et al., 2016). Furthermore, loci specifically imprinted in extraem-
bryonic tissue are enriched for repressive histone modifications such as H3K27me3
and H3K9me2 (Lewis et al., 2004; Umlauf et al., 2004), and the active histone
mark H3K4me2 (Lewis et al., 2006). However, some placentally imprinted genes
appear to be biallelically expressed in mTSCs and blastocysts, and as such show
no enrichment for histone modifications at these loci (Lewis et al., 2006). More-
4.3. Discussion

over, following *in vitro* differentiation, there was no change in allelic expression or histone modifications.

In summary, the cell lines I have derived as a model for X chromosome imprinting in the X^P^O embryo require validation to confirm the expression of known imprinted genes. It would therefore be useful to perform whole genome bisulphite sequencing on F1 mESC, mEpiLC, and mTSC lines. This would enable confirmation of the methylation status of known ICRs based on allele calling data and, furthermore, in conjunction with RNA-seq data, would facilitate the identification of any imprinted genes on the X chromosome in mESCs, mEpiLCs, and mTSCs. ChIP-seq analysis in all cell types for histone modifications such as H3K27me3 would also be important (Inoue et al., 2017; Lewis et al., 2006), with the caveat that mTSCs have been shown to recapitulate genes imprinted at the blastocyst stage but no later. If imprinting candidates are not found at this early stage, alternative models will be required.

4.3.6 Alternative models for investigating the X^P^O embryonic growth deficit

Whilst stem cells *in vitro* undoubtedly contribute a significant amount to our understanding of biological systems and, importantly, facilitate a reduction in the use of live animals in research, it is sometimes challenging to accurately recapitulate the *in vivo* phenotype in a plate or dish. In order to investigate the X^P^O embryonic growth deficit question further, a number of *in vivo* options could complement the current cell focused approach.

As it has previously been shown that the EPC of X^P^O conceptuses is significantly smaller than the EPC of XX littermates (Jamieson et al., 1998), it would be useful to reinvestigate this model with modern techniques. A reporter for *Ascl2*, expressed only in the EPC of the extraembryonic tissue at E7.5 (Guillemot et al., 1994), could be generated by CRISPR-Cas genome editing (Yang et al., 2013) and, following
dissection of the conceptus, used to FACS sort the EPC cells for RNA-seq or proteomic analysis.

A secondary alternative to access the same population would be to fix or freeze the conceptuses and utilise laser capture microdissection (DeCarlo et al., 2011). This technique would enable the isolation of the morphologically identifiable EPC cells in the E7.5 conceptus for use in downstream RNA-seq or proteomic analysis.

A third model could be utilised for a direct comparison with in vitro derived mTSCs. Using laser assisted microdissection on the blastocyst stage embryo, it is possible to separate the mural TE from the polar TE and ICM. These two different cell populations could then be dissociated into single cells for RNA-seq analysis. Using lineage specific marker genes such as Cdx2 and Eomes for the TE and Pou5f1 and Nanog for the ICM, it would be possible to bioinformatically separate the polar TE/ICM population into trophectoderm and ICM (Blakeley et al., 2015); and possibly further separate the ICM into epiblast and primitive endoderm. Finally, differential gene expression analyses could be performed, both between genotypes, as described for mTSCs during this study, and within embryos between cell types. It has previously been shown that transcriptional differences exist between polar and mural TE at E3.5 (Nakamura et al., 2015), and it would be interesting to observe if these differences change with sex chromosome genotype.

4.3.7 Potential use of proteogenomics to further investigate the XPO growth deficit phenotype

The identification and relative measurement of mRNA species by RNA-seq enables the study of transcriptional activity within a given cell or tissue sample. By comparing transcriptomes between samples from different experimental conditions, we are able to determine the impact of the condition on gene expression. RNA species are not generally, however, the final effectors of cellular output; this task falls to protein. Conservative estimates suggest that the correlation between cellular pro-
tein concentration and abundance of their corresponding mRNA is around 40% (de Sousa Abreu et al., 2009; Maier et al., 2009). The 60% variance is attributed to post transcriptional, translational, and degradation regulation (Vogel and Marcotte, 2012). Recent work characterised human colon and rectal tumours by proteogenomic analysis (Zhang et al., 2014). They noted that whilst copy number alterations showed cis and trans effects on mRNA abundance, mRNA transcript abundance did not generally predict protein abundance differences between tumours. The combined study of mRNA and protein profiles helped the authors to prioritise candidate driver genes, and provided functional context to genetic abnormalities, such as duplication of chromosome 20q (Zhang et al., 2014). Such a combined mRNA/protein approach might enabled more effective identification of candidate genes when looking to understand a complex phenotype, such as the growth deficit observed in XP embryos at E10.5.
Chapter 5

Assessing the contribution of X-linked dosage sensitive genes to the postnatal growth deficit in XO female mice

5.1 Introduction

Short stature is one of the most recognisable features associated with Turner syndrome, and is a major cause of disability during childhood and adolescence (Kappeggaard and Laursen, 2011; Lee and Conway, 2014). Haploinsufficiency of \textit{SHOX}, a gene located within PAR1, is likely a major contributor to the phenotype (Rao et al., 1997; Ellison et al., 1997). \textit{SHOX} is a transcription factor highly expressed in chondrocytes and bone marrow fibroblasts (Rao et al., 1997; Ellison et al., 1997; Rao et al., 2001; Aza-Carmona et al., 2011). Mutations in the gene are one of the primary genetic causes of Leri-Weill syndrome, which itself encompasses a short stature phenotype (Belin et al., 1998), and have also been associated with many other cases of short stature (Marstrand-Joergensen et al., 2016; Fukami et al., 2016). However, women with Turner syndrome are generally shorter than patients
with mutations in SHOX, suggesting that other factors may contribute (Ross et al., 2001; Fukami et al., 2016; Zinn and Ross, 1998).

Similar to Turner syndrome in humans, female mice with an XO karyotype are smaller than those with a normal XX karyotype during early postnatal life (Burgoyne et al., 1983a). Although originally suggested to result from genomic imprinting, this phenotype has been shown to be independent of the parental origin of the X chromosome (Burgoyne et al., 2002). Furthermore, mice have no SHOX orthologue, and their PAR is significantly reduced in size and gene content compared to humans. Burgoyne and colleagues (2002) showed that female mice with one normal X chromosome and an additional PAR-only chromosome are small compared to XX littermates. Interestingly, the addition of a second sex chromosome, in the form of a Y chromosome lacking the sex determining gene Sry (XY−), fully rescued the phenotype. The authors therefore concluded that haploinsufficiency for an X linked gene that is located outside the PAR, and has a widely-expressed Y-linked homologue, is the most likely cause of the growth deficit. This gene would also be expected to escape XCI.

In this Chapter I set out to confirm the results from the Burgoyne et al. dataset, and to identify the genes involved in the growth retardation phenotype. A small number of candidates were highlighted following a Y-linked gene expression analysis and RFLP assay to test for escape from XCI. I then sought to address each candidate’s role in the XO postnatal growth deficit by creating mutant lines by CRISPR/Cas9 genome engineering.

5.2 Results

5.2.1 Recapitulation of previous results

I first sought to reproduce the results of Burgoyne et al. (2002) that, relative to XX littermates, XP O and XM O female mice grow at a reduced rate during the first
five weeks of postnatal life. In the Burgoyne et al. study, XO female mice with an additional PAR, denoted XY*X, also grew at a reduced rate across the same period, whereas XY females grew at the same rate as XX littermates.

In order to generate XPO females alongside XX littermates, I crossed In(X)/X female mice with XYRIII males (Figure 5.1A(i)). In addition to XPO females, XX and In(X)/X females were present in the offspring. It has previously been reported that there is no difference in postnatal weight gain between XX and In(X)/X females, so results were pooled (Burgoyne et al., 2002). XMO females were produced by crossing XX female mice with XPOY* males (Figure 5.1B(i)). XO female mice with an additional PAR (XY*X) were generated by crossing XX females with XY* males (Figure 5.1C(i)). Finally, we generated XY females by utilising a male with this variant Y chromosome, in addition to an autosomal Sry transgene, crossed with a wildtype female (Figure 5.1D(i)).

I first compared weights at birth. XPO females showed no significant birth weight difference when compared with XX littermates (Table 5.1, \( p=3.86 \times 10^{-1} \)). Interestingly, XMO females were significantly lighter than XX littermates at birth (Table 5.1, \( p=2.20 \times 10^{-2} \)), as were XY*X females (Table 5.1, \( p=6 \times 10^{-3} \)). XY females were significantly heavier than XX littermates at birth (Table 5.1, \( p=2.80 \times 10^{-3} \)).

In order to estimate the rate of growth across the periods 0-3 weeks and 3-5 weeks, a regression line was fitted to the weight data for each animal. Relative to XX littermates, XPO females grew at a reduced rate during the 0-3 week period (Figure 5.1A(ii), Table 5.2, \( p=2.00 \times 10^{-2} \)). XMO females (Figure 5.1B(ii), Table 5.2: \( p=4.60 \times 10^{-2} \)) and XY*X females (Figure 5.1C(ii), Table 5.2: \( p=2.80 \times 10^{-2} \)) also grew at a significantly reduced rate when compared to XX littermates between 0 and 3 weeks. In contrast, and consistent with previous results, XY females showed no difference in rate of growth when compared to XX littermates across the 0-3 week period (Figure 5.1D(i), Table 5.2: \( p=1.39 \times 10^{-1} \)).
5.2. Results

Across the 3-5 week period, X\textsuperscript{P}O females grew at a rate not significantly different from XX littermates ($p=7.12\times10^{-1}$). Similarly, relative to XX littermates, X\textsuperscript{MO} females (Table 5.2: $p=3.16\times10^{-1}$) and XY\textsuperscript{X} females (Table 5.2: $p=2.64\times10^{-1}$) grew at an equivalent rate. Interestingly, XY\textsuperscript{-} females slowed a slight, but significant, reduction in weight gain during the 3-5 week period in comparison to XX littermates (Table 5.2: $p=3.60\times10^{-2}$): this is in keeping with the trend shown previously (Burgoyne et al., 2002). Figure 5.1E utilises MWD from XX to bring results from all genotypes together. In summary, I confirmed that XO female mice grow at a reduced rate relative to XX littermates during the 0-3 week period, and that this is not a result of haploinsufficiency for PAR genes. I conclude that haploinsufficiency for genes present on both the X and Y chromosomes, and that map outside the PAR, is the most likely cause of a reduced rate of postnatal growth in XO mice.
5.2. Results

A (i)  

B (i)  

C (i)  

D (i)  

E
Figure 5.1: Recapitulation of postnatal growth retardation phenotype in XO female mice. (A) (i) Mouse cross used to generate XX and X\textsuperscript{P}O littermates, and (ii) body weight data of these animals, measured weekly for five weeks. Each data point represents an individual animal, and the line depicts the mean for each genotype. (B) (i) Mouse cross used to generate XX and X\textsuperscript{MO} littermates, and (ii) body weight data. (C) (i) Mouse cross used to generate XX and XY\textsuperscript{aX} littermates, and (ii) body weight data. (D) (i) Mouse cross used to generate XX and XY\textsuperscript{Sry} littermates, and (ii) body weight data. (E) Mean weighted difference between genotypes indicated and XX littermates. Bars represent 95% confidence intervals.
Table 5.1: Birth weight data compared by mean weighted difference (MWD)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Number</th>
<th>Birth weight (g)</th>
<th>MWD±SE (g)</th>
<th>Signif. of MWD (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>XX v X(^P)O</td>
<td>17 6</td>
<td>1.719±0.040 1.699±0.376</td>
<td>0.020±0.045</td>
<td>3.86x10(^{-1})</td>
</tr>
<tr>
<td>XX v X(^M)O</td>
<td>27 21</td>
<td>1.755±0.039 1.747±0.046</td>
<td>0.038±0.042</td>
<td>2.20x10(^{-2})</td>
</tr>
<tr>
<td>XX v XY(^*)X</td>
<td>25 21</td>
<td>1.638±0.037 1.600±0.036</td>
<td>0.041±0.038</td>
<td>6.00x10(^{-3})</td>
</tr>
<tr>
<td>XX v XY(^-)</td>
<td>26 38</td>
<td>1.610±0.035 1.675±0.032</td>
<td>-0.064±0.028</td>
<td>2.80x10(^{-2})</td>
</tr>
</tbody>
</table>

Table 5.2: Means and mean weighted differences (MWD) for phasic weight gain between genotype comparisons shown in Figure 5.1. "a" represents mean of means.
5.2. Results

5.2.2 Verification of potential postnatal growth retardation candidates

Next I sought to narrow down potential candidate genes for the haploinsufficiency hypothesis. As elucidated previously (Burgoyne et al., 2002) and here, such genes will not be imprinted, will not be present in the PAR, and the Y linked homologues will likely exhibit widespread expression. I looked to confirm this pattern by using RT-PCR to assay for expression of all 14 Y linked protein coding genes across multiple tissues representing ectoderm, mesoderm and endoderm (Figure 5.2). Eight Y-linked genes showed expression exclusive to the testis (Zfy1, Ube1y1, Usp9y, Zfy2, Sry, Ssty1, Ssty2, and Sly). Two genes showed expression in the testis and limited expression outside the gonad (H2al2y and Rbmy1a1). The final four genes (Kdm5d, Eif2s3y, Uty, and Ddx3y) showed widespread expression across tissues derived from all three embryonic germ layers. These patterns were consistent with previously reported data (Bellott et al., 2014).

Each of the broadly expressed Y linked genes has a conserved X linked homologue in mouse (Disteche et al., 2002). This conservation not only applies to the mouse sex chromosomes, but also across the eutherian clade, which has been taken to suggest that levels of expression of these X-Y pairs might be important (Bellott et al., 2014; Cortez et al., 2014). In XX female mice, the expression of most X linked genes is controlled by XCI; though it is likely that dosage sensitive genes such as the X linked genes in these X-Y pairs, will escape the chromosome wide silencing. I looked to confirm that the X linked orthologue of each of the widely expressed Y linked genes escapes XCI, as has been previously reported (Disteche et al., 2002). I determined whether one or both alleles of Ddx3x, Eif2s3x, Kdm5c, and Kdm6a was being expressed in adult XX female mice by restriction fragment length polymorphism analysis (RFLP).
Figure 5.2: Expression analysis of genes on the mouse Y chromosome by RT-PCR. Relative positions of genes on the mouse Y chromosome are depicted by cartoon on the left hand side. The positive control was adult wildtype mouse testis; the negative control was testis from an adult male mouse lacking the Y chromosome but carrying Eif2s3y and Zfy2 transgenes.
XCI is usually random in all adult somatic cells, such that either $X^M$ or $X^P$ could be inactivated. As a result, on a tissue level, both alleles from every expressed gene would be detected, regardless of whether genes are subject to or escape from XCI. In order to isolate expression from potential XCI escapees only, I again utilised the $Xist$ mutant allele (Marahrens et al., 1997). The X chromosome carrying this allele cannot be inactivated, and any detected expression from the second X chromosome in XX females could be attributed to escape from XCI.

In order to differentiate between transcripts from the two different X chromosomes, I used an F1 cross between strains bearing known single nucleotide polymorphisms (SNPs) in the genes of interest. After amplifying the locus containing the SNP from cDNA by PCR, I then digested the product with a restriction enzyme specific to one of the variant nucleotides at the SNP. If the gene escaped XCI, I expected to see two products after digest, i.e. one from each allele. If the gene was subject to XCI, I expected to observe only a single product after digest, the product from the active X chromosome.

$Foxo4$, a gene known to be subject to XCI, was assayed first as a negative control. I amplified and digested PCR products from cDNA of the parental strains to test the RFLP assay, expecting to observe a single, digested product in one strain, and a single undigested product in the other strain. In both of the parental strains, following digest, distinct bands were observed on the agarose gel (Figure 5.3A). As a result the $Xist$ mutation on the C57BL/6 X chromosome, the XX female F1 was expected to express $Foxo4$ from the C57BL/6 X chromosome - the active X chromosome - only. This expectation was fulfilled.

I then assayed $Ddx3x$, $Eif2s3x$, $Kdm5c$, and $Kdm6a$. If a gene escapes XCI, I expected to observe expression from both X chromosomes in the XX female F1, in the form of two different allele specific bands. For each of $Ddx3x$, $Eif2s3x$, $Kdm5c$, and $Kdm6a$, this expectation was fulfilled, in both brain and liver tissue samples (Figure 5.3A,B). I therefore confirmed that $Ddx3x$, $Eif2s3x$, $Kdm5c$, and $Kdm6a$ escape XCI across multiple tissues in the mouse. When compiled with the data showing
presence and widespread expression of Y linked orthologues, I conclude that one or more of these four candidates is most likely to be haploinsufficient in XO female mice, resulting in a reduced rate of growth in the early postnatal period.

5.2.3 The generation of mutant alleles by CRISPR Cas9 genome engineering

I next looked to identify which of these genes mediates the postnatal growth retardation phenotype by utilising individual knockout alleles for *Kdm5c*, *Kdm6a*, *Ddx3x* and *Eif2s3x*. As all crosses described up to this point have utilised the MF1 mouse strain, I wanted to maintain consistency and reduce variation attributable to strain specific effects (Latham, 1994) by utilising the same strain for the knockouts. Whilst a number of mutant lines had already been published for *Kdm6a*, none used...
5.2. Results

the MF1 strain. Rather than adopting a lengthy backcrossing strategy, I opted to create new lines via CRISPR/Cas9 genome editing (Jinek et al., 2012; Cong et al., 2013). Guide RNAs were designed utilising the MIT CRISPR Design Tool (Ran et al., 2013), with the aims of removing the translation initiation site, whilst minimizing off-target mutations. Initially, each guide was shown to successfully cut the target locus by digest of a PCR product containing the guide target sequence (Figure 5.4).

To assess guide efficiency, XY mESCs were then electroporated with each guide validated individually, along with Cas9 mRNA. All guides for Ddx3x, Kdm5c, and Kdm6a were able to create mutations in vitro, as assessed by either RFLP or Sanger sequencing (Table 5.3). Targeting efficiency, measured as the percentage of all picked clones carrying a mutation, ranged between 9.7% and 21.9% for Kdm6a, Kdm5c, and Ddx3x guides. Eif2s3x proved more problematic, with the first two guides yielding no mutant clones, and the third guide only one mutant from 40 clones screened. Guides that successfully cut PCR products and showed activity in mESCs were microinjected into pronuclear stage MF1 mouse embryos by Genetic Manipulation Services (Science Technology Platform at Francis Crick Institute), and transferred to pseudopregnant females in order to generate stable mouse lines.

Lines generated for each gene will now be discussed in more detail.

**Figure 5.4: Demonstrating sgRNA cutting of target genomic DNA.** Agarose gel showing Cas9 with sgRNA digest of PCR amplicons derived from target locus. Each PCR amplicon is shown first as uncut (U), followed by results of a incubation with Cas9 and each guide (sgRNA1,2 etc.) specific to the locus. Two different amplicons were utilised for Eif2s3x.
### Table 5.3: Guide efficacy as tested in XY mESCs

<table>
<thead>
<tr>
<th>Gene</th>
<th>sgRNA name</th>
<th>Mutant clones/total clones</th>
<th>Percentage carrying a mutation</th>
<th>Screening method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kdm6a</td>
<td>Kdm6a 1</td>
<td>5/32</td>
<td>15.6</td>
<td>RFLP</td>
</tr>
<tr>
<td></td>
<td>Kdm6a 2</td>
<td>4/22</td>
<td>18.2</td>
<td>RFLP</td>
</tr>
<tr>
<td>Kdm5c</td>
<td>Kdm5c 1</td>
<td>5/30</td>
<td>16.7</td>
<td>RFLP</td>
</tr>
<tr>
<td></td>
<td>Kdm5c 2</td>
<td>3/31</td>
<td>9.7</td>
<td>RFLP</td>
</tr>
<tr>
<td>Ddx3x</td>
<td>Ddx3x 1</td>
<td>7/32</td>
<td>21.9</td>
<td>Sanger seq.</td>
</tr>
<tr>
<td></td>
<td>Ddx3x 2</td>
<td>5/28</td>
<td>17.9</td>
<td>Sanger seq.</td>
</tr>
<tr>
<td>Eif2s3x</td>
<td>Eif2s3x 1</td>
<td>0/40</td>
<td>n/a</td>
<td>Sanger seq.</td>
</tr>
<tr>
<td></td>
<td>Eif2s3x 2</td>
<td>0/40</td>
<td>n/a</td>
<td>Sanger seq.</td>
</tr>
<tr>
<td></td>
<td>Eif2s3x 3</td>
<td>1/40</td>
<td>2.5</td>
<td>RFLP</td>
</tr>
</tbody>
</table>

#### 5.2.4 Kdm5c

Two lines carrying mutations in *Kdm5c* were generated and propagated. One line carried a 189bp deletion and the other a 166bp deletion. Both mutant lines featured the removal of the translation initiation site (Figure 5.5A, red line) and, based on sequencing evidence, also resulted in frameshift mutations within exon 1. All four protein coding isoforms of *Kdm5c* were likely affected (Figure 5.5A, lower). This loss of protein was confirmed by western blot (Figure 5.5B), using an antibody raised to antigen from outside of the deleted region (Figure 5.5A, ‘Ab’). I compared females heterozygous for the mutant allele with wildtype females (Figure 5.5B +/- and +/- respectively), and expected to see a reduction in protein in the mutant: this proved to be correct. I then compared males carrying only a mutant allele with wildtype males (Figure 5.5B -/y, +/-, respectively), and expected to observe complete absence of protein in the mutant: this also proved to be correct. I wanted to compare homozygous mutant females with wildtype females, expecting to see complete absence of protein in the homozygotes. However, I was unable to generate homozygous mutant females. In summary, both mutations resulted in ablation of KDM5C protein expression.

Next I wanted to check for the presence of off target mutations. I decided to proceed with off target analysis for the 189bp line only, as a single line was required for
5.2. Results

the weighing experiment. I chose to screen loci located within exons, and having four or fewer mismatches between guide and target sequence. These were identified using the results from the MIT CRISPR Design Tool. The off target analysis was performed on three male F1s that were used to generate all mutant animals for future experiments. I also sequenced three unrelated wildtype males to capture any background non specific mutations. No mutations were observed at any of the off target loci screened, in either 189bp deletion males or wildtype males (Figure 5.5C), based on PCR and next generation sequencing data. I conclude that I had successfully generated a Kdm5c mutant line with no detectable off target effects.
5.2. Results

A) sgRNA+PAM

B) KDM5C

C) Locus: 12bp seed

189bp del

wildtype

variant
5.2. Results

Figure 5.5: Generation of mutant Kdm5c alleles by CRISPR. (A) Cartoon depiction of the target locus. Wildtype (wt) sequence enables visualisation of the guide and PAM, along with the annotated ATG (red line). Line 1 and line 2 show the origins of the mutations generated following injection of Cas9+sgRNA into pronuclear stage mouse embryos. This is contextualised below, with exons shown as orange boxes. Splice variants of the gene are used to indicate positions of important functional domains, along with the predicted result of the mutations on protein structure. (B) Western blot on protein extracted from brain of seven day old pups with genotypes as indicated. The location of the epitope used to raise the antibody is indicated by Ab in (A). GAPDH was used as a loading control. (C) Off target screening by MiSeq sequencing of amplicon PCRs. The wildtype guide sequence and PAM are shown on the top line, with predicted off target sequences, organised by increasing variation from the guide sequence (orange), shown below. On the right hand side, sequencing results are summarised as filled boxes. Relative abundance of reads per individual animal is depicted by fill level of the squares: for example, the each of the heterozygous females shows 50:50 blue:orange to represent 50% read variation from wildtype. Three mutant males from generation F1 and three unrelated wildtype males were sequenced. Amplicons were sequenced to depths of between 200X and 10000X.

5.2.5 Kdm6a

A Kdm6a mutant line was generated harbouring a 285bp deletion in exon 1, just downstream of the translation initiation site (Figure 5.6A). This deletion was predicted to result in a frameshift, leading to a premature stop codon 13 amino acids after translation initiation and no protein product. I sought to verify this by western blot, using an antibody raised to the C terminal end of KDM6A (Figure 5.6A, red line). I compared protein extracted from wildtype female brain with heterozygous mutant female brain, expecting to see a reduction in the intensity of the protein band in the mutant. Surprisingly, there was no difference between wildtype and heterozygous mutant females (Figure 5.6B). I also compared wildtype male with hemizygous mutant male, expecting to detect no protein in the mutant. Once again, there was no difference between wildtype and mutant. A number of alternative antibodies were tested, but all gave multiple non specific bands under various different conditions. The antibody presented in Figure 5.6B produced a single band of the appropriate size, however, further work is required to confirm the specificity of the blot.
Despite not finding a difference at the protein level, I decided to look for a phenotype in my \textit{Kdm6a} mutant line. I noted that I was not able to generate homozygous mutant females, which was consistent with previous reports of mutations in \textit{Kdm6a} (Shpargel et al., 2012; Welstead et al., 2012). Furthermore, hemizygous mutant males allele were smaller than XY littermates at birth (Figure 5.6C(i)). Although this difference did not reach significance, it was again consistent with previous reports. A subset of these animals were followed through to weaning at three weeks: 100% of the wildtype males survived, whereas only 22% of the hemizygous mutants survived (Table 5.4). These phenotypic results were in complete agreement with the literature, further questioning the specificity of the western blot, and showing that a non functioning \textit{Kdm6a} allele had been generated.

Finally, I screened three F1 males for off target mutations using the same approach as that implemented for the \textit{Kdm5c} line. No mutations were observed at any of the loci sequenced (Figure 5.6C), and I conclude that the targeting of \textit{Kdm6a} was successful.

<table>
<thead>
<tr>
<th></th>
<th>XY</th>
<th>XY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Expected</td>
</tr>
<tr>
<td>PN1</td>
<td>16 (0.53)</td>
<td>15 (0.5)</td>
</tr>
<tr>
<td>PN21</td>
<td>16 (0.53)</td>
<td>15 (0.5)</td>
</tr>
</tbody>
</table>

\textbf{Table 5.4:} Numbers of wildtype and \textit{Kdm6a} hemizygous males recorded across five litters produced from a wildtype male and a heterozygous mutant female cross. Expected numbers and proportions based on male pups only, i.e. the null hypothesis of 50% XY and 50% X\textsuperscript{-}Y per litter.
5.2. Results

A

\[
\text{sgRNA+PAM} \quad \text{wt} \quad \text{ggaggctgctgggttctcagaaatcctgcgttccatgaaatcctg} \\
\text{line 1} \quad \text{ggaggctgctgggttctcagaaatcctg} \quad \text{285bp del} \\
\]

Peptide

Kdm6a-001/2

285bp del. Premature stop codon created

B

KDM6A

GAPDH

C

\[ \text{Weight(g)} \]

\[ \text{XY} \quad \text{X'Y} \]

\[ P = 5.38 \times 10^{-3} \]

D

(i) 12bp seed

OT1 TGGTTCCTCATGAAATCCTGCGG

OT2 TGGTTCCTCATGAAATCCTGCGG

OT3 TGGTTCCTCATGAAATCCTGCGG

OT4 TGGTTCCTCATGAAATCCTGCGG

OT5 TGGTTCCTCATGAAATCCTGCGG

OT6 TGGTTCCTCATGAAATCCTGCGG

(ii) 285bp del wildtype

variant
5.2. Results

Figure 5.6: Generation of mutant Kdm6a allele by CRISPR. (A) Cartoon depiction of the target locus. Wildtype (wt) sequence enables visualisation of the guide and PAM, along with the annotated ATG (red line). Line 1 shows the origin of the mutation generated following injection of Cas9+sgRNA into pronuclear stage mouse embryos. This is contextualised below, with exons shown as orange boxes. Splice variants of the gene are used to indicate positions of important functional domains, along with the predicted result of the mutations on protein structure. (B) Western blot on protein extracted from brain of seven day old pups with genotypes as indicated. The location of the peptide sequence used to raise the antibody is indicated by peptide in (A), GAPDH was used as a loading control. (C) Graph showing birth weight of XY and X Y pups: each point represents an individual, with mean and 95% CI as bars. $P$ value is calculated from Student’s t. (D) Off target screening by MiSeq sequencing of amplicon PCRs. The wildtype guide sequence and PAM are shown on the top line, with predicted off target sequences, organised by increasing variation from the guide sequence (orange), shown below. On the right hand side, sequencing results are summarised as filled boxes. Relative abundance of reads per individual animal is depicted by fill level of the squares: for example, the each of the heterozygous females shows 50:50 blue:orange to represent 50% read variation from wildtype. Three mutant males from generation F1 and three unrelated wildtype males were sequenced. Amplicons were sequenced to depths of between 200X and 10000X.

5.2.6 Ddx3x

A 17bp deletion within exon 1 of Ddx3x was created, downstream of the translation initiation site (Figure 5.7A). This mutation was predicted to result in the generation of a premature stop codon after six amino acids. To verify that the protein was no longer present (Figure 5.7B), western blot was performed using an antibody raised to the whole protein. Wildtype female liver was compared with heterozygous mutant female liver, and a reduction in DDX3X in the mutant tissue was expected. A slight reduction in band intensity was observed. When wildtype XY male liver was compared with hemizygous mutant male liver, complete absence of DDX3X in the mutant was expected. Interestingly, only a slight reduction of DDX3X protein was observed in the mutant. In contrast to the Kdm6a and Kdm5c mutant lines, for Ddx3x I was able to generate homozygous mutant females (Figure 5.7B: -/-). Contrary to predictions, tissue from these animals showed a marked depletion, but not complete loss, of DDX3X protein band by western blot. Based on this result, I concluded that I had generated a DDX3X knockdown line. No off target mutations were observed in the three F1 females screened from this line (Figure 5.7C).
5.2. Results

A

\[
\begin{array}{ll}
\text{wt} & \text{CTTCAGGGATGAGTCATGTGGCAATGGAAAAATGCGGCTGGGCTTGAGCAAGCGGAGTgagccctg} \\
\text{line 1} & \text{CTTCAGGGATGAGTCATGTGGCAATGGAAAAATGCGGCTGGGCTTGAGTGtgcgcctg} \\
\end{array}
\]

\[
\text{17bp del} \rightarrow \text{AGtgagccctg}
\]

Ddx3x-001

17bp del. premature stop codon

B

\[
\text{DDX3X} \\
\text{GAPDH}
\]

C

(i) 12bp seed

(ii) 17bp del

wildtype

variant

<table>
<thead>
<tr>
<th>Locus</th>
<th>AGTGGAAATGGCGGCTGGGCTGG</th>
<th>AGAGGAAAGTCAGCTGGGAAG</th>
<th>AGTTTAAAGCGCGCTGGGCGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>OT1</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
</tr>
<tr>
<td>OT2</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
</tr>
<tr>
<td>OT3</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
</tr>
<tr>
<td>OT4</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
</tr>
<tr>
<td>OT5</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
</tr>
<tr>
<td>OT6</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
</tr>
<tr>
<td>OT7</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
<td>AAGAGAAATGGCGGCTGGGCTGG</td>
</tr>
</tbody>
</table>
5.2. Results

Figure 5.7: Generation of mutant Ddx3x allele by CRISPR. (A) Cartoon depiction of the target locus. Wildtype (wt) sequence enables visualisation of the guide and PAM, along with the annotated ATG (red line), and alternative ATG (*). Line 1 shows the origin of the mutation generated following injection of Cas9+sgRNA into pronuclear stage mouse embryos. This is contextualised below, with exons shown as orange boxes. Splice variants of the gene are used to indicate positions of important functional domains, along with the predicted result of the mutations on protein structure. (B) Western blot on protein extracted from brain of seven day old pups with genotypes as indicated. The location of the peptide sequence used to raise the antibody is indicated by peptide in (A). GAPDH was used as a loading control. (C) Off target screening by MiSeq sequencing of amplicon PCRs. The wildtype guide sequence and PAM are shown on the top line, with predicted off target sequences, organised by increasing variation from the guide sequence (orange), shown below. On the right hand side, sequencing results are summarised as filled boxes. Relative abundance of reads per individual animal is depicted by fill level of the squares: for example, the each of the heterozygous females shows 50:50 blue:orange to represent 50% read variation from wildtype. Three mutant females from generation F1 and three unrelated wildtype females were sequenced. Amplicons were sequenced to depths of between 200X and 10000X.

5.2.7 Eif2s3x

Due to the very high degree of similarity between the sequence around the translation initiation site of Eif2s3x and its Y linked homologue Eif2s3y, I instead chose to target exon 7 of Eif2s3x (Figure 5.8A). Fewer live offspring were recovered when injecting guides targeting Eif2s3x, compared to Kdm5c, Kdm6a, and Ddx3x. Only two pups born from Eif2s3x injections carried frameshifted alleles. These were found at very low frequency (Figure 5.8B, box), and were not transmitted in the germ line. I considered two possibilities to explain these observations. Firstly, the guide may not have been efficient at cutting in the context of the mouse embryo: indeed, efficiency appeared very low in mESCs (Table 5.3). Secondly, Eif2s3x could be critical for early embryo development, and those embryos with a high percentage of frameshift mutations did not survive. This loss of embryos could impact on the development of embryos carrying wildtype or in frame alleles in the same uterus, resulting in a failed pregnancy and explaining the relatively low number of pups born.

To differentiate between the two hypotheses, 39 pronuclear stage embryos were injected with guides targeting Eif2s3x and cultured in vitro for three days to E3.5. 15 embryos arrested before reaching the expected morphology of a blastocyst at
5.2. Results

E3.5, and seven of these failed to give analysable sequencing data. Of the 32 embryos genotyped, all carried mutated alleles, with a very low proportion of wildtype reads (Figure 5.8B). This demonstrated that the guide was highly mutagenic. I also observed that the seven genotyped embryos that failed to develop to blastocyst generally showed a higher proportion of frameshift alleles compared to those successfully reaching this stage (51.4% vs. 30.8%). One embryo arrested at the morula stage (E2.5) and contained only frameshift alleles, and one blastocyst stage embryo had >99% of reads with frameshift mutations. I conclude that mutations in Eif2s3x did not preclude the completion of pre implantation development, as assessed by embryo morphology.

Finally, I attempted to elucidate the timing of embryonic lethality in Eif2s3x mutants by utilising pronuclear injection, followed by embryo transfer, and harvest at E6.5, E9.5 and E12.5. For each of the timepoints at least 50 injected embryos were transferred (Table 5.5). Approximately 50% of recipient females were pregnant and contained implantation sites in the uterus. 50% of these sites contained viable embryos at E6.5, 29% at E9.5, and 50% at E12.5 (Table 5.5). The remaining sites had supported implantation, but there was no embryonic tissue found at the time of dissection due to conceptus reabsorption.

The vast majority of the morphologically normal embryos were entirely wildtype (Figure 5.8B), based on genotyping by next generation sequencing. Two embryos at E6.5 harboured frameshift alleles, but as a low proportion of the overall reads. No frameshift alleles were detected at either E9.5 or E12.5, though in frame mutations were picked up in three embryos at E9.5 and one embryo at E12.5. Taken together, these data suggest that Eif2s3x plays a critical role in mouse embryonic development. As I was unable to identify any mutants with significant contribution from frameshift alleles later than E3.5, I concluded that these conceptuses fail peri implantation.
### Table 5.5: Targeting Eif2s3x in mouse embryos

Microinjection and surgical transfers were carried out by Sophie Wood and Katharine Mankelow in Genetic Manipulation Services.

<table>
<thead>
<tr>
<th>Time</th>
<th>Embryos transferred</th>
<th>Transferred females</th>
<th>Harvested embryos</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Pregnant</td>
<td>Not pregnant</td>
</tr>
<tr>
<td>E6.5</td>
<td>70</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>E9.5</td>
<td>51</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>E12.5</td>
<td>52</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>
5.2. Results

Figure 5.8: Generation of mutant Eif2s3x alleles by CRISPR. (A) Cartoon depiction of the target locus. Wildtype (wt) sequence enables visualisation of the guide and PAM, which is contextualised below, with exons shown as orange boxes. Splice variants of the gene are used to indicate positions of important functional domains. (B) Graphical summary of sequencing results following microinjection of Cas9 and sgRNA into pronuclear stage mouse embryos, followed by harvest at times indicated. Each column represents a single embryo with >200X coverage of the target locus. Mutations were categorised as indicated by the colour scheme.

5.2.8 Haploinsufficiency for Kdm5c, Kdm6a, or Ddx3x does not fully recapitulate the postnatal XO growth deficit

I next sought to utilise these new mutant alleles to test whether haploinsufficiency for Kdm5c, Kdm6a, or Ddx3x would recapitulate the growth deficit phenotype ob-
served in XO female mice. *Eif2s3x* was excluded as I was not able to generate a mutant line, as described above.

In order to generate females heterozygous for a mutant allele alongside wildtype female littermates, I crossed heterozygous females with wildtype males (Figure 5.9A(i), B(i), C(i)). Compared to XX littermates, *Kdm5c* heterozygous mutant females were lighter at birth, though this difference did not reach significance (Table 5.6: *p*=6.90x10⁻²). *Kdm6a* heterozygous mutant females were not significantly different in weight from XX littermates at birth (Table 5.6, *p*=4.46x10⁻¹). *Ddx3x* heterozygous mutant females were slightly, but not significantly, smaller than XX littermates at birth (Table 5.7: *p*=4.67x10⁻¹).

During the 0-3 week period, there was no significant difference in the rate of growth between wildtype XX females and *Kdm5c* heterozygous mutant females (Figure 5.9A(ii), Table 5.7: *p*=1.68x10⁻¹). There was also no significant difference in rate of growth between XX females and *Kdm6a* heterozygous mutant females across 0-3 weeks (Figure 5.9B(ii), Table 5.7: *p*=4.80x10⁻¹); and no significant difference was found between wildtype XX females and *Ddx3x* heterozygous mutant females across the same time period (Figure 5.9C(ii), Table 5.7: *p*=9.03x10⁻¹).

Finally, I compared the rates of growth for heterozygous mutant females to XX wildtype female littermates during the 3-5 week period. *Kdm5c* heterozygous mutant females gained weight at a slightly reduced rate relative to XX littermates, a difference that was borderline statistically significant (Table 5.7: *p*=5.30x10⁻²). There was no difference in rate of growth between *Kdm6a* heterozygous mutant females and their XX wildtype female littermates during this period (Table 5.7: *p*=4.81x10⁻¹). Interestingly, and similar to *Kdm6a* females, *Ddx3x* heterozygous mutant females gained weight at a reduced rate compared to their XX wildtype female siblings (Table 5.7: *p*=3.00x10⁻²). These general trends were reflected when analysing the mean weighted difference from wildtype XX female littermates for each genotype at each timepoint (Figure 5.9D).
5.2. Results

A

(i)

(ii)

B

(i)

(ii)

C

(i)

(ii)

D

(iii)

Kdm5c

Body weight (g)

Time (days)

Kdm6a

Body weight (g)

Time (days)

Ddx3x

Body weight (g)

Time (days)

-1.5

-1.0

-0.5

0.0

0.5

1.0

MWD from XX (g)

Time (days)
Figure 5.9: Assessment of postnatal growth in heterozygous mutant females compared to wildtype littermates. (A-C) Mating schemes to produce heterozygous mutant females alongside wildtype littermates are depicted on the left hand side. Weight data are shown on the right hand side, with each data point representing an individual animal at a single time point. Coloured lines show the mean weight per timepoint for each genotype. (A) shows $Kdm5c$; (B) shows $Kdm6a$; (C) shows $Ddx3x$. (E) Mean weighted difference between genotypes indicated and XX littermates. Bars represent 95% confidence intervals.
<table>
<thead>
<tr>
<th>Comparison</th>
<th>Number</th>
<th>Birth weight (g)</th>
<th>MWD±SE (g)</th>
<th>Signif. of MWD (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A v B</td>
<td>na</td>
<td>nb A</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>XX v XX$^{Kdm5c}$</td>
<td>38</td>
<td>33</td>
<td>1.669±0.032</td>
<td>1.596±0.033</td>
</tr>
<tr>
<td>XX v XX$^{Kdm6a}$</td>
<td>27</td>
<td>43</td>
<td>1.650±0.044</td>
<td>1.612±0.026</td>
</tr>
<tr>
<td>XX v XX$^{Ddx3c}$</td>
<td>19</td>
<td>20</td>
<td>1.735±0.028</td>
<td>1.688±0.029</td>
</tr>
</tbody>
</table>

**Table 5.6:** Birth weight compared by mean weighted difference (MWD) for XX vs. XX$^c$

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Number</th>
<th>0-3 weeks: mean ±SE (g)</th>
<th>3-5 weeks: mean ±SE (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A v B</td>
<td>na</td>
<td>nb xA$^a$ xB$^a$</td>
<td>xA$^a$ xB$^a$</td>
</tr>
<tr>
<td>XX v XX$^{Kdm5c}$</td>
<td>38</td>
<td>33 4.140±0.088 4.074±0.073</td>
<td>0.093±0.067</td>
</tr>
<tr>
<td>XX v XX$^{Kdm6a}$</td>
<td>27</td>
<td>43 4.212±0.173 4.095±0.154</td>
<td>0.060±0.085</td>
</tr>
<tr>
<td>XX v XX$^{Ddx3c}$</td>
<td>19</td>
<td>20 3.615±0.200 3.638±0.195</td>
<td>0.013±0.114</td>
</tr>
</tbody>
</table>

**Table 5.7:** Means and mean weighted differences (MWD) for phasic weight gain between genotype comparisons shown in Figure 5.8. "a" represents mean of means.
5.3 Discussion

The results presented here show that XO female mice gain weight at a reduced rate relative to their XX littermates during the first three weeks of postnatal life. This effect is not attributable to imprinting, as it was observed in both X\(^{P}\)O and X\(^{M}\)O females. Furthermore, the growth deficit remains following the addition of a second PAR, and is only rescued by the presence of a second sex chromosome. These data are consistent with previous data (Burgoyne et al., 2002), and implicated haploinsufficiency for an X linked gene that escapes XCI and has a broadly expressed Y linked homologue. There were, however, a number of inconsistencies with other reports, primarily concerning birth weights.

5.3.1 Comparison of birth weight data

It has previously been reported that relative to XX females, X\(^{P}\)O females are underweight at birth (Burgoyne et al., 1983a, 2002), though here I found no significant difference. These results likely reflect low sample size and reduced statistical power. If the number of animals were increased, it is highly likely that standard error would decrease and the magnitude of the difference between genotypes would be more apparent, becoming statistically significant.

X\(^{M}\)O and XY\(^{*X}\) female embryos were significantly smaller than XX littermates at birth, where no difference was reported previously (Burgoyne et al., 2002). Here, it is unlikely that increasing sample size would have a significant effect on reducing the standard error, the value of which is relatively in line with the standard error in XX littermates for both genotypes. There are two further datasets with some relevance to the current discrepancy. Thornhill & Burgoyne (1993) reported that at E10.5, X\(^{M}\)O embryos were significantly larger than XX littermates; and the same observation was made at E17.5 (Ishikawa et al., 2003). Additionally, Ishikawa et al. noted no difference between XX and XY\(^{*X}\) embryo weights at E17.5, but X\(^{P}\)O embryos were significantly smaller than XX littermates. Taken together, these results suggest that the birth weight deficit noted in X\(^{M}\)O and XY\(^{*X}\) females is unlikely to be due to carryover of an in utero phenotype, though this cannot formally be
excluded without collecting further data from E18.5. The discrepancy is also unlikely to be the result of genetic differences between the animal cohorts, because mice used in the present study were direct descendants of those used in the previous study, maintained on the same strain, and generated from the same crosses. It is difficult to formally exclude whether unknown genetic differences were present then, and/or are now present in these mice, and could impact on the results. I could test this hypothesis, and therefore confirm an $X^M O$ and $X Y^* X$ birth weight deficit, by repeating the experiment in a different mouse strain.

5.3.2 $X Y^*$ females at birth and post weaning

Two non significant trends in the Burgoyne et al. dataset became significant differences in the present dataset. Burgoyne et al. reported that $X Y^*$ females were slightly heavier than $X X$ littermates at birth, and here I utilised an increased sample size to show that this difference was statistically significant. Moreover, our larger sample also showed that $X Y^*$ females then lag behind $X X$ littermates in rate of growth during the 3-5 week period. These results could be due to a $Y$ linked acceleration effect during embryonic and perinatal growth, which then drops off postnatally to the extent that a growth rate deficit ensues. Most $Y$ chromosomes have a mechanistically undefined accelerating effect on mouse embryonic development, first identified at the pre implantation stages (Burgoyne, 1993; Tsunoda et al., 1985). Increased growth relative to $X X$ female littermates is carried over into the postimplantation period, and has been shown present up to E17.5 (Ishikawa et al., 1999). Analysis of birth weight data collected from $X Y$ males during the present study extends this trend further, showing that $X Y$ males are also significantly larger than $X X$ females at birth (Table 5.8). Moreover, of most relevance here, $X Y^*$ females are heavier than $X X$ littermates at E10.5 (Burgoyne et al., 1995). In order to test whether the increased birth weight results from a $Y$ linked accelerating effect in $X Y^*$ females, a $Y$ chromosome variant that lacks the preimplantation acceleration could be modified by removal of $S r y$, using CRISPR Cas9 genome engineering, to generate $X Y^{RIII/-}$ females. As described in Chapter 3, the RIII variant $Y$
5.3. Discussion

5.3.3 Increasing the XO:XX weight difference

The magnitude of the difference in both absolute weights and rates of growth between XO females and XX littermates is very small, and this was acknowledged in the previous report (Burgoyne et al., 2002). A proportional, reproducible increase in this difference could increase the sensitivity of subsequent experiments to detect potentially incremental contributions from Kdm5c, Kdm6a, Ddx3x, or Eif2s3x. This might be achieved by more faithful modeling of the human phenotype, or through adding a stressor to the experimental system that may preferentially impact on XO embryos.

A significant proportion of the human Turner syndrome phenotype occurs as a result of haploinsufficiency of SHOX, a gene located within the PAR and involved in chondrogenesis (Ellison et al., 1997; Rao et al., 1997, 2001). Whilst there is no SHOX orthologue in the mouse, an orthologue for the closely related human gene SHOX2 is present on chromosome 3 (Rovescalli et al., 1996) (Clement-Jones et al., 2000). Mouse SHOX2 is 99% homologous to human SHOX2 at the amino acid level, and shows very similar expression patterns (Blaschke et al., 1998; Semina et al., 1998). Shox2 mutations in mouse have been shown to result in embryonic

---

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Number</th>
<th>Birth weight (g)</th>
<th>MWD±SE (g)</th>
<th>Signif. of MWD (p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>XX v XY</td>
<td>29 v 28</td>
<td>1.650±0.032</td>
<td>1.732±0.045</td>
<td>0.090±0.026</td>
</tr>
</tbody>
</table>

*Table 5.8: Birth weight compared by mean weighted difference (MWD) for XX vs. XY*
lethality around E12.5, due to cardiac and vascular defects (Yu, 2005). Further work targeting the mutation to the limb using a specific Cre driver revealed that complete lack of the protein results in almost complete absence of humerus and femur (Cobb et al., 2006). Whilst heterozygous Shox2 mutants had no obvious abnormal limb phenotype, no quantitative data were provided (Cobb et al., 2006), and it would be interesting to analyse these mice more thoroughly. Perhaps a Shox2 mutant would synergise with the existing XO growth deficit, resulting in a more easily distinguishable phenotype. Such a model would further facilitate the detection of incremental contributions from the haploinsufficiency candidates Kdm5c, Kdm6a, Ddx3x, and Eif2s3x.

An alternative method to achieve a similar result would be caloric restriction of the pregnant dam for a defined period during pregnancy. A 50% reduction in caloric intake from E12.5 to E18.5 has been shown to cause reduced weight in utero at E16.5 and birth, followed by a catch up at around 3-4 weeks (Jimenez-Chillaron et al., 2005). Notably, there was also a significant reduction in litter size at birth, and other physiological effects result from in utero under nutrition, such as glucose intolerance (Jimenez-Chillaron et al., 2005). Taking these effects into consideration, it would nevertheless be interesting to compare XO with XX females in undernourished litters, as the physiological stress could have a more significant effect on those already compromised by XO genotype. A previously described example of this effect highlights the preferential loss of embryos carrying a single paternal X over XX littermate in the context of an XO uterine environments (Hunt, 1991).

5.3.4 The postnatal growth phenotype of heterozygous females mutant for Kdm5c, Kdm6a, or Ddx3x

One of the main aims of this chapter was to generate females heterozygous for each of the haploinsufficiency hypothesis candidate genes, and assess for recapitulation of the postnatal growth retardation phenotype. Whilst such females were generated for three out of four genes, these animals did not fully reproduce the phenotype
reported here and previously when comparing XO and XX littermates. Female mice heterozygous for a mutation in either Kdm5c, Kdm6a, or Ddx3x were born with a reduced weight relative to XX littermates, though the difference reached borderline significance only for Kdm5c. There was no significant difference in rate of growth between females heterozygous for either Kdm5c, Kdm6a, or Ddx3x and XX littermates across the 0-3 week period. During 3-5 weeks, Ddx3x heterozygous mutants gained weight at a significantly reduced rate relative to their XX siblings. The difference in rate of growth was borderline significant for Kdm5c heterozygous mutants relative to XX littermates and there was no difference in Kdm6a mutants. These observations could be explained by a number of factors, including the impact of random XCI on escapee expression, the need for mutants carrying mutations in all four genes concurrently, or an incorrect hypothesis.

XX female mice undergo XCI, inactivating one X chromosome in every somatic cell. Most X linked genes are subject to this process, though a small number escape and are expressed from both X chromosomes, often in specific tissues and/or at a set developmental time (Disteche et al., 2002). We showed here, as has also been shown previously, that Kdm5c, Kdm6a, Ddx3x, and Eif2s3x, all escape XCI in two different adult tissues. Extensive data in the literature extend this to multiple tissues throughout the embryo and adult, across developmental time (Li and Carrel, 2008; Berletch et al., 2015; Disteche, 2016; Balaton and Brown, 2016). Interestingly, a number of these escapees show significant female expression bias at the RNA and protein levels, implying that two active copies of the gene results in higher expression (Reinius et al., 2012; Xu et al., 2002, 2008a,b; Isensee et al., 2007). However, it seems that expression from the escapee alleles on the active X chromosome is not equivalent to alleles on the inactive X chromosome. For example, Kdm5c expression was lower from the inactive X chromosome at between E6.5 and E13.5, though this difference reduced to almost 50:50 in adult tissues, as assayed by single nucleotide primer extension assay (Lingenfelter et al., 1998). More recent work utilising RNA-seq has shown decreased relative Kdm5c expression from the inactive X chromosome in 4 cell through to early blastocyst stage pre implantation embryos.
5.3. Discussion

(Deng et al., 2014). Using the same technique, similar patterns were observed in adult tissues for Kdm5c, Kdm6a, Ddx3x, and Eif2s3x (Berletch et al., 2015), and this has also been noted in microarray data (Reinius et al., 2012). Taken together, these data show that escapee genes are expressed more highly in XX females when compared to XY males, and a significantly higher proportion of the total transcripts originate from the active allele.

This XX:XY comparison should theoretically be similar to the XX:XO comparison regarding escapee gene expression, as both have a single X chromosome. If expression of a given escapee from the active X chromosome is arbitrarily represented as 1.0, and from the inactive X chromosome as 0.5, total expression in XX female somatic cells is 1.5 (Figure 5.10). In XO female somatic cells, there is no XCI, therefore expression from the single active X chromosome is 1.0, and total expression is 1.0 (Figure 5.10). XX⁻ females have two genetically distinct X chromosomes (due to the mutation) and, as a result, each somatic cell will inactivate one of these two X chromosomes at random: the mouse is mosaic. 50% of cells will inactivate the mutant X chromosome, therefore expression from the active wildtype X chromosome is 1.0 (Figure 5.10). In the other 50% of somatic cells, the mutant X chromosome will remain active, with expression of 0, and inactivation of the wildtype X chromosome leads to expression equivalent of the inactive X chromosome in XX cells, i.e. 0.5 (Figure 5.10). When these two levels are averaged across the animal, expression in XX⁻ females is therefore expected to be lower than in XO females, an arbitrary 0.75.

Whilst this example is purely theoretical, and does not take into account X chromosome upregulation (Sangrithi et al., 2017), or genome wide dosage modulation in aneuploidies (Birchler, 2014), the difference could still partly explain the discrepancy between observations of growth rate in XO females and XX⁻ females. Were this hypothesis to be correct, the presently described XX⁻ females have lower expression of each given gene than XO females.
5.3. Discussion

Loss of a single copy of a single gene does not, in itself, appear enough to reproduce the birth weight deficit or the 0-3 weeks rate of growth deficit, but may manifest as the novel post weaning phenotype. I could test the hypothesis, and eliminate random XCI as a variable, by making use of the \textit{Xist}- allele (Marahrens et al., 1997). XCI would become biased towards a specific, known X chromosome, allowing the dissection of effects relating to expression of either a single wildtype allele, or a single mutant allele. Such a model with multiple, reproducible expression levels could also serve to facilitate greater understanding of the relationship between gene dosage and phenotype, as these genes are widely acknowledged to be sensitive to dosage modulation (Bellott et al., 2014; Cortez et al., 2014).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5_10.png}
\caption{Model depicting XCI escapee expression in XX, XO, and XX$^{-}$ females. \textbf{Top:} cartoon bar graphs representing arbitrary escapee expression levels from each sex chromosome in XX, XO, and XX$^{-}$ female somatic cells. \textbf{Bottom:} Representation of the composition of mice from which the above cells are taken, i.e. block colour = 100\%.}
\end{figure}

Each of the heterozygous mutants used in the present study carries one mutant allele for a single gene, although four genes have been identified and targeted as part of the haploinsufficiency hypothesis. It therefore seems logical, and likely, that a female carrying a mutation all four genes will recapitulate the XO phenotype to a greater extent. Whilst this strategy was initially considered, the likelihood of being able to generate enough female animals carrying four mutations on the same chromosome, alongside wildtype littermates, to make biologically relevant conclusions is very small. Instead, utilising the XY$^{-}$ female would be more straightforward (notwithstanding the experiments proposed in 5.3.2), as the Y chromosome does not un-
5.3. Discussion

dergo recombination or XCI. Even taking into account the relatively poor breeding performance of XY females (A. Ojarikre, personal communication), these experiments would undoubtedly be more achievable than the compound X linked mutant. Generating mutations in Y linked homologues of the haploinsufficiency candidates would also increased understanding of the roles of X:Y pairs not only in the context of growth, but also more generally in terms of survival. I briefly touched on this topic during the present chapter by demonstrating that \textit{Eif2s3x} is essential for postimplantation embryonic development (see 5.3.5).

5.3.5 Targeting and re targeting \textit{Eif2s3x}

I was unable to generate a mutant line for the fourth candidate highlighted in the Y expression and XCI escapee screens, \textit{Eif2s3x}. After a number of initial \textit{in vivo} attempts, I failed to generate mutant pups, and doubted the veracity of the guide RNA. I used injected embryos cultured to blastocyst primarily to test guide effectiveness and confirm that mutations were being generated. Only secondarily was I interested to see if there was any correlation between developmental progression and the mutation suite observed following sequencing. Whilst I achieved the former, I did not utilise a sufficient number of control embryos to report the success rate of culture to the blastocyst stage of embryos injected with a scramble guide. It is clear that the suite of mutations observed in preimplantation blastocysts is no longer present in later stage embryos, and the embryos carrying such mutations must therefore be eliminated prior to collection. However, based on the present blastocyst culture data and paucity of control data, I cannot conclude whether this elimination occurs pre or postimplantation, and further work is required.

The inability to generate a \textit{Eif2s3x} mutant line was not surprising, given the reported dosage sensitivity of the four haploinsufficiency candidates. However, it was notable because targeting of \textit{Kdm5c}, \textit{Kdm6a}, and \textit{Ddx3x} was almost immediately successful \textit{in vivo}. In hindsight, the apparently low success rate of the initial targeting of \textit{Eif2s3x} in mESCs could be interpreted as highly efficient targeting of a gene necessary for mESC survival. The discrepancy between success and failure is likely
a result of different guide design strategies and efficiencies: a highly efficient guide may increase the proportion of frameshift mutations (Haeussler et al., 2016), and facilitate observation of the dominant phenotype, i.e. lethality, more often. In terms of design, for this guide I targeted an exon in the centre of the coding sequence instead of the initiator methionine (ATG), due to high homology between \textit{Eif2s3x} and \textit{Eif2s3y}. Such a strategy may also be more efficient at ablating protein function (Li et al., 2016).

I did not see any \textit{Eif2s3x} mutant male animals born. It would be interesting to look at the sex of the embryos analysed at E3.5, E6.5, E9.5 and E12.5, as I would expect to see mutant males surviving over mutant females, given the presence of a 97% similar homologue on the Y chromosome. Indeed, it has previously been shown that \textit{Eif2s3x} is able to partly compensate for the absence of its Y linked homologue in spermatogenesis (Yamauchi et al., 2016). Based on the data presented here, no \textit{Eif2s3x} mutants with mutation suites reflecting those observed in the preimplantation embryo survived postimplanation development, let alone birth. I must therefore conclude that males and females are equally affected by the loss of \textit{Eif2s3x}, and \textit{Eif2s3y} is not able to provide functional compensation. The roles of \textit{Kdm5c}, \textit{Kdm6a}, and \textit{Ddx3x}, and their Y linked homologues, in postnatal survival are the subject of the next chapter.

Finally, as only a single guide sequence was utilised, and an off target screen was not carried out, it is formally possible that the lethality phenotype observed could be explained by an off target mutation. As methods for the detection of off target mutations is a highly contentious area (Iyer et al., 2015; Haeussler et al., 2016; Schaefer et al., 2017), the use of a number of different guides to achieve the same phenotype would provide the most robust confirmation. The work to re target \textit{Eif2s3x} by generating a conditional allele, followed by an off target screen, is ongoing.
Chapter 6

The contribution of X-Y homologous genes to postnatal survival in the mouse

6.1 Introduction

The mammalian sex chromosomes are thought to have originated from a pair of autosomes around 180-200 million years ago (Cortez et al., 2014). The proto Y chromosome evolved a dominant male determining locus, which resulted in the acquisition of sexually antagonistic alleles nearby, and this in turn led to the suppression of recombination with the proto X (Wright et al., 2016). Over time, mammalian Y chromosomes generally, and the mouse Y specifically, has lost, and gained, a significant amount of gene-encoding DNA (Soh et al., 2014); though a small number of genes have been conserved. These genes are the Y linked homologues of the X linked genes targeted in Chapter 5: Ddx3y, Eif2s3y, Kdm5d, and Uty. Both X-Y copies are broadly expressed, the X linked partner escapes XCI, and it is under more intense purifying selection than neighbours on the X chromosome (Bellott et al., 2014). Collectively, these characteristics suggest the longevity of X-Y pairs can be attributed to selection pressure resulting from dosage sensitivity: copies have been
maintained on both sex chromosomes because the organism cannot survive with only a single copy. This is likely true for humans, evidenced by the X chromosome monosomy Turner syndrome. Approximately 99% of XO female foetuses die in utero, and those that survive are often mosaic for at least part of a second sex chromosome (Cockwell et al., 1991; Hook and Warburton, 1983, 2014). In contrast, XO female mice survive, albeit with a slight growth deficit (Chapter 5) and reduced fertility (Burgoyne and Baker, 1981, 1985). The difference in survival between these two mammalian species may result from the relatively small number of mouse X linked genes that require two doses: there are only nine murine X-Y pairs, but 17 pairs in human (Bellott et al., 2014). The survival of XO female mice does, nevertheless, provide the opportunity to address X-Y pair functional divergence in mouse using this model.

Our current understanding of X-Y pair divergence in the mouse stems from a number of transgenic mouse models. Male mice lacking the whole Y chromosome are viable (Yamauchi et al., 2016). Yamauchi and colleagues showed that replacement of the Y chromosome with transgenic copies of $Eif2s3x$ and $Sox9$ could result in males capable of siring offspring. In contrast, results from experiments targeting X linked genes from X-Y pairs suggest functional divergence, as assessed by postnatal survival. Numerous $Kdm6a$ mutants have been published, and in all cases homozygous mutant females die in utero (Lee et al., 2012; Wang et al., 2012; Shpargel et al., 2012; Wesstead et al., 2012; Mansour et al., 2012; Thieme et al., 2013). Hemizygous mutant males showed reduced perinatal viability, and survivors exhibited a lifelong growth deficit. Similarly, females carrying homozygous mutations in $Ddx3x$ die in utero, but in contrast to $Kdm6a$, hemizygous mutant males do not survive beyond E6.5 (Chen et al., 2016a). Taken together, these data show that the Y linked genes from X-Y pairs are no longer required for normal postnatal survival in mouse; however, at least one copy of the X linked genes $Ddx3x$ and $Kdm6a$ are necessary. In this Chapter, I looked to first confirm these results using the mutant lines generated in house. I then utilised these animals, along with sex chromosome
variant lines, to explore the functional redundancy of X-Y genes pairs in postnatal survival.

6.2 Results

6.2.1 *Kdm6a* and *Kdm5c* are required for postnatal survival in female mice

It has previously been shown that XX females homozygous for mutations in either *Kdm6a* or *Ddx3x* die in utero (Lee et al., 2012; Shpargel et al., 2012; Welstead et al., 2012; Chen et al., 2016a). I first sought to confirm this result, and test whether the same conclusion could be applied to *Kdm5c*. To exclude XCI as a confounding factor when comparing the postnatal survival of heterozygous mutant females with homozygous mutant females, I used XO female mice, that do not undergo XCI. I crossed a heterozygous mutant female with a XYO male (Burgoyne et al., 1998; Odorisio et al., 1998), which could produce mutant X-O females along with wild-type XO females within the same litter (Figure 6.1A; Table 6.5, Comparison 1). I predicted that X-O females would not survive, consistent with previous data for *Kdm6a* and *Ddx3x*. Such a result would serve to emphasise the importance of at least one copy of each X linked gene from the X-Y gene pairs. I assayed survival by genotyping pups at PN7. There were no *Kdm6a* X-O mutant females present, in contrast to 27 XO wildtype females (Figure 6.1B, \( p = 1.54 \times 10^{-4} \); Table 6.1, adjusted \( p = 1.02 \times 10^{-4} \)). I conclude that at least one copy of *Kdm6a* is essential for postnatal survival.

For *Ddx3x* I noted the survival of 12 Ddx3x X-O mutant females, compared with 3 XO wildtype females (Figure 6.1C, \( p = 0.1207 \); Table 6.1, adjusted \( p = 0.202 \)). Although the ratio of X-O mutant to XO wildtype females was undoubtedly impacted by the low number of animals generated for the cross, the presence of any mutant X-O mutant females was unexpected, based on previously reported data (Chen et al., 2016b). This will be discussed further later.
For *Kdm5c*, I recovered a single X' O mutant female at PN7, alongside 26 XO wild-type females (Figure 6.1D, \( p=1.03 \times 10^{-3} \); Table 6.1, adjusted \( p=1.21 \times 10^{-3} \)). In conclusion, in XO females, the loss of the single copy of *Kdm6a* and *Kdm5c*, but not *Ddx3x*, compromises postnatal survival.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Number</th>
<th>XO</th>
<th>X O</th>
<th>adjusted ( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Kdm6a</em></td>
<td>27</td>
<td>0</td>
<td>1</td>
<td>1.02 \times 10^{-4}</td>
</tr>
<tr>
<td><em>Ddx3x</em></td>
<td>3</td>
<td>12</td>
<td>2</td>
<td>2.02 \times 10^{-1}</td>
</tr>
<tr>
<td><em>Kdm5c</em></td>
<td>26</td>
<td>1</td>
<td>1</td>
<td>1.21 \times 10^{-3}</td>
</tr>
</tbody>
</table>

*Table 6.1:* Postnatal survival for XO vs X' O females.
6.2. Results

A (i)

\[ \begin{array}{c|c|c|c}
X^- & X^Y & 0 \\
\hline
X^O & X^Y & 1 \\
\hline
O & & \\
\end{array} \]

(ii)

\[
\begin{array}{c|c|c|c}
0.0 & 0.5 & 1.0 \\
F & M & \\
\hline
0.5 & 0.5 & 0.0 \\
\end{array}
\]

B (i)

\[
\begin{array}{c|c}
Kdm6a & \\
\hline
0 & 27 \\
12 & 10 \\
\end{array}
\]

\[ p = 1.54 \times 10^{-4} \]

C (i)

\[
\begin{array}{c|c}
Ddx3x & \\
\hline
12 & 3 \\
12 & 3 \\
\end{array}
\]

\[ p = 0.1207 \]

D (i)

\[
\begin{array}{c|c}
Kdm5c & \\
\hline
1 & 26 \\
11 & 8 \\
\end{array}
\]

\[ p = 1.03 \times 10^{-2} \]
Figure 6.1: Postnatal survival of XO female mice carrying a single mutated allele, relative to XO littermates. (A)(i) Mouse cross used to generate X:O and XO female mice, and (ii) mosaic plot showing proportion of each genotype expected under the null hypothesis. The Y axis represents the relative proportion of female: male as assessed by gonadal sex, and the X axis breaks down the proportional contribution of each genotype per gonadal sex. The hatched line shows 0.5 on the X axis. (B)(i) Number of animals per genotype alive at postnatal day 7 for Kdm6a mutant line; numbers relate to genotypes in equivalent location in A(i). The p value is derived from Fisher’s exact test. Orange ring represents comparison of interest: see text. (C) As (B), for Ddx3x mutant line. (D) As (B), for Kdm5c mutant line.
6.2.2 Y linked genes partially compensate for loss of their X linked homologues

Next I looked to determine whether Y linked genes from X-Y pairs could compensate for loss of their X linked homologues. For this experiment, I wanted to compare survival of XX homozygous mutant females to survival of X-Y hemizygous mutant males. Given the results of the previous section, whereby Kdm6a and Kdm5c XO mutant females did not survive, survival of any hemizygous mutant males would suggest a degree of functional compensation. To generate homozygous mutant females alongside hemizygous mutant males, heterozygous mutant females were crossed with hemizygous mutant males (Figure 6.2A(i)).

For Kdm6a, as predicted, I observed no homozygous mutant females, and 9 hemizygous mutant males (Figure 6.2B, \( p=2.92 \times 10^{-2} \); Table 6.2, adjusted \( p=4.50 \times 10^{-2} \)). I conclude that the presence of the Y chromosome compensated for the loss of Kdm6a.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Number</th>
<th>X X</th>
<th>X Y</th>
<th>adjusted ( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kdm6a</td>
<td>0</td>
<td>9</td>
<td>4.50 \times 10^{-2}</td>
<td></td>
</tr>
<tr>
<td>Ddx3x</td>
<td>8</td>
<td>26</td>
<td>1.73 \times 10^{-2}</td>
<td></td>
</tr>
<tr>
<td>Kdm5c</td>
<td>1</td>
<td>23</td>
<td>3.46 \times 10^{-4}</td>
<td></td>
</tr>
</tbody>
</table>

*Table 6.2:* Postnatal survival for X X vs X Y animals

For Ddx3x, there were 8 homozygous mutant females, compared with 26 hemizygous mutant males (Figure 6.2C, \( p=9.95 \times 10^{-4} \); Table 6.2, adjusted \( p=1.73 \times 10^{-2} \)). This showed that the presence of the Y chromosome compensated for the loss of Ddx3x.

For Kdm5c, one homozygous mutant female was observed, compared with 23 hemizygous mutant males (Figure 6.2D, \( p=1.07 \times 10^{-4} \); Table 6.2, adjusted \( p=3.46 \times 10^{-4} \)). Taken together, these data suggested that the presence of a Y chromosome could compensate for the loss of X linked genes Kdm6a, Ddx3x, and Kdm5c.
In order to assess whether this compensation was complete or partial, I compared the survival of hemizygous mutant males with wildtype males (Table 6.5, Comparison 3). Complete compensation was predicted to result in no difference in postnatal survival between wildtype and hemizygous mutant males. I looked to crosses that would enable a direct comparison between X\textsuperscript{−}Y hemizygous mutant and XY wildtype males. As data were available from the X\textsuperscript{−}X x X\textsuperscript{−}Y cross detailed in Figure 6.2, and the X\textsuperscript{−}X x XY cross in Figure 6.3, a pool was created to increase statistical power (Table 6.3).

For \textit{Kdm6a}, hemizygous mutant males survived in significantly reduced numbers compared to XY males (Table 6.3, \(p=1.00\times10^{-2}\)) as described previously (Lee et al., 2012; Welstead et al., 2012; Shpargel et al., 2012) and in Chapter 5. I conclude that loss of \textit{Kdm6a} was not fully compensated for by its Y linked homologue, \textit{Uty}.

Hemizygous males mutant for either \textit{Ddx3x} or \textit{Kdm5c} survived in fewer numbers than XY wildtype males, though this difference did not reach significance (Table 6.3, \(p=2.89\times10^{-1}\) and \(3.31\times10^{-1}\) respectively). Taken together, these data show that Y linked genes \textit{Uty}, \textit{Ddx3y} and \textit{Kdm5d} are able to at least partially compensate for the loss of \textit{Kdm6a}, \textit{Ddx3x}, and \textit{Kdm5c}, respectively. This compensation may be more complete for the \textit{Kdm5c} and \textit{Ddx3x}, as there was no evidence of a significant difference in postnatal survival between hemizygous mutant males and XY wildtype males for these comparisons.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Number</th>
<th>XY</th>
<th>X\textsuperscript{−}Y</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{Kdm6a}</td>
<td>52</td>
<td>20</td>
<td></td>
<td>1.00\times10^{-2}</td>
</tr>
<tr>
<td>\textit{Ddx3x}</td>
<td>51</td>
<td>36</td>
<td></td>
<td>2.89\times10^{-1}</td>
</tr>
<tr>
<td>\textit{Kdm5c}</td>
<td>77</td>
<td>59</td>
<td></td>
<td>3.31\times10^{-1}</td>
</tr>
</tbody>
</table>

\textbf{Table 6.3:} Postnatal survival for pooled XY vs X\textsuperscript{−}Y males
### 6.2. Results

#### A(i)

![Graphical representation](image)

#### B(i)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

\[ p = 2.92 \times 10^{-2} \]

#### C(i)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>69</td>
<td>8</td>
<td>26</td>
</tr>
<tr>
<td>26</td>
<td>38</td>
<td></td>
</tr>
</tbody>
</table>

\[ p = 9.95 \times 10^{-4} \]

#### D(i)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>23</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

\[ p = 1.07 \times 10^{-4} \]
Figure 6.2: Postnatal survival of female mice carrying two mutated alleles, relative to males with one mutated allele and an intact Y chromosome. (A)(i) Mouse cross used to generate X^X^-, XX^-, X^-Y, and XY wildtype littersmates, and (ii) mosaic plot showing proportion of each genotype expected under the null hypothesis. The Y axis represents the relative proportion of female:male as assessed by gonadal sex, and the X axis breaks down the proportional contribution of each genotype per gonadal sex. The hatched line shows 0.5 on the X axis. (B) (i) Number of animals per genotype alive at postnatal day 7 for Kdm6a mutant line; numbers relate to genotypes in equivalent location in A(i). The p value is derived from Fisher’s exact test. Orange ring represents comparison of interest: see text. (C) As (B), for Ddx3x mutant line. (D) As (B), for Kdm5c mutant line.
6.2. Results

A (i) [Diagram showing sex chromosomes and gene expression]

B (i) [Table showing counts]

\[
\begin{array}{c|c|c}
Kdm6a & 41 & 54 \\
  & 11 & 40 \\
\end{array}
\]

\[ p = 4.68 \times 10^{-4} \]

C (i) [Table showing counts]

\[
\begin{array}{c|c|c}
Ddx3x & 20 & 18 \\
  & 10 & 13 \\
\end{array}
\]

\[ p = 0.5506 \]

D (i) [Table showing counts]

\[
\begin{array}{c|c|c}
Kdm5c & 52 & 52 \\
  & 36 & 54 \\
\end{array}
\]

\[ p = 0.5073 \]
Figure 6.3: Postnatal survival of male and female mice carrying a single mutated allele, relative to wildtype littermates. (A)(i) Mouse cross used to generate X^X, X^Y and wildtype littermates, and (ii) mosaic plot showing proportion of each genotype expected under the null hypothesis. The Y axis represents the relative proportion of female:male as assessed by gonadal sex, and the X axis breaks down the proportional contribution of each genotype per gonadal sex. The hatched line shows 0.5 on the X axis. (B)(i) Number of animals per genotype alive at postnatal day 7 for Kdm6a mutant line; numbers relate to genotypes in equivalent location in A(i). The p value is derived from Fisher’s exact test. Coloured ring represents comparison of interest: see text. (C) As (B), for Ddx3x mutant line. (D) As (B), for Kdm5c mutant line.
6.2.3 Some Y linked genes can compensate for loss of their X linked homologue independent of the hormonal milieu

Phenotypic sex differences between male and female mammals arise from the unequal effects of the sex chromosomes. Such effects can occur either directly, due to differential gene expression from the X and Y chromosomes in XX and XY non gonadal cells; or indirectly, via the gonads, resulting in hormonal differences throughout life (Burgoyne and Arnold, 2016). In the previous experiment, I compared postnatal survival between males and females without controlling for gonadal sex, and thus not taking into account potential confounding effects of sex hormones. I therefore looked to compare X-Y gene function within males and within females separately.

As shown in Table 6.3, I had already established that hemizygous mutant males survived in reduced numbers compared to wildtype males for all three mutant lines, though this difference was only statistically significant for Kdm6a. I then looked to test the female hormonal environment for the same X-Y pair effects. I assessed whether the Y linked gene could compensate for the loss of its X linked homologue within females.

For this experiment, I separated karyotypic sex from gonadal sex by using a male carrying a Y chromosome engineered to lack the sex determining gene Sry, carrying this gene instead as a transgene on an autosome (Gubbay et al., 1990). This XY<sup>Tdy</sup>Sry male (denoted XY<sup>-</sup> in Chapter 5) is thus able to produce four different genotypes of sperm. Two of these genotypes will, following fertilisation, produce female offspring, carrying either an X chromosome or a Y<sup>Tdy</sup> chromosome. The two others will result in male offspring post fertilisation, and these sperm carry either an X chromosome and the Sry transgene, or the Y<sup>Tdy</sup> chromosome and the Sry transgene. I crossed this male with a female heterozygous for one of the mutant alleles (Figure 6.4A(i,ii)). I compared postnatal survival in hemizygous mutant fe-
males ($X^{Y_{Tdy}}$) with $XY^{Tdy}$ females (Table 6.5, Comparison 4). If the Y linked
gene could fully compensate, I would expect no significant difference in survival
between the two genotypes.

For $Kdm6a$, only one $X^{Y_{Tdy}}$ mutant female survived, whereas there were four
$XY^{Tdy}$ females (Figure 6.4B(i,ii)). This cross also produced one hemizygous mu-
tant male and five wildtype males. I tentatively conclude that in both males and
females, the loss of $Kdm6a$ is not fully compensated for by the presence of the Y
chromosome.

I did not recover any $XY^{Tdy}$ control females from the $Ddx3x$ cross, therefore no
conclusions can yet be drawn.

For $Kdm5c$, there was no significant difference in survival of $X^{Y_{Tdy}}$ mutant fe-
ales survived compared to $XY^{Tdy}$ females (Figure 6.4D(i,ii); Table 6.4, adjusted
$p=9.20 \times 10^{-1}$). Consistent with a previous cross (Figure 6.3D(i,ii)) there were fewer
surviving ($X^{Y_{Tdy}}Sry$) mutant males than $XY^{Tdy}Sry$ males, though again this differ-
ence was not significant (Table 6.3). I conclude from these data that in both males
and females, the loss of $Kdm5c$ is fully compensated for by the presence of the Y
chromosome.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Number</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$XY^{Tdy}$</td>
<td>$X^{Y_{Tdy}}$</td>
<td></td>
</tr>
<tr>
<td>$Kdm6a$</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>$Ddx3x$</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$Kdm5c$</td>
<td>15</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 6.4: Postnatal survival for pooled $XY^{Tdy}$ vs X $Y^{Tdy}$ females

The total number of animals used for each experiment described here was sub opti-
mal, partly due to the complexity of the mouse cross and the resulting eight possible
offspring genotypes. Taking into account this caveat, I conclude that in the female
hormonal environment $Uty$ partially compensates for loss of its X linked homologue
6.2. Results

*Kdm6a*. *Kdm5d* appears to fully compensate for the loss of its X linked homologue *Kdm5c* in the female hormonal environment. Further work is required to make any conclusions regarding *Ddx3y* compensation for *Ddx3x*.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Genotype</th>
<th>Number of copies</th>
<th>X-linked allele</th>
<th>Y-linked allele</th>
<th>Sum X+Y</th>
<th>PAR</th>
<th>Male gonads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wildtype</td>
<td>XX</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>XY</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>XO</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>X'O</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>X'X'</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>X'Y</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>XY</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>X'Y</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>XY'Tdy</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>X'Y'Tdy</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Table 6.5:* X-Y pair copy number summary for all comparisons
6.2. Results

A (i)

B (i)

\[
\begin{array}{cccc}
Kdm6a & 2 & 1 & 8 & 4 \\
      & 5 & 1 & 3 & 5 \\
\end{array}
\]

\[p = 0.6019\]

C (i)

\[
\begin{array}{cccc}
Ddx3x & 2 & 2 & 0 & 0 \\
      & 4 & 1 & 4 & 1 \\
\end{array}
\]

\[p = 0.6768\]

D (i)

\[
\begin{array}{cccc}
Kdm5c & 17 & 23 & 15 & 15 \\
      & 13 & 12 & 20 & 23 \\
\end{array}
\]

\[p = 0.8215\]
Figure 6.4: Postnatal survival of XX and XY male, and XX and XY female mice carrying a single mutated allele, relative to wildtype littermates. (A)(i) Mouse cross used to generate X X, X Y and wildtype littermates, and (ii) mosaic plot showing proportion of each genotype expected under the null hypothesis. The Y axis represents the relative proportion of female: male as assessed by gonadal sex, and the X axis breaks down the proportional contribution of each genotype per gonadal sex. The hatched line shows 0.5 on the X axis. (B) (i) Number of animals per genotype alive at postnatal day 7 for Kdm6a mutant line; numbers relate to genotypes in equivalent location in A(i). The p value is derived from Fisher’s exact test. Orange ring represents comparison of interest: see text. (C) As (B), for Ddx3x mutant line. (D) As (B), for Kdm5c mutant line.
6.3 Discussion

Around 180-200 million years ago, one chromosome from a pair of ancestral mammalian autosomes acquired a male determining locus, setting it on the path to becoming the present day Y chromosome via a process of degeneration and evolutionary divergence. A small number of genes remained constant, and changed very little, between the current X and Y chromosomes; and it has been suggested that these genes were retained to maintain dosage.

6.3.1 The $Ddx3x$ mutant allele is a hypomorph, not a knockout

Here, I have utilised mutant alleles generated previously (Chapter 5) to interrogate the necessity of both members of these X-Y pairs in postnatal survival in mouse. I have shown that at least one copy of $Kdm6a$ is essential for survival; and this is likely also true for $Kdm5c$ and $Eif2s3x$. My data for $Ddx3x$ are different to that published by Chen and colleagues (2016). Their mutant allele consisted of an excised block of 14 exons. In contrast, I removed 17bp within exon 1 of $Ddx3x$, which was predicted to result in a frameshift and premature stop codon. Based on western blot data from homozygous female tissue, I conclude that the mutation resulted in almost complete ablation of protein expression (Chapter 5, Fig. 5.7). Furthermore, homozygous mutant females showed significantly reduced survival compared to X-Y males. Together, these data suggest that the 17bp mutation impairs DDX3X protein function.

Interestingly, there are two alternative ATG sites downstream of the annotated ATG, both in different reading frames (* in Chapter 5, Fig. 5.7). Whilst translation initiation sites can be accurately predicted using classical methods (Kozak, 1984), mutations have been associated with alternative translation products (Greene et al., 2003). Moreover, recent work has suggested that mutations induced by CRISPR/Cas can cause exon skipping, whereby the exon containing the deletion is omitted from the final protein product (Kapahnke et al., 2016; Prykhozhij et al.,...
6.3. Discussion

Taken together, these data would suggest that an alternative translation initiation site or splicing event could be giving rise to reduced levels of DDX3X expression in these mutant animals. This protein would function in a similar way to the wildtype DDX3X protein, such that only a mild phenotype is observed in females homozygous for the mutant allele. As this phenotype is inconsistent with the published phenotype, I will use the data reported by Chen et al. (2016) for further discussion of the roles of Ddx3x in postnatal survival.

6.3.2 Conservation of protein coding sequence does not equate to conservation of gene function

Based on the results presented here, there exists a spectrum of functional divergence between members of X-Y gene pairs in the mouse. Whilst Kdm5c undoubtedly contributes to postnatal survival, this function can be at least partially, if not fully, compensated for by the presence of the Y linked homologue Kdm5d in both males and females. If at least one copy of either gene is present, the animal will survive. This is in contrast to the evidence for Kdm6a, Ddx3x, and Eif2s3x. Regardless of gonadal sex, the Y linked gene Uty can only partly compensate for loss of its X linked homologue, Kdm6a. One copy of Ddx3x is essential for postnatal survival. In the absence of Ddx3x, Ddx3y supports development until E6.5 in XY males (Chen et al., 2016a). Further work is required to address the same question in hemizygous mutant females. Finally, utilising the preliminary embryo data from Chapter 5, Eif2s3x is also required in at least one copy for development to progress further than E6.5.

Previous work has reported that X-Y pair genes are under more intense purifying selection than other genes on the X chromosome, as evaluated by the ratio of non synonymous to synonymous substitutions between human and mouse (Bellott et al., 2014). Based on this result, it could be asserted that a high degree of protein coding sequence conservation between X-Y pair genes would support functional interchangeability. Interestingly, I have observed that there is a negative corre-
lation between phenotype in hemizygous mutant males and amino acid sequence homology in X-Y pairs (Table 6.6). \textit{Kdm5c} and \textit{Kdm6a} have moderate-high degrees of similarity to their Y linked homologues, and their loss of function gives rise to a mild phenotype when the Y homologue is present. \textit{Ddx3x} and \textit{Eif2s3x} are almost identical to their Y linked homologues at the amino acid level, and their loss of function gives rise to a severe embryonic lethality phenotype when the Y homologue is present. This suggests that, whilst homology in the protein coding sequence is important to facilitate functional conservation, other factors certainly contribute to the inability of these highly similar proteins to compensate for loss of their homologue.

### 6.3.3 Differences in gene regulation likely underlie the apparent functional differences in X-Y gene pairs

Recent work has begun to highlight the influence of \textit{cis} and \textit{trans} acting regulatory mechanisms on gene expression and, therefore, phenotype. The concept of gene regulation differences accounting for a significant proportion of phenotypic variation is not new (King and Wilson, 1975), however, the development of microarray and RNA-seq technology has allowed a greater depth of study than was previously possible. A well characterised example of this is alternative splicing, which can be species, sex, organ or tissue specific (Blekhman et al., 2010; Merkin et al., 2012). A mouse model of Down syndrome carrying a single human chromosome 21 was shown to exhibit human specific exon skipping frequencies on this chromosome (Barbosa-Morais et al., 2012). Such a result is likely explained by the activity of \textit{cis} acting regulatory elements specific to the human chromosome.

Of more direct relevance to X-Y pair genes is the evolution of testis specific expression by regulatory decay (Cortez et al., 2014). Across therian mammals, many Y linked genes have undergone an almost complete reduction in expression in somatic tissues than in the testis, thus becoming testis specific. Interestingly, the genes that remained ubiquitously expressed, i.e. those from X-Y pairs, also have decreased
expression levels (Cortez et al., 2014). This regulatory decay based reduction in
Y linked gene expression could potentially explain the observations that Ddx3y, 
Eif2s3y and, to a certain extent Uty, cannot compensate for loss of X linked homologues, despite a high degree of protein coding sequence similarity. Whilst all three
genes show widespread expression, the levels may not be equivalent to the X linked
copy. In support of this hypothesis, it was observed that in XO male mice, a single transgenic copy of Eif2s3x could replace Eif2s3y in the initiation of spermatogene-
sis; but at least four transgenic copies were required for progression through meiosis (Yamauchi et al., 2016). Kdm5d remains an anomaly; it is possible that this gene
has specifically evolved increased expression for an alternative function. A qRT-
PCR based approach assaying all four X-Y gene pairs across multiple tissues would
test this hypothesis and potentially show correlation.

In order to then determine the possible impact of an X chromosome based regulatory
environment versus a Y chromosome based environment, a further experiment
would be to exchange the whole protein coding sequence of X and Y genes. Whilst transgenic approaches have been used previously, these have typically utilised ran-
domly inserted BAC constructs, often in multiple copy arrays (Mazeyrat et al.,
2001). By taking advantage of the increased efficiency shown by CRISPR assisted homologous recombination, it would be possible to replace, for example, the endogenous protein coding sequence for Eif2s3x with the protein coding sequence
for Eif2s3y (Yoshimi et al., 2016). A conceptually similar experiment has previ-
ously been used to show that regulatory sequences immediately surrounding a gene
can directly control its expression, independent of the chromosomal context. BAC
based Kdm5c transgenes were inserted into multiple differences locations on the X chromosome in single copies, and were able to escape XCI similar to the endoge-
nous gene (Li and Carrel, 2008). It would therefore be of significant interest to
observe how the Y linked gene sequence behaves within the X linked regulatory
context, and vice versa. Eventually, this would provide a focus for future work by
concisely addressing whether the survival phenotypes observed result from slight
changes in protein coding sequence, or if we should increase efforts to understand the sex chromosome gene regulatory landscape.

<table>
<thead>
<tr>
<th>Gene</th>
<th>XY male phenotype</th>
<th>XY pair amino acid similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kdm5c</td>
<td>Normal</td>
<td>76.6</td>
</tr>
<tr>
<td>Kdm6a</td>
<td>Reduced viability</td>
<td>67.7</td>
</tr>
<tr>
<td>Ddx3x</td>
<td>Embryonic lethal</td>
<td>90.2</td>
</tr>
<tr>
<td>Eif2s3x</td>
<td>Embryonic lethal</td>
<td>97.9</td>
</tr>
</tbody>
</table>

Table 6.6: Phenotype: amino acid similarity correlation for X linked mutants

Multiple studies in recent years have focused on protein coding sequence variation between species in order to infer the evolutionary path of the mammalian sex chromosomes (eg. Bellott et al., 2014; Cortez et al., 2014). We now know that Kdm5d, Ddx3y and Uty are conserved across all eutherian mammals; with Eif2s3y present in all non primate species, and possibly having been replaced by Eif1ay in primates (Cortez et al., 2014; Hughes et al., 2015). When combined with significantly increased annotation data on the X linked partners, ascribing functions such as nucleic acid binding, transcription, and translation, we can infer that these genes have been widely conserved, and maintain ubiquitous expression, because they regulate expression of target genes across the genome (Bellott et al., 2014). As I have shown during this Chapter, significant homology at the protein coding level does not, however, automatically lead to functional conservation. It is becoming increasingly clear that changes in the regulation of gene expression may be of equal or even greater importance than variation in protein coding sequence for understanding speciation, or even intra species phenotypic differences between sexes. With the ultimate goal of being able to predict how changes in sequences at regulatory elements effect changes in gene expression, tractable model systems are required. I believe that X-Y gene replacement model hypothesised here would provide the opportunity to address some of these questions both in vitro and in vivo.
Chapter 7

Summary

In this thesis I have reported experiments I performed to begin to address the role of the X chromosome in growth regulation in the mouse. This work was based around the XO mouse, a model that has previously been shown to exhibit two distinct periods of growth deficit relative to XX littermates (Thornhill and Burgoyne, 1993; Burgoyne et al., 2002).

In Chapter 3, I hypothesised that the growth deficit previously observed in X\textsuperscript{P}O embryos could result from inappropriate expression of \textit{Xist} from the single X chromosome. Using a previously published mutant \textit{Xist} allele (Marahrens et al., 1997) I found that, in the genetic absence of \textit{Xist}, X\textsuperscript{P}O embryos were still growth retarded relative to XY\textsuperscript{RIII} littermates. I concluded that \textit{Xist} is not the cause of the growth deficit, and put forward an alternative hypothesis that the phenotype could result from differential expression of imprinted genes affecting embryonic growth. As embryos with an active X\textsuperscript{M} in extraembryonic tissue do not show the growth deficit phenotype, and X\textsuperscript{P}O embryos have only an active X\textsuperscript{P} in extraembryonic tissue, I suggested that an extraembryonic defect would be more likely than an embryonic defect.

In Chapter 4, I used \textit{in vitro} stem cell models to facilitate the search for candidate imprinted genes differentially expressed between X\textsuperscript{P}O, X\textsuperscript{M}O, and XX geno-
types. Initially I derived cell lines from both extraembryonic and embryonic compartments, mTSCs and mESCs respectively, and the latter were differentiated towards mEpiLCs in order to model the epiblast. I differentiated the mTSCs to model terminally differentiated placental cells, and found no difference between the sex chromosome genotypes in propensity towards or away from a given lineage.

Using RNA-seq followed by differential gene expression analysis, I identified a small number of genes differentially expressed between genotypes within cell types for both mESCs and mEpiLCs. Most interestingly, I highlighted around 3800 genes differentially expressed between either X\textsuperscript{PO} or X\textsuperscript{MO} and XX mTSCs. These genes were physically distributed across the whole genome, suggesting global transcriptional dysregulation in the aneuploid state. There were potential confounding factors in the bioinformatic analysis that could change the current interpretation of the results, and further work is undoubtedly warranted.

In Chapter 5, I set out to investigate the genetic basis underlying the postnatal growth deficit of both X\textsuperscript{PO} and X\textsuperscript{MO} female mice. Firstly I recapitulated previous results: both X\textsuperscript{PO} and X\textsuperscript{MO} female mice had a period of reduced rate of growth in the 0-3 week period, as did female mice with a single X chromosome and a second copy of the PAR. I further confirmed that XY\textsuperscript{-} females do not show the reduced rate of growth. I then used RT-PCR and RFLP analyses to identify four candidate genes on the mouse Y chromosome that may underlie this apparent rescue: \textit{Kdm5d}, \textit{Uty}, \textit{Ddx3y}, and \textit{Eif2s3y}. I hypothesised that carrying only a single X linked copy of these genes, as per the XO female, may be detrimental to the growth of the animal. I used CRISPR-Cas genome editing to generate new mutant lines for the X linked genes, aiming to compare XX females, carrying two copies of a given target gene, with X\textsuperscript{X} females, carrying a single copy of the gene. The results for \textit{Kdm5c}, \textit{Kdm6a}, or \textit{Ddx3x} did not recapitulate the postnatal growth deficit phenotype, likely due to a combination of factors, as discussed in the Chapter. I was not able to generate a mutant line for \textit{Eif2s3x}, and I found that embryos carrying mutations in this gene do not survive beyond E6.5. This gene is therefore embryonic lethal,
and not functionally compensated for by the presence of its Y linked homologue \textit{Eif2s3y}.

In Chapter 6, I sought to address functional conservation between X-Y gene pairs on the mouse sex chromosomes, utilising the mutant models reported in Chapter 5. It has previously been shown that at least one copy of both \textit{Kdm6a} and \textit{Ddx3x} are required for postnatal survival, suggesting that the Y linked homologues of these genes have functionally diverged from their former X linked pairing partners. (Shpargel et al., 2012; Chen et al., 2016a). I confirmed that \textit{Kdm6a} is required for postnatal survival in females, and that \textit{Uty} only partially compensates for the loss of \textit{Kdm6a} in males. Furthermore, I showed that \textit{Kdm5c} is also required for postnatal survival in females, whereas in males, \textit{Kdm5d} fully compensates for loss of \textit{Kdm5c}. The data for \textit{Ddx3x} were not consistent with the previously published phenotype, and this led me to conclude that the mutant allele reported in this thesis is a hypomorph. Taken together with the results from Chapter 5 regarding \textit{Eif2s3x} and Chen and colleagues study on \textit{Ddx3x}, I concluded that \textit{Kdm6a} and \textit{Uty}, \textit{Ddx3x} and \textit{Ddx3y}, and \textit{Eif2s3x} and \textit{Eif2s3y} are functionally divergent in the mouse. In contrast, \textit{Kdm5c} and \textit{Kdm5d} appear to have conserved function, as assessed by postnatal survival. Given that both the X and Y linked genes in the X-Y gene pairs are highly similar at the amino acid level, it is likely that changes in gene regulation underlie the functional divergence reported here.

In summary, I have generated data supporting a role for the X chromosome in both embryonic and postnatal growth in the mouse, though further work is undoubtedly required to confirm the individual loci involved. I have also provided initial evidence regarding the significance of sex chromosome aneuploidy in the mouse placenta, and I am excited to continue investigations in this area. Finally, preliminary data addressing functional convergence between X-Y gene pairs suggest that three of the four Y linked genes have diverged, and are no longer able to functionally compensate for loss of the X linked homologues. These results are at odds with
recent hypotheses regarding mammalian sex chromosome evolution in the literature (eg. Bellott et al., 2014), and certainly merit further work.
Appendix A

Cell culture media and solutions

<table>
<thead>
<tr>
<th>Component</th>
<th>Concentration (mM)</th>
<th>Follicle holding medium (FHM)</th>
<th>Potassium-supplemented simplex optimised medium (KSOM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>95</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>KH$_2$PO$_4$</td>
<td>2.5</td>
<td>2.5</td>
<td>2.5</td>
</tr>
<tr>
<td>MgSO$_4$</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Lactate</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Pyruvate</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Glucose</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Glutamine</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Bovine serum albumin</td>
<td>1mg/ml</td>
<td>1mg/ml</td>
<td>1mg/ml</td>
</tr>
<tr>
<td>EDTA</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>NaHCO$_3$</td>
<td>4</td>
<td>25</td>
<td>n/a</td>
</tr>
<tr>
<td>HEPES</td>
<td>20</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>CaCl$_2$</td>
<td>1.71</td>
<td>1.71</td>
<td>1.71</td>
</tr>
</tbody>
</table>

Table A.1: Components of embryo culture media used in this thesis
<table>
<thead>
<tr>
<th>Medium</th>
<th>Component</th>
<th>Source</th>
<th>Product code</th>
<th>Volume</th>
<th>Final conc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2iLIF</td>
<td>Penicillin-streptomycin</td>
<td>Thermo Fisher Scientific</td>
<td>150700A973163</td>
<td>5ml</td>
<td>50U/ml and 50ug/ml</td>
</tr>
<tr>
<td></td>
<td>Glutamax</td>
<td>Thermo Fisher Scientific</td>
<td>35050038</td>
<td>5ml</td>
<td>2mM</td>
</tr>
<tr>
<td></td>
<td>Human recombinant serum albumin</td>
<td>Sigma-Aldrich (St Louis, USA)</td>
<td>A9731</td>
<td>500ul</td>
<td>50ug/ml</td>
</tr>
<tr>
<td></td>
<td>PD 0325901</td>
<td>Axon Medchem (Groningen, Netherlands)</td>
<td>1408</td>
<td>167ul</td>
<td>1uM</td>
</tr>
<tr>
<td></td>
<td>CHIR 99021</td>
<td>Axon Medchem (Groningen, Netherlands)</td>
<td>1386</td>
<td>167ul</td>
<td>3uM</td>
</tr>
<tr>
<td></td>
<td>Human recombinant LIF 2-mercaptoethanol</td>
<td>Merck-Millipore()</td>
<td>ESG1107</td>
<td>50ul</td>
<td>1000U/ml</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sigma-Aldrich (St Louis, USA)</td>
<td>M6250</td>
<td>3ul</td>
<td>0.1uM</td>
</tr>
<tr>
<td></td>
<td>NDiff227</td>
<td>Cellartis/Takara-Clontech (Tokyo, Japan)</td>
<td>Y40002</td>
<td>(to 500ml)</td>
<td>n/a</td>
</tr>
<tr>
<td>Trophoblast stem cell</td>
<td>Foetal bovine serum</td>
<td>Thermo Fisher Scientific</td>
<td>16000044</td>
<td>75ml</td>
<td>15% (v/v)</td>
</tr>
<tr>
<td></td>
<td>Penicillin-streptomycin</td>
<td>Thermo Fisher Scientific</td>
<td>150700A973163</td>
<td>5ml</td>
<td>50U/ml and 50ug/ml</td>
</tr>
<tr>
<td></td>
<td>Sodium pyruvate</td>
<td>Thermo Fisher Scientific</td>
<td>11360070</td>
<td>5ml</td>
<td>1mM</td>
</tr>
<tr>
<td></td>
<td>Recombinant human FGF4</td>
<td>Peprotech (London, UK)</td>
<td>100-31</td>
<td>500ul</td>
<td>25ng/ml</td>
</tr>
<tr>
<td></td>
<td>Heparin</td>
<td>Sigma-Aldrich (St Louis, USA)</td>
<td>H3149</td>
<td>50ul</td>
<td>1ug/ml</td>
</tr>
<tr>
<td></td>
<td>2-mercaptoethanol</td>
<td>Sigma-Aldrich (St Louis, USA)</td>
<td>M6250</td>
<td>3ul</td>
<td>0.1uM</td>
</tr>
<tr>
<td></td>
<td>RPMI1640 +Glutamax</td>
<td>Thermo Fisher Scientific</td>
<td>61870010</td>
<td>(to 500ml)</td>
<td>n/a</td>
</tr>
<tr>
<td>Mouse embryonic fibroblast</td>
<td>Foetal bovine serum</td>
<td>Thermo Fisher Scientific</td>
<td>16000044</td>
<td>50ml</td>
<td>10% (v/v)</td>
</tr>
<tr>
<td></td>
<td>Penicillin-streptomycin</td>
<td>Thermo Fisher Scientific</td>
<td>150700A973163</td>
<td>5ml</td>
<td>50U/ml and 50ug/ml</td>
</tr>
<tr>
<td></td>
<td>2-mercaptoethanol</td>
<td>Sigma-Aldrich (St Louis, USA)</td>
<td>M6250</td>
<td>3ul</td>
<td>0.1uM</td>
</tr>
<tr>
<td></td>
<td>DMEM +Glutamax</td>
<td>Thermo Fisher Scientific</td>
<td>61965026</td>
<td>(to 500ml)</td>
<td>n/a</td>
</tr>
</tbody>
</table>

**Table A.2:** Components of cell culture media used in this thesis
<table>
<thead>
<tr>
<th>Solution</th>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>Tris (pH8)</td>
<td>15mM</td>
</tr>
<tr>
<td></td>
<td>EDTA</td>
<td>10mM</td>
</tr>
<tr>
<td></td>
<td>RNase A</td>
<td>100ug/ml</td>
</tr>
<tr>
<td>P2</td>
<td>NaOH</td>
<td>200mM</td>
</tr>
<tr>
<td></td>
<td>Sodium dodecyl sulphate</td>
<td>1%</td>
</tr>
<tr>
<td>P3</td>
<td>KOAc (pH5.5)</td>
<td>3M</td>
</tr>
<tr>
<td>2X saline sodium salt (SSC)</td>
<td>NaCl</td>
<td>300mM</td>
</tr>
<tr>
<td></td>
<td>Na$_3$C$_6$H$_5$O$_7$</td>
<td>30M</td>
</tr>
<tr>
<td></td>
<td>HCl</td>
<td>to pH 7.0</td>
</tr>
<tr>
<td>2X hybridisation buffer</td>
<td>SSC</td>
<td>4X</td>
</tr>
<tr>
<td></td>
<td>Dextran sulphate</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>Bovine serum albumin</td>
<td>2mg/ml</td>
</tr>
<tr>
<td></td>
<td>Vanadyl ribonucleoside</td>
<td>2M</td>
</tr>
<tr>
<td>Radioimmunoprecipitation assay (RIPA) buffer</td>
<td>Tris-Cl (pH8.0)</td>
<td>10mM</td>
</tr>
<tr>
<td></td>
<td>EDTA</td>
<td>1mM</td>
</tr>
<tr>
<td></td>
<td>EGTA</td>
<td>0.5mM</td>
</tr>
<tr>
<td></td>
<td>Triton-X100</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>Sodium deoxycholate</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>Sodium dodecyl sulphate</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>140mM</td>
</tr>
<tr>
<td></td>
<td>Phenylmethane sulphonyl fluoride (PMSF)</td>
<td>1mM</td>
</tr>
<tr>
<td>Laemnlli buffer</td>
<td>Tris-Cl (pH6.8)</td>
<td>60mM</td>
</tr>
<tr>
<td></td>
<td>Sodium dodecyl sulphate</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Glycerol</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>2-mercaptoethanol</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>Bromophenol blue</td>
<td>0.01%</td>
</tr>
<tr>
<td>Western blot running buffer</td>
<td>Tris base</td>
<td>3g/L</td>
</tr>
<tr>
<td></td>
<td>Glycine</td>
<td>14.4g/L</td>
</tr>
<tr>
<td></td>
<td>Sodium dodecyl sulphate</td>
<td>1g/L</td>
</tr>
<tr>
<td>Western blot transfer buffer</td>
<td>Tris base</td>
<td>3g/L</td>
</tr>
<tr>
<td></td>
<td>Glycine</td>
<td>14.4g/L</td>
</tr>
<tr>
<td></td>
<td>Sodium dodecyl sulphate</td>
<td>50mg/L</td>
</tr>
<tr>
<td>Tris buffered saline/Tween (TBS/T)</td>
<td>Tris-Cl (pH7.5)</td>
<td>20mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>150mM</td>
</tr>
<tr>
<td></td>
<td>Tween-20</td>
<td>0.2%</td>
</tr>
<tr>
<td>KT lysis buffer</td>
<td>Tris Cl (pH9.1)</td>
<td>67mM</td>
</tr>
<tr>
<td></td>
<td>(NH$_4$)$_2$SO$_4$</td>
<td>16mM</td>
</tr>
<tr>
<td></td>
<td>MgCl$_2$</td>
<td>3.5mM</td>
</tr>
<tr>
<td></td>
<td>Bovine serum albumin</td>
<td>150ug/ml</td>
</tr>
<tr>
<td></td>
<td>Proteinase K</td>
<td>800ug/ml</td>
</tr>
<tr>
<td></td>
<td>Triton-X100</td>
<td>0.5%</td>
</tr>
</tbody>
</table>

**Table A.3:** Components of solutions used in this thesis
### Appendix B

**Oligonucleotide sequences**

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Oligo</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>InX_F</td>
<td>CTGCAATGCCTGCTGTTTTA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>InX_R</td>
<td>CCGGAGTACAAAGGGAGTCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ymt2b_F</td>
<td>CTGGAGCTCTACAGTGATGA</td>
<td>Bishop and Hata, 1987</td>
</tr>
<tr>
<td></td>
<td>Ymt2b_R</td>
<td>CAGTTACCAATCAACACATCAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Xist_mut_F</td>
<td>TCTTTTTCTTTTGCCCCATCG</td>
<td>Marahrens et al., 1997</td>
</tr>
<tr>
<td></td>
<td>Xist_mut_R</td>
<td>TACCGGTGGATGGGAATGT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Xist_RT_F</td>
<td>CGACCTATTCCTTTTGACGA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Xist_RT_R</td>
<td>GTTGAATCTCGGTCATTTA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GAPDH_RT_F</td>
<td>TGTCGTTGGATCTCTGGTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GAPDH_RT_R</td>
<td>TTCACACCCATCACAAACAT</td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Oligo</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Cdx2_F</td>
<td>TCCTGCTGACTGCTTTCTGA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cdx2_R</td>
<td>CCCTTCTGATTTGTGGAGA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Eomes_F</td>
<td>GGAAGTGACAGAGGACGGTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Eomes_R</td>
<td>GCCGTGTACAGGAATCGTAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elf5_F</td>
<td>GGACTCCGTAACCCCATAGCA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elf5_R</td>
<td>TACTGTTCGACGAGAATTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tfp2c_F</td>
<td>TGCACACAGGTATGAAAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tfp2c_R</td>
<td>CGTCAACCACACAAACTAAA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bmp4_F</td>
<td>GAGTTTCCATCAGAAGAACA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bmp4_R</td>
<td>GCTCACAATCGAAGTTTCCC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Esrrb_F</td>
<td>AACAGCCCCCTACCTGAACCT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Esrrb_R</td>
<td>CTCACTCTGTCGCCCAAGTGT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gata3_F</td>
<td>GGTTCGGATGTAAGTCGAG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gata3_R</td>
<td>CCACATGGGGTAAAGGTTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rhox4b_F</td>
<td>AGGGAAGCACTCAAGGACATG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rhox4b_R</td>
<td>TTTGTCCAATTCCACTGCTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Syna_F</td>
<td>TTGAATACACCTTCTCACGC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Syna_R</td>
<td>TGGTGCACAGAACAGGTA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Synb_F</td>
<td>CTTCACCACACCCTACGGGT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Synb_R</td>
<td>TGAACCTTGAGTGTTGAGG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sna1_F</td>
<td>CAGCTGCTTGTGCTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sna1_R</td>
<td>AGACTCTTGTCCTTGGTA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gcm1_F</td>
<td>CTGCTCAAGCTGAAAAACG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gcm1_R</td>
<td>TTGTGTTTGTGTTGTTGCG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asc12_F</td>
<td>TAAAGGCTGACACAGGAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asc12_R</td>
<td>CGTACCAGTCAAGGTTGCTT</td>
<td></td>
</tr>
<tr>
<td>Chapter</td>
<td>Oligo</td>
<td>Sequence</td>
<td>Reference</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
<td>----------------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Tpbpa_F</td>
<td>CCAGCACAGCTTTGGACATCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tpbpa_R</td>
<td>AGCATCCAAACTGCGCTCTCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pcdh12_F</td>
<td>GCTGCTTTTTCGGAACGGAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pcdh12_R</td>
<td>TTTGGGCTGGAATGCGCCCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hand1_F</td>
<td>CAAGATCAAGACTCTGCGCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hand1_R</td>
<td>TTAATCCCTCTCTCCGCCGGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pl1_F</td>
<td>GGAGCCCTACATTTGTTGTTGGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pl1_R</td>
<td>TCCCTATACACATCTGGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prl2c2_F</td>
<td>TGCTCCTGGAACCTGCTCCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prl2c2_R</td>
<td>TGAGCAAAACTGCGGCTAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prl3d3_F</td>
<td>GGCTCCGGAATGCAATTGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prl3d3_R</td>
<td>TCCCTATACACATCTGGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gapdh_qrt_F</td>
<td>TGCTCGTGAGTCTACTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gapdh_qrt_R</td>
<td>TTACACACCACTACAAACAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Srp72_F</td>
<td>GGAGTGAGGTAACCGTTATG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Srp72_R</td>
<td>AGGCAAACACTTTACAGTGCAAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ctsq_probe</td>
<td>TAQMAN MM01349948_M1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gapdh_probe</td>
<td>TAQMAN MM99999915_G1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Srp72_probe</td>
<td>TAQMAN MM01211622_G1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nanog_F</td>
<td>AGCAGATGCAAGACTCTCTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nanog_R</td>
<td>AAGTGGGTTGTACCAAGTCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pou5f1_F</td>
<td>CCCCATGCGTGAAGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pou5f1_R</td>
<td>TCAGCAGCTTGCGCAACTGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fgf5_F</td>
<td>AAAGTCAATGGCTCCCAGAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fgf5_R</td>
<td>CTTCAGGTCGGACTCCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wnt3_F</td>
<td>CAAGCACAACAATGAAACG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Oligo</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wnt3_R</td>
<td>TCGGGACTCACGGTGTTTCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dmmt3b_F</td>
<td>CTCGCAAGGTGTTGCGTTTGTAAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dmmt3b_R</td>
<td>CTGGGCATCTGTCATCTTTGCACC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sox2_probe</td>
<td>TAQMAN MM03053810_S1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Zfy1_RT_F</td>
<td>CACGCAGTAGAGAGTAGAGCACC</td>
<td></td>
</tr>
<tr>
<td>Zfy1_RT_R</td>
<td>GGTAATGTTTGAGATTCAAGGTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ube1y1_RT_F</td>
<td>CTTACGAGCTCTGGAGCCTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ube1y1_RT_R</td>
<td>CTTCTCAAGTCAATGGGGTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5d_RT_F</td>
<td>GAGACCCGTGAGGCTCTAAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5d_RT_R</td>
<td>TTAGATCTCATTCAATCACAGCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2s3y_RT_F</td>
<td>GTACAGTAGCCGGAAGAGCCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2s3y_RT_R</td>
<td>ACCACCTGGAGCACATTACT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uty_RT_F</td>
<td>TTCTGGAAGGAGACAGAAAGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uty_RT_R</td>
<td>GAACACCCCAAATAAATCTTCAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3y_RT_F</td>
<td>GTTCTGATGATCATGGTCGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3y_RT_R</td>
<td>CTTCTAGCCTCTTCAAGGCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usp9y_RT_F</td>
<td>GATGGTCAACTCTGAGCCTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usp9y_RT_R</td>
<td>CTGATTAGCTTGCAATTTGGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zfy2_RT_F</td>
<td>GATGGACTAGAGCAGAACCACAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zfy2_RT_R</td>
<td>CATAAAGCTGCTGCTCCATGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2al2y_RT_F</td>
<td>CCGTGATAGCTGTTTCTTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2al2y_RT_R</td>
<td>AGAGTGGGAGAGTCTGTTCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sry_RT_F</td>
<td>GAGGGCCATGTCAGCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sry_RT_R</td>
<td>AGGTGCCCAGTGCGGATAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rbmy1a1_RT_F</td>
<td>GCGTCCTGCCAAGAGAGATGAGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rbmy1a1_RT_R</td>
<td>GAGTGGTAAATTGCCATAGTCACAG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Oligo</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ssty1_RT_F</td>
<td>GAGCCCTGTACCCCAATGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssty1_RT_R</td>
<td>CCTCAGAGCCATTTCTCCCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssty2_RT_F</td>
<td>GCCATCATATTCTAGTGCAACTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ssty2_RT_R</td>
<td>TAAAACACCAGCAATTTGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sly_RT_F</td>
<td>GTGCCATGGACCTTCATGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sly_RT_R</td>
<td>GGTCCATGGAGGGGCTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_RFLP_F</td>
<td>AAAAGCCCGAGTTCTCCCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_RFLP_R</td>
<td>TGCTGCTGACTTCTCCCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_RFLP_F</td>
<td>AGAGGAGCTGGGTTGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_RFLP_R</td>
<td>AGCTGTAGAGGGGCTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_RFLP_F</td>
<td>AGCTGAAGGGGAAGTGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_RFLP_R</td>
<td>AGCCCCATGAAATGCATTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foxo4_RFLP_F</td>
<td>AGCTCGGAGAGTAGGAAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foxo4_RFLP_R</td>
<td>GCCCATCCAGTAGCTCTAGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2x3x_RFLP_F</td>
<td>ATCCCCGTGACACAGAAAATGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2x3x_RFLP_R</td>
<td>GTTGCGTGTGCTACTTTTGGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_g1_F</td>
<td>CACCGCCGCCCTAAAGCCCTTGACGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_g1_R</td>
<td>AAACCCCGAGGCTTTAGGGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_g2_F</td>
<td>CACCGCCGGCGCCATGCTCTTAAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_g2_R</td>
<td>AAACCTAACGCGTCATGCGCGGCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_g1_T7</td>
<td>TTAATTACGACTCTATAGGGCCCTTAGCCCCTGACGGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_g2_T7</td>
<td>TTAATTACGACTCTATAGGGCCCGGCCATGCTCTTTAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_screen_F</td>
<td>TTTCCCGATGAGCAAGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_screen_R</td>
<td>AGGCTTTGACAGTCTTTGAGGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_geno_F</td>
<td>TGTTGAGGATACAGAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_geno_R</td>
<td>CGGCCAGTGGCTGTATTTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_miseq_F</td>
<td>TCGTCGGCCAGCGTCAGATGTCTATAGAGACAGATGTGGTGAGGTACAGAGGTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chapter</td>
<td>Oligo</td>
<td>Sequence</td>
<td>Reference</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>--------------------------------------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Kdm5c_miseq_R</td>
<td>GTCTCGTGCGCGGCTCGGAGATGTGTATAAGAGACAGGAGTTTCGCGATGTAGCCAAGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_g1_F</td>
<td>CACCCTGGTCGTCCATGGAATCTCCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_g1_R</td>
<td>AAACAGAGATTTTCATGGAACGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_g2_F</td>
<td>CACCCTGACGAGCGACACTCCGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_g2_R</td>
<td>AAACGCGAGGTGTCCTCCTAGTACCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_g1_T7</td>
<td>TTAATACGACTACTATAGGGTGGTGCTTCCTGACATGGAATACCTC</td>
<td>Kdm6a_g1_T7</td>
<td></td>
</tr>
<tr>
<td>Kdm6a_g2_T7</td>
<td>TTAATACGACTCTATAGGGTGAGCGCGACACTCCGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_screen_F</td>
<td>CCTTCACCTCTACCGCGTTA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_screen_R</td>
<td>GACCCTGTTGCGAAAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_miseq_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCACTGTCACCAAGTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_g1_F</td>
<td>CACCAGCGGCTGCTGAGCCGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_g1_R</td>
<td>AAACGCCCGAGCGCATTTTACCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_g2_F</td>
<td>CACCCTGCTGAGCGACACTCCGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_g2_R</td>
<td>AAACGCGAGGTGTCCTCCTAGTACCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_g1_T7</td>
<td>TTAATACGACTCTATAGGGTGAGCGCGACACTCCGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_g2_T7</td>
<td>TTAATACGACTCTATAGGGTGAGCGCGACACTCCGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_screen_F</td>
<td>GTACGAGCTCCTCCAGACGTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_screen_R</td>
<td>CATCCAGCTCCTCCAGACG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_miseq_F</td>
<td>TCGTCGGCACTGCTAGATGTGTATAAGAGACAGGAGGAGGACAGCAGTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_miseq_R</td>
<td>GTCTCGTGCGCGGCTCGGAGATGTGTATAAGAGACAGGAGCTCCGGGCAAGGACAAATAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif23x_g1_F</td>
<td>CACCCTGCAGAGTCCTGCCCACAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif23x_g1_R</td>
<td>AAACATTGTCGAGCTCCTTGACGTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif23x_g2_F</td>
<td>CACCCTGCAGAGTCCTGCCCACAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif23x_g2_R</td>
<td>AAACATTGTCGAGCTCCTTGACGTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif23x_g3_F</td>
<td>CACCCTGCAGAGTCCTGCCCACAT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Oligo</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eif2s3x_g3_R</td>
<td>AAACCTCTGCTACGGTTCCTACAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2s3x_g1_T7</td>
<td>TTAATACGACTCACTATAGGACGATTCGTCCGACAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2s3x_g2_T7</td>
<td>TTAATACGACTCACTATAGGAAAGCCAGGGCTACCAATG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2s3x_g3_T7</td>
<td>TTAATACGACTCACTATAGGTATGTAGGAACCGTAGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2s3x_screen_F</td>
<td>TGAGTTTGAATCTACAGGGGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2s3x_screen_R</td>
<td>AGAAATTTGATCGAGGGGCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2s3x_miseq_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTGACGAGAGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eif2s3x_miseq_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTTCCGACGCAGACGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT1_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTAAACAACCCTGGGCAGTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT1_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGTCTAGCCAGGAGTGACAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT2_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGACTCTCTGCACCTGACTCAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT2_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTGAAGGTGCCACTTGAACC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT3_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCAGCTGGACACTAGCATTAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT3_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAGGGCTGGAATGATGTTTACA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT4_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCCTTCCCTCCCATCTGAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT4_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTTTCGTCTCCTCCTGAGAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT5_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGTTGGAACGCAGAGGTTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT5_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAGGAGTTTACACAGGCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT6_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTGCCTTCTGTGTTATGCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm5c_OT6_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTTTCGTCTCCTCGGAGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT1_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTCCTTCCCTCCCATCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT1_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTTTCGTCTCCTCAGACA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT2_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTTCGTGGGCGCTTCTGAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT2_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTTTCGTCTCCTCAAGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT3_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTTCGTGGGCGCTTCTTGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT3_R</td>
<td>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCTTCGTGGGCGCTTCTTGAG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Oligo</th>
<th>Sequence</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kdm6a_OT4_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTGGGACTGAGACTTTAGGCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT4_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGAAGTGGTGATGTCTGTGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT5_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAACCGAGTTCATGGTGGTGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT5_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGACAGACATCCACAAATGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT6_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAACAAGTTTCAGGGCAGCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kdm6a_OT6_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGAAGTGCTGTGAAGACGACCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT1_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGTGAGTGGTGATGTCTGTGT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT1_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGCCCGATGCTCTGAGTGAAAAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT2_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGTTGACGTGGCCAGTAAAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT2_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGACAACTCCAGTACAGCCTCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT3_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAAAAGCCCCACTCCAAACG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT3_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGGGTCTTCCTTGCTGTGTGA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT4_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCGGTAGCCTCCTCGTCAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT4_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGGTGGTGAGGTGGCTGCTGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT5_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAGGCCCACACTTCAGAATGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT5_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGCCATTTGTGCTCCTGGCTTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT6_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCCAAGTCAGACAGAAGCCA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT6_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGGTGTGGTGGTGGGTAGGTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT7_F</td>
<td>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCCACACACTTCAGAATGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ddx3x_OT7_R</td>
<td>GTCTCGTGAGCTCGGAGATGTGTATAAGAGACAGCCATTTGTGCTCCTGGCTTC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C

Antibodies and BACs

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Species</th>
<th>Supplier</th>
<th>Product code</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>KDM5C (JARID1C)</td>
<td>Rabbit mAb</td>
<td>Cell Signaling Technology (Boston, USA)</td>
<td>5361</td>
<td>1:1000</td>
</tr>
<tr>
<td>DDX3X</td>
<td>Rabbit pAb</td>
<td>Abcam (Cambridge, UK)</td>
<td>ab128206</td>
<td>1:1000</td>
</tr>
<tr>
<td>KDM6A (UTX)</td>
<td>Rabbit pAb</td>
<td>Proteintech (Manchester, UK)</td>
<td>23984</td>
<td>1:1000</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Rabbit pAb</td>
<td>Abcam (Cambridge, UK)</td>
<td>ab9485</td>
<td>1:10000</td>
</tr>
<tr>
<td>Anti rabbit, HRP linked</td>
<td>Goat</td>
<td>Cell Signaling (Boston, USA)</td>
<td>7074S</td>
<td>1:4000</td>
</tr>
<tr>
<td>Anti mouse, HRP linked</td>
<td>Goat</td>
<td>Santa Cruz Biotechnology (Dallas, USA)</td>
<td>sc-20005</td>
<td>1:4000</td>
</tr>
</tbody>
</table>

**Table C.1:** Antibodies used in this thesis

<table>
<thead>
<tr>
<th>Target</th>
<th>CHORI BAC clone</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLX (X chromosome)</td>
<td>RP23 470D15</td>
</tr>
<tr>
<td>SLY (Y chromosome)</td>
<td>RP24 502P05</td>
</tr>
<tr>
<td>PAR</td>
<td>RP24 255O24</td>
</tr>
</tbody>
</table>

**Table C.2:** BAC clones used in this thesis
Appendix D

Colophon

This document was set in the Times Roman typeface using \texttt{\LaTeX} and \texttt{Bib\LaTeX}, composed with a text editor.
Abidi, F. E., L. Holloway, C. A. Moore, D. D. Weaver, R. J. Simensen, R. E. Stevenson, R. C. Rogers, and C. E. Schwartz

Adam, M. P. and L. Hudgins

Adegbola, A., H. Gao, S. Sommer, and M. Browning


Agulnik, A. I., M. J. Mitchell, J. L. Lerner, D. R. Woods, and C. E. Bishop

Agulnik, A. I., M. J. Mitchell, M. G. Mattei, G. Borsani, P. A. Avner, J. L. Lerner, and C. E. Bishop

Aitken, R. J. and J. Graves

Akimoto, C., H. Kitagawa, T. Matsumoto, and S. Kato

Anderson, C. L. and C. J. Brown

Arcipowski, K. M., C. A. Martinez, and P. Ntziachristos

Ariumi, Y., M. Kuroki, K. Abe, H. Dansako, M. Ikeda, T. Wakita, and N. Kato

Armoskus, C., D. Moreira, K. Bollinger, O. Jimenez, S. Taniguchi, and H.-W. Tsai

P. J. Scambler, and K. E. Heath
2011. SHOX interacts with the chondrogenic transcription factors SOX5 and SOX6 to activate the aggrecan enhancer. *Human Molecular Genetics*, 20(8):1547–1559.

Bacher, C. P., M. Guggiari, B. Brors, S. Augui, P. Clerc, P. Avner, R. Eils, and E. Heard

Bachtrog, D.


Bailey, J. A., L. Carrel, A. Chakravarti, and E. E. Eichler

Balaton, B. P. and C. J. Brown

Ballabio, A., R. Carrozzo, A. Gil, B. Gillard, N. Affara, M. A. Ferguson-Smith, N. Fraser, I. Craig, M. Rocchi, and G. Romeo

Banzai, M., K. Omoe, H. Ishikawa, and A. Endo

Banzai, M., K. Omoe, H. Ishikawa, and A. Endo

Barakat, T. S., F. Loos, S. van Staveren, E. Myronova, M. Ghazvini, J. A. Groote-goed, and J. Gribnau


Barton, S. C., M. A. Surani, and M. L. Norris


Beach, R. R., C. Ricci-Tam, C. M. Brennan, C. A. Moomau, P.-h. Hsu, B. Hua, R. E. Silberman, M. Springer, and A. Amon

Beckham, C., A. Hilliker, A.-M. Cziko, A. Noueiry, M. Ramaswami, and R. Parker

Belin, V., V. Cusin, G. Viot, D. Girlich, A. Toutain, A. Moncla, M. Vekemans, M. Le Merrer, A. Munnich, and V. Cormier-Daire

Bellott, D. W. and D. C. Page  


Berletch, J. B., W. Ma, F. Yang, J. Shendure, W. S. Noble, C. M. Disteche, and X. Deng  

Birchler, J. A.  

Birtle, Z., L. Goodstadt, and C. Ponting  

Bishop, C. E. and D. Hatat  
Bishop, D. V., E. Canning, K. Elgar, E. Morris, P. A. Jacobs, and D. H. Skuse
2000. Distinctive patterns of memory function in subgroups of females with
Turner syndrome: evidence for imprinted loci on the X-chromosome affecting

Blackburn, D. G. and A. F. Flemming

Blakeley, P., N. M. E. Fogarty, I. del Valle, S. E. Wamaitha, T. X. Hu, K. Elder,
P. Snell, L. Christie, P. Robson, and K. K. Niakan
2015. Defining the three cell lineages of the human blastocyst by single-cell

Blakeslee, A. F., J. Belling, and M. E. Farnham
1920. Chromosomal duplication and Mendelian phenomena in Datura mutants.

Blank, H. M., J. M. Sheltzer, C. M. Meehl, and A. Amon
2015. Mitotic entry in the presence of DNA damage is a widespread property of


Bol, G. M., M. Xie, and V. Raman


Borsani, G., R. Tonlorenzi, M. C. Simmler, L. Dandolo, D. Arnaud, V. Capra, M. Grompe, A. Pizzuti, D. Muzny, and C. Lawrence

Bouma, G. J., Q. J. Hudson, L. L. Washburn, and E. M. Eicher


Bunt, J., N. A. Hasselt, D. A. Zwijnenburg, J. Koster, R. Versteeg, and M. Kool

Burgoyne, P. S.

Burgoyne, P. S.

Burgoyne, P. S. and A. P. Arnold

Burgoyne, P. S. and T. G. Baker

Burgoyne, P. S. and T. G. Baker

Burgoyne, P. S. and J. D. Biggers
1976. The consequences of X-dosage deficiency in the germ line: impaired de-

Burgoyne, P. S. and E. P. Evans

Burgoyne, P. S., E. P. Evans, and K. Holland


Burgoyne, P. S., O. A. Ojarikre, and J. M. A. Turner

Burgoyne, P. S., P. Tam, and E. P. Evans


Calabrese, J. M., J. Starmer, M. D. Schertzter, D. Yee, and T. Magnuson
Carney, E. W., V. Prideaux, S. J. Lye, and J. Rossant

Carrel, L., P. A. Hunt, and H. F. Willard

Carrel, L. and H. F. Willard

Carrel, L. and H. F. Willard

Carver, E. A. and L. Stubbs


Cattanach, B. M.

Cattanach, B. M.
Chang, P.-C., C.-W. Chi, G.-Y. Chau, F.-Y. Li, Y.-H. Tsai, J.-C. Wu, and Y.-H. Wu Lee


Charlesworth, B.


Chen, T.-C., H. Waldmann, and P. J. Fairchchild


Chomez, P., O. De Backer, M. Bertrand, E. De Plaen, T. Boon, and S. Lucas  


Chureau, C., S. Chantalat, A. Romito, A. Galvani, L. Duret, P. Avner, and C. Rougeulle

Chureau, C., M. Prissette, A. Bourdet, V. Barbe, L. Cattolico, L. Jones, A. Eggen, P. Avner, and L. Duret

Cleaton, M. A. M., C. A. Edwards, and A. C. Ferguson-Smith


Coan, P. M., G. J. Burton, and A. C. Ferguson-Smith

Coate, J. E. and J. J. Doyle

Cobb, J., A. Dierich, Y. Huss-Garcia, and D. Duboule
Cockwell, A., M. MacKenzie, S. Youings, and P. Jacobs


Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini, and F. Zhang

Conway, E., E. Healy, and A. P. Bracken

Cooper, D. W., J. L. VandeBerg, G. B. Sharman, and W. E. Poole

Corbel, C., P. Diabangouaya, A.-V. Gendrel, J. C. Chow, and E. Heard

Cordin, O., J. Banroques, N. K. Tanner, and P. Linder

Cortez, D., R. Marin, D. Toledo-Flores, L. Froidevaux, A. Liechti, P. D. Waters, F. Grützner, and H. Kaessmann


Cremer, T., M. Cremer, S. Dietzel, S. Müller, I. Solovei, and S. Fakan

Cross, J. C., Z. Werb, and S. J. Fisher


Dekker, J. and E. Heard  

Deng, Q., D. Ramskold, B. Reinius, and R. Sandberg  


Disteche, C. M.  
Disteche, C. M.

Disteche, C. M., G. N. Filippova, and K. D. Tsuchiya

2004. The AZFa gene DBY (DDX3Y) is widely transcribed but the protein is limited to the male germ cells by translation control. *Human Molecular Genetics*, 13(19):2333–2341.

Dixon, J. R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren

Donohoe, M. E., S. S. Silva, S. F. Pinter, N. Xu, and J. T. Lee


Dupont, C., F. Loos, J. Kong-A-San, and J. Gribnau


Eissenberg, J. C., M. G. Lee, J. Schneider, A. Ilyvarsonn, R. Shiekhattar, and A. Shilatifard
2007. The trithorax-group gene in Drosophila little imaginal discs encodes a

Elliott, D. J.

Ellis, N. and P. N. Goodfellow

Ellison, J. W., Z. Wardak, M. F. Young, P. G. Robey, M. Laig-Webster, and W. Chiong

Emerson, J. J., H. Kaessmann, E. Betrán, and M. Long

Evans, E. P. and R. J. Phillips


Ferguson-Smith, M. A.


Fukami, M., A. Seki, and T. Ogata

Fukuda, A., A. Mitani, T. Miyashita, T. Sado, A. Umezawa, and H. Akutsu

Fukuda, A., M. Tanino, R. Matoba, A. Umezawa, and H. Akutsu


Garbelli, A., S. Beermann, G. Di Cicco, U. Dietrich, and G. Maga

Gardner, R. L., V. E. Papaioannou, and S. C. Barton

Gaspar, N. J., T. G. Kinzy, B. J. Scherer, M. Hümbelin, J. W. Hershey, and W. C. Merrick
Gayen, S., E. Maclary, E. Buttigieg, M. Hinten, and S. Kalantry

Gaztelumendi, N. and C. Nogués

Gee, S. L. and J. G. Conboy

Glynn, L. M. and C. A. Sandman


Goto, Y. and N. Takagi

Goto, Y. and N. Takagi

Schwartz, and R. Weksberg


Greenberg, M. V. C., J. Glaser, M. Borsos, F. E. Marjou, M. Walter, A. Teissandier, and D. Bourc’his

Greene, M. E., G. Mundschau, J. Wechsler, M. McDevitt, A. Gamis, J. Karp, S. Gurbuxani, R. Arceci, and J. D. Crispino


1990. A gene mapping to the sex-determining region of the mouse Y chromo-
some is a member of a novel family of embryonically expressed genes. *Nature*, 346(6281):245–250.


Jansen, J. Walter, I. G. Gut, H. Marks, and H. G. Stunnenberg

Habuchi, H., N. Nagai, N. Sugaya, F. Atsumi, R. L. Stevens, and K. Kimata
2007. Mice Deficient in Heparan Sulfate 6- O-Sulfotransferase-1 Exhibit Defec-
tive Heparan Sulfate Biosynthesis, Abnormal Placentation, and Late Embryonic

Habuchi, H., M. Tanaka, O. Habuchi, K. Yoshida, H. Suzuki, K. Ban, and K. Ki-
mata
2000. The occurrence of three isoforms of heparan sulfate 6-O-sulfotransferase
having different specificities for hexuronic acid adjacent to the targeted N-

Hadjantonakis, A. K., L. L. Cox, P. P. Tam, and A. Nagy
2001. An X-linked GFP transgene reveals unexpected paternal X-chromosome
activity in trophoblastic giant cells of the mouse placenta. *genesis*, 29(3):133–
140.

Haeussler, M., K. Schönig, H. Eckert, A. Eschstruth, J. Mianné, J.-B. Renaud,
S. Schneider-Maunoury, A. Shkumatava, L. Teboul, J. Kent, J.-S. Joly, and J.-
P. Concordet
2016. Evaluation of off-target and on-target scoring algorithms and integration

Haig, D. and M. Westoby

Hamlett, W. C.
1989. Evolution and morphogenesis of the placenta in sharks. *Journal of Exper-
Hassold, T., F. Benham, and M. Leppert

Hassold, T., E. Kumlin, N. Takaesu, and M. Leppert

Hassold, T., D. Pettay, A. Robinson, and I. Uchida

Hayashi, K., H. Ohta, K. Kurimoto, S. Aramaki, and M. Saitou

Hayashi, K., H. Ohta, K. Kurimoto, S. Aramaki, and M. Saitou

Hayashi, K. and M. Saitou

Heard, E., C. Kress, F. Mongelard, B. Courtier, C. Rougeulle, A. Ashworth, C. Vousc’h, C. Babinet, and P. Avner

Heard, E., F. Mongelard, D. Arnaud, and P. Avner
Heard, E., F. Mongelard, D. Arnaud, C. Chureau, C. Vourc’h, and P. Avner
1999b. Human XIST yeast artificial chromosome transgenes show partial X in-
activation center function in mouse embryonic stem cells. *Proceedings of the
National Academy of Sciences of the United States of America*, 96(12):6841–
6846.

Hemberger, M.
2002. The role of the X chromosome in mammalian extra embryonic develop-

Hemmi, H., T. Kaisho, O. Takeuchi, S. Sato, H. Sanjo, K. Hoshino, T. Horiuchi,
H. Tomizawa, K. Takeda, and S. Akira
2002. Small anti-viral compounds activate immune cells via the TLR7

Henery, C. C., J. B. Bard, and M. H. Kaufman
1992. Tetraploidy in mice, embryonic cell number, and the grain of the develop-

Herz, H. M., L. D. Madden, Z. Chen, C. Bolduc, E. Buff, R. Gupta, R. Davuluri,
A. Shilatifard, I. K. Hariharan, and A. Bergmann
2010. The H3K27me3 Demethylase dUTX Is a Suppressor of Notch-
and Rb-Dependent Tumors in Drosophila. *Molecular and Cellular Biology*,
30(10):2485–2497.

Himeno, E., S. Tanaka, and T. Kunath
2008. Isolation and manipulation of mouse trophoblast stem cells. *Current Pro-

Hofstetter, C., J. M. Kampka, S. Huppertz, H. Weber, A. Schlosser, A. M. Mu ller,
and M. Becker
2016. Inhibition of KDM6 activity during murine ESC differentiation induces
Holm, S.


Hook, E. B. and D. Warburton

Hook, E. B. and D. Warburton


Horvath, L. M., N. Li, and L. Carrel

Hu, D. and J. C. Cross


Hughes, J. F., H. Skaletsky, N. Koutseva, T. Pyntikova, and D. C. Page


Hunt, P. A.
Huynh, K. D. and J. T. Lee

Inoue, A., L. Jiang, F. Lu, T. Suzuki, and Y. Zhang

Isensee, J., H. Witt, R. Pregla, R. Hetzer, V. Regitz-Zagrosek, and P. Ruiz Noppinger

Ishikawa, H., M. Banzai, and T. Yamauchi

Ishikawa, H., Á. Rattigan, R. Fundele, R. Fundele, P. S. Burgoyne, and P. S. Burgoyne

Issaeva, I., Y. Zonis, T. Rozovskaia, K. Orlovsky, C. M. Croce, T. Nakamura, A. Mazo, L. Eisenbach, and E. Canaani

Iurlaro, M., F. von Meyenn, and W. Reik


Jackman, S. M., X. Kong, and M. E. Fant

Jacobs, P., P. Dalton, R. James, K. Mosse, M. Power, D. Robinson, and D. Skuse

Jacobs, P. A.

Jacobs, P. A., A. G. Baikie, W. M. Brown, T. N. Macgregor, N. Maclean, and D. G. Harnden


Jacobs, P. A. and J. A. Strong
Jamieson, R. V., P. P. Tam, and M. Gardiner-Garden

Jamieson, R. V., S. S. Tan, and P. Tam


Ji, X., S. Jin, X. Qu, K. Li, H. Wang, H. He, F. Guo, and L. Dong

Jiang, L., F. Schlesinger, C. A. Davis, Y. Zhang, R. Li, M. Salit, T. R. Gingeras, and B. Oliver

Jiang, W., J. Wang, and Y. Zhang


Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier

John, R. and M. Hemberger

Johnson, M. A., I. Hernandez, Y. Wei, and N. Greenberg


Kanai, Y., N. Dohmae, and N. Hirokawa

Kapahnke, M., A. Banning, and R. Tikkanen

Kappelgaard, A.-M. and T. Laursen

Kaufman, M. H., S. C. Barton, and M. A. Surani

Kaufman, M. H. and S. Webb

Keverne, E. B.

Khil, P. P., N. A. Smirnova, P. J. Romanienko, and R. D. Camerini-Otero

Kim, D., B. Langmead, and S. L. Salzberg
Kim, I. W., A. C. Khadilkar, E. Y. Ko, and E. S. Sabanegh

King, M. C. and A. C. Wilson


Klose, R. J., E. M. Kallin, and Y. Zhang

Kobayashi, S., Y. Fujihara, N. Mise, K. Kaseda, K. Abe, F. Ishino, and M. Okabe


Kobayashi, S., Y. Totoki, M. Soma, K. Matsumoto, Y. Fujihara, A. Toyoda, Y. Sakaki, M. Okabe, and F. Ishino
Koehler, K. E., E. A. Millie, J. P. Cherry, P. S. Burgoyne, E. P. Evans, P. A. Hunt, and T. J. Hassold

Koller, P. C. and C. D. Darlington


Koopman, P., J. Gubbay, N. Vivian, P. Goodfellow, and R. Lovell-Badge

Kortschak, R. D., P. W. Tucker, and R. Saint


Kozak, M.


Lahn, B. T.

Lahn, B. T. and D. C. Page

Lahn, B. T. and D. C. Page


Lanfranco, F., A. Kamischke, M. Zitzmann, and E. Nieschlag
Larizza, D., M. Martinetti, J. M. Dugoujon, C. Tinelli, V. Calcaterra, M. Cuccia, L. Salvaneschi, and F. Severi

Latham, K. E.

Latham, K. E., B. Patel, F. D. Bautista, and S. M. Hawes

Latham, K. E. and L. Rambhatla

Lazic, S. E. and L. Essioux


Lee, J. T.
Lee, J. T., L. S. Davidow, and D. Warshawsky

Lee, J. T. and N. Lu

Lee, J. T., W. M. Strauss, J. A. Dausman, and R. Jaenisch

Lee, M. C. and G. S. Conway

Lee, M. G., J. Norman, A. Shilatifard, and R. Shiekhattar

Lee, N., J. Zhang, R. J. Klose, H. Erdjument-Bromage, P. Tempst, R. S. Jones, and Y. Zhang


Lefebvre, L., S. Viville, S. C. Barton, F. Ishino, E. B. Keverne, and M. A. Surani


Li, N. and L. Carrel 2008. Escape from X chromosome inactivation is an intrinsic property of the...


Liu, J., J. Henao-Mejia, H. Liu, Y. Zhao, and J. J. He


Livak, K. J. and T. D. Schmittgen


Lovell-Badge, R. and E. Robertson

Lyon, M. F.

Maclary, E., E. Buttigieg, M. Hinten, S. Gayen, C. Harris, M. K. Sarkar, S. Purushothaman, and S. Kalantry


1998. Mouse homologues of the human AZF candidate gene RBM are expressed in spermatogonia and spermatids, and map to a Y chromosome deletion inter-

Maier, T., M. Güell, and L. Serrano

Mak, W., J. Baxter, J. Silva, A. E. Newall, A. P. Otte, and N. Brockdorff

Mak, W., T. B. Nesterova, M. de Napoles, R. Appanah, S. Yamanaka, A. P. Otte, and N. Brockdorff

Mamiya, N. and H. J. Worman

Manes, C. and P. Menzel


Marahrens, Y., B. Panning, J. Dausman, W. Strauss, and R. Jaenisch


Matsui, J., Y. Goto, and N. Takagi

Mazeyrat, S. and M. J. Mitchell


Mazeyrat, S., N. Saut, M. G. Mattei, and M. J. Mitchell

McDonald, L. E., C. A. Paterson, and G. F. Kay
McGrath, J. and D. Solter

Meaburn, K. J. and T. Misteli

Menke, D. B. and D. C. Page

Merkin, J., C. Russell, P. Chen, and C. B. Burge

Merrick, W. C.

Miller, J. R., S. Koren, and G. Sutton

Minkovsky, A., T. S. Barakat, N. Sellami, M. H. Chin, N. Gunhanlar, J. Gribnau, and K. Plath

Minkovsky, A., S. Patel, and K. Plath


Moore, Z. Weng, and W. Xue


Müller, H. J.

Nagano, T., J. A. Mitchell, L. A. Sanz, F. M. Pauler, A. C. Ferguson-Smith, R. Feil, and P. Fraser


Nagy, A., J. Rossant, R. Nagy, W. Abramow-Newerly, and J. C. Roder

Nakagawa, O., M. Arnold, M. Nakagawa, H. Hamada, J. M. Shelton, H. Kusano, T. M. Harris, G. Childs, K. P. Campbell, J. A. Richardson, I. Nishino, and E. N. Olson


Namekawa, S. H., B. Payer, K. D. Huynh, R. Jaenisch, and J. T. Lee


Navarro, P., I. Chambers, V. Karwacki-Neisius, C. Chureau, C. Morey, C. Rougeulle, and P. Avner

Nesterova, T. B., C. E. Senner, J. Schneider, T. Alcayna-Stevens, A. Tattermusch, M. Hemberger, and N. Brockdorff

Nguyen, D. K. and C. M. Disteche


Okamoto, I., D. Arnaud, P. Le Baccon, A. P. Otte, C. M. Disteche, P. Avner, and E. Heard  

Okamoto, I., A. P. Otte, C. D. Allis, D. Reinberg, and E. Heard  


Okamoto, I., S. Tan, and N. Takagi  


Otter, M., C. T. Schrander-Stumpel, and L. M. Curfs  

Owsianka, A. M. and A. H. Patel

Pal Bhadra, M.

Palmer, S., J. Perry, D. Kipling, and A. Ashworth

Park, C., L. Carrel, and K. D. Makova

Park, S. H., S. G. Lee, Y. Kim, and K. Song


Passerini, V., E. Ozeri-Galai, M. S. de Pagter, N. Donnelly, S. Schmalbrock, W. P. Kloosterman, B. Kerem, and Z. S. aaacute
2016. The presence of extra chromosomes leads to genomic instability. *Nature Communications*, 7:1–12.

Patrat, C., I. Okamoto, P. Diabangouaya, V. Vialon, P. Le Baccon, J. Chow, and E. Heard
2009. Dynamic changes in paternal X-chromosome activity during imprinted


Pfeifer, G. P., S. D. Steigerwald, R. S. Hansen, S. M. Gartler, and A. D. Riggs  

Pfeifer, G. P., R. L. Tanguay, S. D. Steigerwald, and A. D. Riggs  

Phillips, R. J. and M. H. Kaufman  

Pimentel, H., N. L. Bray, S. Puente, P. Melsted, and L. Pachter  

Plass, C., S. M. Pfister, A. M. Lindroth, O. Bogatyrova, R. Claus, and P. Lichter  

Plath, K., J. Fang, S. K. Mlynarczyk-Evans, R. Cao, K. A. Worringer, H. Wang, C. C. de la Cruz, A. P. Otte, B. Panning, and Y. Zhang  

Poupon, V., B. Bègue, J. Gagnon, A. Dautry-Varsat, N. Cerf-Bensussan, and A. Benmerah  
1999. Molecular cloning and characterization of MT-ACT48, a novel mi-

Power, M. A. and P. P. L. Tam


Prudhomme, J., A. Dubois, P. Navarro, D. Arnaud, P. Avner, and C. Morey


Prudhomme, J. and C. Morey


Prykhozhij, S. V., S. L. Steele, B. Razaghi, and J. N. Berman


Pomeroy, and Y.-J. Cho

Radford, E. J., S. R. Ferrón, and A. C. Ferguson-Smith

Raefski, A. S. and M. J. O’Neill


Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang

Ranke, M. B. and P. Saenger

Rao, E., R. J. Blaschke, A. Marchini, B. Niesler, M. Burnett, and G. A. Rappold

Rastan, S.

Rebuzzini, P., T. Neri, G. Mazzini, M. Zuccotti, C. A. Redi, and S. Garagna

Rebuzzini, P., M. Zuccotti, C. A. Redi, and S. Garagna

Reik, W., M. Constancia, A. Fowden, N. Anderson, W. Dean, A. Ferguson-Smith, B. Tycko, and C. Sibley


Renfree, M. B., S. Suzuki, and T. Kaneko-Ishino

Richardson, B. J., A. B. Czuppon, and G. B. Sharman

Riddell, S. R., M. Murata, S. Bryant, and E. H. Warren


Roeleveld, N., G. A. Zielhuis, and F. Gabreëls


Rossant, J. and J. C. Cross

Rossant, J., J. P. Sanford, V. M. Chapman, and G. K. Andrews

Rovescalli, A. C., S. Asoh, and M. Nirenberg

Rozen, S., H. Skaletsky, J. D. Marszalek, P. J. Minx, H. S. Cordum, R. H. Waterston, R. K. Wilson, and D. C. Page

Rugarli, E. I., D. A. Adler, G. Borsani, K. Tsuchiya, B. Franco, X. Hauge, C. Distefche, V. Chapman, and A. Ballabio


Russell, L. B.

Sado, T., M. H. Fenner, S.-S. Tan, P. Tam, T. Shioda, and E. Li

Sado, T., Y. Hoki, and H. Sasaki

Sado, T., Z. Wang, H. Sasaki, and E. Li

Saldarriaga, W., H. Andrés García-Perdomo, J. Arango-Pineda, and J. Fonseca

Samal, S. K., S. Routray, G. K. Veeramachaneni, R. Dash, and M. Botlagunta

Sanger, R., P. Tippett, J. Gavin, P. Teesdale, and G. L. Daniels


Santos, C., L. Rodriguez-Revenga, I. Madrigal, C. Badenas, M. Pineda, and M. Milà


Scott, I. C., L. Anson-Cartwright, P. Riley, D. Reda, and J. C. Cross

Sekiguchi, T., H. Iida, J. Fukumura, and T. Nishimoto
2004. Human DDX3Y, the Y-encoded isoform of RNA helicase DDX3, rescues

Sekiguchi, T., Y. Kurihara, and J. Fukumura


Semina, E. V., R. S. Reiter, and J. C. Murray


Sengoku, T. and S. Yokoyama


Shao, C. and N. Takagi


Sharma, D. and J. Bhattacharya


Sharman, G. B.


Sheardown, S., D. Norris, A. Fisher, and N. Brockdorff

1996. The mouse Smcx gene exhibits developmental and tissue specific variation


Shibata, S. and J. T. Lee

Shih, J.-W., T.-Y. Tsai, C.-H. Chao, and Y.-H. Wu Lee


Shpargel, K. B., T. Sengoku, S. Yokoyama, and T. Magnuson


Simmons, D. G., A. L. Fortier, and J. C. Cross

Simmons, D. G., D. R. C. Natale, V. Begay, M. Hughes, A. Leutz, and J. C. Cross


Simpson, E.

Simpson, E., P. Chandler, E. Goulmy, C. M. Disteche, M. A. Ferguson-Smith, and D. C. Page


Solari, A. J.

Soma, M., Y. Fujihara, M. Okabe, F. Ishino, and S. Kobayashi

Soneson, C., M. I. Love, and M. D. Robinson


Stavropoulos, N., N. Lu, and J. T. Lee

Stern, C.


Sugawara, A., K. Goto, Y. Sotomaru, T. Sofuni, and T. Ito

Sugawara, O., N. Takagi, and M. Sasaki

Sun, C., C. T. Pager, G. Luo, P. Sarnow, and J. H. D. Cate

Sun, M., L. Song, Y. Li, T. Zhou, and R. S. Jope
Sun, M., T. Zhou, E. Jonasch, and R. S. Jope

Sun, S., B. C. Del Rosario, A. Szanto, Y. Ogawa, Y. Jeon, and J. T. Lee


Surani, M. A. and S. C. Barton

Surani, M. A., S. C. Barton, and M. L. Norris


Takagi, N. and G. R. Martin
Takagi, N. and M. Sasaki

Takagi, N., O. Sugawara, and M. Sasaki

Tam, P. P.

Tam, P. P. L.

Tanaka, S., T. Kunath, A. K. Hadjantonakis, A. Nagy, and J. Rossant


Tease, C. and B. M. Cattanach


Thornhill, A. R. and P. S. Burgoyne

Tian, D., S. Sun, and J. T. Lee


Tie, F., R. Banerjee, P. A. Conrad, P. C. Scacheri, and P. J. Harte

Traub, S., O. Demaria, L. Chasson, F. Serra, B. Desnues, and L. Alexopoulou

Trent, S., A. Dennehy, H. Richardson, O. A. Ojarikre, P. S. Burgoyne, T. Humby, and W. Davies

Tsai, C.-L., R. K. Rowntree, D. E. Cohen, and J. T. Lee
Tsunoda, Y., T. Tokunaga, and T. Sugie


Tunster, S. J., A. B. Jensen, and R. M. John

Turner, H. H.

Turner, J. M. A.

Tybulewicz, V. L. J.


Ullrich, O.
Umlauf, D., Y. Goto, R. Cao, F. Cerqueira, A. Wagschal, Y. Zhang, and R. Feil  
2004. Imprinting along the Kcnq1 domain on mouse chromosome 7 involves  
repressive histone methylation and recruitment of Polycomb group complexes.  


Urbach, A. and N. Benvenisty  

Vakilian, H., M. Mirzaei, M. Sharifi Tabar, P. Pooyan, L. Habibi Rezaee, L. Parker,  
P. A. Haynes, H. Gourabi, H. Baharvand, and G. H. Salekdeh  

Van der Meulen, J., V. Sanghvi, K. Mavrakis, K. Durinck, F. Fang, F. Matthijssens,  
P. Rondou, M. Rosen, T. Pieters, P. Vandenbergh, E. Delabesse, T. Lammens,  
B. De Moerloose, B. Menten, N. Van Roy, B. Verhasselt, B. Poppe, Y. Benoit,  
T. Taghon, A. M. Melnick, F. Speleman, H.-G. Wendel, and P. Van Vlierberghe  

van Haaften, G., G. L. Dalgliesh, H. Davies, L. Chen, G. Bignell, C. Greenman,  
S. Edkins, C. Hardy, S. O’Meara, J. Teague, A. Butler, J. Hinton, C. Latimer,  
M. Jia, D. Jones, C. Y. Kok, C. Leroy, M.-L. Lin, D. J. McBride, M. Maddison,  
S. Maquire, K. McLay, A. Menzies, T. Mironenko, L. Mulderrig, L. Mudie,  
E. Pleasance, R. Shepherd, R. Smith, L. Stebbings, P. Stephens, G. Tang, P. S.


Veitia, R. A. and J. A. Birchler

Vicoso, B. and B. Charlesworth

Vogel, C. and E. M. Marcotte

Vogt, M. H., E. Goulmy, F. M. Kloosterboer, E. Blokland, R. A. de Paus, R. Willemze, and J. H. Falkenburg
Vulpe, C. D., Y. M. Kuo, T. L. Murphy, L. Cowley, C. Askwith, N. Libina, J. Gitschier, and G. J. Anderson


Wang, J., S. Vasaikar, Z. Shi, M. Greer, and B. Zhang

Wang, P. J., J. R. McCarrey, F. Yang, and D. C. Page

Wang, Q., J. Wei, P. Su, and P. Gao
2015. Histone demethylase JARID1C promotes breast cancer metastasis cells


West, J. D., W. I. Frels, V. M. Chapman, and V. E. Papaioannou

Wilsker, D., A. Patsialou, P. B. Dallas, and E. Moran

Wisniewski, A., K. Milde, R. Stupnicki, and J. Szufladowicz-Wozniak


Wolf, J. B. and R. Hager

Wright, A. E., R. Dean, F. Zimmer, and J. E. Mank

Wu, J. A., B. L. Johnson, Y. Chen, C. T. Ha, and G. S. Dveksler

Wutz, A., T. P. Rasmussen, and R. Jaenisch


Xiong, Y., X. Chen, Z. Chen, X. Wang, S. Shi, X. Wang, J. Zhang, and X. He

Xu, J.

Xu, J., P. S. Burgoyne, and A. P. Arnold

Xu, J., X. Deng, and C. M. Disteche
Xu, J., X. Deng, R. Watkins, and C. M. Disteche

Xu, J., R. Watkins, and A. P. Arnold

Xu, N., M. E. Donohoe, S. S. Silva, and J. T. Lee


Yamanaka, Y., F. Lanner, and J. Rossant

Yamauchi, Y., J. M. Riel, V. Ruthig, and M. A. Ward


Yang, F., T. Babak, J. Shendure, and C. M. Disteche
Yang, H., H. Wang, C. S. Shivalila, A. W. Cheng, L. Shi, and R. Jaenisch  


Ying, Q.-L., J. Wray, J. Nichols, L. Batlle-Morera, B. Doble, J. Woodgett, P. Cohen, and A. Smith  

Yoshimi, K., Y. Kunihiro, T. Kaneko, H. Nagahora, B. Voigt, and T. Mashimo  

Yu, L.  


Zechner, U., M. Reule, A. Orth, F. Bonhomme, B. Strack, Guénet, H. Hameister, and R. Fundele  

Zhang, B., S. Kirov, and J. Snoddy  


Zhao, J., B. K. Sun, J. A. Erwin, J. J. Song, and J. T. Lee  

Zhao, L., T. Svingen, E. T. Ng, and P. Koopman  


Zhou, J., J. R. McCarrey, and P. J. Wang  

Zinn, A. R. and J. L. Ross  

Zuffardi, O., P. Maraschio, F. Lo Curto, U. Müller, A. Giarola, and L. Perotti  
Zvetkova, I., A. Apedaile, B. Ramsahoye, J. E. Mermoud, L. A. Crompton, R. John, R. Feil, and N. Brockdorff