A genetic modifier of B6.Y^{POS} gonadal sex reversal on chromosome 13

A thesis presented for the degree of Doctor of Philosophy (D. Phil)

Linacre College, University of Oxford

Catherine Livermore

Mammalian Genetics Unit, MRC Harwell Institute

Department of Biochemistry, University of Oxford

Supervisors: Dr Andy Greenfield (MGU, MRC Harwell Institute)

Prof Francis Barr (Biochemistry, Oxford)

Trinity term 2018

Word count: 41,283

A genetic modifier of B6.Y^{POS} gonadal sex reversal on chromosome 13

Catherine Livermore

Linacre College, Oxford
Mammalian Genetics Unit, MRC Harwell Institute
Department of Biochemistry, University of Oxford

A thesis presented for D. Phil, Trinity term, 2018

Abstract

Sex determination is the process by which the embryonic gonad differentiates into either a testis or an ovary. When the mammalian gonad first arises in the embryo, the cells comprising it are bipotential: they can adopt a testiscular or ovarian fate. Such commitment is governed chromosomally, principally by the presence or absence of the Y chromosome. Testis determination in the mouse is initiated by the expression of the Y-linked gene, *Sry*. The timing of *Sry* expression is critical: if there is any delay in the onset of expression, such that a threshold level is not reached before 11.5 days *post coitum* (dpc), then the outcome will be XY ovotestis or ovary development. Disrupted *Sry* expression and gonadal sex reversal can arise due to mutations in genes encoding key regulators (WT1, GATA4, MAP3K4, GADD45γ etc.) and also from a putative functional 'mismatch' between the Y chromosome and the autosomes/X chromosome. Such a situation arises following introgression of the *domesticus* Y chromosome of *Mus poschiavinus* into the C57BL/6J (B6) genetic background, resulting in so-called B6.Y^{POS} sex reversal.

While examining the molecular basis of B6.Y^{POS} sex reversal, a region of chromosome 13 in the B6.Y^{POS} colony at MRC Harwell was found to be heterozygous for sequences not derived from B6. This region modifies the B6.Y^{POS} sex reversal phenotype, promoting testis development and male fertility. This thesis describes experiments aimed at understanding this locus, attempting to identify its origins, and elucidating the mechanism by which it acts. It also reports on an investigation into the genes within the region and their possible role in sex determination.

Acknowledgments

First and foremost, I would like to express my sincere thanks to my supervisor, Dr Andy Greenfield, for your guidance and support throughout my D.Phil. I am extremely grateful to you for challenging me and for giving me faith in myself. Thank you for your help in the writing of this thesis and for improving my use of commas.

Special thanks go to Michelle Simon for your advice and expertise throughout this project. Thanks go to all the members of the Sexual Development group, both past and present. Particular thanks to Nick Warr for always being happy to discuss my ideas and problems, and to Pam Siggers for teaching me all my good lab habits. To Nicolas Chalon for your invaluable help. To Maddy Pope, Abi Harris and Dani Sagar for being wonderful friends in and out of the lab. Additional thanks to Becca Walker, Ashleigh Wilcox and Rosie Little for ensuring my 4 years at Harwell were enjoyable. Thanks to Isabelle Stévant and Serge Nef for sharing their wonderful data with me.

I am extremely grateful to all the staff in the Mary Lyon centre. Particular thanks go to Jackie Harrison and Sara Wells for sharing your wisdom and for being patient when I made mistakes, and to Alison Gallop and Gemma Atkins, whose dedication to this project as well as my mice ensured everything ran smoothly as it did. Thanks also to the core services at MRC Harwell, particularly to the Histology team, to GEMS and to the CRISPR/Cas9 team. And to Jeremy Sanderson, for inspiring my interest in microscopy and for always being a source of entertainment.

Thank you to my family, particularly my parents, for always being supportive and believing in me. My sincerest thanks go to Ben, for encouraging me through every step of my education. You are the rock that holds me steady.

Finally, thank you to the Medical Research Council for funding this DPhil.

Table of contents

Abstract ii
Acknowledgments iii
Table of contents iv
List of figures ix
List of tables xi
Abbreviations xii
1.1. The central dogma of sex development 2
1.2. The mammalian gonad: origin and formation 4
1.3. Testis development 9
1.4. *Sry*: the sex-determining region Y 10
1.5. Regulation of *Sry* 13
1.5.1. Epigenetic regulation 15
1.5.2. MAP kinase signalling 16
1.5.3. Transcription factors 19
1.5.4. Insulin/IGF signalling 20
1.6. Downstream of SRY: the testis-determining pathway 20
1.7. The ovary-determining pathway 24
1.7.1. RSPO1-WNT-β-catenin 24
1.7.2. FOXL2 26
1.8. Antagonism between the male and female pathways 28
1.9. Disorders/differences of sex development (DSDs) 34
1.9.1. Sex chromosome DSDs 34
1.9.2. 46,XX DSDs 35
1.9.3. 46,XY DSDs 36
1.9.4. Diagnosis of DSDs 36
1.9.5. The oligogenic nature of DSDs 38
1.10. Aims of this thesis 40
2.1. General solutions 43

2.2. Mouse lines and genotyping....43
2.2.1. Chromosomal sexing....44
2.2.2. B6.Y^{POS}....44
2.2.3. B6.Y^{POS}.*Mod13*....44
2.2.4. B6.Y^{AKR}....45
2.2.5. *Ptch1*^{tm1Mps}....45
2.2.6. *Dhh*^{tm1Amc}....46
2.2.7. *Adamts16*$^{em2(IMPC)H}$....46
2.3. DNA methods....47
2.3.1. DNA extraction for genotyping....47
2.3.2. Genotyping using standard PCR/gel electrophoresis....47
2.3.3. Genotyping using LunaProbes^{TM}....52
2.3.4. Sanger sequencing....54
2.3.5. *Ptch1* high-fidelity sequencing....54
2.3.6. Next generation sequencing of B6.Y^{POS}....55
2.4. Embryo analysis....57
2.4.1. Embryo collection and staging....57
2.4.2. Classification of 14.5 dpc gonads....58
2.4.3. Tissue processing....58
2.4.4. Wholemount *in situ* hybridisation (WMISH)....59
2.4.5. qRT-PCR....62
2.4.6. Wholemount kidney immunohistochemistry....64
2.5. Adult analysis....65
2.5.1. Classification of adult mice....65
2.5.2. Fertility testing....65
2.5.3. Necropsy and histology....66
3.1. Introduction....68
3.1.1. B6.Y^{DOM} sex reversal....68
3.1.2. C57BL/6 (B6) sensitivity....70
3.1.3. The *POS* gene theory....71
3.1.4. A genetic modifier of sex determination on chromosome 11....72
3.1.5. B6.Y^{POS} at MRC Harwell....73

3.1.6. Aims 74
3.2. Results 74
3.2.1. Whole genome analysis of B6.Y^{POS} 74
3.2.2. A genetic modifier on chromosome 13 protects against sex reversal in B6.Y^{POS} mice 77
3.2.3. Analysis of adult B6.Y^{POS}.*Mod13* mice 81
3.2.4. The *Mod13* region shares similarities with other mouse strains 83
3.2.5. The role of *Mod13* in male fertility 85
3.3. Discussion 88
3.3.1. Summary 88
3.3.2. The origin of *Mod13* 89
3.3.3. The persistence of *Mod13* 91
3.3.4. B6 sensitivity to gonadal sex reversal 93
4.1. Introduction 96
4.1.1. Summary so far 96
4.1.2. *Sry* underlies B6.Y^{POS} sex reversal 96
4.1.3. Hypotheses for the role of *Mod13* 99
4.1.4. Aims 100
4.2. Results 100
4.2.1. Examination of embryonic B6.XY^{POS}.*Mod13* gonads 100
4.2.2. Gene expression analyses in B6.XY^{POS}.*Mod13* gonads 104
4.2.3. *Sry* expression analysis 107
4.2.4. *Sox9* expression analysis 112
4.2.5. The effect of *Mod13* in XX gonads 113
4.3. Discussion 117
4.3.1. *Mod13* impacts primary sex determination 117
4.3.2. *Mod13* downregulates *Rspo1* and *Foxl2* 118
4.3.3. The effect of *Mod13* on ovarian development and function 121
4.3.4. *Mod13* directly upregulates *Sry,* but not *Sox9* 122
4.3.5. Concluding remarks 124
5.1. Introduction 127
5.1.1. Summary so far 127

5.1.2. Strategies 127
5.1.3. Predicted properties of a candidate gene 128
5.1.4. PTCH1/DHH signalling 130
5.1.5. Aims 135
5.2. Results 136
5.2.1. Characterising SNPs in *Mod13* 136
5.2.2. Narrowing down the region 138
5.2.3. Analysis of *Mod13(60-71)* 141
5.5.3. Candidate genes 144
5.5.3. The role of *Ptch1*/*Dhh* 148
5.3. Discussion 150
5.3.1. Summary 150
5.3.2. PTCH1/DHH signalling 150
6.1. Introduction 154
6.1.1. The ADAMTS family 154
6.1.2. ADAMTS16 expression and function 156
6.1.3. Cryptorchidism and sterility in *Adamts16* knockout rats 159
6.1.4. Descent of the testes 159
6.1.5. Human mutations in *ADAMTS16* 163
6.1.6. Aims 163
6.2. Results 164
6.2.1. Generation of a null *Adamts16* allele using CRISPR/Cas9 164
6.2.2. The role of *Adamts16* in testis development 167
6.2.3. The role of *Adamts16* in testicular descent 172
6.2.4. The role of *Adamts16* in ovary development 174
6.2.5. The role of *Adamts16* in kidney development 174
6.3. Discussion 177
6.3.1. Loss of *Adamts16* leads to cryptorchidism 177
6.3.2. Differences between rats and mice 180
6.3.3. *Adamts16* is not required for testis development 181
6.3.4. Implications for clinicians 182
6.3.5. Future work 184

7.1. Overview 186
7.2. Key conclusions 186
7.2.1. One or more sex-determining genes are located on chromosome 13 186
7.2.2. The *Mod13* locus may contribute to B6 sensitivity to XY gonadal sex reversal 187
7.2.3. *Adamts16* in the mouse is required for complete testis descent (but not testis determination) 188
7.3. Candidate genes 189
7.3.1. *Gas1* 189
7.3.2. *Zfp367* 190
7.3.3. *Med10* 190
7.4. The relationship between *Mod13* and male fertility 191
7.4.1. *Hsd17b3* 192
7.4.2. *Srd5a1* 193
7.4.3. Spermatogenesis and sperm function genes 193
7.5. The importance of genetic background in mice and humans 195
7.6. Future work 197
7.6.1. The search for candidate genes 198
7.6.2. Validation of candidate genes 199
7.6.3. Further molecular analysis of *Mod13* 200
7.6.4. Understanding the effect of *Mod13* in XX mice 202
7.7. Closing remarks 202
References 204
Appendix 1. Validation of SNPs in B6.Y^{POS} 225
Appendix 2. Alignment of *Ptch1* sequences 233
Appendix 3. Genes in *Mod13* 263

List of figures

Figure 1.1 Formation of the primordial gonad 5

Figure 1.2 Organisation of the developing testis and ovary at 14.5 dpc 8

Figure 1.3 Regulation of *Sry* 14

Figure 1.4 The genetic network controlling sex determination 33

Figure 2.1 Lightscanner DNA melting analysis 53

Figure 2.2 Tail somite staging of mouse embryos 57

Figure 3.1 Distribution of SNPs across the B6.Y^{POS} genome 76

Figure 3.2 The *Mod13* genetic modifier region 78

Figure 3.3 Examination of B6.XY^{POS}.*Mod13*(41-71) adult mice 80

Figure 3.4 Circos diagram of SNPs found in chromosome 13 for 11 different mouse strains 84

Figure 3.5 The effect of *Mod13* on male fertility 87

Figure 4.1 Examination of embryonic B6.XY^{POS}.*Mod13* gonads 102

Figure 4.2 Gene expression analyses in B6.XY^{POS}.*Mod13(41-71)* gonads 106

Figure 4.3 *Sry* expression in B6.XY^{POS}.*Mod13* gonads 109

Figure 4.4 *Sox9* expression in B6.XY^{POS}.*Mod13* gonads 110

Figure 4.5 *Sox9* expression in B6.XX.*Mod13* gonads 114

Figure 4.6 Expression of pro-ovarian genes in B6.XX.*Mod13* gonads 116

Figure 5.1 Recombinant alleles of *Mod13* 140

Figure 5.2 Comparison of *Mod13(41-71)* and *Mod13(60-71)* 142

Figure 5.3 Gene expression analyses in B6.XY^{POS}.*Mod13(60-71)* gonads 143

Figure 5.4 Expression profiles of candidate genes in *Mod13(60-71)*. 146-147

Figure 5.5 Analysis of gonads lacking members of the Hedgehog signalling pathway 149

Figure 6.1 The ADAMTS family 155

Figure 6.2 Expression of *Adamts16* in the embryonic gonad and kidneys 158

Figure 6.3 Testis descent in rodents 161

Figure 6.4 Generation of a null allele of *Adamts16* 165

Figure 6.5 No gross developmental defects in *Adamts16* null embryos 166

Figure 6.6 Gene expression analyses in *Adamts16* null XY gonads 168

Figure 6.7 Fertility testing of *Adamts16* null males 170

Figure 6.8 Examination of *Adamts16* null adult males 171

Figure 6.9 Testicular descent in *Adamts16* null mice 173

Figure 6.10 Gene expression analyses in *Adamts16* null XX gonads 175

Figure 6.11 Examination of *Adamts16* null adult females 176

Figure 6.12 Branching in *Adamts16* null kidneys 177

List of tables

Table 2.1 List of general solutions 43

Table 2.2 Gel-based genotyping PCR primers and assay conditions 49

Table 2.3 Lunaprobe™ genotyping primers and assay conditions 50

Table 2.4 SNP sequencing genotyping primers and assay conditions 51

Table 2.5 PCR assays for amplifying/sequencing *Ptch1* 56

Table 2.6 Riboprobes used for WMISH 60

Table 2.7 List of primers used for qRT-PCR 63

Table 3.1 Correlation between fertility and testis weight 86

Table 5.1 *DHH* mutations in 46,XY gonadal dysgenesis patients 132

Table 5.2 Damaging SNPs in *Mod13* 136

Table 5.3 Candidate genes on chromosome 13 between 60 Mb and 71 Mb 145

Table S1 Chromosome 13 variants confirmed by Sanger sequencing 225

Table S2 Chromosome 4 variants confirmed by Sanger sequencing 232

Table S3 Sequence variation between inbred mouse strains in *Ptch1* 249

Table S4 List of protein-coding genes in *Mod13* 263

Abbreviations

5α-RD	5α-reductases
17β-HSD3	17β hydroxysteroid dehydrogenase type 3
ADAMTS	A disintegrin and metalloproteinase with thrombospondin motifs
AMH	Anti-Müllerian hormone
B6	C57BL/6
bp	Base pairs
CAH	Congenital adrenal hyperplasia
CBX	Chromobox homolog
cDNA	Complementary DNA
CGD	Complete gonadal dysgenesis
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation sequencing
CSL	Cranial suspensory ligament
Ct	Cycle threshold
DASA	Disorder in androgen synthesis and action
DHH	Desert hedgehog
DIG	Digoxygenin
dpc	Days *post coitum*
DSD	Disorder/difference of sex development
ECM	Extracellular matrix
ENU	*N*-ethyl *N*-nitrosurea
EMX2	Empty spiracles homeobox 2
EMT	Epithelial-to-mesenchymal transition
FGF9	Fibroblast growth factor 9
FGFR2	FGF receptor 2
FOG2	Friend of GATA 2
FOXL2	Forkhead box L2
FST	Follistatin
Fwd	Forward
GAS1	Growth arrest specific 1
GATA4	GATA-binding protein 4
GD	Gonadal dysgenesis
HH	Hedgehog
HISS	Heat inactivated sheep serum
HMG	High mobility group
IGF1R	Insulin-like growth factor type I receptor
INSL3	Insulin Like 3
INSR	Insulin receptor
JMJD1A	Jumonji domain-containing 1A
LGR8	Leucine-Rich Repeat-Containing G Protein-Coupled Receptor 8
LHX9	Lim homeobox protein 9
MAPK	Mitogen activated protein kinase
MGP	Mouse Genomes Project

MMP	Matrix metalloproteinases
NR5A1	Nuclear receptor subfamily 5 group a member 1 (also known as SF1)
PBS	Phosphate buffered saline
PBT	PBS with 0.1% Tween
PcG	Polycomb group
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
PGC	Primordial germ cell
PGD	Partial gonadal dysgenesis
PGD_2	Prostaglandin D_2
PTCH1	Patched 1
QTL	Quantitative trait locus
qRT-PCR	Quantitative real-time reverse-transcriptase PCR
Rev	Reverse
RSPO1	R-spondin 1
scRNA-seq	Single-cell RNA sequencing
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SEM	Standard error of the mean
SF1	Steroidogenic factor 1 (also known as NR5A1)
SIFT	Sorting intolerant from tolerant
SIX(1/4)	Sine oculis homeobox (1/4)
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SOX9	SRY-related HMG box 9
SRY	Sex-determining region of the Y chromosome
SSC	Saline-sodium citrate buffer
STRA8	Stimulated by retinoic acid 8
TAD	Trans-activation domain
TBST	Tris-buffered saline with Tween
TDF	Testis-determining factor
TES	Testis-specific enhancer of Sox9
TESCO	TES core
ts	Tail somites
UGR	Urogenital ridge
WMISH	Wholemount in situ hybridisation
WNT	Wingless-type MMTV integration site family
wt	Wild-type
WT1	Wilm's tumour 1
ZNRF3	Zinc and ring finger 3

Chapter 1: Introduction

1.1. The central dogma of sex development

Sexual reproduction is critically important for nearly all multicellular organisms. It commonly requires the generation of two discrete types of gamete, within two discrete sexes. In mammals, sex development is a complex process, generally described as comprising three sequential steps. The first step occurs at fertilisation with the inheritance of the sex chromosomes, i.e. a zygote's 'chromosomal sex'. An XX chromosome complement typically results in female development, while an XY chromosome complement leads to male development. The second step in mammalian sex development is the commitment of the gonad primordium to either the testicular or ovarian fate, i.e. 'gonadal sex', in a process often termed primary sex determination. In the mouse, this step occurs at 11.5 days *post coitum* (dpc), whilst in humans it occurs at around 6 weeks of fetal development. Irrespective of chromosomal sex, the gonad is initially bipotential: it is capable of differentiating into two different structures. In a typical XY embryo, the transient expression of the Y-linked gene, *SRY*, provides the trigger that pushes the gonad to differentiate into a testis. In a typical XX embryo, the absence of a Y chromosome, and thus *SRY*, results in the gonad differentiating into an ovary. In the third and final step of sex development, hormones secreted by the differentiated testes guide the development of internal reproductive structures and external genitalia. This is the embryo's 'phenotypic sex'. All embryos initially develop both female (Müllerian) and male (Wolffian) ducts i.e. reproductive tracts. Anti-Müllerian hormone (AMH), produced by Sertoli cells of the testes, causes the Müllerian ducts to regress; testosterone, produced by testicular Leydig cells, promotes

Wolffian duct differentiation into epididymides, vasa deferentia and seminal vesicles. Testosterone also drives masculinisation of the external genitalia. In females, the Müllerian ducts differentiate into fallopian tubes, uterus and upper vagina, while the Wolffian ducts regress, due to the absence of testosterone (Biason-Lauber, 2010; Eicher and Washburn, 1986). Wolffian duct regression in females requires the transcription factor COUP-TFII (Zhao et al., 2017).

Differentiation of the two distinct gonads is tightly controlled by distinct networks of genes. One genetic pathway, initiated by *SRY*, promotes the development of a testis and simultaneously suppresses the genes and/or gene products that promote the ovarian pathway. The pro-ovary genetic pathway, comprising canonical WNT signals and the transcription factor FOXL2, opposes this pro-testis pathway and promotes development of an ovary. This mutual antagonism between the ovary- and testis-determining pathways is a core aspect of sex determination, even continuing after the commitment of the bipotential gonad to either fate, and into adulthood. Disruption to the network of genes required for mammalian sex determination can result in a discrepancy between the chromosomal, gonadal and phenotypic sex of an individual. In humans, such disruption leads to a heterogeneous group of conditions known as disorders/differences of sex development (DSD, see section 1.9). In mice, disruption to sex determination can result in comparable conditions, referred to as gonadal sex reversal. Because of the unique nature of gonad development, some cases of partial sex reversal involve the formation of an unusual and fascinating organ – called an ovotestis – in which both ovarian tissue and testicular tissue are present alongside one another within a single gonad. The

existence of such structures highlights the fact that sex determination is a balancing act, occurring at a molecular and cellular level, as each genetic pathway competes for dominance.

1.2. The mammalian gonad: origin and formation

In the mouse, the urogenital system derives from intermediate mesoderm, forming against the body wall, on the dorsal side of the embryo. From approximately 9.0 dpc, a pair of mounds is visible running along the anterior-posterior length of the embryo, either side of the dorsal aorta and neural tube. This urogenital system is divided into three regions: pronephros, mesonephros, and metanephros. The central region – the mesonephros – gives rise to the gonad. The gonads appear at approximately 10.0 dpc as thickenings of the coelomic epithelium on the ventromedial surface of the mesonephroi. The Müllerian (paramesonephric) ducts, which have the potential to form the upper portion of the female reproductive tract, also arise from the mesonephroi. The Wolffian (mesonephric) ducts, which have the potential to differentiate into the male reproductive tract, arise in the pronephros regions and elongate caudally through the mesonephroi (Shaw and Renfree, 2014; Swain and Lovell-Badge, 1999). The gonad is composed of somatic precursor cells derived from the mesonephros, and primordial germ cells (PGCs) that originate outside of the urogenital system. The PGCs arise at around 6.25 dpc in a region of extraembryonic mesoderm posterior to the primitive streak, before proliferating and migrating to the endoderm of the hindgut, then anteriorly to populate the gonads shortly after they form (Ginsburg et al., 1990).

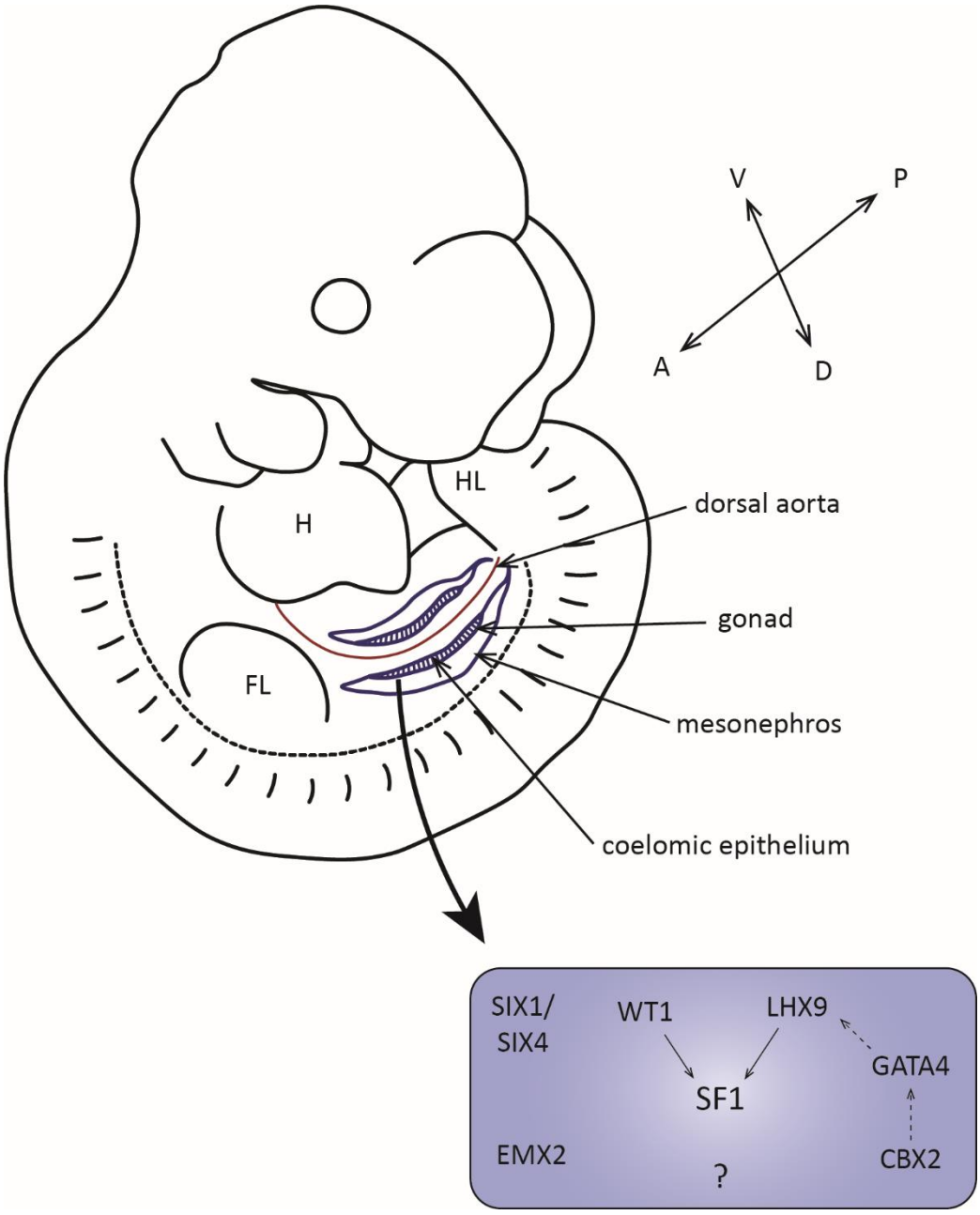


Figure 1.1 Formation of the primordial gonad. The urogenital system arises from intermediate mesoderm on the dorsal body wall, as two mounds running on either side of the dorsal aorta. The bipotential gonad forms at 10.0 dpc as thickenings of the coelomic epithelium, in a process that requires a network of genes, mainly encoding transcription factors. FL = fore-limb bud, HL = hind-limb bud, H = heart, A↔P = anteroposterior axis, D↔V = dorsoventral axis.

Several genes essential for early gonad development have been identified using mouse knockout studies, summarised in **Fig 1.1**. SF1 (steroidogenic factor 1, also known as NR5A1) is an orphan nuclear receptor implicated in the proliferation, survival and differentiation of somatic precursor cells in the gonad. SF1-positive cells in the coelomic epithelium proliferate and undergo an epithelial-to-mesenchymal transition (EMT) to provide most of the somatic cells of the gonad of both sexes (Karl and Capel, 1998; Stevant et al., 2018). Mice lacking *Sf1* completely lack gonads and adrenal glands (Hatano et al., 1996; Luo et al., 1994). *Wt1* (Wilms' tumour 1) is one of the earliest markers of gonadal somatic cells, expressed from 9.0 dpc in the intermediate mesoderm (Armstrong et al., 1993). *Wt1*-deficient mice fail to form gonads and kidneys (Kreidberg et al., 1993), and fail to express *Sf1* (Wilhelm and Englert, 2002). WT1 and LHX9 (Lim homeobox protein 9), act cooperatively to activate *Sf1* expression (Wilhelm and Englert, 2002). Loss of *Lhx9* disrupts somatic cell proliferation during gonad formation and the establishment of *Sf1* expression (Birk et al., 2000). *Gata4* (GATA-binding protein 4), which encodes another transcription factor, is required for thickening of the coelomic epithelium. Conditional deletion of *Gata4* impairs gonad formation and reduces expression of *Lhx9* and *Sf1* (Hu et al., 2013). *Cbx2* (chromobox homolog 2, also known as *M33*) also plays a role in early gonad development, possibly through upregulation of *Gata4* (and thus *Lhx2* and *Sf1*) (Katoh-Fukui et al., 2012; Katoh-Fukui et al., 1998). *Six1*/*Six4* (sine oculis homeobox 1/4) (Fujimoto et al., 2013) and *Emx2* (empty spiracles homeobox 2) (Kusaka et al., 2010) regulate EMT and the subsequent migration of somatic precursor cells in the gonad.

Immediately after its formation the gonad is bipotential: it is capable of developing into an ovary or a testis regardless of its sex chromosome complement. The gonad is a heterogeneous population of cells, composed of four cell lineages: three somatic cell lineages – the supporting cells, the steroidogenic cells and associated connective tissue – and the germ cells (illustrated in **Fig 1.2**). Each lineage is bipotential and differentiates in a sex-specific manner according to signals during or after primary sex determination. Sex determination is initiated in the supporting cell lineage. In the XY gonad, the supporting cell lineage differentiates into Sertoli cells, which orchestrate the formation of testis cords. In the XX gonad, the supporting cell lineage differentiates into granulosa cells, which later form the basis of the ovarian follicles. The role of the supporting cells (as the name suggests) is to provide the correct microenvironment to support the maturation of the germ cells in either the ovary or the testis. After differentiation of the supporting cells, the other cell lineages commit to their sex-specific fate. The steroidogenic cells produce the hormones required for development of secondary sexual characteristics; Leydig cells of the testis produce INSL3 and androgens, while theca cells of the ovary produce the androgen precursor that is then converted into oestrogen by the granulosa cells. The germ cells, surrounded by the supporting cells, later differentiate into either spermatogonia or oogonia. The connective tissue includes stromal cells, endothelial cells and peritubular myoid cells, which facilitate the organisation, structure and vascularisation of the testis or ovary following sex determination (Svingen and Koopman, 2013; Swain and Lovell-Badge, 1999). Prior to sex determination, XX and XY gonads are transcriptionally indistinguishable

(Munger et al., 2013; Nef et al., 2005). Between 11.0 dpc and 11.5 dpc in the mouse, large-scale changes in the transcriptional landscape occur, driving the sex-specific differentiation of the gonad. By 12.5 dpc the morphological differences between male and female gonads are clearly visible.

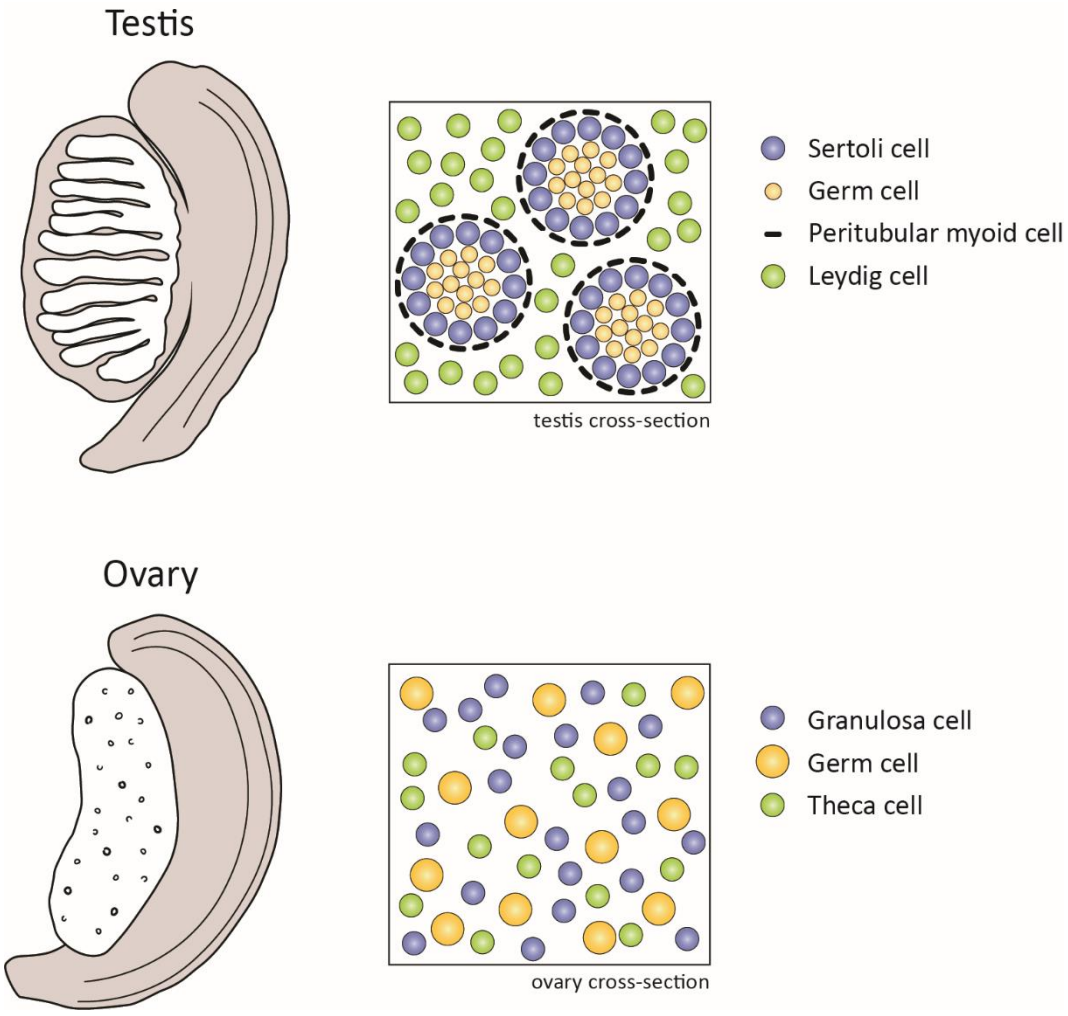


Figure 1.2 Organisation of the developing testis and ovary at 14.5 dpc. Prior to 11.5 dpc the bipotential gonad is indistinguishable between sexes, containing populations of precursor cells that differentiate into one of four cell lineages after sex determination: supporting cells (blue), steroidogenic cells (green), germ cells (yellow), and connective tissue (black). By 14.5 dpc the XY gonad has a distinct structure: Sertoli cells, encased by peritubular myoid cells, form a layer around the germ cells (shown by coronal cross-section of testis cords), which enter mitotic arrest; Leydig cells differentiate within the interstitium between cords. Within the XX gonad, granulosa and theca cell precursors differentiate and germ cells enter meiotic arrest at 13.5 dpc.

1.3. Testis development

The defining feature of the differentiated XY gonad is the testis cords: Sertoli cells form tubular structures, encasing aggregates of germ cells and segregating them from the interstitium. A layer of smooth muscle-like peritubular myoid cells surrounding the Sertoli cells produces the basement membrane required for testis cord integrity. The XY gonad recruits and patterns vasculature to form a distinctive blood vessel, known as the coelomic vessel (Coveney et al., 2008). Endothelial cells migrating from the mesonephros guide the formation of the cords (Combes et al., 2009). The testis cords will form the sperm-producing seminiferous tubules of the adult testis. Within the microenvironment of the testis cords the germ cells enter mitotic arrest, where they will remain quiescent until after birth when they will begin spermatogenesis (Kocer et al., 2009). The interstitium – the area between cords – contains Leydig cells. These steroidogenic cells produce androgens which promote Wolffian duct maturation and development of male secondary sexual characteristics. They also produce INSL3, which is required for testis descent.

The bipotential gonad differentiates in the abdomen, next to the developing kidney. During development, the ovaries maintain their position, while the testes descend into the scrotum. This process is essential for male reproductive health. Testicular descent is typically described as occurring in two phases. The first phase – the transabdominal phase – requires the action of INSL3 binding to LGR8 receptors (Gorlov et al., 2002; Klonisch et al., 2004; Nef and Parada, 1999). In mice, this process occurs between 14.5 dpc and 16.5 dpc. In humans, it occurs between week 10 and week 15 of gestation. The second phase – the

inguinoscrotal phase – is dependent on androgens (Hutson et al., 2013). It generally occurs before birth in humans, and within the first two weeks after birth in mice.

1.4. *Sry*: the sex-determining region Y

In the late 1950s and early 1960s a number of studies of sex chromosome aneuploidy in mice and humans showed that the male-determining factor is located on the Y chromosome. Individuals with an XO karyotype develop as female (Ford et al., 1959; Russell et al., 1959), while individuals with an XXY karyotype develop as males (Jacobs and Strong, 1959; Russell and Chu, 1961). The search began for the gene (or genes) on the Y chromosome responsible for maleness, termed the testis-determining factor (*TDF*). After 30 years, and a number of false starts (Palmer et al., 1989), a series of papers described the best candidate for *TDF*. The gene was named *SRY* (sex-determining region Y) and is now known to be a dominant male determinant, sufficient for testis differentiation in most mammals. Sinclair et al. (1990) identified *SRY* by searching for evolutionarily conserved sequences in the portion of translocated Y chromosome found in human XX male patients. Meanwhile, Gubbay et al. (1990) identified the mouse homolog, *Sry*, by studying a line of sex-reversed female mice with a deletion in the Y chromosome. Further reports quickly confirmed that *SRY/Sry* satisfied important criteria for a testis determining gene. *Sry* transcript was detected specifically in somatic cells of the XY genital ridge between 10.5 dpc and 12.5 dpc, the timepoint just prior to visible gonadal differentiation (Gubbay et al., 1990; Koopman et al., 1990). Supporting the

importance of *SRY* in mammalian testis determination, analyses of 46,XY female patients identified a number of mutations in *SRY*, particularly in the putative functional domain, later confirmed as the DNA-binding HMG domain (Berta et al., 1990; Jager et al., 1990; Jager et al., 1992). It is now estimated that deleterious mutations in *SRY* account for around 15% of 46,XY complete gonadal dysgenesis cases (Bashamboo and McElreavey, 2016). Finally, a 14 kb transgene encompassing mouse *Sry* was incorporated into the genome of XX embryos, leading to testis development and phenotypic, but sterile, transgenic males (Koopman et al., 1991).

SRY is a transcription factor expressed in somatic cell precursors destined to become Sertoli cells (Palmer and Burgoyne, 1991). The only functional domain conserved between species is the DNA binding domain, a 79 amino acid high mobility group (HMG) box (Kashimada and Koopman, 2010). The HMG domain binds specifically to (A/T)ACAA(T/A) sequences, causing a 60-85° bend in the DNA (Harley and Goodfellow, 1994). Most of the *SRY* mutations found in 46,XY gonadal dysgenesis patients are located in the HMG box (Assumpcao et al., 2002). This domain was originally thought to be the only functional domain. Human SRY contains no obvious transactivation domain (TAD) so it presumably acts as an architectural transcription factor, recruiting other transcriptional activator proteins to the target enhancer (Dubin and Ostrer, 1994). The N- and C-terminal domains surrounding the HMG box are poorly conserved between mammals (Zhao et al., 2014a). However, loss of the C-terminus of human SRY leads to reduced DNA binding *in vitro* (Sánchez-Moreno et al., 2008). The C-terminal region in rodents is functional (Bowles et al., 1999); a glutamine repeat

(polyQ) domain encoded by tandem CAG repeats is essential to prevent proteasomal degradation and, interestingly, has transactivation activity (Zhao et al., 2014a). Despite this, the number of glutamine repeats varies dramatically between mouse species; *Mus domesticus* strains have 8 blocks of 2 to 13 glutamine residues while *Mus musculus* strains have 21 blocks, interspersed with histidine-rich spacers (Albrecht and Eicher, 1997; Coward et al., 1994). The specific sequence of the C-terminal domain is apparently irrelevant for SRY stability; degradation can still be prevented when it is replaced by mCherry sequence (Zhao et al., 2014a).

Mouse *Sry* is expressed transiently in a specific spatiotemporal pattern. Some experiments detected *Sry* in the gonad as early as 10.75 dpc (Hacker et al., 1995). Using in situ hybridisation, expression is first detected after 11.0 dpc, beginning in the centre (and anterior pole) of the gonad before spreading in a wave-like pattern to the posterior pole over the next 4 to 8 hours. Expression levels are at their highest at approximately 11.5 dpc, before they rapidly decline in the centre, then in the anterior pole, then finally in the posterior pole (Bullejos and Koopman, 2001). Transcript is undetectable in the gonad by 12.5 dpc (Bullejos and Koopman, 2001; Hacker et al., 1995; Koopman et al., 1990). The exact timing of *Sry* expression is dependent on mouse strain (Bullejos and Koopman, 2005). This transient expression pattern has been described as a window of opportunity: failure of the *Sry* transcript to reach a particular threshold at the correct time leads to partial or complete disruption to testis determination (Bullejos and Koopman, 2005; Nagamine et al., 1999). Use of an inducible *Sry* transgene showed that delaying the expression of *Sry* by a mere 6 hours after the

normal onset of expression results in failure to activate the testis-determining pathway (Hiramatsu et al., 2009).

The primary role of SRY is to initiate the differentiation of Sertoli cells through upregulation of the related gene, *Sox9* (SRY-box 9) (Sekido et al., 2004). A particular threshold of somatic cells within the gonad must express *Sry* and commit to the Sertoli cell fate if proper testis development is to continue. Key experiments using chimaeras comprising combinations of XX and XY mouse embryos showed that only when 35-40% of somatic gonadal cells were XY would a testis form. When 10-35% of somatic cells were XY, an ovotestis would develop. When fewer than 10% of somatic cells were XY, the gonad developed as an ovary (Burgoyne et al., 1988). While *Sry* expression rapidly declines after 11.5 dpc, *Sox9* expression continues into adulthood (Kent et al., 1996). The ultimate purpose of SRY is to activate *Sox9* in enough gonadal somatic cells to promote the testis-specific programme of gene expression and enable testis development. *Sry* does not need to be expressed in all cells destined to become Sertoli cells; a paracrine signalling mechanism is able to upregulate *Sox9* in adjacent cells and recruit them to the Sertoli cell fate, even if they do not or cannot express *Sry* (Wilhelm et al., 2005). Events downstream of *Sry* will be covered in more detail later in this chapter.

1.5. Regulation of *Sry*

Timely production of SRY is crucial for its ability to upregulate *Sox9* and initiate testis determination. As such, the molecular mechanism regulating *Sry* expression is likely to be tightly controlled and involve a complex network of

genes (see **Fig 1.3**). The exact mechanism by which this network interacts with the *Sry* locus to upregulate its expression is not fully understood. However, use of loss-of-function mouse models has provided a list (but not yet a complete list) of genes essential for robust expression of *Sry*, in the gonad, at the required time.

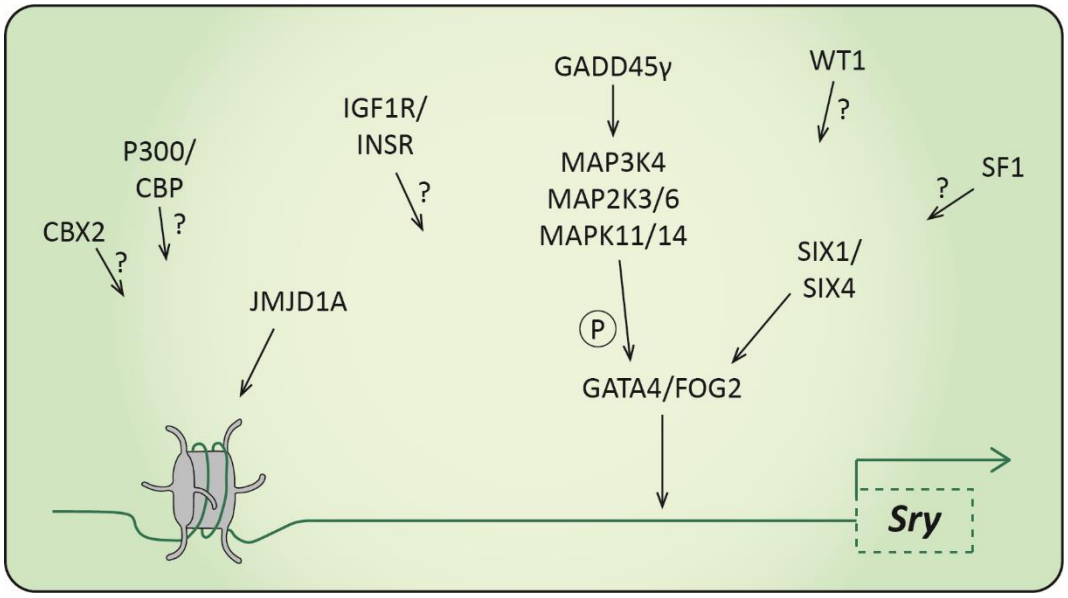


Figure 1.3 Regulation of *Sry*. Schematic of factors implicated in control of *Sry* expression. *Sry* expression must reach a threshold level within the appropriate time to successfully initiate testis determination. Question marks indicate where direct mechanism of *Sry* regulation has not been fully determined.

1.5.1. Epigenetic regulation

The polycomb group (PcG) is a family of proteins involved in chromatin remodelling, initially identified in *Drosophila* as repressors of *Hox* gene expression. As key components of the polycomb-repressive complex 1 (PRC1), CBX (chromobox homolog) proteins are thought to mediate the recognition of methylated histones through their chromodomains (Di Croce and Helin, 2013). CBX2 is a member of the PcG required for *Sry* expression. *Cbx2*-deficient mouse embryos exhibit multiple developmental defects, including hypoplastic gonads in both XX and XY embryos, and XY gonadal sex reversal (Katoh-Fukui et al., 1998). The sex-reversal phenotype was associated with near complete loss of *Sry* expression at 11.5 dpc, and could be rescued by transgenic overexpression of *Sry* or *Sox9* (Katoh-Fukui et al., 2012). Chromatin immunoprecipitation (ChIP) analysis did not detect CBX2 bound to putative regulatory sites upstream of *Sry*, possibly hindered by regions of repetitive sequence (Katoh-Fukui et al., 2012). Whether CBX2 directly regulates *Sry* expression therefore remains unclear. CBX2 has also been implicated in human sex determination; compound heterozygous mutations in *CBX2* were found in a 46,XY female with apparently normal ovaries and female anatomy (Biason-Lauber et al., 2009).

More recently, a direct link between action of a histone demethylase and *Sry* regulation was determined. Mouse embryos lacking *Jmjd1a* (Jumonji domain-containing 1A), which encodes a H3K9 demethylase, underwent complete or partial XY gonadal sex reversal, attributed to a reduction in *Sry* expression at 11.5 dpc (Kuroki et al., 2013). The authors showed that *Jmjd1a* expression increases in gonadal somatic cells just before the onset of *Sry* expression, a period of time

during which levels of H3K9me2 (a histone mark associated with transcriptional repression) were low in wild-type gonads. Gonadal cells from *Jmjd1a*-deficient embryos had an approximately 2-fold increase in levels of H3K9me2 at 11.5 dpc. ChIP analysis found JMJD1A bound to regulatory regions within the *Sry* locus in wild-type gonads at 11.5 dpc, providing strong evidence for the epigenetic control of *Sry* expression. The histone acetyl-transferases, p300 and CBP (CREB binding protein), have also been implicated in regulation of *Sry* expression through modifications to chromatin at the *Sry* promoter (Carre et al., 2018). Conditional loss of two copies of *p300* and one copy of *Cbp* resulted in XY gonadal sex reversal, associated with loss of H3K27Ac (a histone mark associated with transcriptional activation) and dramatic reduction in *Sry* expression at 11.5 dpc.

1.5.2. MAP kinase signalling

Mitogen-activated protein kinase (MAPK) signalling comprises a triple kinase cascade of phosphorylation, with each protein phosphorylating and thereby activating the next component of the pathway. A MAP3K (MAP kinase kinase kinase) phosphorylates Ser/Thr residues in a MAP2K (MAP kinase kinase), which in turn phosphorylates Thr/Tyr residues in a MAPK (MAP kinase). Once activated the MAPK phosphorylates a variety of targets in the cytoplasm or the nucleus, including those involved in transcriptional regulation. Efficient signal transduction requires co-localisation of components, and therefore often involves interaction with scaffold proteins (Yang et al., 2003).

Bogani et al. (2009) identified a loss-of-function mutation in *Map3k4* (MAP kinase kinase kinase 4, also known as MEKK4) during a phenotype-driven mouse ENU mutagenesis screen. Embryos on a C57BL/6J background had XY gonadal sex reversal and a dramatic reduction in *Sry* expression. Loss of MAP3K4 function resulted in reduced levels of phosphorylated p38 MAPK, an important family of MAPKs (Warr et al., 2012). The p38 family has four members, all of which are expressed in the gonads of both sexes between 11.5 dpc and 13.5 dpc (Ewen et al., 2010). Individual loss of *Mapk14* (*p38α*) or *Mapk11* (*p38β*) in mice did not result in a gonadal phenotype, but loss of both genes resulted in complete XY gonadal sex reversal, associated with reduced *Sry* expression (Warr et al., 2012). Further study aimed to identify the MAP2K member(s) of this signalling cascade required to regulate *Sry* expression, but could not conclusively determine which of the seven MAP2Ks play a role. Loss of two copies of *Map2k6* and one copy of *Map2k3* resulted in disrupted testis development, but not enough to cause complete XY sex reversal (Warr et al., 2016). Double homozygosity for *Map2k3* and *Map2k6* causes embryonic lethality at around 11.5 dpc, precluding examination of gonad development. However, MAP2K6, with some compensation by MAP2K3, is currently the best candidate for the activation of p38 MAPK signalling during mouse testis determination.

Three members of the growth arrest and DNA damage response protein family – GADD45α, GADD45β and GADD45γ – have been shown to interact with and activate MAP3K4 (Takekawa and Saito, 1998). Targeted loss of *Gadd45g* produced a very similar phenotype to loss of *Map3k4*; *Gadd45g*-deficient mouse embryos exhibit XY gonadal sex reversal, have reduced levels of phosphorylated

p38 MAPK, and have very low levels of *Sry* transcript and protein at 11.5 dpc, caused by a delay in *Sry* expression (Gierl et al., 2012; Warr et al., 2012). Transgenic overexpression of *Map3k4* rescued defects in testis determination caused by haploinsufficiency of *Gadd45g*, confirming a genetic interaction between *Map3k4* and *Gadd45g* during testis determination (Warr et al., 2012). Expression of *Map3k4* is fairly ubiquitous, while *Gadd45g* is expressed in somatic cells of the gonad in a specific and dynamic pattern, reminiscent of the *Sry* expression profile. In situ hybridisation first detects *Gadd45g* at 10.5 dpc, just prior to the onset of *Sry* expression, in both XX and XY gonads. Expression then spreads across the gonad and is maintained until 12.5 dpc. It is then downregulated, reaching very low levels by 13.5 dpc (Gierl et al., 2012; Warr et al., 2012). It has therefore been suggested that regulation by GADD45γ offers spatiotemporal specificity to MAP3K4, activating gonadal MAPK signalling in the time-specific pattern necessary for upregulation *Sry* expression.

These findings combine to support a model whereby a GADD45γ-MAP3K4-MAP2K3/6-MAPK11/14 signalling pathway acts to regulate the spatiotemporal expression of *Sry*, via phosphorylation of target proteins in gonadal somatic cells. One proposed target is the transcription factor, GATA4. Evidence suggests that activity of GATA4 is regulated by direct phosphorylation at serine 105, by p38 MAPKs (Charron et al., 2001; Kerkela et al., 2002). Additionally, loss of *Map3k4* or *Gadd45g* results in reduced levels of phosphorylated GATA4 in the gonad at the time of sex determination (Warr et al., 2012). There are likely other targets of phosphorylation by p38 MAPK during testis determination that have not been identified.

1.5.3. Transcription factors

As described previously, GATA4 is required for gonad formation (Hu et al., 2013). A mutation in *Gata4* that abolishes the interaction between GATA4 and its co-factor, FOG2 (Friend of GATA 2), does not appear to perturb genital ridge formation but instead causes XY gonadal sex reversal. Loss of *Fog2* also results in XY sex reversal, associated with reduced *Sry* expression at 11.5 dpc (Tevosian et al., 2002). These data indicate that GATA4 and FOG2, and the interaction between them, are critical for appropriate expression of *Sry*. ChIP analysis showed GATA4 occupying one of the GATA consensus motifs proximal to the *Sry* transcriptional start-site in gonadal cells. This interaction was lost after chemical inhibition of p38 MAPK signalling (Gierl et al., 2012). Expression of *Fog2* is thought to be regulated by two homeobox transcription factors, SIX1 and SIX4. XY mouse embryos deficient for both *Six1* and *Six4* undergo sex reversal. Gonads from these embryos have reduced expression of both *Fog2* and *Sry*, suggesting an indirect role for these proteins in *Sry* regulation (Fujimoto et al., 2013). *SF1* has been implicated in direct regulation of *SRY* expression in humans, but binding sites in mice have not yet been identified (Larney et al., 2014). Particular isoforms of WT1 have also been implicated in regulation of *Sry*, in addition to its established role in cell proliferation and formation of the early gonad. Mice lacking the +KTS isoforms (which have the additional lysine-threonine-serine (KTS) tripeptide that other isoforms lack) of WT1 had reduced levels of SRY per individual cell (Bradford et al., 2009). This combined with a reduce number of *Sry*-expressing cells results in XY sex reversal (Hammes et al., 2001).

1.5.4. Insulin/IGF signalling

Insulin/IGF signalling is critical for adrenogenital development (Nef et al., 2003; Pitetti et al., 2013). Loss of both the insulin receptor (INSR) and the insulin-like growth factor type I receptor (IGF1R) results in disruption to gonad differentiation in both XX and XY mouse embryos. XY embryos exhibit gonadal sex reversal, associated with a delay in *Sry* expression and subsequent very low levels of *Sry* at 11.5 dpc. However, the mechanism by which insulin/IGF signalling might regulate *Sry* expression remains to be determined.

1.6. Downstream of SRY: the testis-determining pathway

Sry expression is the first of a series of sex-specific genetic events that drives development of the testis. During its short window of expression, SRY upregulates *Sox9* in a wave-like pattern akin to expression of *Sry*: cells in the centre of the gonad upregulate *Sox9* first, followed by the anterior then posterior poles (Bullejos and Koopman, 2005). SOX9 is a transcription factor, containing the same HMG DNA binding motif as SRY (Wright et al., 1993). *Sox9* is the only confirmed target of SRY, although recent data from ChIP combined with whole-genome microarrays suggest that SRY binds to multiple genes (Li et al., 2014). The study indicated that SRY and SOX9 share many genetic targets. The authors proposed that SRY directly activates testis differentiation genes and represses ovary differentiation genes, before passing the role on to SOX9. Interestingly, induction of testis development is possible by duplication or overexpression of *SOX9* in chromosomally female humans (Cox et al., 2011; Huang et al., 1999) and mice (Vidal et al., 2001), despite the absence of *Sry*. These data strengthen the

theory that *Sox9* is the primary target of SRY and is in itself sufficient for testis determination.

As with *Sry*, the discovery of *Sox9* as an essential testis-determining factor involved a combination of human and mouse genetics. In humans, campomelic dysplasia, a skeletal malformation syndrome associated with gonadal dysgenesis and female development in XY individuals, was attributed to haploinsufficiency of *SOX9* (Foster et al., 1994; Wagner et al., 1994). The mouse homolog, *Sox9*, was mapped to chromosome 11 (Wright et al., 1995), then shown to be weakly expressed in both XX and XY gonads from 10.5 dpc, before being dramatically upregulated specifically in XY gonads by 11.5 dpc (Morais da Silva et al., 1996). Transcript and protein were both localised to Sertoli cells. Expression of *Sox9* was also detected in the postnatal testis (Kent et al., 1996). In contrast to humans, heterozygous loss of *Sox9* in mice did not result in XY gonadal sex reversal (Bi et al., 2001). However, homozygous loss of *Sox9* in the XY gonad caused complete sex reversal, including a failure of Sertoli cells and Leydig cells to differentiate. Mutant XY gonads were morphologically and molecularly comparable to wild-type XX gonads (Barrionuevo et al., 2006; Chaboissier et al., 2004; Lavery et al., 2011).

The first characterised enhancer of *Sox9* was the 3.2 kb TES (Testis-specific enhancer of *Sox9*), which includes a core 1.4 kb element, TESCO. SRY, working in synergy with SF1, has been shown to bind to TESCO and drive *Sox9* expression in the gonad (Sekido and Lovell-Badge, 2008). More recently CRISPR/Cas9-mediated deletion of TES was shown to cause a reduction in *Sox9* expression to

45% relative to wild-type levels in XY embryonic gonads. Deletion of TESCO led to a reduction to 60% of wild-type. However, neither deletion caused XY sex reversal, indicating that TES/TESCO are not the only gonadal enhancers of *Sox9* (Gonen et al., 2017). Further work suggested a complex landscape of enhancer elements in the 2 Mb gene desert upstream of the transcription start-site of *Sox9*. This included the presence of an enhancer element, Enh13, critical for early testis determination and thought to be the binding site for SRY initiating *Sox9* upregulation (Gonen et al., 2018).

Given that *Sox9* expression in the gonad continues long after *Sry* is undetectable, autoregulatory loop mechanisms exist to maintain *Sox9* after its initial upregulation. Once translated, SOX9 is capable of binding to TESCO using its HMG box, providing further upregulation (Sekido and Lovell-Badge, 2008). FGF9 (fibroblast growth factor 9) is essential for maintaining *Sox9* expression; *Fgf9*-deficient XY mice exhibit normal *Sry* expression and early expression of *Sox9*, but completely lack SOX9 in the gonads by 12.5 dpc, resulting in XY gonadal sex reversal (Colvin et al., 2001; Kim et al., 2006). Gonadal deletion of *Fgfr2* (fibroblast growth factor receptor 2), which encodes a receptor of FGF9, also resulted in disruption to testis differentiation (Bagheri-Fam et al., 2008; Kim et al., 2007). *Fgf9* expression was dramatically reduced in *Sox9*-deficient gonads at 11.5 dpc, prompting Kim et al. (2006) to propose that *Sox9* upregulates *Fgf9*, generating a positive feedforward loop to maintain its expression. However, more recent findings suggest that the role of FGF9 signalling is primarily to negatively regulate *Sox9*-antagonist, *Wnt4*, rather than to positively regulate *Sox9* itself (Jameson et al., 2012a) (see section 1.8 for details). *Sox9* also

upregulates *Ptgds* (prostaglandin D synthase), which encodes the enzyme that produces PGD_2 (prostaglandin D_2) to reinforce its own expression (Wilhelm et al., 2007). PGD_2 activates *Sox9* and promotes the nuclear translocation of SOX9 in the XY gonad (Moniot et al., 2009). This amplification loop acts independently of *Fgf9/Fgfr2* and provides a back-up mechanism for establishing testicular fate (Moniot et al., 2009). Paracrine signalling by PGD_2 also recruits neighbouring cells to the Sertoli cell fate, ensuring sufficient canalisation of testis determination (Wilhelm et al., 2005).

SOX9 promotes the testis-specific genetic programme required for the differentiation of Sertoli cells, other cell lineages, and the male reproductive tract. Despite its well established importance in testis determination, the direct targets of regulation by SOX9 are not well understood. There is strong evidence that SOX9 activates *Amh*, which encodes the hormone secreted by Sertoli cells responsible for the regression of the Müllerian ducts in male fetuses (Arango et al., 1999; Behringer et al., 1994). As the main driver of testis differentiation and male development (once SRY is absent), SOX9 is expected to have a plethora of target genes. Indeed, ChIP-microarray experiments using testes from 12.5 dpc mouse embryos (Li et al., 2014) and ChIP-seq experiments using testes from 13.5 dpc mouse embryos (Rahmoun et al., 2017) identified thousands of binding sites across the genome, many of which are genes with known roles in sex differentiation. Notable targets identified included *Dhh* (desert hedgehog), the gene encoding a signalling molecule secreted from Sertoli cells to specify the nearby steroidogenic cell precursors to differentiate into Leydig cells (Yao et al., 2002). Another target was *Cyp26b1*, encoding a retinoic acid-metabolising

enzyme, required to prevent germ cells prematurely entering meiosis in embryonic XY gonads, and therefore required for normal spermatogenesis (Bowles et al., 2006). However, many of the targets implicated by these studies require further validation to determine whether SOX9 up- or down-regulates their expression and what their function is within the developing testis.

1.7. The ovary-determining pathway

In the absence of *Sry* and the subsequent upregulation of *Sox9* and the testis-determining pathway, the XX gonad develops into an ovary. A number of genes have been identified that are required for normal ovarian development; but, in contrast to during testis determination, no master regulator of ovary determination has been identified. In mice, loss of either or both of the known pro-ovary pathways is not sufficient to initiate complete testis determination and Sertoli cell differentiation at 11.5 dpc in the absence of *Sry*, but perinatal/postnatal male-to-female sex reversal has been described (Auguste et al., 2011; Ottolenghi et al., 2007). At present, two genetic cascades have been shown to play distinct, but complementary, roles in ovary development and function: the RSPO1-WNT-β-catenin pathway and a pathway driven by FOXL2 (Pannetier et al., 2016).

1.7.1. RSPO1-WNT-β-catenin

Canonical WNT signalling, and the subsequent stabilisation of β-catenin, is required for normal ovarian development. *Wnt4*, the gene encoding the extracellular signalling protein Wnt Family Member 4, is expressed in the emerging mouse gonad in both XX and XY embryos, prior to primary sex

determination. From 11.5 dpc its expression is maintained in the somatic cells of the XX gonad, and downregulated in the XY gonad (Vainio et al., 1999). *Wnt4*-deficient XX embryos have partially masculinised gonads and internal reproductive tracts. When examined at 18.5 dpc, the XX gonad appeared more rounded than a wild-type ovary, with testis-like encapsulation. Ectopic expression of *3β-Hsd* and *P450c17* (genes associated with male steroidogenesis) was evident in *Wnt4*$^{-/-}$ ovaries, in addition to partial Wolffian duct development and failed Müllerian duct development (Vainio et al., 1999). The sex reversal observed in these embryos was not complete; external genitalia were not masculinised and no ectopic expression of *Sox9* was detected in *Wnt4*$^{-/-}$ ovaries before or after birth (Ottolenghi et al., 2007; Vainio et al., 1999). Conditional deletion of *Ctnnb1* (the gene encoding β-catenin) in somatic gonadal cells causes a similar partial female-to-male sex reversal phenotype, including ectopic development of a coelomic vessel in XX gonads and persistence of Wolffian ducts. However, no activation of the testis-determining pathway was observed, evidenced by a lack of *Sox9* or *Fgf9* upregulation at 11.5 dpc or 12.5 dpc (Liu et al., 2009). Constitutive activation of canonical WNT signalling, achieved by stabilisation of β-catenin in somatic cells of the gonad, is sufficient to cause male-to-female sex reversal (Maatouk et al., 2008).

R-spondin1 (RSPO1) is an activator of WNT signalling, expressed in the mouse urogenital ridge from 10.5 dpc, becoming upregulated in the XX supporting cell lineage from 11.5 dpc (Jameson et al., 2012b; Parma et al., 2006). It was first implicated in ovary determination when a recessive mutation in *RSPO1* was identified in a familial case of complete 46,XX (*SRY*-negative) testicular disorder

of sex development i.e. XX male development (Parma et al., 2006). Mouse loss-of-function models of *Rspo1* exhibit partial XX sex reversal (Chassot et al., 2008; Tomizuka et al., 2008). *Rspo1*$^{-/-}$ XX mice had feminised external genitalia, but harboured ductal structures resembling the male epididymides and vasa deferentia, as well as apparently normal vagina and uterine horns (Tomizuka et al., 2008). Gonadal sex reversal was not evident at 13.5/14.5 dpc, but examination perinatally and postnatally revealed the presence of Sertoli-like cells and seminiferous tubule formation, although some regions of the gonad maintained an ovarian identity (Chassot et al., 2008; Tomizuka et al., 2008). Loss of RSPO1 resulted in dramatic reduction in *Wnt4* and *Axin2* (a readout of canonical WNT signalling) expression (Chassot et al., 2008), while loss of β-catenin did not cause a reduction in *Rspo1* expression (Liu et al., 2009). This indicates that *Rspo1* is genetically upstream and required to establish expression of *Wnt4* in the XX gonad. Interestingly, the sex reversal phenotype and loss of canonical WNT signalling is more severe in *Rspo1*$^{-/-}$ XX embryos, compared to *Wnt4*$^{-/-}$ embryos (Chassot et al., 2008), suggesting that RSPO1 has other targets that have yet to be identified that play a role in sex development.

1.7.2. FOXL2

FOXL2 (forkhead box L2) is a member of the forkhead family of transcription factors. Mutations in *FOXL2* in humans are associated with ovarian dysgenesis and a syndromic form of premature ovarian failure (Crisponi et al., 2001). In mice, expression of *Foxl2* is first detected in the XX gonad at around 11.5 - 12.0 dpc. It is strongly upregulated in the XX gonad but not the XY gonad, peaking at 13.5 dpc, before declining after 14.5 dpc. *Foxl2* is then upregulated again in

granulosa cells of the adult ovary (Auguste et al., 2011). Examination of *Foxl2*-deficient mice revealed the importance of FOXL2 for granulosa cell differentiation; granulosa cells in these mice failed to progress from a squamous to cuboidal shape, resulting in a lack of follicle maturation and ovary atresia, similar to the ovarian failure seen in humans with *FOXL2* mutations (Schmidt et al., 2004; Uda et al., 2004). Given that *Foxl2* is reportedly not expressed until after commitment to the ovarian fate (Auguste et al., 2011; Schmidt et al., 2004), and the fact that *Foxl2* null ovaries are morphologically normal until after birth (Ottolenghi et al., 2005), it has been suggested that FOXL2 does not play a role in primary sex determination in mice. Loss of *Foxl2* does result in female-to-male sex reversal postnatally, with ectopic expression of *Sox9* and trans-differentiation of ovarian tissue to testicular tissue (Ottolenghi et al., 2005; Uhlenhaut et al., 2009). These data indicate that FOXL2 is required for maintenance of ovarian identity.

RNA profiling in *Wnt4*-deficient and *Foxl2*-deficient gonads indicates an overlap in the genes regulated by the two pathways (Ottolenghi et al., 2007). For example, both FOXL2 and WNT4 have been implicated in the upregulation *Fst* (Kashimada et al., 2011; Yao et al., 2004). *Fst* (follistatin) is expressed in the XX gonad from 11.5 dpc, reaches a peak at 12.5 dpc, and continues in the post-natal ovary (Kashimada et al., 2011; Yao et al., 2004). Early upregulation of *Fst* is thought to be dependent on WNT4, while its continued expression relies on FOXL2 (Kashimada et al., 2011). Loss of *Fst* in mice results in ectopic formation of a coelomic vessel in XX gonads and failure in meiotic germ cell survival, similar to the phenotype observed in *Wnt4*-deficient embryos (Yao et al., 2004).

The expression of *Foxl2* is thought to be somewhat dependent on RSPO1 (Auguste et al., 2011). This is supported by the observation that constitutive stabilisation of β-catenin is sufficient to upregulate *Foxl2* in XY gonads (Maatouk et al., 2008). However, analysis in *Rspo1*-deficient gonads shows that expression of *Foxl2* is not completely dependent on RSPO1 (Chassot et al., 2008). The roles of the two genetic cascades are complementary; loss of both *Wnt4* and *Foxl2* (Ottolenghi et al., 2007) or both *Rspo1* and *Foxl2* (Auguste et al., 2011) results in earlier onset and greater severity of female-to-male sex reversal, compared to loss of these genes individually. Gonads from new-born *Wnt4*$^{-/-}$;*Foxl2*$^{-/-}$ XX mice exhibited testis-like tubules, containing spermatogonia and strongly expressing *Sox9* and *Amh*. The sex reversal in these double knockout mice was still not complete, and these gonads contained both male and female differentiated germ cells (Ottolenghi et al., 2007). Testis-like tubules were also observed in *Rspo1*$^{-/-}$;*Foxl2*$^{-/-}$ XX gonads at 18.5 dpc (Auguste et al., 2011). Given that complete XX sex reversal has not been achieved in the absence of *Sry* or *Sox9* translocation, additional ovary-determining genes presumably remain undiscovered.

1.8. Antagonism between the male and female pathways

Unlike *Sry* in the testis determination pathway, no dominant female master switch has been identified. It is therefore tempting to describe ovary determination as a passive process, occurring simply because of a lack of SRY. However, an ovary is not just 'not a testis' but an organ with its own genetic identity from 11.5 dpc, whose genetic programme actively pushes back against

the testis-determining pathway. This is most clearly apparent in cases where the expression of *Sry* is delayed by only a few hours, leading to development of ovarian tissue at the gonad poles or even throughout the gonad (Bullejos and Koopman, 2005; Hiramatsu et al., 2009). A crucial role of SOX9 and FGF9 in canalising the male pathway is repression of pro-ovarian genes. In fact, most of the key sex-determining players function to promote one pathway whilst simultaneously inhibiting the other.

Antagonism between WNT signalling (in the form of RSPO1, β-catenin and WNT4) and the male-determining pathway was identified early on, with the discovery that overexpression of *WNT4* by duplication can cause disruption to testis development in 46,XY humans (Jordan et al., 2001). In mice, ectopic stabilisation of β-catenin is capable of dramatically reducing levels of SOX9 in XY gonads from 12.5 dpc, leading to male-to-female sex reversal (Maatouk et al., 2008). This effect is thought to be primarily due to opposition between pro-testis signalling by FGF9 and pro-ovary signalling by WNT4. *Fgf9* is ectopically expressed in *Wnt4*$^{-/-}$ XX gonads (Kim et al., 2006), while *Fgf9*$^{-/-}$ XY gonads exhibit de-repression of ovarian genes, including *Wnt4*, *Rspo1* and *Foxl2*, comparable to wild-type XX gonads, as early as 11.5 dpc (Jameson et al., 2012a). Additionally, exogenous FGF9 is capable of repressing *Wnt4* in cultured XX gonadal cells (Kim et al., 2006). It has been proposed that the primary role of FGF9 when maintaining *Sox9* expression is the suppression of *Wnt4*. Jameson et al. (2012a) examined gonads in *Fgf9*$^{-/-}$;*Wnt4*$^{-/-}$ and *Fgfr2*$^{-/-}$;*Wnt4*$^{-/-}$ embryos. While embryos lacking *Fgf9* or *Fgfr2* alone have been shown to undergo male-to-female sex reversal (Bagheri-Fam et al., 2008; Colvin et al., 2001; Kim et al., 2007; Siggers et

al., 2014), deletion of *Wnt4* completely rescued the single-mutant phenotypes, with normal development of XY gonads into functional testes. Jameson et al. (2012a) also reported that deletion of *Fgf9* did *not* rescue the female-to-male sex reversal of *Wnt4*$^{-/-}$ XX gonads. Together these data suggest that FGF9 functions solely as a suppressor of WNT signalling, and it's maintenance of *Sox9* is dependent on this suppression, not activation of *Sox9*. In contrast, WNT4 has multiple roles in promoting ovarian development, not limited to suppression of FGF signalling.

Antagonism between SOX9 and WNT signalling components has also been described. There is convincing evidence that β-catenin inhibits SF1 binding to the TESCO enhancer of *Sox9*, preventing its transcription (Bernard et al., 2012). Additionally, studies of *Sox9*;*Foxl2* double knockouts suggest that male-to-female sex reversal in *Sox9*$^{-/-}$ XY gonads is a result of failure to repress *Rspo1*, based on the observation that XY double knockout mice develop hypoplastic testes (Lavery et al., 2012). *Sox9*-deficient embryos exhibit ectopic expression of *Wnt4* (Barrionuevo et al., 2006) and SOX9 has been shown to directly inhibit canonical WNT signalling during chondrocyte differentiation (Topol et al., 2009). ZNRF3 has been proposed to mediate the antagonistic relationship between WNT signalling and the male-determining pathway. ZNRF3 is a known antagonist of WNT signalling and a target of inhibition by RSPO1. Loss of *Znrf3* causes XY gonadal sex reversal, associated with ectopic canonical WNT signalling (Harris et al., 2018).

Despite the critical role of *Foxl2* in XX gonads not being evident until after birth, *Foxl2* is expressed as early as 11.5 dpc in XX gonads (Auguste et al., 2011;

Jameson et al., 2012b). Evidence suggests that FOXL2 has a role to play in antagonising the male-determining pathway, and vice versa. *Foxl2* expression is elevated in *Sox9*-deficient XY gonads (Barrionuevo et al., 2006), although no evidence of direct regulation of *Foxl2* by SOX9 has been reported. Conversely, ChIP data indicates that FOXL2 is capable of binding to the TESCO enhancer, to directly negatively regulate *Sox9* (Uhlenhaut et al., 2009). Examination of double *Fgfr2c*;*Foxl2* knockout mice show that loss of *Foxl2* can rescue the XY sex reversal caused by loss of FGF signalling, although the gonadal phenotype observed at 15.5 dpc was extremely variable (Bagheri-Fam et al., 2017). This suggests that FGF9/FGFR2 signalling is required to repress both *Wnt4* and *Foxl2* for the maintenance of *Sox9* expression during testis determination and indicates a role for *Foxl2* in feminising the XY mutant fetal gonad.

All together, these data paint a complex picture of the interaction between the two sex-determining pathways, as illustrated in **Fig 1.4**. Given that it appears that most key proteins in the sex network have a part to play in directly or indirectly repressing the alternative pathway, while also promoting the pathway of which they are member, when describing a 'pro-ovary' gene there is almost an inherent assumption that the gene is also 'anti-testis', and vice versa. This mutual antagonism is a key aspect of gonad development, a mechanism likely developed to ensure correct development of only testicular or ovarian tissue in the initially bipotential gonad i.e. canalisation. Interestingly, this mutual antagonism continues into adulthood. Conditional ablation of *Foxl2* in 8 week old female mice caused loss of ovarian identity and reprogramming of ovarian granulosa and theca cells into Sertoli-like and Leydig-like cells. Within 3 weeks almost all

follicles had trans-differentiated into seminiferous tubules (Uhlenhaut et al., 2009). The reverse occurred when *Dmrt1* was deleted in male mice; *Dmrt1*-deficient XY embryos develop normal testes (Raymond et al., 2000), but from 2 weeks of age expression of SOX9 was lost in Sertoli cells, to be gradually replaced by FOXL2. Conditional deletion of *Dmrt1* in Sertoli cells of adult males caused trans-differentiation to granulosa-like cells, appearance of theca-like cells and large changes in gonadal gene expression (Matson et al., 2011). Conditional deletion of both *Sox9* and *Sox8* in Sertoli cells of adult testes also resulted in trans-differentiation to granulosa-like cells, genetic reprogramming and loss of seminiferous tubule integrity. This was associated with a loss of *Dmrt1* expression so it has been proposed that *Sox9/8* act to maintain *Dmrt1* expression in the adult Sertoli cells, which in turn maintains the testicular fate (Barrionuevo et al., 2016). These data indicate that gonads remain plastic into adulthood; ongoing mutual antagonism between the pro-male and pro-female pathways means that SOX9/8/DMRT1 and FOXL2 are required for maintaining their respective sex identity postnatally.

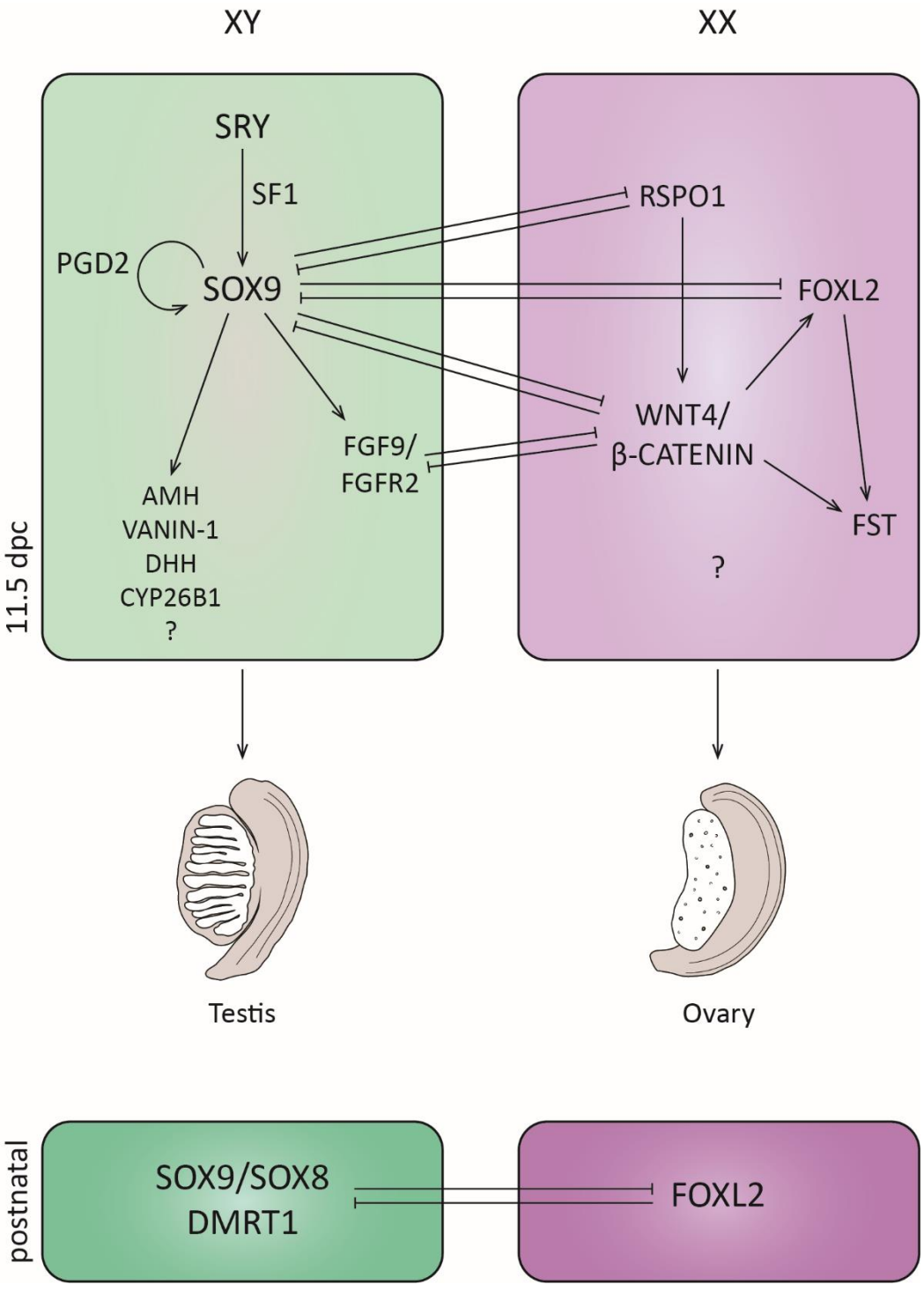


Figure 1.4 The genetic network controlling sex determination. In the XY gonad the transient presence of SRY, which peaks at 11.5 dpc, initiates testis determination by upregulation of *Sox9* and the male pathway. In the absence of SRY, the XX gonad undergoes ovarian differentiation, controlled by the canonical WNT signalling pathway and FOXL2. Mutual antagonism between the male and female pathways drives the gonad down one fate or the other. This mutual antagonism continues into adulthood.

1.9. Disorders/differences of sex development (DSDs)

This introduction has detailed a number of genes required for normal differentiation of the ovaries and testes, as well as secondary sex determination. Mutations or other alterations in these genes can lead to disorders of sex development (DSDs). DSDs are very heterogeneous conditions, defined by the 2006 Consensus Statement on Management of Intersex Disorders as any condition where the chromosomal, gonadal or anatomical sex of an individual is atypical (Lee et al., 2006). The use of the term *disorders* of sex development has been questioned in recent years; a consensus among patients thought the phrase carried unnecessary negative connotations and feelings of stigma (Johnson et al., 2017). It has therefore been suggested that *differences* of sex development should be used in its place, especially when a DSD does not continually impact patient health. DSDs are generally classified into one of three groups, according to the karyotype of the patient: sex chromosome DSDs, 46,XY DSDs and 46,XX DSDs (Lee et al., 2006).

1.9.1. Sex chromosome DSDs

Chromosomal DSDs are conditions caused by sex chromosome aneuploidy; this includes Turner syndrome, Klinefelter syndrome, 45,X/46,XY mosaicism and 46,XX/46,XY chimerism (Lee et al., 2006). Turner syndrome is a result of a 45,X (45,XO) karyotype, occurring in 1 in 2,500 live female births. Patients typically have ovarian failure and do not spontaneously undergo puberty, and 95-98% are infertile. The syndrome is also associated with renal and cardiac anomalies, short stature, distinctive physical features and chronic otitis media (Bondy, 2007).

Klinefelter syndrome is a result of a 47,XXY karyotype (or a rare variation, such as 48,XXXY or 49,XXXXY) and occurs in estimated frequency of 1 in 500-700 males, although it is suspected to be underdiagnosed (Bojesen et al., 2003; Herlihy et al., 2011). It is characterised by feminine distribution of adipose tissue, sparse facial and body hair, small testes, and sterility due to azoospermia. 45,X/46,XY mosaicism (sometimes referred to as mixed gonadal dysgenesis) has a highly variable outcome. The gonadal phenotype ranges from complete gonadal dysgenesis (i.e. streak gonads composed mostly of fibrous tissue) to partial testicular development or asymmetric gonadal development (for example, a streak gonad on one side and a dysgenetic testis on the other) (Ocal et al., 2012). It is a common cause of ambiguous genitalia at birth (Thyen et al., 2006).

1.9.2. 46,XX DSDs

46,XX DSDs describe the incomplete gonadal or phenotypic feminisation of individuals with a normal female chromosome complement. The most common form of 46,XX DSD is congenital adrenal hyperplasia (CAH), a condition in which an adrenal enzymatic defect causes the adrenal glands to produce high levels of androgens, resulting in ambiguous or fully masculinised external genitalia. Patients with CAH have ovaries and Müllerian structures, and most often identify as female (Achermann et al., 2015; Dessens et al., 2005). A rarer subgroup is 46,XX testicular or ovotesticular DSD (i.e. XX male development), often caused by a translocation of the *SRY* gene to the X chromosome, resulting in partial or complete testis development (Abbas et al., 1993; McElreavey et al., 1993). Other causes include loss-of-function mutations in key ovary-determining genes such as

WNT4 (Mandel et al., 2008) and *RSPO1* (Parma et al., 2006), or duplication of *SOX9* (Cox et al., 2011).

1.9.3. 46,XY DSDs

46,XY DSDs can generally be split into two classes: those conditions that arise from disrupted testis determination and those conditions that arise from under-masculinisation of an XY fetus due to disrupted androgen synthesis or action. The latter group includes 5α-RD2 deficiency, androgen insensitivity syndrome and defects in testosterone synthesis (Achermann et al., 2015). The former group, 46,XY gonadal dysgenesis (GD), is the most relevant to this thesis. Complete gonadal dysgenesis (CGD) is associated with bilateral streak gonads (fibrotic tissue with no testicular function), Müllerian structures and completely female external genitalia. Partial gonadal dysgenesis (PGD) is associated with varying degrees of testis formation, a mixture of Wolffian and Müllerian structures and usually ambiguous genitalia. Loss-of-function mutations in *SRY* account for 15% of 46,XY CGD cases (Bashamboo and McElreavey, 2016). Mutations in *NR5A1* (*SF1)* and *MAP3K1* are each estimated to contribute to 10-15% of 46,XY GD cases (Baxter et al., 2015). Other critical sex-determining genes regularly implicated in 46,XY GD include *SOX9*, *DHH*, *WT1*, *GATA4*, *DMRT1*, *WNT4* and *NR0B1* (Bashamboo et al., 2017). This gene list is consistent with the conserved nature of mouse and human sex determination.

1.9.4. Diagnosis of DSDs

Providing a diagnosis to patients with DSDs is an important aspect of clinical management. Some DSDs are syndromic, associated with a risk of potentially life-

threatening adrenal (Camats et al., 2018) or renal problems (Pelletier et al., 1991). Many conditions also increase the risk of gonadal tumours (Cools et al., 2006). A diagnosis in these cases allows the patient and clinician to assess and minimise these risks. A diagnosis in infancy can assist in sex assignment; while surgical assignment of sex is less advocated than in previous decades, information about likely gender identity, options for fertility preservation, and prognoses for sexual and endocrine functions are vital for making the best decisions about care and treatment (Achermann et al., 2015). An additional benefit to providing a diagnosis, especially a detailed genetic one, is empowerment. By removing the uncertainty surrounding a patient's condition it allows them to accept, understand and take control of their treatment, self-manage when possible and search for support groups of similarly affected people.

Among DSDs, non-syndromic forms of 46,XY gonadal dysgenesis are particularly difficult to diagnose. As previously mentioned, mutations in *SRY*, *NR5A1* and *MAP3K1* account for many of these cases, but approximately half of patients do not receive a definitive genetic diagnosis. With the advent of cheaper and faster next generation sequencing techniques, clinicians and researchers are improving the methodologies used when searching for causal mutations in DSD patients (Barseghyan et al., 2018a). Sanger sequencing of individual candidate genes has proven ineffective for diagnosis, especially when biochemical patient data is not enlightening. Recently developed gene panels, simultaneously sequencing 50+ candidate genes, are more efficient (Barseghyan et al., 2018a). However, such panels still rely on our prior knowledge of the genetic pathways underlying

mammalian sex determination, which remains incomplete. Exome sequencing is a method used to simultaneously analyse protein coding regions across the whole genome to identify variants. By sequencing more than just known sex-determining genes it has the potential to identify novel DSD genes. Exome sequencing has been used with some success: pathogenic and likely pathogenic variants were identified in 25% of a cohort of 46,XY DSD patients, most of whom had previously tested negative for mutations in *SRY* or *NR5A1* (Baxter et al., 2015). Difficulties arise when interpreting the many variants identified; exome sequencing typically identifies 21,000 variants per person that differ from the reference human genome (Lee et al., 2014). Establishing the causality of a particular variant requires caution and validation through further research using *in vitro* and *in vivo* models. Combining exome sequencing results and research using mice can identify novel sex-determining genes (Harris et al., 2018) and improve the rate of genetic diagnosis.

1.9.5. The oligogenic nature of DSDs

One problem with current methods for determining a genetic cause in patients with DSDs, is that the search has been limited to significantly disruptive mutations, in pivotal sex-determining genes. One step when identifying a causative variant is to search for the polymorphism in ancestry-matched, healthy control individuals, and rejecting those polymorphisms that appear 'common' (e.g. a frequency of over 1%) in the general population. However, this could lead to the rejection of a polymorphism that, while not phenotypically disruptive on its own, could have caused a small reduction in protein function or stability. Recently it has been suggested that DSDs should be considered as oligogenic

traits (Bashamboo and McElreavey, 2016; Mazen et al., 2016), especially when no definitive, causative mutation can be found. A patient may carry multiple, subtle polymorphisms that together lead to a DSD, but individually do not cause severe defects. Such mutations may be overlooked during exome sequencing because of their low but significant prevalence in the population.

Mazen et al. (2016) and Allali et al. (2011) reported two individuals with the same pathogenic *de novo* missense mutation in *NR5A1*: one 46,XY male patient exhibited distal hypospadias, while the other presented with 46,XY gonadal dysgenesis, a far more severe condition of disrupted testis determination. Interestingly, the latter patient was also found to harbour a missense mutation in *MAP3K1*, a gene known to cause 46,XY GD (Pearlman et al., 2010). Mazen et al. (2016) suggest that the combination of the two mutations is responsible for the more severe gonadal phenotype, and that the phenotypic variability seen in patients with mutations in *NR5A1* is due to digenic or oligogenic inheritance of mutations in other testis-determining genes. A further study sought to assign an oligogenic mode of inheritance to the varied phenotypes in four other cases of 46,XY DSD associated with *NR5A1* mutation (Camats et al., 2018). The authors used whole genome/exome sequencing to identify at least one other potentially deleterious variant in the four patients, and loosely correlated the numbers of these mutations to the severity of the DSD phenotype.

In mice, the concept of a genetic background that affects the phenotypic outcome of mutations is well accepted; different inbred laboratory strains are predisposed to particular disorders, including strain-specific differences in

susceptibility to sex reversal. The C57BL/6J (B6/J) strain is particularly sensitive to XY sex reversal when testis-determining genes are disrupted. For example, loss of *Gadd45γ* on a mixed genetic background does not cause sex reversal and homozygotes are fertile (Hoffmeyer et al., 2001), whereas on a congenic B6/J or B6/N genetic background, complete male-to-female sex reversal is reported (Gierl et al., 2012; Johnen et al., 2013; Warr et al., 2012). Exacerbation of sex reversal phenotypes on a B6/J background, compared to a mixed genetic background, was also observed during investigations into the loss of *Map3k4* (Bogani et al., 2009) and *Fgfr4* (Bagheri-Fam et al., 2017). The importance of genetic background is a common theme throughout this thesis. Whilst a "genetic background" affecting phenotype caused by mutations in human disorders is difficult to determine, it is likely to be incredibly important given that the human population is far more genetically heterogeneous than inbred mice. Therefore, considering how a mutation affects gonad development within the context of an individual's genomic variation is imperative when attempting to identify a genetic cause for a DSD. Sex determination is a balancing act; a polymorphism here and a polymorphism there may be enough to tip the balance in favour of the ovarian fate in an XY individual.

1.10. Aims of this thesis

Many of the genes associated with DSDs were identified or confirmed using gain- and loss-of-function approaches in mice, including the characterisation of *SRY* and *SOX9* and the crucial role they play in testis determination (Gubbay et al., 1990; Koopman et al., 1991; Morais da Silva et al., 1996). However, there are

likely many genes involved in sex determination that are still undiscovered. In this thesis, a fascinating and long-standing mouse model of 46,XY gonadal dysgenesis, known as B6.Y^{POS} sex reversal, will be used to investigate the genetic pathways and antagonisms underlying sex determination. This thesis takes advantage of the intricate differences both between inbred mouse strains, and between individual colonies of the same strain, maintained in different facilities. How genetic background affects the severity of sex reversal will be examined, with the ultimate aim to identify novel sex-determining genes.

Chapter 2: Materials and Methods

2.1. General solutions

Table 2.1 List of general solutions

Reagent	Composition
TE-Tween	50 mM Tris-HCl (pH 8.0), 1 mM EDTA, 0.5% Tween-20 in H_2O
10xTBE	1 M Tris base, 1 M boric acid, 0.02 M EDTA disodium salt in H_2O
PBT	PBS + 0.1% Tween-20
TBST	25 mM Tris-HCl (pH 7.5), 140 mM NaCl, 2.7 mM KCl, 0.1% Tween-20 in H_2O
20xSSC	2 M sodium citrate in H_2O, pH 4.5, autoclaved
WMISH prehybridisation solution	50% formamide, 5xSSC, 0.1% Tween-20, 50 µg/ml Heparin
WMISH hybridisation solution	Prehybridisation solution + 100 µg/ml yeast RNA, 100 µg/ml salmon sperm DNA
Formamide/SSC/SDS mix	50% formamide, 2xSSC, 1% SDS in H_2O
NTMT	50 mM Tris-HCl (pH 9.5), 100 mM NaCl, 50 mM $MgCl_2$, 0.1% Tween-20 in H_2O
Milk block	0.2% milk powder, 1% BSA, 0.3% tritonX in PBS

2.2. Mouse lines and genotyping

All mouse breeding and experimentation was approved by the Animal Welfare and Ethical Review Body (AWERB) at Harwell and conducted with licensed approval from the UK Home Office (PPL 30/2877). Mouse husbandry was carried out by the staff of the Mary Lyon Centre (MLC). Mice were housed in individually ventilated cages (IVCs) in a specific pathogen-free (SPF) environment.

All experimental mice were maintained on a C57BL/6J (B6/J) genetic background, unless otherwise stated. This inbred strain of *Mus musculus* has a known sensitivity to disruption in early testis development (Bouma et al., 2007; Correa et al., 2012) making it an ideal background for studying sex determination.

2.2.1. Chromosomal sexing

All experimental embryos and adults suspected of sex reversal were genotyped for chromosomal sex using a gel-based PCR assay. Primers simultaneously amplify the *Ube1x* and *Ube1y1* genes, present on the X and Y chromosomes respectively. The Y-linked amplicon is larger than the X-linked amplicon, providing a read-out of chromosomal sex (Warr et al., 2009). See **Table 2.2** for primers and reaction conditions.

2.2.2. B6.Y^{POS}

B6.Y^{POS} is a line of mice generated my E. Eicher in the 1980s (Eicher et al., 1982), from crossing a male wild mouse from the Val Poschiavo region to an inbred B6/J female. Male offspring were crossed to B6/J females for multiple generations, to produce a line in which the X chromosome and autosomes are considered congenic B6/J, and the Y chromosome derives from a "*Mus poschiavinus*" male. It should be noted that "*M. poschiavinus*" is not an accepted subspecies or species definition; phylogenetic studies indicate that these mice are likely a subspecies of *Mus musculus domesticus* (Silver, 1995). However, *M. poschiavinus* is the term used throughout sex determination literature when discussing B6.Y^{POS}. In this thesis I will refer to the Y chromosome carried by these mice as Y^{POS}, to distinguish it from other *M. m. domesticus* Y chromosomes and for the sake of consistency with the published literature.

2.2.3. B6.Y^{POS}.*Mod13*

A B6.Y^{POS} colony was brought to MRC Harwell in the 1990s. This line was re-derived from the frozen sperm archive in 2011, and has since been maintained by continued back-crossing of B6.Y^{POS} males to unrelated B6/J females. Historically, 35% of B6.XY^{POS} mice at Harwell develop as males or

hermaphrodites, and 65% develop as females (M. Pope, personal communication). The presence of the Y^{POS} chromosome can be confirmed using a PCR assay that amplifies the Y-linked gene *Zfy2* because $Zfy2^{B6}$ has an 18 bp deletion (**Table 2.2**).

Since the discovery of the *Mod13* genetic modifier in the colony, males heterozygous for *Mod13* have been selected for back-crossing to B6/J females. B6.Y^{POS}.*Mod13* mice were genotyped for the presence of individual SNPs by Sanger sequencing (section 2.3.4) or Lunaprobe™ assay (section 2.3.3). Primer sequences are described in **Tables 2.3** and **2.4**.

2.2.4. B6.Y^{AKR}

Similar to the B6.Y^{POS} line, B6.Y^{AKR} mice carry the autosomes and X chromosome from B6/J and the *Mus domesticus* Y chromosome from AKR/J (Nagamine et al., 1987). The line was maintained by crossing a B6.Y^{AKR} male to B6/J females. All B6.XY^{AKR} mice develop as male, but the AKR Y chromosome confers a strong sensitivity to XY sex reversal (Bouma et al., 2007; Eicher, 1988). Y^{AKR} can be differentiated from Y^{B6} using the same *Zfy2* PCR assay described in the section 2.2.3.

2.2.5. *Ptch1*^{tm1Mps}

This *Ptch1* knockout allele was generated by Goodrich et al. (1997), using homologous recombination to replace the start codon, exon 1 and part of exon 2 with *LacZ* and a neomycin resistance gene. Mice were maintained as heterozygotes. *Ptch1*$^{-/-}$ embryos die between 9.0 dpc and 10.5 dpc. The colony at MRC Harwell is congenic on B6/J. Adults were genotyped by the MRC Harwell

GEMS team (Deen Quaileed, Ramakrishna Kurapati). Embryos were genotyped using a PCR assay designed to amplify a *Ptch1-lacZ* product (**Table 2.2**).

2.2.6. *Dhh*^{tm1Amc}

This knockout allele of *Dhh* was generated by Bitgood et al. (1996). Most of the protein-coding sequence (exons 1 and 2 and part of exon 3) of *Dhh* is replaced by a neomycin phospotransferase selection cassette. Mice were imported to the Mary Lyon Centre, Harwell, in 2004 and backcrossed to B6/J until congenic, then archived via sperm freezing. The line was re-derived in late 2016. Mice were maintained as heterozygotes and genotyped using separated PCR assays provided by the Jackson Laboratory (**Table 2.2**).

2.2.7. *Adamts16*$^{em2(IMPC)H}$

Mice lacking functional *Adamts16* were generated using CRISPR/Cas9 genome editing. A single- guide RNA, designed by the transgenics team at MRC Harwell, and targeting a critical exon, was injected into 1-cell egg of B6/NTac females, along with Cas9 RNA. A founder was identified carrying two different deletions, one of 906 nucleotides and one of 923 nucleotides. Both alleles result in a complete loss of exon 5. Splicing from exon 4 to exon 6 is predicted to result in a frameshift leading to a premature stop codon. The predicted protein would contain 252 amino acids of the full 675 amino acids, and lack the N terminus of the Peptidase M12B domain.

Offspring carrying the 923 nucleotide deletion were bred to produce an adult cohort, which was analysed as part of the International Mouse Phenotyping Consortium. Offspring carrying the 906 nucleotide deletion were used to

generate a heterozygous colony, maintained on a B6/NTac background, and are the focus of chapter 6 of this thesis. All adults were genotyped by MRC Harwell GEMS team. Embryos were genotyped using a 3 primer PCR assay: a common forward primer binds upstream of the deletion, the WT reverse primer binds to the deleted sequence and the del reverse primer binds downstream of the deletion. Differing band sizes are detected by gel electrophoresis. See **Table 2.2** for details.

2.3. DNA methods

2.3.1. DNA extraction for genotyping

Genotyping was performed using genomic DNA extracted from embryonic material collected during dissection or ear biopsies taken from animals after 3 weeks of age. Tissue was treated with 27 µl 0.75 mg/ml Proteinase K in TE-Tween and heated to 55°C for 90 min, then to 95°C for 5 min to denature the enzyme. The tissue was vortexed and debris pelleted by centrifugation at 1,500 x g for 5 min. The supernatant was diluted in ddH_2O for use in genotyping assays: 1:20 for standard PCR assays or 1:100 for LunaProbe^{TM} assays.

2.3.2. Genotyping using standard PCR/gel electrophoresis

Primers for new assays were designed using the NCBI Primer-BLAST online tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Primers were synthesised by Sigma Aldrich. Each standard genotyping reaction contained:

- 1x Maxima Hot Start PCR Master Mix (Thermo Scientific)
- 0.25 µM of each primer

- 1 µl diluted DNA (approximately 25 ng/µl)
- H_2O to a total volume of 10 µl

1x Maxima Hot Start PCR Master Mix contains Maxima Hot Start Taq polymerase, Hot Start PCR buffer, dNTPs (0.2 mM each of dATP, dTTP, dCTP and dGTP) and 2mM Mg^{2+}. Reactions were performed using a Biorad tetrad2 thermal cycler. Primers and their thermocycling reaction conditions are listed in **Table 2.2**.

Table 2.2 Gel-based genotyping PCR primers and assay conditions

Assay	Primer		PCR conditions	Product size	Assay design
XY sexing	Fwd: Rev:	TGGATGGTGTGGCCAATG CACCTGCACGTTGCCCTT	95˚C 3 min 34 x [95˚C 10 sec, 57˚C 10 sec, 72˚C 15 sec] 72˚C 2 min	X: 252 bp Y: 334 bp	N. Warr (Warr et al., 2009)
Adamts16-del906	Fwd: WT: Del:	GAGTCCTGCTCTGCAAGTCC AGGCATGTCTATGGAGGAACAT TGTGTGAAGATATCAGTGAGCCC	95˚C 4 min 35 x [95˚C 20 sec, 56˚C 20 sec, 72˚C 15 sec] 72˚C 5 min	WT: 399 bp Del: 296 bp	C. Livermore
Ptch1-tm1Mps	Fwd: Rev:	CAATGTGGCAATGGAAGGCG TAGATGGGCGCATCGTAACC	95˚C 4 min 35 x [95˚C 20 sec, 57˚C 20 sec, 72˚C 30 sec] 72˚C 5 min	508 bp	C. Livermore
Dhh-tm1Amc	WT: Del: Rev:	ATCCACGTATCGGTCAAAGC GGCATGCTGGGGATGCGGTG GGTCCAGGAAGAGCAGCAC	95˚C 4 min 35 x [95˚C 30 sec, *56/63˚C 30 sec, 72˚C 30 sec] 72˚C 5 min	WT: 442 bp Del: 110 bp	Jax.org/strain/002784
Zfy	Fwd: Rev:	CATTAAGACAGAAAAGACCACCG GTGAGGAAATTTCTTCCTGTGG	95˚C 5 min 36 x [95˚C 30 sec, 57˚C 30 sec, 72˚C 30 sec] 72˚C 5 min	POS/AKR: 202bp (Zfy1 and 2) B6: 202bp (Zfy1) + 184bp (Zfy2)	Prager et al. (1997)

*Assay requires two separate reactions: rev + wt fwd primers anneal at 56 ˚C, rev + mut fwd primers anneal at 63˚C.

Table 2.3 Lunaprobe^{TM} genotyping primers and assay conditions

SNP	SNP position	Primers and probe		Annealing temp	Assay design
rs29231478	13:40958356	Fwd: Rev: Probe (fwd):	AGCTGGTATTTCTGACCCGC TTCAGCATAGCCAAGACCAGAGT CAGCAGCTCCGGCCTAAATG[AmC3]	58°C	C. Livermore
rs13481788	13:43632644	Fwd: Rev: Probe (fwd):	GACATTAGGAACAGACCACACTAT CTTAATTATGGCCCACTGTAGAAA CAATGTTTTTTTCCCCTCTGTCTCTA[AmC3]	60°C	C. Livermore
rs29568118	13:48119885	Fwd: Rev: Probe (rev):	ATGTCACAGTTGACTTAATCATCGT ATGGACCATGAGACAAATAATGGAA GACAGTGTCATGACAGATATTAAAATCTC[AmC3]	60°C	C. Livermore
rs48656322	13:52288276	Fwd: Rev: Probe (fwd):	ATGGAATCTCTGGCTCACG AAACAAACTATGGGTTAGAAAGAATGC GCAAATTTGAACGATAATTCCTCTGTC[AmC3]	60°C	C. Livermore
rs29238368	13:55400536	Fwd: Rev: Probe (fwd):	CAACATGATGTCCTACAGCGA AGGTCTCAGGGTATAGGAGTAAT GGTGCATGGGACAACCTTTGCTTATG[AmC3]	60°C	R. Kurapati (Harwell GEMS core)
rs241205325	13:60767201	Fwd: Rev: Probe (fwd):	GCAATTCAGAGTTTACCCATTGCAG TGTCAGATGCCAACTTGTCAGGAG GAAATGATGGACCATCCCCAGAAAAG[AmC3]	60°C	C. Livermore
rs224110566	13:62773685	Fwd: Rev: Probe (fwd):	GGCTCTATCACGTTGATTACATTAGTAGG CACAACACAGTTCTCTTCAAACTC CACACTAATTACATTTGAAGGGTTTCTCT[AmC3]	60°C	R. Kurapati (Harwell GEMS core)
rs29245244	13:69504739	Fwd: Rev: Probe (rev):	TGGCTGAGCACAATGTAAGCGTA CTGCTGTAGTTGGGAGGCATTT GGTGCCTTTGTATTGCAGGAAAAGATG[AmC3]	60°C	R. Kurapati (Harwell GEMS core)
rs47173256	13:71076570	Fwd: Rev: Probe (fwd):	ATGGAATCTCTGGCTCACG AAACAAACTATGGGTTAGAAAGAATGC GCAAATTTGAACGATAATTCCTCTGTC[AmC3]	58°C	C. Livermore

PCR programme: 95°C for 2 min, 56 cycles of [95°C 30 sec, anneal 30 sec, 72°C 30 sec] then 95°C for 30 sec, 25°C for 30 sec, 15°C for 20 sec.

Table 2.4 SNP sequencing genotyping primers and assay conditions

SNP	SNP position	Primers		Product size	Assay design
rs29231478	13:40958356	Fwd: Rev:	TCCTGGTGGACTAGGGTAGC TTGCCAGATCCTCACGGTTC	465 bp	C. Livermore
rs13481788	13:43632644	Fwd: Rev:	GCAAGGCACACGTTAGTAGG AGAGTTACTCATCAAACGAGCGA	490 bp	C. Livermore
rs29568118	13:48119885	Fwd: Rev:	GCACACTCCAGCCGAATCA TCCTAAATGATCCCACTCGGC	331 bp	C. Livermore
rs48656322	13:52288276	Fwd: Rev:	TCGATTCCAAGCCCTTTACCA TCAAAAACCGCCCACAATTCA	449 bp	C. Livermore
rs29238368	13:55400536	Fwd: Rev:	CAGCAACTGTCAGATTTGTCCC CCAGCTTGGTAGGTAGTGGG	416 bp	N. Warr
rs241205325	13:60767201	Fwd: Rev:	ACAGGAAAGGCATATCGACAACAGG GGGTTAGCTTTAGGCATGGGTG	479 bp	C. Livermore
rs224110566	13:62773685	Fwd: Rev:	GTGAAAAGGCTCTATCACGTTG TCACGACACTTTCACCTCCAAAATCA	256 bp	N. Warr
rs29245244	13:69504739	Fwd: Rev:	AGAGGAGCCTAGAACCCAGG TCGCTCGGGTCTGTGTTTAC	278 bp	N. Warr
rs47173256	13:71076570	Fwd: Rev:	GGCTGGAGAGATGGCTTTGT TTCCCTGCTGGCTTTCTTGT	252 bp	C. Livermore

PCR programme: 95°C for 4 min, 36 cycles of [95˚C 30 sec, 61˚C 20 sec, 72˚C 15 sec] then 72˚C for 5 min.

The primers listed only include those used to genotype for the *Mod13* region. A full list of sequencing primers can be found in **Table S1**.

PCR products were loaded into a 2% Ultrapure™ Agarose (Invitrogen)/ 1x TBE gel containing 1:10,000 SYBR Safe DNA gel stain (Invitrogen). A 100 bp DNA ladder (NEB) was loaded for comparison. The gel was electrophoresed for 30 - 40 minutes at 80 V, and products were visualised using a BioRad UV Gel Doc XR system.

2.3.3. Genotyping using LunaProbes™

Lightscanner high-resolution DNA melting analysis, using the double-stranded DNA-binding dye LCGreen, is a high throughput method for genotyping known SNPs. Primers and probes were designed according to the Lightscanner primer design software and accompanying user guide. Briefly, a primer pair was designed to amplify a 100 - 250 bp amplicon containing the target SNP. A 20 - 30 bp unlabelled oligo (the Lunaprobe™) was designed to bind over the target nucleotide on either the sense or antisense strand. This probe is synthesised with a 3'-amino-modifier-C3 block to prevent extension by the polymerase. An asymmetric PCR is performed in which the primer in the same orientation as the probe is present in 1/5th concentration of the other primer. This produces an excess of the target strand, to which the Lunaprobe™ binds. Oligo sequences and thermocycling conditions are listed in **Table 2.3**. Each PCR reaction contained:

- 1x KAPA Taq ReadyMix (KAPA Biosystems Inc)
- 1x LCGreen Plus dye (Idaho Technology Inc)
- 1 µM Lunaprobe™
- 1 µM primer 1 (opposite orientation as the probe)

- 0.2 µM primer 2 (same orientation as the probe)
- 2 µl diluted DNA (approximately 5 ng/µl)
- H_2O to a total volume of 10 µl

1x KAPA Taq Readymix contains 0.1 U KAPA Taq DNA Polymerase, KAPA Taq Buffer, dNTPs (0.2 mM of each dNTP), 1.5 mM $MgCl_2$ and stabilizers. Each reaction was overlaid with 20 µl mineral oil to prevent evaporation.

A Lightscanner instrument was used to produce melting curves. Polymorphisms in the PCR product cause a mismatch with the probe, thus lowering the melting temperature of the probe:amplicon duplex. Genotypes were therefore determined by the shape of the melting curve. An example is shown in **Fig 2.1**.

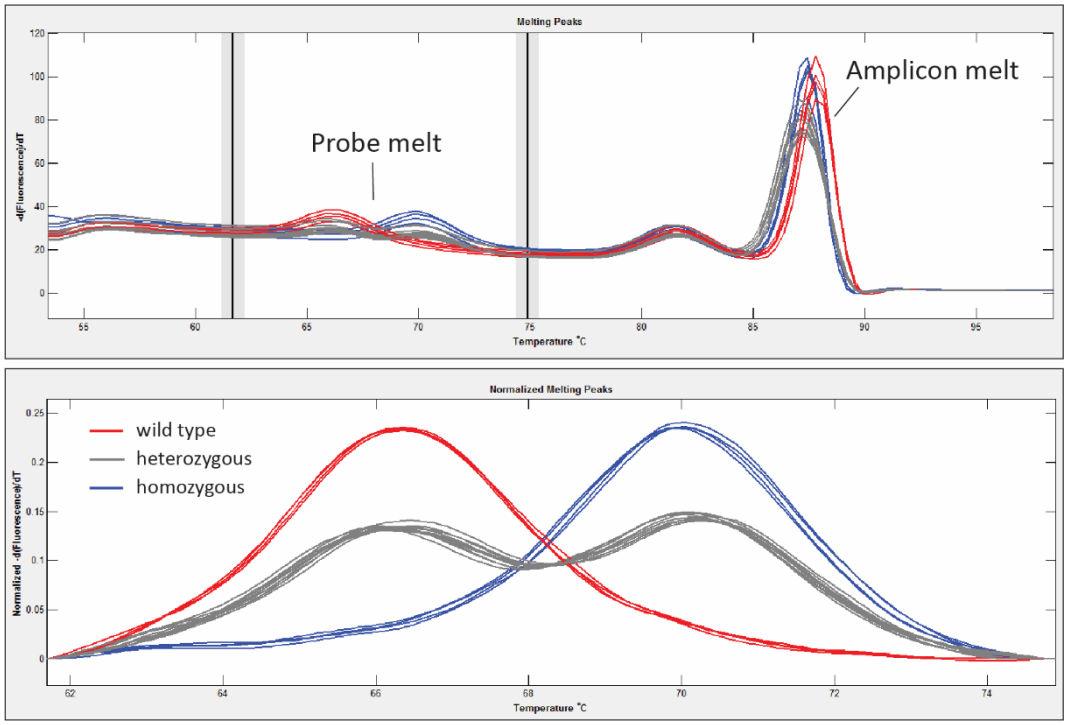


Figure 2.1 Lightscanner DNA melting analysis. The three genotypes are easily distinguishable by the shape of the probe melting curve.

2.3.4. Sanger sequencing

Primers were designed to amplify 200 - 300 bp flanking either side of a target SNP using the NCBI Primer-BLAST online tool and sequence data available from Ensembl.org. Standard PCR was performed as described in section 2.3.2 with a reaction volume of 100 µl. PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN) and eluted in 50 µl of H_2O. DNA was sent to GATC Biotech (Konstanz, Germany) along with 10 µM sequencing primer. The forward primer used for amplification was also used for sequencing unless otherwise stated. DNA trace files were examined for the presence of SNPs. Primers and their thermocycling reaction conditions are listed in **Table 2.4**.

2.3.5. *Ptch1* high-fidelity sequencing

Primers were designed to amplify overlapping sections of the 5' end of *Ptch1*. Primers were used to amplify this region using gDNA extracted (section 2.3.1) from Harwell B6/J, and B6.Y^{POS}.$Mod13^{mod/mod}$ ear tissue. Primers and thermocycling conditions are detailed in **Table 2.5**. Each PCR reaction contained:

- 1x Platinum PCR SuperMix High Fidelity (Thermo Scientific)
- 0.2 µM of each primer
- 5 µl diluted DNA (approximately 50 ng/µl)
- H_2O to a total volume of 90 µl

1x Platinum PCR SuperMix High Fidelity contains 20 U/ml complexed recombinant Taq DNA polymerase, Pyrococcus species GB-D thermostable polymerase, Platinum™ Taq Antibody, 60 mM Tris-SO_4 (pH 8.9), 18 mM $(NH_4)_2SO_4$, 2.2 mM $MgSO_4$, 200 µM dNTPs, and stabilizers.

PCR products were purified and sequenced as described in section 2.3.4. Sequence reads for each strain were assembled manually and variants identified by eye. Sequences were then aligned using Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/) to one another and to the other inbred strain sequences available from Ensembl.org.

2.3.6. Next generation sequencing of B6.Y^{POS}

Genomic DNA was extracted from tail tissue of a fertile B6.Y^{POS} male using the Illustra Nucleon genomic DNA extraction kit. DNA concentration and quality was quantified using a nanodrop and diluted to 50 ng/µl. Sequencing was performed by the Wellcome Trust Centre for Human Genetics High Throughput Genomics facility at Oxford University (Oxford, UK), generating a mean read depth of 13x from a single lane of 100 nucleotide paired end reads. Analysis was performed by Michelle Simon in the MRC Harwell bioinformatics group. Paired-end Illumina reads from the B6.Y^{POS} mouse was aligned to the reference genome (NCBIM38/mm10) using BWA aligner software. Detection of SNPs in each alignment was made using a customised version of the Genome Analysis Toolkit (GATK) with default parameters. A filtering strategy was adopted to identify high-confidence SNPs and reduce false positives. SNP sites that failed GATK's internal status check and a mapping quality <200 were removed from any further analysis. The remaining SNPs were compared to the precompiled SNPs found in 36 strains from the Mouse Genome Project.

Table 2.5 PCR assays for amplifying/sequencing *Ptch1*

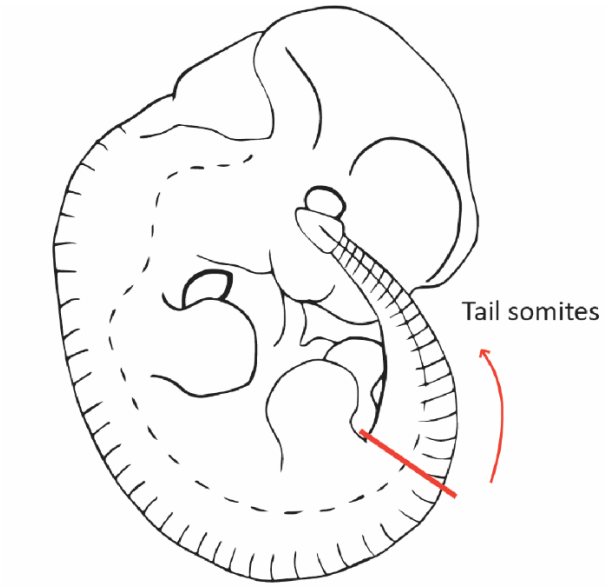
Assay	Primers	Product size	Annealing temp
Ptch1_Seq_1	Fwd: CTCCATCATCAGCTGTACCAGG Rev: TCTCGTCTGCCAATTTGGGT	938 bp	58°C
Ptch1_Seq_2	Fwd: TCTAGAAACAAGCACATAAACCAAC Rev: AGTTTGTACAGCCTCTCGGC	886 bp	58°C
Ptch1_Seq_3	Fwd: TGGCATGACAAATTCTGGAGG Rev: CCTTCCCCAAGGTTTGGGTT	779 bp	58°C
Ptch1_Seq_4	Fwd: TCCATGAGCACTAGGCTTCC Rev: GGTGGAAAGGTGACCCCAAT	764 bp	58°C
Ptch1_Seq_5	Fwd: ATCTGGAGTATGGGGGTGGG Rev: AGCATATGTCGCTGCTGCTTA	749 bp	62°C
Ptch1_Seq_6a	Fwd: CCACCTAGGCCAGAGCTAAGA Rev: AGAGCTGTGCTGTGTAAGGC	780 bp	60°C
Ptch1_Seq_6b	Fwd: CTTGGTTGTGGCACAGAGAC Rev: TTTGTGGGTGGAACATCACCA	1316 bp †	62°C
Ptch1_Seq_7*	Fwd: CCTGAAACTCTGACTGGGGTG Rev: CCTGATCCTACCTCCTGTTGAC	901 bp	62°C
Ptch1_Seq_v1	Fwd: GTCTCTGCGCGGTCTGTAG Rev: AGTTGAGATCTATGGGCAGTGG	319 bp	58°C
Ptch1_Seq_8	Fwd: ATGGTGGTGATGTTCCACCC Rev: GTGAATTGTGTCCCCGTCCT	637 bp	60°C
Ptch1_Seq_9*	Fwd: AGGACCGTGACCAACTTCAAA Rev: CCCGTAGCTTTGTGGCTTCA	462 bp	58°C
Ptch1_Seq_10a	Fwd: GGACGGGGACACAATTCACT Rev: CCGGGGTCAAATCACAGGAA	677 bp	60°C
Ptch1_Seq_10h*	Fwd: CCCGGTTCAGTCCACAAGAT Rev: TGGGTCTGATTCCTCCGACT	397 bp	58°C
Ptch1_Seq_10g	Fwd: CCAGGTGACTCACAGTGGAC Rev: ACAGGTGTCTCCAGCCAATG	272 bp	62°C
Ptch1_Seq_11*	Fwd: ACAATCACTTGGTGGGCAGG Rev: GGCGTGGAATGCTCTCCTTC	800 bp	65°C
Ptch1_Seq_12*	Fwd: CTCAGGACACGGTCCAAAGA Rev: GGGAAAGCCTTACCTGATGC	667 bp	58°C
Ptch1_Seq_13	Fwd: CTCTCGAGAACCCAGCAGAAG Rev: AGCAAGACATTGGCAGCTTG	888 bp	60°C
Ptch1_Seq_14	Fwd: ACAGTGAGTTCTGCGTTCCC Rev: GTGTCGTCTCTTGGCTCTGT	936 bp	60°C
Ptch1_Seq_15	Fwd: ACCACAAAGCAGAGATGGGTT Rev: GCTGGCCAAGTAGGGTAAGTT	661 bp	60°C
Ptch1_Seq_16*	Fwd: TCAGAAAGACCGACATGCCA Rev: TGGCATCACAAGGCCAGATT	995 bp	58°C
Ptch1_Seq_17	Fwd: GAGCATCGTACACAGCCTCT Rev: TCTCCTTTGTTCAGGCCTTTCT	999 bp	60°C
Ptch1_Seq_18	Fwd: CAAGCTGTGGAGTTGCCTCTA Rev: AGATGGCTCTGTCCACCTACA	638 bp	60°C
Ptch1_Seq_19	Fwd: TATTTCTGGAGGGCCGACGTT Rev: TGTGCATTCTGTCGCCTCTC	993 bp	60°C
Ptch1_Seq_20	Fwd: TTTCCAGTCTCCACAGGCATC Rev: AGGCTTGGTGGGGATAGCAT	697 bp	60°C

PCR programme: 94°C for 2 min, 36 cycles of [94°C 30 sec, anneal 30 sec, 72°C 1 min]
*Product sequenced with both fwd and rev primers due to repetitive elements
† Larger product required extension time of 1 min 30 sec

2.4. Embryo analysis

2.4.1. Embryo collection and staging

To obtain embryos at specific ages, timed mates were set up and females were checked for vaginal plugs the next morning. Noon on the day of plug was counted as 0.5 dpc. Pregnant females were culled by cervical dislocation (Schedule One) under Home Office regulations. Uteri were removed through an opening in the abdomen and embryos removed. Embryos 14.0 dpc or older were immediately decapitated in accordance with Home Office regulations. Embryos were dissected in ice cold PBS using a Nikon SMZ800 microscope. Between 11.0 dpc and 12.5 dpc embryos were accurately staged by counting the number of tail somites (ts) posterior to the hindlimb bud (**Fig 2.2**). Embryos over 12.5 dpc were staged using morphological characteristics, guided by the Theiler staging criteria for mouse embryo development.



Stage	11.0 dpc						11.5 dpc					
Tail somite number	9	10	11	12	13	14	15	16	17	18	19	20

Stage	12.0 dpc						12.5 dpc					
Tail somite number	21	22	23	24	25	26	27	28	29	30	31	32

Figure 2.2 Tail somite staging of mouse embryos. Somites are segments of mesoderm that form sequentially on either side of the midline approximately every 2 hours (Tam, 1981). Between 11.0 dpc and 12.5 dpc somites in the tail region of the embryo are clearly visible and can therefore be used to precisely stage the embryo. The table shows how tail somite (ts) number relates to age in days post *coitum* (dpc).

2.4.2. Classification of 14.5 dpc gonads

Gonads were removed from 14.5 dpc embryos and viewed on a Leica MZ16F stereomicroscope with strong bottom lighting. Gonads were grouped into one of five categories according to the proportion of testicular tissue; Ovary (0% testicular tissue), **O**votestis (less than 40% testicular tissue) ovotestis (40-60% testicular tissue), ovo**T**estis (>60% testicular tissue) and testis (100% testicular tissue). This categorisation was based on gonad shape and morphology and the presence of a coelomic vessel. Ovarian tissue is thin, transparent and speckled. Testicular tissue is rounded with visible cord structure and an organised blood vessel running along the coelomic epithelium. All classifications were performed at the time of dissection, blind to the genotype of the embryo. Images were captured using a camera attached to the Leica MZ16F and used to confirm classifications.

2.4.3. Tissue processing

2.4.3.1. WMISH

Urogenital ridges (UGRs) from 11.5 dpc embryos and gonads with attached mesonephroi from 14.5 dpc embryos were collected in microcentrifuge tubes and fixed in 4% paraformaldehyde/PBS overnight. The tissue was taken through an ethanol dehydration series (5 minute washes each; 2 x PBS, 1 x 25% ethanol in PBS, 1 x 50%, 1 x 75%, 2 x 100%) and stored in 100% ethanol at -20˚C.

2.4.3.2. Quantitative real time PCR (qRT-PCR)

Urogenital ridges were collected from 11.5 dpc embryos and further dissected using a needle to separate the gonad from the mesonephros. Gonad pairs were

placed in microcentrifuge tubes in a minimal volume of PBS. Kidneys from 14.5 dpc embryos were collected in pairs in a minimal volume of PBS. Tissue was snap frozen on dry ice and stored at -70°C.

2.4.4. Wholemount *in situ* hybridisation (WMISH)

2.4.4.1. Riboprobe synthesis

All riboprobes were previously designed and validated and are listed in **Table 2.6**. Riboprobes were synthesised from cDNA clones. To generate antisense template, approximately 15 µg of plasmid DNA was linearised with 20 U of the relevant enzyme in 1x buffer and a total volume of 50 µl (all reagents from NEB). The reaction was incubated at 37˚C for 2 hours. The linearised DNA was precipitated with 12.5 µl 10 M ammonium hydroxide and 125 µl -20˚C 100% ethanol. It was centrifuged at 13,000 x g, 4˚C for 20 minutes, the supernatant was removed and replaced with 100 µl -20˚C 70% ethanol, before being centrifuged again for 2 minutes. The supernatant was removed and the DNA pellet was briefly air dried before resuspension in 20 µl RNase-free water.

2 µl linearised DNA was transcribed using 1x DIG labelling mix, 1x transcription buffer, 40 U Protector RNase inhibitor and 40 U RNA polymerase, in a total volume of 20 µl (all reagents from Roche). The reaction was incubated at 37˚C for 2 hours. Any remaining DNA was digested by adding 20 U DNase I (RNase-free; Roche) and incubating at 37˚C for a further 20 minutes. The reaction was stopped using 2 µl 0.2 M EDTA. The RNA was precipitated at -20˚C overnight or on dry ice for 30 minutes in 2.5 µl 4 M LiCl and 75 µl 100% ethanol. The reaction was centrifuged at 13,000 x g, 4˚C for 30 minutes, washed with 500 µl 70%

ethanol, before centrifuging again for 5 minutes. The supernatant was removed and the RNA pellet resuspended in 100 µl RNase-free water. The probe was stored at -20˚C ready for use.

Table 2.6 Riboprobes used for WMISH

Gene of interest	Linearising enzyme	RNA polymerase	Reference
Sry	SacI	T3	(Bogani et al., 2009)
Sox9	HindIII	T7	(Wright et al., 1995)
Stra8	HindIII	T7	(Bogani et al., 2009)
Insl3	SalI	T7	(Warr et al., 2009)

2.4.4.2. WMISH protocol

WMISH was used to analyse the temporal and spatial location of mRNAs in 11.5 - 14.5 dpc gonad explants. Deionised RNase-free AnalaR water was used as standard. All steps were performed at room temperature unless stated otherwise. Tissues were placed in 15 mm Netwell^{TM} inserts (Corning) in 12-well culture dishes.

Tissue collected as described in section 2.4.3.1 was rehydrated through an ethanol series (5 min/ wash) and washed twice in PBT (5 min), then bleached in 6% H_2O_2 in PBT for 45-60 minutes, depending on the size of the tissue. Tissue was washed 3 x 5 minutes PBT, and then digested in 10 µg/ml Proteinase K (Roche) for 15 minutes. The reaction was immediately stopped with a 5 minute wash in 2 mg/ml glycine. After 2 x 5 minute PBT washes, the tissue was post-fixed in 0.2% gluteraldehyde in fresh 4% PFA/PBT for 20 minutes. The tissue was washed 2 x 5 minute in PBT then placed in pre-hybridisation solution (see **Table 2.1**) at 70˚C for 4 hours. Riboprobe was denatured at 80˚C for 10 minutes. The probe was cooled on ice for 1 minute before being added to hybridisation

solution at a concentration of 5 µl/ml. Tissue was incubated at 70˚C overnight in hybridisation solution and probe.

The following day, tissue was washed in the formamide/SSC/SDS mix at 70˚C for 2 x 30 minutes, then at 65˚C for 2 x 30 minutes. The tissue was washed in 1x TBST for 3 x 5 minutes, then transferred to 10% heat inactivated sheep serum (HISS) in 1x TBST for at least 3 hours pre-blocking. Meanwhile 2ml/ well of 1% HISS/TBST was prepared. Approximately 20 µl embryo powder (prepared from homogenised 12.5 dpc – 14.5 dpc embryos, washed with acetone, ground to a powder and dried) was heat inactivated at 70˚C for 30 minutes in 1 ml TBST, before being pulsed down. The TBST was removed and replaced with 1 ml 1% HISS. 1 µl/ well of anti-digoxigenin-AP-Fab fragments (Roche) was added to the embryo powder solution, which was rotated at 4˚C for at least 1 hour. Embryo powder was pulsed down and the 1% HISS containing the antibody was transferred back into the remaining 1% HISS. Tissue was then incubated in this antibody solution at 4˚C overnight.

The next day, tissue was washed in 1x TBST for 3 x 5 minutes, then 5 x 1 hour, before being washed at 4˚C in TBST overnight. The next day tissue was washed 3 x 10 minutes in NTMT, and then transferred into glass dishes. Tissue was incubated in the dark in alkaline phosphatase stain (20 µl/ml NBT/BCIP (Roche) in NTMT) until the stain had sufficiently developed (normally 2 – 6 hours). The reaction was stopped by washing in PBT. The tissue was allowed to destain in the dark in PBT for up to 3 days until background staining was gone, before being fixed in 4% PFA/PBS for 1 hour. The tissue was then stored in PBT with 0.4% PFA

at 4˚C. WMISHs were photographed on 1% agarose/PBT plates, in PBT, using a Leica MZ16F stereomicroscope with a camera attached.

2.4.5. qRT-PCR

2.4.5.1. RNA extraction

Total RNA was extracted from frozen embryonic tissue using an RNeasy Plus Micro kit (QIAGEN), following the manufacturer's instructions. In short, tissue was lysed by vortexing for 15 seconds in a lysis buffer containing 1% β-mercaptoethanol and 57 pg/µl poly-A carrier RNA. Lysate was applied to a gDNA Eliminator mini spin column to separate RNA from DNA. The RNA in the flow-through was precipitated with 70% ethanol then applied to an RNeasy MinElute spin column which binds RNA but not proteins. The bound RNA was washed multiple times with specialist buffers and then eluted in 16µl RNase-free water. RNA concentration and quality was determined by measuring absorption at 260 nm using a Nanodrop ND-8000.

2.4.5.2. Reverse transcription

200ng of RNA was used to synthesise cDNA using a High-Capacity cDNA Reverse Transcription kit (Applied Bioscience) according to the manufacturer's instructions. Synthesis was carried out in a thermocycler with the following conditions: 25°C for 10 minutes, 37°C for 120 minutes, 85°C for 5 minutes. Following synthesis, cDNA was diluted 1:6 with H_2O.

2.4.5.3. qRT-PCR

Real-time PCR was performed using the Fast SYBR™ Green assay (Applied Biosystems). SYBR™ Green dye incorporates into double stranded DNA as it is

amplified by *Taq* polymerase using specific primers. With each amplification cycle SYBR™ Green incorporates into the new DNA and an increase in fluorescence is detected by the machine. The fluorescence intensity is directly proportional to the amount of double-stranded PCR product produced which allows the quantification the cDNA present. Real-time PCR measures the number of cycles it takes to reach a specific threshold of fluorescence, called the cycle threshold (Ct). The lower the Ct value, the higher the quantity of cDNA (and therefore mRNA) at the beginning.

Table 2.7 List of primers used for qRT-PCR

Gene	Forward primer	Reverse primer
Hprt1	CATTATGCCGAGGATTTGGAA	CACACAGAGGGCCACAATGT
Sry	TTATGGTGTGGTCCCGTGGT	GGCCTTTTTTCGGCTTCTGT
Sox9	AAGAAAGACCACCCCGATTACA	CAGCGCCTTGAAGATAGCATT
Rspo1	GCTCCAACTCTCGGAGACAC	GTCACTGTGCCCAGGTAGGT
Foxl2	TCATAGCCAAGTTCCCGTTC	TGATGAAGCACTCGTTGAGG
Fst	AGAGGTCGCTGCTCTCTCTG	AGCTTCCTTCATGGCACACT
Adamts16	GCCGAGTATAACAGCAAGCG	GGACATTGTCACAGCCAACC

Reactions were carried out in MicroAmp™ Fast Optical 96-well reaction plates (Applied Biosystems). All primer pairs used have previously been validated in the lab and are listed in **Table 2.7**. Each sample was analysed in duplicate. Each 20µl PCR reaction contained:

- 1x Fast SYBR™ Green Master Mix
- 0.2 µM of each primer
- H_2O to a volume of 15 µl
- 5 µl cDNA

Fast SYBR™ Green Master Mix contains SYBR™ Green 1 Dye, AmpliTaq Gold™ Fast DNA Polymerase LD, dNTPs with dUTP/dTTP blend, and optimized buffer components.

Real-time PCR was performed using the 7500 Fast real time PCR system (Applied Biosystems). The PCR programme was 95°C for 20 seconds then 40 cycles of [95°C for 3 seconds, 60°C for 30 seconds]. On completion of the assay the cycle threshold (Ct) for each reaction was calculated using the 7500 Fast system software, before being analysed in Microsoft Excel 2010 and Graphpad Prism 7. Relative expression levels were calculated using the comparative method, giving the ratio (R) of the Ct value of the target gene versus the endogenous gene, *Hprt1*:

$$R = \frac{2^{\wedge} - Ct_{target}}{2^{\wedge} - Ct_{endogenous}}$$

2.4.6. Wholemount kidney immunohistochemistry

14.5 dpc embryos were collected as in section 2.4.1 and kidneys harvested. Kidneys were fixed in 4% PFA/PBS for 1 hour at 4°C then washed 3 x 5 minutes in PBT at room temperature. Kidneys were blocked for 1 hour, room temperature, in milk block, before being incubated with primary antibody – 1:50 TROMA-1 (Rat monoclonal anti-cytokeratin 8, from DSHB) – in milk block, overnight at 4°C. Kidneys were washed 3 x 1 hour in PBT, rocking at room temperature, and then left overnight at 4°C.

The following day kidneys were incubated with 1:200 secondary antibody (Goat anti-Rat IgG, Alexa Fluor 488, Life Technologies) in milk block, for 1 hour, room

temperature. Kidneys were washed 3 x 30 minutes in PBT, rocking at room temperature, before being mounted on a glass slide in Hydromount (National Diagnostics), and gently flattened with a coverslip. Kidneys were imaged using a Zeiss LSM710 upright multiphoton microscope equipped with ZEN software..

2.5. Adult analysis

2.5.1. Classification of adult mice

Mice were classified as male or female at weaning by analysis of the anogenital distance by Alison Gallop, Gemma Atkins and Jackie Harrison (Mary Lyon Centre, MRC Harwell). Males that later appeared feminized (ambiguous anogenital distance, mammary-associated hair pigmentation around the nipples, genital swelling) were culled by cervical dislocation and their reproductive organs examined by myself or one of the animal technicians listed above. Mice with a testis and male reproductive structures present on at least one side continued to be classified as male.

2.5.2. Fertility testing

Male experimental mice were tested at 2 - 4 months of age by mating with C57BL/6 or CD1 wildtype females. Females were checked for vaginal plugs to confirm normal mating behaviour. Plugged females were culled and opened at 14.5 dpc, and the number of embryos recorded. Males that failed to produce any embryos after confirmed mating with 3 females were scored as infertile. Female experimental mice were tested at 3 - 4 months of age by mating with C57BL/6 wildtype males. Females were checked for vaginal plugs and culled at 14.5 dpc to

record the number of embryos. Cohorts that produced no embryos from 6 females were scored as infertile.

2.5.3. Necropsy and histology

Mice were culled by cervical dislocation (Schedule One) under Home Office regulations. External genitalia were examined and photographed. Internal reproductive organs were removed and photographed. Testes were weighed, then immersed in Bouin's fixative at room temperature for 4 hours, cut in half (transverse), then immersed in Bouin's for a further 2 hours. Testes were then washed in 70% ethanol and processed by the MRC Harwell Histology team (Caroline Barker, Adele Austin and Emily File) for paraffin wax sectioning. Ovaries were immersed in Bouin's fixative at room temperature for 6 hours then processed as above. 4 µm sections were stained with haematoxylin and eosin (H&E) by the histology team.

Chapter 3: Characterisation of a genetic modifier of sex determination on chromosome 13

3.1. Introduction

3.1.1. B6.Y^{DOM} sex reversal

There is increasing evidence that DSDs are oligogenic and that genetic background affects sensitivity to genetic disruption during gonad development, both in humans and in mice. One mouse model for this genetic background sensitivity has been studied now for nearly forty years. B6.Y^{DOM} sex reversal is a phenomenon whereby mice carrying the Y chromosome from a *Mus musculus domesticus* (*DOM*) strain, but the X chromosome and autosomes from the inbred strain C57BL/6J (B6/J), are unable to undergo normal testis development. The house mouse, *Mus musculus*, has four well-defined subgroups: *M. m. domesticus*, *musculus*, *castaneus* and *bactrianus* (Silver, 1995). Due to their interesting ancestry the genomes of classical inbred mouse strains (such as C57BL/6J) are a mixture of *Mus musculus* species. Other strains – the so-called wild-derived strains – are considered genetically purer than the classical inbred strains. They descend from mice captured in the wild, then inbred, and are thought to genetically represent single *M. musculus* subgroups (Frazer et al., 2007).

The most commonly studied example of B6.Y^{DOM} sex reversal is B6.Y^{POS}. After crossing POS-A males (a semi-inbred strain descended through the male line to a wild mouse caught in the Val Poschiavo region of Switzerland, hence the name POS, short for *poschiavinus*) to XX B6/J females for three generations, it was noted that there was an abnormal sex ratio in favour of females. Upon investigation, a number of phenotypic females were discovered to be

chromosomally XY, and some male offspring were 'true hermaphrodites', a term that is now out of favour but which describes individuals with both ovarian and testicular tissue. Phenotypic males often had testes smaller than those of B6 males (Eicher et al., 1982). This outcome was replicated by crossing B6 females to different males carrying the *M. m. d. poschiavinus* Y chromosome (Y^{POS}) (Eicher and Washburn, 1983). Examination of B6.XY^{POS} embryos, aged between 14.5 dpc and 16.0 dpc, revealed a gonadal sex reversal phenotype. The majority of XY embryos had two morphologically normal ovaries. The rest had one ovary and a contralateral ovotestis, or bilateral ovotestes. Ovotestes, gonads that contained both ovarian and testicular tissue, were variable in their morphology. No individual had a normal testis (Eicher et al., 1982). After multiple generations of backcrossing to B6, 75% of B6.XY^{POS} mice developed as females with bilateral ovaries and the remaining 25% developed as hermaphrodites. No males with bilateral testes were reported (Eicher and Washburn, 2001).

The Y chromosomes from other *M. m. domesticus* strains were also transferred to the B6 background by consecutive crosses to B6 females, each conferring some degree of disruption to testis determination. B6.Y^{TIR} – whose Y chromosome was derived from males caught in Villa di Tirano, Italy – and B6.Y^{ORB} – from Alpie Orobie, Italy – mice also developed predominantly as females and hermaphrodites with ovarian tissue (Eicher, 1988; Nagamine et al., 1987). Y^{TIR} conveyed a slightly less severe defect than Y^{POS}, since some males with bilateral testes were observed (Nagamine et al., 1987). XY embryos carrying the Y chromosome from AKR/J and SJL (both *M. m. domesticus*-derived Y chromosomes) had small, transient, ovarian-like regions at the poles of the

gonads, but all XY individuals were males with two testes when examined postnatally (Nagamine et al., 1987). This had led to the B6.Y^{AKR} strain, in particular, being a useful tool in the study of sex determination, since it is highly sensitised to further disruptions to testis determination and can be used to study heterozygous sex determination phenotypes when homozygotes are non-viable (Bogani et al., 2009; Bouma et al., 2007; Harris et al., 2018; Warr et al., 2016).

3.1.2. C57BL/6 (B6) sensitivity

It is theorised that B6.Y^{DOM} sex reversal is caused by misregulation of the *M. m. domesticus Sry* allele by autosomal genes derived from *M. m. musculus* (Eicher et al., 1996). The ancestry of inbred mouse strains is complex and the genomes of most laboratory mouse strains (including C57BL/6J) are neither pure *M. m. domesticus* nor *M. m. musculus*, but a mixture of *Mus musculus* subspecies (Frazer et al., 2007). It might be expected that other inbred strains are similarly incompatible with *M. m. domesticus* Y chromosomes, but C57BL/6J appears to be the exception. Y^{POS} has been transferred to multiple other inbred strains without disruption to testis development, namely BALB/cBy, C58/J, DBA/2J (Eicher, 1988) and 129S1/SvImJ (Whitney et al., 2000). Studies using a different cause of XY sex reversal, the hair-pin tail (T^{hp}) deletion on chromosome 16, showed that the C3H/HeJ inbred strain is also resistant to testis development disruptions, whilst C57BL/6J is not (Washburn and Eicher, 1983).

Multiple studies have attempted to explain the reason for B6's sensitivity to sex reversal. Comparison of the transcriptomes of B6 and 129S1/SvImJ embryonic XY gonads at 11.5 dpc revealed significant differences between strains in the

expression of autosomal sexually dimorphic genes (Munger et al., 2009; Munger et al., 2013). Expression levels of pro-ovarian genes, such as *Wnt4* and *Fst*, were higher in B6 than in 129S1. This is likely due to an approximate 5-hour delay in the onset of male-enriched genes, which repress the female pathway. Similar findings were observed when comparing B6 to DBA/2J expression levels in fetal gonads at 12.5 dpc (Correa et al., 2012). Thus B6 gonads are "primed" for testis disruption if the expression or activity of *Sry* and/or *Sox9* is disrupted.

3.1.3. The *POS* gene theory

After many tens of generations of backcrossing to B6, B6.Y^{POS} should be considered a consomic strain. That is, all genetic material comes from B6, apart from the segregated Y^{POS} chromosome. However, there has always been the question of why a genetically homogenous colony has such a variety of phenotypes. The majority of B6.XY^{POS} mice are completely sex-reversed, exhibiting bilateral ovaries, but each generation still produces individuals with sufficient testicular tissue to permit successful breeding. One explanation could be stochasticity: it is simply a matter of chance whether a B6.XY^{POS} gonad will develop some testicular tissue. Stochasticity is a known contributor to sex-reversal phenotypes (Harris et al., 2018). An alternative explanation, termed the *POS* gene theory (Eicher and Washburn, 2001), suggests that the inheritance of a *POS* allele of an autosomal gene is responsible. In the homozygous B6 state, the gene is unable to induce testicular development, but in the *POS*/B6 heterozygous state testicular tissue can develop at least some of the time. Because the B6.Y^{POS} line can only be maintained through breeding of a fertile male or hermaphrodite, an autosomal locus like this would be kept in a forced heterozygous state. The

authors concluded that the *POS* gene theory was false; when the selection pressure was removed using a *M. musculus Sry* transgene, which prevented sex reversal in all XY^{POS} carriers, variability in gonadal phenotype was still seen. However, recent findings suggest that they were premature in their rejection of the *POS* gene idea.

3.1.4. A genetic modifier of sex determination on chromosome 11

A B6.Y^{POS} colony at UCLA was found to harbour a dominant, autosomal modifier of sex reversal on chromosome 11 (Arboleda et al., 2014; Nikolova et al., 2008). The original 60 Mb region contained multiple areas of non-B6 derived SNPs. This region was narrowed down to a single region at position 110 Mb – 113.5 Mb of chromosome 11. This "110 region" provided 80% protection against sex reversal when in the heterozygous state, and 100% protection when in the homozygous state. Without the 110 region, 87.5% of B6.XY^{POS} developed exclusively ovarian tissue. Further analysis using two recombinant alleles identified a 1.62 Mb section of this region capable of conveying the same level of protection (Arboleda et al., 2014). This region encompassed the 5' regulatory elements of *Sox9* and was thought to significantly promote *Sox9* expression, thereby rescuing carriers from sex reversal.

At first, the protective region was thought to derive from the 129S1/SvImJ genome, introgressed into the B6.Y^{POS} genome during previous crosses to the inbred strain (Nikolova et al., 2008; Whitney et al., 2000). However, detailed examination indicated that the SNPs in the 110 region did not resemble SNPs in

129S1/SvImJ. Instead, SNP diversity array analysis was performed with data from 60 wild-derived strains. The authors inferred that the protective region derived from the original POS-A genome (Arboleda et al., 2014). It was predicted that the region had been passed down from the original POS-A male to male offspring over a period of 30 years. Because only 12.5% of XY mice without the 110 region developed any testicular tissue, and no mice developed as males with bilateral testes, the 110 region was selected at each generation due to its positive impact on maleness and fertility, allowing its presence to persist for so long in an otherwise congenic B6.Y^{POS} colony.

3.1.5. B6.Y^{POS} at MRC Harwell

A B6.Y^{POS} colony has been bred at MRC Harwell since the 1990s. Sperm from a B6.Y^{POS} male was frozen and this line was re-derived from the sperm archive in 2011. It has since been maintained by continued backcrossing of B6.Y^{POS} males to unrelated B6/J females. Prior to this present study, 65% of B6.XY^{POS} mice at Harwell developed as females, and 35% of developed as males or hermaphrodites (M. Pope, personal communication).

To investigate if the chromosome 11 protective region was present in the B6.Y^{POS} colony at MRC Harwell, Valerie Arboleda (in the Vilain lab at UCLA) kindly sent three SNP sequencing assays and control DNA samples from the UCLA colony; the B6.Y^{POS} animals at Harwell were sequenced by Madeleine Pope, but none of the individuals tested carried the protective SNPs on chromosome 11, in contrast to the controls. According to Arboleda et al., absence of the 110 protective region should result in the extinction of the line due to XY sex reversal. They

hypothesised that only through co-inheritance of the POS-A regulatory region of *Sox9* with the POS Y chromosome will male sex determination and fertility occur. This begs the questions: how is it that 35% of B6.XY^{POS} mice at Harwell developed as males? And how has the colony been maintained without the 110 region?

3.1.6. Aims

The aim of this chapter is to determine if a genetic modifier, different to the one described by Arboleda et al., is present in the Harwell B6.Y^{POS} genome and is protecting against sex reversal in B6.XY^{POS} males. This will be achieved through next generation sequencing and bioinformatic analysis of the B6.Y^{POS} genome at Harwell, and examination of Harwell's B6.XY^{POS} mice.

3.2. Results

3.2.1. Whole genome analysis of B6.Y^{POS}

DNA from a fertile B6.Y^{POS} stud male in the Harwell colony was extracted and used for whole genome sequencing. The chosen male was one of three brothers from which the current B6.Y^{POS} colony at Harwell is descended. The sequence was analysed by Michelle Simon from the bioinformatics team at MRC Harwell. Details of the analysis performed can be found in section 2.3.6 of the materials and methods. For the purpose of this thesis, we will use the term 'SNP' (single nucleotide polymorphism) for sequence variants rather than 'SNV' (single nucleotide variant) because we think that the majority are polymorphic i.e. variants between inbred strains. The sequence was compared to the reference C57BL/6J genome sequence. 45,396 heterozygous non-B6 SNPs and 14,107 homozygous non-B6 SNPs were identified across the B6.Y^{POS} genome **(Fig 3.1)**.

12,565 SNPs were located on the Y chromosome; as expected given the divergence between Y^{B6} and Y^{POS} chromosomes. 112 heterozygous and 36 homozygous SNPs were located on the X chromosome. Since the X chromosome in this male must have originated from a B6 female from the Harwell colony and not the B6.Y^{POS} father, these SNPs indicate a degree of genetic drift between the reference B6 and Harwell B6 genomes, in the case of homozygous SNPs, or sequencing errors. This is important when considering the SNPs found within the autosomes.

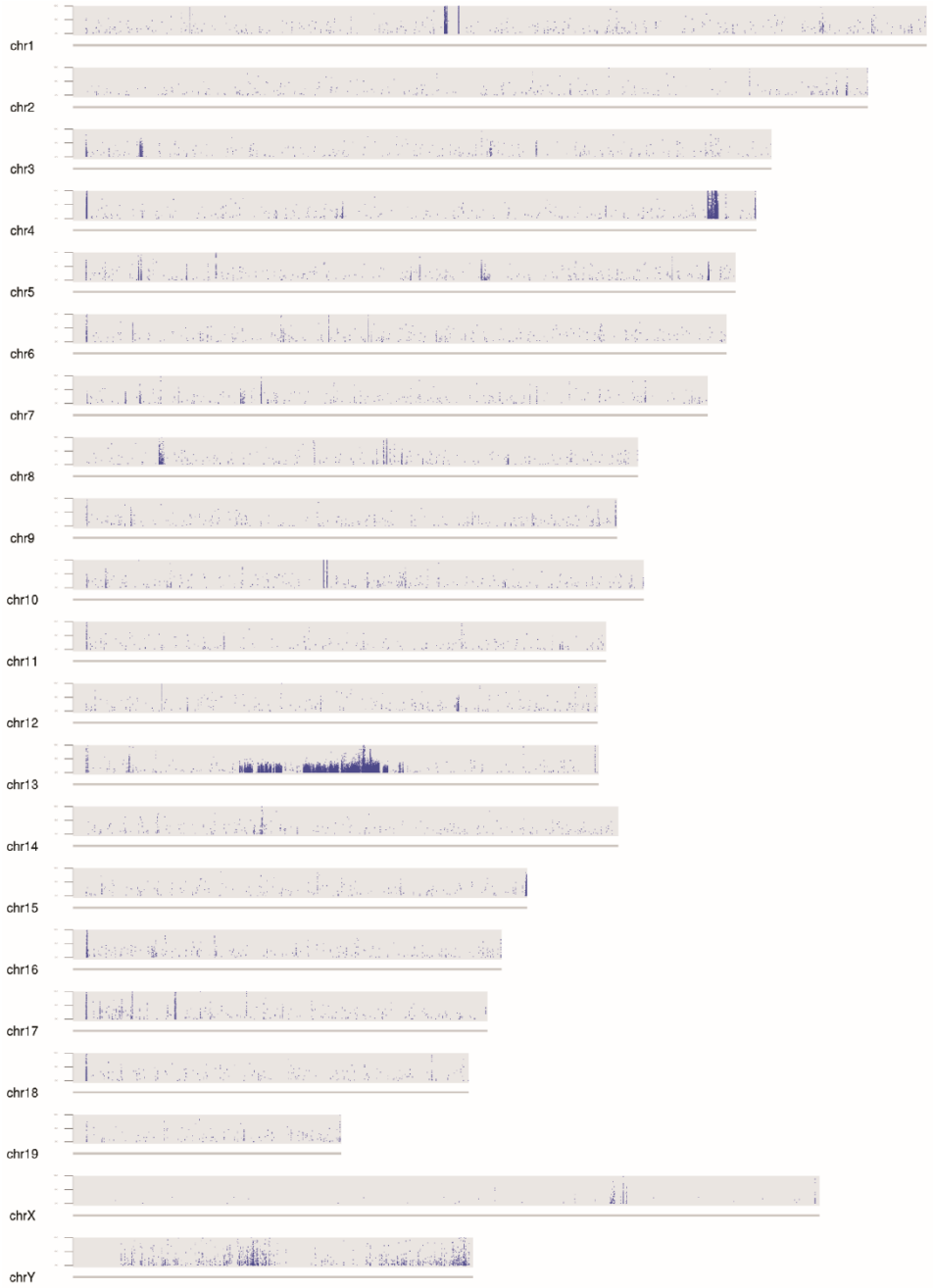


Figure 3.1 Distribution of SNPs across the B6.Y^{POS} genome. Using whole genome sequencing, SNPs were detected in the B6.Y^{POS} genome against the C57BL/6J reference. Various clusters of high density SNPs were found throughout the genome; the largest cluster was found between approximately 38 Mb and 76 Mb of chromosome 13. Figure courtesy of Michelle Simon.

To make the process of validation practical, we sought significant regions of high density SNP heterozygosity, differing between B6 and B6.Y^{POS}, and two such regions were identified: 4,211 (11 of which homozygous) SNPs located between 145.4 Mb and 147.6 Mb of chromosome 4 and 23,523 (190 of which homozygous) SNPs located between 38.0 Mb and 76.0 Mb of chromosome 13. Sanger sequencing of putative SNPs in both regions was carried out using Harwell B6 and Harwell B6.Y^{POS} DNA. Of the six SNPs on chromosome 4 that were selected for validation, two were in regions too repetitive for successful sequencing, two were found to be heterozygous in the Harwell B6 genome, and two did not differ between B6 and B6.Y^{POS} sequences (i.e. were false positives). The reference genome sequence of this region of chromosome 4 contains a high proportion (approximately 25%) of ambiguous base reads (Ns), which can cause poor sequence alignment and increase the probability of errors in the analysis. In contrast, 89% of sequenced SNPs on chromosome 13 were validated as differing between Harwell B6 and B6.Y^{POS} genomes. Full lists of sequenced SNPs can be found in **Tables S1** and **S2** in Appendix 1.

3.2.2. A genetic modifier on chromosome 13 protects against sex reversal in B6.Y^{POS} mice

The B6.Y^{POS} colony (233 XY adult mice) was genotyped for the presence of SNPs on chromosome 13; this genotype was then correlated to the phenotypic sex of XY mice, as determined by their anogenital distance **(Fig 3.2B)**. Significantly more XY mice heterozygous for these SNPs were categorised as male (68%) compared to XY mice homozygous for the B6 allele (33%) (Fisher's exact *p*-value < 0.0001).

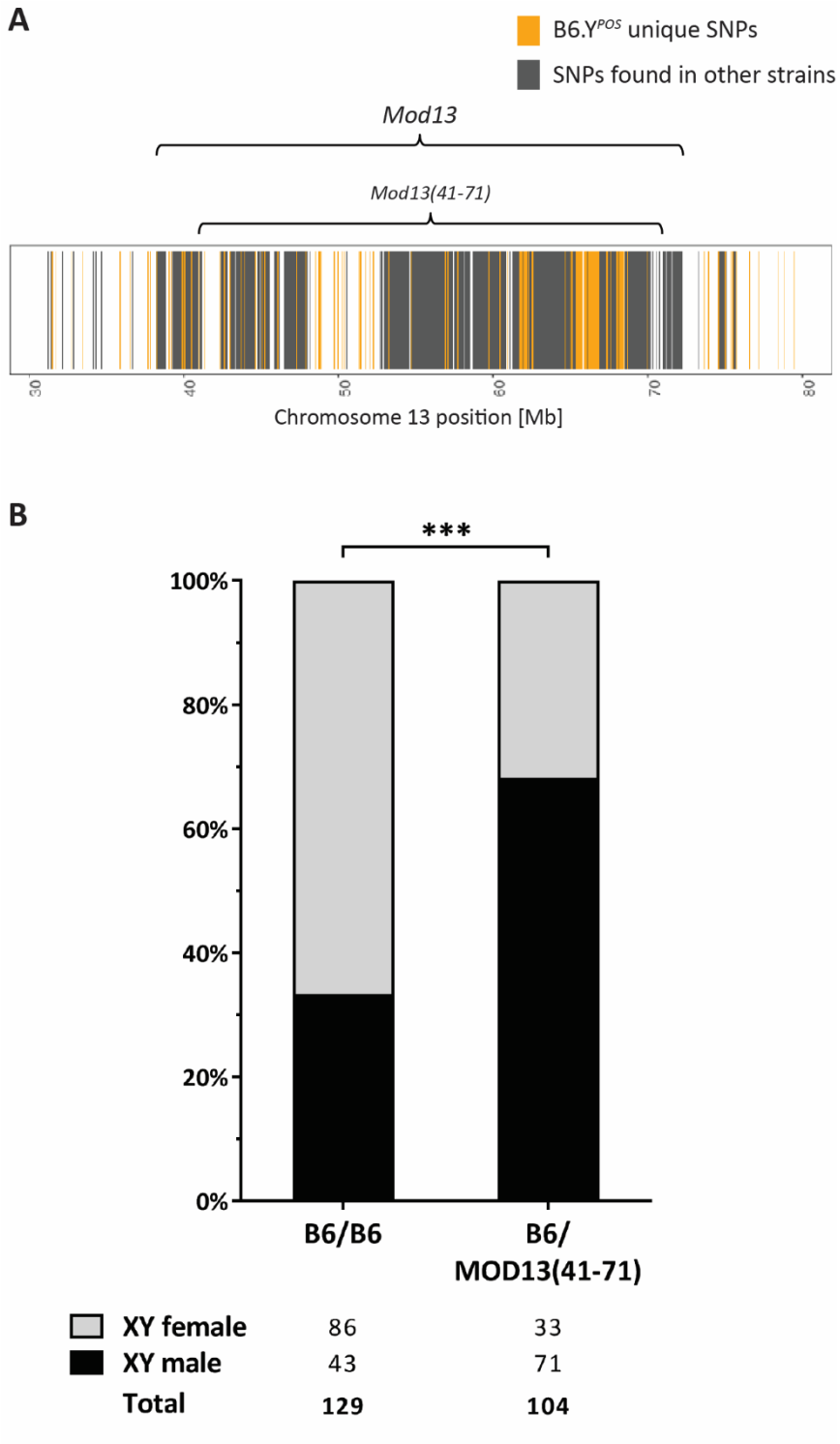


Figure 3.2 The *Mod13* genetic modifier region. A) Density of SNPs found in the *Mod13* region. The colours denote whether the SNPs are unique to the B6.Y^{POS} genome at Harwell compared to 36 other mouse strains (orange), or present in at least one other mouse strain (grey). This will be described in more detail in section 3.2.4. Figure courtesy of Michelle Simon. **B)** Adult phenotypic sex analysis. Ratio of B6.XY^{POS} males and females scored at weaning by anogenital distance, grouped by genotype of chromosome 13 region, *Mod13(41-71)*. A significant difference was found between B6/B6 and B6/mod mice ($p < 0.0001$), using Fisher's exact test.

I conclude that this region of chromosome 13 modifies the sex reversal phenotype of B6.Y^{POS} mice.

This genetic modifier region was named *Mod13*. SNPs were not spread evenly across the approximately 38 Mb of *Mod13*. Instead, regions of high density SNPs were interspersed in segmental fashion with regions containing very few SNPs, illustrated as areas of white in **Fig 3.2A**. Presumably these are regions of B6 genome that have introgressed into *Mod13* by double recombination events over many generations of backcrossing. The boundaries of *Mod13*, excluding these B6 regions, were defined as 38.2 Mb – 41.3 Mb, 42.0 Mb – 48.9 Mb, 52.7 Mb – 72.16 Mb, and 74.5 Mb – 75.7 Mb.

Ten SNPs selected from across the region were initially chosen for genotyping. The genotyping SNPs on either end – rs30098847 at 39.1Mb and rs29535792 at 74.7Mb – were regularly homozygous B6 while all other SNPs were heterozygous, indicating that recombination events were common in these outlying areas of the region. Absence of these SNPs had no effect on phenotypic sex. The remaining 8 SNPs – at 41.0 Mb, 43.6 Mb, 48.1 Mb, 52.3 Mb, 55.4 Mb, 62.8 Mb, 69.5 Mb and 71.1 Mb (details available in **Table 2.4** in Chapter 2) – were used for future genotyping. For simplicity, this allele was named *Mod13(41-71)*. All experiments described in this chapter and the next use individuals carrying *Mod13(41-71)*. Selected individuals were genotyped for additional SNPs, at approximately 1-2 Mb intervals, to regularly check for introgression of B6 by recombination. Recombination events within this region were rare, although some recombinant alleles are described in Chapter 5 of this thesis.

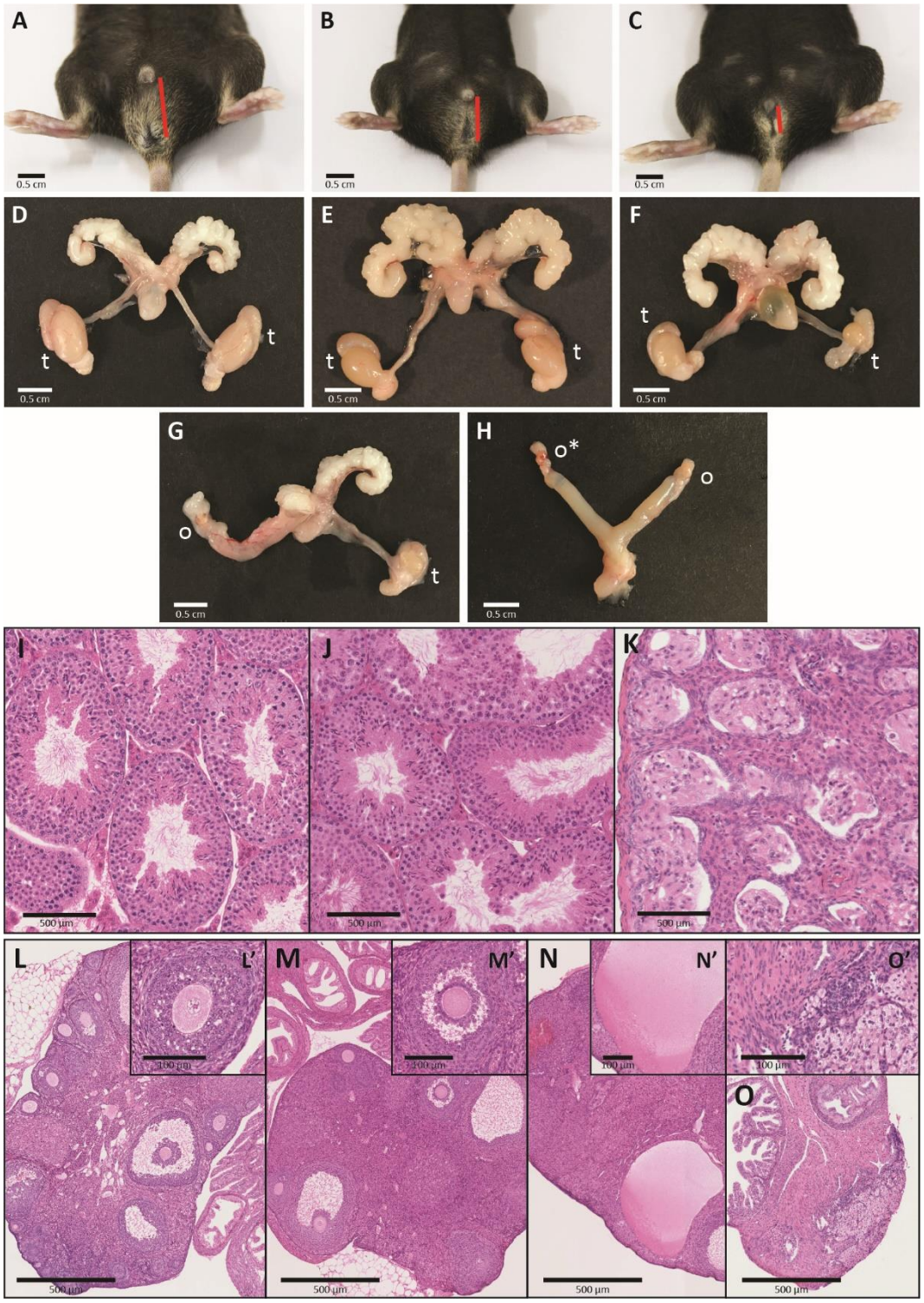


Figure 3.3 Examination of B6.XY^{POS}.*Mod13*(41-71) adult mice. For full figure legend, please see the following page.

Figure 3.3 Examination of B6.XY^{POS}.*Mod13(41-71)* adult mice. Mice were classified at weaning as male or female. Males had a long anogenital distance **(red line, A)** and no pigmentation around the nipples. Females had a short anogenital distance and mammary-associated hair pigmentation **(C)**. Males that later appeared feminized (B) were culled and examined. Mice with a testis and male reproductive structures present on at least one side were still classified as male. Adults were culled for necropsy and histology between 3 and 5 months of age. Images of wild-type B6 animals are included for comparison **(D, I and L)**. Examination of reproductive organs revealed a range of phenotypes: bilateral testes that appeared even and normal **(E)** but smaller than B6.Y^{B6} **(D)**; bilateral testes that were uneven in size **(F);** unilateral testis paired with an ovary **(G)**; and bilateral ovaries **(H)** that were normal in appearance or that contained cysts **(asterisk in H)**. Examination of H&E-stained sections indicated that most testes contained seminiferous tubules with normal morphology and visible spermatids **(J)**. The smallest testes contained only dysgenetic tissue **(K)**. Ovaries from XY females appeared histologically normal **(M)** and contained maturing follicles **(M')** comparable to XX ovarian tissue **(L and L')**, or contained cysts **(N)**. Ovaries from hermaphrodites were typically dysgenetic, containing ovarian-type stroma and no follicles **(O)**. t = testis, o = ovary.

3.2.3. Analysis of adult B6.Y^{POS}.*Mod13* mice

Following the observation that B6.XY^{POS} mice heterozygous for *Mod13* had a higher probability of being scored as male when phenotyped by external appearance, the reproductive anatomy of 17 homozygous B6 (B6/B6) and 20 heterozygous (B6/mod) XY mice, a mix of phenotypic males and females, were examined in more detail. Adult mice (approximately 12 – 16 weeks of age) were culled and their reproductive organs were removed and photographed. Gonads were fixed in Bouin's fixative and processed for histological analysis.

It has been previously reported that B6.XY^{POS} animals exhibit a range of gonadal phenotypes (Eicher et al., 1982) and this full range was seen in both genotypes of mice examined. All animals classified as male at weaning had unilateral or bilateral testes with associated seminal vesicles, epididymides and vasa

deferentia **(Fig 3.3E-G)**. Testes were smaller than those found in control B6 males **(Fig 3.3D)**, but appeared otherwise normal. Bilateral testes were regularly uneven in size **(Fig 3.3F)**. Analysis of H&E-stained sections revealed normal seminiferous tubule morphology **(Fig 3.3J)**. Spermatids were visible in testis sections of both genotypes. Some testes – identified by the presence of male reproductive structures – contained only dysgenetic tissue **(Fig 3.3K).** Some individuals exhibited a unilateral testis and contralateral ovary **(Fig 3.3G)**. In all cases, these mice were classified as male at weaning by an anogenital distance in the masculine range, but these often developed genital swellings or mammary gland-associated pigmentation in adulthood. Many mice with both ovarian and testicular tissue were still capable of impregnating females. All mice with at least one testis and accompanying male reproductive structures were classified as males for this thesis. XY females examined had bilateral ovaries and accompanying uterine horns **(Fig 3.3H)**, although in one case the left gonad was absent. Ovarian morphology varied widely; some ovaries contained multiple follicles at various stages of maturation **(Fig 3.3M)**, comparable to the morphology seen in XX ovaries **(Fig 3.3L)**. Some ovaries contained few follicular structures and regularly had visible cysts **(Fig 3.3N)**, and some (typically those found with a contralateral testis) consisted of mainly ovarian-type stroma. All gonadal phenotypes were seen in both genotypes, but the proportion of males with bilateral testes was higher in B6/mod **(Fig 3.2B)**. These data indicate that *Mod13* promotes the development of testicular tissue, but does not completely rescue sex reversal. To this extent, it behaves like a quantitative trait locus (QTL).

3.2.4. The *Mod13* region shares similarities with other mouse strains

The 110 protective region described by Arboleda et al is thought to derive from the original POS-A founder. An extensive search of the MRC Harwell breeding archive found no evidence of any outcross of B6.Y^{POS} males to other inbred strains, so it is plausible that *Mod13* may have the same POS-A-derived origin. The POS-A strain is long extinct (Eicher et al., 1982) and DNA was unfortunately not available for comparison, despite numerous attempts to identify a source. However, SNP data from 36 laboratory mouse strains were available from the Sanger Institute's Mouse Genomes Project (MGP). A comparison to these common inbred strains identified an interesting phenomenon; the majority of SNPs found in *Mod13* were found to be present in at least one other mouse inbred strain. Only 18% of SNPs were unique to Harwell's B6.Y^{POS} genome. Those SNPs that were unique to *Mod13* clustered between 65 Mb and 67 Mb **(Fig 3.2A)**. Unexpectedly, no one strain stood out as being particularly similar to *Mod13*. The majority of inbred sequences available for analysis contained these SNPs that differ from B6. The extent of this similarity to others strains is shown in **Fig 3.4**.

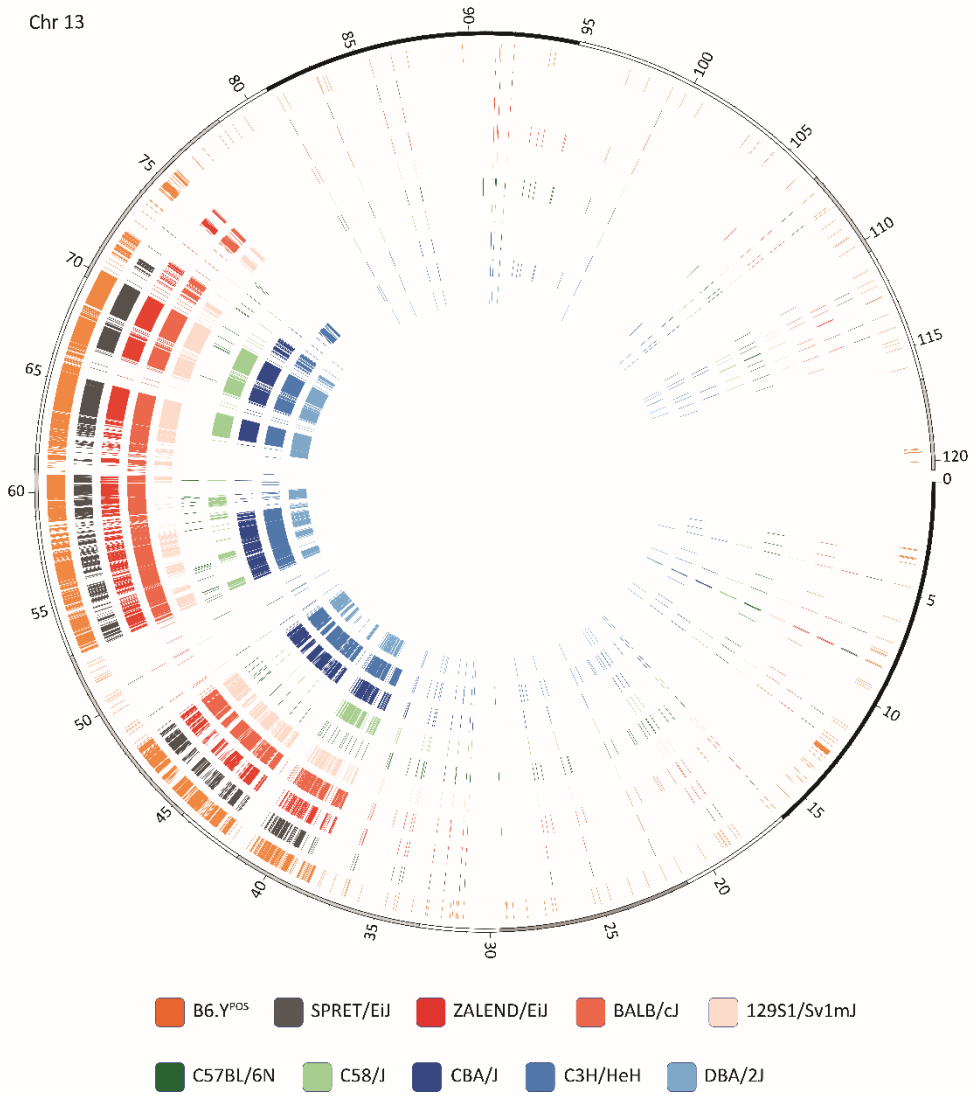


Figure 3.4 Circos diagram of SNPs found in chromosome 13 for 11 different mouse strains. Chromosome 13 is represented as a circle, also termed a circos diagram, with the chromosome positions shown every 5 Mb on the outside. Each track in the circos diagram is an individual mouse genome with the B6.Y^{POS} genome on the outer circumference track. Each line in the circos diagram is a SNP. A high density of SNPs was found between 38 Mb and 76 Mb for the majority of mouse strains. C57BL/6N contains the least number of SNPs in this region and is therefore the least conserved with B6.Y^{POS}.*Mod13*. SNP data from 36 laboratory mouse genomes are available from the Mouse Genomes Project. The 10 strains chosen for this diagram were selected from across the laboratory mouse genealogy, and include those strains known to be resistant to Y^{POS} sex reversal (see section 3.1.2). Figure courtesy of Michelle Simon.

To examine further the similarity to multiple inbred strains revealed by bioinformatic analyses, high-fidelity Sanger sequencing was performed of approximately 12 Kb of *Ptch1* (a gene present at 63.5 Mb on chromosome 13 and known to function in mouse sexual development, see Chapter 5), using DNA from Harwell's B6/J and B6.Y^{POS} colonies. The alignment of these two sequences identified 24 SNPs and five small indels that varied between *Mod13* and the B6 allele across the sequenced area. These variants were compared to variant data from the 36 laboratory mouse strains sequenced as part of the MGP. A list of SNPs in a selection of mouse strains is provided in **Table S3** in Appendix 2. The closest match was BALB/cJ, whose sequence was identical to the *Mod13* sequence bar one indel. This was confirmed by alignment of the BALB/cJ sequence to *Mod13* **(Appendix 2)**. This sequence analysis corroborated the result from bioinformatic analyses.

3.2.5. The role of *Mod13* in male fertility

A large number of spermatogenesis-associated genes are present in *Mod13*. This is of interest, because a role for *Mod13* in promoting spermatogenesis, and therefore male fertility, might explain how this genetic modifier has persisted in the colony despite 33% of XY mice without *Mod13* still developing as phenotypic males.

7 B6/B6 and 11 B6/mod B6.XY^{POS} males were fertility tested by mating to wild-type females. Normal mating behaviour was confirmed by the presence of a vaginal plug. 2 of the 7 B6/B6 males produced no offspring from matings with at least 3 females, compared to 1 of the 11 B6/mod males tested. The 8 B6/B6

males were culled and their testes weighed, sectioned and examined histologically. 4 of the B6/mod males were also culled and their testes weighed and sectioned, including the only mod/B6 male to produce no offspring.

The testis weights of the males examined are listed in **Table 3.1**. All fertile males, of either genotype, had at least one testis weighing over 50 mg (50% of the average testis weight of a young adult male), whereas the infertile males did not. Morphological examination of H&E-stained sections showed that infertile males had some well-formed seminiferous tubules **(Fig 3.5B** and **D)**, within the lumen of which mature spermatids were visible **(Fig 3.5B'** and **D')**. This suggests that a failure of spermatogenesis was not the cause of infertility. Small regions of seminiferous tubule atrophy were visible in the B6.XY^{POS} testes examined **(Fig 3.5A'')**; however, the regions of atrophy appeared larger in infertile males **(Fig 3.5B''** and **D'')**, especially in the area surrounding the rete testis. A possible contributor to infertility may be obstructive lesions caused by testicular malformation and degeneration.

Table 3.1 Correlation between fertility and testis weight

Genotype	Fertile	Testis weight [mg]	
		Testis 1	Testis 2
B6/B6	Y	67	80
B6/B6	Y	17	62
B6/B6	Y	56	30
B6/B6	Y	13	59
B6/B6	N	<10	44
B6/B6	Y	47	70
B6/B6	N	<10	49
B6/mod	Y	86	84
B6/mod	N	34	24
B6/mod	Y	62	26
B6/mod	Y	53	57

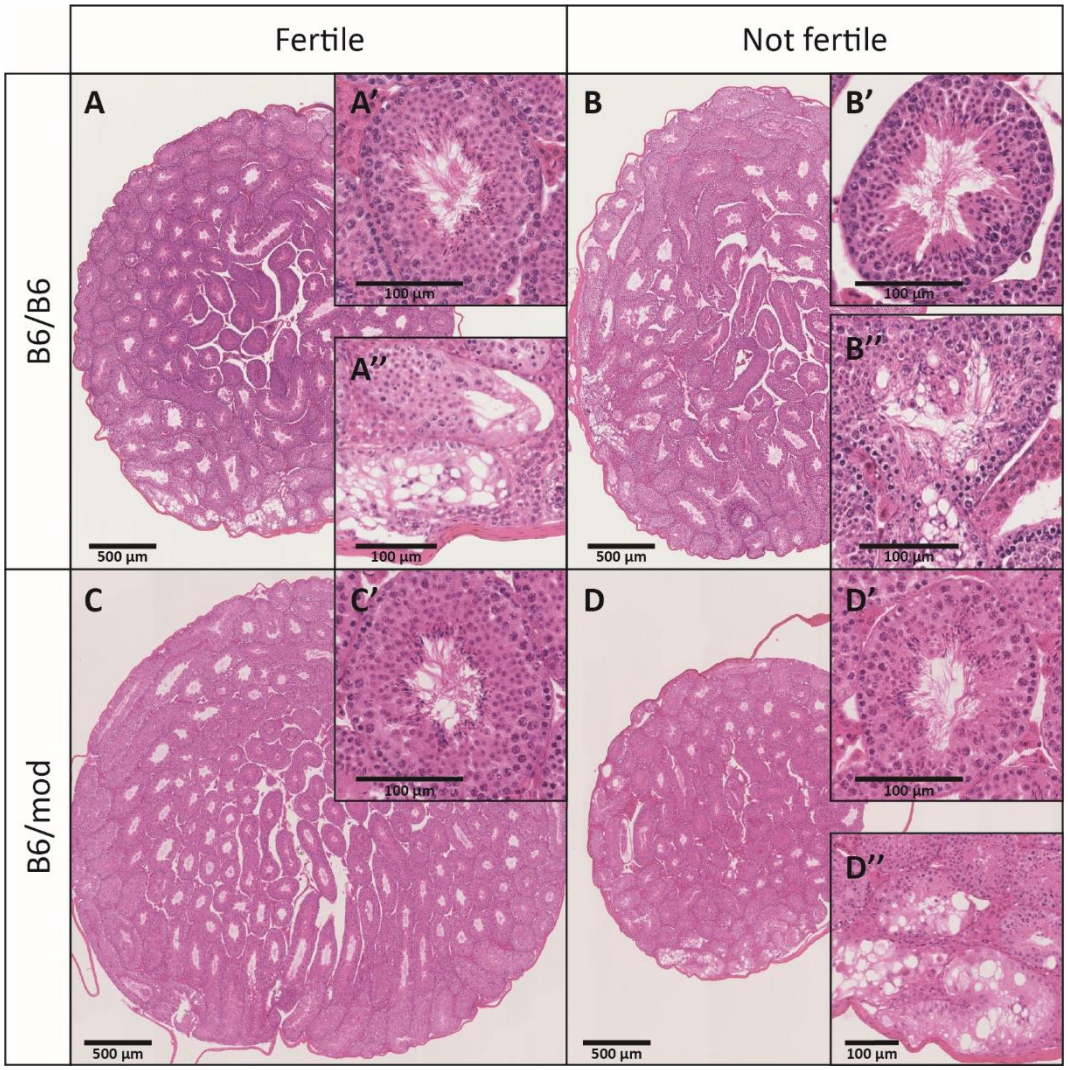


Figure 3.5 The effect of *Mod13* on male fertility. B6.XY^{POS} male mice were culled for necropsy and histology at 4-5 months of age, after fertility testing. H&E stained sections of testes from fertile **(A and C)** and infertile **(B and D)** males reveal normal seminiferous tubules, containing spermatids **(A'-D')**. Areas of tubule degeneration are visible in some fertile males **(A'')** and to a greater degree in infertile males **(B'' and D'')**.

The only differences observed between fertile and infertile males were reduced testis weights and larger degrees of tubular atrophy seen in the infertile males. There is, therefore, currently no evidence to suggest that *Mod13* impacts male fertility, independent of its effect on promoting testicular development. However, insufficient data is available to rule out a role of *Mod13* in spermatogenesis or secondary male function, and this possibility will be discussed in depth in chapter 7.

3.3. Discussion

3.3.1. Summary

A novel genetic modifier of B6.Y^{POS} sex reversal has been identified in the colony at MRC Harwell. The 38 Mb region of chromosome 13, named *Mod13*, contains 23,523 SNPs of non-B6 origin. In a heterozygous state it provides partial protection against sex reversal by promoting the development of testicular tissue, leading to the development of more fertile males. This phenomenon is strikingly similar to the POS gene theory suggested by Eicher and Washburn in 2001, who predicted that an autosomal gene in a non-B6 state that was required for testicular development, would be kept in a forced heterozygous state by virtue of the breeding scheme used in B6.Y^{POS} colonies.

Multiple B6.Y^{POS} colonies are maintained at different locations around the world. The discovery of *Mod13* and its contrast to the 110 region at UCLA shows that each B6.Y^{POS} colony should be considered as a separate sub-strain with a distinct genetic background. Even independent of the protective regions, the B6.Y^{POS} mice at Harwell appear less prone to sex reversal than the mice at UCLA. 33% of

unprotected (B6/B6) XY mice at Harwell are male, compared to 12.5% at UCLA Arboleda et al. (2014). This suggests that the *Sry*^{POS} alleles may differ between locations, with the Harwell *Sry*^{POS} having perhaps evolved to be more compatible with the B6 genome, "softening" the sex reversal phenotype.

3.3.2. The origin of *Mod13*

The origin of *Mod13* remains uncertain due to the lack of POS-A genomic information available for comparison. One theory for the origin is that a shared wild-mouse ancestor carried the *Mod13*-like variants on chromosome 13. The C57BL/6 branch of the mouse family tree then lost these SNPs by introgression with another unknown strain, after 1921 when the C57 group of strains separated from the C58 group, thus explaining why the C57BL/6 sequences are unusual. This idea is supported by the fact that wild-derived strains such as SPRET/EiJ have *Mod13*-like SNPs. It is therefore possible that *Mod13* does derive from POS-A, like the 110 region found in UCLA.

However, given significant sequence similarities to inbred strains with no reported shared ancestry with POS-A, it is worth considering an alternative explanation, suggesting that *Mod13* does not derive from the original POS-A male. Instead, it could be proposed that it was introduced into the B6.Y^{POS} genome at Harwell by outcrossing to another inbred strain. There are no records in the Harwell archive of an outcross to an inbred strain other than B6/J. However, it is conceivable that at some point during its nearly thirty year history at Harwell, the B6.Y^{POS} colony was struggling to produce enough fertile males to continue the colony. This is in accordance with Arboleda et al. stating that

B6.Y^{POS} mice require some genetic protection for the development of males. In the case of a failing colony, a B6.Y^{POS} hermaphrodite may have been outcrossed to a female of a different strain for a single generation to rescue the colony. *Mod13* could have been introduced into the genome and since been carried forward with each generation through the selection of fertile males. Given the similarity between *Mod13* and the equivalent region in BALB/cJ, and the availability of this strain of mice at MRC Harwell, BALB/cJ is the best candidate for the origin of *Mod13*.

The main point not answered by the BALB/cJ-origin theory is the presence of SNPs unique to *Mod13* clustered between 65 Mb and 67 Mb. This sub-region is unlike any of the available strains, including BALB/cJ. The density of *Mod13*-unique SNPs in this area, compared to other areas, suggests this is not a consequence of randomly accumulated mutations i.e. genetic drift, but a result of inheritance from a strain whose genome has not been sequenced. The most obvious candidate for the origin of this sub-region is POS-A. POS-A was a semi-inbred strain generated from crossing a *M. domesticus poschiavinus* male to a random-bred Swiss NMRI female, followed by sibling matings (Eicher et al., 1982). The sequenced mouse strain most closely related is likely ZALENDE/EiJ, a wild-derived *M. domesticus* strain descended from mice caught in the Val Poschiavo region of Switzerland, the same region as the founding POS-A male. The Y chromosome from ZALENDE/EiJ males have been shown to have the same effect on testis determination when bred onto B6 as the Y chromosome from POS-A (Eicher, 1988). The *Mod13* unique sub-region does not resemble ZALENDE/EiJ sequence. It could derive from the NMRI female, however, the

NMRI mice are now outbred and at least 40 years removed from the founding POS-A female, so comparison to this line is unlikely to answer the origin of *Mod13*.

It appears that *Mod13* does not derive entirely from one strain, but is made up of haplotype blocks that originate from likely three different mouse strains: C57BL/6J, BALB/cJ and POS-A being the best candidates. Over its long history, B6.Y^{POS} has accumulated the present *Mod13* allele from multiple sources and because of its promotion of male development and fertility, *Mod13* has been selected at each generation.

3.3.3. The persistence of *Mod13*

Assuming *Mod13* has not been acquired recently (records since the colony's re-derivation in 2011 show only strict matings to B6/J and no other strains) the continued existence of *Mod13* poses a question. While *Mod13* promotes testis development, its presence is not essential for the development of fertile B6.XY^{POS} males at Harwell. 33% of homozygous B6 XY mice are male and the majority of these males are fertile. This is in contrast to the protective region on chromosome 11 reported by Arboleda et al. (2014), which is thought to be close to necessary for development of fertile males in B6.XY^{POS} mice at UCLA.

It remains possible that a second unknown modifier region exists within the colony that enhances the effect of *Mod13*, or acts independently, and is responsible for the development of some fertile males in the absence of *Mod13*. The most fertile males may be heterozygous for both regions; heterozygosity for one but not the other region might promote male development, and B6

homozygosity for both leads to female development. However, the existence of a second region is unlikely. Firstly, the requirement for XY males to co-inherit two unlinked regions would reduce the availability of fertile males and is less probable than the selection of a single modifier locus, especially in a medium-sized colony, over multiple generations. Secondly, no other regions were identified by whole genome analysis. A 2 Mb cluster of SNPs was identified on chromosome 4, but is likely an artefact of a difficult-to-sequence area of the genome. The SNPs chosen for sequencing were not validated. Of course, I cannot exclude a very small region containing just a handful of SNPs, which would be very difficult to distinguish from the general drift-based divergence or sequencing errors seen on every chromosome, including the X. Additionally, the ratio of males to females has remained remarkably constant across the three years this region has been studied. Since the discovery of *Mod13*, only heterozygous males have been selected for breeding. This would be expected to remove at least some of the selection pressure on maintaining a hypothetical second region. Pedigree mapping did not identify any variation between branches of the colony that would suggest the loss of a second region in some individuals.

The most likely reason for *Mod13*'s persistence in the colony prior to its discovery, despite the fact that its existence is not necessary for maleness, is a combination of random chance and the influence of *Mod13* on the proportion of fertile males in the colony. While it is possible for *Mod13* to become extinct, a B6/B6 fertile male who cannot pass on *Mod13* is less likely to sire future fertile males than a B6/mod male. Therefore, a pedigree of heterozygotes would likely

be more productive than a pedigree of B6 homozygotes. In fact, a bottleneck resulting from an unproductive colony is common; the current colony descends from a single male only one generation removed from the male whose genome was sequenced, because other branches of the colony died out.

3.3.4. B6 sensitivity to gonadal sex reversal

An interesting finding from the discovery of *Mod13* is that C57BL/6 (B6) strains are the outliers in their allele of this region. C57BL/6 strains are also unique in their sensitivity to sex reversal. Most studies on B6 sensitivity have used C57BL/6J (Eicher et al., 1996; Munger et al., 2009; Munger et al., 2013); however, there is evidence that C57BL/6N also has a greater propensity to sex reverse than other strains (Johnen et al., 2013). BALB/cBy, C58/J, DBA/2J, C3H/HeJ and 129S1/SvImJ mice are less sensitive to XY sex reversal than C57BL/6 strains. Strains whose genomic sequence is available, and who are closely related to these "protected" strains, show resemblance to *Mod13*. Even the C58/J strain, the closest relative of the C57 branch, has *Mod13*-like SNPs from 38 Mb – 41 Mb and 63 Mb – 70.5 Mb **(Fig 3.4)** and is resistant to Y^{POS} sex reversal (Eicher, 1988). It is therefore possible that *Mod13* identifies a genomic region that contributes to the phenomenon of B6 sensitivity to sex reversal when in the homozygous B6 state.

Studies have shown that multiple loci combine to confer B6's predisposition to ovarian development (Eicher et al., 1996; Munger et al., 2009; Nikolova et al., 2008). Eicher et al. (1996) used genome-wide linkage analysis in C57BL/J-DBA/2J mixed Y^{POS} embryos to correlate ovarian tissue development to regions of B6

homozygosity. They identified regions on chromosomes 2, 4 and 5 that likely contributed to B6.Y^{POS} sex reversal. Another experiment identified regions of the genome associated with differential gene expression in C57BL/6J and 129S1 gonads, but found no association between higher or lower gene expression in B6 gonads and an area of chromosome 13 (Munger et al., 2009). Nikolova et al. (2008) and later Arboleda et al. (2014) showed that the *Sox9* locus on chromosome 11 needs to be in a heterozygous B6/*POS* state to allow for testis development.

Munger et al. (2009) showed that the entire transcriptional landscape is different in B6 gonads at 11.5 dpc, in general promoting the expression of "female" genes, thus making B6 vulnerable to further disruption to testis development. While none of the previous studies have linked chromosome 13 to B6 sensitivity, the results described in this chapter show that homozygosity for the B6 allele of a 38 Mb region of chromosome 13 is associated with a greater chance of ovarian development in B6.XY^{POS} mice. To determine if the B6 allele of *Mod13* is responsible (at least in part) for the sensitivity of B6 to sex reversal, B6.Y^{POS} x BALB/cJ F2s could be generated and examined, looking for a correlation between carrying the BALB/cJ allele of parts of chromosome 13 and XY male development.

Chapter 4: The role of *Mod13* in primary sex determination

4.1. Introduction

4.1.1. Summary so far

A genetic modifier has been identified in the B6.Y^{POS} colony at MRC Harwell. The region of non-B6 SNPs spans 38 Mb of chromosome 13 and provides partial protection against XY gonadal sex reversal in this strain. The presence of *Mod13* increases the probability of testicular, and thus male, development from 33% to 68%. The remainder of this thesis focuses on understanding the mechanism by which *Mod13* promotes testis development. If *Mod13* influences the expression of genes during initial commitment to the testicular fate, examination of its effects, and the genes contained within the 38 Mb region, represents an opportunity to identify a novel sex determining gene.

4.1.2. *Sry* underlies B6.Y^{POS} sex reversal

The B6.Y^{POS} phenotype is characterised by gonadal sex reversal; failure to activate the male pathway leads to the development of ovaries and ovotestes in place of testes (Eicher et al., 1982). As described in section 3.1.2, the B6 background is sensitised to XY sex reversal due to delayed pro-testicular gene expression and subsequently higher expression of pro-ovarian genes at 11.5 dpc (Munger et al., 2013). Despite being capable of initiating testis development on its native genetic background, the Y chromosome from *M. d. poschiavinus* is then unable to interact appropriately with B6 autosomal loci in order to overcome the female pathway, pushing the gonad towards an ovarian fate.

The molecular events underpinning this incompatibility are unclear. Given its central role in testis determination and being the main genetic difference

between B6.Y^{B6} and B6.Y^{POS}, a difference in *Sry* is the obvious answer. An SRY isoform that cannot bind to the *Sox9* promoter efficiently cannot activate the male pathway. *M. domesticus* strains do indeed have different SRY protein isoforms when compared to *M. musculus* strains. Sequencing of many *Mus Sry* alleles shows that the HMG domain is well conserved between sub-species, but the C-terminal region is quite variable (Albrecht and Eicher, 1997; Coward et al., 1994). An earlier STOP codon in the *Sry*^{DOM} CAG trinucleotide repeat region leads to a greatly truncated glutamine repeat (polyQ) domain. Deletion of the sequences encoding the polyQ domain in mouse SRY prevented a transgene from inducing testis determination in XX embryos (Bowles et al., 1999); this region is required both to prevent proteasomal degradation and to bind to the TESCO enhancer element to transactivate *Sox9* (Zhao et al., 2014a). *In vitro* work by Zhao et al. (2014a) showed that the *M. domesticus* polyQ domain was capable of stabilising SRY and transactivating TESCO; however, its transactivation activity was weaker than the *M. musculus* polyQ domain. The number of CAG repeats varies slightly between *M. domesticus* strains (Coward et al., 1994). However, this variation between strains does not correlate with the severity of B6.Y^{DOM} sex reversal (Albrecht and Eicher, 1997; Carlisle et al., 1996), indicating that these peptide sequence differences are not responsible for the incompatibility.

Instead, reduced expression of *Sry* at the critical timepoint of testis determination has been reported in B6.XY^{POS} gonads and proposed as an explanation of B6.Y^{POS} sex reversal (Nagamine et al., 1999). Bullejos and Koopman (2005) reported a significant delay in *Sry* expression in B6.XY^{POS} gonads; when analysed by WMISH on a limited number of samples, the onset of

Sry^{POS} in B6 XY gonads was approximately 4.8 hours later than *Sry*^{B6}, and the peak of *Sry*^{POS} expression was reached 9.6 hours after the normal peak in B6. A matching delay in *Sox9* expression was also observed. No *Sry* delays were detected in B6.Y^{AKR} or B6.Y^{FVB} gonads, two *M. domesticus* Y chromosomes that cause only mild testicular disruption on B6. One explanation for this impact on *Sry*^{POS} expression is that B6 gene products responsible for activating *Sry* at the correct time are less able to interact with the *Sry*^{POS} regulatory regions. The authors reported that *Sry* expression peaks unusually early in B6 compared to other strains and that its window of expression is shorter. In the case of B6.Y^{POS}, the later onset of *Sry* misses the 6 hour window of opportunity reported by Hiramatsu et al. (2009), meaning the testis-determining pathway is not upregulated in time to successfully inhibit the ovarian pathway.

Albrecht et al. (2003) produced an *Sry*^{POS} transgene driven by *M. musculus*-derived regulatory regions. The transgene was capable of inducing testicular development in XX embryos, but despite the normal transcriptional regulation and expression levels 5 times higher than in XY^{B6}, a delay in cord development was observed. This suggests that even very high levels of *Sry*^{POS} expression cannot completely compensate for the reduced activity of the SRY^{POS} protein isoform. B6.Y^{POS} sex reversal is therefore likely a result of delayed *Sry* expression combined with an inefficient SRY protein. This is why *Sry*^{POS} is known as a 'weak' *Sry* allele. It suggests that no amount of upregulation of *Sry*^{POS} by a genetic modifier would be capable of completely rescuing B6.Y^{POS} gonadal sex reversal.

4.1.3. Hypotheses for the role of *Mod13*

There are three points in the sex determination genetic pathway where *Mod13* could be acting to promote testicular development. 1) Upregulation of *Sry*: the timely expression of *Sry* requires a network of signals (see section 1.5) and *Mod13* variants in any genes in this network could promote *Sry* expression, thereby directly rescuing the delayed expression of Sry^{POS} thought to be responsible for the B6.Y^{POS} phenotype. 2) Upregulation of the expression or activity of *Sox9* or other members of the pro-testis pathway, independent of effects on *Sry* expression: the 110 region found in the UCLA colony directly impacts *Sox9* expression, thereby compensating for the weak SRY^{POS} isoform and delayed *Sry* expression to induce testis development (Arboleda et al., 2014). Such an effect might also include upregulation of paracrine signals that activate *Sox9* in SRY-negative cells (Wilhelm et al., 2005). 3) Downregulation of the pro-ovarian pathway: the sensitivity of B6 to perturbations to testis development has been attributed to higher levels of pro-ovarian genes such as *Wnt4* and *Fst* (Munger et al., 2009; Munger et al., 2013). By reducing levels of these genes to 'normal', non-B6 levels, XY embryos would be less predisposed to sex reversal.

Mod13 encompasses a large region, spanning nearly a third of chromosome 13. Such an extensive area of non-B6 genome has been maintained despite tens of generations of crosses to B6, which should gradually replace the unnecessary non-B6 sequences with B6 sequences. The persistence of this large non-B6 region suggests the majority of the region is required for the full rescue effect. *Mod13* likely behaves as a quantitative trait locus (QTL), with multiple variants across the region making small contributions to promote testis development,

which in combination accounts for the partial protective effect. *Mod13* could therefore be acting on multiple points in the sex determination pathway, simultaneously upregulating *Sry* and/or *Sox9* and downregulating pro-ovarian gene expression.

4.1.4. Aims

The aim of this chapter is to investigate how *Mod13* affects gene expression and testis development at the early stages of sex determination. This will be achieved by:

- Analysis of B6.XY^{POS}.*Mod13* gonads at 14.5 dpc.
- Analysis of expression of *Sry* and *Sox9* in B6.XY^{POS}.*Mod13* gonads at 11.5 dpc.
- Analysis of gene expression in B6.XX.*Mod13* embryonic gonads.

4.2. Results

4.2.1. Examination of embryonic B6.XY^{POS}.*Mod13* gonads

It is well established that B6.Y^{POS} sex reversal occurs at the early stage of primary sex determination, with a failure to activate the testicular pathway (Bullejos and Koopman, 2005; Eicher et al., 1982). To confirm that *Mod13* rescues sex reversal by promoting testicular development in the early gonad, individual XY gonads from B6.Y^{POS}.*Mod13(41-71)* heterozygous intercross litters were examined at 14.5 dpc. This cross generates fetuses that are either heterozygous or homozygous for *Mod13* or are homozygous B6, allowing an assessment of the impact of *Mod13* zygosity on testis determination. This stage is ideal for the

identification of ovotestes; ovarian tissue present in the gonad at this age generally regresses and is not visible at later embryonic and postnatal stages (Eicher and Washburn, 1986). Gonads were categorised into one of five groups based on overall morphology and an estimation of the proportion of testicular tissue present. Ovaries (O) were gonads with completely normal ovarian shape and morphology, comparable to XX ovaries of the same stage **(Fig 4.1A)**. Ovotestes (Ot) were predominantly ovarian tissue, with a small region of enlarged testicular tissue in the centre **(Fig 4.1B)**. Testis cords, if present, were poorly structured. A coelomic vessel was often visible, but was disorganised and only extended over a small area **(Fig 4.1B arrow)**. Gonads with around 40% - 60% testicular tissue were classified as ovotestes (ot). The testicular structures were more organised, but still with significant areas of ovarian tissue at one or both poles of the gonad **(Fig 4.1C)**. Gonads with more than 60% testicular tissue were classified as ovoTestes (oT). Tissue with normal testicular morphology made up the majority of the gonad, but small regions of ovarian tissue were visible at one or both poles **(Fig 4.1D)**. The final category was testis (T), where testis cords extended all the way to both gonad poles **(Fig 4.1E)**; however, no testes were found in any of the B6.XY^{POS} embryos examined.

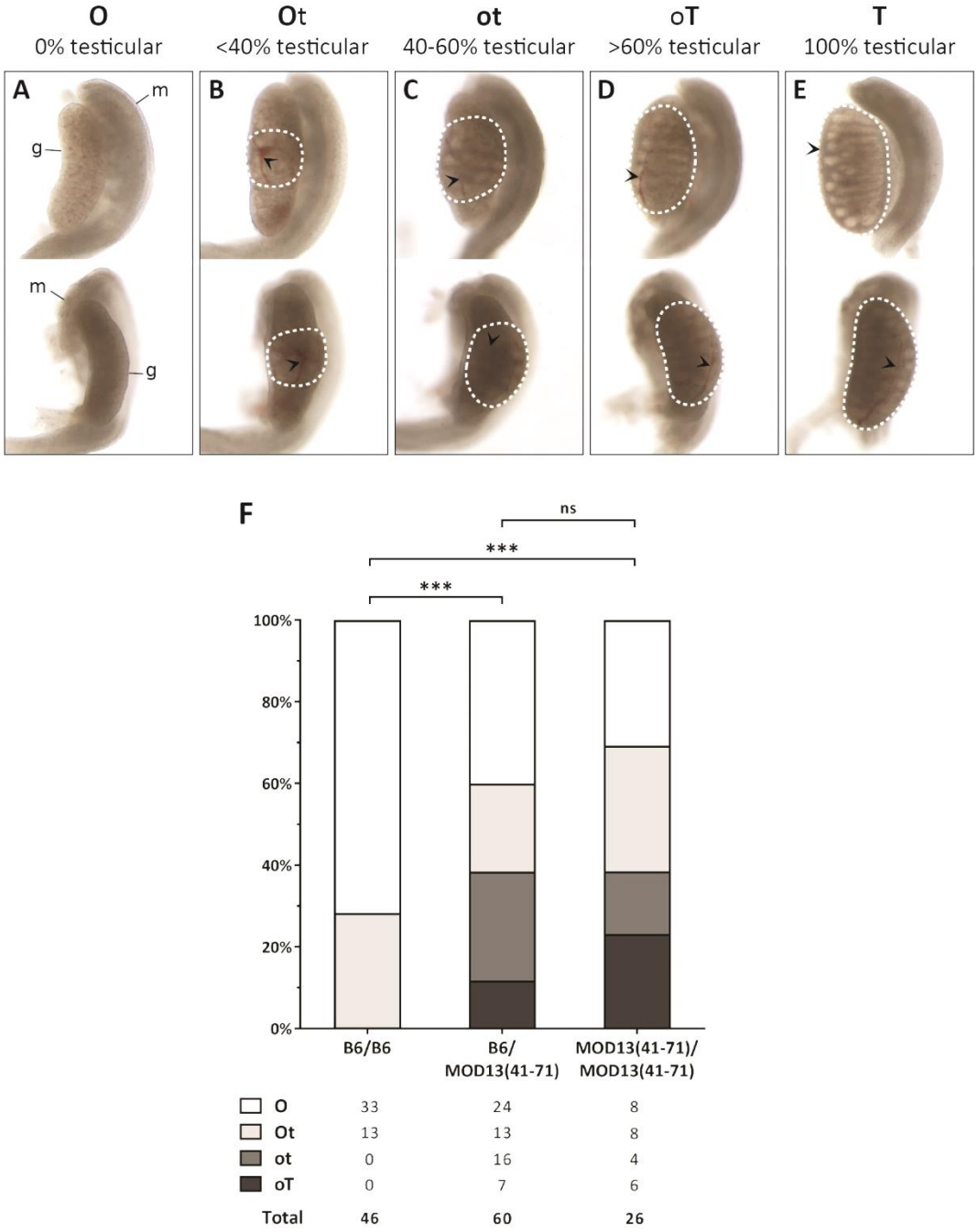


Figure 4.1 Examination of embryonic B6.XY^{POS}.*Mod13* gonads. Individual gonads were classified according to the estimated proportion of testicular tissue at 14.5 dpc, based on gonad shape and morphology **(A-E)**. Upper image shows a ventral view, lower image shows a medial view of the same gonad. Imaged on a Leica MZ16F with attached camera. Classifications were performed at the time of dissection, blind to genotype. White dotted line encircles testicular tissue. Arrow indicates the coelomic vessel. g = gonad, m = mesonephros. **(F)** Proportion of each gonad type by genotype of *Mod13(41-71)*. B6/B6 significantly differs from B6/mod and mod/mod (Fisher's exact *p*-value <0.0001 for both). B6/mod and mod/mod ratios are not significantly different (Fisher's exact *p*-value = 0.322). No testes (T) were observed in any B6.XY^{POS} embryos examined.

After classification, experimental embryos were genotyped for the presence of 8 SNPs across the *Mod13(41-71)* region, as detailed in the previous chapter. Gonads from embryos with recombinant alleles were not included in this analysis. Heterozygosity for *Mod13* was associated with a significant increase in testicular tissue at 14.5 dpc in B6.XY^{POS} embryos **(Fig 4.1F)** (Fisher's exact *p*-value < 0.0001). Only two of the five groups of gonads were represented in homozygous B6 (B6/B6) XY embryos: 72% of gonads were ovaries and the remaining 28% were Ovotestes. These numbers align with the adult data described in chapter 3, which show 33% of B6/B6 adults were males with testicular tissue. This corresponds with the statement from Eicher and Washburn (2001) that any amount of testicular tissue, even if unilateral, is enough to masculinise the reproductive tract and external genitalia of B6.XY^{POS} mice. In contrast to B6/B6 gonads, 40% of heterozygous (B6/mod) XY gonads were categorised as ovaries and the remaining 60% were categorised as a type of ovotestis. 12% of B6/mod gonads examined had over 60% testicular tissue. Homozygosity for *Mod13* (mod/mod) did not cause a significant increase in the proportion of ovotestes observed compared to heterozygotes (Fisher's exact *p*-value = 0.322); however, the overall proportion of gonads with testicular tissue was higher, at 69%. In addition, the proportion of ovoTestes found with the largest amount of testicular tissue was elevated at 23% of mod/mod gonads, compared to 12% of B6/mod gonads.

There was a noticeable degree of variability within individual embryos, regardless of genotype. 23 of the XY embryos examined had an ovary on one side and an ovotestis with up to 60% testicular tissue on the other side. Some

ovoTestes were paired with Ovotestes, despite being genetically identical. Interestingly the left-hand gonad was generally biased towards testicular development: 29 embryos had gonads with approximately equal amounts of testicular tissue, 29 embryos had more testicular tissue on the left-hand side, and only 8 embryos had more testicular tissue on the right-hand side. This same observation was made by Nagamine et al. (1987). This phenotypic variability observed within embryos is an indication of a degree of stochasticity in the earliest events of testis determination.

4.2.2. Gene expression analyses in B6.XY^{POS}.*Mod13* gonads

The same gonads were analysed by WMISH for sex-specific marker gene expression. *Sox9* was used as a marker of testicular Sertoli cells **(Fig 4.2B)**. WMISH analysis showed that regions of *Sox9* expression extend much further from the centre of the gonad in B6/mod and mod/mod ovotestes, compared to B6/B6 ovotestes. It is also noteworthy that many of the *Sox9*-expressing cells were organised into regular testis cords in 2/5th of the B6/mod gonads shown **(Fig 4.2B'')** and 3/5th of the mod/mod gonads shown **(Fig 4.2B''')**. The *Sox9*-expressing cells in the B6/B6 ovotestes were not organised into regular structures **(Fig 4.2B')**, indicating that these gonads, while expressing pro-testis genes, would be less likely to develop into functional testicular tissue in the post-natal mouse. *Stra8* was used as a marker of germ cells entering meiosis, normally only present in ovarian tissue at this stage **(Fig 4.2D)**. *Stra8* expression confirmed the greater degree of ovarian tissue in B6/B6 gonads, compared to B6/mod and

mod/mod. *Insl3* is expressed in Leydig cells of the testis. Analysis of *Insl3* expression in B6.XY^{POS}.*Mod13* gonads indicated that *Mod13* promotes the differentiation of testicular tissue, including the steroidogenic cell lineage into Leydig cells **(Fig 4.2F)**. As observed during initial classification, none of the gonads analysed contained 100% testicular tissue, and none had cord organisation or gene expression pattern comparable to B6.Y^{B6} testes.

Figure 4.2 Gene expression analyses in B6.XY^{POS}.*Mod13(41-71)* gonads. WMISH analysis was performed on B6.XY^{B6}, B6 XX and B6.XY^{POS} gonads at 14.5 dpc for expression of *Sox9* **(A-B)**, *Stra8* **(C-D)** and *Insl3* **(E-F)**. B6.XY^{POS} B6/mod and mod/mod gonads contained larger regions of testicular tissue (indicated by *Sox9* and *Insl3* expression) and smaller regions of ovarian tissue (indicated by *Stra8* expression) than B6/B6 gonads. **(B'-B''')** Close-up of *Sox9*-expressing cell organisation. No gonads contained 100% testicular gene expression. Scale bars: white = 500 µm, black = 200 µm.

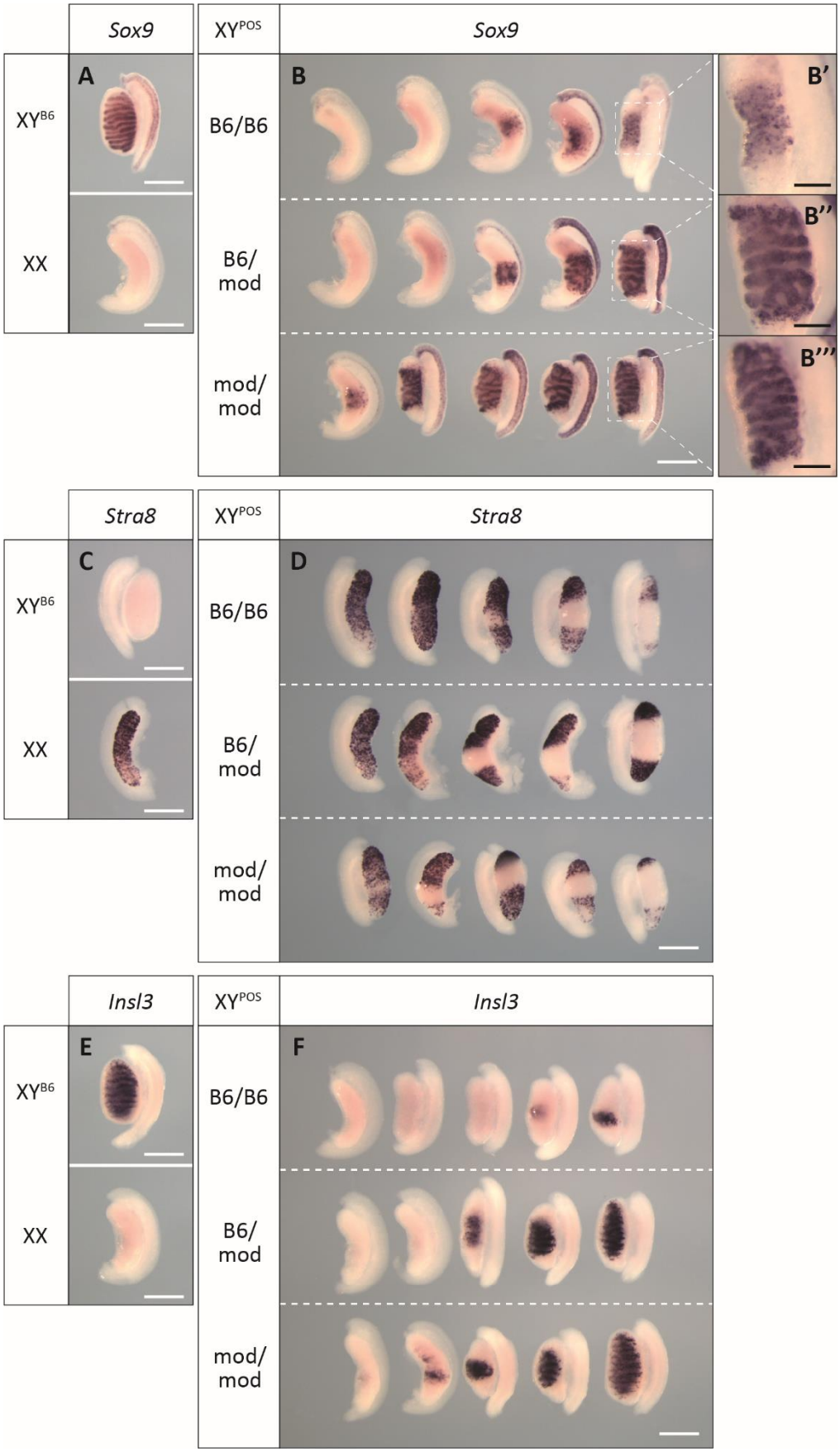


Figure 4.2 Gene expression analyses in B6.XY^{POS}.*Mod13(41-71)* gonads. For full figure legend, please see the previous page.

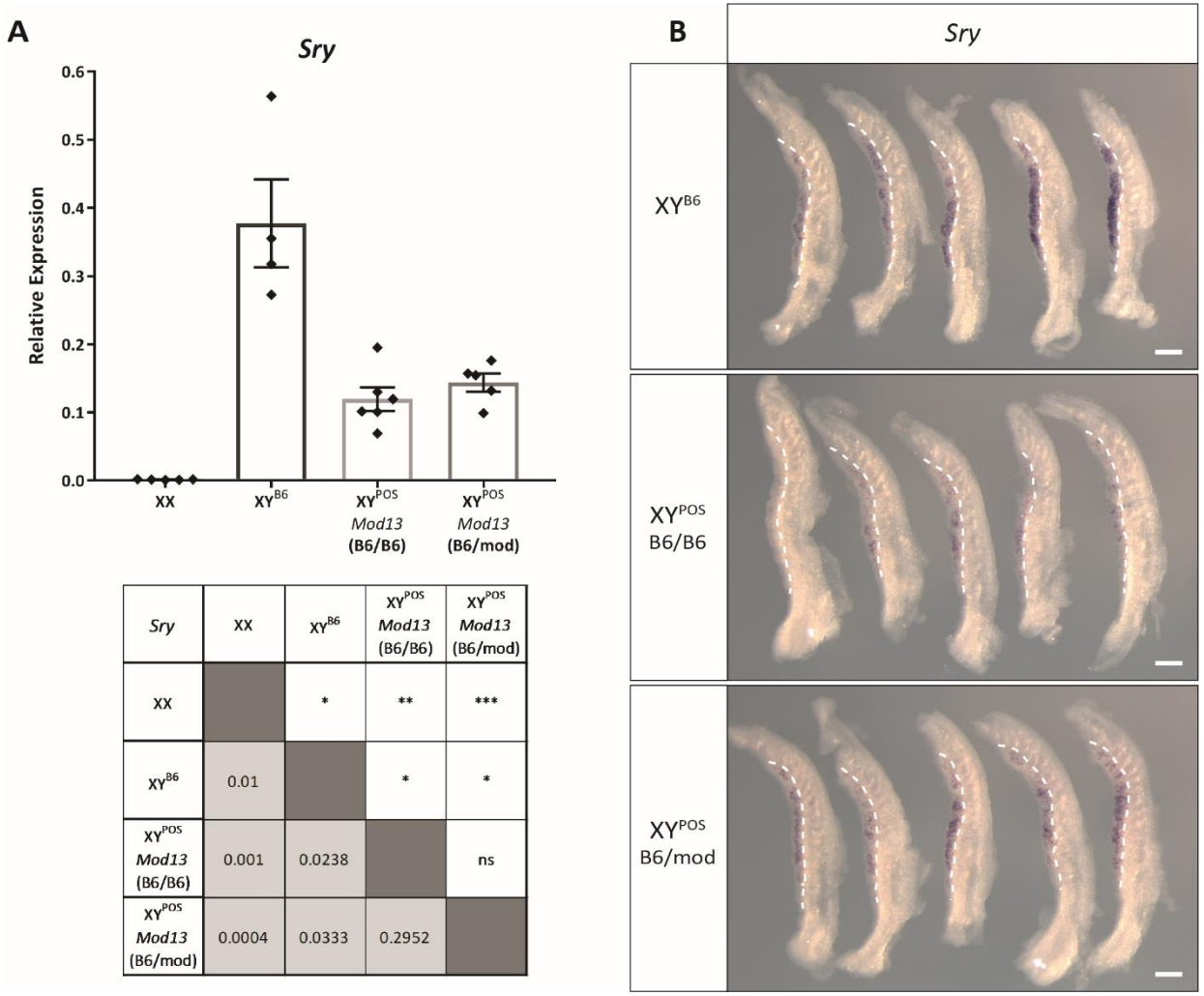
4.2.3. *Sry* expression analysis

Gene expression analysis at 14.5 dpc indicated that the testicular pathway is upregulated in carriers of *Mod13*, but the questions remain: when and how does *Mod13* act? Given the reported delay in *Sry* expression in B6.XY^{POS} gonads, restoration of *Sry* expression is a good candidate for the underlying mechanism of *Mod13* in rescuing testicular development. Using low numbers of gonads, Bullejos and Koopman (2005) reported that *Sry* expression peaks between 16 ts and 18 ts (11.5 dpc) in B6 gonads, before rapidly declining. They reported *Sry* expression present at this stage in B6.Y^{POS}, but at much lower levels, expression instead peaking between 20ts and 22ts. The same observation has been made in B6.XY^{POS} embryos at Harwell, albeit before the discovery of *Mod13* (M. Pope, unpublished). To examine the impact of *Mod13* on *Sry* expression, stage-matched gonads from B6.XY^{POS} B6/B6 and B6/mod embryos were collected, along with B6.XY^{B6} controls, and analysed by both qRT-PCR and WMISH. Due to the limited availability of embryos, a complete time-course could not be performed to fully assess the temporal profile of *Sry* expression. Instead a single time-point was selected, when *Sry* expression peaks in B6.Y^{B6}.

Sub-dissected gonads from 16 ts and 17 ts embryos were used together for analysis by qRT-PCR **(Fig 4.3A)**. As expected, *Sry* was detected in B6.XY^{B6} gonads but not in XX gonads. *Sry* expression was detected in both B6.XY^{POS} genotypes but at significantly lower levels than in B6.XY^{B6}, in accordance with the sex reversal phenotype. *Sry* was 3.2-fold higher in B6.XY^{B6} gonads than in B6.XY^{POS} B6/B6 and 2.6-fold higher than in B6.XY^{POS} B6/mod. There was no significant

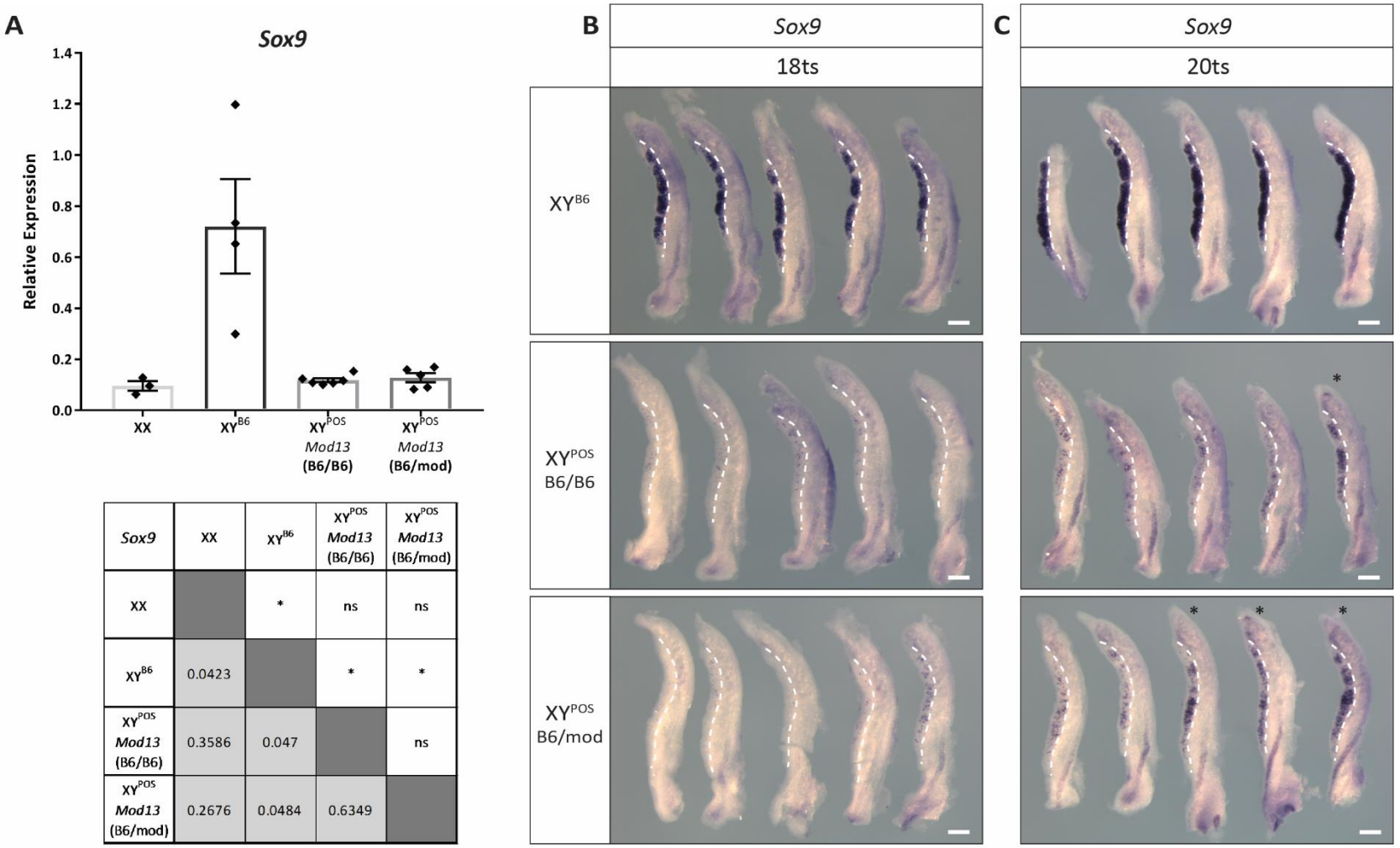
difference between B6/B6 and B6/mod, although the mean relative expression was higher in B6/mod (0.1437) than in B6/B6 (0.1194).

The complexity of the situation was revealed when gonads were analysed by WMISH, specifically at 16ts **(Fig 4.3B)**. Again, *Sry* expression was strong throughout the gonad in B6.XY^{B6}, and weak in B6.XY^{POS} gonads. In B6/B6 gonads *Sry* expression was limited to the centre of the gonad, in line with a delayed expression profile as described by Bullejos and Koopman. However in B6/mod gonads *Sry* expression had extended to the poles of some of the gonads analysed. Levels of *Sry* appeared stronger in some B6/mod gonads than in B6/B6 gonads, although there was a considerable degree of variability. Such variability may be missed when low numbers of samples are analysed. This variability is in line with the phenotypic variability observed at 14.5 dpc. These data suggest that *Mod13* makes a subtle, positive impact on *Sry* expression in B6.XY^{POS} gonads at the time of testis determination.



Sry	XX	XY^{B6}	XY^{POS} *Mod13* (B6/B6)	XY^{POS} *Mod13* (B6/mod)
XX		*	**	***
XY^{B6}	0.01		*	*
XY^{POS} *Mod13* (B6/B6)	0.001	0.0238		ns
XY^{POS} *Mod13* (B6/mod)	0.0004	0.0333	0.2952	

Figure 4.3 *Sry* expression in B6.XY^{POS}.*Mod13* gonads. *Sry* expression was analysed by qRT-PCR **(A)** and WMISH **(B)**. qRT-PCR was performed on subdissected gonads from embryos at 16/17 ts (11.5 dpc). Relative expression was calculated relative to *Hprt1*.Two-tailed Welch's t-tests were performed. *P*-values are detailed in accompanying table. No significant difference between B6.XY^{POS} B6/B6 and B6/mod gonads was found. Error bars show SEM. In contrast, WMISH showed elevated *Sry* expression in some B6/mod gonads compared to B6/B6, but expression was not as high as in XY^{B6}, where variability is still observed. *N* = 5 pairs per genotype, stage matched at 16 ts. Only left gonad shown. Dotted line indicates the gonad-mesonephros boundary, with the gonad on the left. Scale bars 200 µm.



Sox9	XX	XY^{B6}	XY^{POS} *Mod13* (B6/B6)	XY^{POS} *Mod13* (B6/mod)
XX		*	ns	ns
XY^{B6}	0.0423		*	*
XY^{POS} *Mod13* (B6/B6)	0.3586	0.047		ns
XY^{POS} *Mod13* (B6/mod)	0.2676	0.0484	0.6349	

Figure 4.4 *Sox9* expression in B6.XY^{POS}.*Mod13* gonads. For full figure legend, please see the following page.

Figure 4.4 *Sox9* expression in B6.XY^{POS}.*Mod13* gonads. *Sox9* expression was analysed by qRT-PCR **(A)** and WMISH **(B-C)**. qRT-PCR was performed on subdissected gonads from embryos at 16/17 ts (11.5 dpc). Relative expression was calculated relative to *Hprt1*.Two-tailed Welch's t-tests were performed. *P*-values are detailed in accompanying table. Neither B6.XY^{POS}, B6/B6 nor B6/mod were significantly different to XX or each other. Error bars show SEM. WMISH analysis performed on UGRs from embryos at 18 ts **(B)** showed no difference in *Sox9* expression between B6.XY^{POS} B6/B6 or B6/mod gonads. WMISH analysis at 20 ts **(C)** showed elevated *Sox9* expression in 1 out of 5 B6/B6 gonads and 3 out 5 B6/mod gonads (indicated by asterisks). *N* = 5 pairs per genotype, stage matched. Only left gonad shown. Dotted line indicates the gonad-mesonephros boundary, with the gonad on the left. Scale bars 200 μm.

4.2.4. *Sox9* expression analysis

Analysing the timing of *Sox9* expression should provide an indication of when *Mod13* is acting to promote the testis-determining pathway. Sub-dissected gonads from 16/17 ts embryos were used for analysis by qRT-PCR **(Fig 4.4A)**. At this stage the mean relative *Sox9* expression was 6-fold higher in B6.XY^{B6} than in both B6.XY^{POS} genotypes, which were not significantly different from XX samples. A similar result was found when analysed at 18 ts by WMISH **(Fig 4.4B)**. Strong *Sox9* levels were evident throughout B6.XY^{B6} gonads, but very low levels were seen in B6.XY^{POS} gonads, with some gonads only expressing *Sox9* in a handful of cells. Only one of the B6.XY^{POS} B6/mod gonads analysed appeared to have *Sox9*-positive cells across the full length of the gonad, and at much lower levels than in B6.XY^{B6} gonads. These data show that rescue by *Mod13* has not taken affect by a positive impact on *Sox9* expression as early as 11.25-11.5 dpc.

When analysed by WMISH at 20 ts, approximately 4 hours later, the impact of *Mod13* was more evident **(Fig 4.4C)**. While expression was still not as high as in B6.XY^{B6} (as was to be expected given *Mod13* does not convey complete protection), 3/5th of B6.XY^{POS} B6/mod gonads expressed *Sox9* at a noticeably higher level **(Fig 4.4C asterisk)**, in contrast to 1/5th of B6.XY^{POS} B6/B6 gonads. This strongly correlates with the ratios of ovotestes observed at 14.5 dpc. I would speculate that these gonads have sufficient *Sox9* expression to drive some degree of testicular development, while the remaining gonads would develop as ovaries. *Sox9* is directly upregulated by SRY, and therefore increases correspondingly with increased *Sry* expression. The increased *Sox9* expression in

B6.XY^{POS} B6/mod gonads at 11.75 dpc is consistent with a subtle but significant increase in *Sry* expression in some gonads caused by the presence of *Mod13*.

4.2.5. The effect of *Mod13* in XX gonads

To examine if *Sox9* expression was elevated in B6/mod gonads, independently of upregulation of *Sry*, XX gonads from B6.*Mod13(41-71)* embryos were collected and analysed by qRT-PCR at 11.5 dpc and compared to B6.XY^{B6} of the same stage **(Fig 4.5A)**. As expected, *Sox9* expression was significantly higher in XY tissue than in either of the XX groups. There was no detectable increase in *Sox9* expression in B6/mod XX gonads compared to B6/B6 XX gonads. WMISH analysis was performed on XX gonads at 11.5 dpc **(Fig 4.5B)** and 14.5 dpc **(Fig 4.5C)**. There was no evidence of ectopic *Sox9* expression at either 14.5 dpc or 11.5 dpc. XX B6/mod and mod/mod gonads had apparently normal morphology and extensive expression of the meiotic germ cell marker *Stra8* at 14.5 dpc. This suggests that ovarian development had progressed successfully to this stage, with no overt disruption by the pro-testicular pathway. These data support the theory that *Mod13* acts via upregulation of *Sry*, since XX gonads, which do not have *Sry*, were unaffected.

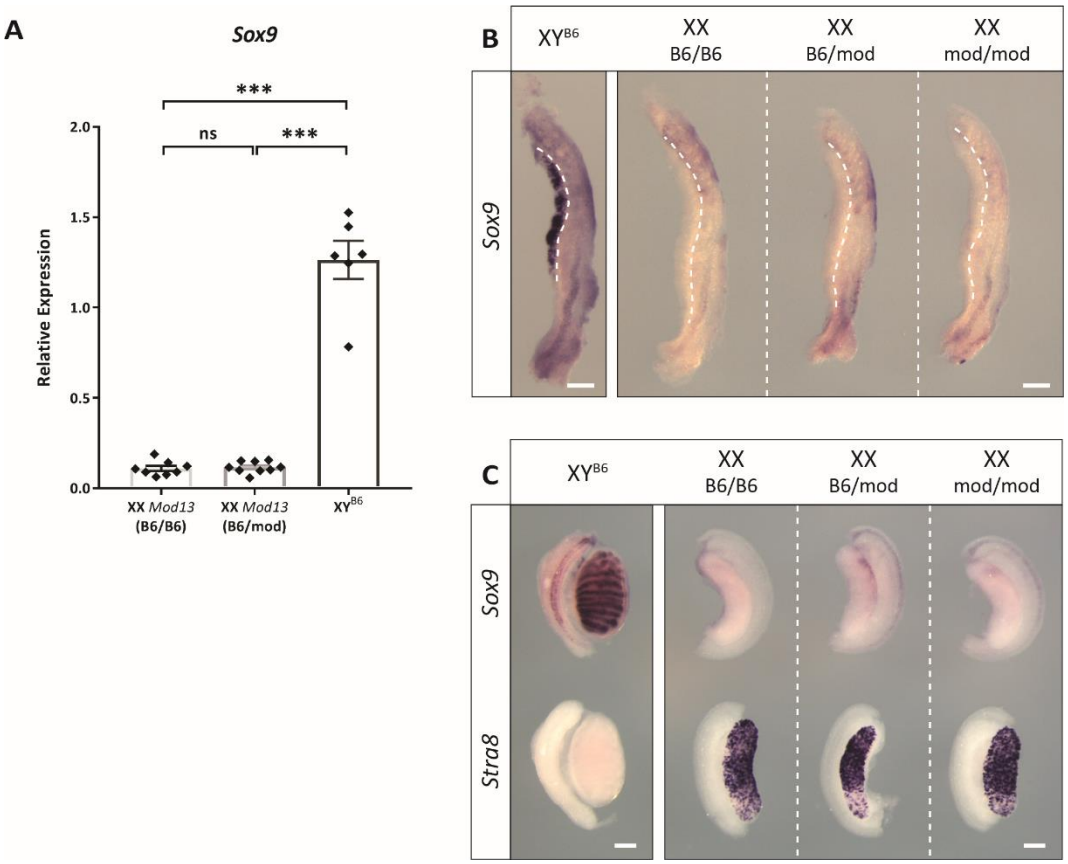
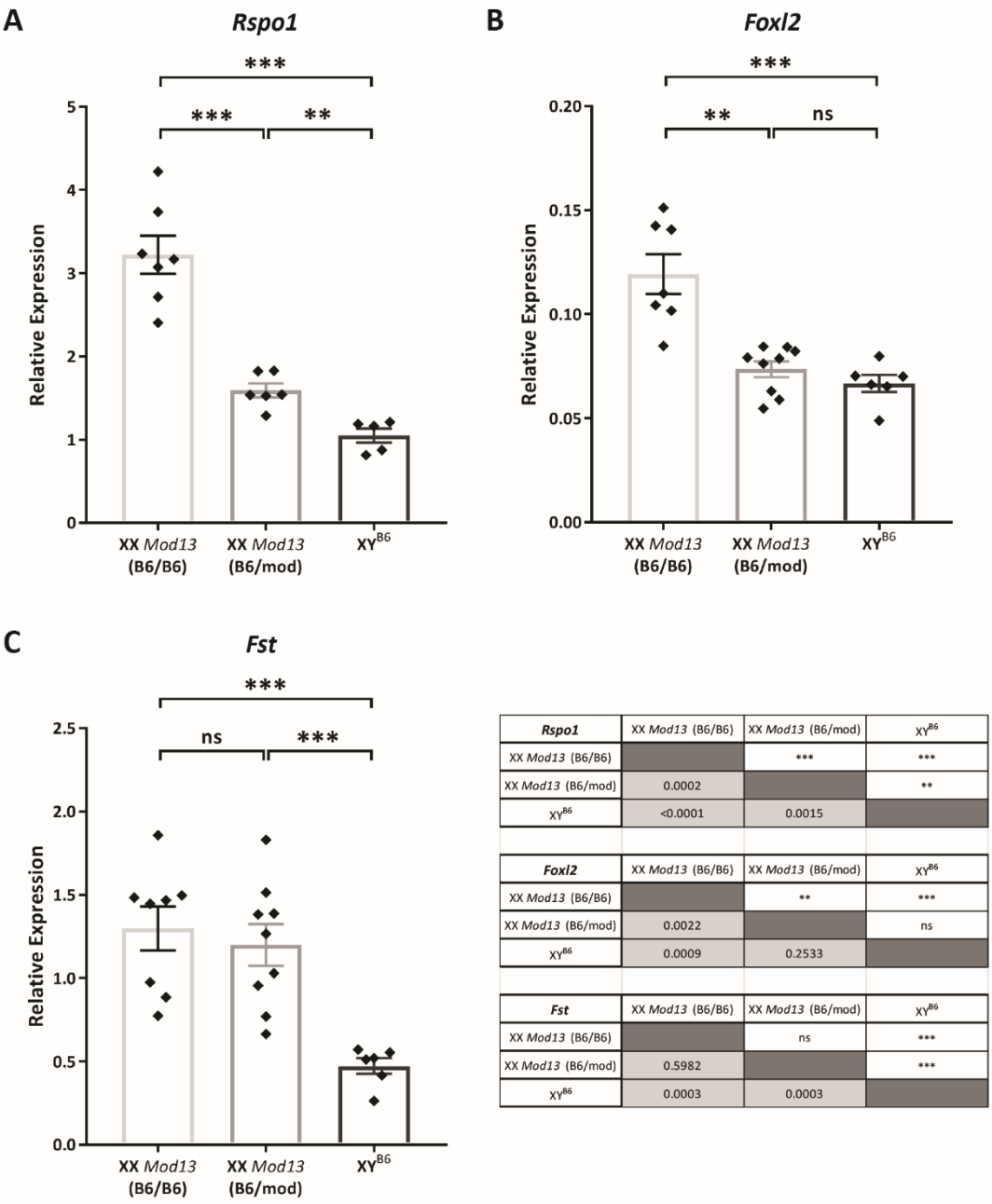


Figure 4.5 *Sox9* expression in B6.XX.*Mod13* gonads. *Sox9* expression was analysed by qRT-PCR at 11.5 dpc (17 ts) **(A)** and WMISH at 11.5 dpc (18 ts) **(B)** and 14.5 dpc **(C, first row)**. *Stra8* expression was also analysed by WMISH at 14.5 dpc **(C, second row) (A)** qRT-PCR was performed on subdissected gonads. Relative expression was calculated relative to *Hprt1*.Two-tailed Welch's t-tests were performed. XX B6/B6 and B6/mod was significantly different from XY^{B6} (*p*-values <0.0001 and 0.0001 respectively) but not different from each other (*p*-value = 0.7305). Error bars show SEM. No difference between B6/B6 and B6/mod was seen by WMISH. **(B)** *N* = 3 gonad pairs per genotype. Dotted line indicates the gonad-mesonephros boundary, with the gonad on the left. **(C)** *N* = 7 pairs per genotype. Scale bars 200 µm.

The expression of pro-ovarian genes was not analysed in B6.XY^{POS} gonads because of the difficulty in separating any direct effect of *Mod13* on these from its positive impact on *Sry* and *Sox9*, which would be predicted to also inhibit pro-ovarian genes. Elevated expression of these genes could be a direct consequence of the presence or absence of *Mod13*, or a consequence of the gonad developing as an ovary, which is also complicated by phenotypic variability in the degree of sex reversal. Therefore, the expression of pro-female genes was instead analysed in XX gonads, from B6.*Mod13(41-71)* embryos, by qRT-PCR **(Fig 4.6)**. At 17 ts (11.5 dpc), the expression of both *Rspo1* and *Foxl2* was significantly reduced in B6/mod gonads compared to B6/B6 gonads. *Rspo1* expression was 2-fold higher in XX B6/B6 gonads than in XX B6/mod gonads, which was 1.5-fold higher than in XY B6 gonads. Levels of *Foxl2* were not significantly different between XX B6/mod gonads and XY B6 gonads, but were 1.6-fold higher in XX B6/B6 gonads, compared to XX B6/mod gonads. The levels of *Fst* were unchanged in XX B6/mod gonads, compared to XX B6/B6 gonads, which were both significantly elevated, compared to XY controls.



Rspo1	XX *Mod13* (B6/B6)	XX *Mod13* (B6/mod)	XY^{B6}
XX *Mod13* (B6/B6)		***	***
XX *Mod13* (B6/mod)	0.0002		**
XY^{B6}	<0.0001	0.0015	

Foxl2	XX *Mod13* (B6/B6)	XX *Mod13* (B6/mod)	XY^{B6}
XX *Mod13* (B6/B6)		**	***
XX *Mod13* (B6/mod)	0.0022		ns
XY^{B6}	0.0009	0.2533	

Fst	XX *Mod13* (B6/B6)	XX *Mod13* (B6/mod)	XY^{B6}
XX *Mod13* (B6/B6)		ns	***
XX *Mod13* (B6/mod)	0.5982		***
XY^{B6}	0.0003	0.0003	

Figure 4.6 Expression of pro-ovarian genes in B6.XX.*Mod13* gonads. Expression of *Rspo1* **(A)**, *Foxl2* **(B)** and *Fst* **(C)** was analysed by qRT-PCR. qRT-PCR was performed on subdissected gonads from 11.5 dpc (17ts) embryos. Relative expression was calculated relative to *Hprt1*.Two-tailed Welch's t-tests were performed. *P*-values are detailed in accompanying tables. *Rpso1* was significantly differentially expressed between the three genotypes, with XX B6/mod expression an intermediate between XX B6/B6 and XY^{B6}. *Foxl2* expression in XX B6/mod gonads was significantly different to expression in XX B6/B6 gonads, but not different compared to XY^{B6}. *Fst* expression was not significantly different between XX B6/B6 and B6/mod, which were both significantly higher than in XY^{B6}. Error bars show SEM.

4.3. Discussion

4.3.1. *Mod13* impacts primary sex determination

Mod13 provides partial protection against sex reversal in B6.XY^{POS} embryos by promoting testis differentiation; the presence of *Mod13* results in a greater degree of testicular tissue developing in XY^{POS} gonads following inspection at 14.5 dpc. *Mod13* is a dominant modifier with additive effects; heterozygosity for *Mod13* is sufficient to promote testis development, with some evidence of dose-dependency. 28% of homozygous B6 gonads develop as ovotestes rather than ovaries, compared to 60% of heterozygous gonads and 69% of homozygous mod gonads. The impact of *Mod13* on testis determination is not fully penetrant; even with two copies of *Mod13*, 31% of gonads develop as ovaries. It was also observed that the presence of *Mod13* increased the quality of testicular tissue; ovotestes containing greater testicular organisation and well-defined tubule structure were found most often in embryos carrying one or two copies of *Mod13*. This was likely due to the greater numbers of differentiating Sertoli cells (identified by WMISH analysis for *Sox9* expression) in these gonads. Examination of ovotestes revealed that gonads with small regions of *Sox9*-positive cells typically exhibited no apparent organisation into testis cords. This suggests that a minimum number of cells must receive signals from SRY/SOX9 and differentiate into Sertoli cells in order for testis cords to form properly. The restricted *Sox9* expression often observed in homozygous B6 gonads at this stage was likely not conducive to testis cord formation, hence the decrease in testicular quality observed in gonads lacking *Mod13*.

The ovotestes seen here are stereotypical in their expression of testicular markers in the centre of the gonad, and their expression of ovarian markers restricted to the poles (particularly the posterior pole) of the gonad. This pattern of sex reversal is thought to be a consequence of the wave-like pattern of *Sry* expression (Bullejos and Koopman, 2001). Cells in the centre of the gonad receive peak SRY signals first and, therefore, effect upregulation of *Sox9* first. Cells in the anterior then posterior poles of the gonad are exposed to SRY approximately 4 to 8 hours later, meaning they do not exhibit upregulated *Sox9* until after the central region. Therefore, in the case of delayed *Sry* expression, the cells in the poles are more susceptible to being overcome by pro-ovarian signals. It was observed that *Sox9*-positive cells were almost always disorganised on the boundary between testicular and ovarian areas in these B6.XY^{POS} ovotestes. This suggests that despite the testis-determining pathway being activated in these cells, they were unable to join the nearby organised testis cords. Perhaps there are opposing, pro-ovarian, signals also present in these cells (or adjacent cells), confusing the signals from SOX9 and associated paracrine signals and suppressing their ability to migrate and/or communicate with other Sertoli cells. This would be an interesting hypothesis to investigate, using immunohistochemistry to look carefully for overlap of pro-testis and pro-ovary proteins, and how such co-localisation affected tissue morphology.

4.3.2. *Mod13* downregulates *Rspo1* and *Foxl2*

Expression levels of the pro-ovarian genes *Rspo1* and *Foxl2* were lower in XX heterozygous gonads at 11.5 dpc, likely as a consequence of modification by genes in the *Mod13* region. Mouse *Rspo1* is located on chromosome 4, while

Foxl2 is located on chromosome 9, therefore these results were not a consequence of direct changes to the gene sequence, but of changes in the regulation of their expression. It may be that *Mod13* SNPs present in the regulatory or coding region of an antagonist of ovarian gene expression increase its expression or activity. Such an antagonist would be independent of *Sox9*, which appears unaffected by *Mod13* in XX gonads. An alternative hypothesis may be that SNPs in a pro-ovarian gene reduce its expression or function. The expression of *Foxl2* is thought to be dependent on RSPO1 in some cells of the developing ovary (Auguste et al., 2011); however, expression analysis in *Rspo1*-deficient gonads show that expression of *Foxl2* is not completely dependent on RSPO1 (Chassot et al., 2008). Therefore, the downregulation of *Foxl2* observed in heterozygous gonads may represent the independent regulation of *Foxl2* by a gene in *Mod13*, or may be a consequence of the downregulation of *Rspo1*. If the downregulation of *Foxl2* is independent of *Rspo1*, the fact that two female genes were significantly downregulated by the presence of *Mod13* could suggest the existence of a master regulator of ovarian gene expression on chromosome 13. There is not currently sufficient evidence to support this hypothesis, however. Interestingly, *Fst* expression was not significantly affected in heterozygous gonads. *Fst* is regulated early by WNT signalling and later by FOXL2 (Kashimada et al., 2011; Yao et al., 2004). The expression analyses described here only measured levels of *Fst* at 11.5 dpc. If expression was analysed in older gonads, at 12.5 dpc for example, downregulation of *Fst* in B6/mod gonads might be observed. Further elucidation of the role *Mod13* plays in downregulation of pro-ovarian genes would require expression analyses of more genes downstream of

Rpso1 and *Foxl2*, such as *Wnt4* and *Axin2* (a transcriptional target of canonical WNT signalling (Chassot et al., 2008)).

At present, two genetic pathways have been shown to play complementary roles in ovary differentiation: one driven by RSPO1 and one driven by FOXL2 (Pannetier et al., 2016). Both these genes appear to be downregulated in carriers of *Mod13*. Expression levels of these ovarian genes were not measured in XY^{POS} gonads, due to the difficulty of determining cause and effect. Homozygous B6 B6.XY^{POS} gonads have been shown to develop most often as ovaries and, therefore, would likely express ovarian genes at higher levels during sex determination. It is possible to isolate the impact of *Mod13* on pro-ovary gene expression from its impacts on *Sry* expression by examining XX gonads. However, this approach assumes that the same downregulation of *Rspo1* and *Foxl2* occurs in XY gonads, in order to extrapolate to the role of *Mod13* in rescuing B6.Y^{POS} sex reversal. RSPO1 activates canonical WNT signalling through the stabilisation of β-catenin in the embryonic XX gonad, and is required for normal ovarian development (Chassot et al., 2008; Tomizuka et al., 2008). Loss of *Rspo1* results in dramatic reduction in expression of *Wnt4* (Chassot et al., 2008; Tomizuka et al., 2008). WNT4 antagonises the testis-determining pathway via downregulation of *Fgf9* (Kim et al., 2006). Evidence also suggests that β-catenin, which is stabilised as a result of RSPO1 activating canonical WNT-signalling, interferes with SF1 binding to the *Sox9* enhancer TESCO, inhibiting transcription of *Sox9* (Bernard et al., 2012). It follows that reduction in the expression of *Rspo1* could result in upregulation of *Sox9* and enhanced testis determination.

FOXL2 is required for differentiation of the supporting cell lineage into granulosa cells and the maintenance of ovarian fate (Schmidt et al., 2004; Uda et al., 2004; Uhlenhaut et al., 2009). Given that loss of *Foxl2* does not result in female-to-male sex reversal in the fetal period, it is not thought to play a role in primary sex determination. Intriguingly, however, loss of *Foxl2* in embryos also lacking *Rspo1* exacerbates the partial gonadal sex reversal observed in postnatal gonads only lacking *Rspo1*, resulting in sex reversal occurring earlier (Auguste et al., 2011). Recently, it has been reported that loss of *Foxl2* can protect against XY sex reversal caused by loss of FGF signalling (Bagheri-Fam et al., 2017). The authors suggest that FOXL2 cooperates with WNT4 in the repression of *Sox9* and the feminisation of the gonad. It appears that the role of FOXL2 in the early gonad is complex and not fully understood. As such, the impact of reduced *Foxl2* in B6/mod gonads also remains unclear.

4.3.3. The effect of *Mod13* on ovarian development and function

Loss-of-function models of *Rspo1* exhibit partial female-to-male sex reversal in the perinatal/postnatal period, exacerbated by additional loss of *Foxl2* (Auguste et al., 2011; Chassot et al., 2008; Tomizuka et al., 2008). No ovarian defects were observed in XX B6/mod or mod/mod embryos at 14.5 dpc. This indicates that the reduction in *Rspo1* and *Foxl2* expression at 11.5 dpc as a result of *Mod13* was not sufficient to significantly impact ovarian development in XX embryos. However, reports by Chassot et al. (2008) and Tomizuka et al. (2008) describe defects in the supporting cell lineage of *Rspo1*-deficient ovaries not becoming apparent

until later in development. By 18.5 dpc Sertoli-like cells were reported to have differentiated and seminiferous tubule-like structures were visible in *Rspo1*$^{-/-}$ ovaries. In this thesis, ovaries from XX females were not examined later than 14.5 dpc. XX heterozygous females were fertile during intercross matings to B6.Y^{POS} males, but regularly produced very low numbers of pups per litter, fewer than would be expected from wild-type B6 females. Histological examination of postnatal and adult ovaries from XX B6/mod and XX mod/mod females would be of interest, to test for ovarian or follicular defects as a result of *Mod13*. Formal fertility testing of XX B6/mod females, using B6.Y^{B6} studs to remove the confounding factor of B6.Y^{POS} subfertility, would also be informative.

4.3.4. *Mod13* directly upregulates *Sry*, but not *Sox9*

In addition to the down-regulation of *Foxl2* and *Rspo1*, *in situ* hybridisation data suggest that *Mod13* modestly elevates *Sry* expression in XY^{POS} gonads, partially reversing the reduced expression thought to cause B6.Y^{POS} sex reversal. The peak of *Sry* expression is reported to be delayed by 9.6 hours in B6.Y^{POS} gonads relative to B6.Y^{B6} gonads (Bullejos and Koopman, 2005). It should be noted, however, that all previous studies of B6.Y^{POS} sex reversal require re-evaluation given the recent identification of the chromosome 11 (Arboleda et al., 2014) and chromosome 13 (this thesis) modifiers, which were not known to exist at that time. Individual embryos in those studies may have differed in their genotype at these loci, confounding interpretation of the data. Phenotypic variation also dictates that conclusions should not be made when only small sample numbers are available, as in some earlier studies. Nevertheless, these earlier results suggest that *Sry*^{POS} misses the 'window of opportunity' for activating *Sox9*,

allowing the ovarian pathway to dominate. How *Mod13* influences the temporal profile of *Sry* was not examined here. Further analysis at different timepoints would determine if *Mod13* is bringing the onset of *Sry* forward or simply increasing overall expression levels. Transgenic overexpression of *Map3k4* can rescue B6.Y^{POS} sex reversal via upregulation of *Sry* (N. Warr, M. Pope, unpublished data). MAP3K4 is a component of the MAPK signalling cascade required for timely expression of *Sry* (Bogani et al., 2009; Gierl et al., 2012; Warr et al., 2012). GADD45γ-MAP3K4-MAP2K3/6-MAPK11/14 act as a pathway, resulting in phosphorylation of target proteins – such as the transcription factor GATA4 – that activate *Sry*. Other known regulators of *Sry* expression include FOG2 (the co-factor of GATA4), homeobox transcription factors SIX1 and SIX4, histone demethylase JMJD1A, histone acetyl-transferases p300 and CBP, and polycomb group protein CBX2 (see chapter 1, section 1.5.1). *Mod13* polymorphisms that increase the level or function of any of these proteins, or other unknown pro-*Sry* proteins, may be responsible for the subtle increase in *Sry* expression.

In wild-type XY B6 gonads, the strong male-specific expression of *Sox9* was already apparent by 11.5 dpc, in accordance with the published expression profile (Kent et al., 1996). Expression of *Sox9* was increased in 3 of the 5 heterozygous XY^{POS} gonads when analysed at 20 ts (11.75 dpc) using WMISH. However, no evidence of *Sox9* upregulation was apparent when heterozygous gonads were examined 4-6 hours earlier at 16/17 ts (using qRT-PCR) or at 18 ts (using WMISH), and levels of *Sox9* expression were not significantly different from XX samples. No ectopic expression of *Sox9* was observed in XX

heterozygous gonads either at 11.5 dpc or at 14.5 dpc. The delay in rescue of *Sox9* expression levels in XY gonads, and the absence of detectable change in XX gonads, suggests that *Mod13* is not directly upregulating *Sox9*. The later elevation is likely a by-product of modest elevation in *Sry* expression levels and the reduction of pro-ovarian gene expression, which would otherwise repress *Sox9*. Neither of these is sufficient, however, to drive strong, early *Sox9* expression as in wild-type controls.

4.3.5. Concluding remarks

From the data described in this chapter it can be surmised that *Mod13* contains a gene or genes that modify sex determination. By identifying candidate genes on chromosome 13 that are responsible for the testis-promoting effect of *Mod13*, a novel sex-determining gene may be identified. The fact that *Mod13* appears to be affecting at least two points in the sex determination pathway (upregulation of *Sry*, downregulation of *Foxl2*/*Rspo1*), supports the theory of *Mod13* behaving like a QTL: the combined effect on both these pathways promotes testis determination. While it is possible for a single gene to be simultaneously upregulating *Sry* and downregulating *Foxl2* and *Rspo1*, it is most likely the effect of variants in two or more genes. Antagonism by the pro-ovarian pathway centres on repression of *Sox9* and *Fgf9*; no female proteins are known to directly downregulate *Sry*. A mechanism by which these genes repress *Sry* does not make sense in the context of evolving a system canalising the XY gonad down the testicular pathway and canalising the XX gonad down the ovarian pathway. In cultured cells and cultured gonadal explants, β-catenin inhibits expression of *Sox9*, but not *Sry* (Bernard et al., 2012). While it remains possible that the

upregulation of *Sry* is secondary to the downregulation of *Foxl2* and *Rspo1*, the search for candidate genes will assume that these impacts are independent.

Chapter 5: Candidate genes in the *Mod13* region

5.1. Introduction

5.1.1. Summary so far

Mod13 is a non-B6-derived genetic region in the centre of chromosome 13, identified in the B6.Y^{POS} genome at Harwell, which protects against XY gonadal sex reversal. B6.XY^{POS} mice without *Mod13* develop predominantly as females, while a significantly higher proportion of B6.XY^{POS} mice heterozygous for *Mod13* develop as males. The presence of *Mod13* promotes primary testis determination: partially via the upregulation of *Sry* and partially via the downregulation of the pro-ovarian genes, *Rspo1* and *Foxl2*, which antagonise the testis-determining pathway. The impact of *Mod13* is significant but subtle, tipping the balance in favour of the male pathway rather than the female pathway; but the rescue effect is far from 100%. None of the genes in the non-B6 segments of *Mod13* have previously been reported as sex-determining, making this an interesting opportunity to identify a novel regulator(s) of the sex-determining pathways. It is worth reiterating that *Mod13* behaves like a QTL, and likely contains several genes that each contribute to its protective effects. When I describe a gene as a candidate, I am by no means suggesting that it is the only gene in *Mod13* that might impact sex determination.

5.1.2. Strategies

Mod13 spans approximately 38 Mb of chromosome 13 and has a segmental structure, as described in Chapter 3. 268 genes classified as protein-coding are present in the non-B6 regions of *Mod13*, and an additional 77 genes lie in the B6 regions, whose regulatory elements could conceivably lie in the non-B6 areas. A

list of these genes can be found in **Table S4** in Appendix 3. There are also 493 non-coding RNAs located across the region that should not be discounted. In order to identify candidate genes, the most important task is to narrow down the region of interest. Genetic recombination occurs during meiosis. Exchange between the maternal chromosome 13 (carrying only B6 sequence) and the paternal chromosome 13 (carrying *Mod13*), can result in the generation of smaller non-B6 regions. Examining how such recombinant alleles of *Mod13* impact sex development may assist in narrowing down the area of chromosome 13 that harbours a gene (or genes) with pro-testis function. However, it should be remembered that in the case of examining a QTL such as *Mod13*, inferring the relationship between a specific recombinant chromosome and the testis determination phenotype it is associated with, in a small number of animals, can be a risky process. Nevertheless, once the region has been limited to a more practical size, selection of potential candidates can be based on gonadal expression patterns and relevant protein functions, using both extensive literature review and available transcriptomic or related functional data.

5.1.3. Predicted properties of a candidate gene

A gene capable of regulating testis determination would need to be expressed with a specific temporal profile and in particular gonadal cells. Differential gene expression in carriers of *Mod13* is already apparent by 11.5 dpc, with the downregulation of *Rspo1* and *Foxl2* in XX gonads. Given this, a candidate gene is expected to be expressed in the gonad at this key sex-determining stage. Also, given its proposed role in positively impacting on *Sry*, a gene of interest would be expected to be expressed prior to or at 11.5 dpc, in cells of the supporting

lineage of the XY gonad, i.e. pre-Sertoli cells. Sexually dimorphic expression at later stages of gonad development could be indicative of a pro-testis or pro-ovary gene, so would also be of interest. It may also more clearly reveal the cellular specificity of expression. However, regulators of *Sry* are not required to be sexually dimorphic. For instance, *Gadd45g* is an essential component of the network of genes regulating *Sry* expression, but is expressed in the supporting cells of both XX and XY gonads, from 10.5 dpc to 13.5 dpc (Gierl et al., 2012; Warr et al., 2012). Indeed, given that an *Sry* transgene, consisting simply of a Y chromosome-derived genomic fragment containing the gene, can cause testis development in an XX fetus (Koopman et al., 1991), it is predicted that all gene products required for *Sry* expression are available in the developing XX gonad.

Despite knowing approximately where and when *Mod13* must act in the genetic and cellular pathways controlling testis determination, the exact mechanism underlying the modifier region's ability to promote testicular development remains difficult to determine. The intricacies that connect each genetic element of the sex determination network are not fully understood and many components are still unknown. The fact that the testis- and ovary-determining pathways are mutually antagonistic ensures that working to understand *Mod13* will be complex, and that the potential role of candidate genes remains wide open. Protein functions of interest include intracellular and intercellular signalling. WNT4 and FGF9 are both secreted signalling factors that play antagonistic roles to drive ovarian and testicular development respectively (Kim et al., 2006). The p38 MAP kinase signalling pathway is required for timely expression of *Sry* (Bogani et al., 2009; Warr et al., 2012). Epigenetic mechanisms

are also known to play a role in regulation of *Sry* expression (Carre et al., 2018; Kuroki et al., 2013), making histone-modifying proteins interesting candidates. A protein with known physical or functional interactions with these pathways or with any of the sex-determining genes described in chapter 1, in gonads or in other systems, is a good candidate for an impact on sex determination.

5.1.4. PTCH1/DHH signalling

Although no known sex-determining genes are present in *Mod13*, one gene with a known role in testis development is *Patched 1* (*Ptch1*), a gene located at position 63 Mb of chromosome 13. *Ptch1* is a key component of the Hedgehog signalling pathway, which plays many diverse roles in tissue patterning and development. Hedgehog (HH) proteins are secreted molecules that transmit signals to adjacent cells by binding to Patched receptors. Two Patched proteins – PTCH1 and PTCH2 – have been identified in vertebrates, and three HH proteins have been identified in mammals: Sonic Hedgehog (SHH), Indian Hedgehog (IHH) and Desert Hedgehog (DHH) (Lee et al., 2016). *Hh* expression rarely overlaps during embryogenesis, with each ligand acting in distinct systems (Bitgood and McMahon, 1995). *Ptch2* is not detected in XY gonads between 11.5 dpc and 13.5 dpc (Yao et al., 2002), and neither *Shh* nor *Ihh* are expressed in the gonad (Bitgood and McMahon, 1995). Hedgehog signalling in the embryonic gonad is therefore transduced via PTCH1 and DHH.

In the absence of HH proteins, PTCH constitutively represses the action of another transmembrane protein, Smoothened (SMO). HH binding to PTCH on the surface of a target cell results in repression of PTCH, thus de-repressing SMO.

SMO accumulates at the cell surface of the primary cilium and activates the GLI proteins, a family of zinc-finger transcription factors, causing up- or downregulation of target gene expression. GLI1 is only capable of activator functions, whereas GLI2 and GLI3 contain both activator and repressor domains. That said, GLI2 generally functions only as an activator of transcription and GLI3 generally functions as a repressor of transcription, in response to HH signalling (Briscoe and Thérond, 2013).

Signalling by PTCH1 and DHH is required for normal testis development in mice and humans. Homozygous mutations in *DHH* have been identified in multiple human patients with 46,XY complete or partial gonadal dysgenesis (Canto et al., 2004; Das et al., 2011; Paris et al., 2017; Umehara et al., 2000; Werner et al., 2015), to such an extent that *DHH* is considered a prime candidate when a patient presents with 46,XY DSD. Such patients are regularly screened for mutations in *DHH* (Baxter et al., 2015). Reported *DHH* mutations are outlined in **Table 5.1**. Patients generally present with primary amenorrhea, feminine external genitalia and reproductive tracts, and unilateral or bilateral streak gonads, identified by ultrasound. Tumours of the germ cells or somatic cells are common. Four patients were described with neuropathy accompanying gonadal dysgenesis. Histopathological analysis of gonads from a patient with partial gonadal dysgenesis revealed some seminiferous tubules (including Sertoli cells and some germ cells), a nodule of hyperplastic Leydig-like cells, but large areas of irregular fibrotic tissue. Histopathological analysis of a gonad from the same patient's 46,XY sister, who shared an identical mutation, revealed a streak gonad with no seminiferous tubules or Leydig cells (Werner et al., 2015).

Table 5.1 *DHH* mutations in 46,XY gonadal dysgenesis patients

Karyotype	Phenotype	Gonads	DHH mutation	Protein	Reference
46, XY	Partial GD and poly-neuropathy	Testis + streak gonad	c.2T>C	p.M1?	Umehara et al. (2000)
46, XY	Complete GD	Bilateral streak gonads	c.485T>C	p.L162P	Canto et al. (2004)
46, XY	Complete GD	Bilateral streak gonads with gonadoblastoma	c.1086delG	p.L363CfsX4	Canto et al. (2004)
46, XY	Complete GD	Bilateral streak gonads with dysgerminoma	c.1086delG	p.L363CfsX4	Canto et al. (2004)
46, XY	Complete GD	Bilateral streak gonads	c.271_273delGAG	p.E90del	Das et al. (2011)
46, XY	Complete GD	Streak ovaries	c.57-60dupAGCC	p.Q21fsX71	Das et al. (2011)
46, XY	Partial GD and neuropathy	Testis on one side, with seminoma	c.371G>A	p.R124Q	Werner et al. (2015)
46, XY	Partial GD and neuropathy	Bilateral testes, seminoma and gonadoblastoma	c.371G>A	p.R124Q	Werner et al. (2015)
46,XY	Complete GD, facial dysmorphia and neuropathy	Intra-abdominal gonads	c.519G>T	p.W173C	Paris et al. (2017)
46,XY	Complete GD	Intra-abdominal gonads	c.519G>T	p.W173C	Paris et al. (2017)

Initial reports of XY mice lacking functional *Dhh* described sterile males, with disruption to spermatogenesis but relatively normal Sertoli cells, Leydig cells and seminiferous tubule structure (Bitgood et al., 1996). On a different genetic background, 92.5% of XY mice lacking functional *Dhh* had small, undescended testes and presented with female external genitalia. The remaining 7.5% of XY mice had a masculine external appearance but still had undersized testes. Examination of *Dhh* null testes revealed severely malformed seminiferous tubules, evident at birth, which became progressively worse with age. Small aggregations of immature Leydig cells were observed but no adult-type Leydig cells were apparent. Spermatogenesis in the seminiferous tubules was greatly disrupted and some germ cells were seen outside the tubules (Clark et al., 2000). This phenotype was attributed to a failure of differentiation in Leydig cells; analysis of $Dhh^{-/-}$ gonads at 13.5 dpc showed a dramatic reduction in the expression of *Scc*, a marker of fetal Leydig cells (Yao et al., 2002).

Homozygous deletion of *Ptch1* in mice causes embryonic lethality, with embryos dying between 9 dpc and 10.5 dpc due to neural tube defects (Goodrich et al., 1997). This precludes the examination of gonads from homozygous *Ptch1* embryos, since the gonadal ridge appears at 10.5 dpc and genetic and morphological sexual dimorphism does not become apparent until 11.5 dpc and 12.5 dpc, respectively. Heterozygous mice often develop medullablastomas, but males and females are fertile with no gonadal phenotype.

Dhh is expressed in pre-Sertoli cells from 11.5 dpc, while analysis of *Ptch1* places its expression from 12.5 dpc exclusively in the cells of the interstitium (Bitgood et

al., 1996; Yao et al., 2002). The proposed role of hedgehog signalling in the gonad is specification of the Leydig cell fate; DHH produced by Sertoli cells binds to PTCH1 on nearby steroidogenic cell precursors, signalling to the cells to undergo Leydig cell specification. However, transcriptomic data from scRNA-seq analyses (Stevant et al. (2018), see **Fig 5.4C**) and microarray analysis (Jameson et al., 2012b) show that *Ptch1* is expressed in the supporting cell lineage of the XY gonad (i.e. pre-Sertoli cells) from 11.5 dpc. *Ptch1* is also expressed in the supporting cell lineage of the XX gonad at 11.5 dpc, before being downregulated.

The complete gonadal dysgenesis seen in patients with mutations in *DHH* suggests a potential requirement for Hedgehog signalling in primary testis determination, a role not yet identified in mice. Examinations of *Dhh* null mice have produced conflicting phenotypic descriptions, some of which do not seem fully explained by *Dhh* and *Ptch1* only playing a role in Leydig cell specification, particularly the malformation of the seminiferous tubules. Previous studies of the role of *Dhh* used have different genetic backgrounds and observed varied phenotypes. XY mice on a mixed genetic background (129/Sv, C57BL/6 and Swiss Webster) were externally feminised, with severely malformed seminiferous tubules and Leydig cell defects (Clark et al., 2000). XY mice on a 129/Sv inbred background developed as males with relatively normal seminiferous tubules, but arrested spermatogenesis and no spermatids. XY mice on a 129/Sv-C57BL/6J F_1 hybrid background developed as males and had some germ cells progressed to late stages of spermatogenesis (Bitgood et al., 1996). No reports of the use of a congenic C57BL/6J (B6) background to study the *Dhh* null allele have been made. As previously discussed in chapter 3, the B6 background is sensitised for

disruption to primary testis determination, making it an important background for studying the effect of gene deletions on Sertoli cell differentiation. The B6.Y^{AKR} strain is also a useful tool in the study of sex determination, since it has even greater sensitivity to disruptions to testis determination than the B6 background. It can be used to study the effect of heterozygous mutations on sex determination, when homozygotes are non-viable (Bogani et al., 2009; Bouma et al., 2007; Harris et al., 2018; Warr et al., 2016). Given the expression of *Ptch1* in gonadal supporting cells of both sexes at 11.5 dpc, its presence in the testis-promoting core of chromosome 13, and the somewhat unexplained gonadal dysgenesis phenotype in mice and humans lacking *DHH* function, *Ptch1* has been selected as a potential candidate gene for the effect of *Mod13*. A congenic B6 genetic background will be used to re-examine *Dhh*-deficient embryonic gonads, and the B6.Y^{AKR} genetic background will be used to examine *Ptch1* heterozygous embryonic gonads.

5.1.5. Aims

The aim of this chapter is to identify candidate genes in *Mod13* responsible for the effect on primary sex determination. The following strategies will be used:

- Detailed analysis of a recombinant allele of *Mod13* and its ability to rescue sex reversal in B6.Y^{POS}.
- Review of gonadal expression and function of genes present in the non-B6 regions of *Mod13*.

- Re-examination of embryonic gonads lacking copies of *Ptch1* or *Dhh* on a congenic sensitised genetic background, focusing on the specification of Sertoli cells.

5.2. Results

5.2.1. Characterising SNPs in *Mod13*

One additional method to identify candidate genes involves examining the nature of the SNPs present in *Mod13*. A missense mutation with predicted deleterious effect would provide a starting point when short-listing genes. SIFT (Sorting Intolerant from Tolerant) is a software tool that predicts whether an amino acid substitution affects protein function based on sequence homology and the physical properties of amino acids (Sim et al., 2012). 68 of the SNPs in *Mod13* were categorised as missense variants. Using SIFT, 8 of these SNPs were predicted to be damaging. **(Table 5.2)**

Table 5.2 Damaging SNPs in *Mod13*

SNP ID	SNP position	Nucleotide change	Residue change	SIFT score	Affected transcript
rs29976898	13:63061185	C→T	A→V	0.028	RIKEN cDNA 2010111I01
rs29976898	13:63061185	C→T	A→V	0.018	
rs49034569	13:64787819	T→C	H→R	0.03	*Cntnap3*
rs48769465	13:67292855	G→A	T→I	0.008	*Zfp457*
rs49822874	13:67294037	A→G	V→A	0.039	
rs49836738	13:67294062	A→G	S→P	0.025	
rs255071065	13:67389964	G→T	P→T	0.009	*Zfp429*
rs29554455	13:67408329	T→C	N→D	0.002	*Zfp259*

As discussed in chapter 3, the majority of SNPs found within the region of interest are also found in other laboratory mouse strains. This means that most SNPs are variants commonly found in mice, and are therefore unlikely to be deleterious. As expected, all 8 of the predicted damaging substitutions are found in the BALB/cJ sequenced genome, and in the genome of at least one other mouse strain. In some cases, the *Mod13* nucleotide was more common than the B6 nucleotide. This suggests that the SNPs in question are unlikely to dramatically affect the development or health of the mice, either because the gene is functionally redundant, or because SIFT incorrectly categorised the SNP(s) as damaging.

Additionally, single cell RNA seq (scRNA-seq) data, providing gene expression data of all the genes in *Mod13* for gonadal progenitor cells and the supporting cell lineage of XX and XY gonads, was kindly provided by Isabelle Stevant (Stevant et al., 2018). These data indicated that of the genes containing predicted damaging missense SNPs, only *Zfp429* and *Zfp459* are expressed in the progenitors or supporting cells of the gonad between 11.5 dpc and 16.5 dpc, and each is only expressed in a small proportion of cells (data not shown). Neither of the zinc finger proteins encoded by these genes are well characterised.

Only 0.68% of the SNPs identified in *Mod13* during the whole genome analysis are located within coding sequences, and 58.3% of SNPs are located more than 5 kb either upstream or downstream of a transcript. The SNPs responsible for the phenotype observed in B6.Y^{POS}.*Mod13* mice are most likely regulatory, rather than affecting protein function. Unless SNPs lie in characterised regulatory

regions, their potential impact is difficult to determine. In conclusion, characterisation of individual SNPs located in *Mod13*, did not identify any SNPs or genes that should be prioritised for further study.

5.2.2. Narrowing down the region

The first approach to identifying candidate genes is narrowing *Mod13* down to a smaller region of interest. Naturally occurring recombination events within *Mod13* were identified during SNP genotyping of members of the B6.Y^{POS} colony. Additional SNPs were sequenced (detailed in **Tables S1** in Appendix 1) to identify approximate breakpoints in these individuals. The recombinant alleles identified over a 2.5 year period are summarised in **Fig 5.1**. 40 XY mice were found to harbour recombination events within *Mod13(38-76)*: 21 phenotypic males and 19 phenotypic females. Multiple animals carried independently-arising alleles with similar approximate breakpoints.

Given the effect of *Mod13* is not fully penetrant, direct phenotype-genotype correlations are difficult; however, a trend is certainly noticeable and informative. The majority of alleles found in males contain larger stretches of non-B6 sequences, compared to the alleles in sex-reversed females where only small segments of non-B6 SNPs remain. Examination of regions of *Mod13* that are mostly retained in males and are mostly lost in females, identified a 10 Mb speculative "core" *Mod13*. Full analysis of each allele would be required to confirm its effectiveness at protecting against sex reversal. Unfortunately time and resources were too limited to allow the breeding of multiple recombinant colonies. Therefore a single male, whose non-B6 region spanned most of the

"core" *Mod13*, was selected for further breeding and analysis. The selected allele was named *Mod13(60-71)*, referring to position of outermost SNPs used for genotyping. The *Mod13(60-71)* founder male was born in early 2016. He was successfully back-crossed to generate a sub-colony for adult and embryo analyses.

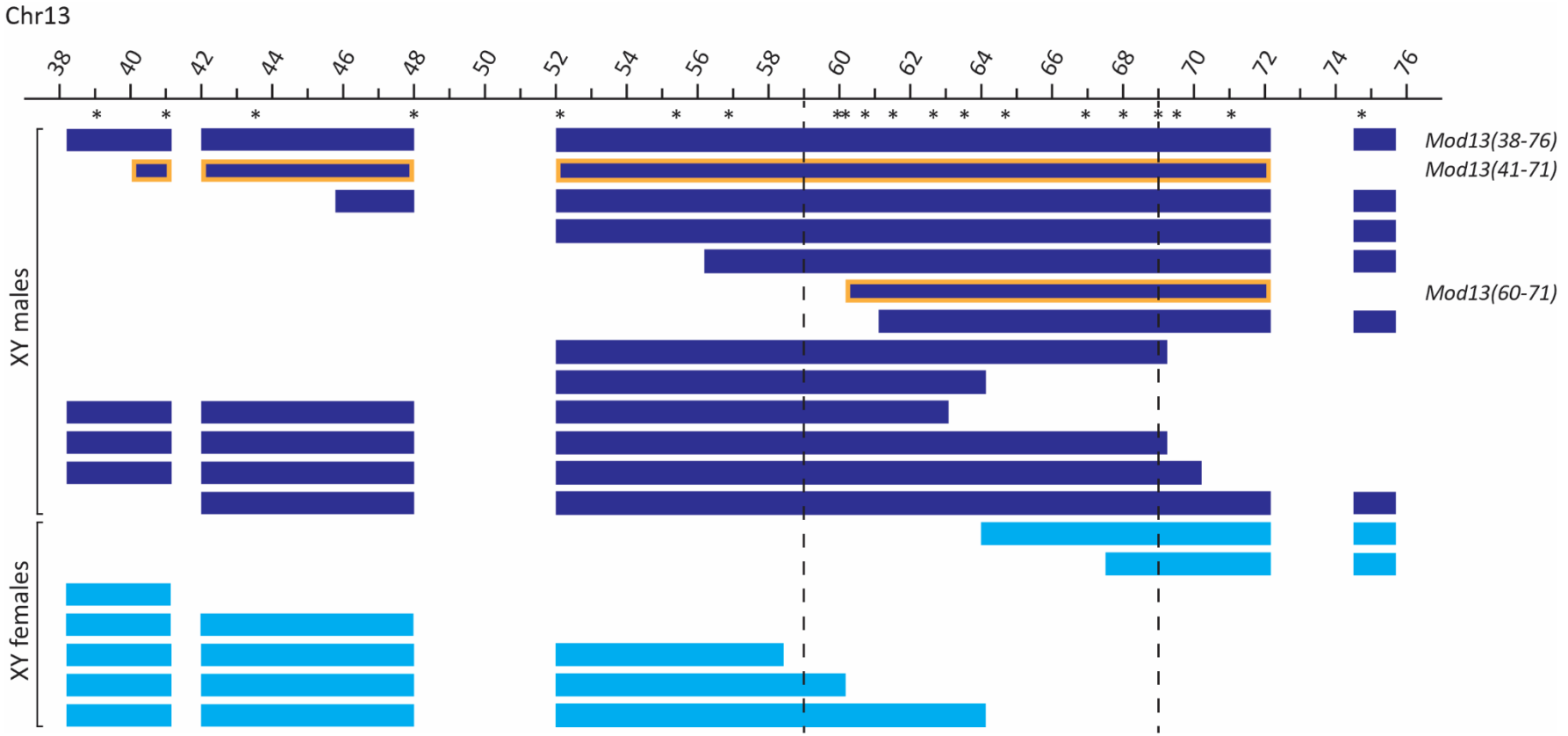


Figure 5.1 Recombinant alleles of *Mod13*. Schematic representing the recombinant alleles that arose in the B6.Y^{POS} colony over a 2.5 year period, coloured by phenotypic sex associated with that allele: dark blue = XY males, light blue = XY females. *Mod13(38-76)* denotes the original region of SNPs identified during genomic analysis of the sequenced stud male. Asterisks indicate position of SNPs sequenced for genotyping recombinant animals. When the same allele arose independently in multiple mice, the allele was grouped according to whichever sex contained the allele most often. Alleles found in equal numbers of males and females were excluded from this figure. For the purposes of this schematic, breakpoints are assumed to be located equidistant between genotyped SNPs. Dashed lines surround a proposed "core" *Mod13* region of interest from 59 Mb – 69 Mb. Yellow outline indicates those alleles that have been analysed in depth in this thesis.

5.2.3. Analysis of *Mod13(60-71)*

Mice in the *Mod13(60-71)* colony were genotyped using four of the distal SNPs used for *Mod13(41-71)*, and an additional SNP at position 60 Mb. As with *Mod13(41-71)*, genotype was correlated to phenotypic sex, as classified at weaning. The proportion of males and females in B6/B6 and B6/mod mice was nearly identical between the original 41-71 and the recombinant 60-71 allele **(Fig 5.2A)**, giving Fisher's exact *p*-values of zero when comparing sex ratios in B6/B6 homozygotes from each colony and when comparing sex ratios in heterozygotes from each colony.

Mod13(60-71) heterozygotes were inter-crossed to generate embryos at 14.5 dpc. Gonads were categorised using the same criteria as described in chapter 4. No significant differences were found in distribution of gonad classification between the two alleles, across all genotypes **(Fig 5.2B)**. B6.XY^{POS}.*Mod13(60-71)* gonads were again analysed by WMISH for expression of *Sox9*, *Stra8* and *Insl3* **(Fig 5.3)**. The results were highly comparable to those from the *Mod13(41-71)* gonads. Gonads with one or two copies of *Mod13(60-71)* expressed *Sox9* and *Insl3* (markers of testicular Sertoli cells and Leydig cells respectively) in larger areas of the gonad, and expressed *Stra8* (a marker of meiotic germ cells found in ovarian tissue) in smaller areas, indicative of the development of more testicular tissue. The improved quality of testis cord formation was again evident in heterozygous and homozygous mod gonads. These data show that *Mod13(60-71)* provides the same protection against sex reversal as the full *Mod13(41-71)*,

and suggest that the testis-determining gene(s) present in *Mod13* is located between 60 Mb and 71 Mb of chromosome 13.

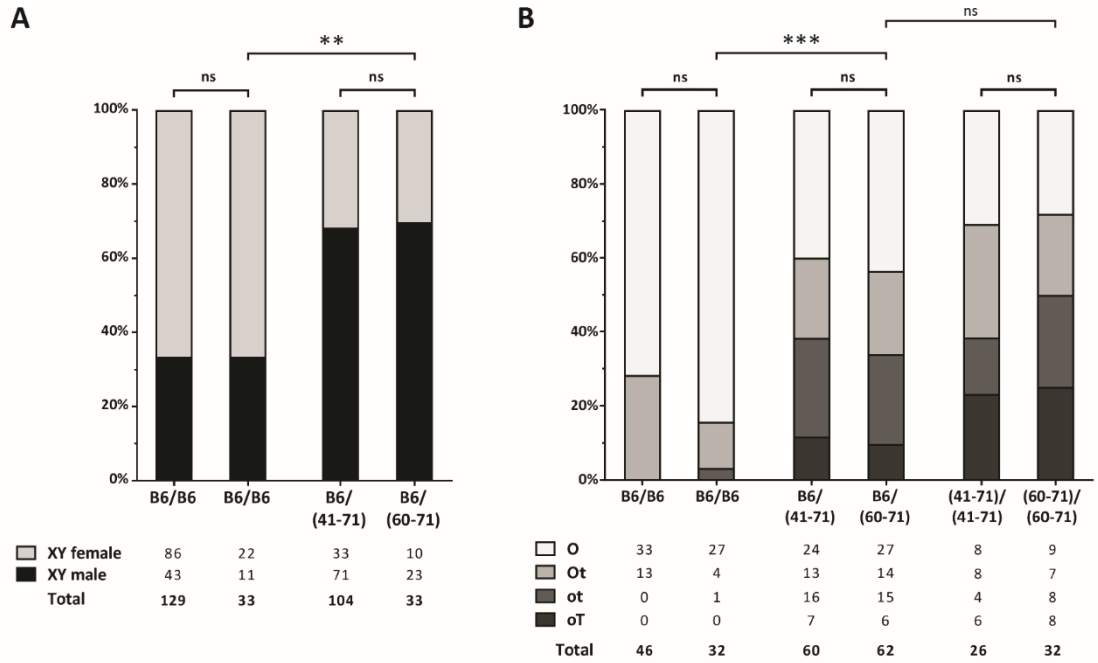


Figure 5.2 Comparison of *Mod13(41-71)* and *Mod13(60-71)*. No significant difference was found between the two *Mod13* alleles. **(A)** Ratio of B6.XY^{POS} males and females scored at weaning by anogenital distance, grouped by genotype of chromosome 13 region. Fisher's exact *p*-values: B6/B6:B6/B6 $p = 1$, B6/*Mod13(41-71)*:B6/*Mod13(60-71)* $p = 1$, B6/B6:B6/*Mod13(60-71)* $p = 0.0063$. **(B)** Classification of XY gonads at 14.5 dpc, grouped by genotype of chromosome 13 region. Fisher's exact *p*-values: B6/B6:B6/B6 $p = 0.09541$, B6/*Mod13(41-71)*:B6/*Mod13(60-71)* $p = 0.9728$, *Mod13(41-71)*/*Mod13(41-71)*: *Mod13(60-71)*/*Mod13(60-71)* $p = 0.8227$, B6/B6:B6/*Mod13(60-71)* $p < 0.001$, B6/B6:*Mod13(60-71)*/*Mod13(60-71)* $p < 0.0001$, B6/*Mod13(60-71)*:*Mod13(60-71)*/*Mod13(60-71)* $p = 0.2116$.

Figure 5.3 Gene expression analyses in B6.XY^{POS}.*Mod13(60-71)* gonads. WMISH analysis was performed on B6.XY^{B6}, B6 XX and B6.XY^{POS} gonads at 14.5 dpc for expression of *Sox9* **(A)**, *Stra8* **(B)** and *Insl3* **(C)**. B6.XY^{POS} B6/mod and mod/mod gonads contained larger regions of testicular tissue (indicated by *Sox9* and *Insl3* expression) and smaller regions of ovarian tissue (indicated by *Stra8* expression) than B6/B6 gonads. These results are comparable to those from analyses performed on B6.XY^{POS}.*Mod13(41-71)* gonads, found in Fig 4.2. Scale bars 500 µm.

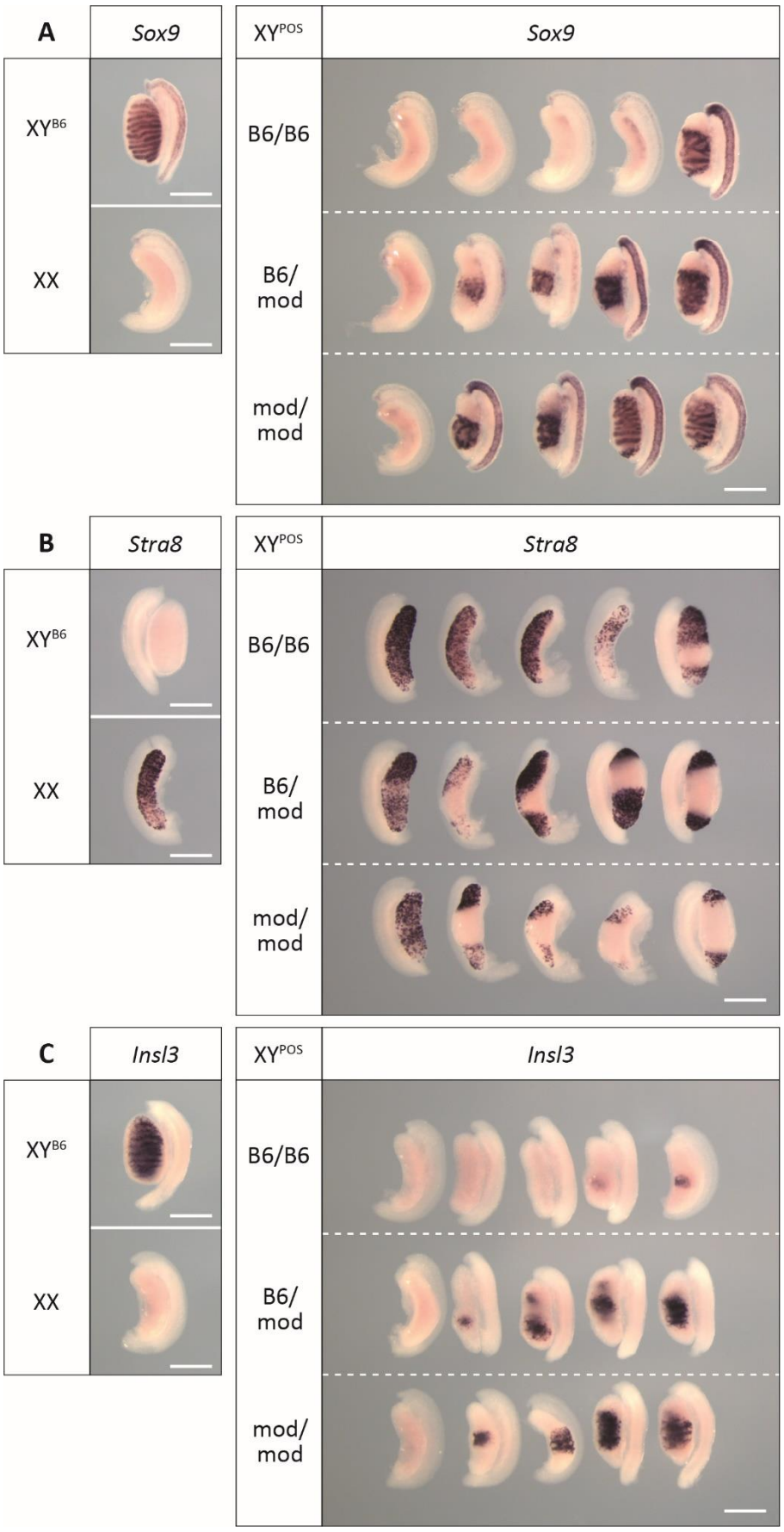


Figure 5.3 Gene expression analyses in B6.XY^{POS}.*Mod13(60-71)* gonads. For full figure legend, please see the previous page.

5.5.3. Candidate genes

A comprehensive review of the genes in *Mod13* was carried out, focussing on the genes between 60 Mb and 71 Mb. The review included examination of published literature and use of knockout data available from the International Mouse Phenotyping Consortium (IMPC). In addition, gonadal expression data was available from microarray analysis (Jameson et al., 2012b) and scRNA-seq analyses (Stevant et al., 2018). Genes expressed in at least one sex in the supporting cell lineage at 11.5 dpc (i.e. with a median expression level above 1 across all cells of that sex and lineage, according Stevant et al. scRNA-seq data) were selected. Using the functional criteria described in 5.1.3, these genes were then shortlisted. Details of the top 8 candidate genes identified by these methods are listed in **Table 5.3**, and expression profiles of each of those 8 genes in progenitor cells and the supporting cell lineage of the developing gonad, generated from Stévant et al's scRNA-seq data, are shown in **Fig 5.4**. One candidate gene – *Adamts16* – is the focus of the next chapter of this thesis. A second candidate gene – *Ptch1* – was briefly investigated by examination of heterozygous knockout embryos and homozygous knockout embryos of its gonadal ligand, *Dhh*, described below.

Table 5.3 Candidate genes on chromosome 13 between 60 Mb and 71 Mb

Gene symbol	Gene name	Position	Protein function	Gonadal expression at 11.5 dpc	Knockout phenotype
Gas1	Growth arrest specific 1	60174405	Proliferation regulation, Hedgehog signalling co-receptor	Both sexes, higher in females	Partial post-natal lethality. Abnormal limb, eye and brain development
Dapk1	Death associated protein kinase 1	60601947	Serine-threonine kinase, cell death	Both sexes, decrease in (pre-)Sertoli cells after 11.5 dpc	Neurological and renal defects
Ptch1	Patched 1	63508328	Hedgehog signalling receptor, embryonic development, fetal Leydig cell specification	Both sexes, supporting lineage	Embryonic lethal with neural tube defects
Zfp367	Zinc finger protein 367	64133022	DNA binding	Both sexes, higher in females	Increased circulating creatinine
Cdc14b	CDC14 cell division cycle 14B	64192545	Tyrosine phosphatase, DNA repair	Both sexes, higher in somatic cells	Early-onset aging, sub-fertility
Med10	Mediator complex subunit 10	69802395	Transcriptional coactivator	Both sexes	Pre-weaning lethality
Ice1	Interactor of little elongation complex ELL subunit 1	70588689	snRNA transcription	Both sexes	N/a
Adamts16	A disintegrin-like and metallo-peptidase with thrombospondin type 1 motif, 16	70727802	Extracellular protease, cell proliferation and migration	Males, (pre-)Sertoli cells	Tremors. Male rats are sterile with cryptorchidism

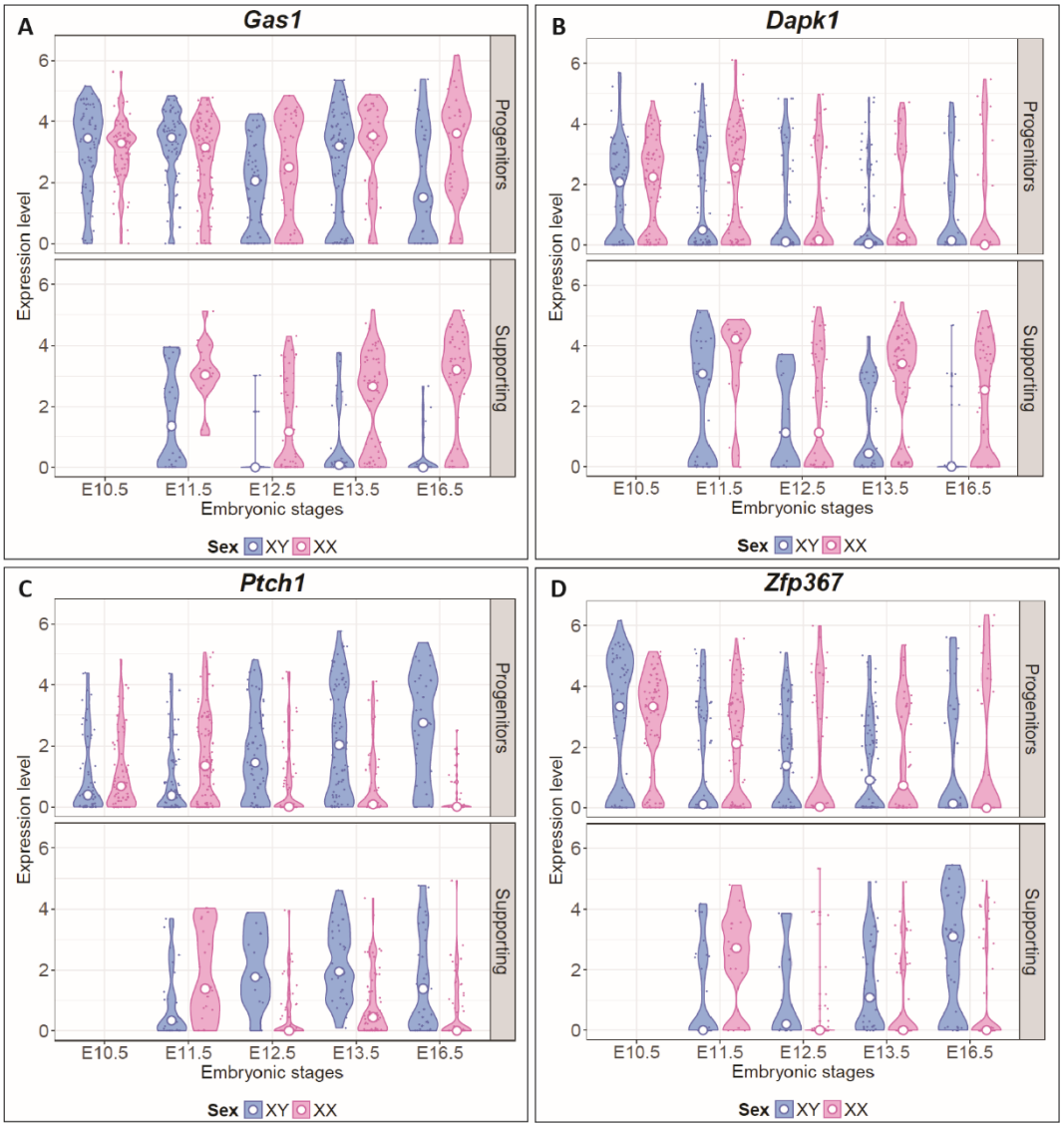


Figure 5.4 Expression profiles of candidate genes in *Mod13(60-71)*. Part 1. Single cell RNA seq data for 8 candidate genes, kindly provided by Isabelle Stevant (Stevant et al., 2018). Gene expression levels measured in single progenitor cells and supporting-lineage cells from 10.5 dpc to 16.5 dpc gonads. Violin plots show the expression value for each single cell (small dots) and the median of expression of all the cells for the given gene (large white dot). The width of the violin represents the proportion of cells that express the given gene at a certain level of expression. Figure continues on following page.

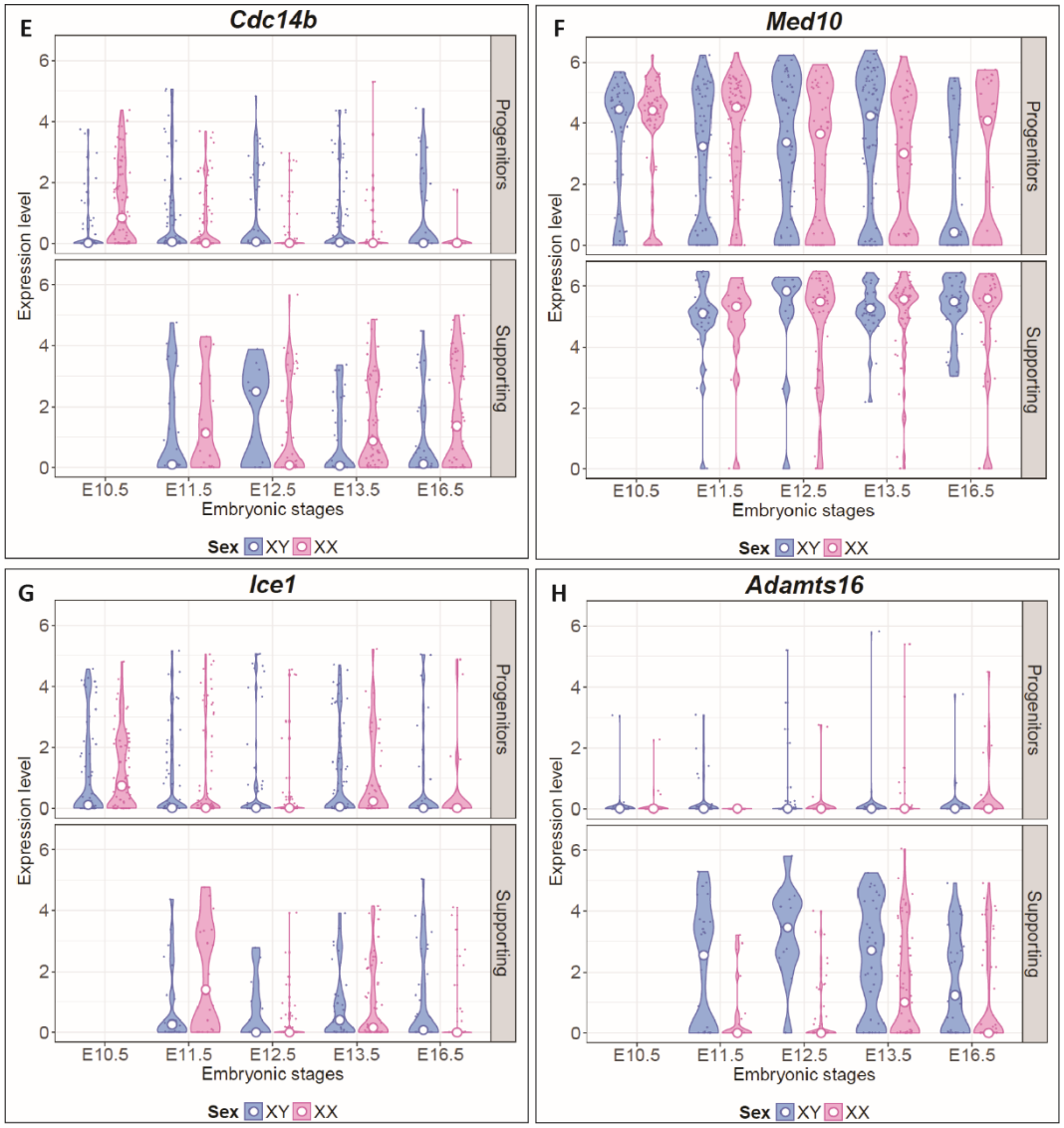


Figure 5.4 Expression profiles of candidate genes in *Mod13(60-71)*. Part 2. Continued from previous page.

5.5.3. The role of *Ptch1*/*Dhh*

The knockout lines of *Ptch1* and *Dhh* used were those previously published (Bitgood et al., 1996; Goodrich et al., 1997). However, for this investigation the effects of haploinsufficiency for *Ptch1* and loss of *Dhh* were analysed on congenic B6 genetic backgrounds. Because *Ptch1*$^{-/-}$ embryos die too early to analyse the effect of complete loss of *Ptch1* on gonad development, heterozygotes were generated with a Y chromosome from the strain AKR/J (the B6.Y^{AKR} genetic background). As described above, this conveys additional sensitivity to XY sex reversal. The aim of this investigation was to identify novel defects in sex determination not previously observed when using non-sensitised, mixed genetic backgrounds.

Embryos were generated by crossing *Ptch1*$^{+/-}$ B6 females to B6.Y^{AKR} males, and gonads were analysed by WMISH at 14.5 dpc **(Fig 5.5A)**. *Sox9* expression analysis showed no apparent differences in testis cord structure or specification of Sertoli cells, between heterozygotes and their controls. Interestingly, the development of the mesonephros appeared abnormal in heterozygotes. *Stra8* expression analysis indicated no sex reversal phenotype, beyond the ovarian-like tissue development at the poles of the gonad that is typical of B6.Y^{AKR} testes at this stage. These transient ovotestis-like gonads go on to form functioning adult testes of reduced size (Harris et al., 2018).

Embryos were generated by inter-crossing *Dhh*$^{+/-}$ mice, and gonads were analysed at 14.5 dpc. Homozygous gonads were noticeably smaller in width and more crescent shaped than wild-type controls. However, WMISH analysis **(Fig**

5.5B) showed that this reduction in size is likely caused by the loss of *Insl3*-expressing fetal Leydig cells; very view *Insl3*-positive cells are visible in homozygous testes, as expected from published literature (Yao et al., 2002). *Sox9*-expressing Sertoli cells were present throughout the gonad and correctly organised into regular testis cords. In conclusion, the use of sensitised genetic backgrounds did not reveal any novel roles for *Ptch1* or *Dhh* in testis determination.

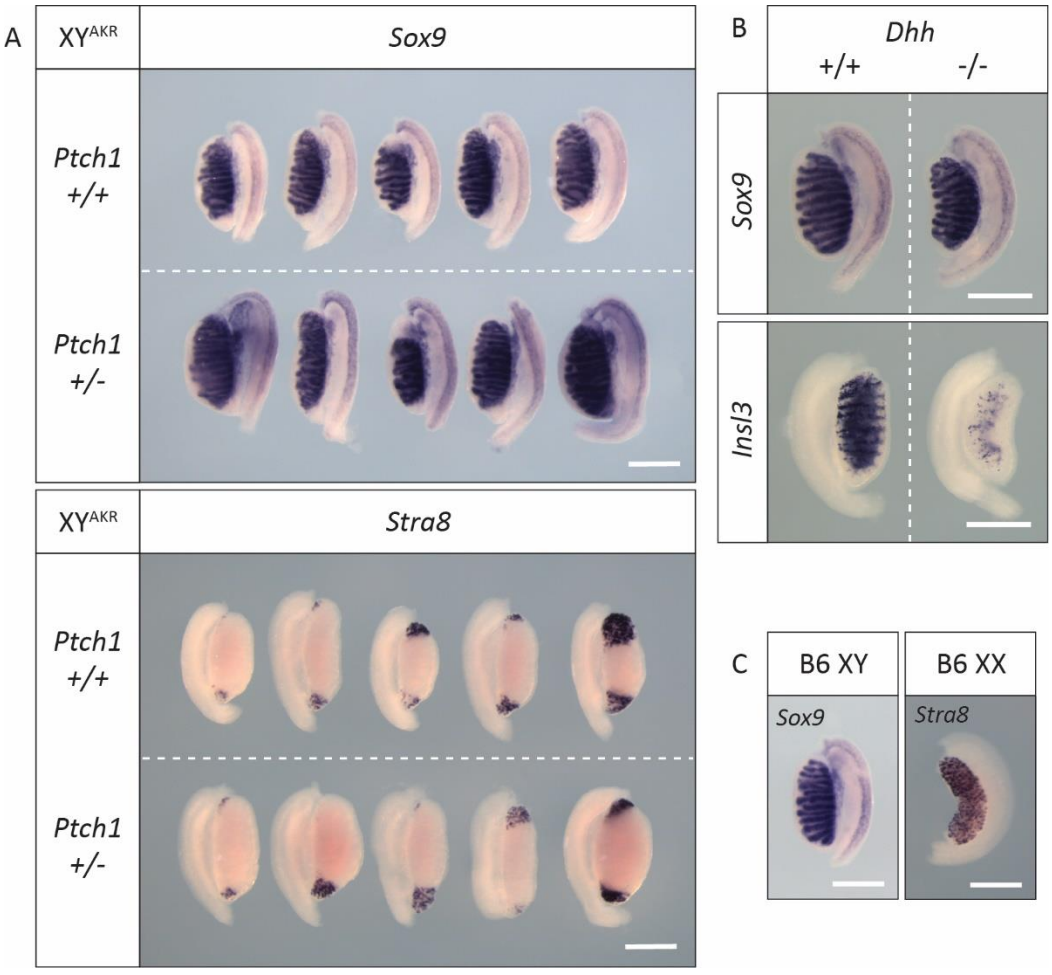


Figure 5.5 Analysis of gonads lacking members of the Hedgehog signalling pathway. WMISH analysis of *Ptch1*$^{+/-}$ gonads on a B6.Y^{AKR} genetic background **(A)**, *Dhh*$^{-/-}$ gonads on a B6.Y^{B6} genetic background **(B)** and wild-type B6 gonads **(C)** at 14.5 dpc. Testis cord morphology is normal in *Ptch1*$^{+/-}$ Y^{AKR} gonads, relative to *Ptch1*$^{+/+}$ Y^{AKR} controls. Small loss of *Sox9* expression at the poles, replaced by *Stra8* expression, is expected in B6.Y^{AKR} testes. Cord morphology and *Sox9* expression is normal in *Dhh*$^{-/-}$ gonads. Reduced size is caused by loss of fetal Leydig cells, shown by *Insl3* expression. *N* = 5 for **A** and 3 for **B** and **C**. Scale bar 500 µm.

5.3. Discussion

5.3.1. Summary

Examination of a recombinant allele of *Mod13* has narrowed the region of interest to 11 Mb of chromosome 13, down from 31 Mb. The effect on primary sex determination appears identical between the full-length allele and the reduced allele, indicating that the gene(s) responsible for upregulating *Sry* expression and downregulating *Rspo1* and *Foxl2* expression are likely to be located in this narrowed region, between positions 60 Mb and 71 Mb of chromosome 13. Review of the published literature and analysis of novel single-cell transcriptomic data has identified potential candidate genes, which will be described in more detail in chapters 6 and 7. The possibility that genes lying outside of the 60-71 Mb region might also have a role to play will also be discussed in chapter 7.

5.3.2. PTCH1/DHH signalling

A gene of interest located within the narrowed modifier region is *Ptch1*. PTCH1, along with its gonadal receptor, DHH, has a previously established role in Leydig cell specification (Clark et al., 2000; Yao et al., 2002). Re-examination of embryos lacking two copies of *Dhh* or one copy of *Ptch1* on a congenic sensitised genetic background did not reveal any new gonadal phenotypes. *Dhh* is not required for Sertoli cell differentiation and no evidence of additional, ectopic ovarian gene expression was observed in $Ptch1^{+/-}$ Y^{AKR} gonads when compared to wild-type controls. Interestingly, defects in mesonephros development were noted. *Ptch1*

expression has been reported in the anterior region of the mesonephric tubules in 11.5 – 13.5 dpc embryos (Yao et al., 2002).

46,XY humans with mutations in *DHH* have been reported to exhibit complete or partial gonadal dysgenesis, with evidence in some patients of complete disruption to seminiferous tubule development **(Table 5.1)**. This suggests a requirement for *DHH* in human Sertoli cell specification, which is contrary to the phenotype seen in *Dhh* null mice. One possibility is that the mutations in *DHH* were not completely causative in these patients, and that they each harboured another mutation in a testis-determining gene (or cumulative mutations in multiple genes). An alternative explanation is that *PTCH1/DHH* signalling is not required for Sertoli cell specification, but the absence of Leydig cells in the patients' testes caused a cascade effect leading to complete degeneration of the seminiferous tubules. Pierucci-Alves et al. (2001) reported that *Dhh* null mice have defects in the basal lamina that surrounds the developing testis cords, an additional explanation for the gonadal dysgenesis seen in patients. Alternatively, these discrepancies may represent a genuine difference between the two species in respect of DHH function.

Further studies of PTCH1/DHH signalling in Sertoli cell specification might involve gene expression analyses in $Ptch1^{+/-}$ and $Dhh^{-/-}$ gonads at 11.5 dpc. If expression levels of *Sry, Sox9* or pro-ovarian genes remained unchanged in these gonads it would confirm that Hedgehog signalling plays no role in primary commitment of gonadal cells to the testicular or ovarian fate. Conclusively ruling out a role of *Ptch1* would require use of a gonad-specific conditional knockout of *Ptch1*, to

circumvent the embryonic lethality of *Ptch1*$^{-/-}$ embryos. While B6.Y^{AKR} provides the genetic background most sensitised to XY sex reversal, the heterozygote analysis performed here is less ideal than examining gonads completely lacking expression of *Ptch1* in order to completely understand its role in testis development. Given the lack of evidence for *Ptch1* playing a role in primary sex determination, it is unlikely to be the gene responsible for the testis-promoting effect of *Mod13*. The search for candidate genes continues.

Chapter 6: The role of *Adamts16* in the mouse urogenital system

6.1. Introduction

6.1.1. The ADAMTS family

ADAMTS16 is a member of the ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) family of secreted, zinc-binding, extracellular enzymes. There are 19 ADAMTS proteins encoded by mammalian genomes, which can be grouped on the basis of their domain organisation and known substrates, as illustrated in **Fig 6.1** (Kelwick et al., 2015). The general structure of ADAMTS proteins includes a propeptide domain, a metalloproteinase M12B domain, a disintegrin-like domain, variable numbers of thrombospondin type 1 motifs and a cysteine rich domain (Porter et al., 2005). They contribute to extracellular matrix (ECM) remodelling and have known functions in cleavage of pro-collagen and extracellular proteoglycans such as aggrecan, versican, brevican and neurocan (Stanton et al., 2011; Wang et al., 2003). Additional roles of ADAMTS proteins include anti-angiogenesis (Vazquez et al., 1999), inflammation (Worley et al., 2003) and fertility. ADAMTS1 is expressed is multiple tissues in the developing embryo, including in ovarian granulosa cells (Thai and Iruela-Arispe, 2002). *Adamts1* null female mice are infertile, as well as being smaller than normal, having severe kidney degeneration and adrenal abnormalities (Shindo et al., 2000). Affected ovaries contain fewer mature follicles than in wild-type females and uterine horns contain cysts. Loss of ADAMTS2, a procollagen I N-proteinase (Wang et al., 2003), causes skin fragility and infertility in males due to deficiency in spermatogenesis (Li et al., 2001).

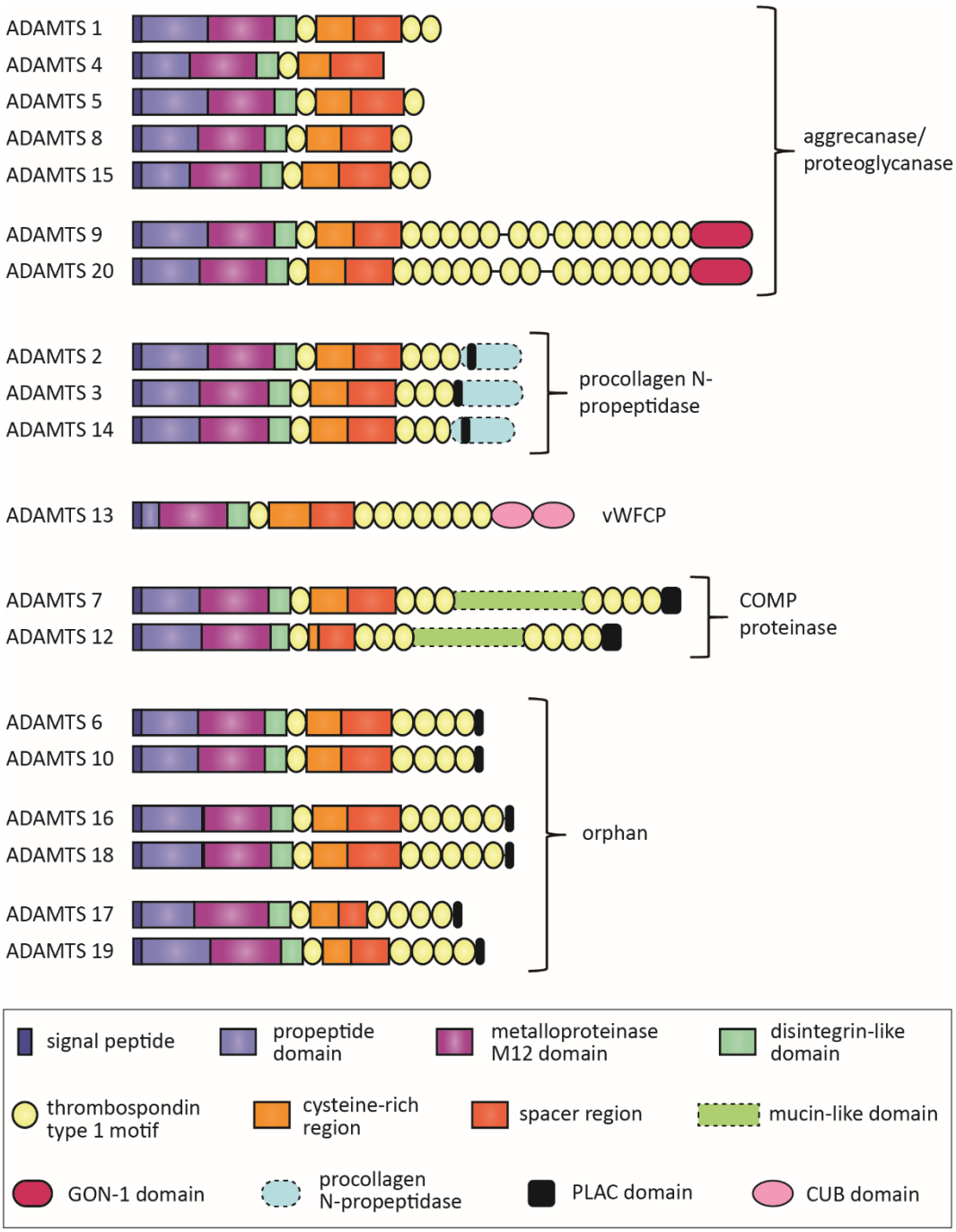


Figure 6.1 The ADAMTS family. The 19 human ADAMTS proteins are grouped on the basis of their domain organisation and known substrates. They are numbered 1 to 20 (*ADAMTS5* and *ADAMTS11* were discovered to be the same gene). They share a common N-terminal domain structure that includes the zinc-binding M12B proteinase domain, but vary more at the C-terminus. Several ADAMTS proteins have additional modules, for example ADAMTS13 is unique in having two CUB domains, a motif found in other extracellular proteins and associated with embryogenesis and organogenesis. COMP proteinases = cartilage oligomeric proteins (i.e. proteoglycans); vWFCP = von-Willebrand factor cleaving protease. Figure adapted from Kelwick et al. (2015).

6.1.2. ADAMTS16 expression and function

The precise molecular function of ADAMTS16 is unknown, with little information available about its physiological substrates and mechanism of action. There is *in vitro* evidence that ADAMTS16 is secreted and processed (Schnellmann et al., 2018), and that it has proteolytic ability. Recombinant protein is capable of cleaving α_2-macroglobulin, a protein component of blood plasma (Gao et al., 2007) and overexpression of ADAMTS16 in cells leads to increased fibronectin cleavage (Schnellmann et al., 2018). Fibronectin, a large glycoprotein of the ECM, and its proteolysis and turn-over are important for ECM remodelling and morphogenesis. This suggests that ADAMTS16 may function in biological processes similar to other members of the protein family.

ADAMTS16 has been associated with multiple pathologies – hypertension in humans and rats (Gopalakrishnan et al., 2012; Joe et al., 2009), renal abnormalities in rats and mice (Gopalakrishnan et al., 2012; Jacobi et al., 2013), osteoarthritis (Davidson et al., 2006), oesophageal squamous cell carcinoma (Sakamoto et al., 2010), and, most recently, disorders/differences of sex development (Barseghyan et al., 2018b). Expression analyses found *ADAMTS16* in human adult brain, kidney, pancreas and ovary, and fetal lung and kidney (Cal et al., 2002; Gao et al., 2007). Ovarian expression is specific to the granulosa cells of pre-ovulatory follicles, and can be stimulated by follicle-stimulating hormone (Gao et al., 2007). In rodents, *Adamts16* is expressed in fetal gonads of both males and females, in spermatids of adult males and in follicular granulosa cells in the ovaries of adult females (Jacobi et al., 2013).

Microarray analysis shows *Adamts16* is elevated in XY fetal gonads but not in XX fetal gonads (Jameson et al., 2012b). Single-cell RNA-seq data, shown in **Fig 5.4H** of the previous chapter, indicate that *Adamts16* is expressed in the supporting cell lineage of the early XY gonad. WMISH analysis was performed on wild-type B6 gonads at multiple embryonic time points **(Fig 6.2)**. It was confirmed that *Adamts16* is expressed at 11.5 dpc in cells of the XY gonad, but not the XX gonad. Expression level is increased in XY gonads at 12.5 dpc. At 13.5 and 14.5 dpc it is clear that *Adamts16* expression is restricted to the developing testis cords, where the supporting Sertoli cells and germ cells are located. Some expression is seen in the ovary at 14.5 dpc. In agreement with published literature, expression was also observed in the tubule of the kidney at 14.5 dpc.

The gonadal expression pattern of *Adamts16* is typical of a gene upregulated by SRY or SOX9 during testis determination. ChIP-microarray data indicate that SRY binds to the promotor of *Adamts16* in the gonad at 11.5 dpc and SOX9 binds to the promotor at 12.5 dpc (Li et al., 2014). Additionally, *Adamts16* was identified as a direct downstream target of WT1, a transcription factor with a known role in gonad development and testis determination, and the two genes are co-expressed in cells of the developing urogenital system (Jacobi et al., 2013). The observation that *Adamts16* is expressed in a sexually dimorphic fashion in supporting cells of the developing gonad at the time of sex determination (11.5 dpc), along with its regulatory links to molecules known to function in testis determination, suggest that it is an excellent candidate testis-determining gene in mice.

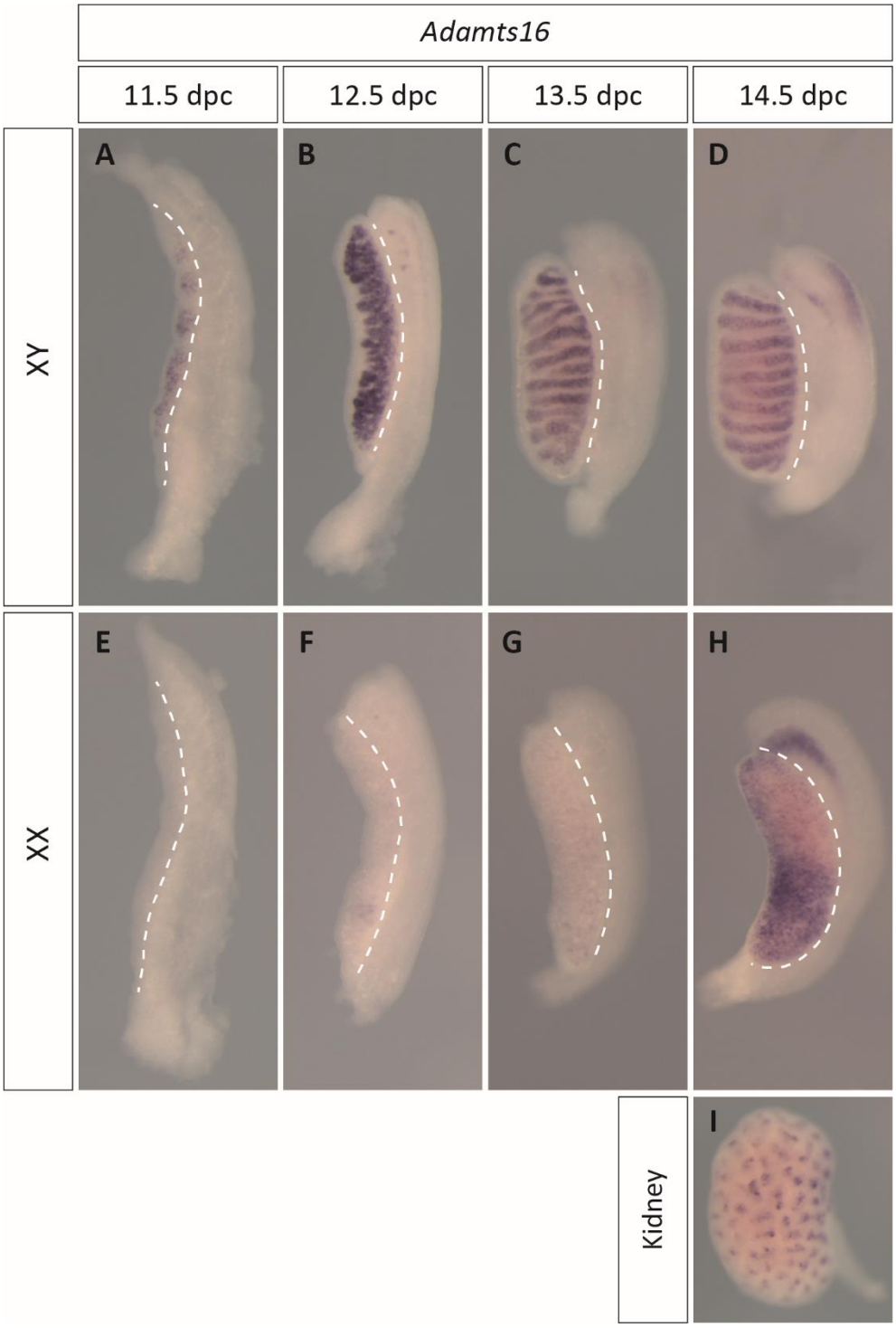


Figure 6.2 Expression of *Adamts16* in the embryonic gonad and kidneys. WMISH analysis of *Adamts16* in gonads between 11.5 dpc and 14.5 dpc **(A-H)** and kidney at 14.5 dpc **(I)**. *Adamts16* is detected at all ages studied in the XY gonad (first row), but only at 14.5 dpc in the XX gonad (second row). Dotted line indicates the gonad-mesonephros boundary, with the gonad on the left. *Adamts16* is expressed in the tubules of the kidney at 14.5 dpc. Data courtesy of Nicolas Chalon.

6.1.3. Cryptorchidism and sterility in *Adamts16* knockout rats

Other compelling evidence for *Adamts16* having an important role in testis development and function comes from a loss-of-function allele of *Adamts16*, generated in rats using a deletion in the first exon (Abdul-Majeed et al., 2014). Homozygous male rats were sterile with bilateral cryptorchidism. Testes were located in the lower abdomen and were significantly smaller than controls, with normalised testes weights measuring a third of age-matched wild-types and heterozygotes at 12 weeks of age. The phenotype was characterised by a progressive degeneration of seminiferous tubule integrity, interstitial hyperplasia and a complete lack of spermatogenesis by 12 weeks of age. The authors hypothesised that ADAMTS16 is required for ECM remodelling, which is needed for testicular descent, to explain why loss of the proteinase leads to cryptorchidism. They also hypothesised a direct role of ADAMTS16 in testicular ECM and its support of spermatogenesis. However the study did not investigate how loss of *Adamts16* impacted embryonic gonad development. Disruption to testis determination could explain the degeneration of testis structures and subsequent sterility.

6.1.4. Descent of the testes

The gonads of both sexes originate and differentiate in the abdomen, next to the developing kidney. During development, the ovaries maintain their position, while the testes descend into the scrotum. This process is essential for male reproductive health. Failure of one or both testes to descend to the scrotum (a

condition known as cryptorchidism) occurs in 3% of full-term new-born males and 21% of premature males. While most cases of undescended testes will rectify themselves within the first 9 months after birth, approximately 1% of boys require medical or surgical intervention to locate the testes and fix them in the correct scrotal position. This need for surgery (known as an orchidopexy) is due to cryptorchidism being strongly linked to infertility and germ cell tumours. Bilateral cryptorchidism results in reduced or loss of fertility if the testes remain undescended beyond puberty (Trussell and Lee, 2004).

The general consensus describes the descent of the testes as occurring in two distinct phases that require different signals from the testes **(Fig 6.3)** (Hutson and Hasthorpe, 2005). The first phase – the transabdominal phase – is androgen-independent. The second phase – the inguinoscrotal phase – is androgen-dependent. Two ligaments are responsible for gonad positioning and undergo sexually dimorphic remodelling during development. The genitoinguinal ligament, known as the gubernaculum, connects the caudal pole of the gonad to the lower abdominal wall (Heyns, 1987). The cranial suspensory ligament (CSL) connects the cranial pole of the mesonephros to the dorsal abdominal wall (Van der Schoot and Emmen, 1996).

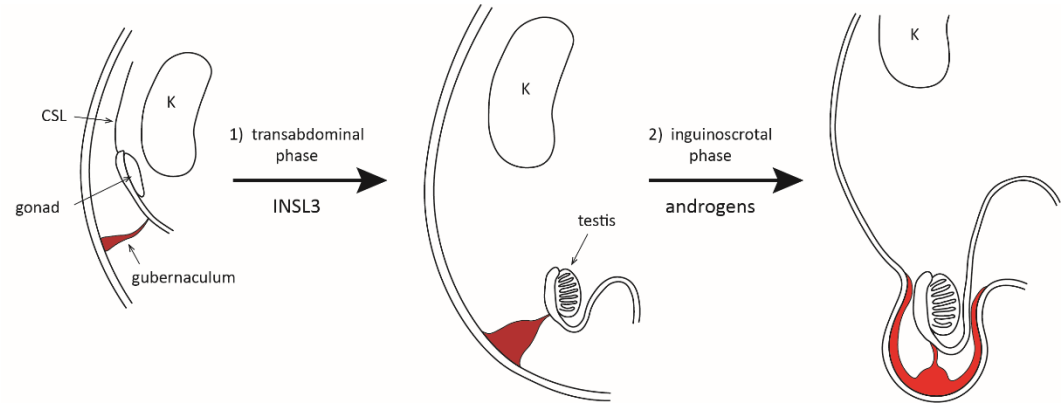


Figure 6.3 Testis descent in rodents. The first phase of testis descent – the transabdominal phase – requires the insulin-like hormone INSL3 and occurs between 14.5 dpc and 16.5 dpc in mice. The gubernaculum swells, anchoring the testis to the abdominal wall as the abdominal cavity grows. In the second phase of testis descent – the inguinoscrotal phase – the gubernaculum undergoes remodelling: it everts and migrates through the inguinal canal, pulling the testis into the scrotum. This step is androgen-dependent and occurs in the first two weeks after birth in mice.

6.1.4.1. Transabdominal descent

The transabdominal phase of descent requires the expression of *Insl3*, which encodes an insulin-like hormone secreted by the fetal Leydig cells. INSL3 binds to its receptor LGR8 (also known as GREAT or RXFP2), which is expressed on cells of the gubernaculum (Bogatcheva et al., 2003). Male mice lacking either *Insl3* or *Lgr8* have bilateral cryptorchidism, with testes located high in the abdomen (Gorlov et al., 2002; Nef and Parada, 1999). Mutations in both *INSL3* and *LGR8* have been found in human patients with cryptorchidism (Ferlin et al., 2003). Overexpression of *Insl3* in female embryos leads to descent of the ovaries (Adham et al., 2002).

INSL3/LGR8 signals cause the gubernaculum to undergo enlargement in a process known as the swelling reaction (Adham et al., 2002; Emmen et al., 2000). The gubernaculum is embedded in the abdominal wall and tethers the gonad

close to the future inguinal canal. As the abdominal cavity of the embryo grows, the gubernaculum remains attached to the inguinal wall, anchoring the testes in place. Simultaneously, androgens produced by the fetal testis cause the regression of the CSL (Van der Schoot and Elger, 1992). In female embryos, the gubernaculum does not swell and simply elongates, providing no tension to anchor the ovary. The CSL persists, holding the ovary in its high, para-renal position.

6.1.4.2. Inguinoscrotal descent

The inguinoscrotal phase of testis descent involves complex remodelling, reorientation, proliferation, migration and contraction of the gubernaculum (Nation et al., 2011). In rodents, the gubernaculum initially points cranially, towards the abdomen. The gubernaculum then undergoes eversion through the inguinal canal, causing it to point caudally. Mesenchymal cells in the gubernacular tip proliferate extensively and the gubernaculum migrates through the inguinal canal and into the scrotum, pulling the testes downwards. This process has been likened to the growth of a limb bud (Hutson et al., 2013).

The inguinoscrotal phase is androgen-dependent (Hutson, 1986), and conditional ablation of androgen receptors in the gubernaculum of mice leads to a failure of the gubernaculum to undergo eversion, resulting in cryptorchidism (Kaftanovskaya et al., 2012). However, the precise molecular mechanism of androgen action in the gubernaculum remains unclear. Transcriptional regulation by canonical/β-catenin WNT signalling has been proposed to mediate the androgen signals, leading to mesenchymal cell migration and secretion of matrix

metalloproteinases needed for extracellular matrix remodelling (Chen et al., 2011). Additionally, expression of *Hoxa-11* has been shown to be required for proliferation of mesenchymal cells in the gubernacular tip, possibly independent of androgens (Harisis et al., 2013).

6.1.5. Human mutations in *ADAMTS16*

More recently, evidence has been reported of a role for ADAMTS16 in human testis determination. Exome sequencing of 46,XY DSD patients identified variants in *ADAMTS16* that are found in 0.1% or less of the general population (Barseghyan et al., 2018b). Heterozygous missense mutations were found in two patients with 46,XY complete gonadal dysgenesis (i.e. XY female presentation). One patient had an Arg to Trp change in the propeptide domain and another patient had a Val to Ile change in the cysteine rich domain. A third heterozygous missense mutation was found in a 46,XY patient with ambiguous genitalia at birth. The resulting Phe to Val change in the metalloproteinase domain was predicted to impair proteolytic function. Interestingly, the authors cross-referenced the list of genes found to contain variants in 46,XY DSD patients to a list of genes differentially expressed in B6.Y^{POS} gonads at 11.5 dpc, which also identified *Adamts16* as a gene of interest due to its lowered expression in gonads with disrupted testis development.

6.1.6. Aims

Adamts16 is located at position 70 Mb of mouse chromosome 13, within the region of *Mod13* responsible for promoting testis determination. *Adamts16* is expressed in pre-Sertoli cells at the time of testis determination and mutations in

human patients and in rats suggest a requirement for functional ADAMTS16 in testis development and function. All this points to *Adamts16* playing an important role in testis determination, thus making it an excellent candidate gene for the role of *Mod13*. The aims of this chapter are to generate a null allele of *Adamts16* in mice, and to study the impact of this loss of function on the urogenital system of embryos and adults.

6.2. Results

6.2.1. Generation of a null *Adamts16* allele using CRISPR/Cas9

A knockout allele of *Adamts16* was generated by the Molecular and Cellular Biology team at MRC Harwell. *Adamts16* consists of 23 exons. The critical metallopeptidase M12B domain is encoded by exons 5 through 10. Cas9 RNA and guide RNA was microinjected into 1-cell C57BL/6NTac embryos, targeting exon 5 **(Fig 6.4A)**. A founder male was identified carrying two different deletions, one of 906 nucleotides and one of 923 nucleotides. Both alleles result in a complete loss of exon 5; splicing from exon 4 to exon 6 is predicted to result in a frameshift followed by a premature stop codon. Such founders generated by CRISPR/Cas9 are commonly mosaic and so allele transmission was not guaranteed. The founder was bred and both alleles were transmitted to the next generation. Offspring carrying the 923 nucleotide deletion were bred to produce an adult cohort, which was analysed as part of the International Mouse Phenotyping Consortium (IMPC). Offspring carrying the 906 nucleotide deletion were bred for the analyses described in this chapter.

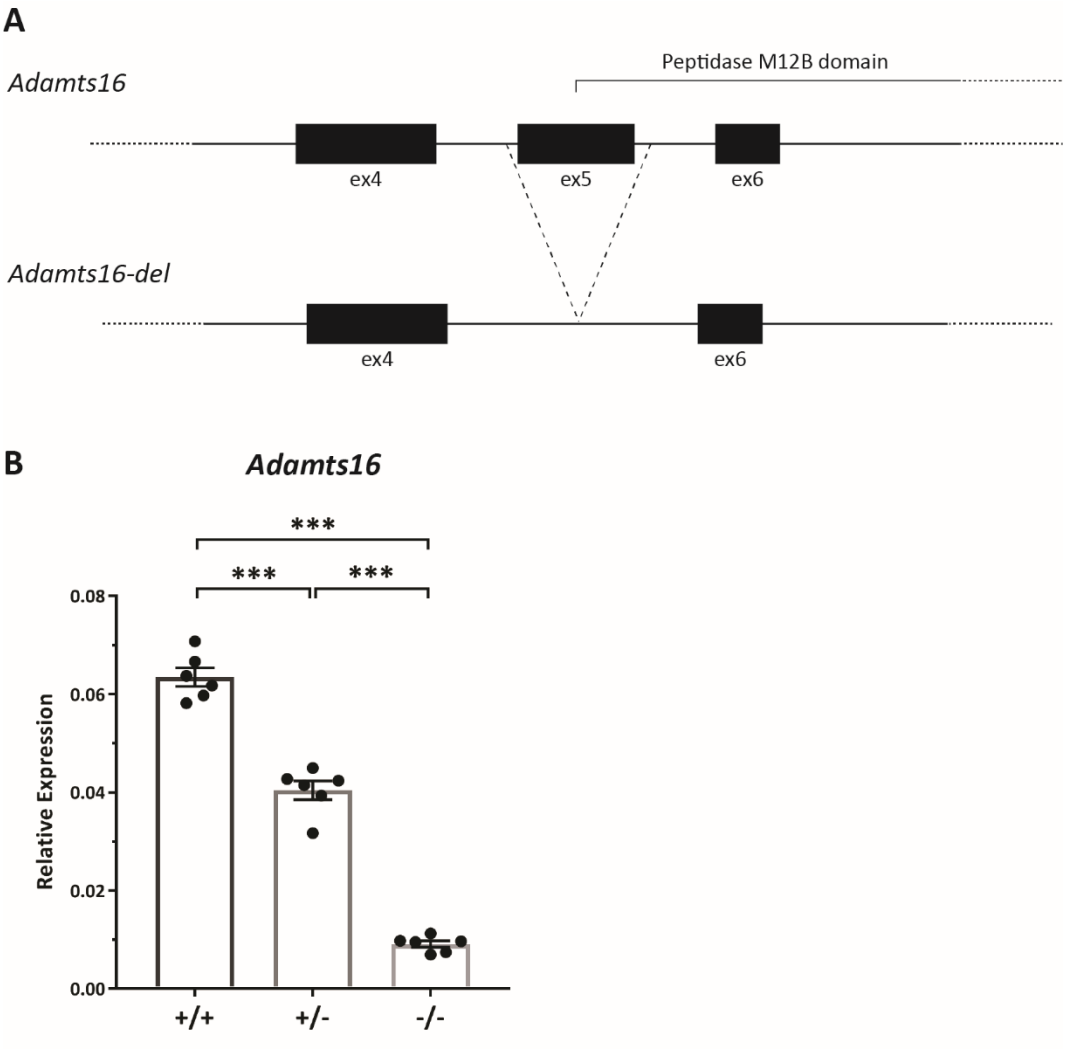


Figure 6.4 Generation of a null allele of *Adamts16*. (A) *Adamts16-del* allele, generated by CRISPR-Cas9, contains a targeted deletion of exon 5, a critical exon for the peptidase domain. Deletion causes a frameshift followed by a premature stop codon. **(B)** *Adamts16* mRNA levels in 14.5 dpc kidneys measured by qRT-PCR. Expression was calculated relative to *Hprt1*. Two-tailed Welch's t-tests were performed. *P*-value <0.0001 across all groups. Error bars show SEM.

Heterozygote males and females were normal and fertile. Heterozygotes were intercrossed to generate homozygous embryos. Levels of *Adamts16* transcript were measured by qRT-PCR in 14.5 dpc kidneys. Homozygote embryos expressed *Adamts16* at negligible levels, a 7-fold reduction compared to wild-types **(Fig 6.4B)**. Expression levels in heterozygote embryos were intermediate between wild-types and homozygotes. This suggests that the mutant mRNA undergoes nonsense-mediated decay. I infer from this that the *Adamts16-del* allele is a null.

Homozygous embryos were examined from 13.5 dpc to 16.5 dpc. No developmental defects were observed **(Fig 6.5)**. Heterozygotes were intercrossed to generate a cohort of homozygous adults for phenotyping. Homozygotes were viable and born in the expected ratios.

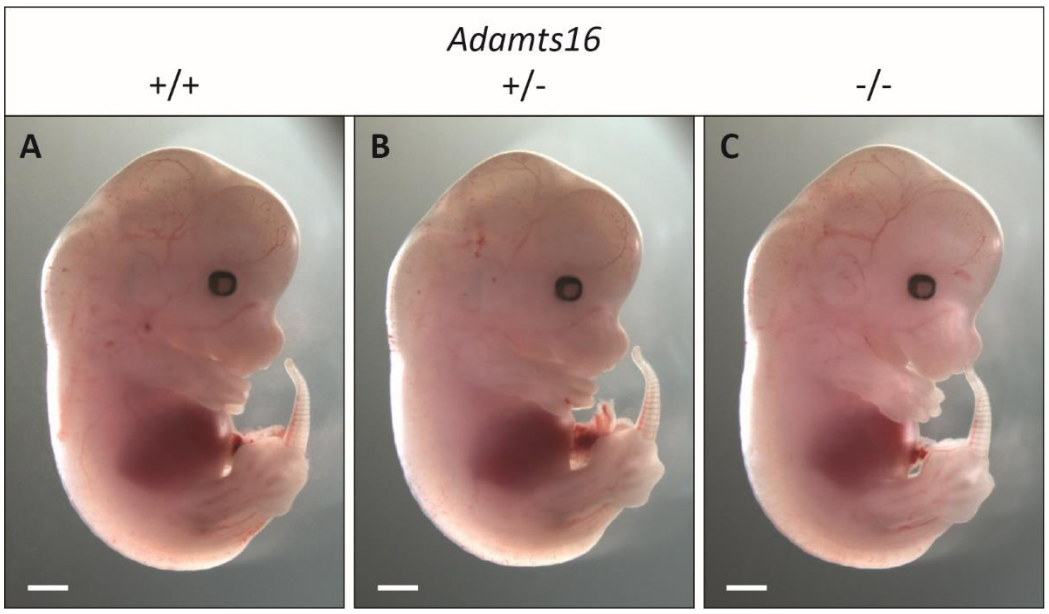


Figure 6.5 No gross developmental defects in *Adamts16* null embryos. 13.5 dpc littermate embryos imaged on a Leica MZ16F with attached camera. No developmental defects were observed at any embryonic ages examined (14.5 – 16.5 dpc not shown). Scale bar 1 mm.

6.2.2. The role of *Adamts16* in testis development

XY gonads from 14.5 dpc embryos were collected and analysed by WMISH **(Fig 6.6)**. Homozygous testes appeared normal; Sertoli cells (identified by strong expression of *Sox9*, **Fig 6.6A**) were arranged in testis cords throughout the gonad. No abnormalities were observed in the fetal Leydig cells (identified by expression of *Insl3*, **Fig 6.6B**), which occupied the interstitium as expected. No *Stra8* expression was observed **(Fig 6.6C)** indicating that there was no ectopic meiotic entry of germs cells, i.e. ovarian development. The only difference detected between homozygotes and the wild-type and heterozygote controls was an apparent delay in testis development; homozygote testes were on average 12% wider and 10% longer. The average number of externally visible testis cords in homozygote testes was 12, compared to 10 in wild-type testes. The testis structure seen in homozygotes at 14.5 dpc is typical of younger (12.5 – 13.5 dpc) embryos (Combes et al., 2009). This could suggest an overall delay in embryonic development, although no delay in growth was observed between littermates **(Fig 6.5)**. No reduction in testis size was observed at 16.5 dpc **(Fig 6.9D)**, suggesting that the delay in testis development is transient.

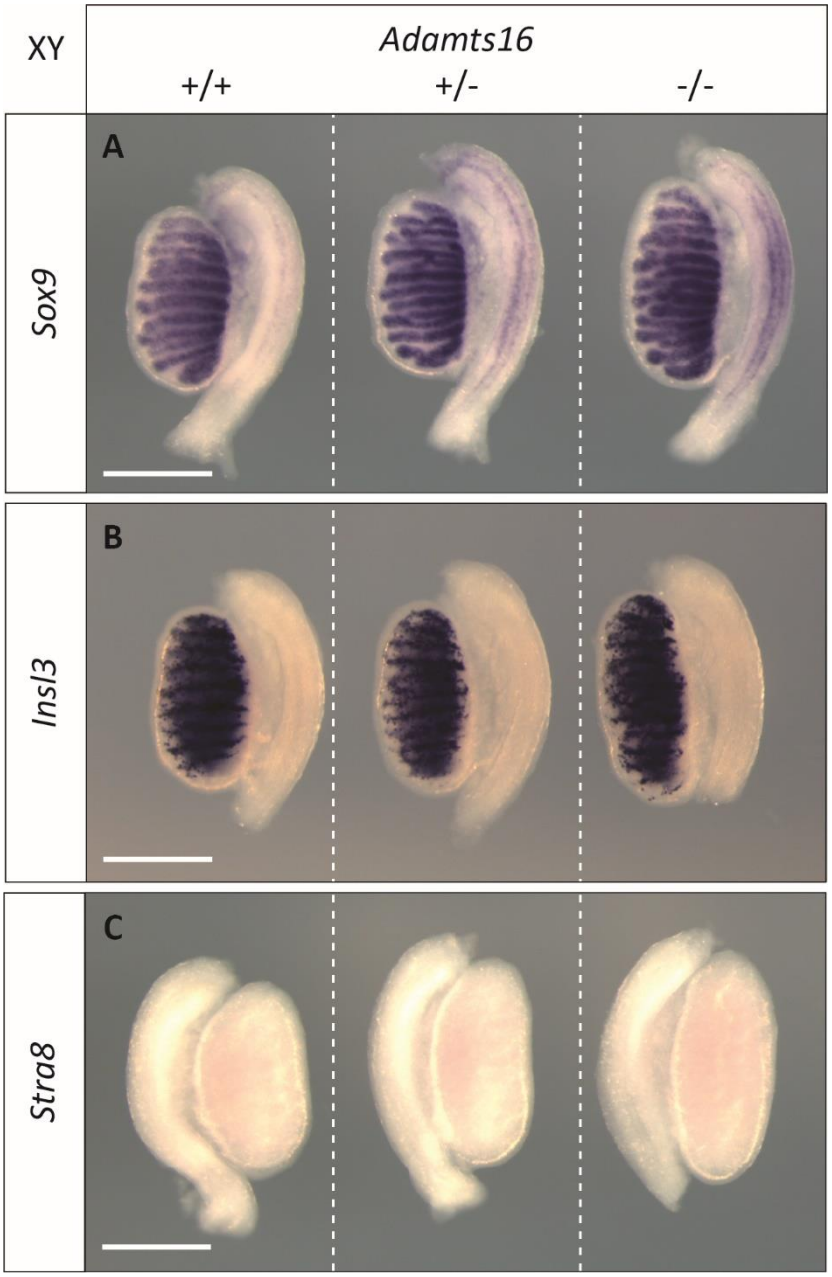


Figure 6.6 Gene expression analyses in *Adamts16* null XY gonads. WMISH analysis was performed on XY gonads at 14.5 dpc for expression of *Sox9* **(A)**, *Insl3* **(B)** and *Stra8* **(C)**. No abnormalities in gene expression were observed in *Adamts16*$^{-/-}$ testes. *N*s: +/+ *Sox9*, *Insl3* and *Stra8* = 8, *+/- Sox9* and *Stra8* = 5, +/- *Insl3* = 8, -/- *Sox9* and *Stra8* = 9, -/- *Insl3* = 7. Scale bar 500 µm.

Given the sterility of knockout rats, five homozygous adult males and five wild-type littermate males were fertility tested by mating to six B6NTac females each, from 12 weeks of age. In contrast to the data in rats, all the males were fertile. The rate of pregnancy (after matings confirmed by the presence of a vaginal plug) was not significantly different between wild-types and homozygotes (wild-type: 70%, homozygote: 73%, $p > 0.999$ using Fisher's exact test). The mean litter size of homozygous males was 7.0 pups; that of wild-type control males was 8.4 pups. This represents a significant reduction in litter size in homozygotes **(Fig 6.7)**. Males were culled at 14-15 weeks of age and their reproductive systems examined. External appearance and anogenital distance was normal **(Fig 6.8B)**. Homozygotes had well-formed seminal vesicles, vasa deferentia and epididymides **(Fig 6.8D)**. Testes appeared normal, although some homozygotes had smaller than average testes. The mean testis weight of homozygotes was 91 mg, a reduction of 9% compared to wild-type **(Fig 6.8E)**.

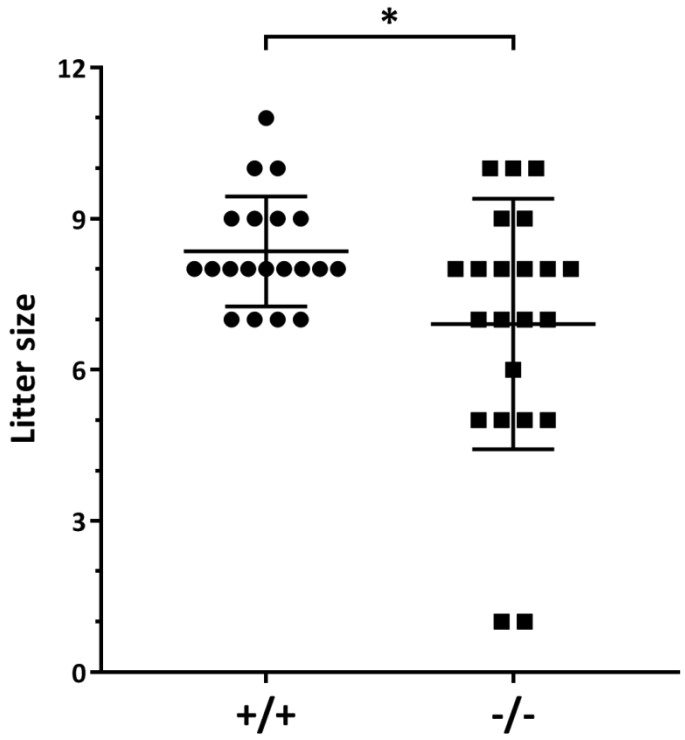


Figure 6.7 Fertility testing of *Adamts16* null males. 12-week old males were mated to wild-type females. 6 females were used per male. Matings were determined by the presence of a vaginal plug. Females were culled and opened at 13.5 -15.5 dpc and the number of pups counted. *P* –value = 0.0455 using a Mann-Whitney test. Error bars show SD. *N* = 5 males per genotype.

H&E-stained sections of homozygous testes revealed no gross abnormalities. Seminiferous tubules were well-formed and spermatids were visible in the majority of tubule lumens **(Fig 6.8G)**. However, closer inspections revealed areas of seminiferous tubule atrophy **(Fig 6.8H)**; tubules had hypospermatogenesis and contained cell debris, vacuoles and immature spermatocytes. Some tubules had mostly Sertoli cells. Occasional tubule defects are normal in wild-type testes; however, the degree of degeneration appeared abnormal in homozygotes.

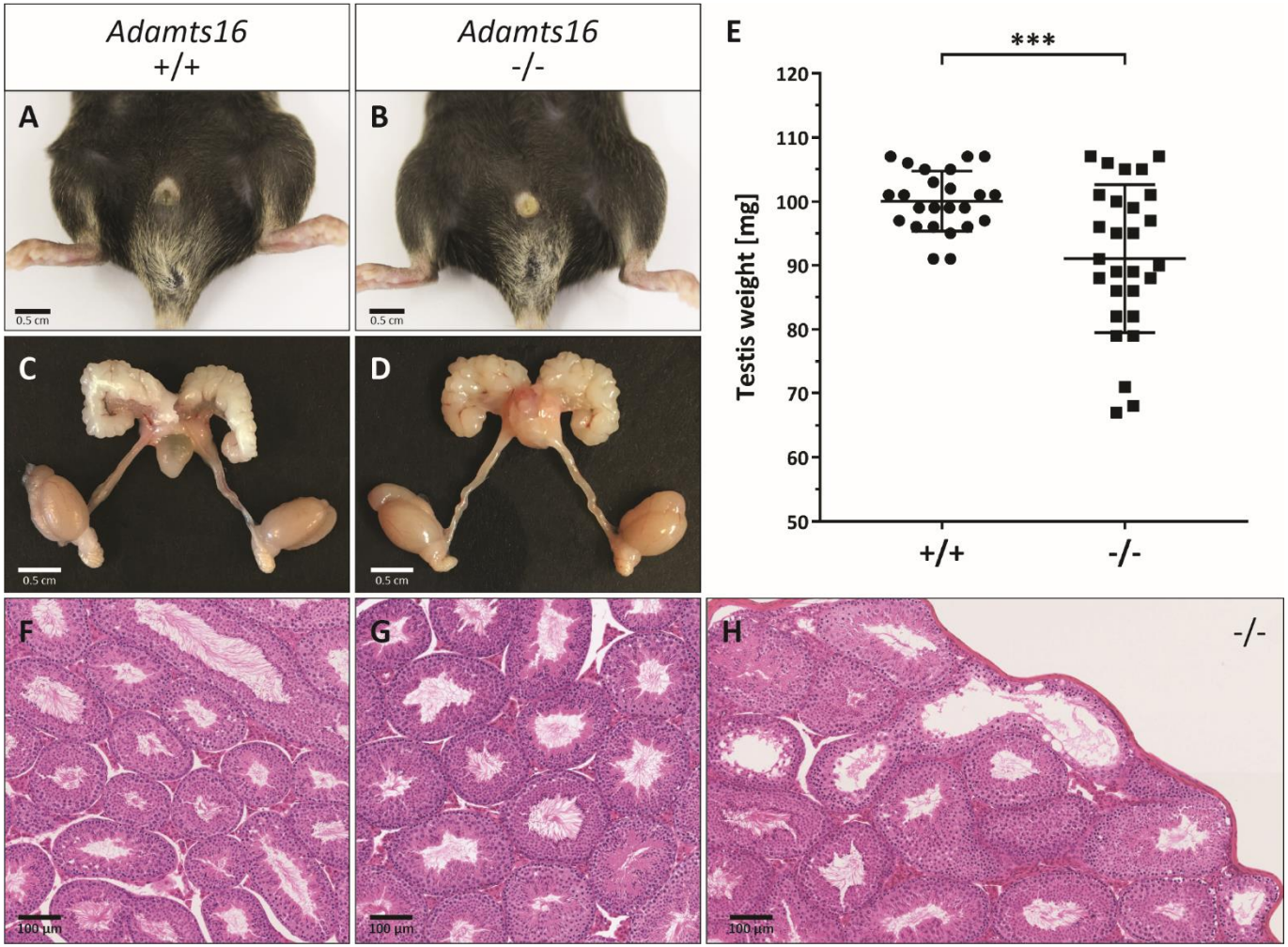


Figure 6.8 Examination of *Adamts16* null adult males. Age-matched control and homozygous males were culled at 14 weeks old. Reproductive systems were examined externally **(A-B)** and internally **(C-D).** Homozygous males have a normal anogenital distance and bilateral seminal vesicles and vasa deferentia. Testes were weighed **(E)** and a two-tailed Welch's t-test was performed. The mean testis weight of homozygous males was significantly lower than wild-type controls. *P*-value = 0.0006. Error bars show SD. Examination of H&E-stained testis sections **(F-H)** showed normal testicular morphology with well-formed seminiferous tubules and maturing spermatids in homozygous testes **(G)**, but all homozygous testes contained small regions of tubule atrophy **(H)**.

6.2.3. The role of *Adamts16* in testicular descent

Testes in mutant *Adamts16* rats failed to descend to the scrotum. Given that homozygous male mice were fertile, a severe cryptorchidism phenotype was not expected. Embryos were examined at 16.5 dpc **(Fig 6.9A-D)**. Homozygous testes had undergone the first (transabdominal) phase of descent **(Fig 6.9D)** and were located in the same position as wild-type and heterozygous testes **(Fig 6.9B-C)**. A defect in testicular descent was found in adult males examined post-mortem; the majority of homozygous males had both testes in a suprascrotal position **(Fig 6.9G)**, higher than the bilateral scrotal position in most wild-type males **(Fig 6.9E)**. Unlike in humans, testes in mice are known to temporarily retract throughout adulthood (Kaufman et al., 2009). Some testes in wild-type controls were located in a suprascrotal position **(Fig 6.9F)** but these were located, on average, in a more caudal position than testes in homozygotes **(Fig 6.9H)**. These data suggest that *Adamts16* is required for complete descent of the testes into the scrotum.

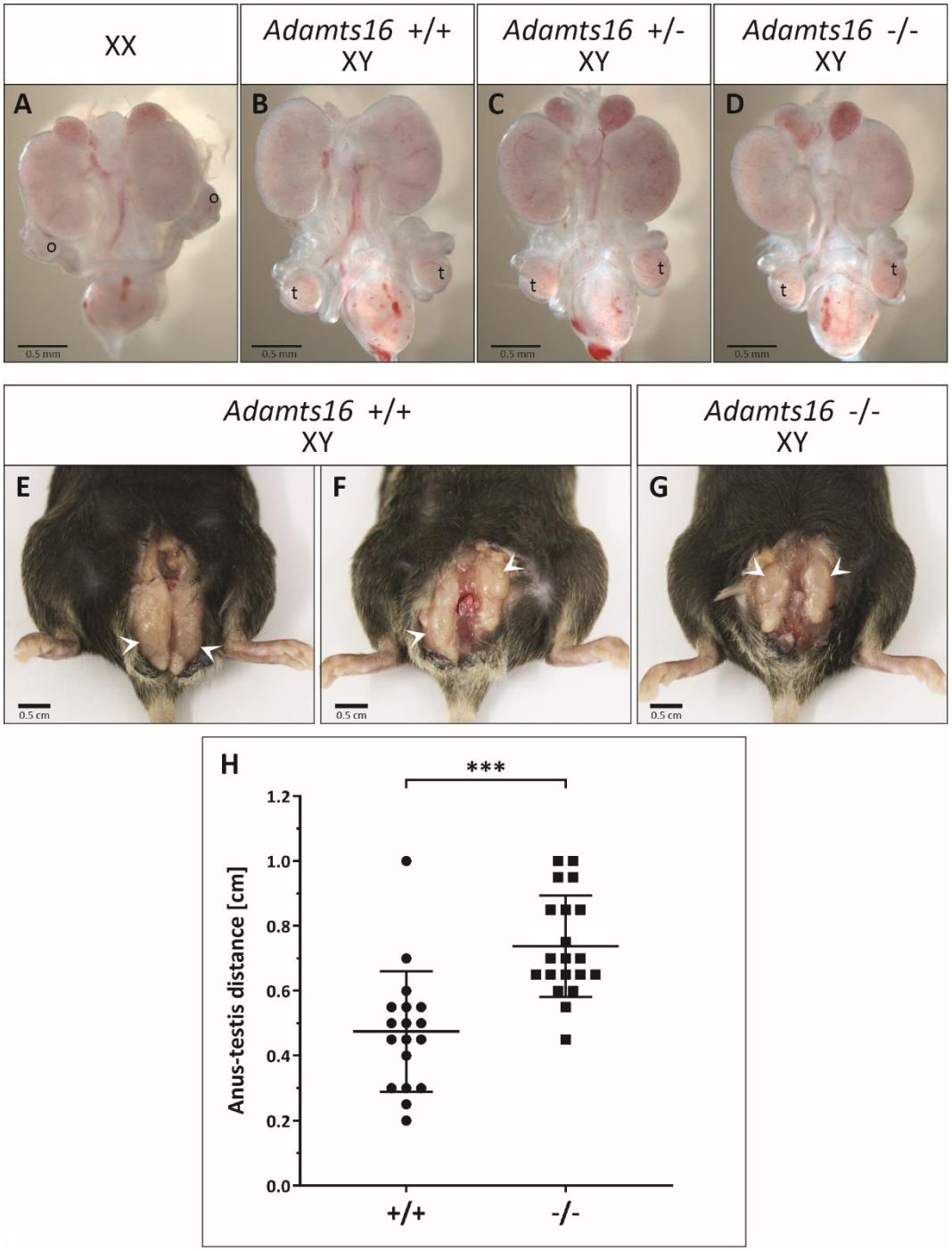


Figure 6.9 Testicular descent in *Adamts16* null mice. (A-D) Urogenital systems from 16.5 dpc embryos, imaged on a Leica MZ16F with attached camera. No defect in the transabdominal phase of testicular descent was observed in *Adamts16*$^{-/-}$ embryos. t = testis, o = ovary. **(E-G)** Example images of testis location of adult males; two scrotal testes **(E)**, uneven testes **(F)** and two suprascrotal testes **(G)**. Arrows indicate location of testes. Testes in *Adamts16*$^{+/+}$ were most often found in the scrotal position, but the occasional testis was located in the suprascrotal position. The majority of testes in *Adamts16*$^{-/-}$ mice were in the suprascrotal position. **(H)** Distance from bottom of testis to anus, measured in age-matched adult males. Testes located significantly more caudal in *Adamts16*$^{+/+}$ than in *Adamts16*$^{-/-}$ males. *P* –value <0.0001 using a Mann-Whitney test. Error bars show SD.

6.2.4. The role of *Adamts16* in ovary development

Adamts16 is expressed in the ovary from 14.5 dpc **(Fig 6.3)** and in follicular granulosa cells in the ovaries of adult females (Jacobi et al., 2013). Gonads from 14.5 dpc XX homozygote embryos were analysed by WMISH. These gonads undergo apparently normal ovarian differentiation, shown by expression of *Stra8* **(Fig 6.10A)** and absence of *Sox9* expression **(Fig 6.10B)**.

Homozygote adult females were culled at 13 weeks of age and examined. Their external genitalia were normal **(Fig 6.11B)** and uterine horn and ovaries appeared unaffected **(Fig 6.11D)**. Analysis of H&E stained ovary sections confirmed no abnormalities **(Fig 6.11F)**; ovarian morphology was normal and multiple follicles were visible at various stages of maturation. Homozygous females were fertile: six of the seven females tested became pregnant after the first mating, with an average litter size of 9 pups. *Adamts16* is not required for normal ovarian development and function.

6.2.5. The role of *Adamts16* in kidney development

Adamts16 is expressed in the developing kidney tubules. Jacobi et al. (2013) reported that knockdown of *Adamts16* by antisense morpholinos in 13.5 dpc mouse kidney organ cultures caused abnormal kidney branching morphogenesis. Wholemount immunohistochemistry was performed on 14.5 dpc kidneys, using an anti-Cytokeratin 8 antibody. Cytokeratin 8 localises to the kidney tubules. No branching defects were observed in homozygotes **(Fig 6.12)**. *Adamts16* is not required for kidney tubule morphogenesis.

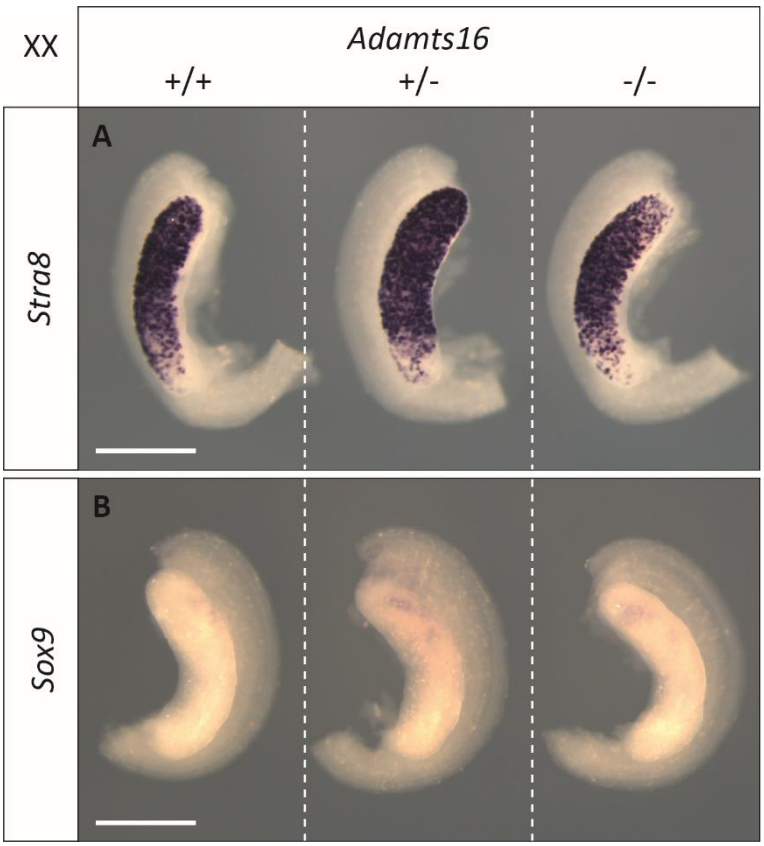


Figure 6.10 Gene expression analyses in *Adamts16* null XX gonads. WMISH analysis was performed on XX gonads from 14.5 dpc embryos, for expression of *Stra8* **(A)** and *Sox9* **(B)**. No abnormalities observed in *Adamts16*$^{-/-}$ ovaries. *N*s: +/+ = 11, *+/-* = 8, -/- = 11. Scale bar 500 µm.

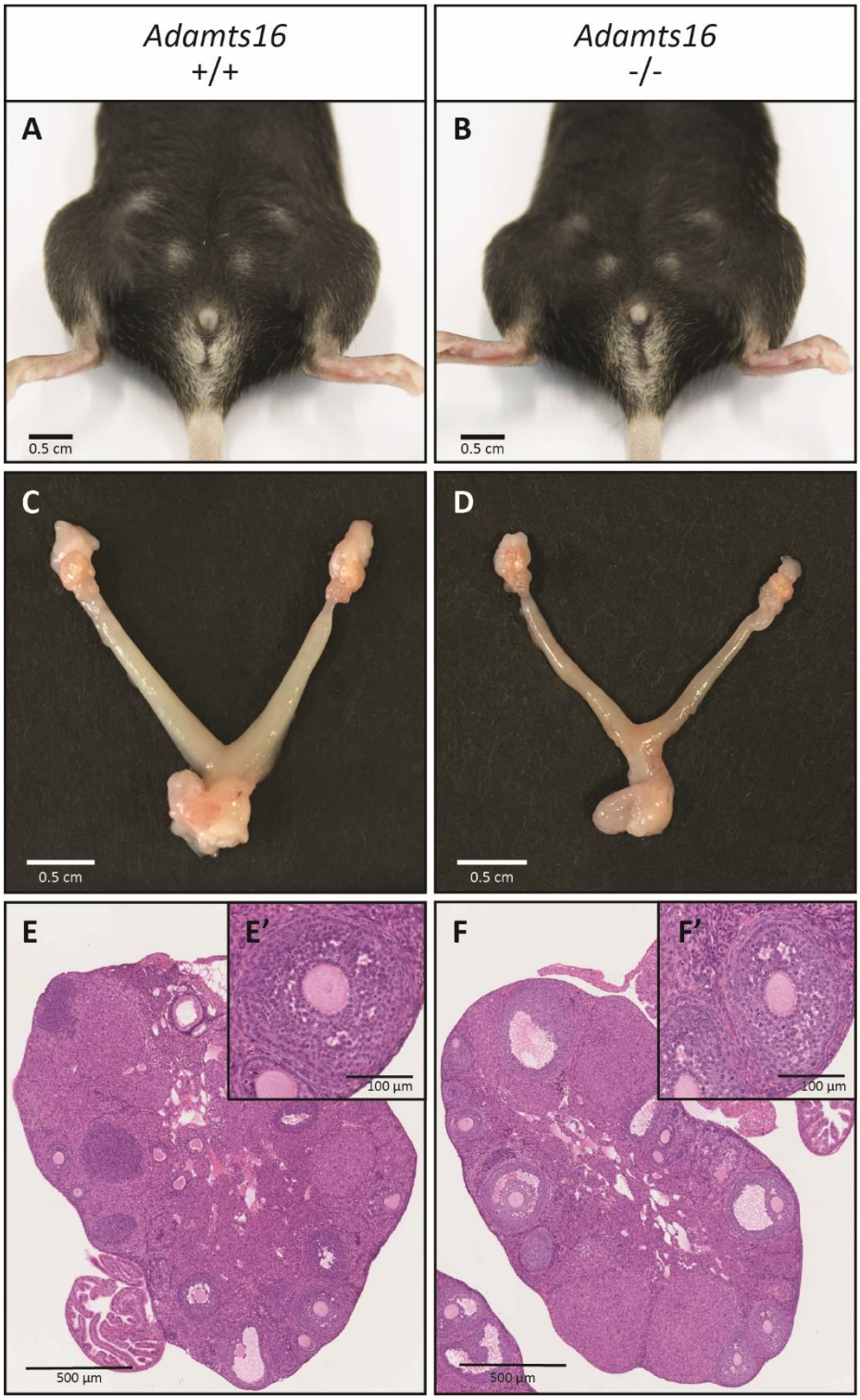


Figure 6.11 Examination of *Adamts16* null adult females. Females were culled at 13 weeks old. Reproductive systems were examined externally **(A-B)** and internally **(C-D).** Homozygous females have normal external genitalia, uterine horns and ovaries. Examination of H&E-stained ovary sections **(E-F)** showed no morphological abnormalities. All homozygote ovaries contained follicles at various stages of maturation **(F')**, comparable to wild-type controls **(E')**. *N* = 8 per genotype.

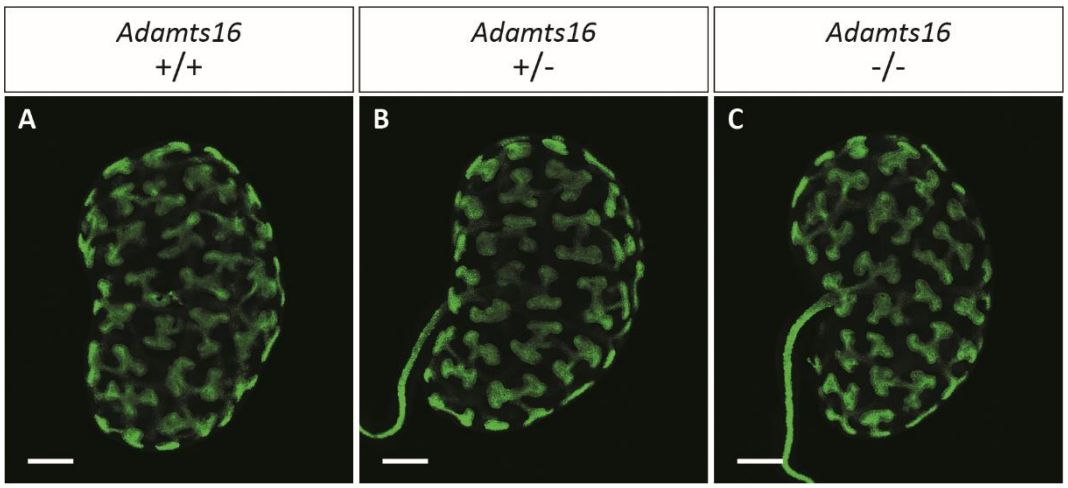


Figure 6.12 Branching in *Adamts16* null kidneys. Wholemount immunohistochemistry of 14.5 dpc kidneys from age-matched littermates, using an anti-Cytokeratin 8 antibody. Imaged on a Zeiss LSM710 upright confocal microscope using a 10x Plan-Apochromat objective. *N* = 11 pairs per genotype. Scale bar 200 µm.

6.3. Discussion

6.3.1. Loss of *Adamts16* leads to cryptorchidism

Adult male rodents with disruption to *Adamts16* have bilateral cryptorchidism; the testes in homozygous rats are located in the lower abdomen (Abdul-Majeed et al., 2014) and this chapter has demonstrated that loss of *Adamts16* in mice also leads to cryptorchidism, albeit in a milder form with testes of homozygous males located in the suprascrotal position. Testicular descent occurs in two phases. The first phase – the transabdominal phase – requires the action of *Insl3* produced by fetal Leydig cells binding to LGR8 receptors (Gorlov et al., 2002; Klonisch et al., 2004; Nef and Parada, 1999). *Insl3* expression is normal in *Adamts16*$^{-/-}$ mouse gonads and testes do descend from the pararenal position, as seen at 16.5 dpc. The second phase – the inguinoscrotal phase – is mediated by androgens and requires complex remodelling, reorientation, proliferation,

migration and contraction of the gubernaculum, guiding and anchoring the testes in the scrotum (Heyns, 1987; Hutson et al., 2013). It is this phase that is affected in *Adamts16*$^{-/-}$ mice. It is unclear which phase of testis descent is affected in rats with disruption to *Adamts16*.

ADAMTS16 is a metalloproteinase capable of cleaving ECM components such as fibronectin (Schnellmann et al., 2018). Abdul-Majeed et al. (2014) hypothesised that ADAMTS16 is required for ECM remodelling during gubernaculum maturation. The gubernaculum is a mesenchymal structure that connects the base of the testis and the epididymis to the scrotum. It is mainly composed of fibroblasts, smooth muscle cells and abundant ECM (Costa et al., 2002). Its structure and ECM composition undergoes extensive remodelling during its maturation, as do the surrounding tissues. At the end of the transabdominal phase in human fetuses the gubernaculum contains a high concentration of proteoglycans and glycosaminoglycan side chains. Later in development these proteoglycans are less abundant, replaced by a high collagen content (Soito et al., 2011), a transformation that would require extracellular proteinases, such as ADAMTS proteins. Rapid remodelling of the ECM within the gubernaculum has also been observed in rodents; just prior to inguinoscrotal descent the majority of the gubernacular ECM is degraded, which is thought to allow the gubernaculum to evert (Nation et al., 2011). The loss of ADAMTS16 could prevent the remodelling and maturation of the gubernaculum, thus weakening it or otherwise hindering its ability to direct or maintain the position of the testes in the scrotum.

An alternative hypothesis that accounts for the observed cryptorchidism phenotype is that ADAMTS16 is involved with changing the inguinal extracellular environment. Churchill et al. (2011) suggest that another family of matrix metalloproteinases – appropriately named the MMPs – are required for inguinoscrotal descent based on their expression in and around the gubernaculum. They propose that MMPs present near the gubernaculum clear a path for migration by digesting the collagen-rich inguinal fat pad. Many ADAMTS proteins are collagenases and could play a similar role (Porter et al., 2005). Androgen receptors were found in the subcutaneous inguinal fat pad during late gestation (Hutson et al., 2013), which supports the theory that the inguinal fat pad plays an active role in inguinoscrotal testis descent and could be a target of ADAMTS16 proteolysis. Canonical WNT signalling has been implicated in the response to androgens during inguinoscrotal descent, and has a known role in upregulating expression of MMPs and ADAMTSs (Buchtova et al., 2015; Chen et al., 2011).

Preliminary qRT-PCR experiments detected only very low *Adamts16* expression in the gubernaculae of 14.5 dpc embryos or two-week old pups (N. Chalon, N. Warr, personal communication), and no morphological abnormalities were apparent in the gubernacula of homozygous *Adamts16* males at 2 weeks (N. Warr, personal communication). Interestingly, weak *Adamts16* expression was detected in the mesonephric duct mesenchyme and the developing cranial suspensory ligament (CSL) of both sexes at 14.5 dpc. This suggests that ADAMTS16 may not be involved in gubernaculum maturation but might instead play a role in CSL degeneration during male development. The mechanism

underlying CSL regression is not well understood (Kassim et al., 2010), but it is androgen-dependent (Van der Schoot and Elger, 1992) and presumably involves ECM remodelling and the action of proteinases. Persistence of the CSL in males has been proposed to interfere with testis descent, so abnormalities in this process could contribute to the observed cryptorchidism, but whether this is the case remains uncertain (Adham et al., 2002; Kassim et al., 2010; Van der Schoot and Emmen, 1996).

6.3.2. Differences between rats and mice

The knockout phenotype seen here in mice is milder than the phenotype seen in rats (Abdul-Majeed et al., 2014). Homozygous rats have testes in an abdominal position compared to the suprascrotal position observed in mice. Rat testes were examined at multiple time points from 3 weeks of age; testis degeneration was seen to progress rapidly and males were completely sterile by 12 weeks of age due to a lack of spermatogenesis. In contrast, mice have only small regions of degeneration when examined at 14 weeks of age. Homozygous mice had reduced fertility but none were infertile. This difference is presumably due to the less severe cryptorchidism phenotype.

A possible explanation for the phenotypic differences between species is a divergence between rat and mouse ADAMTS16. Alignment of the peptide sequences shows a 92% identity between the rat and mouse proteins. An alternative explanation is that an additional ADAMTS family protein functions in mice and compensates for the lack of ADAMTS16, but is not expressed in the same location in rats or does not have the same catalytic ability. As described in

section 6.1.1, there are 19 members of the ADAMTS family, which often have overlapping functions and expression patterns. A genome-wide association study identified a link between SNPs in *ADAMTS20* and the incidence of cryptorchidism in Siberian husky dogs (Zhao et al., 2014b). There are also broader differences in the reproductive biology of mice and rats to consider, such as the size of organs and matters of developmental timing.

6.3.3. *Adamts16* is not required for testis development

Adamts16 is expressed in the supporting cell lineage in the developing XY gonad from 11.5 dpc, possibly downstream of the activity of SRY/SOX9 (Li et al., 2014). However, as demonstrated by the loss-of-function studies of *Adamts16* described here, it is not required for testis determination or the development of a functional testis in the mouse. In the absence of *Adamts16* the XY gonad still develops normally, with strong expression of testicular markers *Sox9* and *Insl3*. Testes have normal tubule morphology, seen both in the embryo and in adulthood, which supports spermatogenesis.

Loss of *Adamts16* may be compensated for by other members of the ADAMTS family, which often have overlapping expression profiles and catalytic functions (Kelwick et al., 2015). Microarray data of gene expression in the early mouse gonad (Jameson et al., 2012b) indicate that over half of the 19 *Adamts* genes (*1*, *2*, *5*, *6*, *7*, *9*, *10*, *12*, *16*, *17* and *19*) are expressed in some gonadal cells of one or both sexes. *Adamts5* and *Adamts12* are both expressed in the interstitial/stromal cell lineage. *Adamts19* is expressed in supporting and stromal cells from 11.5 dpc, specifically in the XX gonad (Menke and Page, 2002). *Adamts9* is expressed

in the supporting lineage of XY gonads. ADAMTS9 cleaves the proteoglycans aggrecan and versican (Somerville et al., 2003), and has multiple roles in developmental processes and morphogenesis (Dubail et al., 2014). Null *Adamts9* embryos die at 7.5 dpc, meaning information on its necessity for testis development is unavailable (Kern et al., 2010). Interestingly, however, *Adamts9* is the closest mammalian orthologue of the *C. elegans* gene *gon-1*. GON-1 is a metalloproteinase required for cell migration and subsequent morphogenesis of the tubular somatic gonadal structures (Blelloch et al., 1999; Somerville et al., 2003). There may be functional redundancy between ADAMTS9 and ADAMTS16 that accounts for the lack of gonadal phenotype seen in *Adamts16* null mice.

6.3.4. Implications for clinicians

Barseghyan et al. (2018b) described missense mutations in *ADAMTS16* found by exome sequencing in 3 out of 32 patients with 46,XY DSD, mainly complete gonadal dysgenesis, of unknown genetic aetiology. The authors categorised these polymorphisms as 'variants of unknown significance', but asserted that they were likely to causative. They backed up their conclusion by identifying *Adamts16* as expressed at lower levels in 11.5 dpc B6.Y^{POS} gonads with disrupted testis development. This differential gene expression is, however, entirely predicted given that *Adamts16* expression is strictly sexually dimorphic at that stage. Caution is always advised when interpreting the impact of such variants (Richards et al., 2015). Despite *Adamts16* not being required for testis development in mice, *ADAMTS16* may well be required for testis development in humans, and the mutations described Barseghyan et al. (2018b) may be causative of the 46,XY gonadal dysgenesis and ambiguous genitalia phenotypes.

The difference in phenotype seen even between rodent species highlights how genes do not always play the same roles in distinct mammals, and how mice may not always make good models for human disorders.

The testis undergoes large morphological changes after the onset of *SOX9* expression, predominantly the migration and arrangement of pre-Sertoli cells into organised testis cords, the pre-cursors to functioning seminiferous tubules (Combes et al., 2009). A possible hypothesis is that ADAMTS16, secreted by the pre-Sertoli cells, supports the morphogenesis of the testis cords via proteolysis of extracellular proteoglycans and remodelling of the extracellular matrix. This theory is supported by the observation that homozygote testes appear transiently underdeveloped at 14.5 dpc.

That said, it is unlikely that *Adamts16* is the gene (or one of the genes) responsible for the rescue seen in B6.Y^{POS}.*Mod13* mice. *Mod13* appears to act upstream of *Sry*, promoting its expression at 11.5 dpc, while ChiP-microarray data places *Adamts16* downstream of SRY and SOX9 (Li et al., 2014). *Mod13* also downregulates pro-ovarian gene expression at 11.5 dpc, but a role for *Adamts16* here is also unlikely given that the *Adamts16*$^{-/-}$ XY embryos described here had no evidence of ovarian gene expression or disruption to *Sox9* expression in the gonad. Gene expression analysis was not performed at 11.5 dpc in *Adamts16*$^{-/-}$ embryos. Expression analysis of *Sox9* or pro-ovarian genes such as *Rspo1* and *Foxl2* could confirm that *Adamts16* is not required for antagonism of the ovarian pathway.

6.3.5. Future work

Many questions remain to be answered regarding the role of *Adamts16* in testis descent, and the phenotypic differences observed between loss-of-function mutants in mice and rats. It is unclear how and where the gene acts to contribute to testicular descent. The similarities between the phenotypes in mice and rats suggest the gubernaculum as the most likely place for ADAMTS16 function. Further analysis of ADAMTS16 localisation in that area would be required, as well as investigating the expression of other *Adamts* genes in the search for compensatory proteins.

The mice described here were fertility tested from 10 weeks of age and culled and examined at 14 weeks of age. Small areas of testis were beginning to undergo degeneration, but this degree of atrophy only reduced testis weight by 10%. There was a reduction in litter size produced by homozygous males, but not a complete loss of fertility. Mice with conditional inactivation of the androgen receptor gene in the gubernaculum have testes located in the suprascrotal position, and became sterile with age due to progressive degeneration of the seminiferous tubules (Kaftanovskaya et al., 2012). Given this, a new cohort of *Adamts16*$^{-/-}$ males has been generated to determine if the phenotype described here becomes more severe with age. Males will be bred and then examined after 6 months of age. I predict that testicular degeneration will have progressed due to the extended time the testes are located in the suprascrotal position.

Chapter 7: General Discussion

7.1. Overview

This thesis has demonstrated that an 11 Mb QTL exists on mouse chromosome 13 that, on an otherwise sex-reversing genetic background, promotes testis differentiation. The presence of this genetic modifier causes subtle elevation of *Sry* expression, reduction of *Rspo1* and *Foxl2* expression, and subsequent upregulation of *Sox9* and increased testicular development. The region does not contain a known sex-determining gene, making this an exciting opportunity to discover novel genes involved in the sex determination genetic network, including a possible new regulator of *Sry* expression. A number of candidate genes have been identified, but the definitive genes responsible remain elusive. This thesis has also described the examination of the role of *Adamts16* in male and female sex development, in an endeavour to corroborate the proposal that mutations in *ADAMTS16* in humans can cause 46,XY gonadal dysgenesis. *Adamts16* does not appear to play a role in primary sex determination, but is required for complete testis descent.

7.2. Key conclusions

7.2.1. One or more sex-determining genes are located on chromosome 13

The location of the gene(s) influencing sex determination has been limited to position 60 Mb – 71 Mb of chromosome 13. 79 characterised protein-coding genes are located in this region, two of which were explored in more depth. Both *Ptch1* and *Adamts16* were promising candidate genes based on their sexually dimorphic expression in pre-Sertoli cells at 11.5 dpc, and their reported

association with mammalian sex development. However, examination of a null allele of *Adamts16* did not reveal any requirement in testis or ovary determination or development. Embryos haploinsufficient for *Ptch1* on a genetic background highly sensitised to XY sex reversal also did not reveal a role for this gene in testis determination. Other candidate genes have been identified and will be discussed in section 7.3. The use of knockout data to inform candidate gene selection is not ideal in this instance. *Mod13* represents a subtle modification of gene expression/function: a slight gain- or loss-of-function or mild alteration in expression regulation is the most likely explanation of its effect. Inspection of the sequence differences in *Mod13* did not suggest any major loss-of-function in relevant candidate genes. Gain-of-function effects in a pro-testis gene or regulator of *Sry* cannot be modelled simply by examination of knockouts; a gene could be responsible for the elevation of *Sry* expression, but not be essential for sex determination because its loss is compensated by other genes. Therefore, a knockout phenotype likely does not convey the full picture regarding a gene's function in rescuing sex reversal. This complexity is important to consider when considering candidate genes.

7.2.2. The *Mod13* locus may contribute to B6 sensitivity to XY gonadal sex reversal

The presence of *Mod13*-like sequences at this locus of chromosome 13 appears to correlate with the propensity of a strain to undergo gonadal sex reversal on a Y^{POS} background. The SNPs found in *Mod13* are also found in most other laboratory mouse strains. The fact that the 'protected' inbred strains (such as

129/Sv, BALB/cJ and including B6.*Mod13*) share these SNPs that differ from B6 suggests that this thesis has identified a genomic region that contributes to the phenomenon of B6 sensitivity to XY sex reversal. This hypothesis is further supported by the observed reduction in expression of pro-ovarian genes in B6.*Mod13* gonads at the sex-determining stage, in contrast to the pro-ovarian transcriptional landscape thought to explain B6's vulnerability to sex reversal (Munger et al., 2009; Munger et al., 2013). Of course *Mod13* is not the only locus to influence B6 sensitivity to Y^{POS} sex reversal; its effect is far from fully penetrant and other potential loci have previously been identified, often using other Y^{POS} colonies (Eicher et al., 1996; Munger et al., 2009; Nikolova et al., 2008). B6.Y^{POS} is an historic mouse model of 46,XY gonadal dysgenesis, used for many decades to study the regulation of testis determination (Eicher et al., 1996; Nikolova et al., 2008), to study the antagonism between the pro-testis and pro-ovary pathways (Wilhelm et al., 2009), and more recently to validate candidates of 46,XY DSDs (Barseghyan et al., 2018b). The discovery of *Mod13* adds an intriguing layer to the already complex story of B6.Y^{POS}.

7.2.3. *Adamts16* in the mouse is required for complete testis descent (but not testis determination)

The testes of *Adamts16*-deficient male mice develop normally but are mostly located in the suprascrotal position rather than the scrotal position, in contrast to control males, suggesting a role of *Adamts16* in the inguinoscrotal phase of testis descent or in the maintenance of testis position. This, along with the more severe cryptorchidism phenotype observed in *Adamts16* knockout rats (Abdul-

Majeed et al., 2014), suggests that mutations in *ADAMTS16* should be considered a candidate cause of failed testis descent in humans, rather than the cause of gonadal dysgenesis, as proposed by Barseghyan et al. (2018b). Analysis of *Adamts16* expression outside of the testis (i.e. in the gubernaculum or CSL) has generated only preliminary data, and use of an antibody to investigate ADAMTS16 protein localisation was unsuccessful. Therefore, efforts must be continued in order to elucidate the mechanism underlying the role of ADAMTS16.

7.3. Candidate genes

7.3.1. *Gas1*

GAS1 (growth arrest specific 1) is a membrane-associated protein, originally identified as a regulator of cell proliferation. It has pleiotropic functions during development, most notably as an essential component of hedgehog signalling. GAS1 acts as a co-receptor, forming a complex with PTCH1 to positively modulate HH signalling (Izzi et al., 2011). *Gas1* is a direct target of WT1, and is important for the maintenance of nephron progenitor cells during kidney morphogenesis (Kann et al., 2015). GAS1 is an activator of p38 MAPK (Leem et al., 2011), an important regulator of *Sry* expression (Warr et al., 2012). According to the scRNA-seq data presented in chapter 5, *Gas1* is highly expressed in gonadal somatic progenitor cells of both sexes at all stages analysed (from 10.5 dpc to 16.5 dpc), and in the supporting cell lineage in XX gonads from 11.5 dpc to 16.5 dpc. *Gas1* is expressed in the majority of supporting cells in XY gonads at 11.5 dpc, but appears to be downregulated in these cells by 12.5 dpc. This

expression profile suggests that regulation of *Sry* is a possibility. Also, expression in pre-granulosa cells but downregulation in Sertoli cells could suggest a possible function in ovarian differentiation. Interestingly, it was recently reported that *Gas1* regulates gene expression, including the expression of *Wnt4*, in the adult ovary during ovulation (Ren et al., 2016). Together, these data suggest *Gas1* as an interesting candidate gene.

7.3.2. *Zfp367*

Zfp367 is one of the numerous zinc-finger protein genes located in the genetic modifier region of chromosome 13. Sc-RNA-seq analysis shows an interesting expression profile. At 10.5 dpc, *Zfp367* is highly expressed in the somatic gonadal progenitor cells of both sexes, before reducing in this lineage. At 11.5 dpc, *Zfp367* expression in the supporting cell lineage appears sexually dimorphic in favour of XX gonads, but this gradually reverses over the next 5 days. In the adult mouse *Zfp367* is strongly expressed in the ovary (Asano et al., 2004). *In vitro* evidence suggests that ZFP367 localises to the nucleus and acts as a transcription factor (Asano et al., 2004), making it a potential regulator of gene expression.

7.3.3. *Med10*

MED10 is a component of the Mediator complex (Han et al., 2001). Mediator is a large, modular complex of proteins that acts as a coactivator of transcription, serving as a bridge between transcription factors and the transcriptional machinery. It is generally required for transcription by RNA polymerase II. The Mediator complex can control gene expression and thus regulate biological processes (Kagey et al., 2010; Soutourina, 2018). Unsurprisingly, given its

universal function in transcription, *Med10* is highly expressed in somatic progenitor cells and supporting cells of both sexes, at all stages tested by (Stevant et al., 2018). Very little is published about *Med10* and phenotyping data from *Med10* knockout mice, provided by the IMPC, does not include reproductive phenotype because of pre-weaning lethality. In zebrafish, *med10* appears to antagonise the WNT signalling pathway (Lin et al., 2007).

7.4. The relationship between *Mod13* and male fertility

No correlation was found between the presence of *Mod13* and enhanced male fertility, independent of *Mod13* promoting testicular development and thereby providing more, and higher quality, testicular tissue. However, this does not rule out a role for *Mod13* in promoting fertility more directly; rather, the subtle nature of *Mod13*'s impact, and the limited sensitivity of the techniques used here, were unable to detect an effect. Detailed fertility parameters – spermatid head count within the seminiferous tubules, sperm count, motility and morphology, androgen levels – were not measured. The proportion of homozygous B6 males that were fertile was lower than the proportion of heterozygous males. The experimental focus of this thesis was the role of *Mod13* in primary sex determination, not fertility; but here, I devote some discussion to the plethora of fertility-associated genes in *Mod13*. Mouse chromosome 13 itself is associated with male fertility due to the presence of key genes required for biosynthesis and action of testosterone.

One enigma that has persisted throughout my study of *Mod13* is the question of how such a large non-B6 region – spanning nearly a third of chromosome 13 –

has been maintained despite tens of generations of backcrosses to B6. The fact that the region has not organically narrowed down by gradual encroachment of B6 sequences implies a selective pressure to continue transmission of the full-length *Mod13*. Since my research has concluded that the gene(s) required to promote primary testis determination are contained in the smaller 11 Mb region, it is surprising that a reduction to this "core" *Mod13* did not happen sooner. The presence of secondary sex development and/or fertility-promoting genes in the other areas of *Mod13* may have provided this selection pressure. Formal fertility testing of the 60-71 line was not performed, but breeding of the line was a struggle that required determination and patience. The average intercross litter size was equal between the 41-71 and 60-71 lines, but the pregnancy rate (after confirmed matings) was lower at 48% in the 60-71 line compared to 68% in the 41-71 line. Perhaps without such focused efforts to breed the line for this thesis, thus circumventing to some extent this apparent negative impact on fertility, the shorter 60-71 allele would have been lost. I contend that carriers of the smaller 60-71 region have lost an important, although not essential, genetic advocate of male fertility present in the full-length *Mod13* allele. Many of the germ cell/spermatogenesis-associated genes mentioned below (*Catsper3*, *Pfn3*, *Pfn3*, *Spata31d1a*, *b* and *d*) lie outside of *Mod13(60-71)*.

7.4.1. *Hsd17b3*

Hsd17b3 encodes an enzyme involved in the final step in the biosynthesis of testosterone: 17β hydroxysteroid dehydrogenase type 3 (17β-HSD3). Localised to the testes, its role is to convert the precursor hormone androstenedione into testosterone (Baker et al., 1997). 17β-HSD3 deficiency is a known cause of 46,XY

DSD; patients with mutations in *HSD17B3* have testes but, due to reduced testosterone secretion, exhibit undervirilisation of the external genitalia. Phenotypes vary in severity from ambiguous or female external genitalia with undescended testes, to male genitalia with micropenis or hypospadias (Ahmed et al., 2013; Ben Rhouma et al., 2017; Rosler et al., 1996; Sata et al., 2010). Testosterone is important for male fertility; testosterone is required for both testis descent (Bay et al., 2011) and maintenance of spermatogenesis (De Gendt et al., 2004).

7.4.2. *Srd5a1*

Two steroid 5α-reductases – 5α-RD1 and 5α-RD2 – convert testosterone to dihydrotestosterone (DHT), a key step in androgen action. 5α-RD2 deficiency is associated with development of female genitalia in XY individuals, who undergo virilisation at puberty (Ahmed et al., 2013). 5α-RD1, the enzyme of interest here, has not been implicated as the cause of 46,XY DSDs, but is found in human spermatozoa (Aquila et al., 2015). In mice, loss of *Srd5a1* does not affect male development or fertility (Mahendroo et al., 1996).

7.4.3. Spermatogenesis and sperm function genes

Numerous genes associated with sperm form and function are located in the *Mod13* region. *Catsper3* encodes a chemosensory calcium channel thought to be required for hyperactivated motility in sperm (i.e. a type of motility used by sperm for penetrating the zona pellucida during fertilisation) and is essential for male fertility in mice (Qi et al., 2007). Mutations in other *CATSPER* genes have been linked to male infertility in humans (Singh and Rajender, 2015). *Pfn3* is

specifically expressed in spermatids during the later stages of mouse spermatogenesis and encodes a protein with a proposed role in actin dynamics (Behnen et al., 2009; Braun et al., 2002). *Spata31* is a recently described gene also essential for male fertility in mice. SPATA31 is reportedly involved in maintaining Sertoli-germ cell interactions, thereby making it important for successful spermatogenesis (Wu et al., 2015). Four other related but poorly described genes (*Spata31da-d*) are also located in *Mod13*. Interestingly, the five *Spata31*-related genes contained multiple SNPs in the *Mod13* sequence, including 8 of the 68 missense variants in *Mod13(41-71)*. In particular, 27 and 35 SNPs were associated with *Spata31* and *Spata31d1c* respectively. FANCC is involved in DNA repair and stability. Mice lacking *Fancc* suffer loss of fertility in both males and females, due to lack of gametogenesis. Male mice exhibit atrophy of the seminiferous tubules (Chen et al., 1996; Whitney et al., 1996). CTSL, an extracellular proteinase, is secreted by Sertoli cells. *Ctsl*-deficient male mice also exhibit abnormal spermatogenesis, reduced sperm count and seminiferous tubule atrophy (Wright et al., 2003). NSUN2 is an RNA methyltransferase involved in post-transcriptional regulation. NSUN2 is essential for meiotic progression of germ cells (Hussain et al., 2013). The Mouse Genetics Project (MGP) at the Wellcome Trust Sanger Institute reported infertility in *Nsun* knockout males, associated with testicular degeneration (Adissu et al., 2014). *Msx2*, with the related gene *Msx1*, is required for timely migration of primordial germ cells and, therefore, the availability of sufficient germ cell numbers in the gonad (Sun et al., 2016).

7.5. The importance of genetic background in mice and humans

A common theme throughout this thesis has been the importance of genetic background in research using mice; but the same concept can be applied to human traits and disorders. While the majority of 46,XY gonadal dysgenesis cases are caused by *de novo* mutations, a number of cases of familial gonadal dysgenesis have been reported (Berg et al., 1989; Chemke et al., 1970; Sternberg et al., 1968). Some of these familial cases are explained by paternal mosaicism for a mutation in *SRY* (Bilbao et al., 1996; Isidor et al., 2009; Schmitt-Ney et al., 1995). However, there have also been gonadal dysgenesis cases in families, sometimes across multiple generations, with no evidence of mosaicism (Assumpcao et al., 2002; Jager et al., 1992; Vilain et al., 1992). Assumpcao et al. (2002) reported a missense mutation in *SRY* in a large family with variable phenotypes; two siblings with 46,XY partial gonadal dysgenesis, their sister with 46,XY complete gonadal dysgenesis, their three unaffected brothers and their father all shared the same mutation. The mutation was not present in 100 male controls and led to poor phosphorylation of SRY *in vitro*, indicating that this is unlikely to be a neutral polymorphism that is coincidently present in a family affected by DSDs. It seems paradoxical that a deleterious mutation in a key testis-determining gene can be passed from a fertile male to his XY offspring, who then develop gonadal dysgenesis as a result. Most reported mutations in the open reading frame of *SRY* result in complete gonadal dysgenesis, with no testicular development. Clearly other factors are at play in causing the phenotypic variability described in cases of familial DSDs.

It has been suggested that environmental factors can contribute to DSD phenotypes; the increasing incidence of 'testicular dysgenesis syndrome' (TDS), which encompasses undescended testes, hypospadias, testicular germ cell tumours and low sperm count, has been linked to environmental factors such as fetal exposure to phthalates (Sharpe and Skakkebaek, 2008; Skakkebaek et al., 2001). However, genetic factors have been increasingly considered the cause of variable penetrance of mutations in DSD genes. Genetic modifiers, i.e. additional variants in functionally connected genes, are thought to impart wide phenotypic variability in individuals carrying the same mutation in core DSD genes such as *SRY* and *SF1* (Ahmed et al., 2013; Bashamboo and McElreavey, 2016; Camats et al., 2018; Eggers et al., 2016). In the case of the family described above, the mutation might result in a protein only slightly different from wild-type, allowing sufficient SRY function for complete testis determination in some individuals. However, the weaker SRY would increase sensitivity to subtle changes in the expression of other sex-determining genes or changes in the SRY-binding sequence of a target gene, causing gonadal dysgenesis in other individuals. The mouse offers a powerful tool for modelling these instances of background-dependency in DSDs because of the existence of specific mouse strains that are susceptible to sex reversal. The B6 strain is indispensable as a genetic background sensitised to disruptions in the testis-determining pathway. The B6.Y^{AKR} genetic background is an excellent model of a mildly perturbed *SRY* that alone does not cause sex reversal, but strongly predisposes the genetic network to the ovarian fate when the dosage of sex-determining genes is altered.

Mazen et al. (2016) reported the incidence of two separate deleterious mutations, one in *MAP3K1* and one in *SF1*/*NR5A1*, causing 46,XY gonadal dysgenesis when the identical *SF1* mutation alone caused mild hypospadias. This digenic concept is probably too simplistic and the phenotypic variation observed is likely a case of multiple variants having an accumulative effect on sex development. As more researchers consider the impact of multiple, individually non-deleterious mutations on sex determination, it is likely that more instances of this will come to light, and the concept of DSDs as oligogenic traits will take hold. To make the task of understanding complex traits like DSDs even more daunting, it has even been proposed that variants in any gene expressed in a disease-relevant cell are capable of impacting phenotypes because genetic networks are so interlinked within a cell (Boyle et al., 2017). This 'omnigenic' model of complex traits is a tantalising idea. Large genetic modifiers like *Mod13* offer a fantastic opportunity to understand the complex nature of sex determination and DSD inheritance. Interestingly, a large-scale targeted sequencing study of 46,XY DSD patients found a number of cases of individuals harbouring a combination of pathogenic variants in testis-determining and DASA (disorder in androgen synthesis and action, e.g. *HSD17B3* and *SRD5A2*) genes (Eggers et al., 2016), a scenario possibly not dissimilar to *Mod13*.

7.6. Future work

As is often the way with scientific research, this thesis has raised as many questions as it has answered. Given more time, there are several experiments

that may help to expand knowledge of genes in *Mod13* and the mechanism underlying their role in sex determination.

7.6.1. The search for candidate genes

Firstly, confirmation of the same depletion of pro-ovarian gene expression and elevation of *Sry* expression in B6.Y^{POS}.*Mod13(60-71)* embryos as was observed in B6.Y^{POS}.*Mod13(41-71)* embryos is necessary to confirm that the sex-determining genes of interest lie in the 11 Mb region. Secondly, further narrowing of the region of interest would certainly be helpful in the identification of genes responsible. This would require the generation and analysis of additional recombinant alleles, ideally through breeding of the *Mod13(60-71)* line and identification of naturally occurring recombination events. This step may not be possible; in the probable event that multiple genes within *Mod13* combine to promote testis development and maleness, there will be a limit to how small the region of non-B6 SNPs can become before the protective effect is lost. Of course, careful examination of recombinant lines could allow the segregation of causative genes (assuming they are spaced reasonably far apart).

Mod13 acts in a dominant manner. This suggests that its pro-testis effects may be a result of upregulation of a gene or genes in the region, compared to expression levels in B6/J. An obvious experiment to identify candidate genes is to use gene expression analysis techniques to search for genes in the region that have elevated expression in B6/mod gonads, compared to B6/B6 gonads. This could include qRT-PCR of selected candidate genes, or involve transcriptomic techniques such as RNA-seq. A significant problem with analysing gene

expression changes in XY gonads is that huge variability in the degree of sex reversal makes separating any direct effect of *Mod13* from its positive impact on the pro-testis pathway. Elevated expression of any genes could be a direct consequence of variants in *Mod13*, or a consequence of the gonad developing as an ovary. This makes interpreting transcriptomic data from B6.XY^{POS}.*Mod13* gonads very difficult. Analysis of B6.XX.*Mod13* gonads alleviates this problem, but does not provide a complete picture of genes in *Mod13* promoting testis determination. An alternative method to allow analysis of gene expression in XY gonads is to transfer the *Mod13(60-71)* allele to a B6.Y^{B6} genetic background, by crossing a *B6.Mod13(60-71)* XX female to a B6 wild-type male. Performing RNA-seq on XY gonads from such mice might identify genes in the region differentially expressed between B6/B6 and B6/mod XY mice, without the presence of the *poschiavinus* Y chromosome or sex-reversal phenotype cofounding the results.

7.6.2. Validation of candidate genes

Some interesting candidate genes have already been identified, particularly *Gas1*. Validation remains to be performed of the candidate gene expression profiles provided by scRNA-seq; confirmation of the expression/localisation of candidate genes/proteins in the differentiating gonad is important. Later, verification of SNPs in the vicinity of candidate genes would be essential, followed by analysis of candidate gene expression levels in B6.Y^{POS}.*Mod13* gonads at 11.5 dpc to determine if the presence of *Mod13* impacts their transcription. Single-cell transcriptomics, such as the techniques used by Stevant et al., would provide a global view of the genetic impact of *Mod13* and would allow us to precisely define the gonadal cell types in which *Mod13* is acting. This

would be technically challenging, especially given the need for many repeats due to the highly variable nature of B6.Y^{POS}.*Mod13* gonadal phenotypes; but even so, it would be extremely interesting.

The difficulty facing this project is that without further comprehension of the impact of individual SNPs in *Mod13*, the genes responsible for promoting testis development will likely remain elusive. 99% of the SNPs identified in *Mod13* are located within non-protein-coding regions, and more than 50% of SNPs are located more than 5 kb either upstream or downstream of a transcript. Functional interpretation of SNPs in candidate genes requires further understanding of the non-coding elements of the genome, which may require us to await better tools and better analyses to truly understand *Mod13*. Techniques like ATAC-seq – a method of mapping chromatin accessibility across the genome and identifying potential regulatory regions (Buenrostro et al., 2015) – offer the ability to enhance our knowledge of where important non-coding regulatory elements are, but a great deal more work is required to grasp how non-coding variants potentially affect biological functions (Yue et al., 2014; Zhang and Lupski, 2015). The mechanism of *Mod13*'s effects may go beyond the framework of protein-coding genes: non-coding RNAs are present in the region and should also be considered in the future.

7.6.3. Further molecular analysis of *Mod13*

More comprehensive gene expression analysis is required to fully unravel the molecular mechanism underlying the rescue effect of *Mod13* during sex determination. A detailed time-course of the *Sry* expression profile in

heterozygous gonads compared to homozygous B6 gonads would flesh out how *Sry* expression is affected and whether the elevation seen at 11.5 dpc is a result of a delay in expression or in an overall increase in transcript level. To address the mode of *Sry* regulation affected by *Mod13*, analysis of the pathways known to regulate the *Sry* locus would be of great interest. For example, the transcription factor GATA4 is thought to directly regulate *Sry* expression in a phosphorylation-dependent manner (Gierl et al., 2012; Warr et al., 2012). Reduction of *Sry* expression in *Map3k4* or *Gadd45g*-deficient embryos was attributed to reduced levels of phosphorylated GATA4 (Warr et al., 2012); therefore, quantification of levels of phosphorylated GATA4 in B6/mod gonads could inform whether the kinase cascade upstream of *Sry* expression is affected (a possible mechanism of regulation by *Gas1*). Similarly, given the reported importance of histone modification in *Sry* regulation (Carre et al., 2018; Katoh-Fukui et al., 2012; Kuroki et al., 2013), analyses of positive and negative histone marks in key regulatory elements may prove informative. But these are technically daunting experiments due to the limiting amount of tissue available, and its cellular heterogeneity, at these early stages of gonad development.

To further understand the role *Mod13* in downregulation of pro-ovarian genes, expression analyses of additional genes downstream of *Rpso1* and *Foxl2*, such as *Wnt4* and *Axin2* would be of interest. It is necessary to analyse a range of stages, especially later than the 11.5 dpc time-point previously analysed, to assess precisely when *Mod13* is acting to suppress pro-ovarian gene expression and whether the impact continues past the moment of sex determination.

7.6.4. Understanding the effect of *Mod13* in XX mice

The fact that the expression of *Foxl2* and *Rspo1* in XX B6/mod embryonic gonads was close to or the same as XY levels, even at the very early stage examined, suggests the possibility of a delay or disruption to ovarian differentiation and development. This thesis has focused on the effects of *Mod13* in XY individuals. Therefore, an interesting investigation for the future is of the phenotype of adult XX carriers of *Mod13*. This would include fertility testing of heterozygous XX females with B6.Y^{B6} males to identify sub-fertility/ reduction in litter size relative to homozygous B6 XX females. Inspection of ovaries from postnatal and adult mice, young and old, using H&E histological sections to examine follicles and immunohistochemical analyses to examine levels of FOXL2 and other ovarian markers, could also be carried out. It would even be interesting to determine whether SOX9 or other testis-determining proteins could be detected, suggesting transdifferentiation of adult ovarian somatic cells, either due to earlier disruption to ovary development or some direct impact of *Mod13* in the adult ovary.

7.7. Closing remarks

In summary, this thesis has examined a novel genetic modifier of B6.Y^{POS} gonadal sex reversal at MRC Harwell. This modifier region has a fascinating but enigmatic history that adds to an already complex mouse model of human 46,XY gonadal dysgenesis. Analysis of this modifier has identified the location of one or more sex-determining genes. While the identity of the causative genes remains elusive, the search has prompted investigation of *Adamts16*, and clarified its role in mammalian testis descent. As the oligogenic view of DSDs becomes more

widespread, genetic modifiers like *Mod13* may prove invaluable in the identification of genes functioning in different aspects of sex development and in furthering our knowledge of the non-coding genome.

References

Abbas, N., McElreavey, K., Leconiat, M., Vilain, E., Jaubert, F., Berger, R., Nihoul-Fekete, C., Rappaport, R., and Fellous, M. (1993). Familial case of 46,XX male and 46,XX true hermaphrodite associated with a paternal-derived SRY-bearing X chromosome. Comptes rendus de l'Academie des sciences Serie III, Sciences de la vie *316*, 375-383.

Abdul-Majeed, S., Mell, B., Nauli, S.M., and Joe, B. (2014). Cryptorchidism and Infertility in Rats with Targeted Disruption of the <italic>Adamts16</italic> Locus. PLoS ONE *9*, e100967.

Achermann, J.C., Domenice, S., Bachega, T.A., Nishi, M.Y., and Mendonca, B.B. (2015). Disorders of sex development: effect of molecular diagnostics. Nat Rev Endocrinol *11*, 478-488.

Adham, I.M., Steding, G., Thamm, T., Bullesbach, E.E., Schwabe, C., Paprotta, I., and Engel, W. (2002). The overexpression of the insl3 in female mice causes descent of the ovaries. Molecular endocrinology (Baltimore, Md) *16*, 244-252.

Adissu, H.A., Estabel, J., Sunter, D., Tuck, E., Hooks, Y., Carragher, D.M., Clarke, K., Karp, N.A., Project, S.M.G., Newbigging, S*., et al.* (2014). Histopathology reveals correlative and unique phenotypes in a high-throughput mouse phenotyping screen. Disease Models & Mechanisms *7*, 515-524.

Ahmed, S.F., Bashamboo, A., Lucas-Herald, A., and McElreavey, K. (2013). Understanding the genetic aetiology in patients with XY DSD. Br Med Bull *106*, 67-89.

Albrecht, K.H., and Eicher, E.M. (1997). DNA sequence analysis of Sry alleles (subgenus Mus) implicates misregulation as the cause of C57BL/6J-Y(POS) sex reversal and defines the SRY functional unit. Genetics *147*, 1267-1277.

Albrecht, K.H., Young, M., Washburn, L.L., and Eicher, E.M. (2003). Sry expression level and protein isoform differences play a role in abnormal testis development in C57BL/6J mice carrying certain Sry alleles. Genetics *164*, 277-288.

Allali, S., Muller, J.B., Brauner, R., Lourenco, D., Boudjenah, R., Karageorgou, V., Trivin, C., Lottmann, H., Lortat-Jacob, S., Nihoul-Fekete, C*., et al.* (2011). Mutation analysis of NR5A1 encoding steroidogenic factor 1 in 77 patients with 46, XY disorders of sex development (DSD) including hypospadias. PLoS One *6*, e24117.

Aquila, S., Montanaro, D., Guido, C., Santoro, M., Perrotta, I., Gervasi, S., De Amicis, F., and Lanzino, M. (2015). Human sperm molecular anatomy: the enzyme 5alpha-reductase (SRD5A) is present in the sperm and may be involved in the varicocele-related infertility. Histochemistry and cell biology *144*, 67-76.

Arango, N.A., Lovell-Badge, R., and Behringer, R.R. (1999). Targeted mutagenesis of the endogenous mouse Mis gene promoter: in vivo definition of genetic pathways of vertebrate sexual development. Cell *99*, 409-419.

Arboleda, V.A., Fleming, A., Barseghyan, H., Delot, E., Sinsheimer, J.S., and Vilain, E. (2014). Regulation of sex determination in mice by a non-coding genomic region. Genetics *197*, 885-897.

Armstrong, J.F., Pritchard-Jones, K., Bickmore, W.A., Hastie, N.D., and Bard, J.B. (1993). The expression of the Wilms' tumour gene, WT1, in the developing mammalian embryo. Mechanisms of development *40*, 85-97.

Asano, H., Murate, T., Naoe, T., Saito, H., and Stamatoyannopoulos, G. (2004). Molecular cloning and characterization of ZFF29: a protein containing a unique Cys2His2 zinc-finger motif. The Biochemical journal *384*, 647-653.

Assumpcao, J.G., Benedetti, C.E., Maciel-Guerra, A.T., Guerra, G., Jr., Baptista, M.T., Scolfaro, M.R., and de Mello, M.P. (2002). Novel mutations affecting SRY DNA-binding activity: the HMG box N65H associated with 46,XY pure gonadal dysgenesis and the familial non-HMG box R30I associated with variable phenotypes. Journal of molecular medicine (Berlin, Germany) *80*, 782-790.

Auguste, A., Chassot, A.A., Gregoire, E.P., Renault, L., Pannetier, M., Treier, M., Pailhoux, E., and Chaboissier, M.C. (2011). Loss of R-spondin1 and Foxl2 amplifies female-to-male sex reversal in XX mice. Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation *5*, 304-317.

Bagheri-Fam, S., Bird, A.D., Zhao, L., Ryan, J.M., Yong, M., Wilhelm, D., Koopman, P., Eswarakumar, V.P., and Harley, V.R. (2017). Testis Determination Requires a Specific FGFR2 Isoform to Repress FOXL2. Endocrinology *158*, 3832-3843.

Bagheri-Fam, S., Sim, H., Bernard, P., Jayakody, I., Taketo, M.M., Scherer, G., and Harley, V.R. (2008). Loss of Fgfr2 leads to partial XY sex reversal. Dev Biol *314*, 71-83.

Baker, P.J., Sha, J.H., and O'Shaughnessy, P.J. (1997). Localisation and regulation of 17beta-hydroxysteroid dehydrogenase type 3 mRNA during development in the mouse testis. Molecular and cellular endocrinology *133*, 127-133.

Barrionuevo, F., Bagheri-Fam, S., Klattig, J., Kist, R., Taketo, M.M., Englert, C., and Scherer, G. (2006). Homozygous inactivation of Sox9 causes complete XY sex reversal in mice. Biol Reprod *74*, 195-201.

Barrionuevo, F.J., Hurtado, A., Kim, G.J., Real, F.M., Bakkali, M., Kopp, J.L., Sander, M., Scherer, G., Burgos, M., and Jimenez, R. (2016). Sox9 and Sox8 protect the adult testis from male-to-female genetic reprogramming and complete degeneration. eLife *5*.

Barseghyan, H., Delot, E.C., and Vilain, E. (2018a). New technologies to uncover the molecular basis of disorders of sex development. Molecular and cellular endocrinology *468*, 60-69.

Barseghyan, H., Symon, A., Zadikyan, M., Almalvez, M., Segura, E.E., Eskin, A., Bramble, M.S., Arboleda, V.A., Baxter, R., Nelson, S.F.*, et al.* (2018b). Identification of novel candidate genes for 46,XY disorders of sex development (DSD) using a C57BL/6J-Y (POS) mouse model. Biology of sex differences *9*, 8.

Bashamboo, A., Eozenou, C., Rojo, S., and McElreavey, K. (2017). Anomalies in human sex determination provide unique insights into the complex genetic interactions of early gonad development. Clinical genetics *91*, 143-156.

Bashamboo, A., and McElreavey, K. (2016). Mechanism of Sex Determination in Humans: Insights from Disorders of Sex Development. Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation *10*, 313-325.

Baxter, R.M., Arboleda, V.A., Lee, H., Barseghyan, H., Adam, M.P., Fechner, P.Y., Bargman, R., Keegan, C., Travers, S., Schelley, S.*, et al.* (2015). Exome Sequencing for the Diagnosis of 46,XY Disorders of Sex Development. The Journal of Clinical Endocrinology & Metabolism *100*, E333-E344.

Bay, K., Main, K.M., Toppari, J., and Skakkebaek, N.E. (2011). Testicular descent: INSL3, testosterone, genes and the intrauterine milieu. Nature reviews Urology *8*, 187-196.

Behnen, M., Murk, K., Kursula, P., Cappallo-Obermann, H., Rothkegel, M., Kierszenbaum, A.L., and Kirchhoff, C. (2009). Testis-expressed profilins 3 and 4 show distinct functional characteristics and localize in the acroplaxome-manchette complex in spermatids. BMC cell biology *10*, 34.

Behringer, R.R., Finegold, M.J., and Cate, R.L. (1994). Mullerian-inhibiting substance function during mammalian sexual development. Cell *79*, 415-425.

Ben Rhouma, B., Kallabi, F., Mahfoudh, N., Ben Mahmoud, A., Engeli, R.T., Kamoun, H., Keskes, L., Odermatt, A., and Belguith, N. (2017). Novel cases of Tunisian patients with mutations in the gene encoding 17beta-hydroxysteroid dehydrogenase type 3 and a founder effect. The Journal of steroid biochemistry and molecular biology *165*, 86-94.

Berg, F.D., Kurzl, R., Hinrichsen, M.J., and Zander, J. (1989). Familial 46,XY pure gonadal dysgenesis and gonadoblastoma/dysgerminoma: case report. Gynecologic oncology *32*, 261-267.

Bernard, P., Ryan, J., Sim, H., Czech, D.P., Sinclair, A.H., Koopman, P., and Harley, V.R. (2012). Wnt signaling in ovarian development inhibits Sf1 activation of Sox9 via the Tesco enhancer. Endocrinology *153*, 901-912.

Berta, P., Hawkins, J.R., Sinclair, A.H., Taylor, A., Griffiths, B.L., Goodfellow, P.N., and Fellous, M. (1990). Genetic evidence equating SRY and the testis-determining factor. Nature *348*, 448-450.

Bi, W., Huang, W., Whitworth, D.J., Deng, J.M., Zhang, Z., Behringer, R.R., and de Crombrugghe, B. (2001). Haploinsufficiency of Sox9 results in defective cartilage primordia and premature skeletal mineralization. Proc Natl Acad Sci U S A *98*, 6698-6703.

Biason-Lauber, A. (2010). Control of sex development. Best practice & research Clinical endocrinology & metabolism *24*, 163-186.

Biason-Lauber, A., Konrad, D., Meyer, M., DeBeaufort, C., and Schoenle, E.J. (2009). Ovaries and female phenotype in a girl with 46,XY karyotype and mutations in the CBX2 gene. American journal of human genetics *84*, 658-663.

Bilbao, J.R., Loridan, L., and Castano, L. (1996). A novel postzygotic nonsense mutation in SRY in familial XY gonadal dysgenesis. Human genetics *97*, 537-539.

Birk, O.S., Casiano, D.E., Wassif, C.A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J.A., Parker, K.L*., et al.* (2000). The LIM homeobox gene Lhx9 is essential for mouse gonad formation. Nature *403*, 909-913.

Bitgood, M.J., and McMahon, A.P. (1995). Hedgehog and Bmp genes are coexpressed at many diverse sites of cell-cell interaction in the mouse embryo. Dev Biol *172*, 126-138.

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

Blelloch, R., Anna-Arriola, S.S., Gao, D., Li, Y., Hodgkin, J., and Kimble, J. (1999). The gon-1 gene is required for gonadal morphogenesis in Caenorhabditis elegans. Dev Biol *216*, 382-393.

Bogani, D., Siggers, P., Brixey, R., Warr, N., Beddow, S., Edwards, J., Williams, D., Wilhelm, D., Koopman, P., Flavell, R.A*., et al.* (2009). Loss of mitogen-activated protein kinase kinase kinase 4 (MAP3K4) reveals a requirement for MAPK signalling in mouse sex determination. PLoS Biol *7*, e1000196.

Bogatcheva, N.V., Truong, A., Feng, S., Engel, W., Adham, I.M., and Agoulnik, A.I. (2003). GREAT/LGR8 is the only receptor for insulin-like 3 peptide. Molecular endocrinology (Baltimore, Md) *17*, 2639-2646.

Bojesen, A., Juul, S., and Gravholt, C.H. (2003). Prenatal and postnatal prevalence of Klinefelter syndrome: a national registry study. The Journal of clinical endocrinology and metabolism *88*, 622-626.

Bondy, C.A. (2007). Care of girls and women with Turner syndrome: a guideline of the Turner Syndrome Study Group. The Journal of clinical endocrinology and metabolism *92*, 10-25.

Bouma, G.J., Washburn, L.L., Albrecht, K.H., and Eicher, E.M. (2007). Correct dosage of Fog2 and Gata4 transcription factors is critical for fetal testis development in mice. Proceedings of the National Academy of Sciences *104*, 14994-14999.

Bowles, J., Cooper, L., Berkman, J., and Koopman, P. (1999). Sry requires a CAG repeat domain for male sex determination in Mus musculus. Nature genetics *22*, 405-408.

Bowles, J., Knight, D., Smith, C., Wilhelm, D., Richman, J., Mamiya, S., Yashiro, K., Chawengsaksophak, K., Wilson, M.J., Rossant, J*., et al.* (2006). Retinoid signaling determines germ cell fate in mice. Science (New York, NY) *312*, 596-600.

Boyle, E.A., Li, Y.I., and Pritchard, J.K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell *169*, 1177-1186.

Bradford, S.T., Wilhelm, D., Bandiera, R., Vidal, V., Schedl, A., and Koopman, P. (2009). A cell-autonomous role for WT1 in regulating Sry in vivo. Human molecular genetics *18*, 3429-3438.

Braun, A., Aszodi, A., Hellebrand, H., Berna, A., Fassler, R., and Brandau, O. (2002). Genomic organization of profilin-III and evidence for a transcript expressed exclusively in testis. Gene *283*, 219-225.

Briscoe, J., and Thérond, P.P. (2013). The mechanisms of Hedgehog signalling and its roles in development and disease. Nature Reviews Molecular Cell Biology *14*, 416.

Buchtova, M., Oralova, V., Aklian, A., Masek, J., Vesela, I., Ouyang, Z., Obadalova, T., Konecna, Z., Spoustova, T., Pospisilova, T*., et al.* (2015). Fibroblast growth factor and canonical WNT/β-catenin signaling cooperate in suppression of chondrocyte differentiation in experimental models of FGFR signaling in cartilage. Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease *1852*, 839-850.

Buenrostro, J., Wu, B., Chang, H., and Greenleaf, W. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Current protocols in molecular biology / edited by Frederick M Ausubel [et al] *109*, 21.29.21-21.29.29.

Bullejos, M., and Koopman, P. (2001). Spatially dynamic expression of Sry in mouse genital ridges. Dev Dyn *221*, 201-205.

Bullejos, M., and Koopman, P. (2005). Delayed Sry and Sox9 expression in developing mouse gonads underlies B6-Y(DOM) sex reversal. Dev Biol *278*, 473-481.

Burgoyne, P.S., Buehr, M., and McLaren, A. (1988). XY follicle cells in ovaries of XX—XY female mouse chimaeras. Development *104*, 683-688.

Cal, S., Obaya, A.J., Llamazares, M., Garabaya, C., Quesada, V., and Lopez-Otin, C. (2002). Cloning, expression analysis, and structural characterization of seven novel human ADAMTSs, a family of metalloproteinases with disintegrin and thrombospondin-1 domains. Gene *283*, 49-62.

Camats, N., Fernandez-Cancio, M., Audi, L., Schaller, A., and Fluck, C.E. (2018). Broad phenotypes in heterozygous NR5A1 46,XY patients with a disorder of sex development: an oligogenic origin? European journal of human genetics : EJHG.

Canto, P., Soderlund, D., Reyes, E., and Mendez, J.P. (2004). Mutations in the desert hedgehog (DHH) gene in patients with 46,XY complete pure gonadal dysgenesis. The Journal of clinical endocrinology and metabolism *89*, 4480-4483.

Carlisle, C., Winking, H., Weichenhan, D., and Nagamine, C.M. (1996). Absence of correlation between Sry polymorphisms and XY sex reversal caused by the M. m. domesticus Y chromosome. Genomics *33*, 32-45.

Carre, G.A., Siggers, P., Xipolita, M., Brindle, P., Lutz, B., Wells, S., and Greenfield, A. (2018). Loss of p300 and CBP disrupts histone acetylation at the mouse Sry promoter and causes XY gonadal sex reversal. Human molecular genetics *27*, 190-198.

Chaboissier, M.C., Kobayashi, A., Vidal, V.I., Lutzkendorf, S., van de Kant, H.J., Wegner, M., de Rooij, D.G., Behringer, R.R., and Schedl, A. (2004). Functional analysis of Sox8 and Sox9 during sex determination in the mouse. Development *131*, 1891-1901.

Charron, F., Tsimiklis, G., Arcand, M., Robitaille, L., Liang, Q., Molkentin, J.D., Meloche, S., and Nemer, M. (2001). Tissue-specific GATA factors are transcriptional effectors of the small GTPase RhoA. Genes Dev *15*, 2702-2719.

Chassot, A.A., Ranc, F., Gregoire, E.P., Roepers-Gajadien, H.L., Taketo, M.M., Camerino, G., de Rooij, D.G., Schedl, A., and Chaboissier, M.C. (2008).

Activation of beta-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. Human molecular genetics *17*, 1264-1277.

Chemke, J., Carmichael, R., Stewart, J.M., Geer, R.H., and Robinson, A. (1970). Familial XY gonadal dysgenesis. J Med Genet *7*, 105-111.

Chen, M., Tomkins, D.J., Auerbach, W., McKerlie, C., Youssoufian, H., Liu, L., Gan, O., Carreau, M., Auerbach, A., Groves, T.*, et al.* (1996). Inactivation of Fac in mice produces inducible chromosomal instability and reduced fertility reminiscent of Fanconi anaemia. Nature genetics *12*, 448-451.

Chen, N., Harisis, G.N., Farmer, P., Buraundi, S., Sourial, M., Southwell, B.R., Balic, A., and Hutson, J.M. (2011). Gone with the Wnt: the canonical Wnt signaling axis is present and androgen dependent in the rodent gubernaculum. J Pediatr Surg *46*, 2363-2369.

Churchill, J.A., Buraundi, S., Farmer, P.J., Li, R., Southwell, B.R., Hutson, J.M., and Balic, A. (2011). Gubernaculum as icebreaker: do matrix metalloproteinases in rodent gubernaculum and inguinal fat pad permit testicular descent? J Pediatr Surg *46*, 2353-2357.

Clark, A.M., Garland, K.K., and Russell, L.D. (2000). Desert hedgehog (Dhh) Gene Is Required in the Mouse Testis for Formation of Adult-Type Leydig Cells and Normal Development of Peritubular Cells and Seminiferous Tubules. Biology of Reproduction *63*, 1825-1838.

Colvin, J.S., Green, R.P., Schmahl, J., Capel, B., and Ornitz, D.M. (2001). Male-to-female sex reversal in mice lacking fibroblast growth factor 9. Cell *104*, 875-889.

Combes, A.N., Lesieur, E., Harley, V.R., Sinclair, A.H., Little, M.H., Wilhelm, D., and Koopman, P. (2009). Three-dimensional visualization of testis cord morphogenesis, a novel tubulogenic mechanism in development. Dev Dyn *238*, 1033-1041.

Cools, M., Drop, S.L., Wolffenbuttel, K.P., Oosterhuis, J.W., and Looijenga, L.H. (2006). Germ cell tumors in the intersex gonad: old paths, new directions, moving frontiers. Endocr Rev *27*, 468-484.

Correa, S.M., Washburn, L.L., Kahlon, R.S., Musson, M.C., Bouma, G.J., Eicher, E.M., and Albrecht, K.H. (2012). Sex reversal in C57BL/6J XY mice caused by increased expression of ovarian genes and insufficient activation of the testis determining pathway. PLoS genetics *8*, e1002569.

Costa, W.S., Sampaio, F.J., Favorito, L.A., and Cardoso, L.E. (2002). Testicular migration: remodeling of connective tissue and muscle cells in human gubernaculum testis. The Journal of urology *167*, 2171-2176.

Coveney, D., Cool, J., Oliver, T., and Capel, B. (2008). Four-dimensional analysis of vascularization during primary development of an organ, the gonad. Proc Natl Acad Sci U S A *105*, 7212-7217.

Coward, P., Nagai, K., Chen, D., Thomas, H.D., Nagamine, C.M., and Lau, Y.F. (1994). Polymorphism of a CAG trinucleotide repeat within Sry correlates with B6.YDom sex reversal. Nature genetics *6*, 245-250.

Cox, J.J., Willatt, L., Homfray, T., and Woods, C.G. (2011). A SOX9 duplication and familial 46,XX developmental testicular disorder. The New England journal of medicine *364*, 91-93.

Crisponi, L., Deiana, M., Loi, A., Chiappe, F., Uda, M., Amati, P., Bisceglia, L., Zelante, L., Nagaraja, R., Porcu, S*., et al.* (2001). The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. Nature genetics *27*, 159-166.

Das, D.K., Sanghavi, D., Gawde, H., Idicula-Thomas, S., and Vasudevan, L. (2011). Novel homozygous mutations in Desert hedgehog gene in patients with 46,XY complete gonadal dysgenesis and prediction of its structural and functional implications by computational methods. Eur J Med Genet *54*, e529-534.

Davidson, R.K., Waters, J.G., Kevorkian, L., Darrah, C., Cooper, A., Donell, S.T., and Clark, I.M. (2006). Expression profiling of metalloproteinases and their inhibitors in synovium and cartilage. Arthritis Res Ther *8*, R124.

De Gendt, K., Swinnen, J.V., Saunders, P.T., Schoonjans, L., Dewerchin, M., Devos, A., Tan, K., Atanassova, N., Claessens, F., Lecureuil, C*., et al.* (2004). A Sertoli cell-selective knockout of the androgen receptor causes spermatogenic arrest in meiosis. Proc Natl Acad Sci U S A *101*, 1327-1332.

Dessens, A.B., Slijper, F.M., and Drop, S.L. (2005). Gender dysphoria and gender change in chromosomal females with congenital adrenal hyperplasia. Archives of sexual behavior *34*, 389-397.

Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. Nature structural & molecular biology *20*, 1147-1155.

Dubail, J., Aramaki-Hattori, N., Bader, H.L., Nelson, C.M., Katebi, N., Matuska, B., Olsen, B.R., and Apte, S.S. (2014). A new Adamts9 conditional mouse allele identifies its non-redundant role in interdigital web regression. Genesis *52*, 702-712.

Dubin, R.A., and Ostrer, H. (1994). Sry is a transcriptional activator. Molecular endocrinology (Baltimore, Md) *8*, 1182-1192.

Eggers, S., Sadedin, S., van den Bergen, J.A., Robevska, G., Ohnesorg, T., Hewitt, J., Lambeth, L., Bouty, A., Knarston, I.M., Tan, T.Y*., et al.* (2016). Disorders of sex development: insights from targeted gene sequencing of a large international patient cohort. Genome biology *17*, 243.

Eicher, E.M. (1988). Sex determination in mouse and man - Autosomal genes involved in mammalian primary sex determination. Philosophical Transactions of the Royal Society of London B, Biological Sciences *322*, 109-118.

Eicher, E.M., and Washburn, L.L. (1983). Inherited sex reversal in mice: identification of a new primary sex-determining gene. The Journal of experimental zoology *228*, 297-304.

Eicher, E.M., and Washburn, L.L. (2001). Does one gene determine whether a C57BL/6J-Y(POS) mouse will develop as a female or as an hermaphrodite? The Journal of experimental zoology *290*, 322-326.

Eicher, E.M., Washburn, L.L., Schork, N.J., Lee, B.K., Shown, E.P., Xu, X., Dredge, R.D., Pringle, M.J., and Page, D.C. (1996). Sex-determining genes on mouse autosomes identified by linkage analysis of C57BL/6J-YPOS sex reversal. Nature genetics *14*, 206-209.

Eicher, E.M., Washburn, L.L., Whitney, J.B., 3rd, and Morrow, K.E. (1982). Mus poschiavinus Y chromosome in the C57BL/6J murine genome causes sex reversal. Science (New York, NY) *217*, 535-537.

Eicher, E.M.a., and Washburn, L.L. (1986). Genetic Control of Primary Sex Determination in Mice. Annual Review of Genetics *20*, 327-360.

Emmen, J.M., McLuskey, A., Adham, I.M., Engel, W., Grootegoed, J.A., and Brinkmann, A.O. (2000). Hormonal control of gubernaculum development during testis descent: gubernaculum outgrowth in vitro requires both insulin-like factor and androgen. Endocrinology *141*, 4720-4727.

Ewen, K., Jackson, A., Wilhelm, D., and Koopman, P. (2010). A male-specific role for p38 mitogen-activated protein kinase in germ cell sex differentiation in mice. Biol Reprod *83*, 1005-1014.

Ferlin, A., Simonato, M., Bartoloni, L., Rizzo, G., Bettella, A., Dottorini, T., Dallapiccola, B., and Foresta, C. (2003). The INSL3-LGR8/GREAT ligand-receptor pair in human cryptorchidism. The Journal of clinical endocrinology and metabolism *88*, 4273-4279.

Ford, C.E., Jones, K.W., Polani, P.E., De Almeida, J.C., and Briggs, J.H. (1959). A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). Lancet (London, England) *1*, 711-713.

Foster, J.W., Dominguez-Steglich, M.A., Guioli, S., Kwok, C., Weller, P.A., Stevanovic, M., Weissenbach, J., Mansour, S., Young, I.D., Goodfellow, P.N.*, et al.* (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. Nature *372*, 525-530.

Frazer, K.A., Eskin, E., Kang, H.M., Bogue, M.A., Hinds, D.A., Beilharz, E.J., Gupta, R.V., Montgomery, J., Morenzoni, M.M., Nilsen, G.B.*, et al.* (2007). A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. Nature *448*, 1050.

Fujimoto, Y., Tanaka, S.S., Yamaguchi, Y.L., Kobayashi, H., Kuroki, S., Tachibana, M., Shinomura, M., Kanai, Y., Morohashi, K., Kawakami, K.*, et al.* (2013). Homeoproteins Six1 and Six4 regulate male sex determination and mouse gonadal development. Dev Cell *26*, 416-430.

Gao, S., De Geyter, C., Kossowska, K., and Zhang, H. (2007). FSH stimulates the expression of the ADAMTS-16 protease in mature human ovarian follicles. MHR: Basic science of reproductive medicine *13*, 465-471.

Gierl, M.S., Gruhn, W.H., von Seggern, A., Maltry, N., and Niehrs, C. (2012). GADD45G functions in male sex determination by promoting p38 signaling and Sry expression. Dev Cell *23*, 1032-1042.

Ginsburg, M., Snow, M.H., and McLaren, A. (1990). Primordial germ cells in the mouse embryo during gastrulation. Development *110*, 521-528.

Gonen, N., Futtner, C.R., Wood, S., Garcia-Moreno, S.A., Salamone, I.M., Samson, S.C., Sekido, R., Poulat, F., Maatouk, D.M., and Lovell-Badge, R. (2018). Sex reversal following deletion of a single distal enhancer of Sox9. Science (New York, NY) *360*, 1469-1473.

Gonen, N., Quinn, A., O'Neill, H.C., Koopman, P., and Lovell-Badge, R. (2017). Normal Levels of Sox9 Expression in the Developing Mouse Testis Depend on the TES/TESCO Enhancer, but This Does Not Act Alone. PLoS genetics *13*, e1006520.

Goodrich, L.V., Milenkovic, L., Higgins, K.M., and Scott, M.P. (1997). Altered neural cell fates and medulloblastoma in mouse patched mutants. Science (New York, NY) *277*, 1109-1113.

Gopalakrishnan, K., Kumarasamy, S., Abdul-Majeed, S., Kalinoski, A.L., Morgan, E.E., Gohara, A.F., Nauli, S.M., Filipiak, W.E., Saunders, T.L., and Joe, B. (2012). Targeted disruption of Adamts16 gene in a rat genetic model of hypertension. Proc Natl Acad Sci U S A *109*, 20555-20559.

Gorlov, I.P., Kamat, A., Bogatcheva, N.V., Jones, E., Lamb, D.J., Truong, A., Bishop, C.E., McElreavey, K., and Agoulnik, A.I. (2002). Mutations of the GREAT gene cause cryptorchidism. Human molecular genetics *11*, 2309-2318.

Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature *346*, 245-250.

Hacker, A., Capel, B., Goodfellow, P., and Lovell-Badge, R. (1995). Expression of Sry, the mouse sex determining gene. Development *121*, 1603-1614.

Hammes, A., Guo, J.K., Lutsch, G., Leheste, J.R., Landrock, D., Ziegler, U., Gubler, M.C., and Schedl, A. (2001). Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. Cell *106*, 319-329.

Han, S.J., Lee, J.S., Kang, J.S., and Kim, Y.J. (2001). Med9/Cse2 and Gal11 modules are required for transcriptional repression of distinct group of genes. The Journal of biological chemistry *276*, 37020-37026.

Harisis, G.N., Lewis, A.G., Southwell, B.R., and Hutson, J.M. (2013). Hoxa-11 maintains cell proliferation in the mouse gubernaculum to facilitate testicular descent. J Pediatr Surg *48*, 2431-2436.

Harley, V.R., and Goodfellow, P.N. (1994). The biochemical role of SRY in sex determination. Mol Reprod Dev *39*, 184-193.

Harris, A., Siggers, P., Corrochano, S., Warr, N., Sagar, D., Grimes, D.T., Suzuki, M., Burdine, R.D., Cong, F., Koo, B.K.*, et al.* (2018). ZNRF3 functions in mammalian sex determination by inhibiting canonical WNT signaling. Proc Natl Acad Sci U S A *115*, 5474-5479.

Hatano, O., Takakusu, A., Nomura, M., and Morohashi, K. (1996). Identical origin of adrenal cortex and gonad revealed by expression profiles of Ad4BP/SF-1. Genes Cells *1*, 663-671.

Herlihy, A.S., Halliday, J.L., Cock, M.L., and McLachlan, R.I. (2011). The prevalence and diagnosis rates of Klinefelter syndrome: an Australian comparison. The Medical journal of Australia *194*, 24-28.

Heyns, C.F. (1987). The gubernaculum during testicular descent in the human fetus. Journal of anatomy *153*, 93-112.

Hiramatsu, R., Matoba, S., Kanai-Azuma, M., Tsunekawa, N., Katoh-Fukui, Y., Kurohmaru, M., Morohashi, K., Wilhelm, D., Koopman, P., and Kanai, Y. (2009). A critical time window of Sry action in gonadal sex determination in mice. Development *136*, 129-138.

Hoffmeyer, A., Piekorz, R., Moriggl, R., and Ihle, J.N. (2001). Gadd45gamma is dispensable for normal mouse development and T-cell proliferation. Mol Cell Biol *21*, 3137-3143.

Hu, Y.-C., Okumura, L.M., and Page, D.C. (2013). Gata4 Is Required for Formation of the Genital Ridge in Mice. PLoS genetics *9*, e1003629.

Huang, B., Wang, S., Ning, Y., Lamb, A.N., and Bartley, J. (1999). Autosomal XX sex reversal caused by duplication of SOX9. American journal of medical genetics *87*, 349-353.

Hussain, S., Tuorto, F., Menon, S., Blanco, S., Cox, C., Flores, J.V., Watt, S., Kudo, N.R., Lyko, F., and Frye, M. (2013). The mouse cytosine-5 RNA methyltransferase NSun2 is a component of the chromatoid body and required for testis differentiation. Mol Cell Biol *33*, 1561-1570.

Hutson, J.M. (1986). Testicular feminization: a model for testicular descent in mice and men. J Pediatr Surg *21*, 195-198.

Hutson, J.M., and Hasthorpe, S. (2005). Testicular descent and cryptorchidism: the state of the art in 2004. J Pediatr Surg *40*, 297-302.

Hutson, J.M., Southwell, B.R., Li, R., Lie, G., Ismail, K., Harisis, G., and Chen, N. (2013). The regulation of testicular descent and the effects of cryptorchidism. Endocr Rev *34*, 725-752.

Isidor, B., Capito, C., Paris, F., Baron, S., Corradini, N., Cabaret, B., Leclair, M.D., Giraud, M., Martin-Coignard, D., David, A.*, et al.* (2009). Familial frameshift SRY mutation inherited from a mosaic father with testicular dysgenesis syndrome. The Journal of clinical endocrinology and metabolism *94*, 3467-3471.

Izzi, L., Levesque, M., Morin, S., Laniel, D., Wilkes, B.C., Mille, F., Krauss, R.S., McMahon, A.P., Allen, B.L., and Charron, F. (2011). Boc and Gas1 each form distinct Shh receptor complexes with Ptch1 and are required for Shh-mediated cell proliferation. Dev Cell *20*, 788-801.

Jacobi, C.L., Rudigier, L.J., Scholz, H., and Kirschner, K.M. (2013). Transcriptional regulation by the Wilms tumor protein, Wt1, suggests a role of the metalloproteinase Adamts16 in murine genitourinary development. The Journal of biological chemistry *288*, 18811-18824.

Jacobs, P.A., and Strong, J.A. (1959). A case of human intersexuality having a possible XXY sex-determining mechanism. Nature *183*, 302-303.

Jager, R.J., Anvret, M., Hall, K., and Scherer, G. (1990). A human XY female with a frame shift mutation in the candidate testis-determining gene SRY. Nature *348*, 452-454.

Jager, R.J., Harley, V.R., Pfeiffer, R.A., Goodfellow, P.N., and Scherer, G. (1992). A familial mutation in the testis-determining gene SRY shared by both sexes. Human genetics *90*, 350-355.

Jameson, S.A., Lin, Y.T., and Capel, B. (2012a). Testis development requires the repression of Wnt4 by Fgf signaling. Dev Biol *370*, 24-32.

Jameson, S.A., Natarajan, A., Cool, J., DeFalco, T., Maatouk, D.M., Mork, L., Munger, S.C., and Capel, B. (2012b). Temporal Transcriptional Profiling of Somatic and Germ Cells Reveals Biased Lineage Priming of Sexual Fate in the Fetal Mouse Gonad. PLoS genetics *8*, e1002575.

Joe, B., Saad, Y., Dhindaw, S., Lee, N.H., Frank, B.C., Achinike, O.H., Luu, T.V., Gopalakrishnan, K., Toland, E.J., Farms, P.*, et al.* (2009). Positional identification of variants of Adamts16 linked to inherited hypertension. Human molecular genetics *18*, 2825-2838.

Johnen, H., González-Silva, L., Carramolino, L., Flores, J.M., Torres, M., and Salvador, J.M. (2013). Gadd45g Is Essential for Primary Sex Determination, Male Fertility and Testis Development. PLoS ONE *8*, e58751.

Johnson, E.K., Rosoklija, I., Finlayson, C., Chen, D., Yerkes, E.B., Madonna, M.B., Holl, J.L., Baratz, A.B., Davis, G., and Cheng, E.Y. (2017). Attitudes towards "disorders of sex development" nomenclature among affected individuals. Journal of Pediatric Urology.

Jordan, B.K., Mohammed, M., Ching, S.T., Delot, E., Chen, X.N., Dewing, P., Swain, A., Rao, P.N., Elejalde, B.R., and Vilain, E. (2001). Up-regulation of WNT-4 signaling and dosage-sensitive sex reversal in humans. American journal of human genetics *68*, 1102-1109.

Kaftanovskaya, E.M., Huang, Z., Barbara, A.M., De Gendt, K., Verhoeven, G., Gorlov, I.P., and Agoulnik, A.I. (2012). Cryptorchidism in mice with an androgen receptor ablation in gubernaculum testis. Molecular endocrinology (Baltimore, Md) *26*, 598-607.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Kann, M., Bae, E., Lenz, M.O., Li, L., Trannguyen, B., Schumacher, V.A., Taglienti, M.E., Bordeianou, L., Hartwig, S., Rinschen, M.M.*, et al.* (2015). WT1 targets Gas1 to maintain nephron progenitor cells by modulating FGF signals. Development *142*, 1254-1266.

Karl, J., and Capel, B. (1998). Sertoli cells of the mouse testis originate from the coelomic epithelium. Dev Biol *203*, 323-333.

Kashimada, K., and Koopman, P. (2010). *Sry*: the master switch in mammalian sex determination. Development *137*, 3921-3930.

Kashimada, K., Pelosi, E., Chen, H., Schlessinger, D., Wilhelm, D., and Koopman, P. (2011). FOXL2 and BMP2 act cooperatively to regulate follistatin gene expression during ovarian development. Endocrinology *152*, 272-280.

Kassim, N.M., Russell, D.A., and Payne, A.P. (2010). Does the cranial suspensory ligament have a role in cryptorchidism? Cells, tissues, organs *191*, 307-315.

Katoh-Fukui, Y., Miyabayashi, K., Komatsu, T., Owaki, A., Baba, T., Shima, Y., Kidokoro, T., Kanai, Y., Schedl, A., Wilhelm, D.*, et al.* (2012). Cbx2, a polycomb group gene, is required for Sry gene expression in mice. Endocrinology *153*, 913-924.

Katoh-Fukui, Y., Tsuchiya, R., Shiroishi, T., Nakahara, Y., Hashimoto, N., Noguchi, K., and Higashinakagawa, T. (1998). Male-to-female sex reversal in M33 mutant mice. Nature *393*, 688-692.

Kaufman, M., Nikitin, A.Y., and Sundberg, J.P. (2009). Histologic Basis of Mouse Endocrine System Development: A Comparative Analysis (CRC Press, Inc.).

Kelwick, R., Desanlis, I., Wheeler, G.N., and Edwards, D.R. (2015). The ADAMTS (A Disintegrin and Metalloproteinase with Thrombospondin motifs) family. Genome biology *16*, 113.

Kent, J., Wheatley, S.C., Andrews, J.E., Sinclair, A.H., and Koopman, P. (1996). A male-specific role for SOX9 in vertebrate sex determination. Development *122*, 2813-2822.

Kerkela, R., Pikkarainen, S., Majalahti-Palviainen, T., Tokola, H., and Ruskoaho, H. (2002). Distinct roles of mitogen-activated protein kinase pathways in GATA-4 transcription factor-mediated regulation of B-type natriuretic peptide gene. The Journal of biological chemistry *277*, 13752-13760.

Kern, C.B., Wessels, A., McGarity, J., Dixon, L.J., Alston, E., Argraves, W.S., Geeting, D., Nelson, C.M., Menick, D.R., and Apte, S.S. (2010). Reduced versican cleavage due to Adamts9 haploinsufficiency is associated with cardiac and aortic anomalies. Matrix biology : journal of the International Society for Matrix Biology *29*, 304-316.

Kim, Y., Bingham, N., Sekido, R., Parker, K.L., Lovell-Badge, R., and Capel, B. (2007). Fibroblast growth factor receptor 2 regulates proliferation and Sertoli differentiation during male sex determination. Proc Natl Acad Sci U S A *104*, 16558-16563.

Kim, Y., Kobayashi, A., Sekido, R., DiNapoli, L., Brennan, J., Chaboissier, M.C., Poulat, F., Behringer, R.R., Lovell-Badge, R., and Capel, B. (2006). Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination. PLoS Biol *4*, e187.

Klonisch, T., Fowler, P.A., and Hombach-Klonisch, S. (2004). Molecular and genetic regulation of testis descent and external genitalia development. Dev Biol *270*, 1-18.

Kocer, A., Reichmann, J., Best, D., and Adams, I.R. (2009). Germ cell sex determination in mammals. Molecular human reproduction *15*, 205-213.

Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1991). Male development of chromosomally female mice transgenic for Sry. Nature *351*, 117-121.

Koopman, P., Munsterberg, A., Capel, B., Vivian, N., and Lovell-Badge, R. (1990). Expression of a candidate sex-determining gene during mouse testis differentiation. Nature *348*, 450-452.

Kreidberg, J.A., Sariola, H., Loring, J.M., Maeda, M., Pelletier, J., Housman, D., and Jaenisch, R. (1993). WT-1 is required for early kidney development. Cell *74*, 679-691.

Kuroki, S., Matoba, S., Akiyoshi, M., Matsumura, Y., Miyachi, H., Mise, N., Abe, K., Ogura, A., Wilhelm, D., Koopman, P*., et al.* (2013). Epigenetic regulation of mouse sex determination by the histone demethylase Jmjd1a. Science (New York, NY) *341*, 1106-1109.

Kusaka, M., Katoh-Fukui, Y., Ogawa, H., Miyabayashi, K., Baba, T., Shima, Y., Sugiyama, N., Sugimoto, Y., Okuno, Y., Kodama, R*., et al.* (2010). Abnormal epithelial cell polarity and ectopic epidermal growth factor receptor (EGFR) expression induced in Emx2 KO embryonic gonads. Endocrinology *151*, 5893-5904.

Larney, C., Bailey, T.L., and Koopman, P. (2014). Switching on sex: transcriptional regulation of the testis-determining gene Sry. Development *141*, 2195-2205.

Lavery, R., Chassot, A.A., Pauper, E., Gregoire, E.P., Klopfenstein, M., de Rooij, D.G., Mark, M., Schedl, A., Ghyselinck, N.B., and Chaboissier, M.C. (2012). Testicular differentiation occurs in absence of R-spondin1 and Sox9 in mouse sex reversals. PLoS genetics *8*, e1003170.

Lavery, R., Lardenois, A., Ranc-Jianmotamedi, F., Pauper, E., Gregoire, E.P., Vigier, C., Moreilhon, C., Primig, M., and Chaboissier, M.C. (2011). XY Sox9 embryonic loss-of-function mouse mutants show complete sex reversal and produce partially fertile XY oocytes. Dev Biol *354*, 111-122.

Lee, H., Deignan, J.L., Dorrani, N., Strom, S.P., Kantarci, S., Quintero-Rivera, F., Das, K., Toy, T., Harry, B., Yourshaw, M.*, et al.* (2014). Clinical exome sequencing for genetic identification of rare Mendelian disorders. Jama *312*, 1880-1887.

Lee, P.A., Houk, C.P., Ahmed, S.F., and Hughes, I.A. (2006). Consensus statement on management of intersex disorders. International Consensus Conference on Intersex. Pediatrics *118*, e488-500.

Lee, R.T.H., Zhao, Z., and Ingham, P.W. (2016). Hedgehog signalling. Development *143*, 367-372.

Leem, Y.E., Han, J.W., Lee, H.J., Ha, H.L., Kwon, Y.L., Ho, S.M., Kim, B.G., Tran, P., Bae, G.U., and Kang, J.S. (2011). Gas1 cooperates with Cdo and promotes myogenic differentiation via activation of p38MAPK. Cellular signalling *23*, 2021-2029.

Li, S.W., Arita, M., Fertala, A., Bao, Y., Kopen, G.C., Langsjo, T.K., Hyttinen, M.M., Helminen, H.J., and Prockop, D.J. (2001). Transgenic mice with inactive alleles for procollagen N-proteinase (ADAMTS-2) develop fragile skin and male sterility. The Biochemical journal *355*, 271-278.

Li, Y., Zheng, M., and Lau, Y.-Fai C. (2014). The Sex-Determining Factors SRY and SOX9 Regulate Similar Target Genes and Promote Testis Cord Formation during Testicular Differentiation. Cell Reports *8*, 723-733.

Lin, X., Rinaldo, L., Fazly, A.F., and Xu, X. (2007). Depletion of Med10 enhances Wnt and suppresses Nodal signaling during zebrafish embryogenesis. Dev Biol *303*, 536-548.

Liu, C.F., Bingham, N., Parker, K., and Yao, H.H. (2009). Sex-specific roles of beta-catenin in mouse gonadal development. Human molecular genetics *18*, 405-417.

Luo, X., Ikeda, Y., and Parker, K.L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. Cell *77*, 481-490.

Maatouk, D.M., DiNapoli, L., Alvers, A., Parker, K.L., Taketo, M.M., and Capel, B. (2008). Stabilization of β-catenin in XY gonads causes male-to-female sex-reversal. Human molecular genetics *17*, 2949-2955.

Mahendroo, M.S., Cala, K.M., and Russell, D.W. (1996). 5 alpha-reduced androgens play a key role in murine parturition. Molecular endocrinology (Baltimore, Md) *10*, 380-392.

Mandel, H., Shemer, R., Borochowitz, Z.U., Okopnik, M., Knopf, C., Indelman, M., Drugan, A., Tiosano, D., Gershoni-Baruch, R., Choder, M.*, et al.* (2008). SERKAL syndrome: an autosomal-recessive disorder caused by a loss-of-function mutation in WNT4. American journal of human genetics *82*, 39-47.

Matson, C.K., Murphy, M.W., Sarver, A.L., Griswold, M.D., Bardwell, V.J., and Zarkower, D. (2011). DMRT1 prevents female reprogramming in the postnatal mammalian testis. Nature *476*, 101.

Mazen, I., Abdel-Hamid, M., Mekkawy, M., Bignon-Topalovic, J., Boudjenah, R., El Gammal, M., Essawi, M., Bashamboo, A., and McElreavey, K. (2016). Identification of NR5A1 Mutations and Possible Digenic Inheritance in 46,XY Gonadal Dysgenesis. Sexual Development *10*, 147-151.

McElreavey, K., Vilain, E., Abbas, N., Herskowitz, I., and Fellous, M. (1993). A regulatory cascade hypothesis for mammalian sex determination: SRY represses a negative regulator of male development. Proc Natl Acad Sci U S A *90*, 3368-3372.

Menke, D.B., and Page, D.C. (2002). Sexually dimorphic gene expression in the developing mouse gonad. Gene expression patterns : GEP *2*, 359-367.

Moniot, B., Declosmenil, F., Barrionuevo, F., Scherer, G., Aritake, K., Malki, S., Marzi, L., Cohen-Solal, A., Georg, I., Klattig, J.*, et al.* (2009). The PGD2 pathway, independently of FGF9, amplifies SOX9 activity in Sertoli cells during male sexual differentiation. Development *136*, 1813-1821.

Morais da Silva, S., Hacker, A., Harley, V., Goodfellow, P., Swain, A., and Lovell-Badge, R. (1996). Sox9 expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. Nature genetics *14*, 62-68.

Munger, S.C., Aylor, D.L., Syed, H.A., Magwene, P.M., Threadgill, D.W., and Capel, B. (2009). Elucidation of the transcription network governing mammalian sex determination by exploiting strain-specific susceptibility to sex reversal. Genes Dev *23*, 2521-2536.

Munger, S.C., Natarajan, A., Looger, L.L., Ohler, U., and Capel, B. (2013). Fine time course expression analysis identifies cascades of activation and repression and maps a putative regulator of mammalian sex determination. PLoS genetics *9*, e1003630.

Nagamine, C.M., Morohashi, K., Carlisle, C., and Chang, D.K. (1999). Sex reversal caused by Mus musculus domesticus Y chromosomes linked to variant expression of the testis-determining gene Sry. Dev Biol *216*, 182-194.

Nagamine, C.M., Taketo, T., and Koo, G.C. (1987). Studies on the genetics of tda-1 XY sex reversal in the mouse. Differentiation; research in biological diversity *33*, 223-231.

Nation, T.R., Buraundi, S., Farmer, P.J., Balic, A., Newgreen, D., Southwell, B.R., and Hutson, J.M. (2011). Development of the gubernaculum during testicular descent in the rat. Anat Rec (Hoboken) *294*, 1249-1260.

Nef, S., and Parada, L.F. (1999). Cryptorchidism in mice mutant for Insl3. Nature genetics *22*, 295-299.

Nef, S., Schaad, O., Stallings, N.R., Cederroth, C.R., Pitetti, J.L., Schaer, G., Malki, S., Dubois-Dauphin, M., Boizet-Bonhoure, B., Descombes, P.*, et al.* (2005). Gene expression during sex determination reveals a robust female genetic program at the onset of ovarian development. Dev Biol *287*, 361-377.

Nef, S., Verma-Kurvari, S., Merenmies, J., Vassalli, J.D., Efstratiadis, A., Accili, D., and Parada, L.F. (2003). Testis determination requires insulin receptor family function in mice. Nature *426*, 291-295.

Nikolova, G., Sinsheimer, J.S., Eicher, E.M., and Vilain, E. (2008). The chromosome 11 region from strain 129 provides protection from sex reversal in XYPOS mice. Genetics *179*, 419-427.

Ocal, G., Berberoglu, M., Siklar, Z., Ruhi, H.I., Tukun, A., Camtosun, E., Savas Erdeve, S., Hacihamdioglu, B., and Fitoz, S. (2012). The clinical and genetic heterogeneity of mixed gonadal dysgenesis: does "disorders of sexual development (DSD)" classification based on new Chicago consensus cover all sex chromosome DSD? European journal of pediatrics *171*, 1497-1502.

Ottolenghi, C., Omari, S., Garcia-Ortiz, J.E., Uda, M., Crisponi, L., Forabosco, A., Pilia, G., and Schlessinger, D. (2005). Foxl2 is required for commitment to ovary differentiation. Human molecular genetics *14*, 2053-2062.

Ottolenghi, C., Pelosi, E., Tran, J., Colombino, M., Douglass, E., Nedorezov, T., Cao, A., Forabosco, A., and Schlessinger, D. (2007). Loss of Wnt4 and Foxl2 leads to female-to-male sex reversal extending to germ cells. Human molecular genetics *16*, 2795-2804.

Palmer, M.S., Sinclair, A.H., Berta, P., Ellis, N.A., Goodfellow, P.N., Abbas, N.E., and Fellous, M. (1989). Genetic evidence that ZFY is not the testis-determining factor. Nature *342*, 937-939.

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

Pannetier, M., Chassot, A.A., Chaboissier, M.C., and Pailhoux, E. (2016). Involvement of FOXL2 and RSPO1 in Ovarian Determination, Development, and Maintenance in Mammals. Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation *10*, 167-184.

Paris, F., Flatters, D., Caburet, S., Legois, B., Servant, N., Lefebvre, H., Sultan, C., and Veitia, R.A. (2017). A novel variant of DHH in a familial case of 46,XY disorder of sex development: Insights from molecular dynamics simulations. Clinical endocrinology *87*, 539-544.

Parma, P., Radi, O., Vidal, V., Chaboissier, M.C., Dellambra, E., Valentini, S., Guerra, L., Schedl, A., and Camerino, G. (2006). R-spondin1 is essential in sex determination, skin differentiation and malignancy. Nature genetics *38*, 1304-1309.

Pearlman, A., Loke, J., Le Caignec, C., White, S., Chin, L., Friedman, A., Warr, N., Willan, J., Brauer, D., Farmer, C*., et al.* (2010). Mutations in MAP3K1 cause 46,XY disorders of sex development and implicate a common signal transduction pathway in human testis determination. American journal of human genetics *87*, 898-904.

Pelletier, J., Bruening, W., Kashtan, C.E., Mauer, S.M., Manivel, J.C., Striegel, J.E., Houghton, D.C., Junien, C., Habib, R., Fouser, L*., et al.* (1991). Germline mutations in the Wilms' tumor suppressor gene are associated with abnormal urogenital development in Denys-Drash syndrome. Cell *67*, 437-447.

Pierucci-Alves, F., Clark, A.M., and Russell, L.D. (2001). A developmental study of the Desert hedgehog-null mouse testis. Biol Reprod *65*, 1392-1402.

Pitetti, J.L., Calvel, P., Romero, Y., Conne, B., Truong, V., Papaioannou, M.D., Schaad, O., Docquier, M., Herrera, P.L., Wilhelm, D*., et al.* (2013). Insulin and IGF1 receptors are essential for XX and XY gonadal differentiation and adrenal development in mice. PLoS genetics *9*, e1003160.

Porter, S., Clark, I.M., Kevorkian, L., and Edwards, D.R. (2005). The ADAMTS metalloproteinases. The Biochemical journal *386*, 15-27.

Prager, E.M., Boursot, P., and Sage, R.D. (1997). New assays for Y chromosome and p53 pseudogene clines among East Holstein house mice. Mammalian genome : official journal of the International Mammalian Genome Society *8*, 279-281.

Qi, H., Moran, M.M., Navarro, B., Chong, J.A., Krapivinsky, G., Krapivinsky, L., Kirichok, Y., Ramsey, I.S., Quill, T.A., and Clapham, D.E. (2007). All four CatSper ion channel proteins are required for male fertility and sperm cell hyperactivated motility. Proc Natl Acad Sci U S A *104*, 1219-1223.

Rahmoun, M., Lavery, R., Laurent-Chaballier, S., Bellora, N., Philip, G.K., Rossitto, M., Symon, A., Pailhoux, E., Cammas, F., Chung, J.*, et al.* (2017). In mammalian foetal testes, SOX9 regulates expression of its target genes by binding to genomic regions with conserved signatures. Nucleic Acids Research *45*, 7191-7211.

Raymond, C.S., Murphy, M.W., O'Sullivan, M.G., Bardwell, V.J., and Zarkower, D. (2000). Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. Genes Dev *14*, 2587-2595.

Ren, Y.A., Liu, Z., Mullany, L.K., Fan, C.M., and Richards, J.S. (2016). Growth Arrest Specific-1 (GAS1) Is a C/EBP Target Gene That Functions in Ovulation and Corpus Luteum Formation in Mice. Biol Reprod *94*, 44.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E.*, et al.* (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine : official journal of the American College of Medical Genetics *17*, 405-424.

Rosler, A., Silverstein, S., and Abeliovich, D. (1996). A (R80Q) mutation in 17 beta-hydroxysteroid dehydrogenase type 3 gene among Arabs of Israel is associated with pseudohermaphroditism in males and normal asymptomatic females. The Journal of clinical endocrinology and metabolism *81*, 1827-1831.

Russell, L.B., and Chu, E.H.Y. (1961). An XXY male in the mouse. Proceedings of the National Academy of Sciences of the United States of America *47*, 571-575.

Russell, W.L., Russell, L.B., and Gower, J.S. (1959). Exceptional inheritance of a sex-linked gene in the mouse explained on the basis that the X/O sex-chromosome constitution is female. Proceedings of the National Academy of Sciences of the United States of America *45*, 554-560.

Sakamoto, N., Oue, N., Noguchi, T., Sentani, K., Anami, K., Sanada, Y., Yoshida, K., and Yasui, W. (2010). Serial analysis of gene expression of esophageal squamous cell carcinoma: ADAMTS16 is upregulated in esophageal squamous cell carcinoma. Cancer science *101*, 1038-1044.

Sánchez-Moreno, I., Coral-Vázquez, R., Méndez, J.P., and Canto, P. (2008). Full-length SRY protein is essential for DNA binding. MHR: Basic science of reproductive medicine *14*, 325-330.

Sata, F., Kurahashi, N., Ban, S., Moriya, K., Tanaka, K.D., Ishizuka, M., Nakao, H., Yahata, Y., Imai, H., Kakizaki, H.*, et al.* (2010). Genetic polymorphisms of 17 beta-hydroxysteroid dehydrogenase 3 and the risk of hypospadias. The journal of sexual medicine *7*, 2729-2738.

Schmidt, D., Ovitt, C.E., Anlag, K., Fehsenfeld, S., Gredsted, L., Treier, A.C., and Treier, M. (2004). The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. Development *131*, 933-942.

Schmitt-Ney, M., Thiele, H., Kaltwasser, P., Bardoni, B., Cisternino, M., and Scherer, G. (1995). Two novel SRY missense mutations reducing DNA binding identified in XY females and their mosaic fathers. American journal of human genetics *56*, 862-869.

Schnellmann, R., Sack, R., Hess, D., Annis, D.S., Mosher, D.F., Apte, S.S., and Chiquet-Ehrismann, R. (2018). A Selective Extracellular Matrix Proteomics Approach Identifies Fibronectin Proteolysis by A Disintegrin-like and Metalloprotease Domain with Thrombospondin Type 1 Motifs (ADAMTS16) and Its Impact on Spheroid Morphogenesis. Mol Cell Proteomics *17*, 1410-1425.

Sekido, R., Bar, I., Narvaez, V., Penny, G., and Lovell-Badge, R. (2004). SOX9 is up-regulated by the transient expression of SRY specifically in Sertoli cell precursors. Dev Biol *274*, 271-279.

Sekido, R., and Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature *453*, 930-934.

Sharpe, R.M., and Skakkebaek, N.E. (2008). Testicular dysgenesis syndrome: mechanistic insights and potential new downstream effects. Fertility and sterility *89*, e33-38.

Shaw, G., and Renfree, M.B. (2014). Wolffian duct development. Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation *8*, 273-280.

Shindo, T., Kurihara, H., Kuno, K., Yokoyama, H., Wada, T., Kurihara, Y., Imai, T., Wang, Y., Ogata, M., Nishimatsu, H.*, et al.* (2000). ADAMTS-1: a metalloproteinase-disintegrin essential for normal growth, fertility, and organ morphology and function. The Journal of clinical investigation *105*, 1345-1352.

Siggers, P., Carre, G.A., Bogani, D., Warr, N., Wells, S., Hilton, H., Esapa, C., Hajihosseini, M.K., and Greenfield, A. (2014). A novel mouse Fgfr2 mutant, hobbyhorse (hob), exhibits complete XY gonadal sex reversal. PLoS One *9*, e100447.

Silver, L. (1995). (Oxford University Press).

Sim, N.-L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., and Ng, P.C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Research *40*, W452-W457.

Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A.-M., Lovell-Badge, R., and Goodfellow, P.N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature *346*, 240-244.

Singh, A.P., and Rajender, S. (2015). CatSper channel, sperm function and male fertility. Reproductive biomedicine online *30*, 28-38.

Skakkebaek, N.E., Rajpert-De Meyts, E., and Main, K.M. (2001). Testicular dysgenesis syndrome: an increasingly common developmental disorder with environmental aspects. Human reproduction (Oxford, England) *16*, 972-978.

Soito, I.C., Favorito, L.A., Costa, W.S., Sampaio, F.J., and Cardoso, L.E. (2011). Extracellular matrix remodeling in the human gubernaculum during fetal testicular descent and in cryptorchidic children. World J Urol *29*, 535-540.

Somerville, R.P., Longpre, J.M., Jungers, K.A., Engle, J.M., Ross, M., Evanko, S., Wight, T.N., Leduc, R., and Apte, S.S. (2003). Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS subfamily related to Caenorhabditis elegans GON-1. The Journal of biological chemistry *278*, 9503-9513.

Soutourina, J. (2018). Transcription regulation by the Mediator complex. Nature reviews Molecular cell biology *19*, 262-274.

Stanton, H., Melrose, J., Little, C.B., and Fosang, A.J. (2011). Proteoglycan degradation by the ADAMTS family of proteinases. Biochimica et biophysica acta *1812*, 1616-1629.

Sternberg , W.H., Barclay , D.L., and Kloepfer , H.W. (1968). Familial Xy Gonadal Dysgenesis. New England Journal of Medicine *278*, 695-700.

Stevant, I., Neirijnck, Y., Borel, C., Escoffier, J., Smith, L.B., Antonarakis, S.E., Dermitzakis, E.T., and Nef, S. (2018). Deciphering Cell Lineage Specification during Male Sex Determination with Single-Cell RNA Sequencing. Cell Rep *22*, 1589-1599.

Sun, J., Ting, M.C., Ishii, M., and Maxson, R. (2016). Msx1 and Msx2 function together in the regulation of primordial germ cell migration in the mouse. Dev Biol *417*, 11-24.

Svingen, T., and Koopman, P. (2013). Building the mammalian testis: origins, differentiation, and assembly of the component cell populations. Genes Dev *27*, 2409-2426.

Swain, A., and Lovell-Badge, R. (1999). Mammalian sex determination: a molecular drama. Genes Dev *13*, 755-767.

Takekawa, M., and Saito, H. (1998). A family of stress-inducible GADD45-like proteins mediate activation of the stress-responsive MTK1/MEKK4 MAPKKK. Cell *95*, 521-530.

Tam, P.P.L. (1981). The control of somitogenesis in mouse embryos. Journal of Embryology and Experimental Morphology *65*, 103-128.

Tevosian, S.G., Albrecht, K.H., Crispino, J.D., Fujiwara, Y., Eicher, E.M., and Orkin, S.H. (2002). Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. Development *129*, 4627-4634.

Thai, S.N., and Iruela-Arispe, M.L. (2002). Expression of ADAMTS1 during murine development. Mechanisms of development *115*, 181-185.

Thyen, U., Lanz, K., Holterhus, P.M., and Hiort, O. (2006). Epidemiology and initial management of ambiguous genitalia at birth in Germany. Hormone research *66*, 195-203.

Tomizuka, K., Horikoshi, K., Kitada, R., Sugawara, Y., Iba, Y., Kojima, A., Yoshitome, A., Yamawaki, K., Amagai, M., Inoue, A.*, et al.* (2008). R-spondin1 plays an essential role in ovarian development through positively regulating Wnt-4 signaling. Human molecular genetics *17*, 1278-1291.

Topol, L., Chen, W., Song, H., Day, T.F., and Yang, Y. (2009). Sox9 inhibits Wnt signaling by promoting beta-catenin phosphorylation in the nucleus. The Journal of biological chemistry *284*, 3323-3333.

Trussell, J.C., and Lee, P.A. (2004). The relationship of cryptorchidism to fertility. Current urology reports *5*, 142-148.

Uda, M., Ottolenghi, C., Crisponi, L., Garcia, J.E., Deiana, M., Kimber, W., Forabosco, A., Cao, A., Schlessinger, D., and Pilia, G. (2004). Foxl2 disruption causes mouse ovarian failure by pervasive blockage of follicle development. Human molecular genetics *13*, 1171-1181.

Uhlenhaut, N.H., Jakob, S., Anlag, K., Eisenberger, T., Sekido, R., Kress, J., Treier, A.C., Klugmann, C., Klasen, C., Holter, N.I.*, et al.* (2009). Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. Cell *139*, 1130-1142.

Umehara, F., Tate, G., Itoh, K., Yamaguchi, N., Douchi, T., Mitsuya, T., and Osame, M. (2000). A novel mutation of desert hedgehog in a patient with 46,XY partial gonadal dysgenesis accompanied by minifascicular neuropathy. American journal of human genetics *67*, 1302-1305.

Vainio, S., Heikkila, M., Kispert, A., Chin, N., and McMahon, A.P. (1999). Female development in mammals is regulated by Wnt-4 signalling. Nature *397*, 405-409.

Van der Schoot, P., and Elger, W. (1992). Androgen-induced prevention of the outgrowth of cranial gonadal suspensory ligaments in fetal rats. J Androl *13*, 534-542.

Van der Schoot, P., and Emmen, J.M. (1996). Development, structure and function of the cranial suspensory ligaments of the mammalian gonads in a cross-species perspective; their possible role in effecting disturbed testicular descent. Hum Reprod Update *2*, 399-418.

Vazquez, F., Hastings, G., Ortega, M.A., Lane, T.F., Oikemus, S., Lombardo, M., and Iruela-Arispe, M.L. (1999). METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity. The Journal of biological chemistry *274*, 23349-23357.

Vidal, V.P., Chaboissier, M.C., de Rooij, D.G., and Schedl, A. (2001). Sox9 induces testis development in XX transgenic mice. Nature genetics *28*, 216-217.

Vilain, E., McElreavey, K., Jaubert, F., Raymond, J.P., Richaud, F., and Fellous, M. (1992). Familial case with sequence variant in the testis-determining region associated with two sex phenotypes. American journal of human genetics *50*, 1008-1011.

Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasantes, J., Bricarelli, F.D., Keutel, J., Hustert, E.*, et al.* (1994). Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. Cell *79*, 1111-1120.

Wang, W.M., Lee, S., Steiglitz, B.M., Scott, I.C., Lebares, C.C., Allen, M.L., Brenner, M.C., Takahara, K., and Greenspan, D.S. (2003). Transforming growth

factor-beta induces secretion of activated ADAMTS-2. A procollagen III N-proteinase. The Journal of biological chemistry *278*, 19549-19557.

Warr, N., Carre, G.-A., Siggers, P., Faleato, Jessica V., Brixey, R., Pope, M., Bogani, D., Childers, M., Wells, S., Scudamore, Cheryl L*., et al.* (2012). Gadd45γ and Map3k4 Interactions Regulate Mouse Testis Determination via p38 MAPK-Mediated Control of Sry Expression. Developmental Cell *23*, 1020-1031.

Warr, N., Siggers, P., Bogani, D., Brixey, R., Pastorelli, L., Yates, L., Dean, C.H., Wells, S., Satoh, W., Shimono, A*., et al.* (2009). Sfrp1 and Sfrp2 are required for normal male sexual development in mice. Dev Biol *326*, 273-284.

Warr, N., Siggers, P., Carre, G.A., Wells, S., and Greenfield, A. (2016). Genetic Analyses Reveal Functions for MAP2K3 and MAP2K6 in Mouse Testis Determination. Biol Reprod *94*, 103.

Washburn, L.L., and Eicher, E.M. (1983). Sex reversal in XY mice caused by dominant mutation on chromosome 17. Nature *303*, 338-340.

Werner, R., Merz, H., Birnbaum, W., Marshall, L., Schroder, T., Reiz, B., Kavran, J.M., Baumer, T., Capetian, P., and Hiort, O. (2015). 46,XY Gonadal Dysgenesis due to a Homozygous Mutation in Desert Hedgehog (DHH) Identified by Exome Sequencing. The Journal of clinical endocrinology and metabolism *100*, E1022-1029.

Whitney, J.B., Mills, T.M., Lewis, R.W., Wartell, R., and Abney, T.O. (2000). A single genetic determinant that prevents sex reversal in C57BL-YPOS congenic mice. Biochemical genetics *38*, 119-137.

Whitney, M.A., Royle, G., Low, M.J., Kelly, M.A., Axthelm, M.K., Reifsteck, C., Olson, S., Braun, R.E., Heinrich, M.C., Rathbun, R.K*., et al.* (1996). Germ cell defects and hematopoietic hypersensitivity to gamma-interferon in mice with a targeted disruption of the Fanconi anemia C gene. Blood *88*, 49-58.

Wilhelm, D., and Englert, C. (2002). The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1. Genes Dev *16*, 1839-1851.

Wilhelm, D., Hiramatsu, R., Mizusaki, H., Widjaja, L., Combes, A.N., Kanai, Y., and Koopman, P. (2007). SOX9 regulates prostaglandin D synthase gene transcription in vivo to ensure testis development. The Journal of biological chemistry *282*, 10553-10560.

Wilhelm, D., Martinson, F., Bradford, S., Wilson, M.J., Combes, A.N., Beverdam, A., Bowles, J., Mizusaki, H., and Koopman, P. (2005). Sertoli cell differentiation is induced both cell-autonomously and through prostaglandin signaling during mammalian sex determination. Dev Biol *287*, 111-124.

Wilhelm, D., Washburn, L.L., Truong, V., Fellous, M., Eicher, E.M., and Koopman, P. (2009). Antagonism of the testis- and ovary-determining pathways during ovotestis development in mice. Mechanisms of development *126*, 324-336.

Worley, J.R., Baugh, M.D., Hughes, D.A., Edwards, D.R., Hogan, A., Sampson, M.J., and Gavrilovic, J. (2003). Metalloproteinase expression in PMA-stimulated THP-1 cells. Effects of peroxisome proliferator-activated receptor-gamma (PPAR gamma) agonists and 9-cis-retinoic acid. The Journal of biological chemistry *278*, 51340-51346.

Wright, E., Hargrave, M.R., Christiansen, J., Cooper, L., Kun, J., Evans, T., Gangadharan, U., Greenfield, A., and Koopman, P. (1995). The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. Nature genetics *9*, 15-20.

Wright, E.M., Snopek, B., and Koopman, P. (1993). Seven new members of the Sox gene family expressed during mouse development. Nucleic Acids Res *21*, 744.

Wright, W.W., Smith, L., Kerr, C., and Charron, M. (2003). Mice that express enzymatically inactive cathepsin L exhibit abnormal spermatogenesis. Biol Reprod *68*, 680-687.

Wu, Y.Y., Yang, Y., Xu, Y.D., and Yu, H.L. (2015). Targeted disruption of the spermatid-specific gene Spata31 causes male infertility. Mol Reprod Dev *82*, 432-440.

Yang, S.H., Sharrocks, A.D., and Whitmarsh, A.J. (2003). Transcriptional regulation by the MAP kinase signaling cascades. Gene *320*, 3-21.

Yao, H.H., Matzuk, M.M., Jorgez, C.J., Menke, D.B., Page, D.C., Swain, A., and Capel, B. (2004). Follistatin operates downstream of Wnt4 in mammalian ovary organogenesis. Dev Dyn *230*, 210-215.

Yao, H.H., Whoriskey, W., and Capel, B. (2002). Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. Genes Dev *16*, 1433-1440.

Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., Ma, Z., Davis, C., Pope, B.D.*, et al.* (2014). A comparative encyclopedia of DNA elements in the mouse genome. Nature *515*, 355.

Zhang, F., and Lupski, J.R. (2015). Non-coding genetic variants in human disease. Human molecular genetics *24*, R102-R110.

Zhao, F., Franco, H.L., Rodriguez, K.F., Brown, P.R., Tsai, M.J., Tsai, S.Y., and Yao, H.H. (2017). Elimination of the male reproductive tract in the female embryo is promoted by COUP-TFII in mice. Science (New York, NY) *357*, 717-720.

Zhao, L., Ng, E.T., Davidson, T.L., Longmuss, E., Urschitz, J., Elston, M., Moisyadi, S., Bowles, J., and Koopman, P. (2014a). Structure-function analysis of mouse Sry reveals dual essential roles of the C-terminal polyglutamine tract in sex determination. Proc Natl Acad Sci U S A *111*, 11768-11773.

Zhao, X., Onteru, S., Saatchi, M., Garrick, D., and Rothschild, M. (2014b). A genome-wide association study for canine cryptorchidism in Siberian Huskies. Journal of animal breeding and genetics = Zeitschrift fur Tierzuchtung und Zuchtungsbiologie *131*, 202-209.

Appendix 1. Validation of SNPs in B6.Y^{POS}

Table S1 Chromosome 13 variants confirmed by Sanger sequencing

SNP ID	Position	Functional class	Gene	Primers	Product size	Reference C57BL/6J	Harwell C57BL/6J	Harwell C57BL/6J. MOD13	Valid
rs30098847	13 : 39167018	Intergenic variant	.	GACCTCCGAGGGGGCATTTA GATGGAGTCCTGGTTGAGGG	492 bp	G	G	A	yes
rs29231478	13 : 40958356	3 prime UTR variant	*Gcnt2*	TCCTGGTGGACTAGGGTAGC TTGCCAGATCCTCACGGTTC	465 bp	G	G	C	yes
rs13481788	13 : 43632644	Intron variant	*Rnf182*	GCAAGGCACACGTTAGTAGG AGAGTTACTCATCAAACGAGCGA	490 bp	T	T	C	yes
rs29832595	13 : 44920218	3 prime UTR variant	*Jarid2*	TAGGATACAAGCTGTCTTTGCAC TACAAACTCCCTTCGCACATGA	464 bp	T	T	A	yes
rs29715039	13 : 46058241	Intergenic variant	.	TGTGCATACCATGAGCCGATT GTGATTCACACCCTCGCCTC	225 bp	G	G	A	yes
rs29535407	13 : 47027645	Intron variant	*Tpmt*	TTGTGGGGTATGTGTGTGGG GTTTGGGGGAGAAGGTTGCT	235 bp	A	A	G	yes
rs30009714	13 : 47703466	Intergenic variant	.	CAGGGAGCAAACTAAGGCCA CCCACAATGATGCTCCACCT	457 bp	T	T	C	yes
rs29831223	13 : 47703494	Intergenic variant	.			G	G	A	yes
rs45665661	13 : 47703555	Intergenic variant	.			C	C	T	yes
rs49195987	13 :	Intergenic	.			T	T	C	yes

SNP ID	Position	Functional class	Gene	Primers	Product size	Reference C57BL/6J	Harwell C57BL/6J	Harwell C57BL/6J. MOD13	Valid
	47703618	variant							
rs47004668	13 : 47703697	Intergenic variant	.			C	C	A	yes
rs51092835	13 : 47703763	Intergenic variant	.			A	A	T	yes
rs29568118	13 : 48119885	Intron variant	.	GCACACTCCAGCCGAATCA TCCTAAATGATCCCACTCGGC	331 bp	C	C	T	yes
rs48656322	13 : 52288276	Intergenic variant	.	TCGATTCCAAGCCCTTTACCA TCAAAAACCGCCCACAATTCA	449 bp	T	T	C	yes
rs30015565	13 : 53263711	Intron variant	*Ror2*	ATCGGAGGCAGGTCTAGTGG GGTGCCAGCCTATGTTGAAG	459 bp	A	A	G	yes
rs29615746	13 : 54473304	Upstream gene variant	*Thoc3*	CCTCCGACACTGGTGTCTTC AACTGCACCCAGAAATGCCT	208 bp	A	A	G	yes
rs29783542	13 : 54473312	Upstream gene variant	*Thoc3*			G	G	A	yes
rs29244641	13 : 54697741	Upstream gene variant	*Rnf44*	ACAGGACCCAGATCTACCTAGTG TTCGGGCTCCCCAGATGTAAC	273 bp	T	T	C	yes
rs48574973	13 : 55001410	Intron variant	*Unc5a*	GATGGTGGGTATTCCGTGGC CGGAGAGCAGCATTGAGGTC	564 bp	A	A	G	yes
rs48644240	13 : 55001423	Intron variant	*Unc5a*			A	A	C	yes
rs227636889	13 : 55001462	Intron variant	.			C	C	-	yes
rs52545671	13 :	Downstream	*Fgfr4*	GCAAGCGAACACACCTCTTG	395 bp	T	T	C	yes

SNP ID	Position	Functional class	Gene	Primers	Product size	Reference C57BL/6J	Harwell C57BL/6J	Harwell C57BL/6J. MOD13	Valid
	55173589	gene variant		CTCAGACACTGTGGTTCCCC					
rs46080380	13 : 55173677	Downstream gene variant	*Fgfr4*			G	G	A	yes
rs29245839	13 : 55173691	Downstream gene variant	*Fgfr4*			T	T	G	yes
rs29238368	13 : 55400536	Missense variant	*Slc34a1*	CAGCAACTGTCAGATTTGTCCC CCAGCTTGGTAGGTAGTGGG	416 bp	G	G	A	yes
rs29554650	13 : 56852649	Intron variant	*Trpc7*	TACTCTGGGGTCAGCCTTGA AAGTGCATCTGACTGCTCCC	484 bp	G	G	A	yes
rs29530238	13 : 56852756	Intron variant	*Trpc7*			T	T	C	yes
rs29709494	13 : 56852809	Intron variant	*Trpc7*			A	A	G	yes
rs29570819	13 : 58031276	Intron variant	*Klhl3*	CAAAGAGCCCCTTCGGTGTG AGCTCCTAAAGGGCATTGACA	326 bp	G	G	A	yes
rs29781657	13 : 58031477	Intron variant	*Klhl3*			A	A	G	yes
rs265656437	13 : 59973936	Intergenic variant	.	AAGACCCTCCTTCAGTCGAG GATCAGCAAGGCCAGAACATAC	487 bp	G	G	A	yes
rs262331200	13 : 59973988	Intergenic variant	.			C	C	T	yes
rs30239787	13 : 60078238	Intergenic variant	.	TCCTTTCGCACCTTGCAGATT GAACAGTAGGTCTCTCCGGG	469 bp	T	T	A	yes
rs30239786	13 :	Intergenic	.			C	C	A	yes

SNP ID	Position	Functional class	Gene	Primers	Product size	Reference C57BL/6J	Harwell C57BL/6J	Harwell C57BL/6J. MOD13	Valid
	60078303	variant							
rs30239785	13 : 60078525	Intergenic variant	.			T	T	C	yes
rs51454148	13 : 60212468	5 prime UTR variant	.	TGTAACTCATGACGTTCTAACCAAA CTCAGTCTGCTCTCCTCGGT	380 bp	G	G	A	yes
rs107770014	13 : 60767190	Upstream gene variant	.	ACAGGAAAGGCATATCGACAACAG G GGGTTAGCTTTAGGCATGGGTG	479 bp	C	C	T	yes
rs241205325	13 : 60767201	Upstream gene variant	.			G	G	A	yes
rs224742780	13 : 60767514	Upstream gene variant	.			T	T	C	yes
rs108884768	13 : 60767517	Upstream gene variant	.			T	T	A	yes
rs262601145	13 : 60767529	Upstream gene variant	.			T	T	A	yes
rs230181856	13 : 60767540	Upstream gene variant	.			T	T	C	yes
rs244215635	13 : 60767542	Upstream gene variant	.			G	G	A	yes
rs253950006	13 : 61457774	Intergenic variant	.	AGGATGACAGTGTTGGCTACAGGT TCCTTCTACAGTGCCAGTGTCAT	589 bp	T	T	G	yes
rs6331667	13 : 61538538	Synonymous variant	*Ctsm*	GTCTGTTGAACCGAAAGGCA CCCAGTAGCATTTTATTTCCAGGC	384 bp	T	T	C	yes
rs6331675	13 :	Missense	*Ctsm*			A	A	T	yes

SNP ID	Position	Functional class	Gene	Primers	Product size	Reference C57BL/6J	Harwell C57BL/6J	Harwell C57BL/6J. MOD13	Valid
	61538544	variant							
rs6344510	13 : 61538569	Intron variant	*Ctsm*			G	G	A	yes
rs6344577	13 : 61538599	Intron variant	*Ctsm*			A	A	G	yes
rs108034900	13 : 61538604	Intron variant	*Ctsm*			G	G	A	yes
.	13 : 61820491	Downstream gene variant	.	GGTGCCACTGTGGAAACCTT TCAATGGACCTCATAGGGCT	313 bp	C	C	C	no
.	13 : 62127556	Upstream gene variant	*Zfp808*	CGGAAAACCAGTCAGTGCTC CCCACCCTGTTTACCCCATT	338 bp	G	G	G	no
rs224110566	13 : 62773685	Stop gained	*Gm5141*	GTGAAAAGGCTCTATCACGTTG TCACGACACTTTCACCTCCAAAATC A	256 bp	G	G	A	yes
rs46960049	13 : 63512038	Missense variant	*Ptch1*	GTCTCTGCGCGGTCTGTAG AGTTGAGATCTATGGGCAGTGG	319 bp	G	G	T	yes
rs29969295	13 : 64761844	Intron variant	*Cntnap3*	GTTTGCTTCTTGGCCACCAG CACAGGCCGACCATACTCTG	357 bp	T	T	C	yes
rs48004575	13 : 64761911	Splice region variant	*Cntnap3*			A	A	G	yes
rs29829962	13 : 66999945	Downstream gene variant	*Ptdss1*	TTGCTGCAACTGCCTCCTAA GCAGGAGCCAGAGGTCTTTT	486 bp	T	T	C	yes
rs49425444	13 : 67496018	Intron variant	*Zfp58*	AGGAAGAGAACACAGAGCAGG CCCTGGATGCTGGATGTGTG	599 bp	C	C	T	yes

SNP ID	Position	Functional class	Gene	Primers	Product size	Reference C57BL/6J	Harwell C57BL/6J	Harwell C57BL/6J. MOD13	Valid
rs108728730	13 : 68013715	Intergenic variant	.	TCCTTTGGATCGCATGGACA TGGTGTGCATCCACCTTCTT	425 bp	A	A	C	yes
rs46311648	13 : 68591842	Intron variant	*Fastkd3*	CTCAGACTCCCCTCGGTACT TCACCGAGAATCCTGAAGCC	339 bp	G	G	G	no
rs49768875	13 : 68592015	Missense variant	*Fastkd3*			G	G	G	no
rs29631363	13 : 69003102	Upstream gene variant	*Adcy2*	AGGGCAGGTTTACGCACAAG CACAGTGCTAAGTGACCGTGA	472 bp	A	A	G	yes
rs29251922	13 : 69504728	Synonymous variant	*Papd7*	AGAGGAGCCTAGAACCCAGG TCGCTCGGGTCTGTGTTTAC	278 bp	A	A	G	yes
rs29245244	13 : 69504739	Splice region variant	*Papd7*			C	C	T	yes
rs52465159	13 : 70780126	Intron variant	*Adamts16*	GTCTAACGTGAGGGTTTCCCC AGAAAGTCCTTGAAGAGGGCG	338 bp	T	C	C	no
rs29802787	13 : 70164398	Upstream gene variant	.	GCCAGTGGGCACAGAGTTAT GCCATCTTCTCAGGCTGTGT	304 bp	G	G	A	yes
rs50180126	13 : 71076538	Intergenic variant	.	GGCTGGAGAGATGGCTTTGT TTCCCTGCTGGCTTTCTTGT	252 bp	C	C	T	yes
rs49008638	13 : 71076570	Intergenic variant	.			G	G	A	yes
rs49318132	13 : 71076606	Intergenic variant	.			C	C	A	yes
rs13472522	13 : 71959581	Synonymous variant	*Irx1*	TGACAGGTGTACAGTCCGTG ACCAGAGCAACGAAGACGAA	385 bp	A	A	G	yes

SNP ID	Position	Functional class	Gene	Primers	Product size	Reference C57BL/6J	Harwell C57BL/6J	Harwell C57BL/6J. MOD13	Valid
rs51554283	13 : 72107179	Intergenic variant	.	GGCATCCCAAAGGAACACTCT GCTTGGATGTCGTAAGTTCCTC	303 bp	C	C	C	no
rs47725203	13 : 72107295	Intergenic variant	.			G	G	G	no
rs46415947	13 : 72107329	Intergenic variant	.			T	T	T	no
rs52003699	13 : 72107356	Intergenic variant	.			C	C	C	no
rs29223975	13 : 74740014	Intron variant	*Cast*	AGGTCATCCAAGGCAGCATC GCCACATGGCTAATCCTTGC	335 bp	C	C	T	yes
rs29685106	13 : 74740051	Intron variant	*Cast*			T	T	G	yes
rs29765784	13 : 74740122	Intron variant	*Cast*			A	A	T	yes
rs29535792	13 : 74740135	Intron variant	*Cast*			A	A	G	yes
rs30026943	13 : 74740166	Intron variant	*Cast*			C	C	A	yes
rs29772787	13 : 75688132	Upstream gene variant	.	TCTCAGTGAATAACCCCAGTAGC CCAATAACTGAGAAGTGGCTGG	482 bp	C	C	T	yes

Table S2 Chromosome 4 variants confirmed by Sanger sequencing

SNP ID	Position	Functional class	Gene	Primers	Product size	Reference C57BL/6J	Harwell C57BL/6J	Harwell C57BL/6J. MOD13	Valid
.	4 : 147058183	Missense variant	*Rex2*	.	.	G	too repetitive to sequence		
rs47033754	4 : 147063572	Intron variant	.	TATTTCCCTGGTGGTTGGAA GAAGCCATGTTGTGCTCAGA	372 bp	A	T	T	no
rs47172221	4 : 147613029	Stop lost	*2610305D 13Rik*	TTCTCTGATGACTTTTCAGATGGT CCCACAAATTTAATCTGAGAACA	285 bp	A	A/T	A/T	no
rs50412577	4 : 147613220	Missense variant	.	TGCATTATTAAGCAAAGCACTTG GCTTTATGCAACAATCCCTTC	354 bp	C	C	C	no
rs243923553	4 : 147674273	Missense variant	.	AGGCCAAAAGGAAAGGATTC TGCTTTACTACAAAAGGCAATCTG	351 bp	T	A/T	A/T	no
.	4 : 147674467	Missense variant	*Zfp534*	.	.	C	too repetitive to sequence		

Appendix 2. Alignment of *Ptch1* sequences

Key: SNP
Indel

```
C57BL/6J    TTTAAAAAACTTCTCAACTTTAAGTCAGTGTCTCTTTGTGTGTCTGTGTGTGGGGTGTGC  60
mod         TTTAAAAAACTTCTCAACTTTAAGTCAGTGTCTCTTTGTGTGTCTGTGTGTGGGGTGTGC  60
BALB/cJ     TTTAAAAAACTTCTCAACTTTAAGTCAGTGTCTCTTTGTGTGTCTGTGTGTGGGGTGTGC  60
            ************************************************************

C57BL/6J    ACGTGTATGCAGGTGCCAGTAGAGGCCAGAAGAGGGCATCAGATCCCCTGGAGCGTGAGG  120
mod         ACGTGTATGCAGGTGCCAGTAGAGGCCAGAAGAGGGCATCAGATCCCCTGGAGCGTGAGG  120
BALB/cJ     ACGTGTATGCAGGTGCCAGTAGAGGCCAGAAGAGGGCATCAGATCCCCTGGAGCGTGAGG  120
            ************************************************************

C57BL/6J    CAACAGTTCTGAGTGGCTCTCTGTAAGAGTAACTGTCACTCCTAACCAACCACTGATTGA  180
mod         CAACAGTTCTGAGTGGCTCTCTGTAAGAGTAACTGTCACTCCTAACCAACCACTGATTGA  180
BALB/cJ     CAACAGTTCTGAGTGGCTCTCTGTAAGAGTAACTGTCACTCCTAACCAACCACTGATTGA  180
            ************************************************************

C57BL/6J    CCTTTGCAGGGGTCACTTGTGTTTACACTGAATTTGGCTGTATTCACTTCCCGCGTGCCT  240
mod         CCTTTGCAGGGGTCACTTGTGTTTACACTGAATTTGGCTGTATTCACTTCCCGCGTGCCT  240
BALB/cJ     CCTTTGCAGGGGTCACTTGTGTTTACACTGAATTTGGCTGTATTCACTTCCCGCGTGCCT  240
            ************************************************************

C57BL/6J    ACATAGAAAGGTAAGAAAAGGGAACGCTCAGAAGAGTCATTAGGTGACTTCCCCAGTGTC  300
mod         ACATAGAAAGGTAAGAAAAGGGAACGCTCAGAAGAGTCATTAGGTGACTTCCCCAGTGTC  300
BALB/cJ     ACATAGAAAGGTAAGAAAAGGGAACGCTCAGAAGAGTCATTAGGTGACTTCCCCAGTGTC  300
            ************************************************************

C57BL/6J    CACAAAATACACATAAAACAGCTTCCGTAAGGCCAGCAAGAAGAAACGAACAGACGGCCT  360
mod         CACAAAATACACATAAAACAGCTTCCGTAAGGCCAGCAAGAAGAAACGAACAGACGGCCT  360
BALB/cJ     CACAAAATACACATAAAACAGCTTCCGTAAGGCCAGCAAGAAGAAACGAACAGACGGCCT  360
            ************************************************************

C57BL/6J    GGAACTTCTTGTGTATATGTAGTTTAAGAAACATTCTAGAAACAAGCACATAAACCAACA  420
mod         GGAACTTCTTGTGTATATGTAGTTTAAGAAACTTTCTAGAAACAAGCACATAAACCAACA  420
BALB/cJ     GGAACTTCTTGTGTATATGTAGTTTAAGAAACTTTCTAGAAACAAGCACATAAACCAACA  420
            ******************************* ****************************

C57BL/6J    AGTTATATTATCTTTATTCATGTTAGAATTTCTATCAGTGTTTGAACATTAATATTTAGA  480
mod         AGTTATATTATCTTTATTCATGTTAGAATTTCTATCAGTGTTTGAACATTAATATTTAGA  480
BALB/cJ     AGTTATATTATCTTTATTCATGTTAGAATTTCTATCAGTGTTTGAACATTAATATTTAGA  480
            ************************************************************

C57BL/6J    GAGTTTTGTACATTTTCATCTGAGAACCGCACTAGGTCACCATGTTCACACCGATACTTT  540
mod         GAGTTTTGTACATTTTCATCTGAGAACCGCACTAGGTCACCATGTTCACACCGATACTTT  540
BALB/cJ     GAGTTTTGTACATTTTCATCTGAGAACCGCACTAGGTCACCATGTTCACACCGATACTTT  540
            ************************************************************

C57BL/6J    CCAATCACATGGATCACTGCTCACGTGTACTCTCCTAACACTGTTAAAACATACTCTCTT  600
mod         CCAATCACATGGATCACTGCTCACGTGTACTCTCCTAACACTGTTAAAACATACTCTCTT  600
BALB/cJ     CCAATCACATGGATCACTGCTCACGTGTACTCTCCTAACACTGTTAAAACATACTCTCTT  600
            ************************************************************

C57BL/6J    CTCTCCCTTGCTGTCATTTATAACGAAACGTGTCCTCACACCCATGTCTCAGCAAAGTTC  660
mod         CTCTCCCTTGCTGTCATTTATAACGAAACGTGTCCTCACACCCATGTCTCAGCAAAGTTC  660
BALB/cJ     CTCTCCCTTGCTGTCATTTATAACGAAACGTGTCCTCACACCCATGTCTCAGCAAAGTTC  660
            ************************************************************

C57BL/6J    CAGACATTATGGATTTCATCACATATAAATTCTTTAAAAATATACTTCTGTCGAAAGACT  720
mod         CAGACATTATGGATTTCATCACATATAAATTCTTTAAAAATATACTTCTGTCGAAAGACT  720
BALB/cJ     CAGACATTATGGATTTCATCACATATAAATTCTTTAAAAATATACTTCTGTCGAAAGACT  720
            ************************************************************
```

```
C57BL/6J   TGATGACTACTATGTACACTACAAAAATAAATTTTCATATAAATAAATTATATGGCATAC  780
mod        TGATGACTACTATGTACACTACAAAAATAAATTTTCATATAAATAAATTATATGGCATAC  780
BALB/cJ    TGATGACTACTATGTACACTACAAAAATAAATTTTCATATAAATAAATTATATGGCATAC  780
           ************************************************************

C57BL/6J   TTTCATCTTTGTAATTGAAGTGACACAAACTCATTTCCACCCAAATTGGCAGACGAGACC  840
mod        TTTCATCTTTGTAATTGAAGTGACACAAACTCATTTCCACCCAAATTGGCAGACGAGACC  840
BALB/cJ    TTTCATCTTTGTAATTGAAGTGACACAAACTCATTTCCACCCAAATTGGCAGACGAGACC  840
           ************************************************************

C57BL/6J   CAGTTTGAATATTTATTTCCTTTAAACTCCTTACACTAAAACCTACCATGAGTCCTTAAT  900
mod        CAGTTTGAATATTTATTTCCTTTAAACTCCTTACACTAAAACCTACCATGAGTCCTTAAT  900
BALB/cJ    CAGTTTGAATATTTATTTCCTTTAAACTCCTTACACTAAAACCTACCATGAGTCCTTAAT  900
           ************************************************************

C57BL/6J   GTAAATGTTATATACTCTGAACTATTTAACATTAGTAAGCACTCTATACAAATAAAACTT  960
mod        GTAAATGTTATATACTCTGAACTATTTAACATTAGTAAGCACTCTATACAAATAAAACTT  960
BALB/cJ    GTAAATGTTATATACTCTGAACTATTTAACATTAGTAAGCACTCTATACAAATAAAACTT  960
           ************************************************************

C57BL/6J   CTGTTCAAAAGTAAAACATAACTGCAATAAGTGTTTCCAATAGAAATAAAATCCCTGCAA 1020
mod        CTGTTCAAAAGTAAAACATAACTGCAATAAGTGTTTCCAATAGAAATAAAATCCCTGCAA 1020
BALB/cJ    CTGTTCAAAAGTAAAACATAACTGCAATAAGTGTTTCCAATAGAAATAAAATCCCTGCAA 1020
           ************************************************************

C57BL/6J   TTTTTTCTTCTTGTTATAAACTGGCATGACAAATTCTGGAGGACATTCTTATGATACAAA 1080
mod        TTTTTTCTTCTTGTTATAAACTGGCATGACAAATTCTGGAGGACATTCTTATGATACAAA 1080
BALB/cJ    TTTTTTCTTCTTGTTATAAACTGGCATGACAAATTCTGGAGGACATTCTTATGATACAAA 1080
           ************************************************************

C57BL/6J   ATTATAAATATCAGCAAAGATCTTTTCTTTTACAAATTAAAACATCCAAATGAACAAACA 1140
mod        ATTATAAATATCAGCAAAGATCTTTTCTTTTACAAATTAAAACATCCAAATGAACAAACA 1140
BALB/cJ    ATTATAAATATCAGCAAAGATCTTTTCTTTTACAAATTAAAACATCCAAATGAACAAACA 1140
           ************************************************************

C57BL/6J   ATACAGACAAAATAAAACTATGAGAGAGACAGACAGAGACACACACTCTCTGACTTTAGT 1200
mod        ATACAGACAAAATAAAACTATGAGAGAGACAGACAGAGACACACACTCTCTGACTTTAGT 1200
BALB/cJ    ATACAGACAAAATAAAACTATGAGAGAGACAGACAGAGACACACACTCTCTGACTTTAGT 1200
           ************************************************************

C57BL/6J   CTCCTGCCAGCTCTGTGTCCTTCCTGCCTTGGCTGATGGGACTCATGATGATGCACAGAA 1260
mod        CTCCTGCCAGCTCTGTGTCCTTCCTGCCTTGGCTGATGGGACTCATGATGATGCACAGAA 1260
BALB/cJ    CTCCTGCCAGCTCTGTGTCCTTCCTGCCTTGGCTGATGGGACTCATGATGATGCACAGAA 1260
           ************************************************************

C57BL/6J   GCCGAGAGGCTGTACAAACTTCCTTCTGCCACATGCACAAGGCTCAGGCATGAGGGACAC 1320
mod        GCCGAGAGGCTGTACAAACTTCCTTCTGCCACATGCACAAGGCTCAGGCATGAGGGACAC 1320
BALB/cJ    GCCGAGAGGCTGTACAAACTTCCTTCTGCCACATGCACAAGGCTCAGGCATGAGGGACAC 1320
           ************************************************************

C57BL/6J   AATCAAGAACCACAGCCAAGGTGGCTTCCGCACTCTACATCTCCCTCCTCAACTCAGGGT 1380
mod        AATCAAGAACCACAGCCAAGGTGGCTTCCGCACTCTACATCTCCCTCCTCAACTCAGGGT 1380
BALB/cJ    AATCAAGAACCACAGCCAAGGTGGCTTCCGCACTCTACATCTCCCTCCTCAACTCAGGGT 1380
           ************************************************************

C57BL/6J   GCAAGATATCCGAGCAGAGCTGGTGTTGGGTAGGCAGACAGTGGCTTGACAGGCATTAAG 1440
mod        GCAAGATATCCGAGCAGAGCTGGTGTTGGGTAGGCAGACAGTGGCTTGACAGGCATTAAG 1440
BALB/cJ    GCAAGATATCCGAGCAGAGCTGGTGTTGGGTAGGCAGACAGTGGCTTGACAGGCATTAAG 1440
           ************************************************************

C57BL/6J   ACAGTAGCAGTGGATGGCCAGGGCTTCTGGCGCGCACGCGCACACGGATATACAGACAGA 1500
mod        ACAGTAGCAGTGGATGGCCAGGGCCTCTGGCGCGCACGCGCACACGGATATACAGACAGA 1500
BALB/cJ    ACAGTAGCAGTGGATGGCCAGGGCCTCTGGCGCGCACGCGCACACGGATATACAGACAGA 1500
           ********************** *********************************
```

```
C57BL/6J   CACAGACA-CACACACACACGGTAGTCACAGGTGGCTTCCAATTAGTACCTTACAATTT   1558
mod        CACAGACACACACACACACGGTAGTCACAGGTGGCTTCCAATTAGTACCTTACAATTT  1560
BALB/cJ    CACAGACACACACACACACGGTAGTCACAGGTGGCTTCCAATTAGTACCTTACAATTT  1560
           ********  **************************************************

C57BL/6J   GGGGGATAAATTTTAGGATCTGAGGGCACTAGATTATTGGTACTACTAGGGACCCATCTC  1618
mod        GGGGGATAAATTTTAGGATCTGAGGGCACTAGATTATTGGTACTACTAGGGACCCATCTC  1620
BALB/cJ    GGGGGATAAATTTTAGGATCTGAGGGCACTAGATTATTGGTACTACTAGGGACCCATCTC  1620
           ************************************************************

C57BL/6J   CTTAAATTTTTCCATGAAGGATAGGAGGGATTCAGGGAAGCTGCATGCAGACAGAGTACA  1678
mod        CTTAAATTTTTCCATGAAGGATAGGAGGGATTCAGGGAAGCTGCATGCAGACAGAGTACA  1680
BALB/cJ    CTTAAATTTTTCCATGAAGGATAGGAGGGATTCAGGGAAGCTGCATGCAGACAGAGTACA  1680
           ************************************************************

C57BL/6J   TTCCTCACAACAGCTTCCATGAGCACTAGGCTTCCTTCCTCTGAGACAAGTCCTCAATGT  1738
mod        TTCCTCACAACAGCTTCCATGAGCACTAGGCTTCCTTCCTCTGAGACAAGTCCTCAATGT  1740
BALB/cJ    TTCCTCACAACAGCTTCCATGAGCACTAGGCTTCCTTCCTCTGAGACAAGTCCTCAATGT  1740
           ************************************************************

C57BL/6J   TCACTTGCAAGGAATAAGGGCATTTCCTTCCCTATCTACTAGTTGCTGACCAGAGACAAA  1798
mod        TCACTTGCAAGGAATAAGGGCATTTCCTTCCCTATCTACTAGTCGCTGACCAGAGACAAA  1800
BALB/cJ    TCACTTGCAAGGAATAAGGGCATTTCCTTCCCTATCTACTAGTCGCTGACCAGAGACAAA  1800
           ******************************************* ****************

C57BL/6J   GAAACCCAAACCTTGGGGAAGGAGAACACAATCAAACTCTAAAGCCCAACACCTGGTTCA  1858
mod        GAAACCCAAACCTTGGGGAAGGAGAACACAATCAAACTCTAAAGCCTAACACCTGGTTCA  1860
BALB/cJ    GAAACCCAAACCTTGGGGAAGGAGAACACAATCAAACTCTAAAGCCTAACACCTGGTTCA  1860
           ********************************************** *************

C57BL/6J   AAGAGCTTACCCTATATGCTTTTACGAGGTGGCATCTTCCTGCTTTCACAGACATCACAC  1918
mod        AAGAGCTTACCCTATATGCTTTTACGAGGTGGCATCTTCCTGCTTTCACAGACATCACAC  1920
BALB/cJ    AAGAGCTTACCCTATATGCTTTTACGAGGTGGCATCTTCCTGCTTTCACAGACATCACAC  1920
           ************************************************************

C57BL/6J   TGGATGGATGCTAACAGTGGGGAAGAGAAGAAAGTATACACAGGAAAAGACCGTGTGGAA  1978
mod        TGGATGGATGCTAACAGTGGGGAAGAGAAGAAAGTATGCACAGGAAAAGACCGTGTGGAA  1980
BALB/cJ    TGGATGGATGCTAACAGTGGGGAAGAGAAGAAAGTATGCACAGGAAAAGACCGTGTGGAA  1980
           ************************************ ***********************

C57BL/6J   AAGAGTCCTTAGAAACTCTCCCCACACAGTTGACTGAACCACGATTAGCCTAGCCTGCTC  2038
mod        AAGAGTCCTTAGAAACTCTCCCCACACAGTTGACTGAACCACGATTAGCCTAGCCTGCTC  2040
BALB/cJ    AAGAGTCCTTAGAAACTCTCCCCACACAGTTGACTGAACCACGATTAGCCTAGCCTGCTC  2040
           ************************************************************

C57BL/6J   CAACCAGGCAGTCATAGCAAGGGAGGGACTATCATTTGTAAAGGCTCGAATATAGACCTT  2098
mod        CAACCAGGCAGTCATAGCAAGGGAGGGACTATCATTTGTAAAGGCTCGAATATAGACCTT  2100
BALB/cJ    CAACCAGGCAGTCATAGCAAGGGAGGGACTATCATTTGTAAAGGCTCGAATATAGACCTT  2100
           ************************************************************

C57BL/6J   CAGTGAAATAGGTTTTTTTGGGCCTGTATCTGGAGTATGGGGGTGGGAGGCTGTGTGTAT  2158
mod        CAGTGAAATAGGTTTTTTTGGGCCTGTATCTGGAGTATGGGGGTGGGAGGCTGTGTGTAT  2160
BALB/cJ    CAGTGAAATAGGTTTTTTTGGGCCTGTATCTGGAGTATGGGGGTGGGAGGCTGTGTGTAT  2160
           ************************************************************

C57BL/6J   GTGTGATCATCAGAAAGGAACAACCACACACCCCTAGAGGGTAATCTCATGACCCGGCTG  2218
mod        GTGTGATCATCAGAAAGGAACAACCACACACCCCTAGAGGGTAATCTCATGACCCGGCTG  2220
BALB/cJ    GTGTGATCATCAGAAAGGAACAACCACACACCCCTAGAGGGTAATCTCATGACCCGGCTG  2220
           ************************************************************

C57BL/6J   AGTCTCACCATCAGGCAACATTGTCCAATCTTTCACCACTAGTGCTGACTGCTGACTATA  2278
mod        AGTCTCACCATCAGACAACATTGTCCAATCTTTCACCACTAGTGCTGACTGCTGACTATA  2280
BALB/cJ    AGTCTCACCATCAGACAACATTGTCCAATCTTTCACCACTAGTGCTGACTGCTGACTATA  2280
           ************** *********************************************
```

```
C57BL/6J   TAGAGGCCCTAATAGGCAGCACAATGTTTGTGAGAACATCTTCTCTTTGAAACCCAAGCA  2338
mod        TAGAGGCCCTAATAGGCAGCACAATGTTTGTGAGAACATCTTCTCTTTGAAACCCAAGCA  2340
BALB/cJ    TAGAGGCCCTAATAGGCAGCACAATGTTTGTGAGAACATCTTCTCTTTGAAACCCAAGCA  2340
           ************************************************************

C57BL/6J   GAGGTGCTTGGAGGGCTACCCATGACCTCAGAAGCTGCCACACACAGGCATGAAGCAAAT  2398
mod        GAGGTGCTTGGAGGGCTACCCATGACCTCAGAAGCTGCCACACACAGGCATGAAGCAAAT  2400
BALB/cJ    GAGGTGCTTGGAGGGCTACCCATGACCTCAGAAGCTGCCACACACAGGCATGAAGCAAAT  2400
           ************************************************************

C57BL/6J   CTGGCTTCTCTTGAATAGTCCCAGAGAATAGGACAGATGATTGGGGTCACCTTTCCACCA  2458
mod        CTGGCTTCTCTTGAATAGTCCCAGAGAATAGGACAGATGATTGGGGTCACCTTTCCACCA  2460
BALB/cJ    CTGGCTTCTCTTGAATAGTCCCAGAGAATAGGACAGATGATTGGGGTCACCTTTCCACCA  2460
           ************************************************************

C57BL/6J   GCATCACGACAGAAAAAGACCTGGACCTTCCCAGCCCACCTAGGCCAGAGCTAAGATCAT  2518
mod        GCATCACGACAGAAAAAGACCTGGACCTTCCCAGCCCACCTAGGCCAGAGCTAAGATCAT  2520
BALB/cJ    GCATCACGACAGAAAAAGACCTGGACCTTCCCAGCCCACCTAGGCCAGAGCTAAGATCAT  2520
           ************************************************************

C57BL/6J   ACAAAGCAACAAAAAACACAAAAATGAACAACAACAGCAACAACAACAACAAACCACAAT  2578
mod        ACAAAGCAACAAAAAACACAAAAATGAACAACAACAGCAACAACAACAACAAACCACAAT  2580
BALB/cJ    ACAAAGCAACAAAAAACACAAAAATGAACAACAACAGCAACAACAACAACAAACCACAAT  2580
           ************************************************************

C57BL/6J   ACAGTACATTGGAAAGTCACATGCACACACACATCCACCCATGCGAACCATATACAACTA  2638
mod        ACAGTACATTGGAAAGTCACATGCACACACACATCCACCCATGCGAACCATATACAACTA  2640
BALB/cJ    ACAGTACATTGGAAAGTCACATGCACACACACATCCACCCATGCGAACCATATACAACTA  2640
           ************************************************************

C57BL/6J   CCCCTAGAGAAGCCACTGGGATTAACCTGTTTTCTTCATGTCTGCTCGTTAAACAATTAC  2698
mod        CCCCTAGAGAAGCCACTGGGATTAACCTGTTTTCTTCATGTCTGCTCGTTAAACAATTAC  2700
BALB/cJ    CCCCTAGAGAAGCCACTGGGATTAACCTGTTTTCTTCATGTCTGCTCGTTAAACAATTAC  2700
           ************************************************************

C57BL/6J   CACGCATAGCGTGTTCAAAAGTTAACAGTAACAATTCATCCTACCACAGGGTTGTGATAT  2758
mod        CACGCATAGCGTGTTCAAAAGTTAACAGTAACAATTCATCCTACCACAGGGTTGTGATAT  2760
BALB/cJ    CACGCATAGCGTGTTCAAAAGTTAACAGTAACAATTCATCCTACCACAGGGTTGTGATAT  2760
           ************************************************************

C57BL/6J   GCAAATTTAAAATATTTTAAACAGAAACCTTTACAAATAATCCTATTACATAAGCAATAT  2818
mod        GCAAATTTAAAATATTTTAAACAGAAACCTTTACAAATAATCCTATTACATAAGCAATAT  2820
BALB/cJ    GCAAATTTAAAATATTTTAAACAGAAACCTTTACAAATAATCCTATTACATAAGCAATAT  2820
           ************************************************************

C57BL/6J   TTGCATAGAATTATACAAGTAAATTATACAATATTTAAGCAGCAGCGACATATGCTGGGA  2878
mod        TTGCATAGAATTATACAAGTAAATTATACAATATTTAAGCAGCAGCGACATATGCTGGGA  2880
BALB/cJ    TTGCATAGAATTATACAAGTAAATTATACAATATTTAAGCAGCAGCGACATATGCTGGGA  2880
           ************************************************************

C57BL/6J   GATTTTTTTTAAAACTAAGTTAAGCTTGGTTGTGGCACAGAGACCAATGTACACAGGGGA  2938
mod        GATTTTTTTTAAAACTAAGTTAAGCTTGGTTGTGGCACAGAGACCAATGTACACAGGGGA  2940
BALB/cJ    GATTTTTTTTAAAACTAAGTTAAGCTTGGTTGTGGCACAGAGACCAATGTACACAGGGGA  2940
           ************************************************************

C57BL/6J   AAGGGCCCCTGTGGTCTCCCCACTCTGGGGCAGGAGTGGAGAGGCCCCAGGAAATAGACT  2998
mod        AAGGGCCCCTGTGGTCTCCCCACTCTGGGGCAGGAGTGGAGAGGCCCCAGGAAATAGACT  3000
BALB/cJ    AAGGGCCCCTGTGGTCTCCCCACTCTGGGGCAGGAGTGGAGAGGCCCCAGGAAATAGACT  3000
           ************************************************************

C57BL/6J   GTACAGCATTTCCATGTATATTACATGTGTACACCTTTTAAATATTTATAGAAATATTTC  3058
mod        GTACAGCATTTCCATGTATATTACATGTGTACACCTTTTAAATATTTATAGAAATATTTC  3060
BALB/cJ    GTACAGCATTTCCATGTATATTACATGTGTACACCTTTTAAATATTTATAGAAATATTTC  3060
           ************************************************************
```

```
C57BL/6J   ACAAAATAATACAATCAGTTACAGTAACAATGAACTGCCTTCCCATAATTCCAAGCAGTT  3118
mod        ACAAAATAATACAATCAGTTACAGTAACAATGAACTGCCTTCCCATAATTCCAAGCAGTT  3120
BALB/cJ    ACAAAATAATACAATCAGTTACAGTAACAATGAACTGCCTTCCCATAATTCCAAGCAGTT  3120
           ************************************************************

C57BL/6J   CTCTTCAAGCAGTTCTGGAAAGAGGTGGGGGCGGGGCTTTCCAATCTTTGGCCTCTTTGC  3178
mod        CTCTTCAAGCAGTTCTGGAAAGAGGTGGGGGCGGGGCTTTCCAATCTTTGGCCTCTTTGC  3180
BALB/cJ    CTCTTCAAGCAGTTCTGGAAAGAGGTGGGGGCGGGGCTTTCCAATCTTTGGCCTCTTTGC  3180
           ************************************************************

C57BL/6J   TTCAGATTTTAATTACCCTTTAAAAGAGAACAAAAAaGGAGGTCACATTATGGTCACTAC  3238
mod        TTCAGATTTTAATTACCCTTTAAAAGAGAACAAAAAAGGAGGTCACATTATGGTCACTAC  3240
BALB/cJ    TTCAGATTTTAATTACCCTTTAAAAGAGAACAAAAAAGGAGGTCACATTATGGTCACTAC  3240
           ************************************************************

C57BL/6J   CACACTGAACTGTGAGCCTTACACAGCACAGCTCTTTACTCCCAGTGGGCAGCTGAGAGC  3298
mod        CACACTGAACTGTGAGCCTTACACAGCACAGCTCTTTACTCCCAGTGGGCAGCTGAGAGC  3300
BALB/cJ    CACACTGAACTGTGAGCCTTACACAGCACAGCTCTTTACTCCCAGTGGGCAGCTGAGAGC  3300
           ************************************************************

C57BL/6J   CCTGAAACTCTGACTGGGGTGCTAAACTCATTTCATCATTTTTTTCCCCACCCTGTAAAG  3358
mod        CCTGAAACTCTGACTGGGGTGCTAAACTCATTTCATCATTTTTTTCCCCACCCTGTAAAG  3360
BALB/cJ    CCTGAAACTCTGACTGGGGTGCTAAACTCATTTCATCATTTTTTTCCCCACCCTGTAAAG  3360
           ************************************************************

C57BL/6J   TCATCTGCATGGTTGTTAGAACTCTGCTGTAAGTGCCCAGCAAACAGGTACAAGGTTAAA  3418
mod        TCATCTGCATGGTTGTTAGAACTCTGCTGTAAGTGCCCAGCAAACAGGTACAAGGTTAAA  3420
BALB/cJ    TCATCTGCATGGTTGTTAGAACTCTGCTGTAAGTGCCCAGCAAACAGGTACAAGGTTAAA  3420
           ************************************************************

C57BL/6J   CTATTAAACTGGCATTCTAAGGGATTAATAGCAGGGTGGGGGGGGGGgAGATCCAGTTT  3478
mod        CTATTAAACTGGCATTCTAAGGGATTAATAGCAGGGTGGG---GGGGGGAGATCCAGTTT  3477
BALB/cJ    CTATTAAACTGGCATTCTAAGGGATTAATAGCAGGGTGGG---GGGGGGAGATCCAGTTT  3477
           ****************************************   *****************

C57BL/6J   AGAACGGTATTGTTACTCTAAGGATGAGTTTGGCCTCGTCAGCTTTGCCTTGGGAATAAA  3538
mod        AGAACGGTATTGTTACTCTAAGGATGAGTTTGGCCTCGTCAGCTTTGCCTTGGGAATAAA  3537
BALB/cJ    AGAACGGTATTGTTACTCTAAGGATGAGTTTGGCCTCGTCAGCTTTGCCTTGGGAATAAA  3537
           ************************************************************

C57BL/6J   AGGTCCCCCAGGGTCTAGTTTGGGCCACCGCTCAGATGGCCCTTGGAAGCACCACGGGGA  3598
mod        AGGTCCCCCAGGGTCTAGTTTGGGCCACCGCTCAGATGGCCCTTGGAAGCACCACGGGGA  3597
BALB/cJ    AGGTCCCCCAGGGTCTAGTTTGGGCCACCGCTCAGATGGCCCTTGGAAGCACCACGGGGA  3597
           ************************************************************

C57BL/6J   TACCTCTGGCACACACCTCAGTTGGAGCTGCTCCCCCACGGCCTCTCCTCACATTCCACG  3658
mod        TACCTCTGGCACACACCTCAGTTGGAGCTGCTCCCCCACGGCCTCTCCTCACATTCCACG  3657
BALB/cJ    TACCTCTGGCACACACCTCAGTTGGAGCTGCTCCCCCACGGCCTCTCCTCACATTCCACG  3657
           ************************************************************

C57BL/6J   TCCTGTAGCTCTATGACCTCCACCTTTGAGTCCCTCCTCTCACACCTGACATGAAAAGGC  3718
mod        TCCTGTAGCTCTATGACCTCCACCTTTGAGTCCCTCCTCTCACACCTGACATGAAAAGGC  3717
BALB/cJ    TCCTGTAGCTCTATGACCTCCACCTTTGAGTCCCTCCTCTCACACCTGACATGAAAAGGC  3717
           ************************************************************

C57BL/6J   ACATGAGGATCCTCAAATACCCCGTGATCAGTCTCAGGGTAGCTCTCATAGCCTGGACAG  3778
mod        ACATGAGGATCCTCAAATACCCCGTGATCAGTCTCAGGGTAGCTCTCATAGCCTGGACAG  3777
BALB/cJ    ACATGAGGATCCTCAAATACCCCGTGATCAGTCTCAGGGTAGCTCTCATAGCCTGGACAG  3777
           ************************************************************

C57BL/6J   GGCCCCCCTCGGGGGTTGCGCCCAGGTCCAGGCGGGGGATGCACAGCAACAGTCACCGAA  3838
mod        GGCCCCCCTCGGGGGTTGCGCCCAGGTCCAGGCGGGGGATGCACAGCAACAGTCACCGAA  3837
BALB/cJ    GGCCCCCCTCGGGGGTTGCGCCCAGGTCCAGGCGGGGGATGCACAGCAACAGTCACCGAA  3837
           ************************************************************
```

```
C57BL/6J   GCAGAAGCCGTCACAGTGGTGATGGGCTGGCAGTAGCTGGGCACAGAGCTGCCCATGGCG  3898
mod        GCAGAAGCCGTCACAGTGGTGATGGGCTGGCAGTAGCTGGGCACAGAGCTGCCCATGGCG  3897
BALB/cJ    GCAGAAGCCGTCACAGTGGTGATGGGCTGGCAGTAGCTGGGCACAGAGCTGCCCATGGCG  3897
           ************************************************************

C57BL/6J   GTGGACGTTGGGTTCCGAGGGTTGTGAGAACGGGCCCCACGGGGCCCTGAGCGGTCCCTA  3958
mod        GTGGACGTTGGGTTCCGAGGGTTGTGAGAACGGGCCCCACGGGGCCCTGAGCGGTCCCTA  3957
BALB/cJ    GTGGACGTTGGGTTCCGAGGGTTGTGAGAACGGGCCCCACGGGGCCCTGAGCGGTCCCTA  3957
           ************************************************************

C57BL/6J   TTGCTAGGGCCAGAATGCCCTTCAGTAGAAATTTCAAAAGCGTCTCTGCGCGGTCTGTAG  4018
mod        TTGCTAGGGCCAGAATGCCCTTCAGTAGAAATTTCAAAAGCGTCTCTGCGCGGTCTGTAG  4017
BALB/cJ    TTGCTAGGGCCAGAATGCCCTTCAGTAGAAATTTCAAAAGCGTCTCTGCGCGGTCTGTAG  4017
           ************************************************************

C57BL/6J   GGGGGTGGCCGCAAGCCTTCTCTAGGGGGATCCCTTCGAGGCTGCTGGCCTTGCCGTCCA  4078
mod        GGGGGTGGCCGCAAGCCTTCTCTAGGGGGATCCCTTCGAGGCTGCTGGCCTTGCCGTCCA  4077
BALB/cJ    GGGGGTGGCCGCAAGCCTTCTCTAGGGGGATCCCTTCGAGGCTGCTGGCCTTGCCGTCCA  4077
           ************************************************************

C57BL/6J   GGGGACAAGGAGCCAGAGTCCAGGTGGGGCTGTTGCCGAGGGGTCAAGGGAGGCTGATGT  4138
mod        GGGGACAAGGAGCCAGAGTCCAGGTGGGGCTGTTGCCGAGGGTTCAAGGGAGGCTGATGT  4137
BALB/cJ    GGGGACAAGGAGCCAGAGTCCAGGTGGGGCTGTTGCCGAGGGTTCAAGGGAGGCTGATGT  4137
           ****************************************** ******************

C57BL/6J   CTGGAGTCCGGATGGACCACCTGCAGAGGAGAGACATAGGTCAACACAGGGAGGTAGGAT  4198
mod        CTGGAGTCCGGATGGACCACCTGCAGAGGAGAGACATAGGTCAACACAGGGAGGTAGGAT  4197
BALB/cJ    CTGGAGTCCGGATGGACCACCTGCAGAGGAGAGACATAGGTCAACACAGGGAGGTAGGAT  4197
           ************************************************************

C57BL/6J   GGTGGTGATGTTCCACCCACAAAAGAAAACCTATTCCTTTAGAAACCTCCAGGATGTGAA  4258
mod        GGTGGTGATGTTCCACCCACAAAAGAAAACCTATTCCTTTAGAAACCTCCAGGATGTGAA  4257
BALB/cJ    GGTGGTGATGTTCCACCCACAAAAGAAAACCTATTCCTTTAGAAACCTCCAGGATGTGAA  4257
           ************************************************************

C57BL/6J   TCCTGCCTGCACCTGCACAGCTGGCTGGAGGCATATAGCCACTGCCCATAGATCTCAACT  4318
mod        TCCTGCCTGCACCTGCACAGCTGGCTGGAGGCATATAGCCACTGCCCATAGATCTCAACT  4317
BALB/cJ    TCCTGCCTGCACCTGCACAGCTGGCTGGAGGCATATAGCCACTGCCCATAGATCTCAACT  4317
           ************************************************************

C57BL/6J   TACCCTCACAACCAACTGCCCCCAGGCCTAAGTTCTCTGCCTCAAAACTGCCAAGGCCTG  4378
mod        TACCCTCACAACCAACTGCCCCCAGGCCTAAGTTCTCTGCCTCAAAACTGCCAAGGCCTG  4377
BALB/cJ    TACCCTCACAACCAACTGCCCCCAGGCCTAAGTTCTCTGCCTCAAAACTGCCAAGGCCTG  4377
           ************************************************************

C57BL/6J   GATAGCCAAGAGCCTGGGTGTCTTGGAAATATGCAACCATAAATAGTAGCTTTTAGAAGT  4438
mod        GATAGCCAAGAGCCTGGGTGTCTTGGAAATATGCAACCATAAATAGTAGCTTTTAGAAGT  4437
BALB/cJ    GATAGCCAAGAGCCTGGGTGTCTTGGAAATATGCAACCATAAATAGTAGCTTTTAGAAGT  4437
           ************************************************************

C57BL/6J   ATAAGGCTCCTGTTTCTGGGTCATATTAGTGTTGTTTTCACCTGTCCCCAGCCCTAAGCC  4498
mod        ATAAGGCTCCTGTTTCTGGGTCATATTAGTGTTGTTTTCACCTGTCCCCAGCCCTAAGCC  4497
BALB/cJ    ATAAGGCTCCTGTTTCTGGGTCATATTAGTGTTGTTTTCACCTGTCCCCAGCCCTAAGCC  4497
           ************************************************************

C57BL/6J   AGGTGTGGCCAGAAGCAAATGTACTGTAAGAGCAGAGCAAAAACTTCCACACAGATAGTT  4558
mod        AGGTGTGGCCAGAAGCAAATGTACTGTAAGAGCAGAGCAAAAACTTCCACACAGATAGTT  4557
BALB/cJ    AGGTGTGGCCAGAAGCAAATGTACTGTAAGAGCAGAGCAAAAACTTCCACACAGATAGTT  4557
           ************************************************************

C57BL/6J   CTGTTAGGCAATACATCTCTGCCTGACTATTAGGAATCTGGTTTCTGGGTCCTCTGTACA  4618
mod        CTGTTAGGCAATACATCTCTGCCTGACTATTAGGAATCTGGTTTCTGGGTCCTCTGTACA  4617
BALB/cJ    CTGTTAGGCAATACATCTCTGCCTGACTATTAGGAATCTGGTTTCTGGGTCCTCTGTACA  4617
           ************************************************************
```

```
C57BL/6J   AAGCTCGGAGCAACACAGTGGCCACATCAATCAAAAGGACCGTGACCAACTTCAAAGTCG  4678
mod        AAGCTCGGAGCAACACAGTGGCCACATCAATCAAAAGGACCGTGACCAACTTCAAAGTCG  4677
BALB/cJ    AAGCTCGGAGCAACACAGTGGCCACATCAATCAAAAGGACCGTGACCAACTTCAAAGTCG  4677
           ************************************************************

C57BL/6J   GTGAGCTTGTACCTATTTTTAGGCTCCTGCTGAACAGAACCAGATTCACACTACAGCTCA  4738
mod        GTGAGCTTGTACCTATTTTTAGGCTCCTGCTGAACAGAACCAGATTCACACTACAGCTCA  4737
BALB/cJ    GTGAGCTTGTACCTATTTTTAGGCTCCTGCTGAACAGAACCAGATTCACACTACAGCTCA  4737
           ************************************************************

C57BL/6J   GCAGGGCATCGTCACGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGGGGG  4798
mod        GCAGGGCATCGTCACGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT--T----  4791
BALB/cJ    GCAGGGCATCGTCACGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT--T----  4791
           ********************************************************  *

C57BL/6J   GGGGGGGGTGGACAGAGGACGGGGACACAATTCACTGGCCAGCCCTTCTCTCCTTCAAGG  4858
mod        -GGGGGGGTGGACAGAGGACGGGGACACAATTCACTAGCCAGCCCTTCTCTCCTTCAAGG  4850
BALB/cJ    -GGGGGGGTGGACAGAGGACGGGGACACAATTCACTAGCCAGCCCTTCTCTCCTTCAAGG  4850
            ********************************** ***********************

C57BL/6J   AAGGCTGCTCTAGCCTGGGACTGGAATACACATTTCCTGTAAACATGGTGGGGGCCTCAG  4918
mod        AAGGCTGCTCTAGCCTGGGACTGGAATACACATTTCCTGTAAACATGGTGGGGGCCTCAG  4910
BALB/cJ    AAGGCTGCTCTAGCCTGGGACTGGAATACACATTTCCTGTAAACATGGTGGGGGCCTCAG  4910
           ************************************************************

C57BL/6J   GCAAGCCAGAGTTTTGGAGCCTTCCTTAACTCTTCAAGGTGGAGCATCTTGACTTGGAGG  4978
mod        GCAAGCCAGAGTTTTGGAGCCTTCCTTAACTCTTCAAGGTGGAGCATCTTGACTTGGAGG  4970
BALB/cJ    GCAAGCCAGAGTTTTGGAGCCTTCCTTAACTCTTCAAGGTGGAGCATCTTGACTTGGAGG  4970
           ************************************************************

C57BL/6J   GTGGGGGTGCGGGTAAGGAAGGAACCTGTGGACTCCTCACTACCTTCTGGAATTAAGTAC  5038
mod        GTGGGGGTGCGGGTAAGGAAGGAACCTGTGGACTCCTCACTACCTTCTGGAATTAAGTAC  5030
BALB/cJ    GTGGGGGTGCGGGTAAGGAAGGAACCTGTGGACTCCTCACTACCTTCTGGAATTAAGTAC  5030
           ************************************************************

C57BL/6J   TCAGGATGCGTCTGCATACTTTCAATAGATCTCCAGCCATCTACAACAATAAGACCCTGA  5098
mod        TCAGGATGCGTCTGCATACTTTCAATAGATCTCCAGCCATCTACAACAATAAGACCCTGA  5090
BALB/cJ    TCAGGATGCGTCTGCATACTTTCAATAGATCTCCAGCCATCTACAACAATAAGACCCTGA  5090
           ************************************************************

C57BL/6J   AGCCACAAAGCTACGGGAAAGCCTTAGATAGTAATTAAATAACTTTGTGCCAAGGCCACT  5158
mod        AGCCACAAAGCTACGGGAAAGCCTTAGATAGTAATTAAATAACTTTGTGCCAAGGCCACT  5150
BALB/cJ    AGCCACAAAGCTACGGGAAAGCCTTAGATAGTAATTAAATAACTTTGTGCCAAGGCCACT  5150
           ************************************************************

C57BL/6J   TAAAGGTACCAGAGACGAGAGGAAAAGAATAAATGCTTCTTGTGGTGGAATACTACGCTG  5218
mod        TAAAGGTACCAGAGACGAGAGGAAAAGAATAAATGCTTCTTGTGGTGGAATACTACGCTG  5210
BALB/cJ    TAAAGGTACCAGAGACGAGAGGAAAAGAATAAATGCTTCTTGTGGTGGAATACTACGCTG  5210
           ************************************************************

C57BL/6J   TTTGTTACATCATCTCCTCACATGTCCTTCTGAAAGGACAGATAGGGCAAAAGGAGATAA  5278
mod        TTTGTTACATCATCTCCTCACATGTCCTTCTGAAAGGACAGATAGGGCAAAAGGAGATAA  5270
BALB/cJ    TTTGTTACATCATCTCCTCACATGTCCTTCTGAAAGGACAGATAGGGCAAAAGGAGATAA  5270
           ************************************************************

C57BL/6J   AGGCCCTAAAAAGCTGAGTCCATTCTGAATAAGGGTAGGGCTGGGGATGAGCCCCTGGGC  5338
mod        AGGCCCTAAAAAGCTGAGTCCATTCTGAATAAGGGTAGAGCTGGGGATGAGCCCCTGGGC  5330
BALB/cJ    AGGCCCTAAAAAGCTGAGTCCATTCTGAATAAGGGTAGAGCTGGGGATGAGCCCCTGGGC  5330
           ************************************* **********************

C57BL/6J   TTTGGGAATTCCAGGCAAGCGCTTCACCACTAAACTATGCCTCCTTGTAACAATGCTGGC  5398
mod        TTTGGGAATTCCAGGCAAGCGCTTCACCACTAAACTATGCCTCCTTGTAACAATGCTGGC  5390
BALB/cJ    TTTGGGAATTCCAGGCAAGCGCTTCACCACTAAACTATGCCTCCTTGTAACAATGCTGGC  5390
           ************************************************************
```

```
C57BL/6J   ACTGCCCCGGTTCAGTCCACAAGATTAAAAGGAAACACATCCTACAGTCTACAAAGAATG  5458
mod        ACTGCCCCGGTTCAGTCCACAAGATTAAAAGGAAACACATCCTACAGTCTACAAAGAATG  5450
BALB/cJ    ACTGCCCCGGTTCAGTCCACAAGATTAAAAGGAAACACATCCTACAGTCTACAAAGAATG  5450
           ************************************************************

C57BL/6J   GCGCGCTGGTGGTTTCCTGTGATTTGACCCCGGCGTGTACACTTTTTTCTTGCACCCCAT  5518
mod        GCGCGCTGGTGGTTTCCTGTGATTTGACCCCGGCGTGTACACTTTTTTCTTGCACCCCAT  5510
BALB/cJ    GCGCGCTGGTGGTTTCCTGTGATTTGACCCCGGCGTGTACACTTTTTTCTTGCACCCCAT  5510
           ************************************************************

C57BL/6J   CTGTCTACTAGATCTGCTGCTCAGAGTACACAAGAGCCTGCTCCGAACCTCCTGGCCCAC  5578
mod        CTGTCTACTAGATCTGCTGCTCAGAGTACACAAGAGCCTGCTCCGAACCTCCTGGCCCAC  5570
BALB/cJ    CTGTCTACTAGATCTGCTGCTCAGAGTACACAAGAGCCTGCTCCGAACCTCCTGGCCCAC  5570
           ************************************************************

C57BL/6J   TGCCACAGGACAAATGAATCTTCCCCACCGCCC-------CCCCCCCCTTGCCAACCCCC  5631
mod        TGCCACAGGACAAATGAATCTTCCCCACCGCCCCC-----CCCCCCCCTTGCCAACCCCC  5625
BALB/cJ    TGCCACAGGACAAATGAATCTTCCCCACCGCCCCCCCCCCCCCCCCTTGCCAACCCCC  5630
           **********************************       ********************

C57BL/6J   CAGGTGACTCACAGTGGACCGGGCAAAGACAGGGTTTTCTGTGGCTTCCACAATCACTTG  5691
mod        CAGGTGACTCACAGTGGACCGGGCAAAGACAGGGTTTTCTGTGGCTTCCACAATCACTTG  5685
BALB/cJ    CAGGTGACTCACAGTGGACCGGGCAAAGACAGGGTTTTCTGTGGCTTCCACAATCACTTG  5690
           ************************************************************

C57BL/6J   GTGGGCAGGGCCTCCGGCACCCTGCTGTGCTTCGTATTGCCTGAGCTCCTCACTGATGCC  5751
mod        GTGGGCAGGGCCTCCGGCACCCTGCTGTGCTTCGTATTGCCTGAGCTCCTCACTGATGCC  5745
BALB/cJ    GTGGGCAGGGCCTCCGGCACCCTGCTGTGCTTCGTATTGCCTGAGCTCCTCACTGATGCC  5750
           ************************************************************

C57BL/6J   AGACACCGTGGTCTGAGAGCTGTACTCCGAGTCGGAGGAATCAGACCCATTGTTCGTGTG  5811
mod        AGACACCGTGGTCTGAGAGCTGTACTCCGAGTCGGAGGAATCAGACCCATTGTTCGTGTG  5805
BALB/cJ    AGACACCGTGGTCTGAGAGCTGTACTCCGAGTCGGAGGAATCAGACCCATTGTTCGTGTG  5810
           ************************************************************

C57BL/6J   ACCAGGAGGCACGGCAAACCGGACGACACTTGGAGGCGGCTCAGGCGAAGGAGTGGGCAG  5871
mod        ACCAGGAGGCACGGCAAACCGGACGACACTTGGGGGCGGCTCAGGCGAAGGAGTGGGCAG  5865
BALB/cJ    ACCAGGAGGCACGGCAAACCGGACGACACTTGGGGGCGGCTCAGGCGAAGGAGTGGGCAG  5870
           ********************************* ************************

C57BL/6J   TCGGTTTAGGCCATTGGCTGGAGACACCTGTTGAAGGAGGTTTCCAATTAAAAATGTTAC  5931
mod        TCGGTTTAGGCCATTGGCTGGAGACACCTGTTGAAGGAGGTTTCCAATTAAAAATGTTAC  5925
BALB/cJ    TCGGTTTAGGCCATTGGCTGGAGACACCTGTTGAAGGAGGTTTCCAATTAAAAATGTTAC  5930
           ************************************************************

C57BL/6J   GTTTGTTGACCCGCAGGTCACTTTCAAAGTTCAAAGCCCAGGATTATTGGCCTCTGTGAC  5991
mod        GTTTGTTGACCCGCAGGTCACTTTCAAAGTTCAAAGCCCAGGATTATTGGCCTCTGTGAC  5985
BALB/cJ    GTTTGTTGACCCGCAGGTCACTTTCAAAGTTCAAAGCCCAGGATTATTGGCCTCTGTGAC  5990
           ************************************************************

C57BL/6J   CTCGTAGGGGAACCCCATACATTAAACTGTATTTACTTCCCGTCTCCCTGCTATCTGGAA  6051
mod        CTCGTAGGGGAACCCCATACATTAAACTGTATTTACTTCCCGTCTCCCTGCTATCTGGAA  6045
BALB/cJ    CTCGTAGGGGAACCCCATACATTAAACTGTATTTACTTCCCGTCTCCCTGCTATCTGGAA  6050
           ************************************************************

C57BL/6J   GACACTCACGATTCTGGCACAAAAACATTGTCCCAACTTCTAATACCAGAGAAGCCAGAC  6111
mod        GACACTCACGATTCTGGCACAAAAACATTGTCCCAACTTCTAATACCAGAGAAGCCAGAC  6105
BALB/cJ    GACACTCACGATTCTGGCACAAAAACATTGTCCCAACTTCTAATACCAGAGAAGCCAGAC  6110
           ************************************************************

C57BL/6J   ATCGGCTACCTATCTTGTCTGTCCTGTTAAGCCCAAACAGTTATCTGAGGTCGAAAACCT  6171
mod        ATCGGCTACCTATCTTGTCTGTCCTGTTAAGCCCAAACAGTTATCTGAGGTCGAAAACCT  6165
BALB/cJ    ATCGGCTACCTATCTTGTCTGTCCTGTTAAGCCCAAACAGTTATCTGAGGTCGAAAACCT  6170
           ************************************************************
```

```
C57BL/6J   TCCTAAGCTGCCCATCCATTGGTGTAAAACAGTCAAAGGGAAAAATGAAACAGCCATTGT  6231
mod        TCCTAAGCTGCCCATCCATTGGTGTAAAACAGTCAAAGGGAAAAATGAAACAGCCATTGT  6225
BALB/cJ    TCCTAAGCTGCCCATCCATTGGTGTAAAACAGTCAAAGGGAAAAATGAAACAGCCATTGT  6230
           ************************************************************

C57BL/6J   TAATGGATGTTGACCGAAAAGCCCAGACCATCTGCTAATATGGCAACACCTGAATATCTC  6291
mod        TAATGGATGTTGACCGAAAAGCCCAGACCATCTGCTAATATGGCAACACCTGAATATCTC  6285
BALB/cJ    TAATGGATGTTGACCGAAAAGCCCAGACCATCTGCTAATATGGCAACACCTGAATATCTC  6290
           ************************************************************

C57BL/6J   AATCCGGGGTGGCATTGGGGGGCAATCACTGACCTCAGGACACGGTCCAAAGAAGGATAA  6351
mod        AATCCGGGGTGGCATTGGGGGGCAATCACTGACCTCAGGACACGGTCCAAAGAAGGATAA  6345
BALB/cJ    AATCCGGGGTGGCATTGGGGGGCAATCACTGACCTCAGGACACGGTCCAAAGAAGGATAA  6350
           ************************************************************

C57BL/6J   GAGGACAGGCAGCAGAACCAGTCCATTGAGAACCCCCAAGACGGTGAGAATGGCCAGGAC  6411
mod        GAGGACAGGCAGCAGAACCAGTCCATTGAGAACCCCCAAGACGGTGAGAATGGCCAGGAC  6405
BALB/cJ    GAGGACAGGCAGCAGAACCAGTCCATTGAGAACCCCCAAGACGGTGAGAATGGCCAGGAC  6410
           ************************************************************

C57BL/6J   GGCAAAGAAGTATCTGGAAGACAAAAGGGATATGGGACATGTCAGCGAGGAAGGAGAGCA  6471
mod        GGCAAAGAAGTATCTGGAAGACAAAAGGGATATGGTACATGTCAGCGAGGAAGGAGAGCA  6465
BALB/cJ    GGCAAAGAAGTATCTGGAAGACAAAAGGGATATGGTACATGTCAGCGAGGAAGGAGAGCA  6470
           ********************************** ***********************

C57BL/6J   TTCCACGCCCCCGGGCGCCCAAAGGTTACTCTGCACTAAGAGCTCAGCTCAACAAACCTG  6531
mod        TTCCACGCCCCCGGGCGCCCAAAGGTTACTCTGCACTAAGAGCTCAGCTCAACAAACCTG  6525
BALB/cJ    TTCCACGCCCCCGGGCGCCCAAAGGTTACTCTGCACTAAGAGCTCAGCTCAACAAACCTG  6530
           ************************************************************

C57BL/6J   CTGGGCTGAGAAGTCAACCACAAAACACTAGTTCAAACTTAGCATAGATTAGTTTGGGAT  6591
mod        CTGGGCTGAGAAGTCAACCACAAAACACTAGTTCAAACTTAGCATAGATTAGTTTGGGAT  6585
BALB/cJ    CTGGGCTGAGAAGTCAACCACAAAACACTAGTTCAAACTTAGCATAGATTAGTTTGGGAT  6590
           ************************************************************

C57BL/6J   TTTGCTCTTTTTTTTTTTTTTATTTATCCTTGCGGCTCTGAGGACTAGATTCCTCCACACC  6651
mod        TTTGCTCTTTTTTT-TTTTTTTATTTATCCTTGCGGCTCTGAGGACTAGATTCCTCCACACC  6644
BALB/cJ    TTTGCTCTTTTTTT-TTTTTTTATTTATCCTTGCGGCTCTGAGGACTAGATTCCTCCACACC  6649
           ************* ********************************************

C57BL/6J   CTGCAGCGTTACTTCATGAAGGGAAGGGGAAAAAATTGTTAATTTGATGAAATCTCCTGC  6711
mod        CTGCAGCGTTACTTCATGAAGGGAAGGGGAAAAAATTGTTAATTTGATGAAATCTCCTGC  6704
BALB/cJ    CTGCAGCGTTACTTCATGAAGGGAAGGGGAAAAAATTGTTAATTTGATGAAATCTCCTGC  6709
           ************************************************************

C57BL/6J   CTGTGATGTTATTTATGAACGTGTTACACCAACACAGGAATATGGAATGTGAGCTGAATG  6771
mod        CTGTGATGTTATTTATGAACGTGTTACACCAACACAGGAATATGGAATGTGAGCTGAATG  6764
BALB/cJ    CTGTGATGTTATTTATGAACGTGTTACACCAACACAGGAATATGGAATGTGAGCTGAATG  6769
           ************************************************************

C57BL/6J   GCTGAGCACCCTGAAAGTACAGATACAATGCTGGGCGAGGGGCCCCTCTCGAGAACCCAG  6831
mod        GCTGAGCACCCTGAAAGTACAGATACAATGCTGGGCGAGGGGCCCCTCTCGAGAACCCAG  6824
BALB/cJ    GCTGAGCACCCTGAAAGTACAGATACAATGCTGGGCGAGGGGCCCCTCTCGAGAACCCAG  6829
           ************************************************************

C57BL/6J   CAGAAGCAAACAAATGAAGTACAGGGCGTAGAATTTAAACAATTTGGCGCAGGAGGAAAA  6891
mod        CAGAAGCAAACAAATGAAGTACAGGGCGTAGAATTTAAACAATTTGGCGCAGGAGGAAAA  6884
BALB/cJ    CAGAAGCAAACAAATGAAGTACAGGGCGTAGAATTTAAACAATTTGGCGCAGGAGGAAAA  6889
           ************************************************************

C57BL/6J   GCAGCCCACTTCTGCAGGCAGCTAATCCACTGTGAAATGCCCCTGCCTCTTACAGCAGAA  6951
mod        GCAGCCCACTTCTGCAGGCAGCTAATCCACTGTGAAATGCCCCTGCCTCTTATAGCAGAA  6944
BALB/cJ    GCAGCCCACTTCTGCAGGCAGCTAATCCACTGTGAAATGCCCCTGCCTCTTATAGCAGAA  6949
           ************************************************** ********
```

```
C57BL/6J   TCTCTAATTAAGTCTCTAGTGCATCAGGTAAGGCTTTCCCTGGGCTTTCATTAAGCACAA  7011
mod        TCTCTAATTAAGTCTCTAGTGCATCAGGTAAGGCTTTTCCTGGGCTTTCATTAAGCACAA  7004
BALB/cJ    TCTCTAATTAAGTCTCTAGTGCATCAGGTAAGGCTTTTCCTGGGCTTTCATTAAGCACAA  7009
           ********************************** **********************

C57BL/6J   CTTTCTGCCTAGAATTTCTCTGGGTTCCAGTTTTACCAAACAGCTCCAATTGCAACCTTT  7071
mod        CTTTCTGCCTAGAATTTCTCTGGGTTCCAGTTTTACCAAACAGCTCCAATTGCAACCTTT  7064
BALB/cJ    CTTTCTGCCTAGAATTTCTCTGGGTTCCAGTTTTACCAAACAGCTCCAATTGCAACCTTT  7069
           ************************************************************

C57BL/6J   TTAGAACACAGGAACCAGCTTTCTGCAAAGAAAACGCTGCCCATTCTGTGAGAGCCGGGG  7131
mod        TTAGAACACAGGAACCAGCTTTCTGCAAAGAAAACGCTGCCCATTCTGTGAGAGCCGGGG  7124
BALB/cJ    TTAGAACACAGGAACCAGCTTTCTGCAAAGAAAACGCTGCCCATTCTGTGAGAGCCGGGG  7129
           ************************************************************

C57BL/6J   CATTGTGCAGGGCGAGGGCACTCCACAATGGCAGGGCTGCTCTGCAGGCCAATCTTTTGT  7191
mod        CATTGTGCAGGGCGAGGGCACTCCACAATGGCAGGGCTGCTCTGCAGGCCAATCTTTTGT  7184
BALB/cJ    CATTGTGCAGGGCGAGGGCACTCCACAATGGCAGGGCTGCTCTGCAGGCCAATCTTTTGT  7189
           ************************************************************

C57BL/6J   TGAACAATTTAGCATCTTTGAATACTCTTGTTTTGCGTCAGCCAATTCAAACATAAAGGC  7251
mod        TGAACAATTTAGCATCTTTGAATACTCTTGTTTTGCGTCAGCCAATTCAAACATAAAGGC  7244
BALB/cJ    TGAACAATTTAGCATCTTTGAATACTCTTGTTTTGCGTCAGCCAATTCAAACATAAAGGC  7249
           ************************************************************

C57BL/6J   TGAAGCACCATGGTAGAAGAGGGAGAGTGAACAGTGAGTTCTGCGTTCCCACCATTGAGG  7311
mod        TGAAGCACCATGGTAGAAGAGGGAGAGTGAACAGTGAGTTCTGCGTTCCCACCATTGAGG  7304
BALB/cJ    TGAAGCACCATGGTAGAAGAGGGAGAGTGAACAGTGAGTTCTGCGTTCCCACCATTGAGG  7309
           ************************************************************

C57BL/6J   GCTAAAGGTCTGGGATCCTTCAAGAATACCAAGGTAAACCTGTCAGCCAGGGCCCGCCTC  7371
mod        GCTAAAGGTCTGGGATCCTTCAAGAATACCAAGGTAAACCTGTCAGCCAGGGCCCGCCTC  7364
BALB/cJ    GCTAAAGGTCTGGGATCCTTCAAGAATACCAAGGTAAACCTGTCAGCCAGGGCCCGCCTC  7369
           ************************************************************

C57BL/6J   TGGTCAGCACTCTGCACAGCTGGAGGGCGCACACTCCACGCAAGGATGCTTTTCTGGAAC  7431
mod        TGGTCAGCACTCTGCACAGCTGGAGGGCGCACACTCCACGCAAGGATGCTTTTCTGGAAC  7424
BALB/cJ    TGGTCAGCACTCTGCACAGCTGGAGGGCGCACACTCCACGCAAGGATGCTTTTCTGGAAC  7429
           ************************************************************

C57BL/6J   TAATGATTTCTAGGGCTTTGATGGAAAGGTGGGGGCTGGGGGCTGGGGTTAGGATGTCCT  7491
mod        TAATGATTTCTAGGGCTTTGATGGAAAGGTGGGGGCTGGGGGCTGGGGTTAGGATGTCCT  7484
BALB/cJ    TAATGATTTCTAGGGCTTTGATGGAAAGGTGGGGGCTGGGGGCTGGGGTTAGGATGTCCT  7489
           ************************************************************

C57BL/6J   GGAAAGAGAGAGGGACGGCTAGATATCTATCTGTTAATGGTCCCCTGGATCATAAAACTG  7551
mod        GGAAAGAGAGAGGGACGGCTAGATATCTATCTGTTAATGGTCCCCTGGATCATAAAACTG  7544
BALB/cJ    GGAAAGAGAGAGGGACGGCTAGATATCTATCTGTTAATGGTCCCCTGGATCATAAAACTG  7549
           ************************************************************

C57BL/6J   AACAGCTCCCCTACAATGCCATGTCTACGCATGTTTTGACATTAGATGAAAGGAGAGAGA  7611
mod        AACAGCTCCCCTACAATGCCATGTCTACGCATGTTTTGACATTAGATGAAAGGAGAGAGA  7604
BALB/cJ    AACAGCTCCCCTACAATGCCATGTCTACGCATGTTTTGACATTAGATGAAAGGAGAGAGA  7609
           ************************************************************

C57BL/6J   ATCCCTATAAAAATACAAATACCTTACAGCTATTGTAAACACTAAAACCTGCTGAGTTGC  7671
mod        ATCCCTATAAAAATACAAATACCTTACAGCTATTGTAAACACTAAAACCTGCTGAGTTGC  7664
BALB/cJ    ATCCCTATAAAAATACAAATACCTTACAGCTATTGTAAACACTAAAACCTGCTGAGTTGC  7669
           ************************************************************

C57BL/6J   ATTAGTAAAAACCCAAGCTGCCAATGTCTTGCTCTAACATAAACTGTGGAACTTTTGGTG  7731
mod        ATTAGTAAAAACCCAAGCTGCCAATGTCTTGCTCTAACATAAACTGTGGAACTTTTGGTG  7724
BALB/cJ    ATTAGTAAAAACCCAAGCTGCCAATGTCTTGCTCTAACATAAACTGTGGAACTTTTGGTG  7729
           ************************************************************
```

```
C57BL/6J   GAGGCATATATCAACGGCAGGCATTAAGACATTTGCATCCCATCCATCTAGCTGCAACAT  7791
mod        GAGGCATATATCAACGGCAGGCATTAAGACATTTGCATCCCATCCATCTAGCTGCAACAT  7784
BALB/cJ    GAGGCATATATCAACGGCAGGCATTAAGACATTTGCATCCCATCCATCTAGCTGCAACAT  7789
           ************************************************************

C57BL/6J   AAACAGACCAATTAGATTTCTGACGGCTTCATGTATACAGAGTCTGCTCATTTATTCAAC  7851
mod        AAACAGACCAATTAGATTTCTGACGGCTTCATGTATACAGAGTCTGCTCATTTATTCAAC  7844
BALB/cJ    AAACAGACCAATTAGATTTCTGACGGCTTCATGTATACAGAGTCTGCTCATTTATTCAAC  7849
           ************************************************************

C57BL/6J   TCCAGTGAGGACTTTATCCATCCTATTTTAGAAAATGGGAAGCCCTTCTAAGTGCCTGAA  7911
mod        TCCAGTGAGGACTTTATCCATCCTATTTTAGAAAATGGGAAGCCCTTCTAAGTGCCTGAA  7904
BALB/cJ    TCCAGTGAGGACTTTATCCATCCTATTTTAGAAAATGGGAAGCCCTTCTAAGTGCCTGAA  7909
           ************************************************************

C57BL/6J   ACAGAAACCATGATCTCTCTGGGAGGGAATGAGCAGGCCCAGCGAGGACAGGATGAGCGG  7971
mod        ACAGAAACCATGATCTCTCTGGGAGGGAATGAGCAGGCCCAGCGAGGACAGGATGAGCGG  7964
BALB/cJ    ACAGAAACCATGATCTCTCTGGGAGGGAATGAGCAGGCCCAGCGAGGACAGGATGAGCGG  7969
           ************************************************************

C57BL/6J   TTAGGTTGTGTTCCTTAGTATCCTATGTAAACCCGTGCTCCAGCTCCTCAGCTGACTTGA  8031
mod        TTAGGTTGTGTTCCTTAGTATCCTATGTAAACCCGTGCTCCAGCTCCTCAGCTGACTTGA  8024
BALB/cJ    TTAGGTTGTGTTCCTTAGTATCCTATGTAAACCCGTGCTCCAGCTCCTCAGCTGACTTGA  8029
           ************************************************************

C57BL/6J   CCTGCAAGGTCTCCGCACACCCAACTTTCTCTCGGCTTTGAACTGCTTAGAGAGCTATAT  8091
mod        CCTGCAAGGTCTCCGCACACCCAACTTTCTCTTGGCTTTGAACTGCTTAGAGAGCTATAT  8084
BALB/cJ    CCTGCAAGGTCTCCGCACACCCAACTTTCTCTTGGCTTTGAACTGCTTAGAGAGCTATAT  8089
           ******************************* ****************************

C57BL/6J   CTCCCAATCCCTGCAGACTATAACCACAAAGCAGAGATGGGTTTCCTGCTACCTGCCAGT  8151
mod        CTCCCAATCCCTGCAGACTATAACCACAAAGCAGAGATGGGTTTCCTGCTACCTGCCAGT  8144
BALB/cJ    CTCCCAATCCCTGCAGACTATAACCACAAAGCAGAGATGGGTTTCCTGCTACCTGCCAGT  8149
           ************************************************************

C57BL/6J   AATCACAAAAAAAGGCACCTGTGAGGCCAGCATTACCTTTGTGTTGACAGAGCCAAGAGA  8211
mod        AATCACAAAAAAaGGCACCTGTGAGGCCAGCATTACCTTTGTGTTGACAGAGCCAAGAGA  8204
BALB/cJ    AATCACAAAAAAAGGCACCTGTGAGGCCAGCATTACCTTTGTGTTGACAGAGCCAAGAGA  8209
           ************************************************************

C57BL/6J   CGACACCTATTCAGTCTCATTTGGCATTGGGCATTGATATCATTAAGACAAAGCTAGTAT  8271
mod        CGACACCTATTCAGTCTCATTTGGCATTGGGCATTGATATCATTAAGACAAAGCTAGTAT  8264
BALB/cJ    CGACACCTATTCAGTCTCATTTGGCATTGGGCATTGATATCATTAAGACAAAGCTAGTAT  8269
           ************************************************************

C57BL/6J   CTGTGTAAGAGTGTGTGGGAGAAGCAAACACCAAATCAACACTTTAAAACAAAATAATCC  8331
mod        CTGTGTAAGAGTGTGTGGGAGAAGCAAACACCAAATCAACACTTTAAAACAAAATAATCC  8324
BALB/cJ    CTGTGTAAGAGTGTGTGGGAGAAGCAAACACCAAATCAACACTTTAAAACAAAATAATCC  8329
           ************************************************************

C57BL/6J   CAAACCAGGATCTAAGAGCAACTATCTCAATGGTATTTGGCAAATGAGAATCTAATGAGT  8391
mod        CAAACCAGGATCTAAGAGCAACTATCTCAATGGTATTTGGCAAATGAGAATCTAATGAGT  8384
BALB/cJ    CAAACCAGGATCTAAGAGCAACTATCTCAATGGTATTTGGCAAATGAGAATCTAATGAGT  8389
           ************************************************************

C57BL/6J   GGTTCACCAACTTACCAATAGGCTGTGTGAATACTAGCATGCTCAGGCTCACCTGGGAGC  8451
mod        GGTTCACCAACTTACCAATAGGCTGTGTGAATACTAGCATGCTCAGGCTCACCTGGGAGC  8444
BALB/cJ    GGTTCACCAACTTACCAATAGGCTGTGTGAATACTAGCATGCTCAGGCTCACCTGGGAGC  8449
           ************************************************************

C57BL/6J   TGTCTGGAAAGGCCTCCATCAATATGTATCTTAATAGCCCCAGATGATGAGGGAGCATAT  8511
mod        TGTCTGGAAAGGCCTCCATCAATATGTATCTTAATAGCCCCAGATGATGAGGGAGCATAT  8504
BALB/cJ    TGTCTGGAAAGGCCTCCATCAATATGTATCTTAATAGCCCCAGATGATGAGGGAGCATAT  8509
           ************************************************************
```

```
C57BL/6J   TATACAGGAGTTCTAGTCCACACTGAAATGGAGTCAATCTTCTGCCTACTCTACTCTCAA  8571
mod        TATACAGGAGTTCTAGTCCACACTGAAATGGAGTCAATCTTCTGCCTACTCTACTCTCAA  8564
BALB/cJ    TATACAGGAGTTCTAGTCCACACTGAAATGGAGTCAATCTTCTGCCTACTCTACTCTCAA  8569
           ************************************************************

C57BL/6J   TGAGCCAACCATGCCCTCCATCAGAAAGACCGACATGCCAGAAGCTTGGCAATGCCTCAA  8631
mod        TGAGCCAACCATGCCCTCCATCAGAAAGACCGACATGCCAGAAGCTTGGCAATGCCTCAA  8624
BALB/cJ    TGAGCCAACCATGCCCTCCATCAGAAAGACCGACATGCCAGAAGCTTGGCAATGCCTCAA  8629
           ************************************************************

C57BL/6J   GGCTGGTCAATGAACTCACACCCTTTCTTTTTCCTCCTCTCCTCGTGGGGGACTGAATTC  8691
mod        GGCTGGTCAATGAACTCACACCCTTTCTTTTTCCTCCTCTCCTCGTGGGGGACTGAATTC  8684
BALB/cJ    GGCTGGTCAATGAACTCACACCCTTTCTTTTTCCTCCTCTCCTCGTGGGGGACTGAATTC  8689
           ************************************************************

C57BL/6J   ACTTATTCCATGCATTCGGCTAAAACTTGTAGCTAAAAATCCTTTTCTTATTTGAAAATG  8751
mod        ACTTATTCCATGCATTCGGCTAAAACTTGTAGCTAAAAATCCTTTTCTTATTTGAAAATG  8744
BALB/cJ    ACTTATTCCATGCATTCGGCTAAAACTTGTAGCTAAAAATCCTTTTCTTATTTGAAAATG  8749
           ************************************************************

C57BL/6J   TGAACTTACCCTACTTGGCCAGCCTGTGTGTACACACATGTAAAGCTTGATTAGTGCCAG  8811
mod        TGAACTTACCCTACTTGGCCAGCCTGTGTGTACACACATGTAAAGCTTGATTAGTGCCAG  8804
BALB/cJ    TGAACTTACCCTACTTGGCCAGCCTGTGTGTACACACATGTAAAGCTTGATTAGTGCCAG  8809
           ************************************************************

C57BL/6J   ATGTGGTAGCACATGCCTATAATTTTAGCACTTAGAAGGTGGAGGGGGGACAGCAACTCA  8871
mod        ATGTGGTAGCACATGCCTATAATTTTAGCACTTAGAAGGTGGAGGGGGGACAGCAACTCA  8864
BALB/cJ    ATGTGGTAGCACATGCCTATAATTTTAGCACTTAGAAGGTGGAGGGGGGACAGCAACTCA  8869
           ************************************************************

C57BL/6J   AAGCTGTTCTCTACTATGGGAGTTTGAGGCAAGCCTGGGCTACATGAGACCTGTTTTGTA  8931
mod        AAGCTGTTCTCTACTATGGGAGTTTGAGGCAAGCCTGGGCTACATGAGACCTGTTTTGTA  8924
BALB/cJ    AAGCTGTTCTCTACTATGGGAGTTTGAGGCAAGCCTGGGCTACATGAGACCTGTTTTGTA  8929
           ************************************************************

C57BL/6J   AACACACAAACAAAACCCAAAACACCCAACCCCCTAAAAAGTATTTTCTTGTCTTCTTCA  8991
mod        AACACACAAACAAAACCCAAAACACCCAACCCCCTAAAAAGTATTTTCTTGTCTTCTTCA  8984
BALB/cJ    AACACACAAACAAAACCCAAAACACCCAACCCCCTAAAAAGTATTTTCTTGTCTTCTTCA  8989
           ************************************************************

C57BL/6J   CACTTTCCTGCTTTTTAACTTTTTTT---TTTTTTGGTTTTTCGAGACAGGGTTTCTCTG  9048
mod        CACTTTCCTGCTTTTTAACTTTTTTTTTTTTTTTGGTTTTTCGAGACAGGGTTTCTCTG  9044
BALB/cJ    CACTTTCCTGCTTTTTAACTTTTTTTTTTTTTTTGGTTTTTCGAGACAGGGTTTCTCTG  9049
           *************************   *******************************

C57BL/6J   TGTAGTCCTGGCTGTCCTGGAATTCACTCTGTAGACCAGGCTGGCCTTGAACTCAGAAAT  9108
mod        TGTAGTCCTGGCTGTCCTGGAATTCACTCTGTAGACCAGGCTGGCCTTGAACTCAGAAAT  9104
BALB/cJ    TGTAGTCCTGGCTGTCCTGGAATTCACTCTGTAGACCAGGCTGGCCTTGAACTCAGAAAT  9109
           ************************************************************

C57BL/6J   GCGCCTGCCTCTGCCTCCCAAGTGCTAGGATTAAGGGTGTGTGCCACCACGCCTGGCACT  9168
mod        GCGCCTGTCTCTGCCTCCCAAGTGCTAGGATTAAGGGTGTGTGCCACCACGCCTGGCACT  9164
BALB/cJ    GCGCCTGTCTCTGCCTCCCAAGTGCTAGGATTAAGGGTGTGTGCCACCACGCCTGGCACT  9169
           ******* ****************************************************

C57BL/6J   TTCCTGCTTTCTAATAGACACTAGTGCTTTTCCTAAAAAAGTAGCCTGAAGAGTAGCCAG  9228
mod        TTCCTGCTTTCTAATAGACACTAGTGCTTTTCCTAAAAAAGTAGCCTGAAGAGTAGCCAG  9224
BALB/cJ    TTCCTGCTTTCTAATAGACACTAGTGCTTTTCCTAAAAAAGTAGCCTGAAGAGTAGCCAG  9229
           ************************************************************

C57BL/6J   AAGGTACAGGAGTTCTAGCCCCACCAGGGAAGGGCAGAGTGATCCTATTTCACAGAGTAC  9288
mod        AAGGTACAGGAGTTCTAGCCCCACCAGGGAAGGGCAGAGTGATCCTATTTCACAGAGTAC  9284
BALB/cJ    AAGGTACAGGAGTTCTAGCCCCACCAGGGAAGGGCAGAGTGATCCTATTTCACAGAGTAC  9289
           ************************************************************
```

```
C57BL/6J   CAATAAAAGCAAAACTAAACAATGCTACCAAGGACAGCCTGATTAGGTCACCTCTCCAGA  9348
mod        CAATAAAAGCAAAACTAAACAATGCTACCAAGGACAGCCTGATTAGGTCACCTCTCCAGA  9344
BALB/cJ    CAATAAAAGCAAAACTAAACAATGCTACCAAGGACAGCCTGATTAGGTCACCTCTCCAGA  9349
           ************************************************************

C57BL/6J   CATGGAGCATCGTACACAGCCTCTTGCTATTCTAAAGACATACTAGTATCTTAAAACACT  9408
mod        CATGGAGCATCGTACACAGCCTCTTGCTATTCTAAAGACATACTAGTATCTTAAAACACT  9404
BALB/cJ    CATGGAGCATCGTACACAGCCTCTTGCTATTCTAAAGACATACTAGTATCTTAAAACACT  9409
           ************************************************************

C57BL/6J   GGATTTTCCTCCCCAGTGGGTCATGGTTTTAATCAGTGGTCACCCGGGGAAGCCAACATT  9468
mod        GGATTTTCCTCCCCAGTGGGTCATGGTTTTAATCAGTGGTCACCCGGGGAAGCCAACATT  9464
BALB/cJ    GGATTTTCCTCCCCAGTGGGTCATGGTTTTAATCAGTGGTCACCCGGGGAAGCCAACATT  9469
           ************************************************************

C57BL/6J   GCCTTAGAGTTGCTTCTTTCTCGGGTGGTTAAGAGTTTCCAAAATGAAAACTCACTGTCC  9528
mod        GCCTTAGAGTTGCTTCTTTCTCGGGCGGTTAAGAGTTTCCAAAATGAAAACTCACTGTCC  9524
BALB/cJ    GCCTTAGAGTTGCTTCTTTCTCGGGCGGTTAAGAGTTTCCAAAATGAAAACTCACTGTCC  9529
           ************************ **********************************

C57BL/6J   TTTACTTTGCAAAGTCAGCATCTTACCCAAAGGAACAAAATCTGGCCTTGTGATGCCAAG  9588
mod        TTTACTTTGCAAAGTCAGCATCTTACCCAAAGGAACAAAATCTGGCCTTGTGATGCCAAG  9584
BALB/cJ    TTTACTTTGCAAAGTCAGCATCTTACCCAAAGGAACAAAATCTGGCCTTGTGATGCCAAG  9589
           ************************************************************

C57BL/6J   AGGAGTACACAGGCCAAGCTTAGAGTCTCTAAATTTGACCCTGCAGTCAGGTACCCAGCA  9648
mod        AGGAGTACACAGGCCAAGCTTAGAGTCTCTAAATTTGACCCTGCAGTCAGGTACCCAGCA  9644
BALB/cJ    AGGAGTACACAGGCCAAGCTTAGAGTCTCTAAATTTGACCCTGCAGTCAGGTACCCAGCA  9649
           ************************************************************

C57BL/6J   TTTATGAAGCCTCAGAAGGGACCCAAACACAAAATCTGGGTACACTGTATCATTTCCTAT  9708
mod        TTTATGAAGCCTCAGAAGGGACCCAAACACAAAATCTGGGTACACTGTATCATTTCCTAT  9704
BALB/cJ    TTTATGAAGCCTCAGAAGGGACCCAAACACAAAATCTGGGTACACTGTATCATTTCCTAT  9709
           ************************************************************

C57BL/6J   GTGTACTTGAGCTAGGTGACCATATTGGAGGTAGACAATGTTTTCTTTCTGTCTTGCTTT  9768
mod        GTGTACTTGAGCTAGGTGACCATATTGGAGGTAGACAACGTTTTCTTTCTGCCTTGCTTT  9764
BALB/cJ    GTGTACTTGAGCTAGGTGACCATATTGGAGGTAGACAACGTTTTCTTTCTGCCTTGCTTT  9769
           ************************************ ************ ********

C57BL/6J   AAGGCAGGCTCTTAAGTATCCCTGGCTATCCTCAAACTCAATATAGCCAAGGGTTGGGAT  9828
mod        AAGGCAGGCTCTTAAGTATCCCTGGCTATCCTCAAACTCAATATAGCCAAGGGTTGGGAT  9824
BALB/cJ    AAGGCAGGCTCTTAAGTATCCCTGGCTATCCTCAAACTCAATATAGCCAAGGGTTGGGAT  9829
           ************************************************************

C57BL/6J   TACAGAGATAAGCCACCTGCACTTGTTTCACATGATATTGGAGATGGGATTCAAGGGTTC  9888
mod        TACAGAGATAAGCCACCTGCACTTGTTTCACATGATATTGGAGATGGGATTCAAGGGTTC  9884
BALB/cJ    TACAGAGATAAGCCACCTGCACTTGTTTCACATGATATTGGAGATGGGATTCAAGGGTTC  9889
           ************************************************************

C57BL/6J   ACGCATACCTAGCACCCCAAAGGGGTACATCTCTGTCCCCTCCACAGTTCTGAAGGAGAG  9948
mod        ACGCATACCTAGCACCCCAAAGGGGTACATCTCTGTCCCCTCCACAGTTCTGAAGGAGAG  9944
BALB/cJ    ACGCATACCTAGCACCCCAAAGGGGTACATCTCTGTCCCCTCCACAGTTCTGAAGGAGAG  9949
           ************************************************************

C57BL/6J   GGGCCTAGAGCTGATTGTCACTAATTACCAGAAGCCAAAGTCATCTCTCTGTAGATGTTT  10008
mod        GGGCCTAGAGCTGATTGTCACTAATTACCAGAAGCCAAAGTCATCTCTCTGTAGATGTTT  10004
BALB/cJ    GGGCCTAGAGCTGATTGTCACTAATTACCAGAAGCCAAAGTCATCTCTCTGTAGATGTTT  10009
           ************************************************************

C57BL/6J   GCAACTGGAATCTCAACACATGCTCTTTCAAGCTGTGGAGTTGCCTCTACTCTAACCCAA  10068
mod        GCAACTGGAATCTCAACACATGCTCTTTCAAGCTGTGGAGTTGCCTCTACTCTAACCCAA  10064
BALB/cJ    GCAACTGGAATCTCAACACATGCTCTTTCAAGCTGTGGAGTTGCCTCTACTCTAACCCAA  10069
           ************************************************************
```

```
C57BL/6J   GAGACTGCTAGACCAAGATGCCTGTAAGAGACAGGCTGAGATACACCCCCTACAAGACAG  10128
mod        GAGACTGCTAGACCAAGATGCCTGTAAGAGACAGGCTGAGATGCACCCCCTACAAGACAG  10124
BALB/cJ    GAGACTGCTAGACCAAGATGCCTGTAAGAGACAGGCTGAGATGCACCCCCTACAAGACAG  10129
           ************************************** ****************

C57BL/6J   AAAAGGAATAAGCCACGAAGACAATAACGATTTTTGTATCAAGCGTCCTCTCCCATTTCA  10188
mod        AAAAGGAATAAGCCACGAAGACAATAACGATTTTTGTATCAAGCGTCCTCTCCCATTTCA  10184
BALB/cJ    AAAAGGAATAAGCCACGAAGACAATAACGATTTTTGTATCAAGCGTCCTCTCCCATTTCA  10189
           ************************************************************

C57BL/6J   GCTTACCTGACAATGAAATCAAATTCGGACCCTGCAAGCATCAGTACACCCAGCAGAGTG  10248
mod        GCTTACCTGACAATGAAATCAAATTCGGACCCTGCAAGCATCAGTACACCCAGCAGAGTG  10244
BALB/cJ    GCTTACCTGACAATGAAATCAAATTCGGACCCTGCAAGCATCAGTACACCCAGCAGAGTG  10249
           ************************************************************

C57BL/6J   GACACAGCACCGTCCAGAACGGGAGCAAACATGTGCTCCAGAGCGAGCATAGCCCTGTGG  10308
mod        GACACAGCACCGTCCAGAACGGGAGCAAACATGTGCTCCAGAGCGAGCATAGCCCTGTGG  10304
BALB/cJ    GACACAGCACCGTCCAGAACGGGAGCAAACATGTGCTCCAGAGCGAGCATAGCCCTGTGG  10309
           ************************************************************

C57BL/6J   TTCTTGTCCCCAATGGCTGTCAGAAAGGCCTGAACAAAGGAGAAAATTGACACGGTCACA  10368
mod        TTCTTGTCCCCAATGGCTGTCAGAAAGGCCTGAACAAAGGAGAAAATTGACACGGTCACA  10364
BALB/cJ    TTCTTGTCCCCAATGGCTGTCAGAAAGGCCTGAACAAAGGAGAAAATTGACACGGTCACA  10369
           ************************************************************

C57BL/6J   TTCTGGGTGTGGTAAAGTGCTCAGCTGTGTCTATACTTGGGTTTTGTTGACTTTTCTTAT  10428
mod        TTCTGGGTGTGGTAAAGTACTCAGCTGTGTCTATACTTGGGTTTTGTTGACTTTTCTTAT  10424
BALB/cJ    TTCTGGGTGTGGTAAAGTACTCAGCTGTGTCTATACTTGGGTTTTGTTGACTTTTCTTAT  10429
           ****************** ****************************************

C57BL/6J   GTCCCTCAGACCGTTCTCAAACTCACTACTGTACAAGAATGGCCTTGAGCTAAAGCCTCC  10488
mod        GTCCCTCAGACCGTTCTCAAACTCACTACTGTACAAGAATGGCCTTGAGCTAAAGCCTCC  10484
BALB/cJ    GTCCCTCAGACCGTTCTCAAACTCACTACTGTACAAGAATGGCCTTGAGCTAAAGCCTCC  10489
           ************************************************************

C57BL/6J   TTGTTCTATTTCTGGAGGGCCGACGTTGACAGGTGTGCACTACTATGCCTGGTTTTCTGT  10548
mod        TTGTTCTATTTCTGGAGGGCCGACGTTGACAGGTGTGCACTACTATGCCTGGTTTTCTGT  10544
BALB/cJ    TTGTTCTATTTCTGGAGGGCCGACGTTGACAGGTGTGCACTACTATGCCTGGTTTTCTGT  10549
           ************************************************************

C57BL/6J   GGTGCTGAGGATAGAGCCCAGGGCATCCTGCAGACTAGGCATGCACATTGCCGCCTTACC  10608
mod        GGTGCTGAGGATAGAGCCCAGGGCATCCTGCAGACTAGGCAAGCACATTGCCGCCTTACC  10604
BALB/cJ    GGTGCTGAGGATAGAGCCCAGGGCATCCTGCAGACTAGGCAAGCACATTGCCGCCTTACC  10609
           ****************************************** ****************

C57BL/6J   TATAGCCCAGCCCCTACCCTACAAGTTAAGTAATTCTTAAGTTTCTGTAGGTGGACAGAG  10668
mod        TATAGCCCAGCCCCTACCCTACAAGTTAAGTAATTCTTAAGTTTCTGTAGGTGGACAGAG  10664
BALB/cJ    TATAGCCCAGCCCCTACCCTACAAGTTAAGTAATTCTTAAGTTTCTGTAGGTGGACAGAG  10669
           ************************************************************

C57BL/6J   CCATCTCAAAGAGACTTCTCTATCATTAAGGAGAAAAAAAaTCATCCAAAATGTAGCAAT  10728
mod        CCATCTCAAAGAGACTTCTCTATCATTAAGGAGAAAAAAAaTCATCCAAAATGTAGCAAT  10724
BALB/cJ    CCATCTCAAAGAGACTTCTCTATCATTAAGGAGAAAAAAAATCATCCAAAATGTAGCAAT  10729
           ************************************************************

C57BL/6J   AAAGCCAAAAGGATGACTGACTTTTTAATGTCACGATAAAACATCATTTATTAGGTCTTC  10788
mod        AAAGCCAAAAGGATGACTGACTTTTTAATGTCACGATAAAACATCATTTATTAGGTCTTC  10784
BALB/cJ    AAAGCCAAAAGGATGACTGACTTTTTAATGTCACGATAAAACATCATTTATTAGGTCTTC  10789
           ************************************************************

C57BL/6J   TGAAAATAAGAGTGTGTATAGGTGTTTTGCTCGTATGTATGTGCATACGTCTGTTCACCA  10848
mod        TGAAAATAAGAGTGTGTATAGGTGTTTTGCTCGTATGTATGTGCATACGTCTGTTCACCA  10844
BALB/cJ    TGAAAATAAGAGTGTGTATAGGTGTTTTGCTCGTATGTATGTGCATACGTCTGTTCACCA  10849
           ************************************************************
```

```
C57BL/6J   TGTGCTTGCCTGGGGCCCACAGGAGATAAGAGGGCCTTAGATCCACTGGAGCTGGAGTAG  10908
mod        TGTGCTTGCCTGGGGCCCACAGGAGATAAGAGGGCCTTAGATCCACTGGAGCTGGAGTAG  10904
BALB/cJ    TGTGCTTGCCTGGGGCCCACAGGAGATAAGAGGGCCTTAGATCCACTGGAGCTGGAGTAG  10909
           ************************************************************

C57BL/6J   AAGGTGGTTTTGAGCTTATGCTAGCTAGCCAAGCCCTCTACCAATGAGCTTACATTCCAG  10968
mod        AAGGTGGTTTTGAGCTTATGCTAGCTAGCCAAGCCCTCTACCAATGAGCTTACATTCCAG  10964
BALB/cJ    AAGGTGGTTTTGAGCTTATGCTAGCTAGCCAAGCCCTCTACCAATGAGCTTACATTCCAG  10969
           ************************************************************

C57BL/6J   CCTCAGCTCTGGCGGAGATATTAGACACTGCGTCTGCCTTTTCTTCCCCCCAAATCTATG  11028
mod        CCTCAGCTCTGGCGGAGATATTAGACACTGCGTCTGCCTTTTCTTCCCCCCAAATCTATG  11024
BALB/cJ    CCTCAGCTCTGGCGGAGATATTAGACACTGCGTCTGCCTTTTCTTCCCCCCAAATCTATG  11029
           ************************************************************

C57BL/6J   TTCAATAGTCTACTGCTTAGCTCCTACAATGTTTATTTGTTCTCACTGAATCTGGATTCA  11088
mod        TTCAATAGTCTACTGCTTAGCTCCTACAATGTTTATTTGTTCTCACTGAATCTGGATTCA  11084
BALB/cJ    TTCAATAGTCTACTGCTTAGCTCCTACAATGTTTATTTGTTCTCACTGAATCTGGATTCA  11089
           ************************************************************

C57BL/6J   TTACAATGCCCTGTTCATAGTACTGAGCCATAGTTAATGATAACGCATATTTTGAACAAA  11148
mod        TTACAATGCCCTGTTCATAGTACTGAGCCATAGTTAATGATAACGCATATTTTGAACAAA  11144
BALB/cJ    TTACAATGCCCTGTTCATAGTACTGAGCCATAGTTAATGATAACGCATATTTTGAACAAA  11149
           ************************************************************

C57BL/6J   ACCCAAAGCACGGTCTAGGCTTATAATCCCAACAGGCAGATCTCTGGACTTTGCCTATCA  11208
mod        ACCCAAAGCACGGTCTAGGCTTATAATCCCAACAGGCAGATCTCTGGACTTTGCCTATCA  11204
BALB/cJ    ACCCAAAGCACGGTCTAGGCTTATAATCCCAACAGGCAGATCTCTGGACTTTGCCTATCA  11209
           ************************************************************

C57BL/6J   GTCAGCCTAACCTACTTAGTGTGCTCCAGGCCAGTAAGAGAATCTGTGCAAAAAATCCCC  11268
mod        GTCAGCCTAACCTACTTAGTGTGCTCCAGGCCAGTAAGAGAATCTGTGCAAAAAATCCCC  11264
BALB/cJ    GTCAGCCTAACCTACTTAGTGTGCTCCAGGCCAGTAAGAGAATCTGTGCAAAAAATCCCC  11269
           ************************************************************

C57BL/6J   AAGGCAAATGGTGCCTCAGTAACAGTATCTGCTACTTTCCAGTCTCCACAGGCATCTCAC  11328
mod        AAGGCAAATGGTGCCTCAGTAACAGTATCTGCTACTTTCCAGTCTCCACAGGCATCTCAC  11324
BALB/cJ    AAGGCAAATGGTGCCTCAGTAACAGTATCTGCTACTTTCCAGTCTCCACAGGCATCTCAC  11329
           ************************************************************

C57BL/6J   TCGGTTGCTTGCAGACACACAGGCAACTCACATACACGGACACGTCTGAGATACTGTTTC  11388
mod        TCGGTTGCTTGCAGACACACAGGCAACTCACATACACGGACACGTCTGAGATACTGTTTC  11384
BALB/cJ    TCGGTTGCTTGCAGACACACAGGCAACTCACATACACGGACACGTCTGAGATACTGTTTC  11389
           ************************************************************

C57BL/6J   TAAGTCATCATCTTACTAATAGCTTCAAAACACGTTCTGTACATCCACATCCACAGCTGG  11448
mod        TAAGTCATCATCTTACTAATAGCTTCAAAACACATTCTGTACATCCACATCCACAGCTGG  11444
BALB/cJ    TAAGTCATCATCTTACTAATAGCTTCAAAACACATTCTGTACATCCACATCCACAGCTGG  11449
           ********************************* ************************

C57BL/6J   TGAGAGCGGTGACAGCTCTGAGAGGCGACAGAATGCACATACAGTGTGCTACACACAAAG  11508
mod        TGAGAGCGGTGACAGCTCTGAGAGGCGACAGAATGCACATACAGTGTGCTACACACAAAG  11504
BALB/cJ    TGAGAGCGGTGACAGCTCTGAGAGGCGACAGAATGCACATACAGTGTGCTACACACAAAG  11509
           ************************************************************

C57BL/6J   CACAATGCTGGAAAGGCAATGGTAGCATCCCTGGGTTCAGGCACAACTTGGGTATTCCAA  11568
mod        CACAATGCTGGAAAGGCAATGGTAGCATCCCTGGGTTCAGGCACAACTTGGGTATTCCAA  11564
BALB/cJ    CACAATGCTGGAAAGGCAATGGTAGCATCCCTGGGTTCAGGCACAACTTGGGTATTCCAA  11569
           ************************************************************

C57BL/6J   ACACTATCATTTATTCAGTCAATACTCTTTATCAAATGCCTGTTAACAGCCCTGGGTACT  11628
mod        ACACTATCATTTATTCAGTCAATACTCTTTATCAAATGCCTGTTAACAGCCCTGGGTACT  11624
BALB/cJ    ACACTATCATTTATTCAGTCAATACTCTTTATCAAATGCCTGTTAACAGCCCTGGGTACT  11629
           ************************************************************
```

```
C57BL/6J   TAACACTATGAGTACATCTGTCGGCAGAAACTAAACCACACAATGCTACAAGGTTTTTAA  11688
mod        TAACACTATGAGTACATCTGTCGGCAGAAACTAAACCACACAATGCTACAAGGTTTTTAA  11684
BALB/cJ    TAACACTATGAGTACATCTGTCGGCAGAAACTAAACCACACAATGCTACAAGGTTTTTAA  11689
           ************************************************************

C57BL/6J   GTAGTTTTCTCAGTGTGTAAAACAACCAGCTCATCTTTTGGATTCTAGACTCTCTTTCCA  11748
mod        GTAGTTTTCTCAGTATGTAAAACAACCAGCTCATCTTTTGGATTCTAGACTCTCTTTCCA  11744
BALB/cJ    GTAGTTTTCTCAGTATGTAAAACAACCAGCTCATCTTTTGGATTCTAGACTCTCTTTCCA  11749
           ************** ********************************************

C57BL/6J   GTTAACTTCTTCCAACAACCAGGCAGTGTATATATAGAGAGGGGCCTACATCCTGGTGAA  11808
mod        GTTAACTTCTTCCAACAACCAGGCAGTGTATATATAGAGAGGGGCCTACATCCTGGTGAA  11804
BALB/cJ    GTTAACTTCTTCCAACAACCAGGCAGTGTATATATAGAGAGGGGCCTACATCCTGGTGAA  11809
           ************************************************************

C57BL/6J   AGCATCCCTGGTTTTCTTGTTTTCTGCCACCATGGATGCATGCAGAAGCTGGGGGGTCCC  11868
mod        AGCATCCCTGGTTTTCTTGTTTTCTGCCACCATGGATGCATGCAGAAGCTGGGGGGTCCC  11864
BALB/cJ    AGCATCCCTGGTTTTCTTGTTTTCTGCCACCATGGATGCATGCAGAAGCTGGGGGGTCCC  11869
           ************************************************************

C57BL/6J   CAGAGCTGGGTGGGCTGCGGTGGAAGAGCAGGTTCTTCAGGGCTGTGCACCCGAGCTTTC  11928
mod        CAGAGCTGGGTGGGCTGCGGTGGAAGAGCAGGCTCTTCAGGGCTGTGCACCCGAGCTTTC  11924
BALB/cJ    CAGAGCTGGGTGGGCTGCGGTGGAAGAGCAGGCTCTTCAGGGCTGTGCACCCGAGCTTTC  11929
           ******************************** ***************************

C57BL/6J   ATTCACAAGCCTGAGGGCATTTCTGTGGCTGCCAGCTACTTCACCCTCAATTATGCTATC  11988
mod        ATTCACAAGCCTGAGGGCATTTCTGTGGCTGCCAGCTACTTCACCCTCAATTATGCTATC  11984
BALB/cJ    ATTCACAAGCCTGAGGGCATTTCTGTGGCTGCCAGCTACTTCACCCTCAATTATGCTATC  11989
           ************************************************************

C57BL/6J   C  11989
mod        C  11985
BALB/cJ    C  11990
           *
```

Table S3 Sequence variation between inbred mouse strains in *Ptch1*

List of SNPs relative to the C57BL/6J mouse reference genome. SNPs that differ from the reference C57BL/6J but match C57BL/6J.*Mod13* are highlighted yellow. SNPs that differ from both the reference C57BL/6J and C57BL/6J.*Mod13* are highlighted blue. Harwell strain variants identified by comparison of *Ptch1* sequences. All other strain information downloaded from the Mouse Genomes Project, www.sanger.ac.uk/science/data/mouse-genomes-project, release REL-1505, location 13:63507918-63519906. Position coordinates are relative to the GRCm38 assembly. Insertions, deletions and larger structural variations are not shown. - indicates high confidence match to reference. ~ indicates low confidence match to reference.

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs249869728	13:63507926	**A**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs264656256	13:63507948	**T**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs233706103	13:63507977	**C**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs46434257	13:63508002	**G**	-	-	-	-	-	-	-	-	-	-	A	-	A	-	-	-	-	-
rs220513150	13:63508031	**C**	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs581380403	13:63508032	**G**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs245546526	13:63508083	**C**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs213331878	13:63508090	**C**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs6338126	13:63508136	**T**	-	-	C	-	-	-	-	-	C	-	-	C	-	-	C	C	-	C
rs48471164	13:63508144	**C**	-	-	T	-	-	-	-	-	-	-	-	T	-	-	-	-	-	T
rs6338149	13:63508151	**C**	-	-	T	-	-	-	-	-	-	-	-	T	-	-	T	T	-	T
rs50027181	13:63508176	**A**	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-	-	-	T
rs46548499	13:63508220	**C**	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs239386952	13:63508242	**C**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs250693181	13:63508244	**G**	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs46414852	13:63508265	**G**	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs245129678	13:63508298	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-
rs45860437	13:63508310	A	-	T	T	-	T	-	-	-	-	-	T	T	T	-	-	-	T	T
rs234974821	13:63508442	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs247457124	13:63508492	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs259770841	13:63508517	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs46672527	13:63508751	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	T	-	-	-
rs245515005	13:63508803	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs213445670	13:63508947	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs228229275	13:63509099	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs30042039	13:63509253	C	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs212362700	13:63509257	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs243487983	13:63509270	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs13472134	13:63509382	T	-	C	C	-	C	-	-	-	C	-	C	C	C	-	C	C	C	C
rs216183264	13:63509403	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs250611678	13:63509422	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G/C	-	-	-
rs219522079	13:63509464	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs48457688	13:63509606	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs47050445	13:63509624	G	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-
rs47381254	13:63509631	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-
rs250095211	13:63509657	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs29225396	13:63509699	T	-	C	C	-	C	-	-	-	C	-	C	C	C	-	C	C	C	C
rs227183261	13:63509711	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs239587824	13:63509716	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
rs257295287	13:63509736	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
rs49107169	13:63509762	C	-	T	-	-	T	-	-	-	-	-	T	T	T	-	T	T	T	T
rs248988177	13:63509765	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	-	-
rs218667136	13:63509774	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-
rs232864967	13:63509785	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-
rs47366571	13:63509809	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	G	-	-
-	13:63509824	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T
rs216282818	13:63509836	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-
rs29771248	13:63509873	A	-	G	G	-	G	-	-	-	G	-	G	G	G	-	G	G	G	G
rs250767848	13:63509887	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs213940875	13:63509917	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs237392119	13:63509972	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs255706884	13:63510100	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs50151146	13:63510131	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
rs29243318	13:63510150	G	-	A	-	-	A	-	-	-	-	-	A	-	A	-	-	-	A	-
rs253539500	13:63510155	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs51549361	13:63510175	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-
rs239553251	13:63510214	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs257411251	13:63510223	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs46882028	13:63510233	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-
rs243733813	13:63510254	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs264334781	13:63510291	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs233635016	13:63510327	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs578375480	13:63510382	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs582091424	13:63510411	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs245691055	13:63510419	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs258470734	13:63510446	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
-	13:63510472	G	-	-	-	-	-	-	-	-		-	-	-	-	-	-	A	-	-
rs227158337	13:63510527	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs244914028	13:63510545	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs214519628	13:63510857	A	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-
rs48347912	13:63510907	A	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	C	-	-
rs252866267	13:63511152	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs47132539	13:63511153	T	-	-	-	-	-	-	-	-	G	-	-	-	-	-	G	G	-	-
rs235573988	13:63511368	A	-	-	-	-	-	~	-	-		-	-	-	-	-	G	-	-	-
-	13:63511372	T	-	-	-	-	-	~	-	-		-	-	-	-	-	A	~	-	-
rs578779351	13:63511391	A	-	-	-	-	-	~	-	-	-	-	-	-	-	-	G	-	-	-
rs232760861	13:63511484	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
-	13:63511554	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T
rs251241746	13:63511698	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs220907614	13:63511743	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	G	G	-	-
rs241026712	13:63511752	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs49370129	13:63511822	G	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs225556581	13:63511964	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs239668604	13:63512013	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs256836169	13:63512028	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs46960049	13:63512038	G	-	T	T	-	T	-	-	-	T	-	T	T	T	-	T	T	T	T
rs244883163	13:63512047	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs51562740	13:63512061	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs229515304	13:63512140	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs249383094	13:63512186	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs219267906	13:63512224	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs235673752	13:63512271	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs247441709	13:63512327	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs215258162	13:63512332	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs232996754	13:63512368	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs48818207	13:63512385	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	C	-	-
rs212806788	13:63512405	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs50063079	13:63512420	G	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs6163280	13:63512495	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-
rs6163355	13:63512542	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
rs29232668	13:63512752	G	-	A	-	-	A	-	-	-	-	-	A	A	A	-	-	-	A	-
rs583908079	13:63512827	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs251060039	13:63512890	T	-	A	-	-	-	~	-	-	G	-	-	-	-	-	G	G	-	-
rs51104582	13:63512933	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	T	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs237180812	13:63512965	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	A	-	-
rs265162966	13:63512971	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-
rs47984282	13:63513029	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs244531320	13:63513081	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs264549050	13:63513091	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	G	-	-
rs51091189	13:63513115	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-
rs586502414	13:63513166	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
rs29548072	13:63513234	G	-	-	-	-	A	-	-	-	-	-	A	A	A	-	-	-	A	A
rs226588239	13:63513266	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs250806730	13:63513276	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	-	-
rs45748091	13:63513292	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-
rs220367259	13:63513294	G	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs52059800	13:63513300	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	C	-	-
rs45707299	13:63513357	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs233471656	13:63513360	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs250979531	13:63513378	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs219888521	13:63513407	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
-	13:63513408	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
rs48074199	13:63513443	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-
rs50100770	13:63513477	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-
rs219736778	13:63513480	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
-	13:63513522	A	-	-	-	A/C	-	-	-	-	-	-	A/C	-	~	~	~	-	~	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs51509021	13:63513597	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs228723531	13:63513626	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs48679846	13:63513627	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	G	-	-
rs259268460	13:63513702	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs50558238	13:63513741	G	-	-	-	-	-	~	-	-	-	-	-	-	-	-	A	-	-	-
rs255530754	13:63513744	A	-	-	-	-	-	~	-	-	-	-	-	-	-	-	-	G	-	-
rs13481848	13:63513762	A	-	G	G	-	G	-	-	-	G	-	G	G	G	-	G	G	G	G
rs47881459	13:63513861	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	-	-
rs47204533	13:63513913	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-
rs29493976	13:63513965	G	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs233439035	13:63513977	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs245168521	13:63513994	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs214312807	13:63514009	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs49228649	13:63514019	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	A	A	-	-
rs261407337	13:63514033	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs219904761	13:63514037	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs242679274	13:63514042	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs48904135	13:63514059	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs50887990	13:63514104	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs47975699	13:63514105	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-
rs50888186	13:63514176	C	-	-	-	-	-	-	-	-	A	-	-	-	-	-	A	-	-	-
rs223052360	13:63514251	C	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs29517505	13:63514311	G	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs51221562	13:63514329	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs227660960	13:63514355	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs29956298	13:63514364	G	-	T	-	-	T	-	-	-	-	-	T	T	T	-	-	-	T	-
rs258932704	13:63514367	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	G	-	-
rs227530650	13:63514368	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	A	-	-
rs29551107	13:63514374	C	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs214271709	13:63514401	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs237955448	13:63514404	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-
rs249630066	13:63514419	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs48398359	13:63514447	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-
rs46463712	13:63514470	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	G	G	-	-
rs254057749	13:63514486	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs583394551	13:63514529	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
-	13:63514533	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs47569777	13:63514539	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	A	A	-	-
rs235153299	13:63514541	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs48863813	13:63514560	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	T	-	-	-
rs246282584	13:63514615	A	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-
rs262015580	13:63514666	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs51477632	13:63514797	G	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-
rs48345850	13:63514861	C	-	T	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs29781921	13:63514906	C	-	T	-	-	T	-	-	-	-	-	T	T	T	-	-	-	T	-
rs227682563	13:63515057	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	G	G	-	-
rs46569002	13:63515068	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	G	G	-	-
rs256296130	13:63515304	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-
rs229669024	13:63515425	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs29731536	13:63515778	G	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs211809826	13:63515887	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs47371629	13:63515901	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs247814298	13:63515913	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs387208459	13:63515923	G	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-
rs215606195	13:63515927	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	A	A	-	-
rs233472468	13:63515961	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs29225631	13:63515981	C	-	T	T	-	T	-	-	-	T	-	T	T	T	-	-	T	T	T
rs215667499	13:63515997	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs236430099	13:63516148	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-
rs261705676	13:63516168	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-
rs581416332	13:63516323	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-
rs220817249	13:63516371	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-
rs241727050	13:63516390	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-
rs252776734	13:63516397	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-
rs29932711	13:63516500	T	-	-	-	-	-	-	-	-	-	-	C	-	C	-	-	-	-	-
rs239309848	13:63516519	C	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs262751451	13:63516520	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
rs230198622	13:63516586	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	-	-	-
rs242201791	13:63516589	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs264594883	13:63516598	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs229158890	13:63516626	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs585238013	13:63516768	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-
rs247747937	13:63516808	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs215696154	13:63516847	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs579374582	13:63516963	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs580713056	13:63516971	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs584248611	13:63516984	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs29512560	13:63516988	T	-	-	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs245450056	13:63517008	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs213397617	13:63517026	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs238542280	13:63517027	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs30029736	13:63517033	C	-	T	T	-	T	-	-	-	T	-	T	-	T	-	T	T	T	T
-	13:63517040	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T
rs212577846	13:63517060	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	A	-	-
rs235439516	13:63517076	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs252719164	13:63517192	A	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-
rs221571031	13:63517198	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs239270382	13:63517276	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs262483769	13:63517328	A	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	T	-	-
rs222676539	13:63517345	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs244998923	13:63517363	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs261517369	13:63517371	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	A	-	-
rs223313099	13:63517383	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-
rs241901475	13:63517398	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs259648265	13:63517407	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs50656600	13:63517411	T	-	C	C	-	C	-	-	-	C	-	C	C	C	-	C	C	C	C
rs46897152	13:63517472	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-
rs264984392	13:63517580	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs228360993	13:63517595	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs29239373	13:63517664	T	-	C	C	-	C	-	-	-	-	-	C	C	C	-	-	-	C	
rs29495042	13:63517677	T	-	C	C	-	C	-	-	-	C	-	C	C	C	-	C	-	C	C
rs235409092	13:63517807	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs246966660	13:63517936	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs216218281	13:63517937	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs232993182	13:63518010	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs258080237	13:63518026	A	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-
rs29534888	13:63518028	A	-	G	G	-	G	-	-	-	G	-	G	G	G	-	G	G	G	G
rs235511992	13:63518047	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-
rs259768720	13:63518090	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs220951223	13:63518105	A	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	T	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs246596130	13:63518278	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs29245350	13:63518304	G	-	A	-	-	A	-	-	-	-	-	A	A	A	-	-	-	A	A
rs228798657	13:63518397	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	G	-	-	-
rs248883737	13:63518398	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-	-
rs254215342	13:63518427	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs46480049	13:63518438	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-
rs256158675	13:63518489	C	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs29683364	13:63518507	T	-	A	A	-	A	-	-	-	A	-	A	A	A	-	A	A	A	A
rs238197602	13:63518555	G	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-
rs237417116	13:63518556	T	-	-	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A
rs214049591	13:63518667	T	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs29907830	13:63518753	C	-	-	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
rs255547038	13:63518759	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs223544092	13:63518777	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs6161820	13:63518812	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-	-
rs253618772	13:63518823	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	-	-	-	-
rs220868797	13:63518883	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
-	13:63518912	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs6162339	13:63518920	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
rs6162817	13:63518984	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs6162870	13:63519023	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	-	-	-
rs243770107	13:63519043	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs259512595	13:63519076	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs228044549	13:63519111	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-
rs241160386	13:63519144	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	-	-	-	-
rs578488934	13:63519168	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs258513979	13:63519203	A	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	-	-	-
rs232416427	13:63519233	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs251790362	13:63519250	T	-	-	-	-	-	-	-	-	G	-	-	-	-	-	-	G	-	-
rs214561673	13:63519267	G	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-
rs228981445	13:63519283	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	A	-	-	-
rs243735935	13:63519315	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs217481319	13:63519325	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs29584348	13:63519339	G	-	A	A	-	A	-	-	-	-	-	A	A	A	-	-	-	A	A
rs253540984	13:63519355	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs215730134	13:63519363	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs238896316	13:63519372	C	-	-	-	-	-	-	-	-	T	-	-	-	-	-	-	-	-	-
rs256945300	13:63519379	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs220956442	13:63519384	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs240906915	13:63519389	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs253041393	13:63519392	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs221727391	13:63519434	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs239706476	13:63519445	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	C	-	-
rs258460914	13:63519454	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-

SNP ID	Position	Ref_C57BL/6J	Har_C57BL/6J	Har_C57BL/6J.Mod13	129S1_SvImJ	AKR_J	BALB_cJ	C3H_HeH	C57BL_6NJ	C58_J	CAST_EiJ	CBA_J	DBA_2J	FVB_NJ	NOD_ShiLtJ	NZO_HlLtJ	PWK_PhJ	SPRET_EiJ	WSB_EiJ	ZALENDE_EiJ
rs47439519	13:63519478	T	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	C	-	-
rs48588885	13:63519557	G	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	-	-	-
rs262511097	13:63519576	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs29783022	13:63519620	G	-	A	-	-	A	-	-	-	-	-	A	-	A	-	-	-	A	-
rs243651053	13:63519715	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs217591662	13:63519752	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs228571038	13:63519807	G	-	-	-	-	-	-	-	-	A	-	-	-	-	-	A	-	-	-
rs247325320	13:63519812	A	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	G	-	-
rs29980834	13:63519818	T	-	C	C	-	C	-	-	-	C	-	C	C	C	-	C	C	C	C
rs236950674	13:63519820	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-	-
rs262030804	13:63519833	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-
rs212696502	13:63519878	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	T	-	-

Appendix 3. Genes in *Mod13*

Table S4 List of protein-coding genes in *Mod13*

Chromosome 13, position 38.2 Mb – 75.7 Mb (from GRCm38 genome). Shaded genes indicate those located in B6 segments of *Mod13*, unshaded genes indicate those located in non-B6 segments. Genes present in the core segment of *Mod13* (60-71 Mb) are highlighted in yellow. Non-coding RNA genes and unclassified genes were not included.

Start	End	Gene Symbol	Name
38204938	38227665	*Snrnp48*	small nuclear ribonucleoprotein 48 (U11/U12)
38345716	38500314	*Bmp6*	bone morphogenetic protein 6
38500273	38528824	*Txndc5*	thioredoxin domain containing 5
38602706	38635876	*Bloc1s5*	biogenesis of lysosomal organelles complex-1, subunit 5, muted
38645691	38659028	*Eef1e1*	eukaryotic translation elongation factor 1 epsilon 1
38932140	38960875	*Slc35b3*	solute carrier family 35, member B3
40001882	40288011	*Ofcc1*	orofacial cleft 1 candidate 1
40715302	40738237	*Tfap2a*	transcription factor AP-2, alpha
40846739	40847554	*Gm35160*	predicted gene, 35160
40859768	40960892	*Gcnt2*	glucosaminyl (N-acetyl) transferase 2, I-branching enzyme
41001023	41013019	*Pak1ip1*	PAK1 interacting protein 1
41016292	41022586	*Tmem14c*	transmembrane protein 14C
41025008	41079706	*Mak*	male germ cell-associated kinase
41101427	41111035	*Gcm2*	glial cells missing homolog 2
41114307	41171684	*Sycp2l*	synaptonemal complex protein 2-like
41182495	41220405	*Elovl2*	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 2
41249860	41276574	*Smim13*	small integral membrane protein 13
41273803	41273977	*Gm17364*	predicted gene, 17364
41309916	41487360	*Nedd9*	neural precursor cell expressed, developmentally down-regulated gene 9
41606216	41641357	*Tmem170b*	transmembrane protein 170B
41763148	41847684	*Adtrp*	androgen dependent TFPI regulating protein
42051970	42185030	*Hivep1*	human immunodeficiency virus type I enhancer binding protein 1
42301476	42307990	*Edn1*	endothelin 1
42680623	43138526	*Phactr1*	phosphatase and actin regulator 1
43151740	43171501	*Tbc1d7*	TBC1 domain family, member 7
43195245	43304172	*Gfod1*	glucose-fructose oxidoreductase domain containing 1
43365496	43395203	*Sirt5*	sirtuin 5
43398376	43402858	*Nol7*	nucleolar protein 7
43402675	43481282	*Ranbp9*	RAN binding protein 9
43538393	43560191	*Mcur1*	mitochondrial calcium uniporter regulator 1
43615710	43670945	*Rnf182*	ring finger protein 182

Start	End	Gene Symbol	Name
43785112	43803133	*Cd83*	CD83 antigen
44729474	44921643	*Jarid2*	jumonji, AT rich interactive domain 2
44922077	45002096	*Dtnbp1*	dystrobrevin binding protein 1
45078692	45081222	*Gm9817*	predicted gene 9817
45389742	45411940	*Mylip*	myosin regulatory light chain interacting protein
45507444	45546382	*Gmpr*	guanosine monophosphate reductase
45549755	45965008	*Atxn1*	ataxin 1
46273721	46300115	*Stmnd1*	stathmin domain containing 1
46418434	46431095	*Rbm24*	RNA binding motif protein 24
46501848	46649642	*Cap2*	CAP, adenylate cyclase-associated protein, 2 (yeast)
46669522	46678056	*Fam8a1*	family with sequence similarity 8, member A1
46679905	46727940	*Nup153*	nucleoporin 153
46749087	46929867	*Kif13a*	kinesin family member 13A
47012557	47014850	*Nhlrc1*	NHL repeat containing 1
47025170	47044737	*Tpmt*	thiopurine methyltransferase
47043499	47084613	*Kdm1b*	lysine (K)-specific demethylase 1B
47084775	47106201	*Dek*	DEK oncogene (DNA binding)
47122720	47247991	*Rnf144b*	ring finger protein 144B
48261427	48264036	*Id4*	inhibitor of DNA binding 4
48487647	48513451	*Zfp169*	zinc finger protein 169
48513570	48514859	*A830005F24Rik*	RIKEN cDNA A830005F24 gene
48577872	48625664	*Ptpdc1*	protein tyrosine phosphatase domain containing 1
48662998	48666507	*Barx1*	BarH-like homeobox 1
48801750	48870885	*Phf2*	PHD finger protein 2
48879216	48967877	*Fam120a*	family with sequence similarity 120, member A
49036303	49148012	*Wnk2*	WNK lysine deficient protein kinase 2
49187485	49196244	*Ninj1*	ninjurin 1
49202950	49216037	*Card19*	caspase recruitment domain family, member 19
49230690	49248706	*Susd3*	sushi domain containing 3
49261554	49320311	*Fgd3*	FYVE, RhoGEF and PH domain containing 3
49341585	49387026	*Bicd2*	BICD cargo adaptor 2
49421190	49463060	*Ippk*	inositol 1,3,4,5,6-pentakisphosphate 2-kinase
49464023	49652785	*Cenpp*	centromere protein P
49504810	49532789	*Ecm2*	extracellular matrix protein 2, female organ and adipocyte specific
49544443	49567565	*Aspn*	asporin
49582462	49592822	*Omd*	osteomodulin
49608071	49624500	*Ogn*	osteoglycin
49653078	49679016	*Nol8*	nucleolar protein 8
49682100	49734267	*Iars*	isoleucine-tRNA synthetase
49750462	49752536	*Gm19334*	predicted gene, 19334
49764328	49765303	*Gm270*	predicted gene 270

Start	End	Gene Symbol	Name
49784742	49790401	*Gm30302*	predicted gene, 30302
49820728	49821646	*Gm272*	predicted gene 272
49861669	49864129	*Gm8672*	predicted gene 8672
49898842	49902468	*Gm8674*	predicted gene 8674
49976258	49977074	*Gm17611*	predicted gene, 17611
50099026	50104045	*Gm8700*	predicted gene 8700
50161821	50163398	*Gm19470*	predicted gene, 19470
50245181	50250308	*Gm906*	predicted gene 906
50399335	50400423	*Gm19536*	predicted gene, 19536
50417864	50433780	*Fbxw17*	F-box and WD-40 domain protein 17
50443408	50443912	*Gm47429*	predicted gene, 47429
50467307	50475355	*Nutm2*	NUT family member 2
50698312	50703435	*Gm8765*	predicted gene 8765
50885270	50891424	*Gm8777*	predicted gene 8777
51100880	51152546	*Spin1*	spindlin 1
51171022	51175188	*Nxnl2*	nucleoredoxin-like 2
51202727	51203065	*Hist1h2al*	histone cluster 1, H2al
51408639	51422797	*S1pr3*	sphingosine-1-phosphate receptor 3
51431041	51567084	*Shc3*	src homology 2 domain-containing transforming protein C3
51645232	51650660	*Cks2*	CDC28 protein kinase regulatory subunit 2
51651697	51684044	*Secisbp2*	SECIS binding protein 2
51685529	51701041	*Gm15440*	predicted gene 15440
51701246	51793747	*Sema4d*	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4D
51846678	51848468	*Gadd45g*	growth arrest and DNA-damage-inducible 45 gamma
52504375	52531241	*Diras2*	DIRAS family, GTP-binding RAS-like 2
52583173	52648792	*Syk*	spleen tyrosine kinase
52835119	52929681	*Auh*	AU RNA binding protein/enoyl-coenzyme A hydratase
52967209	52981073	*Nfil3*	nuclear factor, interleukin 3, regulated
53109312	53286124	*Ror2*	receptor tyrosine kinase-like orphan receptor 2
53332748	53377397	*Sptlc1*	serine palmitoyltransferase, long chain base subunit 1
53415943	53419345	*Gm38564*	predicted gene, 38564
53466884	53473074	*Msx2*	msh homeobox 2
54051183	54055658	*Drd1*	dopamine receptor D1
54071869	54108342	*Sfxn1*	sideroflexin 1
54192129	54236180	*Hrh2*	histamine receptor H2
54371349	54383917	*Cplx2*	complexin 2
54458837	54468849	*Thoc3*	THO complex 3
54503779	54551290	*Simc1*	SUMO-interacting motifs containing 1
54551218	54565435	*4833439L19Rik*	RIKEN cDNA 4833439L19 gene

Start	End	Gene Symbol	Name
54575015	54581128	*Arl10*	ADP-ribosylation factor-like 10
54584185	54590090	*Nop16*	NOP16 nucleolar protein
54590207	54591158	*Higd2a*	HIG1 domain family, member 2A
54592401	54611344	*Cltb*	clathrin, light polypeptide (Lcb)
54621784	54664068	*Faf2*	Fas associated factor family member 2
54679401	54693907	*Rnf44*	ring finger protein 44
54701461	54736662	*Cdhr2*	cadherin-related family member 2
54736671	54749874	*Gprin1*	G protein-regulated inducer of neurite outgrowth 1
54758860	54766595	*Sncb*	synuclein, beta
54783998	54788459	*Eif4e1b*	eukaryotic translation initiation factor 4E family member 1B
54789377	54796776	*Tspan17*	tetraspanin 17
54949411	55006018	*Unc5a*	unc-5 netrin receptor A
55005985	55021385	*Hk3*	hexokinase 3
55027880	55100300	*Uimc1*	ubiquitin interaction motif containing 1
55105311	55134825	*Zfp346*	zinc finger protein 346
55127839	55128183	*Gm46415*	predicted gene, 46415
55132962	55133548	*Gm17617*	predicted gene, 17617
55152640	55168759	*Fgfr4*	fibroblast growth factor receptor 4
55209782	55318325	*Nsd1*	nuclear receptor-binding SET-domain protein 1
55319743	55321946	*Rab24*	RAB24, member RAS oncogene family
55322055	55325272	*Prelid1*	PRELI domain containing 1
55325168	55329758	*Mxd3*	Max dimerization protein 3
55343833	55362783	*Lman2*	lectin, mannose-binding 2
55369732	55384687	*Rgs14*	regulator of G-protein signaling 14
55399622	55414695	*Slc34a1*	solute carrier family 34 (sodium phosphate), member 1
55414688	55415232	*Pfn3*	profilin 3
55414900	55415591	*BC038268*	cDNA sequence BC038268
55417958	55426793	*F12*	coagulation factor XII (Hageman factor)
55445072	55460927	*Grk6*	G protein-coupled receptor kinase 6
55464267	55473154	*Prr7*	proline rich 7 (synaptic)
55473429	55488111	*Dbn1*	drebrin 1
55495795	55513676	*Pdlim7*	PDZ and LIM domain 7
55523231	55529296	*Dok3*	docking protein 3
55530410	55536658	*Ddx41*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 41
55539316	55571120	*Fam193b*	family with sequence similarity 193, member B
55593135	55597697	*Tmed9*	transmembrane p24 trafficking protein 9
55599896	55610443	*B4galt7*	xylosylprotein beta1,4-galactosyltransferase, polypeptide 7 (galactosyltransferase I)
55623005	55632411	*Caml*	calcium modulating ligand
55635027	55681256	*Ddx46*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 46
55693124	55703500	*B230219D22Rik*	RIKEN cDNA B230219D22 gene
55714650	55726227	*Txndc15*	thioredoxin domain containing 15

Start	End	Gene Symbol	Name
55727368	55776830	*Pcbd2*	pterin 4 alpha carbinolamine dehydratase/dimerization cofactor of hepatocyte nuclear factor 1 alpha (TCF1) 2
55784568	55808998	*Catsper3*	cation channel, sperm associated 3
55825051	55836192	*Pitx1*	paired-like homeodomain transcription factor 1
56073619	56136361	*H2afy*	H2A histone family, member Y
56173704	56178885	*Tifab*	TRAF-interacting protein with forkhead-associated domain, family member B
56250492	56252163	*Neurog1*	neurogenin 1
56288643	56296551	*Cxcl14*	chemokine (C-X-C motif) ligand 14
56426594	56433753	*Gm3045*	predicted gene 3045
56438355	56474239	*Slc25a48*	solute carrier family 25, member 48
56479277	56482246	*Il9*	interleukin 9
56522472	56537898	*Fbxl21*	F-box and leucine-rich repeat protein 21
56542464	56548502	*Lect2*	leukocyte cell-derived chemotaxin 2
56601418	56602562	*Gm10781*	predicted gene 10781
56609603	56639339	*Tgfbi*	transforming growth factor, beta induced
56703010	56742377	*Smad5*	SMAD family member 5
56752371	56753827	*Gm45623*	predicted gene 45623
56773113	56895789	*Trpc7*	transient receptor potential cation channel, subfamily C, member 7
57421195	57908314	*Spock1*	sparc/osteonectin, cwcv and kazal-like domains proteoglycan 1
58004221	58113592	*Klhl3*	kelch-like 3
58125879	58128556	*Hnrnpa0*	heterogeneous nuclear ribonucleoprotein A0
58157649	58164693	*Idnk*	idnK gluconokinase homolog (E. coli)
58176156	58215653	*Ubqln1*	ubiquilin 1
58233346	58274839	*Gkap1*	G kinase anchoring protein 1
58287502	58359122	*Kif27*	kinesin family member 27
58380046	58385225	*2210016F16Rik*	RIKEN cDNA 2210016F16 gene
58391142	58403343	*Hnrnpk*	heterogeneous nuclear ribonucleoprotein K
58402248	58411148	*Rmi1*	RecQ mediated genome instability 1
58545399	58610877	*Slc28a3*	solute carrier family 28 (sodium-coupled nucleoside transporter), member 3
58806569	59133970	*Ntrk2*	neurotrophic tyrosine kinase, receptor, type 2
59449537	59585227	*Agtpbp1*	ATP/GTP binding protein 1
59585259	59634798	*Naa35*	N(alpha)-acetyltransferase 35, NatC auxiliary subunit
59634996	59675784	*Golm1*	golgi membrane protein 1
59687402	59694103	*4921517D22Rik*	RIKEN cDNA 4921517D22 gene
59700083	59706588	*Spata31d1a*	spermatogenesis associated 31 subfamily D, member 1A
59712284	59719289	*Spata31d1b*	spermatogenesis associated 31 subfamily D, member 1B
59725925	59731752	*Spata31d1d*	spermatogenesis associated 31 subfamily D, member 1D

Start	End	Gene Symbol	Name
59740842	59746752	*1700014D04Rik*	RIKEN cDNA 1700014D04 gene
59755409	59769810	*Isca1*	iron-sulfur cluster assembly 1
59758935	59782245	*Gm49354*	predicted gene, 49354
59771876	59823169	*Tut7*	terminal uridylyl transferase 7
60174405	60177365	*Gas1*	growth arrest specific 1
60331449	60333751	*Gm10779*	predicted gene 10779
60601947	60763191	*Dapk1*	death associated protein kinase 1
60781887	60786364	*BC051665*	cDNA sequence BC051665
60798250	60802844	*Ctsll3*	cathepsin L-like 3
60842612	60864416	*4930486L24Rik*	RIKEN cDNA 4930486L24 gene
60842621	60896894	*Gm49392*	predicted gene, 49392
60853230	60935590	*Gm49393*	predicted gene, 49393
60895350	60897447	*Ctla2b*	cytotoxic T lymphocyte-associated protein 2 beta
60895350	60936571	*Gm49391*	predicted gene, 49391
60901295	60904847	*Tpbpb*	trophoblast specific protein beta
60934155	60936625	*Ctla2a*	cytotoxic T lymphocyte-associated protein 2 alpha
60938692	60941935	*Tpbpa*	trophoblast specific protein alpha
61000277	61005958	*Ctsj*	cathepsin J
61035027	61040631	*Ctsq*	cathepsin Q
61159215	61164188	*Ctsr*	cathepsin R
61195132	61203410	*Cts6*	cathepsin 6
61246747	61255348	*Cts8*	cathepsin 8
61247878	61288839	*Gm49398*	predicted gene, 49398
61352461	61358197	*Cts7*	cathepsin 7
61535743	61541839	*Ctsm*	cathepsin M
61538390	61568174	*Gm49352*	predicted gene, 49352
61564630	61570127	*Cts3*	cathepsin 3
61630023	61636282	*Gm21933*	predicted gene, 21933
62015605	62050986	*Gm3325*	predicted gene 3325
62129916	62173419	*Zfp808*	zinc finger protein 808
62368328	62383177	*Gm3604*	predicted gene 3604
62453016	62466812	*Zfp935*	zinc finger protein 935
62453016	62466832	*Gm49359*	predicted gene, 49359
62516795	62579320	*Zfp934*	zinc finger protein 934
62603013	62624182	*6720489N17Rik*	RIKEN cDNA 6720489N17 gene
62671784	62706256	*Gm45951*	predicted gene, 45951
62772200	62785808	*Gm5141*	predicted gene 5141
62836885	62858370	*Fbp2*	fructose bisphosphatase 2
62864753	62888282	*Fbp1*	fructose bisphosphatase 1
63282142	63326096	*2010111I01Rik*	RIKEN cDNA 2010111I01 gene
63304709	63497278	*Fancc*	Fanconi anemia, complementation group C
63508328	63573598	*Ptch1*	patched 1
63815240	63900302	*Ercc6l2*	excision repair cross-complementing rodent repair deficiency, complementation group 6 like 2

Start	End	Gene Symbol	Name
64058266	64089230	*Hsd17b3*	hydroxysteroid (17-beta) dehydrogenase 3
64096308	64129368	*Slc35d2*	solute carrier family 35, member D2
64133022	64153202	*Zfp367*	zinc finger protein 367
64161866	64186537	*Habp4*	hyaluronic acid binding protein 4
64192545	64274988	*Cdc14b*	CDC14 cell division cycle 14B
64291836	64312710	*Aaed1*	AhpC/TSA antioxidant enzyme domain containing 1
64361890	64370748	*Ctsl*	cathepsin L
64432319	64439722	*Cdk20*	cyclin-dependent kinase 20
64481246	64497792	*Fam240b*	family with sequence similarity 240 member B
64737591	64904029	*Cntnap3*	contactin associated protein-like 3
64917406	64923184	*Spata31*	spermatogenesis associated 31
65033058	65038004	*Spata31d1c*	spermatogenesis associated 31 subfamily D, member 1C
65044613	65052760	*Prss47*	protease, serine 47
65064663	65112975	*Mfsd14b*	major facilitator superfamily domain containing 14B
65152226	65153152	*Olfr466*	olfactory receptor 466
65177111	65205716	*Nlrp4f*	NLR family, pyrin domain containing 4F
65259539	65260826	*Gm10775*	predicted gene 10775
65278814	65304221	*Zfp369*	zinc finger protein 369
65396483	65415122	*Gm31714*	predicted gene, 31714
65512678	65512848	*Gm10139*	predicted gene 10139
65689613	65713454	*Gm19428*	predicted gene, 19428
65897183	65907988	*Gm17353*	predicted gene, 17353
65950589	65958927	*Gm46425*	predicted gene, 46425
66062452	66062976	*Gm3844*	predicted gene 3844
66119547	66122836	*Gm10324*	predicted gene 10324
66163779	66165062	*Gm45965*	predicted gene, 45965
66227319	66227573	*Gm10772*	predicted gene 10772
66281716	66284053	*Gm46426*	predicted gene, 46426
66418114	66441118	*2410141K09Rik*	RIKEN cDNA 2410141K09 gene
66527953	66529383	*Gm46427*	predicted gene, 46427
66589169	66598733	*Gm45966*	predicted gene, 45966
66668946	66680803	*Zfp640*	zinc finger protein 640
66792583	66802563	*Gm7896*	predicted gene 7869
66849651	66854871	*Gm10323*	predicted gene 10323
66900617	66905378	*Uqcrb*	ubiquinol-cytochrome c reductase binding protein
66904914	66909236	*Gm10767*	predicted gene 10767
66906968	66933088	*Mterf3*	mitochondrial transcription termination factor 3
66932830	66998401	*Ptdss1*	phosphatidylserine synthase 1
67038596	67061170	*Zfp712*	zinc finger protein 712
67069398	67081545	*Zfp708*	zinc finger protein 708
67069399	67108604	*Gm28557*	predicted gene 28557

Start	End	Gene Symbol	Name
67096613	67116263	*Rslcan18*	regulator of sex-limitation candidate 18
67125124	67142404	*Zfp759*	zinc finger protein 759
67173207	67183126	*Rsl1*	regulator of sex limited protein 1
67194506	67209298	*Zfp455*	zinc finger protein 455
67254918	67278466	*Zfp458*	zinc finger protein 458
67292450	67306467	*Zfp457*	zinc finger protein 457
67312998	67332560	*Zfp595*	zinc finger protein 595
67317942	67360536	*Gm28044*	predicted gene, 28044
67339307	67360605	*Zfp953*	zinc finger protein 953
67345246	67375733	*Gm28041*	predicted gene, 28041
67362113	67375810	*Zfp456*	zinc finger protein 456
67389309	67399767	*Zfp429*	zinc finger protein 429
67405722	67421418	*Zfp459*	zinc finger protein 459
67424549	67451624	*Zfp874a*	zinc finger protein 874a
67471514	67484225	*Zfp874b*	zinc finger protein 874b
67490167	67500555	*Zfp58*	zinc finger protein 58
67491152	67526151	*Gm26965*	predicted gene, 26965
67515781	67526177	*Zfp87*	zinc finger protein 87
67538641	67553138	*Zfp748*	zinc finger protein 748
67589443	67609707	*Zfp729b*	zinc finger protein 729b
67592159	67637769	*Gm49345*	predicted gene, 49345
67617001	67637791	*Zfp729a*	zinc finger protein 729a
67667437	67687071	*Zfp738*	zinc finger protein 738
67705306	67729119	*Zfp65*	zinc finger protein 65
67747801	67755228	*Zfp85*	zinc finger protein 85
67779754	67792512	*Zfp493*	zinc finger protein 493
67813740	67827002	*Zfp273*	zinc finger protein 273
67831106	67844035	*Gm10037*	predicted gene 10037
67863326	67863923	*BC048507*	cDNA sequence BC048507
67885454	67889982	*Gm40991*	predicted gene, 40991
68444523	68447939	*Gm46386*	predicted gene, 46386
68560780	68582149	*Mtrr*	5-methyltetrahydrofolate-homocysteine methyltransferase reductase
68582234	68592338	*Fastkd3*	FAST kinase domains 3
68597439	68614231	*1700001L19Rik*	RIKEN cDNA 1700001L19 gene
68620043	68999541	*Adcy2*	adenylate cyclase 2
69497959	69534617	*Tent4a*	terminal nucleotidyltransferase 4A
69533746	69635780	*Nsun2*	NOL1/NOP2/Sun domain family member 2
69534549	69535451	*A930018O16Rik*	RIKEN cDNA A930018O16 gene
69573449	69611442	*Srd5a1*	steroid 5 alpha-reductase 1
69702832	69739889	*Ube2ql1*	ubiquitin-conjugating enzyme E2Q family-like 1
69802395	69816104	*Med10*	mediator complex subunit 10
70588689	70637840	*Ice1*	interactor of little elongation complex ELL subunit 1

Start	End	Gene Symbol	Name
70727802	70841811	*Adamts16*	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 16
71957921	71963723	*Irx1*	Iroquois homeobox 1
72628820	72634198	*Irx2*	Iroquois homeobox 2
73260479	73269608	*Irx4*	Iroquois homeobox 4
73319874	73328542	*Ndufs6*	NADH:ubiquinone oxidoreductase core subunit S6
73331009	73332178	*Mrpl36*	mitochondrial ribosomal protein L36
73467197	73516422	*Lpcat1*	lysophosphatidylcholine acyltransferase 1
73536747	73578672	*Slc6a3*	solute carrier family 6 (neurotransmitter transporter, dopamine), member 3
73604006	73620605	*Clptm1l*	CLPTM1-like
73626911	73649843	*Tert*	telomerase reverse transcriptase
73661752	73678023	*Slc6a18*	solute carrier family 6 (neurotransmitter transporter), member 18
73679745	73704865	*Slc6a19*	solute carrier family 6 (neurotransmitter transporter), member 19
73733094	73816754	*Slc12a7*	solute carrier family 12, member 7
73818534	73847631	*Nkd2*	naked cuticle 2
73911347	73937767	*Trip13*	thyroid hormone receptor interactor 13
73937811	73960895	*Brd9*	bromodomain containing 9
73963847	73992985	*Zdhhc11*	zinc finger, DHHC domain containing 11
74009407	74035753	*Tppp*	tubulin polymerization promoting protein
74036500	74062299	*Cep72*	centrosomal protein 72
74121457	74169442	*Slc9a3*	solute carrier family 9 (sodium/hydrogen exchanger), member 3
74169802	74208729	*Exoc3*	exocyst complex component 3
74211118	74292331	*Ahrr*	aryl-hydrocarbon receptor repressor
74303509	74317580	*Pdcd6*	programmed cell death 6
74322254	74350280	*Sdha*	succinate dehydrogenase complex, subunit A, flavoprotein (Fp)
74350309	74365783	*Ccdc127*	coiled-coil domain containing 127
74359578	74364005	*Lrrc14b*	leucine rich repeat containing 14B
74369326	74390836	*Zfp72*	zinc finger protein 72
74480054	74493970	*Zfp825*	zinc finger protein 825
74639624	74692537	*Erap1*	endoplasmic reticulum aminopeptidase 1
74692368	74808810	*Cast*	calpastatin
75089826	75132498	*Pcsk1*	proprotein convertase subtilisin/kexin type 1