Analysis of Genes Involved in Gonadal Development: Identification of Novel Sex Determination Candidates

Frances Wong

Presented for the Degree of PhD

University of Edinburgh

2004

Declaration

The experiments described in this thesis were the unaided work of the author except where acknowledgement is made by reference. No part of this work has previously been accepted for any other degree, nor is any part of it being submitted concurrently in candidature for another degree.

Frances Wong, 2004

Acknowledgements

I would like to thank my supervisors, Andreas Schedl and Nick Hastie. I am truly grateful for their guidance and support throughout my PhD, especially for their advice and comments whilst putting this thesis together. Special thanks to Aswin Menke for all his encouragement and continued support, which has been invaluable and very much appreciated.

I would also like to thank Marie-Christine Chaboissier and Simon Dixon for helpful discussions with my real-time PCR work, along with Peter Hohenstein for discussions with my Affymetrix data and siRNA work. A sincere thanks to Kevin Robertson for his generous time and patience in assisting with the GeneSpring software, which produced the beautiful scatter plots and heat maps! Many thanks also to Rob Watson for taking the challenge and demonstrating how "easy" it is to section tiny pieces of tissue! I am grateful to Paul Perry and Patrick Garner for their help with microscopy-related things when I needed it. Not forgetting Lee, as someone who has recently "been there and done it", a big thank you for volunteering to read chapters of this thesis – something which I'm sure you later regretted!

Thanks to all the people I have worked with over the past few years for all their contributions. Many of you have not only been "members of the lab or other labs", but are also friends whom I hope will stay in touch. You have all kept me sane and given me strength when the going was tough. There are simply too many of you to name, so I hope you will forgive me – and I'm sure you know whom you all are!

I owe a big thank you to all my friends outside of the lab for being there for me, trying to understand my crazy lab hours and making sure I was not working 24 hours a day!

Finally, I would like to thank my Dad and my Gran for their unwavering support every step of the way.

Abstract

Genetic and developmental studies have shown that upon the presence or absence of the Y-linked sex-determining gene *SRY*, the bipotential gonad will develop either as a testis or ovary. Since the discovery of *SRY*, several other genes (such as *WT1*, *SF1*, *DAX1* and *SOX9*) have been isolated which play an important role in gonadal development and sex determination. Despite these advances our understanding of the mammalian sex determination process remains incomplete.

To identify new genes involved in gonadal development, a differential screen using Affymetrix GeneChip arrays was performed. Wild-type male and female gonads were dissected from mouse embryos at E12.5 (during the differentiation process) and subjected to microarray analysis. From this extensive analysis, several novel transcripts were identified, which show a sex-specific expression pattern.

The validity of this approach was verified by analysing the expression of these novel transcripts by *in-situ* hybridisation in male and female gonads at E12.5. Transcripts with a confirmed sex-specific expression pattern where then selected for detailed *in-situ* hybridisation of male and female gonads at E11.5, E13.5 and E14.5. Moreover, real-time PCR of these candidates was carried out to establish a sex-specific gene expression profile during gonadal development, ranging from 17 to 36 tail somites.

This thesis also describes the novel approach of applying small interfering RNAs (siRNAs) to a gonadal organ culture system. Several known sex determination genes were successfully targeted and knocked-down by applying this technique. Importantly, siRNA against *Sry* effectively blocked male gonad differentiation resulting in a lack of expression of male specific markers. Taken together these results suggest that the siRNA approach in gonad cultures can be used to efficiently analyse the function of candidate genes isolated in our sex determination screen.

Contents

List of Figures

Chapter 4

Chapter 5

Chapter 6

List of Tables

Abbreviations

°C	degrees centigrade
A	absorbance
AMH	Anti-Müllerian Hormone
β-gal	β-galactosidase
bp	base pairs
BSA	bovine serum albumin
CBLN1	Cerebellin 1 Precursor Protein
cDNA	complementary deoxyribonucleic acid
Cre	Cre Recombinase
dATP	2'-deoxyadenosine 5'-triphosphate
DAX1	Dosage-Sensitive Sex-Reversal Adrenal Hypoplasia Congentia Critical Region on the X Chromosome
dCTP	2'-deoxyctosine 5'-triphosphate
DEPC	diethyl pyrocarbonate
dGTP	2'-deoxyguanosine 5'-triphosphate
dH_2O	distilled water
DHH	Desert Hedgehog
DIG	Digoxygenin
DMEM	Dulbecco's MEM
DMRT1	*Double-sex* and Mab-3 Related Transcription Factor 1
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
dpc	days post coitum
ds	double stranded
dTTP	2'-deoxythymidine 5'-triphosphate
DTT	dithiothreitol
E	embryonic day
EDTA	ethylenediaminetetra-acetic acid di-sodium salt
EMX2	Evenskipped Homologue Gene 2
EST	expressed sequence tag
FCS	foetal calf serum
FGF9	Fibroblast Growth Factor 9
FOG2	Friend of Gata-2
g	gram
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GATA4	GATA binding protein 4
HCL	Hydrochloric acid
KOAc	potassium acetate
l	litres
lacZ	β-galactosidase gene
LHX9	Lim-type Homeobox Containing Gene 9
μ	prefix *micro*
m	prefix *milli*
M	molar

M33	Mouse Homologue of Polycomb
MAS	Microarray Suite
$MgCl_2$	magnesium chloride
MgOAc	magnesium acetate
MIS	Müllerian Inhibiting Substance
mRNA	messanger ribonucleic acid
n	prefix *nano*
NaCl	sodium chloride
NaOAc	sodium acetate
NaOH	sodium hydroxide
PAX2	Paired Box Gene 2
PAX6	Paired Box Gene 6
PBS	phosphate buffered saline
PBT	phosphate buffered saline and Tween
PCR	polymerase chain reaction
PFA	paraformaldehyde
PN-1	Protease Nexin-1
POD-1	Podocyte Expressed 1
PTC	Patched Homologue
RNA	ribonucleic acid
RNAi	RNA interference
RNA Pol II	RNA Polymerase II
rpm	revolutions per minute
RT-PCR	reverse-transcriptase polymerase chain reaction
SDS	sodium dodecyl (lauryl) sulphate
SF1	Steroidogenic Factor 1
siRNA	Small interfering RNA
SMOC2	Secreted Modular Calcium-binding Protein 2
SOX8	Sry-box Gene 8
SOX9	Sry-box Gene 9
SRY	Sex Determining Region of the Y Chromosome
ss	single stranded
SSC	saline sodium citrate
TBE	Tris, boric acid, EDTA
TBST	Tris Buffered Saline and Tween
TE	Tris, EDTA
TPD52L1	Tumour Protein D52-like 1
Tris	tris hydroxymethyl aminomethane
UV	Ultraviolet
VNN1	Vanin-1
WNT4	Wingless-Type MMTV Integration Site Family, Member 4
WT1	Wilms' Tumour Suppressor 1
X-gal	5-bromo-4-chloro-3-indolyl-β-$_D$-galactosidase
ZFY	Zinc Finger Protein, Y-linked

Chapter 1 Introduction

Chapter 1 Introduction

1.1 Sex Determination and Sex Chromosomes

How an individual's sex is determined has been one of the intriguing questions of life for many centuries. In 355 BC, Aristotle claimed that sex was determined by the heat of the semen at the time of copulation: hot semen generated males, whereas cold semen produced females. Since then, greater knowledge of the molecular events of sex determination has been established. Mammalian sex determination is now known to be controlled by genes, which are expressed in a sexually dimorphic manner. In humans and other mammals, the presence of either a second X chromosome or a Y chromosome determines whether the embryo is to be female (XX) or male (XY). However, in birds, the situation is reversed: males have two similar sex chromosomes (ZZ), while females have the unmatched pair (ZW). Interestingly, the gonad is the only organ that can develop into two different organs: either a testis or an ovary. Hence, the path of differentiation taken by this organ determines the future sexual development of the organism.

1.2 Development of the Gonad

Mammalian gonads are formed by proliferation of the coelomic epithelium and condensation of the underlying mesenchyme (Sadler 1995). Initially, these gonads are indifferent or bipotential as they cannot be distinguished to be either testes or ovaries (Gubbay et al., 1990; Sinclair et al., 1990). In humans, gonadal rudiments appear in the intermediate mesoderm during week 4 and remain sexually indifferent until week 7 (Francavilla et al., 1990). These gonadal rudiments form as paired regions adjacent to the developing kidneys and the ventral parts constitute the genital ridge epithelium. In the mouse, this formation begins at approximately E10.5 (Figure 1.1) and the gonads appear as a pair of longitudinal ridges on the ventromedial sides of the mesonephroi, situated on either side of the gut mesentery

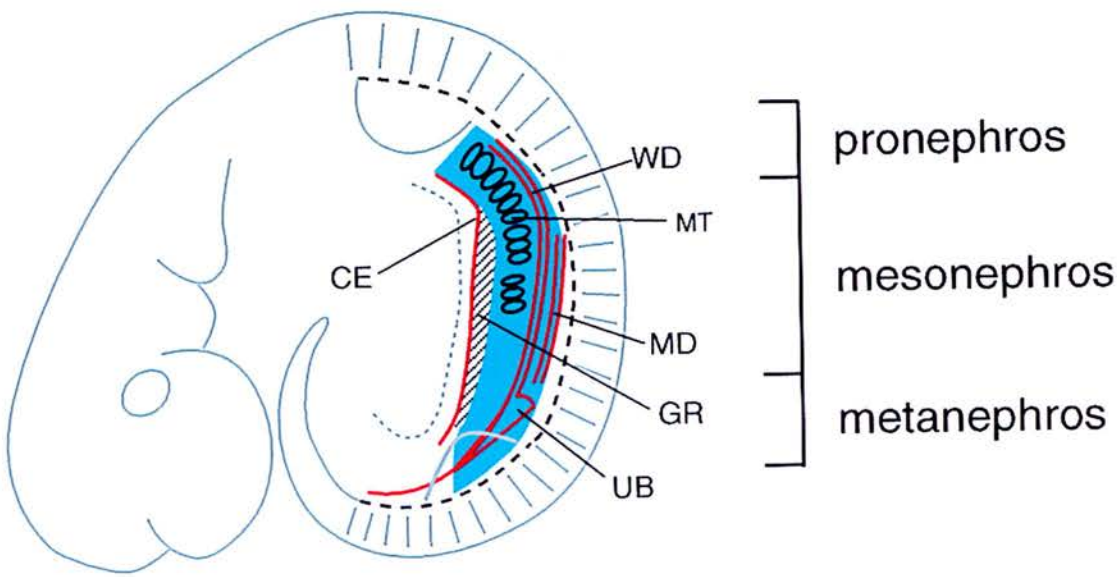


Figure 1.1 Structure of the Urogential System at E10.5

Schematic diagram illustrating the mouse urogenital system at E10.5. Epithelial structures are shown in red and mesenchymal structures are shown in blue. The striped region depicts the genital ridge (GR). The Wolffian duct (WD), mesonephric tubules (MT), Müllerian duct (MD), ureteric bud (UB) and coelomic epithelia (CE) are also shown (Figure adapted from Swain and Lovell-Badge 1999).

(Karl and Capel, 1998). Interestingly, testes and ovaries have corresponding cell types which display similar functions in reproduction. For example, Leydig cells (in testes) and theca cells (in ovaries) are involved in steroidogenesis, Sertoli cells and granulosa cells (from testes and ovaries, respectively) play an important role in germ cell maturation. In the mouse, primordial germ cells (PGC) migrate from the base of the allantois, via the gut mesentery, reaching the gonad between E10.5 and E11.5 (Ginsburg et al., 1990; Gomperts et al., 1994). At approximately E11.5, sexual differentiation between the testis and the ovary commences.

During the early stages of gonad differentiation, the size of the male gonad increases comparatively to the female gonad in mammals and in other vertebrates such as birds and reptiles (Mittwoch, 1989; Schmahl et al., 2003). In the mouse, the size increase of the XY gonad results from increased cell proliferation (Schmahl et al., 2000), which is concentrated in the coelomic epithelium (Karl and Capel, 1998). Throughout the early phase of proliferation, which occurs between E11.0 – E11.5, both Sertoli and interstitial cells are derived from the division of SF1-expressing cells in this layer. Subsequently, SF1 expression is lost from coelomic epithelial cells. As proliferation continues, only interstitial cells are produced. Furthermore, in a study where proliferation inhibitors were used to block cell division prior to the stages of the XY gonad committing to the testicular pathway, the inhibitors stopped proliferation of coelomic epithelial cells, blocked expression of the Sertoli cell marker, *Sox9*, and prevented testis cord formation (Schmahl and Capel, 2003). Moreover, the results revealed that proliferation within a specific 8-hour time frame around E11.0 (the early phase of proliferation) was a critical period for the establishment of the male pathway and the formation of the testis (Schmahl and Capel, 2003).

Following sexual differentiation, in mice, the gonad can be distinguished histologically by E13.5. Cells within the testes are arranged into two distinct compartments: testicular cords and the interstitial region. Testicular cords (precursors of the seminiferous tubules) are composed of primordial germ cells and

epithelialized Sertoli cells (Tilmann and Capel, 1999). The interstitial region, which surrounds the testicular cords, contains the steroidogenic Leydig cells and the peritubular myoid cells (a smooth muscle cell lineage forming the connective tissue of the gonads). In contrast, cells within the ovaries have an undefined "ground-glass" appearance and show little structural differentiation until the later stages of gestation (Parker et al., 2001).

In the mouse, Sertoli cells are the supporting cell lineage of the testis as they maintain the growth and maturation of germ cells. They are also essential for the early stages of testis development. Analysis of cells from the testes of chimeric XX $\leftrightarrow$ XY embryos revealed that Sertoli cells were the only cell type to show a strong bias for the Y chromosome. Approximately 90% of Sertoli cells were XY, whereas other testicular cell types were XX or XY with equal frequency (Palmer and Burgoyne, 1991). This genetic evidence suggested that expression of genes on the Y chromosome (including *Sry*) might only be required in the Sertoli-cell lineage. However, it appears the requirement for *Sry* is not absolute as approximately 10% of Sertoli cells in the adult chimeras were XX (had no Y chromosome or *Sry* gene), which suggests that a downstream paracrine factor could be involved in the recruitment of non-*Sry*-expresssing cells to the Sertoli pathway.

Migration of mesenchymal cells from the mesonephros into the adjacent XY gonad is also important for testicular development and cord formation in mice. Recombinant gonadal organ culture experiments have shown that the recruitment of cells from the mesonephros is male-specific (Martineau et al., 1997) and *in vitro* culture experiments using gonads that normally form ovotestes have shown that migration is absent from ovarian regions and is confined to regions where testis cords develop (Albrecht et al., 2000). In addition, cord formation was found to be impaired when migration was blocked at E11.5 by culturing the gonad without attached mesonephroi (Buehr et al., 1993) and by placing a barrier between the two tissues in culture (Tilmann and Capel, 1999). Furthermore, culturing of an XY gonad on the surface of an XX gonad caused the migration of cells from the

mesonephros to be induced in the XX gonad between E11.5 – E13.5, which resulted in the formation of cord-like structures and expression of male markers in some samples (Tilmann and Capel, 1999).

Vascularization also contributes to the morphogenesis of the XY mouse gonad. At the early bipotential stages of gonad development, vasculature growth from the mesonephros into the gonad is identical in XX and XY gonads (Brennan et al., 2002; Bullejos et al., 2002). However, following *Sry* expression, additional endothelial cells are only recruited to the XY gonad. These migrating cells then assemble to generate the coelomic vessel, an arterial system denoted by the arterial maker ephrin-B2 and elements of the Notch signalling pathway (Brennan et al., 2002). This arterial system in the early mouse testis establishes an alternative pattern of blood flow that is diverted from the mesonephros through the coelomic vessel. It has been suggested that the arterial system increases blood flow through the testis in order to promote the efficient export of testosterone from the early testis, as the delivery of testosterone is essential for establishment of male characteristics in sex determination (Brennan et al., 2002).

Testosterone is produced by Leydig cells in the testis. The main function of Leydig cells is steroidogenesis. There are two distinct types of Leydig cells in mice: foetal and adult lineages (Habert et al., 2001). Foetal Leydig cells are present in the testis from E12.5 until their numbers decrease shortly after birth. The role of these foetal Leydig cells is the masculinization of the male urogenital system. More specifically, foetal Leydig cells produce testosterone to virilize the internal and external genitalia. Adult Leydig cells appear during puberty and testosterone produced on this occasion is required for the onset of spermatogenesis and maintenance of male reproductive function (Habert et al., 2001). However, it should be noted that foetal Leydig cells are more active than adult Leydig cells in testosterone production.

Peritubular myoid cells are believed to be amongst the cells that originate from the mesonephros (Martineau et al., 1997; Tilmann and Capel, 1999), which subsequently

migrate into the developing mouse testis. These cells are found within the interstitial region, surrounding and enclosing the Sertoli and germ cells, creating discrete testis cords. Once peritubular myoid and Sertoli cells associate, they secrete a matrix that forms the basal lamina, which separates the testis cords from the interstitial compartment (Tung et al., 1984). Disruption or delay of testis cord formation may cause gonadal dysgenesis or infertility. A significant deficiency of peritubular myoid cells, along with disorganised and incomplete formation of testis cords, was observed in *Dax1* null mice. This indicates the important role that *Dax1* plays in testis differentiation by regulating the development of peritubular myoid cells and the formation of intact testis cords (Meeks et al., 2003a).

The internal genitalia (derived from the genitourinary tract) are initially the same in both male and female mouse embryos, where they have two identical sets of paired ducts: the Müllerian (paramesonephric) ducts and the Wolffian (mesonephric) ducts (Parker et al., 2001). Upon Y chromosome activation of the male developmental pathway, the gonads develop into testes (Figure 1.2), whereby Sertoli cells within the testicular cords produce a glycoprotein hormone, Müllerian inhibiting substance (MIS) (Tran et al., 1977). This hormone mediates the regression of the Müllerian ducts (Jost, 1953). In parallel, Leydig cells differentiate within the interstitial region and synthesize testicular androgens, which result in stabilisation and differentiation of the Wolffian ducts and formation of the external genitalia. In the absence of testicular hormones, the Wolffian ducts regress and the Müllerian ducts form the oviducts, Fallopian tubes, uterus and upper vagina giving rise to a female phenotype (Figure 1.2) (Jost et al., 1973).

Similar to the internal genitalia, the external genitalia also possess structures that are initially found in both sexes. These include the genital tubercle, the urethral groove and the labial (or labio-scrotal) folds. In the female mouse, at approximately E14.5, the genital tubercle differentiates into the clitoris and glans clitoridis (Kaufman, 1999). While, a stage later, the urethra forms from the urethral groove at E15.5. In addition, the labial folds do not fuse across the ventral midline but form the labia

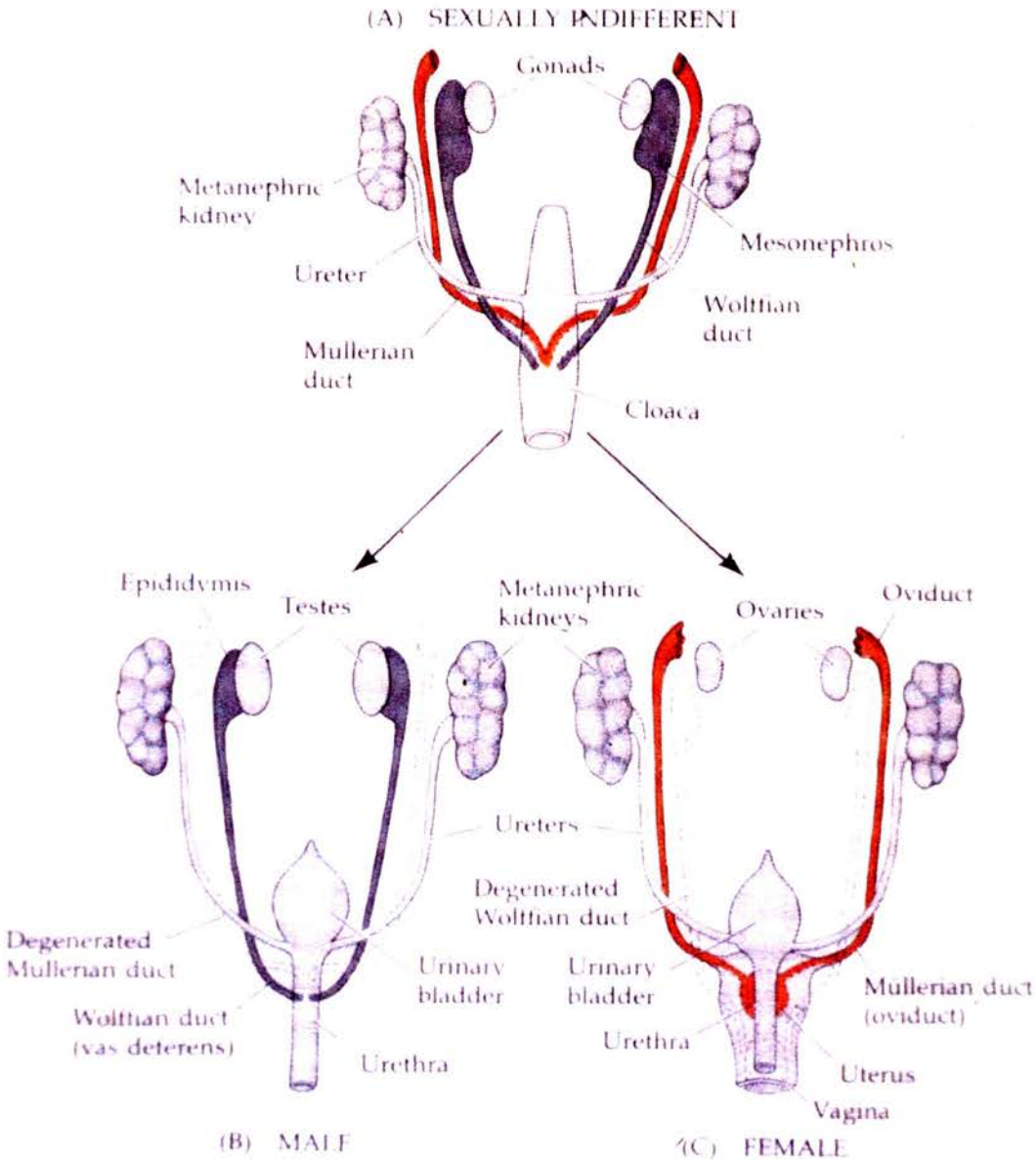


Figure 1.2 Structure of Urogenital System

Schematic diagram of the urogenital system, where (A) represents bipotential/indifferent gonads, (B) illustrates the development of the gonad into testes and (C) the development into ovaries. (Figure adapted from Gilbert, 2000).

majora, whereas the labia minora results from the lack of fusion of the urethral folds. In the male mouse, the glans penis differentiates from the distal part of the genital tubercle at approximately E14.5, the urethral groove forms the penile urethra and the labial folds fuse across the midline in the region of the cloacal membrane to form the scrotum (Kaufman, 1999).

1.3 Genes Involved in Formation and Survival of the Gonad

Over the last decade, numerous studies have been conducted in order to gain a better understanding of the molecular mechanisms underlying the formation of the gonadal primordium. However, to date, very little is known. So far, only four genes have been discovered to be critical in this process. These are: the Wilms' Tumour suppressor gene (*Wt1*), the steroidogenic factor gene (*Sf1*), the Lim-type homeobox containing gene (*Lhx9*) and the evenskipped homologue gene (*Emx2*). It is apparent that all of these genes are essential for the development and survival of the bipotential gonad, as knockout mutations of each of these genes result in mice lacking gonadal tissues.

1.3.1 Wilms' Tumour Suppressor Gene (*Wt1*)

Wilms' tumour, or nephroblastoma, is a paediatric kidney cancer that originates from pluripotent embryonic renal precursors (Bennington, 1975). Positional cloning strategies led to the identification of the *WT1* tumour suppressor gene at chromosome 11p13 (Call et al., 1990; Gessler et al., 1990). The *WT1* gene is highly complex and encodes 24 different protein isoforms through a combination of alternative splicing, RNA editing and differential initiation of translation (Little et al., 1999). All the WT1 proteins contain four zinc fingers of the Kruppel type at the C-terminus. In +KTS isoforms, zinc fingers 3 and 4 are separated by three amino acids which affect their ability to bind DNA (Larsson et al., 1995). These proteins co-localise with

nuclear speckles and splicing factors (Larsson et al., 1995). In contrast, –KTS isoforms show a predominant diffuse nuclear localisation typical of transcription factors (Larsson et al., 1995). This suggests that the different WT1 isoforms may have different biochemical functions (Little et al., 1999).

In humans, individuals affected by Frasier syndrome carry a point mutation on one copy of *WT1*, preventing the production of the +KTS isoform from that allele. As a consequence, the ratio between the +KTS and –KTS isoform is altered (Klamt et al., 1998). In our laboratory, animal models for Frasier syndrome have been generated. XY mice homozygous for the Frasier mutations are completely sex reversed which suggests that *Wt1* is required for sex determination (Hammes et al., 2001). Interestingly, it has been discovered that the +KTS and –KTS isoforms of WT1 have distinct functions in gonad formation and sex determination. –KTS isoforms are necessary for the survival of the gonadal primordium and mice ablated for the –KTS isoform show increased apoptosis (Hammes et al., 2001). Absence or reduced levels of +KTS isoforms, as found in human Frasier patients, show a decrease in *Sry* expression levels and, consequently, *Sox9* and *MIS* fail to be activated (Hammes et al., 2001). In addition transgenic mice genetically depleted of *Wt1* do not develop kidneys, gonads, spleen and adrenals (Kreidberg et al., 1993) indicating that *Wt1* is essential for the formation of these organs.

In the developing mouse, expression of *Wt1* first appears in the urogenital ridge at E9.0 and continues to be expressed after formation of the gonads, particularly in the male sex cords (structures that contain the precursors of Sertoli cells) (Shimamura et al., 1997). After birth, *Wt1* is expressed in the testis, specifically in Sertoli cells, but not in Leydig or germ cells (Shimamura et al., 1997). In addition, Wolffian and Müllerian ducts do not express *Wt1* either. *Wt1*-null mice die at an early stage of embryonic development, E13.5, due to malformations of the heart and diaphragm (Moore et al., 1999). The gonadal anlagen in *Wt1* knockout mice undergoes apoptosis, hence *Wt1*-null embryos lack gonads. This demonstrates that early expression of *Wt1* is necessary for development of the undifferentiated gonad,

although migration of the primordial germ cells remains unaffected (Kreidberg et al., 1993). Recent studies have detected *Wt1* expression in embryonic germ cells beginning at E11.5, following the migration of primordial germ cells into the gonad (Natoli et al., 2004). Furthermore, germ cells isolated from *Wt1*-null embryos show impaired growth in culture (Natoli et al., 2004). These data have led to the discovery of a previously unknown role for *Wt1* in gonadal germ cell proliferation or survival.

1.3.2 Steroidogenic Factor Gene (*Sf1*)

The orphan nuclear receptor steroidogenic factor 1 (*SF1*) plays an important role in the gonadal and adrenal development pathways, and is known to regulate the expression of enzymes involved in steroid biosynthesis (Reviewed in Parker and Schimmer 1997). In mice, *Sf1* is expressed between E9.0 and E12.0 in the gonads of both sexes (Ikeda et al., 1994). However, its expression is then down regulated in developing ovaries (Ikeda et al., 1994). Similar to *Wt1* knockout mice, the gonadal anlagen in *Sf1* knockout mice also seem to undergo apoptosis in both sexes, although the initial stages of genital ridge formation are unaffected (Luo et al., 1994). Thus, the gonads of $Sf1^{-/-}$ embryos do not develop beyond the early indifferent stage and XY *Sf1* mutant embryos lack testes and therefore develop a female urogenital system (Luo et al., 1994).

In humans, a heterozygous 2-bp G35E (GGC→GAA) mutation in exon 3, which encodes part of the DNA-binding domain of SF1, has been shown to cause complete XY sex reversal (including normal female external genitalia and retention of the uterus) and adrenal failure (Achermann et al., 1999). Interestingly, $Sf1^{+/-}$ mice do not display the XY intersex phenotype described in the heterozygous SF1 human patients (Bland et al., 2000). Evidence that SF1 regulates the regression of Müllerian structures, either by directly acting on the *MIS* promoter or through an abnormality of Sertoli cell development or function, has also been found (Achermann et al., 1999). Interestingly, site-directed mutagenesis studies in mice revealed that this

heterozygous 2-bp G35E (GGC→GAA) mutation found in humans only disrupts the binding site of SF1; it does not effect protein translation, stability or nuclear localisation (Giuili et al., 1997).

Recently, in mice, the specific *in vivo* roles of SF1 in discrete gonadal cell lineages was explored using the *Cre-loxP* system to inactivate SF1 in a cell-selective manner at a relatively early stage of gonadal development (Jeyasuria et al., 2004). The results strongly indicate that SF1 in Leydig cells is essential for expression of the cholesterol side-chain cleavage enzyme (Cyp11a), which catalyses the first reaction in testosterone production, and the steroidogenic acute regulatory protein (StAR), a second essential component of steroidogenesis. Thus, providing the first evidence that SF1 is essential for the expression of these target genes *in vivo* (Jeyasuria et al., 2004). Furthermore, the results also indicate important roles of SF1 in estrogen production by granulosa cells *in vivo* (Jeyasuria et al., 2004).

Taken together, these data suggest that SF1 has at least three important functions in sexual development. Firstly, in stabilisation and development of the genital ridge, secondly, in formation of the testis, and finally, in regulation of *MIS* and also in regulation of steroid synthesis in Leydig and granulosa cells (Koopman, 1999).

1.3.3 Lim-type Homeobox Containing Gene 9 (*Lhx9*)

Lim-type Homeobox Containing Gene 9 (*Lhx9*), a member of the LIM homeobox domain gene family of transcription factors, has been mapped to mouse chromosome 1 (Failli et al., 2000). Analysis in the gonadal region found *Lhx9* transcripts to be expressed in the medial surface of the urogenital ridge from 9.5 dpc in both sexes (Birk et al., 2000). At E13.5, *Lhx9* is differentially expressed. In males, the expression pattern of *Lhx9* is high in the mesothelial layer and the outer part of the testis (the developing tunica), lower in the interstitial mesenchyme, and almost absent in the sex cords (Birk et al., 2000). In females, *Lhx9* expression is mainly

restricted to the cortical region (Birk et al., 2000). Although *Lhx9* is normally expressed in the embryonic central nervous system, limbs and pancreas, *Lhx9* knockout mice did not show any abnormalities in these structures (Birk et al., 2000). However, targeted disruption of *Lhx9* resulted in the failure of a discrete gonad to be formed (Birk et al., 2000). More specifically, $Lhx9^{-/-}$ and wild-type urogenital ridges were indistinguishable at E11.5. However, by E12.0, budding was apparent in the wild-type genital ridge, yet no such proliferation was seen in the *Lhx9* mutants. Furthermore, at E13.5, when distinct testes and ovaries can be observed in wild-type embryos, no gonads were visible in $Lhx9^{-/-}$ embryos (Birk et al., 2000). This developmental defect is thought to be caused by the failure of somatic cells in the genital ridge to proliferate, as alkaline phosphatase staining revealed the presence of primordial germ cells within E12.0 $Lhx9^{-/-}$ genital ridges, thus indicating that germ cell migration occurred at the expected time (Birk et al., 2000). In view of the fact that *Lhx9* mutants display a gonadal deficiency, it can therefore be concluded that *Lhx9* is vital for gonad formation.

1.3.4 Evenskipped Homologue Gene 2 *(Emx2)*

The homeobox gene *Emx2* is a mouse cognate of a *Drosophila* head gap gene, empty spiracles (*ems*), and is fundamental for the development of the dorsal telencephalon (Yoshida et al., 1997). Simultaneously, *Emx2* is expressed in the epithelial components of the developing urogenital system and *Emx2* transcripts are specifically observed in the developing gonad. Interestingly, kidneys, ureters, gonads and genital tracts are completely missing in *Emx2* mutant mice (Miyamoto et al., 1997). The ablation of the gonadal tissue is considered to be due to proliferative defects within the coelomic epithelium, as poor thickening of this was seen in E11.5 *Emx2* deficient embryos (Miyamoto et al., 1997). Consequently, *Emx2* may participate in proliferation, differentiation and/or survival of gonadal cells. However, further studies need to be conducted before its specific function in gonadal formation is elucidated.

1.4 Gonad Formation: Developmental Cascade

Evidently, it is important to discover new genes which are involved in the formation of the gonad, but it is also of equal importance to understand the functional role(s) and relationship(s) they have with each other. Considering the above genes are all expressed in the developing gonad at approximately the same time, one can speculate that they are likely to share a common pathway.

As the effects on gonadogenesis in *Wt1* knockout mice are remarkably similar to those observed in *Sf1* knockout mice (the gonads appear normal at early stages but regress through apoptosis as they reach a certain period of development), the possible interactions between *Wt1* and *Sf1* have been examined. Analysis of the genital ridges of *Wt1* knockout mice indicated that *Sf1* expression is completely absent (Wilhelm and Englert, 2002). This finding suggests that the expression of the *Sf1* gene in the early gonad is WT1-dependent. Furthermore, *Sf1* expression, which could first be detected at E9, was dependent on WT1 at this stage. This implies that WT1 is required for the initiation of *Sf1* expression. Interestingly, characterisation binding studies show only the –KTS form of WT1 is able to bind to and transactivate the *Sf1* promoter (Wilhelm and Englert, 2002). This observation is consistent with differential roles for the –KTS and +KTS variants of WT1, whereby the –KTS isoform is required for the survival of the gonadal primordium and the +KTS isoform seems primarily involved in the male sex determination pathway (Hammes et al., 2001).

Investigations into the possible interactions of *Lhx9* with *Wt1* and *Sf1* have also been conducted, by examining their expression in *Lhx9*-null embryos (Birk et al., 2000). No change in the expression pattern of *Wt1* was found. However, a reduced expression of *Sf1* in E11.5 *Lhx9* deficient urogenital ridges was observed. This suggests that WT1 is necessary but not sufficient to activate the expression of *Sf1*. Furthermore, taken together with the continuous expression of *Lhx9* in *Sf1* knockout embryos, these findings imply that *Lhx9* may also be upstream of *Sf1*. Consequently,

it has been found that *Wt1* and *Lhx9* activate the expression of *Sf1* in the developmental cascade leading to early gonadogenesis (Wilhelm and Englert, 2002).

In conclusion, it should be noted that deciphering part of a developmental cascade, and interpreting experimental data based on the impact of a gene knockout on the expression of other genes, is difficult. Direct and indirect effects of unknown factors may influence the results. An alternative approach to test for genetic interactions would be to generate double homozygous mutants by crossing heterozygous knockouts of the genes of interest. These mutants could potentially confirm previous results and provide insight into the phenotypic consequences of the absence of both genes.

1.5 Genes Involved In Sex Determination

1.5.1 Sex Determining Region of the Y Chromosome (*Sry*)

Genetic and developmental studies have shown that, depending on the presence or absence of the Y-linked sex-determining gene, *SRY* (Sex determining Region of the Y chromosome), the bipotential gonad will develop either as a testis or ovary (Figure 1.3). *SRY* plays a crucial role in male sex determination. It encodes a high-mobility group (HMG) box protein that can function as a DNA binding protein (Su and Lau, 1993). It is believed to act within cells of the supporting cell lineage of the indifferent gonad to stimulate a series of events resulting in differentiation of Sertoli rather than follicle cells (Hacker et al., 1995). Further studies have shown that *SRY* initiates a "remarkable" increase in vascularization of the gonad and on somatic cell proliferation at the coelomic epithelium of XY gonads, resulting in a thickening of this later (Schmahl et al., 2000).

In the mouse, *Sry* is expressed in a proportion of cells in the gonadal ridge, which represent the precursors of Sertoli cells in the developing testis only for a brief period

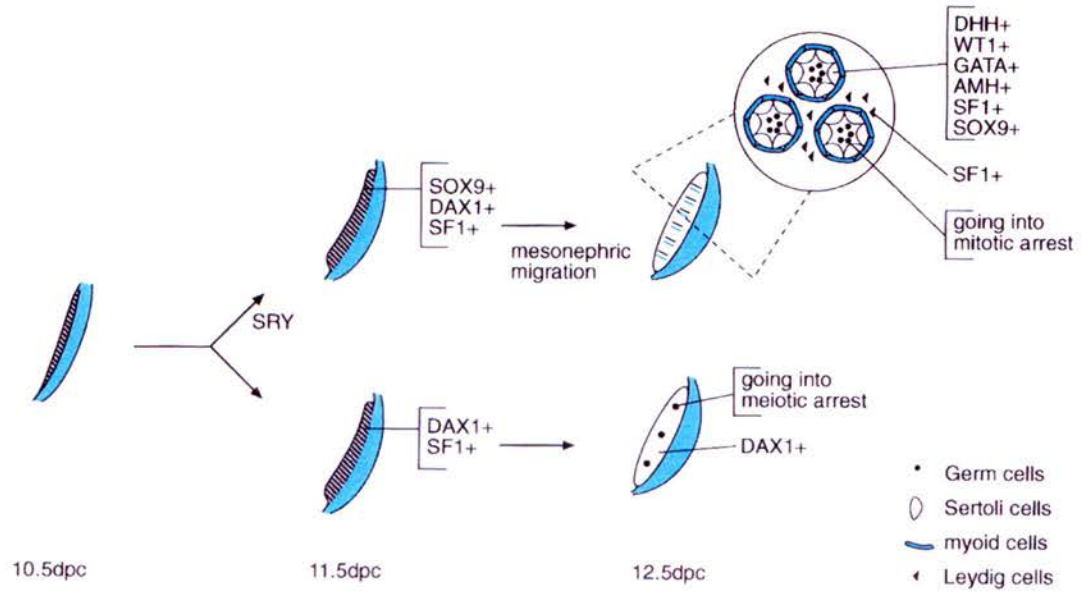


Figure 1.3 Early Development of the Gonad

Schematic diagram of the early developmental stages of the gonad, illustrating some of the essential genes involved and the cell types present. (Figure adapted from Swain and Lovell-Badge 1999).

(Lovell-Badge et al., 2002). Its expression in the genital ridge is wave-like from the central and/or anterior regions, which subsequently extends to both poles and ends in the caudal pole (Bullejos and Koopman, 2001). As previously mentioned, analysis of chimeric mice derived from XX and XY cells demonstrated that nearly all testicular Sertoli cells were XY while other cell lineages did not favour either chromosome (Palmer and Burgoyne, 1991). Therefore, the action of SRY is believed to initiate differentiation of the Sertoli cell lineage in the testis. Once SRY triggers Sertoli cells they in turn guide the differentiation of the other testicular cell types, resulting in structural organisation and formation of the testis cords (Koopman et al., 1991). Thus, the decision of sex determination is fundamentally one of cell fate: a cell, which would otherwise become a granulosa cell, becomes a Sertoli cell due to its activation by SRY.

In mice, transcription of *Sry* in the male (XY) urogenital ridge is switched on at E10.5, reaching a peak at E11.5, prior to overt differentiation, before being completely switched off by E13.0 (Hacker et al., 1995). The expression of *Sry* in pre-Sertoli cells (Albrecht and Eicher, 2001) initiates the differentiation of Sertoli cells, whereby at E12.5 Sertoli cells sequester primordial germ cells within testis cords. By E13.5, as a result of interactions with Sertoli cells, germ cells arrest in the G1 phase of the mitotic cell cycle and do not enter meiosis until well after birth (McLaren, 1988). In the absence of germ cells in the male gonad, the somatic cell lineages appear to differentiate normally (McLaren, 1991), although a number of abnormalities can be detected in empty testis tubules of adult mice. Despite the fact that germ cells are not required for the organisation of the testis cords, they may facilitate this process by responding to the masculinising somatic environment and become committed to spermatogenesis (Adams and McLaren, 2002). The expression of *Sry* also induces the migration of peritubular myoid cells from the mesonephros, which facilitates testis cord formation (Tung et al., 1984).

In the absence of *Sry* expression, ovarian fate occurs (Gubbay et al., 1992; Hawkins et al., 1992). In contrast to the XY gonad, germ cells are important for the formation

and maintenance of the structure of the ovary. In the absence of germ cells, follicle cells do not assemble and deteriorate rapidly (McLaren, 1988). It has been shown that when E14.5 XX somatic cells are recombined with XY somatic cells, testis cord structures form normally. Whereas, when XX germ cells are recombined with XY somatic cells, cord structures are disrupted (Yao et al., 2003). In the XX gonad, germ cells enter meiosis by E13.5 and arrest in prophase I by birth (McLaren, 1988). It has been suggested that once germ cells commit to meiosis (McLaren and Southee, 1997), they reinforce ovarian fate by antagonising the testis pathway (Yao et al., 2003).

Since the discovery of *SRY*, several other genes have been isolated which play a part in the process of gonadal development and sex determination (Figures 1.4 and 1.5). Most of them have been isolated from human patients affected by complete or partial sex reversal in addition to abnormalities in other organs. In fact, some of these genes are essential for normal kidney (*WT1* and *SF1*), adrenal (*DAX1* and *SF1*) and cartilage (*SOX9*) formation, as well as being required for sex determination.

1.5.2 *Sry-box* Gene 9 (*Sox9*)

Sox (*Sry-box*) genes are involved in many developmental processes. For example, chrondrogenesis, hematopoiesis, neurogenesis, lens development and sex determination (Reviewed in Wilson and Koopman 2002). The *Sox* gene family encodes proteins with at least 50% amino acid identity to the 79 amino acid HMG-type DNA-binding motif present in *Sry* (Bowles et al., 2000). The sequence-specific DNA binding and transactivation properties of SOX proteins indicate they function as transcription factors with characteristics of both traditional transcription factors and architectural chromatin factors (Reviewed in Wilson and Koopman 2002).

The *Sox9* gene is distinctly related to *Sry* (59% nucleotide identity within the HMG box) (Bowles et al., 2000). *SOX9* is autosomal, contains introns and has clearly been

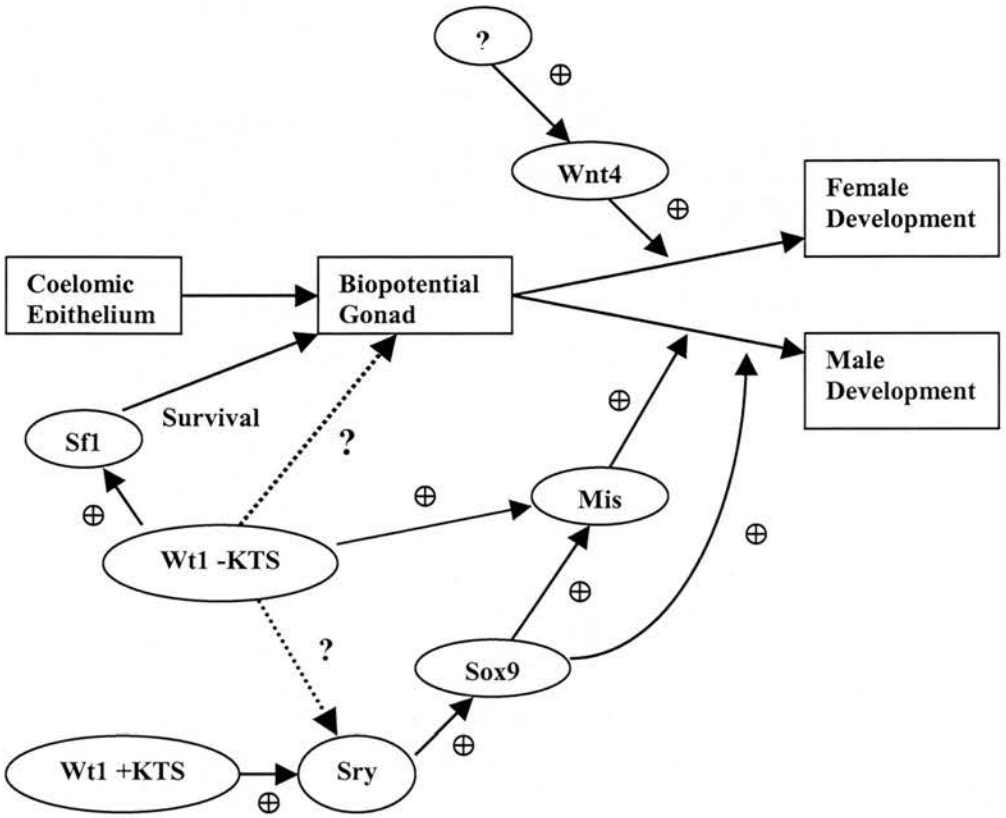


Figure 1.4 Genes Involved in Gonad Formation and Sex Determination

Schematic model of genes that are involved in gonad formation and sex determination, particularly illustrating the different functions of the Wt1 ± KTS isoforms (Figure modified from Guo et al., 2002).

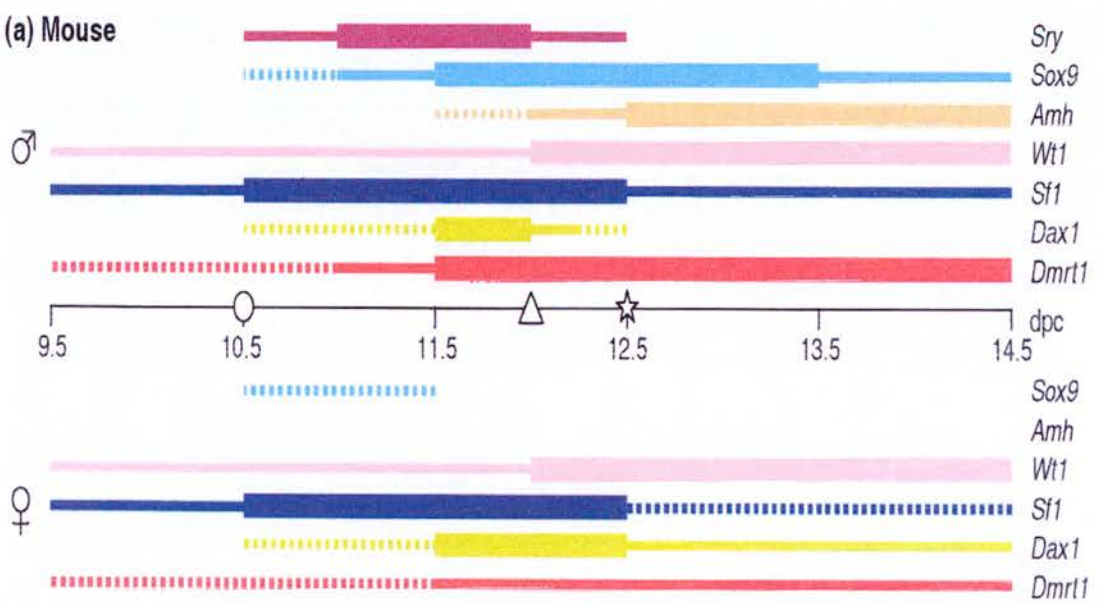


Figure 1.5 Gene Expression Profiles of Sex Determining Genes

Mouse gene expression profiles of sex determining genes in male and female gonads, from E9.5 – E14.5 (Figure adapted from Morrish and Sinclair 2002).

shown to play a key role in testis development (Reviewed in Clarkson and Harley 2002). It is known that in the human syndrome campomelic dysplasia, heterozygous loss-of-function mutations in *SOX9* cause skeletal dysmorpology and repeatedly leads to defects in testicular differentiation or complete XY sex reversal, where ovaries are found instead of testes (Foster, 1996). In contrast, sex determination in heterozygous *Sox9* knockout mice occurs normally, indicating that this pathway is less sensitive to gene dosage in mice than in humans (Bi et al., 2001).

In mice, *Sox9* expression within the developing gonad is sexually dimorphic. Although the level of expression is initially equal in male and female foetal gonads (following *Sry* activation), *Sox9* increases in XY gonads and decreases in XX gonads (Kent et al., 1996). At E12.5-E13.5, *Sox9* expression is localised to the sex cords in the testis, which are composed of Sertoli and germ cells at this stage (Kent et al., 1996; Morais da Silva et al., 1996). In addition, the expression of *Sox9* was also detected within the sex cords of E17.5 male gonads by *in-situ* hybridisation (Kent et al., 1996), which indicates that *Sox9* expression is maintained in the testicular cords as they differentiate into the seminiferous tubules. The early activation of *Sox9* after the expression of *Sry*, its association with sex reversal in man, and its ability to functionally substitute for *Sry* in transgenic mice (Bishop et al., 2000; Vidal et al., 2001) suggest that *Sox9* is a direct downstream target of *Sry* in the sex determination pathway. In addition, high levels of *Sry* expression are found to persist in *Sox9* knockout mice, indicating that the activation of *Sox9* leads to the down regulation of *Sry* (Chaboissier et al., 2004).

Sox9 can function as a transcriptional activator and there is compelling evidence that it regulates *Mis*. Co-transfection studies have shown that Sox9 can bind and transactivate the *Mis* promoter *in vitro* (De Santa Barbara et al., 1998). In addition, mutations in the Sox9-binding site upstream of the *Mis* promoter do not initiate *Mis* transcription *in vivo* (Arango et al., 1999). Recently, an *in vitro* study has revealed that Sox8, a close homologue of Sox9, can also activate the *Mis* promoter, although to a lesser degree (Schepers et al., 2003). The expression of *Sox8* commences at

E12, following the activation of *Sox9*, whereby it also shows a Sertoli cell-specific pattern in the developing testis. However, homozygous *Sox8* knockout mice do not show a gonadal phenotype (Sock et al., 2001). Conversely, *Sox8/Sox9* double knockout analyses suggest that *Sox8* reinforces the function of *Sox9* in testis formation (Chaboissier et al., 2004).

1.5.3 Müllerian Inhibiting Substance (*MIS*)

During mammalian male sexual differentiation, one of the first factors secreted by foetal Sertoli cells is Müllerian inhibiting substance (*MIS*), also known as anti-Müllerian hormone (*AMH*) (Tran et al., 1977). *MIS* is a homodimeric glycoprotein linked by disulphide bonds, belonging to the transforming growth factor-β (TGFβ) superfamily (Picard et al., 1978). It initiates the regression of the Müllerian ducts during sexual development of the male gonad by affecting mesenchymal cells surrounding the Müllerian ducts (Dyche, 1979). Its type II receptor, MIS type II receptor (MISRII), also known as anti-Müllerian hormone receptor (di Clemente et al., 1994), has a very restricted expression pattern and overlaps with the MIS expression pattern in foetal Sertoli and postnatal granulosa cells. It is found in the mesenchymal cells of the Müllerian ducts (Baarends et al., 1994), foetal ovaries and testes, and subsequently in granulosa, Sertoli and Leydig cells (Racine et al., 1998).

MIS signalling is mediated by the single transmembrane-spanning serine/threonine kinase receptors type I and II. The MIS type II receptor binds MIS (Mishina et al., 1999) and then recruits type I receptor forming tetraheteromers. When receptor I is phosphorylated by the MIS type II receptor, downstream signalling occurs. Recently, the widely expressed type I bone morphogenetic protein (BMP) receptor Bmpr1a (also known as Alk3) has been identified as a type I receptor for Mis-induced regression of Müllerian ducts (Jamin et al., 2002). Targeted deletion of *Bmpr1a* in the mesenchymal cells of the Müllerian duct resulted in retention of oviducts and uteri in males (Jamin et al., 2002). Thus, *Mis* has an impact on cell

function by signalling via mesenchymal cells surrounding the Müllerian ducts, which causes it to regress in the presence of the male gonad. Since this receptor is also expressed in the female gonad, it can be concluded that the regulation of Müllerian duct regression is due to the expression of *Mis* by Sertoli cells in the testis.

In the mouse testis, *Mis* expression is first detected at E11.5 and it continues to be expressed throughout foetal development. Several days after birth, the expression level continues and only declines at puberty. In contrast, *Mis* expression is first detectable in granulosa cells of growing follicles in the female several days after birth and persists throughout life (Hirobe et al., 1992). The over-expression of *Mis* in male transgenic mice showed gonadal abnormalities, such as feminized genitalia and fewer Leydig cell numbers, and exhibit reduced serum levels of testosterone (Lyet et al., 1995). Hence, testosterone-dependent organs such as seminal vesicles, epididymis and external genitalia are under-developed or not virilised. The over-expression of *Mis* in female transgenic mice resulted in infertility and most have a blind-ending vagina due to the absence of a uterus and oviducts (Behringer et al., 1990). Although the ovaries are present at birth, the majority of animals have ovaries that do not contain germ cells and cord-like structures develop within 2 weeks after birth (Behringer et al., 1990). In adult females, the ovaries have degenerated. Therefore, in mice, the presence of *Mis* at the foetal stage has a marked impact on ovarian and female germ cell development.

In mice, *Mis*-deficient males possess testes that are fully descended and produce functional sperm. However, they also developed female reproductive organs which interferes with the transfer of sperm into females, resulting in a large majority of these males being infertile (Behringer et al., 1994). Consequently *Mis* does not appear to be essential for the differentiation of the testis in the male, but it may affect subsequent testicular function. For example, in the regulation of postnatal Leydig cell development, as *Mis* knockout male mice develop Leydig cell hyperplasia (Behringer et al., 1994). *Mis* null females exhibit normal fertility but with abnormal

follicle recruitment (Durlinger et al., 1999). Further studies indicated that *Mis* may be able to inhibit the initiation of follicle growth in ovaries.

The *MIS* promoter contains an SF1 binding site which is crucial for high expression levels of the *MIS* gene in Sertoli cells *in vitro* and in transgenic *MIS*-reporter mice (Giuili et al., 1997). However, this SF1 binding site in the *MIS* promoter is not required for the initiation of *Mis* transcription, but is necessary for the upregulation of *Mis* during embryogenesis (Arango et al., 1999). More specifically, WT1 -KTS isoforms associate and synergise with SF1 to promote *MIS* expression (Nachtigal et al., 1998). In addition, it has been proposed that overexpression of *Dax1* may antagonise the synergy between SF1 and WT1 to impair testis development (Nachtigal et al., 1998). Further mutational analysis studies of the *MIS* promoter have found that multiple SF1 binding sites, along with multiple GATA 4 binding sites, are required for normal tissue-specific and developmental expression of *MIS* (Watanabe et al., 2000).

1.5.4 Dosage-Sensitive Sex-Reversal Adrenal Hypoplasia Congenita Critical Region on the X Chromosome (*Dax1*)

The *Dax1* gene encodes an orphan member of the nuclear receptor superfamily that does not contain the typical zinc finger DNA-binding motif, but contains a ligand-binding domain with sequence homology to that of other family members (Zanaria et al., 1994). Mutations of *DAX1* in humans causes an X-linked form of adrenal hypoplasia congenital (AHC) in males (Achermann et al., 2001). AHC patients fail to undergo puberty as they also have hypogonadotropic hypogonadism. These findings led to investigations of *Dax1* in mice, whereby the expression of Dax1 was found to be present in the adrenal cortex, gonads, hypothalamus and anterior pituitary (Guo et al., 1995). During urogential development in the mouse, the expression pattern of Dax1 is sexually dimorphic. Dax1 is expressed in somatic cells of the genital ridge at E10.5-E11.0 and peaks at E12.0 (Swain et al., 1996). The

levels of *Dax1* transcripts in males decrease as testis cords develop, whereas the levels of *Dax1* in females remain unchanged, and the gene continues to be expressed during ovarian development.

In humans, duplication of the *DAX1* gene results in male to female sex reversal (Bardoni et al., 1994). However, in mice, multiple copies of the *Dax1* gene have only been found to cause sex reversal on the background of a weak *Sry* allele (Swain et al., 1998). Therefore, *Dax1* and *Sry* act antagonistically where increasing expression of the former leads to female development and increasing activity of the latter to male development. Although *Dax1* was once considered to be a candidate ovarian determining gene, this hypothesis now seems unlikely as *Dax1* null mutant XX mice have normal ovarian development and function (Yu et al., 1998). Evidence now suggests that the effects of *Dax1* are highly dosage sensitive and dependent upon genetic background (Meeks et al., 2003c), whereby transgenic overexpresssion of *Dax1* causes varying degrees of gonadal dysgenesis and male-to-female sex reversal at the highest dose (Swain et al., 1998). *Dax1* has also been found to interact with *Sf1*, another orphan nuclear receptor, which plays a key role in the development and function of the same tissues. *In vivo* studies have implicated that SF1 regulates the expression of *Dax1* as *Sf1*-deficient mice show reduced levels of Dax1 expression (Hoyle et al., 2002). In addition, a SF1 consensus-binding site found in *Dax1* is required to initiate its expression in the developing gonad. However, the achievement of high levels of Dax1 at this stage requires multiple SF1 consensus binding sites (Hoyle et al., 2002). This is consistent with results from *MIS* promoter studies, whereby SF1 has been shown to be involved and act through several binding sites (Watanabe et al., 2000).

Furthermore, results obtained in *Dax1* knockout mice (*Ahch*, the mouse *Dax1* homologue) suggest that *Dax1* may also play a role in the initiation and maintenance of spermatogenesis (Yu et al., 1998). Spermatogenesis is the process by which spermatogonia (primitive germ cells) undergo mitosis and meiosis to provide a constant supply of male gametes. Mutant male mice were both infertile and

hypogonadal despite normal serum levels of testosterone, gonadotrophins and adrenal steroids (Yu et al., 1998). In wild-type adult male mice, *Dax1* transcripts are detected in both Sertoli and Leydig cells (Ikeda et al., 1996; Swain et al., 1996). However, the relative contributions of Sertoli and Leydig cell dysfunction could not be determined from the *Dax1* knockout mouse. To directly investigate the role of *Dax1* in Sertoli cells, DAX1 was expressed by the MIS promoter (MIS-DAX1) in the genetic background of the *Dax1* knockout model (Jeffs et al., 2001). Sertoli cell-specific expression of the *DAX1* transgene was shown to be sufficient to partially rescue the primary testicular defect of the male *Dax1* knockout mouse (Jeffs et al., 2001). Moreover, fertility was restored to similar levels found in wild-type littermates. However, histopathological examination of the testes only revealed a slight improvement in testicular morphology compared to *Dax1* knockout mice. More specifically, a larger number of seminiferous tubules with round and elongated spermatids were observed in the rescued animals (Jeffs et al., 2001). This finding suggests that *Dax1* may play an important role in other testicular cell types, namely the Leydig cells.

An analysis of Leydig cells from the *Dax1* knockout mice revealed the expression of aromatase (*Cyp19*) was up-regulated (Wang et al., 2001). This enzyme converts testosterone to estradiol. Estradiol acts as a mitogen to somatic cells of the testis, and over-expression of aromatase causes degeneration of the germinal epithelium and Leydig cell hyperplasia (Li et al., 2001). This led to the hypothesis that the phenotype observed in *Dax1* knockout testis may partly be due to steroidogenic dysfunction of the Leydig cells. A 2.1-kilobase fragment of the murine luteinizing hormone receptor 5' promoter (LHR-DAX1) was used to generate transgenic mice that selectively express *Dax1* in Leydig cells (Meeks et al., 2003b). Similar to the Sertoli cell-specific DAX1 rescue, fertility was improved, whereby sperm production was increased. Although testicular size was not increased, testicular pathology was slightly improved in these Leydig cell-specific rescue animals in comparison to the *Dax1* knockout animals (Meeks et al., 2003b).

Taking the Sertoli cell-specific DAX1 rescue along with the Leydig cell-specific DAX1 rescue, the testis phenotype of *Dax1* knockout mice exhibits the combined effects of *Dax1* deficiency in both Sertoli and Leydig cells. Thus, the function of *Dax1* is believed to be critical in both Sertoli and Leydig cells.

1.5.5 Wingless-Type MMTV Integration Site Family, Member 4 (*Wnt4*)

Wnt4 is a member of the Wnt family of secreted molecules that was originally identified as mammalian homologues of the *Drosophila* wingless gene (Nusse et al., 1991; Rijsewijk et al., 1987). Members of this family are locally acting secreted growth factors, which can trigger a cascade of intracellular signals resulting in transcriptional activation of a variety of target genes (Miller, 2002). Thus, they are involved in different developmental processes such as cell differentiation, cell migration, cell polarity and cell proliferation. Wnts bind to members of the Frizzled family of cell-surface receptors (Borello et al., 1999) and are known to take part in at least three separate signalling pathways to effect gene expression (Pandur et al., 2002). The most well known being the Wnt/β-catenin pathway which is highly conserved between species. In vertebrates, Wnt-11 and Wnt-5A have been shown to activate the Wnt/JNK pathway (Yamanaka et al., 2002) and the Wnt/Ca^{2+} pathway has been described in *Xenopus* (Kuhl et al., 2000) and zebrafish (Slusarski et al., 1997).

So far, Wnt4 is the only signalling molecule of the Wnt family known to be involved in sex determination. The expression of *Wnt4* is down-regulated in the developing male gonad after E11.5, but it persists in the developing female gonad (Vainio et al., 1999). It has been shown that targeted deletion of *Wnt4* leads to the masculinization of XX mice, whereby basic development of the Wolffian ducts and degeneration of the Müllerian ducts is found (Vainio et al., 1999). Therefore, *Wnt4* is believed to suppress male development in the developing ovary.

Interestingly, it has been observed that both *Wnt4* and *Dax1* are switched off in the developing testis (at E11.5 and E12.5, respectively), but continue to be expressed in the developing ovary and an excess of both genes can feminize XY embryos (Swain et al., 1998). In addition, it has been shown that, overexpression of *Wnt4* up-regulates *Dax1* which leads to a XY female phenotype (Jordan et al., 2001). These results suggest that *Dax1* and *Wnt4* are both involved in the prevention of testis formation. As *Wnt4* knockout mice show a more severe phenotype compared to the *Dax1* knockout mice, this is consistent with the hypothesis that *Wnt4* acts upstream of *Dax1*. However, it is possible that *Wnt4* causes sex reversal independently of *Dax1*. Further transgenic analyses need to be conducted in order to address this issue.

In humans, an XY female with ambiguous genitalia carrying a large duplication of chromosome 1p35 including the *WNT4* locus was identified (Jordan et al., 2001). This led the authors to hypothesise that this patient's inter-sex phenotype result from a gain of WNT4 function. The authors investigated this hypothesis by generating a gain-of-function mouse model using a clone carrying the human *WNT4* gene controlled by its endogenous promoter (Jordan et al., 2003). This transgenic mouse model showed that WNT4 overexpression interferes with the normal development of testicular vasculature and testosterone synthesis (Jordan et al., 2003), but did not result in an XY sex reversed phenotype as first noted in the human patient carrying a duplication of the *WNT4* locus (Jordan et al., 2001). This may be due to the phenotype in this particular WNT4 transgenic mouse line reflecting a milder form of the symptoms displayed in human patients or other genes which are duplicated within the same chromosome region are necessary for a full XY sex reversal. Alternatively, it may be due to differences between species whereby the transgenic mouse line reacts differently or is less susceptible to WNT4.

It was once thought that *Wnt4* has additional roles in female development, for example, in the suppression of Leydig cell formation. However, it has recently been shown that the role of WNT4 in gonad development (Vainio et al., 1999) is to pattern

the sex-specific vasculature and to regulate steroidogeic cell recruitment (Jeays-Ward et al., 2003). Moreover, WNT4 represses endothelial and steroidogenic cell migration from the mesonephros into the developing XX gonad, which prevents the formation of a male-specific coelomic blood vessel and the production of masculinising sex hormones (Jeays-Ward et al., 2003). Interestingly, it has been demonstrated that WNT4 does not effect Leydig cell differentiation as Sf1:Wnt4 (BAC construct containing Wnt4 cDNA inserted into the 5'untranslated region of the *Sf1* gene by homologous recombination) transgenic embryos showed that WNT4 disrupts the vascular pattern formation of the testis, but the differentiation of Leydig cells in the XY gonad is not affected (Jeays-Ward et al., 2003). Thus, WNT4 is able to repress the molecular pathway that controls the process of migration of at least two different cell types from the mesonephros into the gonad. Further work will be required to establish these molecular pathways in detail, which may lead to a better understanding of the cell movements within the developing gonad.

1.5.6 Fibroblast Growth Factor 9 (*Fgf9*)

Fgf9 is one of at least 22 members of the Fgf family (Ornitz and Itoh, 2001). Members of this family regulate many developmental processes, including embryogenesis of the lung, limb and anterior pituitary. *Fgf9* is widely expressed. *In-situ* hybridisation of mouse embryos detected *Fgf9* expression in skeletal myoblasts, limb apical ectoderm, ventricular myocardium, gut luminal epithelium and the inner ear (Colvin et al., 1999). $Fgf9^{-/-}$ mice die at birth, which is most likely to be caused by lung hypoplasia (Colvin et al., 2001). The observation that females were dramatically overrepresented amongst $Fgf9^{-/-}$ embryos led to the discovery of novel functions for *Fgf9* in sex determination and testicular development.

Mice lacking *Fgf9* were found to be male-to-female sex reversed (Colvin et al., 2001). *Fgf9* is also speculated to act downstream of *Sry* due to reduced mesenchyme in $Fgf9^{-/-}$ XY gonads and the fact that testicular *Fgf9* expression begins shortly after

the onset of *Sry* expression (Colvin et al., 2001). *Fgf9*$^{-/-}$ XY gonads also displayed impaired Sertoli cell development, suggesting that Fgf9 could directly induce Sertoli cell specification, proliferation and/or maintenance of differentiation. In addition, reduced proliferation below the coelomic epithelium at E12.5 was found, indicating that Fgf9 is required for normal proliferation at this stage. Finally, exogenous FGF9 has been shown to induce mesonephric cell migration into E11.5 XX gonads, which suggests that FGF9 could act as a chemotactic factor for mesonephric cells in the early testis (Colvin et al., 2001).

1.5.7 Mouse Homologue of *Polycomb* (*M33*)

M33, a mouse homologue of *Polycomb*, was isolated by means of the structural similarity of its chromodomain (Pearce et al., 1992) to *Polycomb* genes in *Drosophila*. The fifth exon of *M33* contains a region of homology shared by *Drosophila* and *Xenopus* (Pearce et al., 1992; Reijnen et al., 1995). Katoh-Fukui et al. disrupted *M33* in mice by inserting a poly(A) targeting vector into its fifth exon (Katoh-Fukui et al., 1998). More than half of the resultant homozygous mutant mice died before weaning, and intriguingly survivors demonstrated male-to female sex reversal (Katoh-Fukui et al., 1998). Defects in the formation of genital ridges were observed in both XX and XY homozygous mutant embryos (Katoh-Fukui et al., 1998). Furthermore, these defects occurred at approximately the time of *Sry* expression. This suggests that the deficiency of *M33* in mice may cause sex reversal by interfering with events upstream of *Sry*.

1.5.8 Friend of GATA2 (*Fog-2*)

Friend of GATA-2 (*Fog-2*) is an 1,151 amino acid nuclear protein that contains eight zinc finger motifs that are structurally related to those of FOG (Svensson et al., 1999). *Fog-2* is first expressed in the mouse embryonic heart and septum

transversum at E8.5 and is subsequently expressed in the developing neuroepithelium and urogenital ridge (Svensson et al., 1999). In the adult, *Fog-2* is expressed predominately in the heart, brain and testis (Svensson et al., 1999).

Mice homozygous for a null allele of *Fog-2* (*Fog-2*$^{-/-}$) die at mid-gestation from cardiac defects (Svensson et al., 2000). Examination of these mutant XY gonads at E13.5 revealed that testicular cords were absent and semi-quantitative RT-PCR showed a significant decrease in the Sry RNA transcript level in E11.5 *Fog-2*$^{-/-}$ XY gonads when compared with *Fog-2*$^{+/-}$ and *Fog-2*$^{+/+}$ control XY gonads (Tevosian et al., 2002). Fog-2 interacts directly with the N-terminal zinc finger of GATA-4 both *in vitro* and *in vivo* (Svensson et al., 1999). The analysis of the gonads of *Gata-4*$^{-/-}$ embryos is not possible as they die at approximately E7.0-E9.5 (Kuo et al., 1997; Molkentin et al., 1997). However, a *Gata-4* knock-in allele (*Gata-4*^{ki}) approach, which disrupts the interaction between GATA-4 and FOG-2 (Tevosian et al., 2002), allows insight into the effects of the GATA-4/FOG-2 interaction in mammalian gonad development. Analysis of *Gata-4*$^{ki/ki}$ XY gonads also revealed the absence of testicular cords (Tevosian et al., 2002), similar to *Fog-2*$^{-/-}$ XY gonads. Examination of *Fog-2*$^{-/-}$ XY and *Gata-4*$^{ki/ki}$ XY gonads showed that the expression of Sertoli cell markers (*Sox9*, *Mis* and *Dhh*) were absent (Tevosian et al., 2002). This indicates that Sertoli cell differentiation and downstream genes of *Sry* are impaired in the absence of FOG-2 or fully functional GATA-4. Taken together, *Fog-2* and *Gata-4* play a crucial role in the sex development cascade.

1.6 Identification of Genes by Microarray Screens

1.6.1. Production of Microarrays

In the past, comparing expression levels between different tissue samples or cell types was restricted to testing one or a few genes at a time. With the advancement of microarray technology, it is now possible to investigate many genes simultaneously

and to determine which genes are being expressed in a particular cell or tissue type at a specific time and under certain conditions. A microarray consists of multiple features (spots) of DNA that can be used to establish the level of mRNA expression within different tissue samples or cell types. The DNA from each spot represents a gene of interest and is a probe for the mRNA encoded by that gene. Each spot has a unique DNA sequence and each spot will hybridise only to its complementary DNA strand.

At present, microarrays can be produced by two different methods. DNA probes may be synthesized individually, using PCR for cDNAs or chemical synthesis for oligonulceotides. Then, these DNA probes are spotted onto the microarrays, into very small grids, by a robot. Alternatively, DNA oligonucleotides can be synthesized directly onto the microarrays using UV-masks and photoactivated chemistry (Figure 1.6). This is the method currently used by Affymetrix. Probes are synthesized in parallel, whereby an A, C, T or G nucleotide is simultaneously added to multiple growing chains of linker molecules attached to a silane matrix (providing a surface that can be spatially activated by light). Each growing chain is a 'site'. Deprotected sites will have the next base (A, C, T or G) bound to them using UV light, then the next base is bound to those sites and repeated with a different base. To control which sites will be deprotected, a photolithographic mask which allows UV light to activate certain sites is used. This technology by Affymetrix enables large arrays of oligonucleotides to be produced in parallel, however each probe is a maximum of 25 nucleotides in length due to synthesis efficiencies (Affymetrix).

1.6.2 Screening for Sex Determining Genes

Sex determination is a complicated process. Although a large number of sex-determining genes are now known, many more genes remain unidentified. Over the past few years, new genes involved in sex determination in mice have been isolated by cDNA microarray screenings, whereby cDNA libraries have been produced using

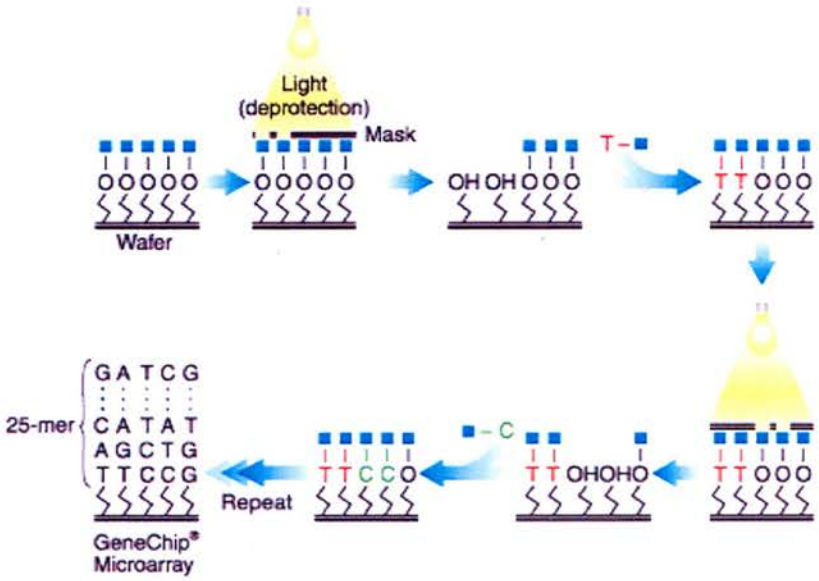


Figure 1.6 Manufacturing of Microarrays Using Photolithography and Combinatorial Chemistry (Affymetrix)

Schematic diagram illustrating the syntheses of DNA oligonucleotides directly onto microarrays using UV-masks and photoactivated chemistry. (Figure adapted from Affymetrix).

microarrays for normalisation of male and female gonadal samples. To date only a few of these screens have been carried out, which have led to the identification of several new genes. For example, a suppression-subtractive hybridisation approach was used to construct normalised cDNA libraries enriched for male-specific or female-specific transcripts in mouse E12.5 urogential ridges and E13.5 gonads (Grimmond et al., 2000). Subsequently, these libraries were efficiently screened for candidate sex determination and gonadogenesis genes, whereby Vanin-1 *(Vnn1)* and Protease nexin-1 *(Pn-1)* were both shown to be play a role in testis development (Grimmond et al., 2000).

Another sex differential screen using a similar PCR-based cDNA subtraction approach has been reported. In this screen, E12.5 XY gonadal cDNA was subtracted with E12.5 XX gonadal cDNA, which yielded 19 genes that were expressed at significantly higher levels in XY gonads (Menke and Page, 2002). Amongst these was *Pn-1*, which again showed higher expression in male gonads than female gonads. In a reciprocal subtraction of E12.5 XX gonadal cDNA with E12.5 XY gonadal cDNA, two genes (*Fst* and *Adamts19*) were identified to be expressed at higher levels in XX gonads (Menke and Page, 2002). The differential expression of all these genes was verified by whole-mount *in-situ* hybridisation.

More recently, another subtractive hybridisation screen using E12.0 – E12.5 mouse testis and ovaries has been conducted (McClive et al., 2003). In this screen, 54 genes were identified to show higher expression in the testis than in the ovary. Both *Vnn1* and *Pn-1* were again identified to be more highly expressed in male gonads than female gonads. From the 19 genes found to be more highly expressed in XY gonads previously identified by Menke and Page (Menke and Page, 2002), 4 of these genes were also identified in the screen by McClive et al. (McClive et al., 2003). These genes were: Cerebellin 1 precursor protein (*Cbln1*), Procollagen, type IX, alpha 3 (*Col9a3*), Prostaglandin D2 synthase (*Ptdgs*) and *Pn-1*. In addition, McClive et al. (McClive et al., 2003) identified 8 genes that were novel or of unknown function.

Plus, analysis of another 18 testis and 9 ovary genes from the same screen is ongoing (McClive et al., 2003).

The above mentioned cDNA screenings have proven to be successful in identifying expression differences which arise during gonadal sex differentiation, leading to the discovery of novel sex determination candidates. However, many other genes involved in the process of gonadal development and sex determination remain to be identified.

1.7 Functional Analysis Approaches

Studying the molecular genetics of mammalian development is most commonly done by the production of knockout mice. While this is an excellent method of functionally analysing genes, this technique is not viable for investigating genes that are necessary at multiple stages of development due to embryonic lethality or abnormalities at the initial stage the gene is required. As a result, this prevents the examination of downstream events. These difficulties can be overcome by advanced transgenic techniques, such as the *cre-lox* system (Gu et al., 1994; Kuhn et al., 1995). However, this is dependent upon having tissue-specific promoters that may not be available. In addition, generation and analysis of knockout mice is time-intensive and expensive. Taking all this into consideration, along with the fact that our aim is to analyse the function of as many potential sex determination candidates as possible from our microarray results, an *in-vitro* approach of gene-specific knockdown in gonadal organ cultures will be attempted. This would be done using the recently developed technique of RNA interference (RNAi) mediated by small interfering RNAs (siRNAs).

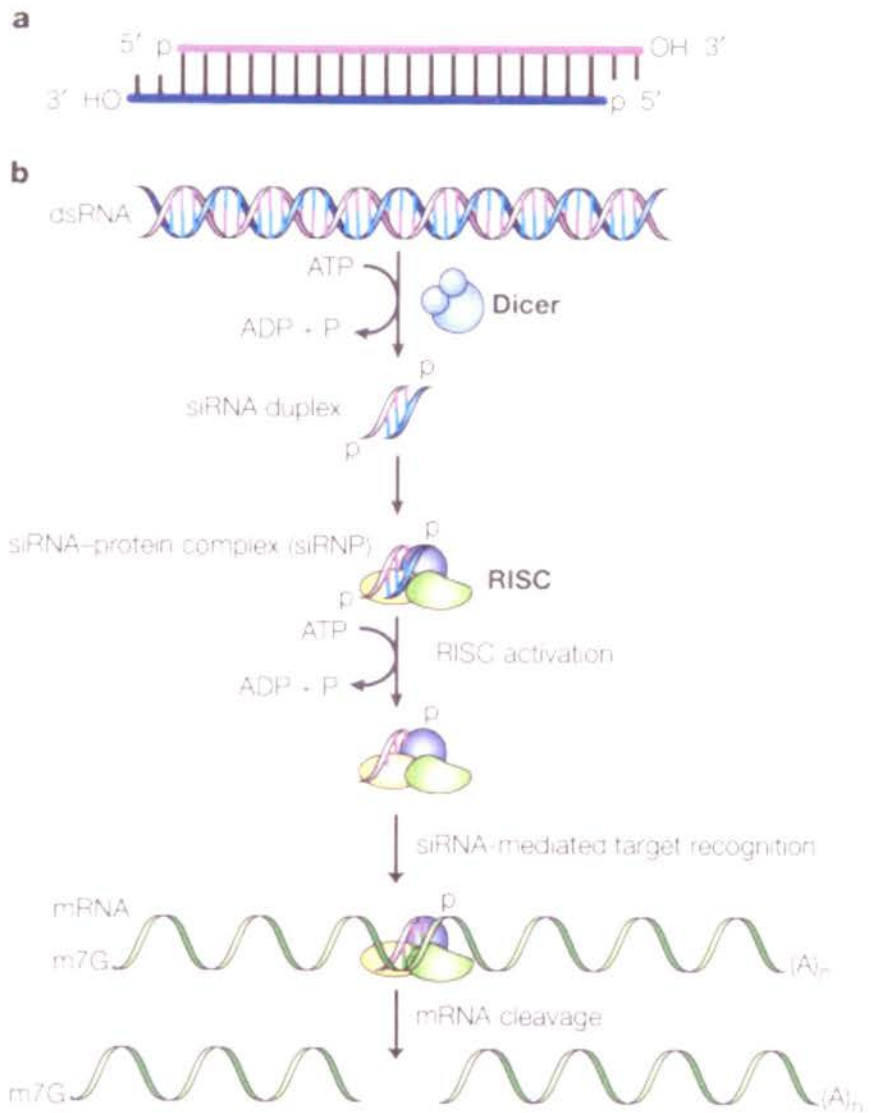


Figure 1.7 Schematic Diagram of the RNA Interference Mechanism (Figure adapted from Dykxhoorn, Novina et al. 2003).

a) Schematic representation of a siRNA which is characterised by 5' phosphorylated ends, a 19 nucleotide duplexed region and a 2 nucleotide unpaired and unphosphorylated 3' ends that are characteristic of RNase III cleavage products.

b) The siRNA pathway. Long dsRNA is cleaved by the Rnase III family member, dicer, into siRNAs in an ATP-dependent reaction. These siRNAs are incorporated into the RNA-inducing silencing complex (RISC) and unwound into separate strands. The anti-sense strand then guides RISC to its homologous target mRNA and then endonucleolytic cleavage of the target mRNA occurs.

1.7.1 Mechanism of RNAi

RNAi is a sequence-specific, post-transcriptional gene silencing process in plants and animals. It is triggered by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene (Sharp, 2001), whereby the degradation is specific to the identified region of the dsRNA (Zamore et al., 2000). This dsRNA is cleaved by the "dicer" enzyme, a highly conserved family member of ribonuclease III (Bernstein et al., 2001), to produce shorter fragments of 21- and 22-nucleotide RNAs. Biochemical characterisation studies have shown that siRNAs are 21-22 nucleotide dsRNA duplexes with symmetric 2-3 nucleotide 3' overhangs and 5' phosphate and 3' hydroxyl groups (Elbashir et al., 2001; Elbashir et al., 2001). These 21- and 22-nucleotide siRNA duplexes are incorporated into a multi-protein RNA-inducing silencing complex (RISC). Only siRNAs which are 5' phosphorylated can enter into RISC (Nykanen et al., 2001) and those lacking a 5' phosphate are rapidly phosphorylated by an endogenous kinase (Schwarz et al., 2002). The duplex siRNA is unwound and separated into single strands. The anti-sense strand then guides RISC to its homologous target mRNA, where it binds to the target mRNA for endonucleolytic cleavage (Figure 1.7), destroying the target mRNA and thus specifically suppressing the expression of that gene (Gura, 2000).

1.7.2 Application of siRNAs

SiRNA represents a novel method in molecular biology for analysing the function of a specific gene in a biological system. The effect of siRNA was established in *Caenorhabditis elegans*, whereby the use of dsRNA was discovered to be significantly more effective at producing interference than either strand alone (Fire et al., 1998). Purified anti-sense and sense RNAs were found to only have a marginal interference activity and required a very high dose of injected RNA to produce any visible effect. In contrast, a mixture of anti-sense and sense RNAs of the same gene segment produced a highly effective interference activity with the endogenous gene

(Fire et al., 1998). Since this discovery, duplex siRNA molecules have been used in all RNAi experiments and the technique has been applied in a variety of organisms.

Recently, in the mouse, duplex siRNA was used in three-dimensional organ culture of kidney rudiments to investigate previously inaccessible aspects of *Wt1* in renal development. Using siRNA to repress *Wt1* expression at different stages of renal development, *Wt1* was found to be required not only at the beginning but also later for nephron differentiation (Davies et al., 2004). In addition, RNAi for the Paired box gene 2 (Pax2) and Wnt4 siRNAs display phenotypes that are in accordance with *in vitro* observations made on the interactions of these genes. Pax2 is known to bind and transactivate the *Wt1* promoter *in vitro* (Dehbi et al., 1996; McConnell et al., 1997). Using Pax2 siRNAs, *Wt1* expression is repressed and both bud growth and nephron differentiation is blocked (Davies et al., 2004). However, Wnt4 siRNAs blocks nephron differentiation without having an affect on *Wt1* expression (Davies et al., 2004). This is in line with *Wnt4*$^{-/-}$ knockout mice, where the basic mesenchymal expression of Wt1 is unchanged (Sim et al., 2002). Thus, the use of siRNAs to knockdown gene expression was successful in this organ culture system.

The use of siRNA in gene expression knockdown is being adapted for various culture systems. However, to date, siRNA has never been applied in gonadal organ cultures. Analysis of a large number of sex-determining genes is time-consuming, therefore an *in vitro* approach of gene specific knockdown would be advantageous. Thus, the development of a system that could be used on a wide-scale was required and such a system could be based on siRNA. We believe that the establishment of siRNA in gonadal organ cultures would be an alternative approach for allowing many sex determination genes to be functionally analysed.

1.8 Aims and Summary of Project

This thesis describes a project that aims to further identify new genes involved in the development of the gonad, which may lead to a better understanding of the male and female developmental pathways. Therefore, a sex determination screen using Affymetrix GeneChip arrays will be performed as they provide the largest, most complete set of mouse transcripts available at the time. Wild-type E12.5 male and female gonads will be used for our screen. More specifically, each embryo will be quantitatively staged to ensure more accuracy of the developmental time point and only gonads staged between 28 to 30 tail somites will be used. The trigger for sex determination is controlled by the expression of *Sry*, which peaks at E11.5. Using E12.5 gonads, which reflects a time point shortly after the initiation of sex determination, will allow the isolation of genes at a very early stage of the sex determination process rather than genes expressed in the already differentiated tissues of the ovary or testis. The mesonephros will be removed from each gonad to increase the sensitivity of our screen and focus on genes directly involved in sex determination.

Following the Affymetrix sex determination screen, microarray data will need to be verified to ensure that the results show sex differentially expressed genes. This will first be done by whole-mount *in-situ* hybridisation of male and female gonads at E12.5. Initially, a known sex differentially expressed gene, also identified from the screen, will be used. If this is confirmed to be sex differentially expressed, a number of novel candidates showing the largest differences between male and female expression will then be selected for further analysis. Extending this analysis, transcripts that are shown to be sex differentially expressed will be selected for a more detailed expression profile analysis. This will involve further *in-situ* hybridisation of male and female gonads at E11.5, E13.5 and E14.5. In addition, *in-situ* hybridisation of E14.5 gonadal sections will be performed in an attempt to identify the specific cell type within which each gene is expressed. Real-time PCR of these candidates will also be carried out to verify the fold changes from the

microarray analysis and to determine a specific gonadal development gene expression profile, ranging from 17 to 36 tail somites.

It would also be interesting to analyse the function of these potential candidates. Due to the expected high number of candidates that may arise, a fast yet efficient approach was needed. It was therefore decided to attempt the novel approach of applying siRNAs to a gonadal organ culture system. Once this system is established, the function of many potential sex determination candidates can be analysed.

In this thesis, chapter 3 describes the Affymetrix sex determination screen, along with the results obtained and analysis of the microarray data. Verification of this microarray data set is described in chapter 4, whilst chapter 5 shows a detailed analysis of three novel sex determination candidates identified from the screen. Finally, chapter 6 describes the novel approach of applying siRNAs to a gonadal organ culture system. It shows the establishment of this system by illustrating several known sex determination genes were successfully targeted and knocked-down by applying this technique. Importantly, siRNA against *Sry* effectively blocked male gonad differentiation resulting in a lack of expression of male specific markers. Taken together these results suggest that the siRNA approach in gonad cultures can be used to efficiently analyse the function of candidate genes isolated in our sex determination screen. This may lead to a better understanding of the sex development pathways.

Chapter 2 Materials and Methods

Chapter 2 Materials and Methods

2.1 Manipulation of Nucleic Acids

2.1.1 General Reagents Used in Molecular Biology Procedures

All chemicals were analytical grade and were supplied by Sigma, Promega, Gibco BRL, Flowgen, Fisher Scientific, Invitrogen, Roche and Boehringer Mannheim. General solutions were prepared, autoclaved and stored at room temperature.

Tris.HCl:

Tris base (tris[hydroxymethyl]aminomethane) was dissolved in sterile water. HCl was used to adjust the pH to the required value.

EDTA:

EDTA (ethyldiaminetetra-acetic acid di-sodium salt) was dissolved in sterile distilled water. The solution was adjusted to pH8.0 by adding solid NaOH pellets.

TE Buffer:

10mM Tris.HCl (pH7.5); 1mM EDTA.

TBE Buffer (20X stock):

Tris base	216g
Boric acid	110g
0.5M EDTA	80mM

Distilled water was added to a final volume of 1 litre. Stock was diluted to 1X with distilled water.

2.1.2 DNA Electrophoresis

DNA molecules were separated according to their size by agarose gel electrophoresis.

Agarose Gel Loading Buffer:

Loading buffer was prepared as a 10X stock and stored at room temperature.

	Final concentration
Ficoll	20%
Orange G (Sigma)	1%
EDTA	100mM

Made to required volume with distilled water.

Gels were prepared by heating and dissolving the required amount of agarose (Q-BioGene) in 1X TBE. Molten agarose was cooled and ethidium bromide added to a final concentration of 50μg/100ml agarose. The percentage of agarose used was dependent on the size range of the DNA molecules to be separated. PCR products were generally run on 1% - 2% agarose gels. DNA size markers were run alongside experimental samples to estimate the size and amount of DNA in the samples. The size marker used routinely was bacteriophage lamda DNA digested with *Hind*III ('λ') (Gibco). Samples were run in 1X loading buffer. Following electrophoresis DNA fragments were visualised on an UV trans-illuminator. If DNA fragments were to be purified from the gel then low melting point agarose (Flowgen) was used and the DNA was recovered using a gel extraction kit (Qiagen), according to the manufacturer's instructions.

2.1.3 DNA Quantification

DNA concentrations were determined either by agarose gel electrophoresis or by measuring the absorbance at 260nm (A_{260}) using a spectrophotometer.

To determine the concentration by electrophoresis, several different volumes of the DNA (e.g. 1, 3 and 5μl) were run alongside standard amounts of DNA (usually DNA size markers). An estimate of the concentration was made by visual comparison of the intensity of the sample's fluorescence to that of the known amounts of DNA under UV illumination.

To determine the concentration with a spectrophotometer, the DNA sample was diluted 1:100 with dH_2O. The spectrophotometer was calibrated using a water only blank sample. The samples were placed in clean cuvettes and the absorbance at 260nm (A_{260}) was measured. The concentration of the original sample was calculated as follows:

Concentration (mg/ml) = A_{260} x 100 (dilution factor) x 50 (conversion factor).

2.1.4 Restriction Digests

Restriction enzymes were either purchased from Amersham, Promega or Boehringer Mannheim. Digestions of DNA with restriction endonucleases were carried out in the appropriate buffer using the optimal conditions as recommended by the particular manufacturer. Digests were performed using 1X buffer, DNA (10ng - 1μg/μl) and 0.5 – 5 units of enzyme per μg of DNA. The total volume of enzyme never exceeded 10% of the reaction volume to avoid glycerol inhibition of digestion. Incubations were at 37°C, unless otherwise instructed, for periods from 1 hour to overnight.

2.1.5 PCR

DNA molecules were amplified by polymerase chain reaction (PCR). Primers were purchased from Invitrogen as lyophilised desalted compounds. Stocks were made up to 500ng/μl using sterile dH_2O. Working stocks were diluted to 10ng/μl (1:50

dilution). Primers were used in PCRs at a final concentration of 1ng/μl (1:10 dilution). REDTaq™ ReadyMix™ PCR Reaction Mix (Sigma) containing Ampli*Taq*, PCR buffer, dNTPs, Mg^{2+} and gel loading buffer was routinely used at 10μl per 20μl reaction. PCRs were done in a MJ Research DNA Engine Tetrad. 100ng-200ng of genomic DNA/cDNA was used for PCR reactions. Generally all PCR programmes were identical (35 cycles), except the annealing temperature was varied according to the primers used.

General PCR Programme:

94°C	3 minutes		
94°C	45 seconds	}	
n°C (variable annealing temperature)	45 seconds	}	x 34 cycles
72°C	1 minute	}	
72°C	9 minutes		

2.1.6 Generation of DNA Templates for Riboprobes

DNA templates for riboprobes were generated by PCR amplification using primers with T3 and T7 promoter sequences incorporated into their design (see Appendix 1). All primers were purchased from Invitrogen as lyophilised desalted compounds and made up as described in section 2.1.5. JumpStart Taq (Sigma) was used in an attempt to reduce non-specific priming and the formation of primer dimers. JumpStart Taq (Sigma) is a Taq DNA polymerase with a neutralizing antibody rendering it inactive at room temperature. Its activity is restored once the reaction temperature reaches 70°C. PCR reactions were set up as follows:

2X JumpStart™ REDTaq™ ReadyMix™ PCR Reaction Mix (Sigma)	10μl
Primer 5' (10ng/μl)	2μl
Primer 3' (10ng/μl)	2μl
cDNA (from dissected gonads)	1μl
dH_2O	5μl
	20μl

JumpStart PCR Programme:

94°C	10 minutes		
94°C	45 seconds	}	
n°C (variable annealing temperature)	45 seconds	}	x 34 cycles
72°C	1 minute	}	
72°C	9 minutes		

The correct PCR products were excised from the gel and the DNA was recovered using a gel extraction kit (Qiagen), according to the manufacturer's instructions. DNA was stored at -20°C until labelling with digoxygenin (section 2.5.1).

2.2 RNA Manipulation

2.2.1 Preparation of RNA

Cells were treated with 0.25% trypsin/EDTA (Gibco) to detach them from the dish and collected into a loose pellet by centrifugation at a low speed (3000rpm) for 5 minutes. The supernatant was removed before cells were flash-frozen in liquid nitrogen and stored at -80°C. Dissected/cultured gonads were also immediately flash-frozen in liquid nitrogen to minimise RNA degradation and then stored at -80°C. Total RNA was isolated from flash-frozen samples using an Absolutely RNA RT-PCR Mini-prep kit (Stratagene). Manufacturer's instructions were followed and eluted RNA samples were stored at -80°C.

2.2.2 cDNA Synthesis and RT-PCR

2-5μg of total RNA from each sample was diluted with sterile dH_2O to bring the volume to 18μl and then 2μl of 1.5mM random hexamer (Amersham) was added. Samples were denatured at 95°C for 5 minutes and then placed on ice. 8μl of 5X reaction buffer (Gibco), 2μl of 100mM DTT (Gibco), 8μl of 2.5mM dNTP mix (Promega), 2μl of RNase inhibitor (Roche) and 3μl of M-MLV reverse transcriptase (Gibco) was added to each sample. Samples were incubated at 37°C for 1 hour, then heated at 95°C for 5 minutes and the resulting cDNA was stored at -20°C until required. RT-PCR was carried out as described in section 2.1.5.

2.3 Mouse Experiments

All animals were maintained in a specific pathogen free (SPF) environment and experiments were carried out under Home Office licence. Wild-type animals were either bred in-house or obtained from Charles River Laboratories. Breeding animals were maintained on a C57/Bl-6 background.

2.3.1 Mouse Strains and Timed Matings

C57/Bl-6 mice were used in all wild-type timed matings. Only *Sox8*$^{+/+}$ mice (from Dr Michael Wegner) were used in *Sox8* experiments (Sock et al., 2001). Embryos for all experiments were generated from timed matings, with the morning of vaginal plug detection being counted as embryonic day 0.5 (E0.5).

2.3.2 Genotyping of Embryos

Genotyping by PCR amplification was performed to determine the sex of each embryo. Genomic DNA derived from embryo heads or embryonic yolk sacs (for intact whole-mount *in-situ* hybridisation embryos) was digested in 300μl of mouse tail buffer/proteinase K at 55°C for 2-3 hours and then stored at -20°C overnight. During setup of PCR reaction mix, DNA samples were centrifuged at 4°C and 0.2μl of genomic DNA (taken from the top) was used.

Mouse Tail Buffer:

	Final concentration
Tris.HCl (pH7.9)	10mM
EDTA	1mM
NaOAc	300mM
SDS	1%

Proteinase K (stock 10mg/ml) (Sigma) was added at time of use to a final concentration of 0.2mg/ml.

The following pairs of primers (purchased from Invitrogen as lyophilised desalted compounds) were used in the same reaction:

Zfy 5' (forward primer): GAC TAG ACA TGT CTT AAC ATC TGT CC
Zfy 3' (reverse primer): CTT ATT GCA TGG ACT GCA GCT TAT G
Pax6 H499 (forward primer): CTT TCT CCA GAG CCT CAA TCT G
Pax6 H500 (reverse primer): GCA ACA GGA AGG AGG GGG AGA

The Y chromosome-linked Zfy primers were used to identify male samples and the Pax6 primers were used as an internal control. Thus, male samples show two DNA bands (Zfy and Pax6, approximately 200bp and 150bp respectively), whilst female samples show one DNA band (Pax6, approximately 150bp).

PCR Reaction Mix:

2X REDTaq™ ReadyMix™ PCR Reaction Mix (Sigma)	10μl
Zfy 5' (10ng/μl)	1μl
Zfy 3' (10ng/μl)	1μl
Pax6 H499 (10ng/μl)	1μl
Pax6 H500 (10ng/μl)	1μl
Genomic DNA	0.2μl
dH_2O	5.8μl
	20μl

PCR Programme:

94°C	3 minutes		
94°C	45 seconds	}	
65°C	45 seconds	}	x 34 cycles
72°C	1 minute	}	
72°C	9 minutes		

2.3.3 Quantitative Staging of Embryos

Embryos were quantitatively staged by tail somite counting, beginning at the first somite posterior to the hind limb towards the tip of the tail. Using this method, E10.5 embryos corresponds to approximately 8 tail somites, E11.5 embryos corresponds to approximately 18 tail somites, and E12.5 embryos corresponds to approximately 30 tail somites (Hacker et al., 1995; Schmahl et al., 2000).

2.4 Affymetrix GeneChip Microarrays

2.4.1 RNA Isolation

Wild-type E12.5 gonads were dissected in PBS, transferred to 1.5ml microfuge tubes, immediately snap-frozen in liquid nitrogen and stored at -80°C. After the genotype of each pair of gonads was determined by PCR amplification (as described in section 2.3.2), gonads of the same sex were pooled together. Only gonads quantitatively staged between 28 to 30 tail somites (as described in section 2.3.3) were used. The Absolutely RNA RT-PCR Mini-prep kit (Stratagene) was used for RNA isolation and manufacturer's instructions were followed. A total of 52 male gonads and 52 female gonads were pooled in screen 1, which yielded 6.00μg and 4.60μg of RNA respectively. For screen 2, a total of 20 male gonads and 20 female gonads were pooled, which yielded 2.64μg and 1.43μg of RNA respectively. RNA was quantified by measuring the absorbance at 260nm (A_{260}) using a spectrophotometer. The concentration of each sample was calculated as follows:

Concentration (μg/μl) = A_{260} x dilution factor x 0.04 (conversion factor).

2.4.2 Screen 1

For screen 1, double-stranded cDNA was synthesized from total RNA using cDNA Synthesis System (Roche). All components were contained in kit. Following manufacturer's instructions, 2μl of oligo $[(dT)_{24}$ T7 promotor$]_{65}$ primer was added to each RNA sample and incubated at 70°C for 10 minutes and then placed on ice. 8μl 5X RT-buffer, 4μl 0.1M DTT, 2μl AMV, 1μl RNase inhibitor and 4μl 10mM dNTP mix were added to each sample, gently mixed and incubated at 42°C for 1 hour. Tubes were then placed on ice to terminate this first strand reaction. For second strand synthesis, 30μl 5X second strand buffer, 1.5μl 10mM dNTP mix, 6.5μl second strand enzyme blend, 72μl redistilled water was added to the first strand reaction.

Samples were gently mixed and incubated at 16°C for 2 hours. 20μl T4 DNA polymerase was added and incubated at 16°C for 5 minutes. The reaction was stopped by adding 17μl 0.2 M EDTA (pH 8.0). To digest the residual RNA, 1.5μl RNase I was added and samples were incubated at 37°C for 30 minutes. Then, 5μl Proteinase K was also added and samples were incubated for another 30 minutes at 37°C. Phenol was used to clean-up the double-stranded cDNA, followed by ethanol precipitation. Each pellet was dissolved in 6μl RNase-free water. The synthesis of double-stranded cDNA was checked by running a small aliquot on a 1% agarose gel.

Following the synthesis of double-stranded cDNA, the *in-vitro* transcription labelling reaction was set up as detailed below:

NTP labelling mix for 1 reaction:
2μl 10X ATP (75mM; Ambion)
2μl 10X GTP (75mM; Ambion)
1.5μl 10X CTP (75mM; Ambion)
1.5μl 10X UTP (75mM; Ambion)
3.75μl Bio-11-CTP (10mM; Affymetrix)
3.75μl Bio-16-UTP (10mM; Affymetrix)

For each *in-vitro* transcription labelling reaction:
6μl double-stranded cDNA
14.5μl NTP labelling mix (as described above)
2.5μl 10X T7 transcription buffer (Ambion)
2μl 10X T7 enzyme mix (Ambion)

Samples were incubated at 37°C for 5 hours. 75μl DEPC-treated water was added to each cRNA product and the RNeasy Mini Kit (Qiagen), following manufacturer's instructions, was used to clean-up the labelled cRNA. A small aliquot was taken to measure the absorbance at 260nm (A_{260}), determining the amount of labelled cRNA obtained. The concentration of each sample was calculated as follows:

Concentration ($\mu g/\mu l$) = A_{260} x dilution factor x 0.04 (conversion factor).

The male pool of gonads resulted in 8.76μg of labelled cRNA and the female pool of gonads resulted in 6.96μg of labelled cRNA. 5μg of labelled cRNA from each pool was fragmented and hybridised onto separate microarrays (section 2.4.4).

2.4.3 Screen 2

For screen 2, the small sample labelling protocol version II produced by Affymetrix was used. Following their recommendations, 1μl T7-Oligo(dT) promoter primer (5μM; Affymetrix) was added to each RNA sample and incubated at 70°C for 6 minutes and then cooled to 4°C for 2 minutes. Meanwhile, the RT_Premix_1 was prepared in a large volume to minimise pipetting errors.

RT_Premix_1:

	For 4 reactions
DEPC-treated water	1.5μl
5X first strand buffer (Invitrogen)	4μl
0.1M DTT (Invitrogen)	2μl
10mM dNTP mix (Invitrogen)	1.5μl
RNase inhibitor (Roche)	1μl
SuperScript II (Invitrogen)	2μl
Total volume	**12μl**

3μl of the RT_Premix_1 was added to each sample, mixed thoroughly and incubated at 42°C for 1 hour. Samples were then heated at 70°C for 10 minutes to inactivate SuperScript II. Tubes were placed on ice to cool while the SS_Premix_1 was prepared.

SS_Premix_1:

	For 4 reactions
DEPC-treated water	91μl
5X second strand buffer (Invitrogen)	30μl
10mM dNTP mix (Invitrogen)	3μl
DNA ligase (Invitrogen)	1μl
DNA polymerase I (Invitrogen)	4μl
RNase H (Invitrogen)	1μl
Total volume	**130μl**

32.5μl of the SS_Premix_1 was added to each first strand reaction. Components were mixed thoroughly and incubated at 16°C for 2 hours. 1μl T4 DNA polymerase (Invitrogen) was added and incubated at 16°C for 10 minutes. This first cycle double-stranded cDNA was cleaned-up by ethanol precipitation and pellets were dried in a speed vacuum.

At room temperature, the following reagents from the MEGAscript T7 kit (Ambion) were added to the dried double-stranded cDNA pellet:

	For 1 reaction
DEPC-treated water	4μl
Pre-mixed NTPs, 18.75mM each	4μl
10X reaction buffer	1μl
10X enzyme mix	1μl
Total volume	**10μl**

Components were mixed thoroughly and incubated at 37°C for 5-6 hours. This first cycle *in-vitro* transcription reaction resulted in amplification of the cRNA. 90μl of RNase-free water was added to each cRNA product and the RNeasy Mini Kit (Qiagen), following the RNeasy Mini Protocol from manufacturer's handbook, was used to clean-up the cRNA.

The entire cRNA elute was used in the second cycle amplification reaction. Beginning with the second cycle first strand cDNA synthesis, 1μl random primers (Invitrogen) was added to each cRNA sample, mixed thoroughly, incubated at 70°C for 10 minutes and then cooled on ice for 2 minutes. Meanwhile, the RT_Premix_2 was prepared.

RT_Premix_2:

	For 2 reactions
5X first strand buffer (Invitrogen)	4μl
0.1M DTT (Invitrogen)	2μl
10mM dNTP mix (Invitrogen)	1μl
RNase inhibitor (Roche)	1μl
SuperScript II (Invitrogen)	2μl
Total volume	**10μl**

5μl of the RT_Premix_2 was added to each sample, mixed thoroughly and incubated at 42°C for 1 hour. To remove the RNA template, 1μl RNase H (Invitrogen) was added and samples were incubated at 37°C for 20 minutes. To inactivate the RNase H, samples were heated at 95°C for 5 minutes and then chilled on ice for 2 minutes.

To synthesize the second strand of the second cycle cDNA, 2μl T7-Oligo(dT) promoter primer (5μM; Affymetrix) was added to each chilled sample, mixed thoroughly and incubated at 70°C for 6 minutes. Tubes were placed on ice to cool while the SS_Premix_2 was prepared.

SS_Premix_2:

	For 1 reaction
DEPC-treated water	43.5μl
5X second strand buffer (Invitrogen)	15μl
10mM dNTP mix (Invitrogen)	1.5μl
DNA polymerase I (Invitrogen)	2μl
Total volume	**62μl**

62μl of the SS_Premix_2 was added to each sample, mixed thoroughly and incubated at 16°C for 2 hours. Then 2μl T4 DNA polymerase (Invitrogen) was added, mixed thoroughly and incubated at 16°C for 10 minutes. This second cycle double-stranded cDNA was cleaned-up by ethanol precipitation and pellets were dried in a speed vacuum. cDNA pellets were stored at -20°C overnight before proceeding to the second cycle of *in-vitro* transcription and labelling reaction.

At room temperature, the following reagents from the ENZO BioArray High Yield RNA Transcript Labelling Kit (Affymetrix) were added to the dried double-stranded cDNA pellet from the previous step:

	For 1 reaction
DEPC-treated water	22μl
10X HY reaction buffer	4μl
10X Biotin labelled ribonucleotides	4μl
10X DTT	4μl
10X RNase inhibitor mix	4μl
20X T7 RNA polymerase	2μl
Total volume	**40μl**

Components were thoroughly mixed and the pellet was completely dissolved before samples were incubated at 37°C for 4-5 hours. 60μl of RNase-free water was added to each cRNA product and the RNeasy Mini Kit (Qiagen), following the RNeasy

Mini Protocol from manufacturer's handbook, was used to clean-up the labelled cRNA. A small aliquot was taken to measure the absorbance at 260nm (A_{260}), determining the amount of labelled cRNA obtained. The concentration of each sample was calculated as follows:

Concentration ($\mu g/\mu l$) = A_{260} x dilution factor x 0.04 (conversion factor).

The male pool of gonads resulted in 17.52μg of labelled cRNA and the female pool of gonads resulted in 16.48μg of labelled cRNA. 15μg of labelled cRNA from each pool was fragmented and hybridised onto separate microarrays (section 2.4.4).

2.4.4 Fragmentation and Hybridisation

2.4.4.1 Solutions for Fragmentation and Hybridisation

5X Fragmentation Buffer:
200mM Tris-acetate (pH 8.1)
500mM KOAc
150mM MgOAc
Filter sterilise (0.2μm)
Store at room temperature

12X MES Stock Buffer:
70.4g MES free acid monohydrate (Sigma)
193.3g MES Sodium Salt (Sigma)
800ml Molecular Biology Grade water (BioWhittaker)
Mix and adjust volume to 1000mL (pH should be between 6.5 and 6.7)
Filter sterilise (0.2μm)
Store at 4°C

2X Hybridisation Buffer:

8.3ml 12X MES Stock Buffer

17.7ml 5M NaCl

4.0ml 0.5M EDTA

0.1ml 10% Tween 20 (Sigma)

19.9 ml dH_20

2.4.4.2 Fragmentation and Hybridisation Procedure

16μl of labelled cRNA (5μg in screen 1, 15μg in screen 2) was fragmented in 4μl 5X fragmentation buffer by incubating at 95°C for 35 minutes and then placed on ice to cool.

Component	Screen 1	Screen 2
Fragmented cRNA	5μg	15μg
Control Oligonucleotide B2 (3nM; Affymetrix)	6μl	5μl
20X Eukaryotic Hybridisation Controls (BioB, BioC, BioD, cre; Affymetrix)	----	15μl
Herring Sperm DNA (10mg/ml; Promega)	2μl	3μl
Acetylated BSA (50mg/ml; Gibco)	2μl	3μl
2X Hybridisation Buffer	100μl	150μl
dH_20	To final volume of 250μl	To final volume of 300μl
Final Volume	**250μl**	**300μl**

Table 2.1 Hybridisation Cocktail for Single Probe Array

The above components were mixed for each fragmented cRNA target for hybridisation to multiple probe arrays (as recommended by Affymetrix).

Probe arrays were equilibrated to room temperature immediately before use. Affymetrix Mouse U74Av1 and U74Bv1 GeneChips were used in screen 1. In screen 2, the complete set of Affymetrix Mouse U74v2 GeneChips (A, B and C) was used. Once equilibrated, arrays were filled with 250µl 1X hybridisation buffer and incubated at 45°C for 10 minutes with rotation. Meanwhile, each hybridisation cocktail was heated to 99°C for 5 minutes and then incubated at 45°C for 5 minutes. Hybridisation cocktails were centrifuged for 5 minutes to remove any insoluble material from the hybridisation mixture. The buffer solution was removed from the probe array cartridge and replaced with the relevant hybridisation cocktail (avoiding any insoluble matter at the bottom of the tube). Probe arrays were placed in a 45°C rotating oven (Affymetrix) and hybridisation occurred for 16 hours.

2.4.5 Washing, Staining and Scanning

2.4.5.1 Solutions for Washing, Staining and Scanning

Non-Stringent Wash Buffer:

300ml 20X SSPE (3M NaCl, 0.2M NaH_2PO_4, 0.02M EDTA; BioWhittaker)

1ml 10% Tween 20 (Sigma)

698ml dH_20

Filter sterilise (0.2μm)

Add 1ml 5 % Antifoam (Sigma)

Stringent Wash Buffer:

83.3ml 12X MES Stock Buffer

5.2ml 5M NaCl

1ml 10% Tween 20 (Sigma)

910.5ml dH_20

Filter sterilise (0.2μm)

2X Stain Buffer:

41.7ml 12X MES Stock Buffer

92.5ml 5M NaCl

2.5ml 10% Tween 20 (Sigma)

112.8ml dH_20

Filter sterilise (0.2μm)

Add 0.5ml 5 % Antifoam (Sigma)

SAPE Stain Solution:

600µl 2X Stain Buffer

540µl dH_20

48µl 50mg/ml acetylated BSA (Gibco)

12µl 1mg/ml SAPE (Molecular Probes)

Mix well and divide into two aliquots of 600µl each to be used for stains 1 and 3 respectively.

Antibody Solution:

300µl 2X Stain Buffer

266.4 dH_20

24µl 50mg/ml acetylated BSA (Gibco)

6µl 10mg/ml normal goat IgG (Sigma)

3.6µl 0.5mg/ml biotinylated anti-streptavidin goat antibody (Vector Laboratories)

10mg/ml Goat IgG Stock (Sigma):

Resuspend 50mg in 5ml PBS

Store at 4°C

2.4.5.2 Washing, Staining and Scanning Protocol

	EukGE-WS2
Post Hyb Wash #1	10 cycles of 2 mixes/cycle with non-stringent wash buffer at 25°C
Post Hyb Wash #2	4 cycles of 15 mixes/cycle with stringent wash buffer at 50°C
Stain	Stain probe array for 10 minutes in SAPE stain solution at 25°C
Post Stain Wash	10 cycles of 4 mixes/cycle with non-stringent wash buffer at 25°C
Second Stain	Stain probe array for 10 minutes in antibody solution at 25°C
Third Stain	Stain probe array for 10 minutes in SAPE stain solution at 25°C
Final Wash	15 cycles of 4 mixes/cycle with non-stringent wash buffer at 30°C. The holding temperature is 25°C.

Table 2.2 Fluidics EukGE-WS2 protocol for washing and staining standard probe arrays

Hybridisation cocktails were removed from each probe array cartridge and stored at -80°C. Probe arrays were filled with non-stringent wash buffer and inserted into the fluidics station (Affymetrix). Following Affymetrix's technical manual, the fluidics

station (Affymetrix) was setup to wash and stain each microarray using the EukGE-WS2 protocol (Table 2.2).

After washing and staining, the probe arrays were ejected from the fluidics station (Affymetrix) and placed into the HP GeneArray scanner (Affymetrix). Probe arrays were scanned and data was acquired. Data for all chips were analysed using Affymetrix Microarray Suite (version 4.0.1 in screen 1 and version 5.0 in screen 2).

2.5 RNA In-situ Hybridisation

2.5.1 Labelling Riboprobes with Digoxygenin

Riboprobes were labelled with digoxygenin for non-radioactive *in-situ* hybridisation. The polymerases, DIG labelling mix, transcription buffer and RNase inhibitor were purchased (Roche). Plasmids containing the riboprobe were linearised with an appropriate restriction enzyme. DNA templates containing the riboprobe generated from T3 and T7 PCR amplification (as described in section 2.1.6) did not require linearisation. The labelling reaction was setup as below and incubated at 37°C for 2 hours.

dH_2O	13μl - xμl
10X Transcription Buffer	2μl
10X NTP mix-DIG labelled UTP	2μl
Linearised DNA	xμl
RNase Inhibitor	1μl
Polymerase (T3 or T7)	2μl
	20μl

Template DNA was digested by adding 2μl DNase 1 (RNase free) (Roche) for 15 minutes at 37°C. The labelled RNA was precipitated with 2.5μl 3M NaOAc and 75μl 100% ethanol at -80°C for 30 minutes. Tubes were then centrifuged at 4°C for 20 minutes at maximum speed. The pellet was washed with 50μl cold 70% ethanol and resuspended in 30μl DEPC-treated dH_2O. 1μl RNase inhibitor was added. The RNA was checked by electrophoresis.

2.5.2 Whole-mount *In-situ* Hybridisation

Wild-type embryos were processed for whole-mount *in-situ* hybridisation. Embryos were dissected in PBS and the head (or embryonic yolk sac for intact embryos) was used for genotyping. All steps were done at room temperature unless stated otherwise. Solutions were made up using DEPC-treated water and PBS for steps up to and including the hybridisation step. These solutions were DEPC-treated by adding 1μl/ml DEPC (Sigma) to sterile water or PBS, gently shaking at room temperature for several hours (usually overnight) prior to re-autoclaving.

2.5.2.1 Solutions for Whole-mount *In-situ* Hybridisation

RIPA Buffer: (not autoclaved, but DEPC-treated solutions were used)

2.5ml 10% SDS

15ml 5M NaCl

5ml Nonidet-P40

25ml 10% Deoxycholate

1ml 0.5M EDTA

25ml 1M Tris.HCl (pH8)

426.5ml DEPC H_2O

PBT:

PBT was prepared fresh each time by adding 0.5ml 10% Triton X-100 to 50ml DEPC-treated PBS (final concentration: 0.1%).

20X SSC:

87.65g NaCl

44.1g Sodium Citrate

Dissolve above in 400ml dH_2O

Adjust to pH7 with conc. HCl

Adjust to 500ml with dH_2O

Treat with 0.01% DEPC

Autoclave

Heparin (Sigma):

100mg/ml in 4X DEPC-SSC

Store at -20°C

Hybe Buffer:

50ml De-ionised Formamide (Sigma)

25ml 20X DEPC-SSC

50μl Heparin

100μl Tween 20 (Sigma)

20ml DEPC H_2O

Adjust to pH6 with 1M citric acid

Adjust final volume to 100ml with DEPC H_2O

Store at -20°C

1X TBST:

1X TBS

Add 0.1% Tween 20 (Sigma) before use

10X TBS:

40g NaCl

1g KCl

125ml 1M Tris.HCl (pH7.5)

Adjust to 500ml with dH_2O

SSC/FA/Tween 20: (prepared fresh each time, not autoclaved)

5ml 20X SSC

25ml Formamide (Sigma)

50μl Tween 20 (Sigma)

Adjust to 50ml with dH_2O

MAB:

11.6g Maleic acid (Sigma)

8.8g NaCl

Add 800ml dH_2O

Adjust with NaOH to pH7.5

Adjust to 1L with dH_2O

MABT:

MAB + 0.1% Tween 20 (Sigma)

NTMT: (prepared fresh each time, not autoclaved)

1ml 5M NaCl

2.5ml 1M $MgCl_2$

50μl Tween 20 (Sigma)

5ml 1M Tris.HCl (pH9.5)

500μl 200mM Levamisole (final conc.: 2mM)

Add 41ml dH_2O

10% Blocking Solution:

Dissolve 10g blocking reagent (Roche) in 100ml MAB

Autoclave and store at -20°C

Add 0.1% Tween 20 (Sigma) before use

2.5.2.2 Procedure for Whole-mount *In-situ* Hybridisation

Embryos were prepared by first fixing in 4% PFA at 4°C overnight. PFA was removed and embryos were washed with PBT twice for 10 minutes each. Embryos were dehydrated through a methanol series of increasing concentrations (25%, 50%, 75% methanol/PBT and twice with 100% methanol) for 10 minutes each, then stored in 100% methanol at -20°C.

Embryos were rehydrated through a decreasing methanol concentration series (75%, 50%, 25% methanol/PBT and twice with PBT) for 10 minutes each, then bleached with 6% hydrogen peroxide in PBT for 30 minutes. Embryos were washed three times in PBT for 5 minutes each and digested in 20μg/ml proteinase K in PBT for appropriate time (2 minutes for E10.5, 4 minutes for E11.5, 6 minutes for E12.5, 8 minutes for E13.5 and 10 minutes for E14.5). Embryos were washed in freshly prepared PBT/glycine twice for 5 minutes each. Then further washed for 5 minutes each with PBT twice, RIPA buffer three times and PBT three times. Embryos were re-fixed in freshly prepared 4% PFA/0.2% glutaraldehyde in PBT for 20 minutes on ice and then washed three times with PBT for 5 minutes each. Embryos were placed in 2ml screw cap vials and washed in hybe buffer/PBT (1:1) for 10 minutes, then with hybe buffer for 10 minutes and then incubated in hybe buffer for 3 hours at 65°C. Probes were diluted in hybe buffer (1:200 – 1:1000 dilution) and denatured at 80°C for 3 minutes. Probes were transferred to corresponding embryos and hybridised at 65°C overnight.

The following day, the probe was removed and embryos were washed twice with hybe buffer at 65°C for 30 minutes each, then once in 1:1 hybe buffer/SSC/FA/Tween 20 at 65°C for 5 minutes. Embryos were then subjected to a series of SSC/FA/Tween 20 washing steps at 65°C: 5 minutes twice, 10 minutes three times, 30 minutes once and 1 hour five times. Embryos were cooled down to room temperature and washed with 1:1 SSC/FA/Tween20/TBST for 10 minutes, then washed twice with TBST and twice with MABT for 10 minutes each. Embryos were transferred into glass vials and pre-blocked with 10% blocking solution/MABT for 1 hour, then incubated with anti-DIG AP fab fragment (1:5000 dilution; Roche) in 1% blocking solution at 4°C overnight while gently shaking.

The next day, the antibody was removed and embryos were washed three times in TBST for 5 minutes each and a further five times in TBST for 2 hours each. Embryos were left gently shaking in TBST at 4°C overnight.

Finally, embryos were prepared for staining by washing three times in NTMT for 20 minutes each. Embryos were stained in BM Purple AP substrate (Roche) with 1% Tween 20 and 2mM Levamisole in the dark. Embryos were checked every 30 minutes until the stain developed. Embryos were fixed using 4% PFA.

2.5.3 RNA *In-situ* Hybridisation on Wax Sections

Wild-type E14.5 embryos were used for *in-situ* hybridisation on wax sections. Embryos were dissected in PBS and the head of the embryos was used for genotyping. Gonads were fixed in 4% PFA at 4°C overnight. PFA was removed and gonads were washed three times in PBT for 5 minutes each on ice. Gonads were dehydrated through a methanol series of increasing concentrations (30%, 50% and three times 80% methanol/PBT) for 10 minutes each and then stored in 80% methanol at -20°C until ready for sectioning.

2.5.3.1 Preparation of Sections

Gonads were hand-processed for paraffin wax-embedding, firstly by 100% methanol, three times for 5 minutes each on ice, then twice in xylene for 10 minutes each at room temperature before three times in pre-warmed wax at 58°C for 10 minutes each. Gonads were embedded in wax immediately after processing. 10μm transverse sections were cut on a microtome (Leica 2030) by Dr Robert Watson and sections were floated out into a 42°C waterbath and collected onto superfrost glass slides (BDH). Sections were air-dried for at least one hour and stored at 4°C until ready for hybridisation.

2.5.3.2 Solutions for RNA *In-situ* Hybridisation on Wax Sections

Pre-hybridisation Solution:

25ml De-ionised Formamide (Sigma)
12.5ml 20X SSC (3M NaCl, 0.3M Na_3Citrate, DEPC-treated)
1g Blocking Powder (Boehringer Mannheim)
0.5ml 10% Trition X-100 (Sigma)
2.5ml 10% CHAPS (Sigma)
1ml Yeast RNA (50mg/ml; Sigma)
0.5ml 0.5M EDTA
250μl Heparin (10mg/ml; Sigma)
Adjust final volume to 50ml with DEPC H_2O

Hybridisation Solution:

Heat 2μl RNA probe in 100μl pre-hybridisation solution to 80°C for 3 minutes and add to slides.

Post-hybridisation Solution:

25ml Formamide (Q-BioGene)

12.5ml 20X SSC

0.5ml 10% Trition X-100 (Sigma)

2.5ml 10% CHAPS (Sigma)

Adjust final volume to 50ml with dH_2O

TNT:

2.5ml 1.5M Tris (pH7.5)

1.5ml 5M NaCl

0.5ml 10% Trition X-100 (Sigma)

Adjust final volume to 50ml with dH_2O

Blocking Solution:

7.5ml Sheep Serum (heat inactivated)

1g BSA (Sigma)

2.5ml 1M Tris (pH7.5)

1.5ml 5M NaCl

0.5ml 10% Trition X-100 (Sigma)

Adjust final volume to 50ml with dH_2O

Antibody Solution:

Add 1:2000 anti-DIG AP fab fragment (Roche) to fresh blocking solution (5μl antibody per 10ml blocking solution).

NMT:

5ml 1M Tris (pH7.5)

1ml 5M NaCl

2.5ml 1M $MgCl_2$

Adjust final volume to 50ml with dH_2O

Staining Solution:

3.3μl/ml NBT (Roche) and 3.5μl/ml BCIP (Roche) in NMT

2.5.3.3 Procedure for RNA *In-situ* Hybridisation on Wax Sections

Slides were placed in coplin jars and sections were de-waxed with two 10-minute washes in xylene, then washed in two 5-minute washes of 100% methanol and rehydrated through a decreasing methanol concentration series (70%, 50%, 30% methanol/PBT and three times with PBT) for 5 minutes each. Sections were fixed in 4% PFA for 10 minutes, washed three times with PBT for 5 minutes each, then digested in 10μg/ml Proteinase K in PBT for 5 minutes and further washed three times with PBT for 5 minutes each. Slides were transferred to a humidified chamber and incubated with 1ml pre-hybridisation solution/slide for 4 hours at 65°C. The pre-hybridisation solution was removed and replaced with 200μl hybridisation solution/slide (plus a coverslip to prevent evaporation). Slides were incubated at 65°C overnight.

The following day, coverslips were removed and slides were transferred to a coplin jar for a series of post-hybridisation washes at 65°C. Beginning with two 10-minute washes in 100% post-hybridisation solution, one 10-minute wash in 75% post-hybridisation solution/2X SSC, one 10-minute wash in 50% post-hybridisation solution/2X SSC and one 10-minute wash in 25% post-hybridisation solution/2X SSC. Then two 30-minute washes in 2X SSC, 0.1% CHAPS, two 30-minute washes in 0.2X SSC, 0.1% CHAPS and, finally, three 5-minute washes in TNT at room temperature. Slides were transferred to a humidified chamber and incubated with 1ml blocking solution/slide for 1 hour at 4°C. The blocking solution was removed and slides were further incubated with blocking solution for two times, 1 hour each at 4°C. 1ml antibody solution was placed on each slide and incubated at 4°C overnight.

The next day, the antibody was removed and slides were transferred to a coplin jar for a series of post-antibody washes at room temperature. This consisted of four 10-minute washes in TNT, 0.1% BSA and four 10-minute washes in NMT, 0.1% Triton X-100. Slides were then placed in staining solution and checked every 30 minutes until the stain developed. Sections were fixed in 4%PFA at 4°C overnight.

The next day, slides were rinsed three times in PBS for 10 minutes each and counter-stained with nuclear fast red (Vector Laboratories) for 30 seconds. Staining was stopped by rinsing slides under tap water and then washing with 95% methanol for 5 seconds, followed by 100% ethanol for 5 seconds. Slides were cleared by washing in xylene twice for 5 minutes each and then mounted using DPX mountant for microscopy (BDH).

2.6 Tissue Culture

2.6.1 Cell Culture

TM4 cells were purchased from ECACC and cultured in Ham's F12 and DMEM (1:1) medium (Gibco) supplemented with 2mM glutamine (Gibco), 5% horse serum (Gibco) and 2.5% FCS (Gibco). Cells were maintained in 75cm^2 canted neck tissue culture flasks (Iwaki) and incubated at 37°C in 5% CO_2. 0.25% trypsin/EDTA (Gibco) was used to detach cells from the bottom of the flask when cultures were 70-80% confluent and required to be split. For experiments, cells were plated in 35mm tissue culture dishes (Iwaki) and cultured as normal.

2.6.2 Gonad Culture

The gonad plus mesonephros (urogenital ridge) were dissected from each embryo and the head was used for genotyping. Isolated gonads and mesonephroi were

placed on 1.5% agar blocks (Difco Bacto Agar in DMEM) in 35mm tissue culture dishes (Iwaki). 1ml culture medium (DMEM supplemented with 10% FCS and 100% penicillin and streptomycin; Gibco) was added to the bottom of each dish. Dishes were incubated at 37°C in 5% CO_2. Cultures were replenished with fresh culture medium after 4-6 hours and every 12 hours thereafter until the desired stage.

2.7 siRNA

2.7.1 Production of siRNAs

siRNAs were synthesised using the *Silencer* siRNA construction kit (Ambion). All reagents were supplied in kit (unless otherwise stated).

Following manufacturer's recommendations, siRNA oligonucleotide templates were designed for *Sry* and *Sox9* (sequences were checked against the mouse genome database to eliminate any sequences with significant homology to other genes). All siRNA oligonucleotide templates were purchased from Invitrogen as HPLC purified compounds.

Sense and anti-sense siRNA oligonucleotide templates were prepared for transcription by hybridisation to the T7 promoter primer (Ambion) at 70°C for 5 minutes, then at room temperature for 5 minutes. For each siRNA, two transcription reactions were assembled at room temperature to synthesize the sense and anti-sense RNA strands of the siRNA separately. For each transcription reaction, the following components were mixed in the order shown:

Component (Ambion)	Amount
Sense or anti-sense siRNA template	2μl
Nuclease-free water	4μl
2X NTP mix	10μl

10X T7 Reaction buffer	2μl
T7 Enzyme mix	2μl

Components were mixed gently and incubated at 37°C for 2 hours. Sense and anti-sense transcription reactions were combined into a single tube and further incubated at 37°C overnight.

The following day, the dsRNA was digested with RNase and DNase by adding the following reagents to the tube:

Component (Ambion)	Amount
Digestion buffer	6μl
Nuclease-free water	48.5μl
RNase	3μl
DNase	2.5μl

Contents were mixed gently and incubated at 37°C for 2 hours. 400μl siRNA binding buffer (Ambion) was added to the nuclease digestion reaction and incubated at room temperature for 5 minutes. The siRNA was purified using filter cartridge columns (Ambion) and eluted in 100μl nuclease-free water (Ambion). siRNAs were stored at -80°C until they were prepared for transfection.

2.7.2 siRNA Protocol

Cells/gonad cultures were rinsed in PBS and then incubated with Opti-MEM I (Gibco) for 1 hour. During this time, the transfection mix was prepared as follows:

For every 100μl of transfection mix: 3μl siRNA (60pmol) was mixed with 50μl Opti-MEM I (Gibco). In a separate tube, 3μl Oligofectamine (Invitrogen) was mixed with 12μl Opti-MEM I (Gibco). Both tubes were incubated at room temperature for

10 minutes, then the two mixes were combined and incubated at room temperature for a further 20 minutes. 32μl Opti-MEM I (Gibco) was added to the combined mix to bring the total volume to 100μl. This mixture was then directly overlaid onto cells/gonad culture and incubated at 37°C in 5% CO_2 for 4 hours.

After the 4-hour incubation, 1ml fresh culture medium was added to each dish. Ham's F12 and DMEM (1:1) medium (Gibco) supplemented with 2mM glutamine (Gibco), 5% horse serum (Gibco) and 2.5% FCS (Gibco) was used for cells. DMEM (Gibco) supplemented with 10% FCS (Gibco) and 100% penicillin and streptomycin (Gibco) was used for gonad cultures. Cells/gonad cultures were then grown as described in section 2.6.

2.8 X-Gal Staining

Following culture, *Sox8*$^{+/+}$ gonads were rinsed in PBS and fixed in fixative for 30 minutes at 4°C. Gonads were then washed three times in β-Gal wash solution at room temperature for 20 minutes each. Subsequently, gonads were stained in freshly prepared X-Gal staining solution overnight at 37°C in the dark and post-fixed in fixative the following day.

Fixative:

	Final concentration
Glutaraldehyde	0.2%
Phosphate buffer (pH7.3)	0.1M
$MgCl_2$	2mM
EGTA	5mM

Phosphate Buffer:

0.1M disodium hydrogen orthophosphate (Na_2HPO_4) : 0.1M sodium dihydrogen orthophosphate (NaH_2PO_4).

β-Gal Wash Solution:

	Final concentration
Phosphate buffer (pH7.3)	0.1M
$MgCl_2$	2mM
Sodium deoxycholate	0.1%
Nonidet-P40	0.02%
BSA	0.05%

X-Gal Staining Solution:

1) 25mg X-Gal was dissolved in 0.5ml dimethyl formamide.

2) 25ml wash solution plus:

6mg spermidine

41mg $K_3Fe(CN)_6$

52.5mg $K_4Fe(CN)_6$

0.4ml 0.09% NaCl

Two solutions were mixed together and filtered before use.

2.9 Real-time PCR

Samples were run on a Roche Light-cycler instrument and analysed using the Roche Light-cycler software package (version 3.5.3).

2.9.1 Sequence-Specific Hybridisation

For sequence-specific real-time PCR, the light-cycler fast start DNA master hybridisation probes kit (Roche) was used. Sequence-specific hybridisation probes were designed and purchased from TIB MOLBIOL, Berlin.

Each reaction was generally set-up as follows:

10X DNA fast start mix	2μl
$MgCl_2$ (25mM)	2.4μl
Sense primer (10ng/μl)	1μl
Anti-sense primer (10ng/μl)	1μl
Fluorescein probe (2ng/μl)	1μl
Light-cycler Red 640 probe (2ng/μl)	2μl
cDNA template	1μl
dH_2O	9.6μl
	20μl

Light-cycler Programme:

95°C	10 minutes	No acquisition		
95°C	3 seconds	No acquisition	}	
55°C	10 seconds	Single acquisition	}	x 45 cycles
72°C	16 seconds	No acquisition	}	
40°C	30 seconds	No acquisition		

2.9.2 SYBR Green System

For SYBR green real-time PCR, the light-cycler fast start DNA master SYBR green I kit (Roche) was used. Primers were purchased from Invitrogen as lyophilised desalted compounds.

Each reaction was generally set-up as follows:

10X SYBR Green fast start mix	2μl
$MgCl_2$ (25mM)	2μl
Sense primer (10ng/μl)	1μl
Anti-sense primer (10ng/μl)	1μl
cDNA template	1μl
dH_2O	13μl
	20μl

SYBR Green Light-cycler Programme:

95°C	10 minutes	No acquisition		
95°C	3 seconds	No acquisition	}	
55°C	5 seconds	No acquisition	}	
72°C	18 seconds	No acquisition	}	x 45 cycles
82°C	30 seconds	Single acquisition	}	
95°C	0 seconds	No acquisition		
65°C	45 seconds	No acquisition		
95°C	0 seconds	Continuous acquisition at a slope of 0.1°C/sec.		
40°C	30 seconds	No acquisition		

2.10. Microscopy and Associated Equipment

Dissections of embryos were done with the aid of a Leica MZ12$_5$ microscope.

Photographs of whole-mount embyros, gonads and gonad cultures were taken using a Zeiss AxioCam HRc digital coloured camera attached to a Zeiss Stemi SV11 Apo microscope. Images were captured using Zeiss Axiovision software.

Images of wax sections were captured using a Photometrics Coolsnap HQ digital monochrome camera (Ropern Scientific, Tucson, USA) attached to a Zeiss Axioplan II microscope equipped with a filterwheel in the transmission lightpath containing Andover rgb colour additive filters (L.O.T. Oriel Ltd, UK). Sequential images were captured using the red, green, and blue filters and these images were then merged together to provide the coloured image. Image capture, control of the filterwheel and merging of images were controlled by in-house scripts written for IP Lab Spectrum (Scanalytics corp, Fairfax VA, USA). Dr Paul Perry wrote all scripts for capturing digital images with IP Lab software.

Chapter 3 Affymetrix Microarray Analysis of Gonadal Development

Chapter 3 Affymetrix Microarray Analysis of Gonadal Development

3.1 Introduction

The work presented in this chapter compares the gene expression profile of male and female developing mouse gonads using Affymetrix GeneChip arrays. Affymetrix GeneChip arrays were used as they provided the largest, most complete set of mouse transcripts available at the time. In this study, microarrays were used as a starting point to screen for possible candidates. These candidates would later be verified by other methods, such as *in-situ* hybridisation and real-time PCR. The outcome of this study may lead to the identification of novel candidates involved in gonadal development and may provide further insight into the male and female sex determination pathways.

3.2 Affymetrix GeneChip Arrays

The Affymetrix murine genome U74 set of microarrays consists of three GeneChip probe arrays (A, B and C). Each array contains oligonucleotide probes specific for approximately 6,000 known mouse genes and approximately 6,000 EST sequences from the UniGene database (Build 74). Different regions of certain genes are represented in different probe sets on the microarray to ensure a more accurate data set. In a preliminary attempt (screen 1), the Affymetrix murine genome U74 version 1 A and B chips were used. Unfortunately, these arrays contained a percentage of sequences assigned in the wrong direction due to submission errors from GenBank and UniGene. Therefore, Affymetrix produced a mask file containing the ambiguous sequences so that they could be excluded from the data set. This problem was rectified in the database and the Affymetrix murine genome U74 version 2 was later produced. Consequently, the A, B and C chips from this new and corrected version were used in screen 2.

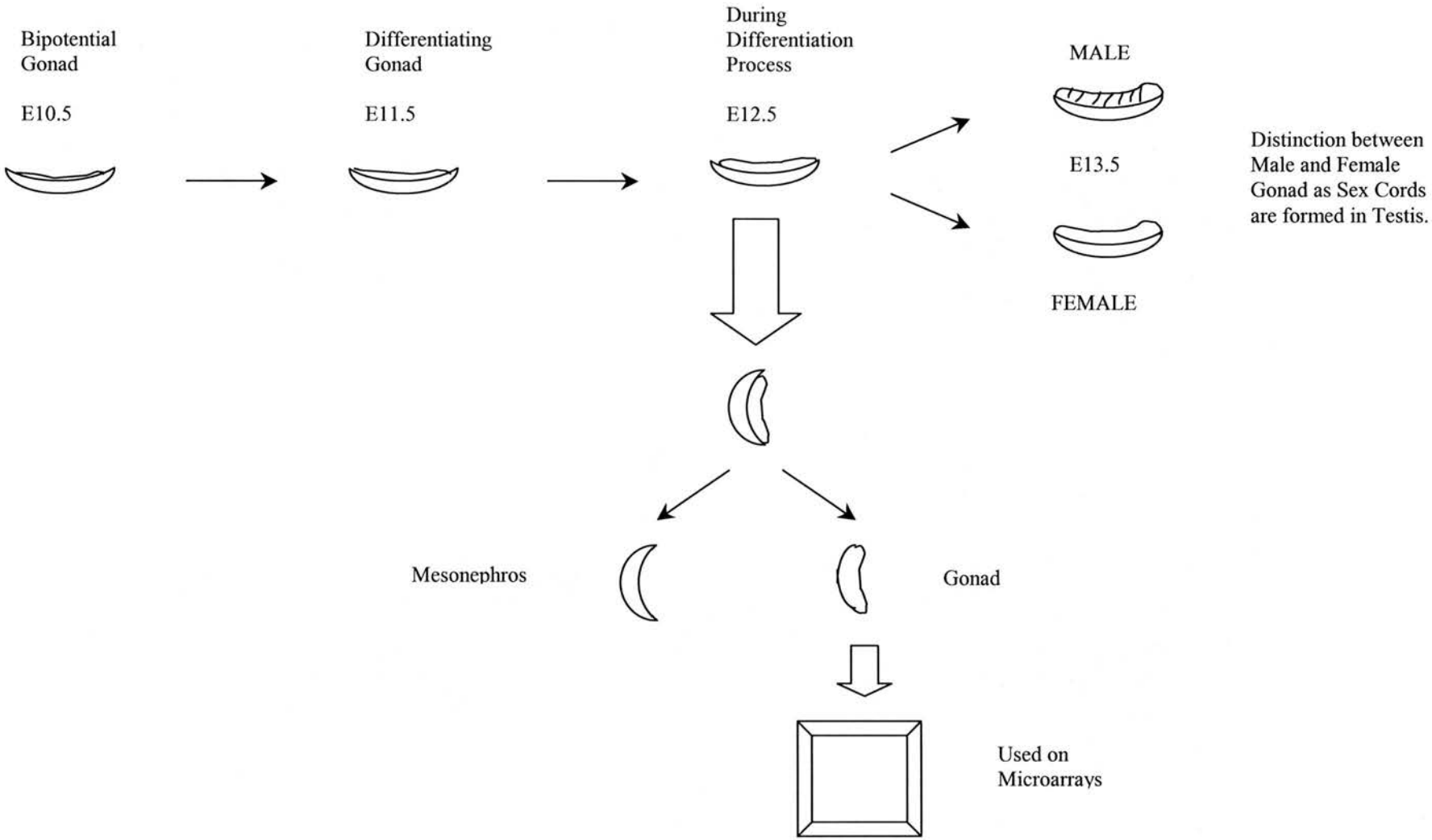


Figure 3.1 Schematic Diagram of the Developing Gonad
Stages of the developing gonad illustrating the use of E12.5 gonads for our microarray analysis.

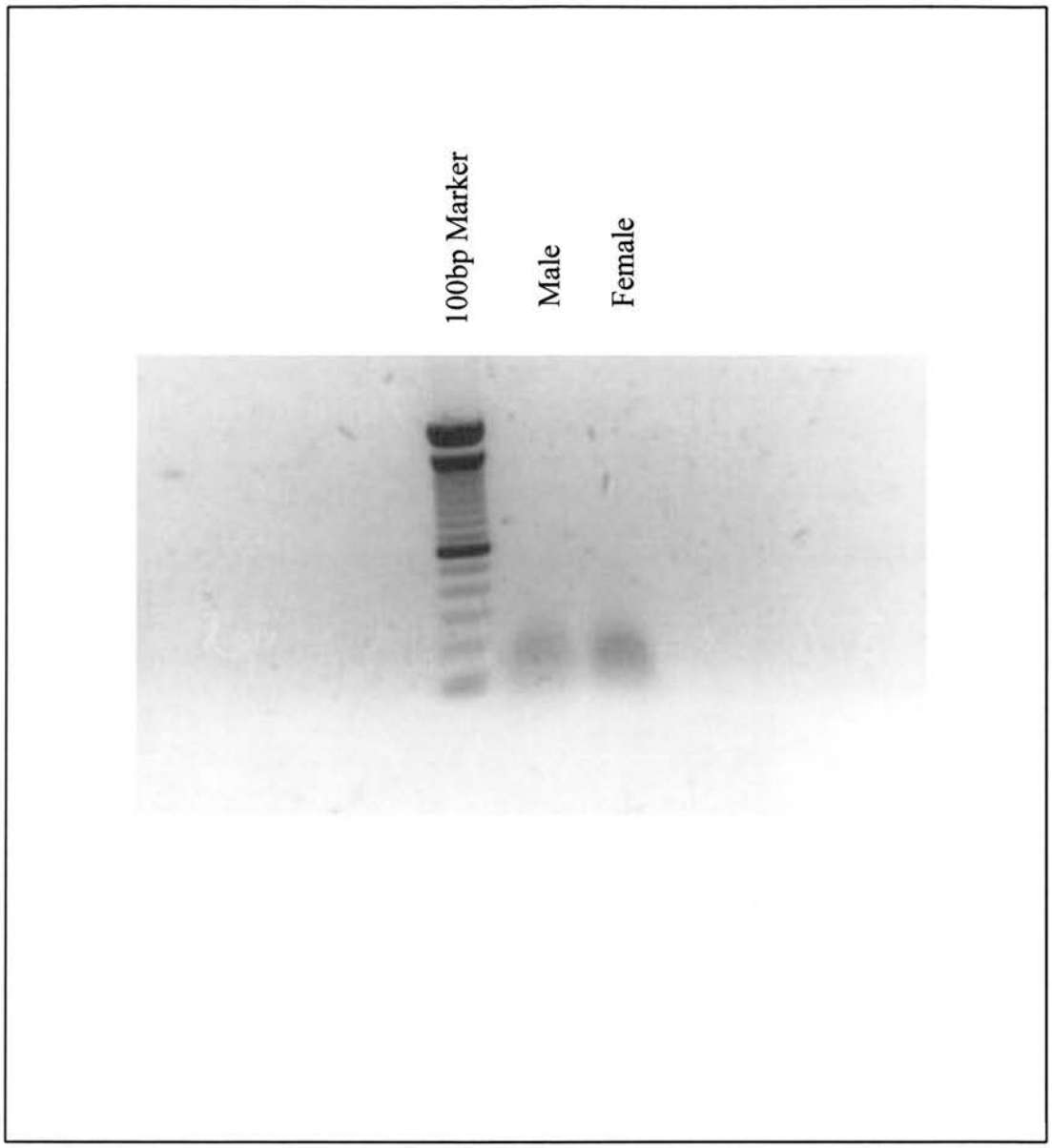


Figure 3.2 Fragmented cRNA Samples for Screen 1

16µl of labelled male and female cRNA samples were fragmented. 1µl of each fragmented cRNA sample was loaded onto an agarose gel prior to hybridisation onto microarrays. Results show both cRNA samples were fragmented and of significant quality.

3.3 Screen 1

3.3.1 Experimental Design

To identify new genes involved in gonadal development and sex determination, wild-type male and female gonads were quantitatively staged (as described in section 2.3.3) and dissected from mouse embryos at E12.5 (28 to 30 tail somites). Induction of sex determination is driven by the expression of *Sry,* which peaks at E11.5. A screen at E12.5 (Figure 3.1), shortly after the initiation of sex determination, should therefore lead to the isolation of genes at a very early stage of the sex determination process rather than genes expressed in the already differentiated ovarian/testicular tissues. In contrast to screens performed in other laboratories, we decided to remove the mesonephros from each gonad. This we hoped would increase the sensitivity of the screen and focus on genes directly involved in the sex determination process. Equal amounts of RNA were hybridised to each GeneChip to reduce the variability between samples and to allow the data sets to be compared. Figure 3.2 shows an aliquot of the male and female fragmented cRNA samples run on an agarose gel to ensure the cRNA was of good quality before hybridisation to microarrays.

3.3.2 Normalisation of Arrays

The Affymetrix GeneChip arrays were scanned and analysed with Affymetrix MAS 4.0.1 software to identify transcripts that were classified as being 'absent', 'present' or 'marginal'. In addition to the unique probes represented on the arrays, each array contains control probes (such as β-actin and GAPDH) that account for the quality of the biotinylated RNA samples and hybridisation protocol. A report file detailing the statistics of each array can be generated. The key features of this report are the percentage of probe sets which are 'present' and the 3'-5' GAPDH ratio. As a guideline, Affymetrix propose users to expect approximately a minimum of 30% of

the probe sets to be reported as 'present' and a 3'-5' GAPDH ratio no greater than 3 shows full-length transcripts were hybridised to arrays.

In our studies, 31.5% of the probe sets was scored as 'present' (after the mask file was applied) using the male sample on the MG_U74A chip and the 3'-5' GAPDH ratio was 1.77 (Appendix 4a). The corresponding MG_U74A chip hybridised with the female sample was scored as 28.5% probe sets 'present' (after the mask file was applied) and the 3'-5' GAPDH ratio was 2.45 (Appendix 4b). Considering only 5μg of RNA (instead of the recommended 15μg) was used for each sample, the statistics obtained fell approximately within the manufacturer's guidelines. Hybridisation to the control probes was also checked and found to be 'present' in both arrays. Based upon these findings, the two GeneChips were of comparable quality. Next the expression data from the two arrays was normalised to account for the individual hybridisation intensities. Thus, meaningful comparisons can be made between the male and female arrays.

Due to the positive percentages and ratios obtained from the A chips, suggesting that the male and female RNA samples were of reasonable quality, progression onto the B chips occurred. The male and female hybridisation cocktails from the A chips were re-used for the B chips. The MG_U74B chip hybridised with the male sample was scored as 12.1% probe sets 'present' (after the mask file was applied) and the 3'-5' GAPDH ratio was 2.22 (Appendix 4c). The corresponding MG_U74B chip hybridised with the female sample was scored as 14.7% probe sets 'present' (after the mask file was applied) and the 3'-5' GAPDH ratio was 3.53 (Appendix 4d). The percentage of probe sets scored as 'present' was extremely low in both arrays as there were over 40% incorrect genes and ESTs represented on the B arrays, which dramatically reduced the number of correct genes and ESTs in both data sets. Applying the mask files removed these incorrect sequences and they were classified as percentage 'absent' probe sets, thus accounting for the low percentage of 'present' probe sets obtained. The 3'-5' GAPDH ratio has increased in both male and female B arrays compared to their A array equivalent. This was caused by the re-use of the

hybridisation cocktails as RNA quantity and quality will be reduced. Therefore, a higher 3'-5' GAPDH ratio was to be expected. To ensure good hybridisation of the samples to the arrays, the hybridisation to the control probes was checked. Both arrays reported these to be 'present'. Subsequently, the expression data from these two B arrays was normalised in the same way as for the A arrays, thus allowing meaningful comparisons to be made between both sets of male and female arrays.

3.3.3 Data Analysis

The differences in gene expression profiles between male and female gonads were first analysed by scatter plot analysis. This graphical representation gives an indication of the overall similarities or differences between the two RNA samples, where each point on the graph represents a single annotated gene or EST on the microarray. The position of each point on the scatter plot is determined by the difference in signal intensity of this transcript from one RNA sample to the other. Figure 3.3 shows the scatter plot analysis (GeneSpring 6.1, Silicon Genetics) of the female gonadal pool of RNA versus the male gonadal pool of RNA (chips U74A and U74B with mask file applied). The signal intensities of the female array (X axis) are compared to the signal intensities of the male array (Y axis). The distance with which a point lies from the line of no change (line of intercept passing through the point of origin) represents the differential expression of a gene between the male and female samples. For example, a point lying close to the line of no change represents a gene that is expressed at similar levels in both pools of RNA, whereas a point lying away from the line signifies a gene to be differentially expressed between the two pools of RNA. Depending upon the point lying closer towards the X or Y axis, the sex specific expression of the gene can be easily distinguished as female specific points will be found closer to the X axis and male specific points will be found closer to the Y axis. The location of a point within the scatter plot, in relation to the origin, indicates its level of expression in relation to the other genes.

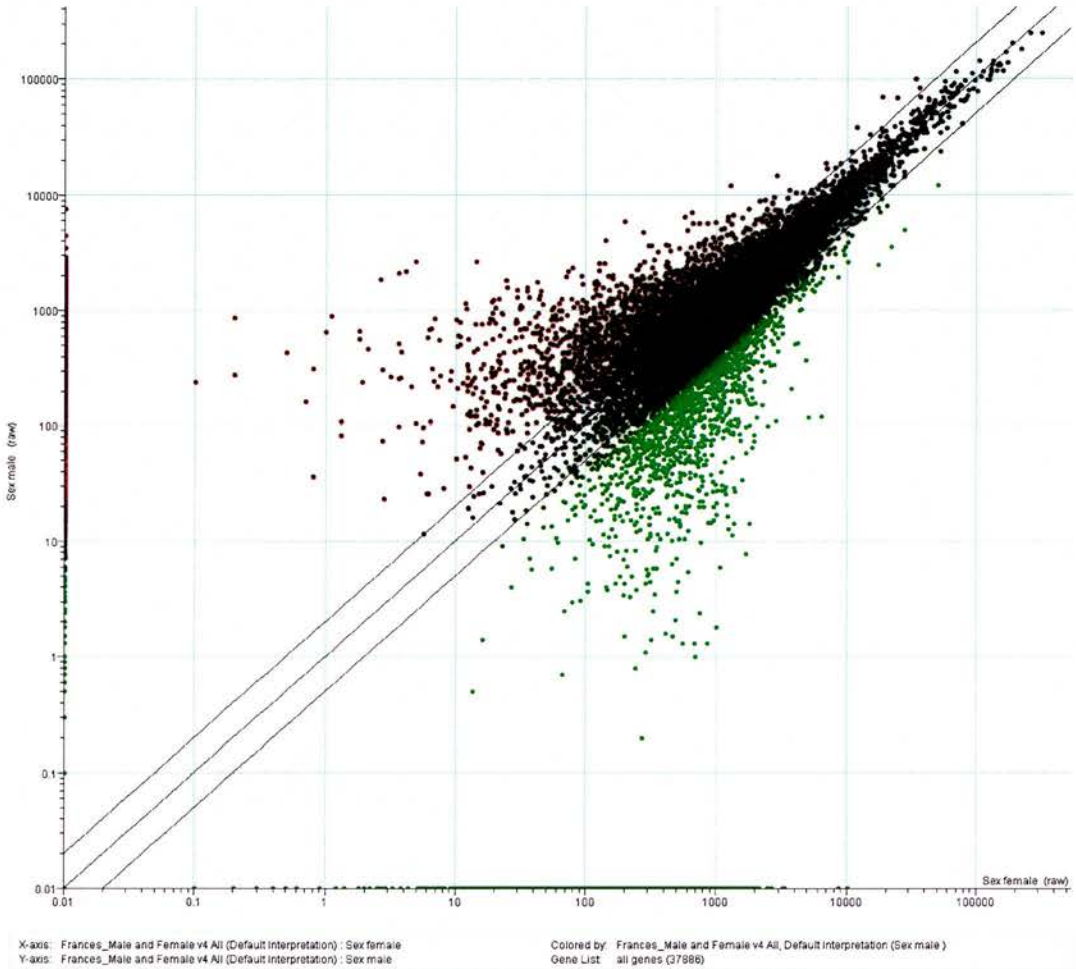


Figure 3.3 Scatter Plot Analysis of the Gene Expression Profile of Wild-type E12.5 Gonads

Each point represents the normalised expression level of an individual transcript within male and female RNA samples. Genes that are more highly expressed in male lie closer to the Y axis and genes that are more highly expressed in female lie closer to the X axis.

The two black lines parallel to the line of no change (one above and one below) refer to the significant expression level of each gene or EST. These lines represent a fold change of plus or minus 2 (top or bottom line respectively). Therefore, points which lie above the top line have a fold change of greater than 2 and are more highly expressed in male gonads. Whereas, points which lie below the bottom line have a fold change more than -2 and are more highly expressed in female gonads. Points that lie between these two lines have a fold change less than 2 and –2 which, in this case, deems them to be differentially expressed to an insignificant extent.

The points on the scatter plot are also colour-coded by a gradient system ranging from red to green. This colour-coded system reflects the expression levels of each gene or EST as a ratio to the median (a standard normalised value). Black points represent transcript expression levels that have not deviated from the median or have only deviated slightly from it. These black points are classed as having unchanged or marginally changed expression levels. Red and green points represent transcript expression levels that have deviated from the median. The brighter the red or green, the greater the deviation. It should be noted that the red and green colours do not represent sex specificity. It should also be noted that negative signal intensities are not accurately represented in the scatter plot as the GeneSpring 6.0 software does not allow negative values to be plotted against the median. Therefore, all negative signal intensities are considered to have a value of zero and are either plotted on the X axis or Y axis depending on their sex specificity. In summary, the colour of each point is related to signal intensity and the position of each point refers to the sex specificity.

To begin to analyse the data set, an arbitrary value of fold changes of greater than 2 or –2 was chosen to represent genes that are differentially expressed between male gonads and female gonads. Plus, to focus on the more reliable sex differentiated genes or ESTs within the data set, genes and ESTs marked as 'present' in at least one sample (male or female) were selected. Using these criteria in GeneSpring 6.1 (Silicon Genetics), 1096 genes and ESTs were found to be differentially expressed. Of the 1096 genes and ESTs differentially regulated in the data set, 521 of these were

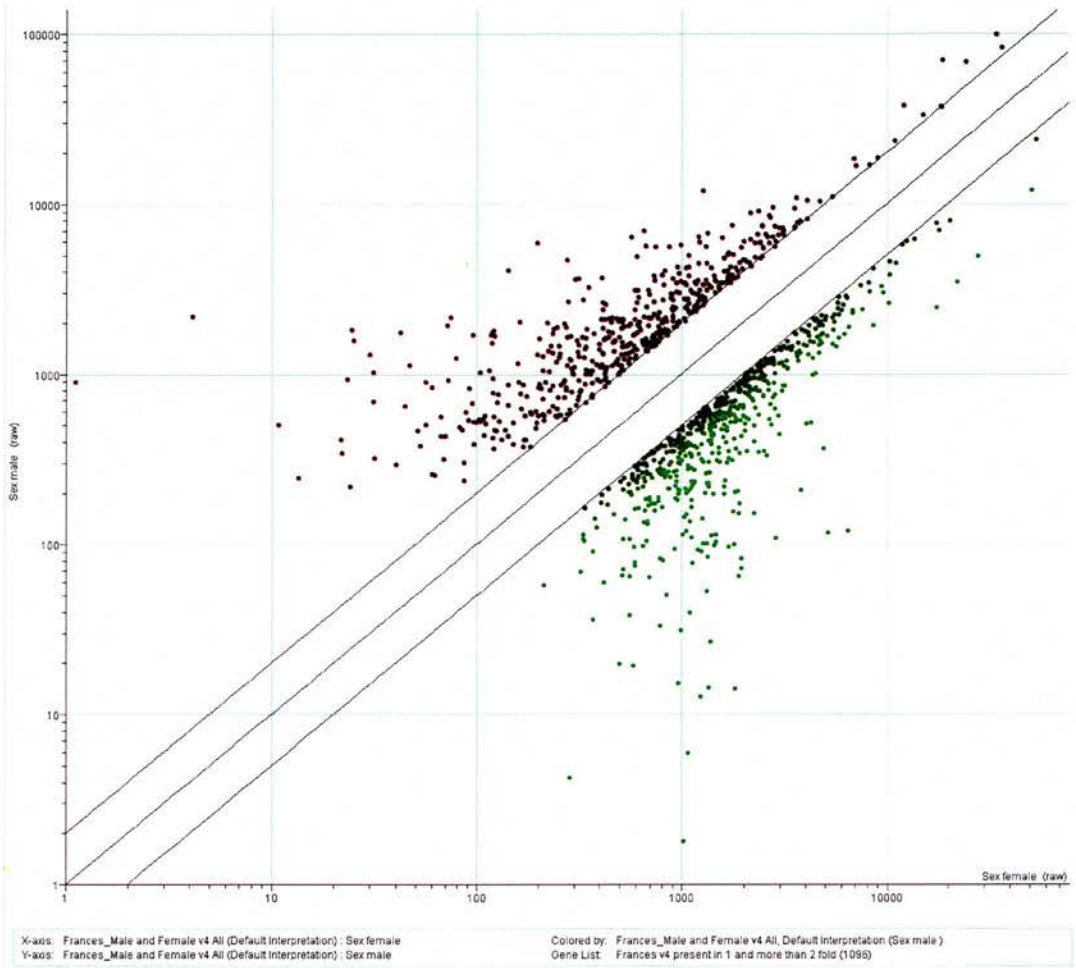


Figure 3.4 Scatter Plot Analysis of Genes Expressed Greater than 2 or –2 and Present in at Least One Sample of Wild-type E12.5 Gonads

Each point represents the normalised expression level of an individual transcript within male and female RNA samples. Genes that are more highly expressed in male lie closer to the Y axis and genes that are more highly expressed in female lie closer to the X axis.

up-regulated (more highly expressed in male) and 575 were down-regulated (more highly expressed in female). This is represented in a scatter plot in figure 3.4 (GeneSpring 6.1, Silicon Genetics). It should be noted that using GeneSpring 6.1 software (Silicon Genetics), a larger number of differentially regulated genes were obtained compared to analysis using Affymetrix MAS 4.0.1 software. This is due to the different normalisation standards within each software programme. The data set was initially analysed using Affymetrix MAS 4.0.1, then later analysed with GeneSpring 6.1 (Silicon Genetics).

3.3.4 Initial Verification of Microarray Data Set

To first verify the microarray data set was representing sex specific genes, transcripts with the highest fold changes were closely examined for known sex differentiating genes. Using Affymetrix MAS 4.0.1, one of the highest male specific fold changes was the patched homologue (*Ptc*). This gene was up-regulated by 11.4-fold and was deemed to be increased in male gonads (high-lighted in red in Appendix 5). This was found to be consistent with previous data, whereby Ptc was observed to be localised to Leydig cells and peritubular cells (Clark et al., 2000). Ptc is also known to be the receptor for the Desert hedgehog (*Dhh*) gene, which is a signalling molecule expressed by Sertoli cells. Interestingly, homozygous *Dhh*-null female mice have no obvious phenotype whereas homozygous *Dhh*-null male mice are sterile and the expression of Ptc is lost in *Dhh* mutants (Bitgood et al., 1996). Therefore, *Ptc* had already been implicated in the testicular development pathway. Interestingly, Dhh/Ptc signalling has been shown to play an important role in testis formation as it triggers Leydig cell differentiation (Yao et al., 2002). Thus, *Ptc* was the chosen candidate to assess the validity of our sex specific data set.

This was carried out by RNA *in-situ* hybridisation on eviscerated male and female E12.5 embryos using a Ptc probe (kindly donated by Walter Birchmeier's Laboratory, Berlin). Figure 3.5 shows a pair of male E12.5 gonads and a pair of

a)

b)

Figure 3.5 *In-situ* Hybridisation of Wild-type E12.5 Gonads Using Ptc Homologue Probe

a) Male E12.5 gonads showing strong staining within the interstitial region.
b) Female E12.5 gonads showing no specific staining.

female E12.5 gonads. Both were hybridised with the Ptc probe under the same conditions. A distinct staining pattern within the interstitial region of the male gonads can be observed, whereas no specific staining is seen in the female gonads. This confirmed that *Ptc* is indeed male specific and our initial screen has pulled out sex differentially expressed genes and ESTs.

It should be noted that testatin (high-lighted in blue in Appendix 5) also showed a large fold change value of 7.3 and deemed to be increased in male gonads. Testatin is a cystatin-related gene and has been shown to be expressed during early testis development in mice (Tohonen et al., 1998). Furthermore, testatin has been found to be expressed during testis cord formation in pre-Sertoli cells (the site of *Sry* action) immediately following the peak of *Sry* expression (Tohonen et al., 1998). The human orthologue of this gene has now been isolated (Eriksson et al., 2002). Similar to mouse testatin, human testatin is also specifically expressed in the testis which suggests that testatin has a role in male reproduction. Interestingly, testatin was also later identified in the subtraction hybridisation screen by Menke and Page (Menke and Page, 2002), but before that in a signal peptide differential display by Tohonen et al. (Tohonen et al., 1998). The identification of testatin, along with *Ptc*, in our initial screen further corroborates the isolation of sex differentially expressed genes and ESTs in the data set.

3.3.5 Sex Differentially Expressed Genes

Using Affymetrix MAS 4.0.1, the data set was further analysed for other sex differentially expressed genes, whereby comparisons were made against other genes found to be sex differentially expressed. Table 3.1 shows a list of genes that were examined and their expression levels within the data set. These data are also graphically represented in a scatter plot (Figure 3.6) along with heat maps of their expression between male and female (Figure 3.7) (GeneSpring 6.1, Silicon Genetics). It should be noted that colours of genes and ESTs on the scatter plot are

Affy Probe Set ID	M_Avg Diff	M_Abs Call	Diff Call	Fold Change	Descriptions
104030_at	6499	P	I	11.4	Cluster Incl AI848841:UI-M-AJ1-ahb-c-08-0-UI.s1 Mus musculus cDNA, 3 end
103296_at	5003	P	I	7.3	Cluster Incl Y18243:Mus musculus mRNA for testatin
92932_at	3730.8	P	I	~6.7	Cluster Incl X61448:M.musculus mRNA for D3 clone
102105_f_at	3801.9	P	I	5.2	Cluster Incl AI840733:UI-M-AM0-adn-b-04-0-UI.s1 Mus musculus cDNA, 3 end
101446_at	6043.1	P	I	4.3	Cluster Incl AF004428:Tumor protein D52-like 1
96926_at	1994.1	P	I	~4.2	Cluster Incl AA980164:ua31a05.r1 Mus musculus cDNA, 5 end
98480_s_at	1763.7	P	I	~4.2	Cluster Incl M32352:Mouse renin (Ren-1-d) gene, complete cds
97487_at	1857.4	P	I	~4.0	Cluster Incl X70296:Serine protease inhibitor 4
97519_at	1396.8	P	I	~3.9	Cluster Incl X13986:Mouse mRNA for minopontin
104165_at	1762.3	P	I	~3.6	Cluster Incl AJ132098:Mus musculus mRNA for Vanin-1
102070_at	3457.5	P	I	3.4	Cluster Incl AW212495:uo89g01.x1 Mus musculus cDNA, 3 end
109521_at	4125.2	P	I	~3.3	Cluster Incl AW123386:UI-M-BH2.1-apf-h-01-0-UI.s1 Mus musculus cDNA, 3 end
105120_f_at	2053.8	P	I	~3.1	Cluster Incl AI550379:vx13a04.x1 Mus musculus cDNA, 3 end
92545_f_at	1306.9	P	I	~3.0	Cluster Incl AB006361:Mus musculus mRNA for prostaglandin D synthetase, complete cds
102416_at	673.9	A	NC	~2.7	Cluster Incl M64863:Cytochrome P450
92546_r_at	1205.3	A	NC	~2.7	Cluster Incl AB006361:Mus musculus mRNA for prostaglandin D synthetase, complete cds
116952_at	1582	P	NC	~2.4	Cluster Incl AW061185:UI-M-BH1-anc-e-06-0-UI.s1 Mus musculus cDNA, 3 end
103674_f_at	695.3	P	NC	~2.2	Cluster Incl AJ006584:Mus musculus mRNA for translation initiation factor eIF2 gamma Y
106061_at	764.1	A	NC	~2.1	Cluster Incl AI842602:UI-M-AQ1-aed-b-10-0-UI.s1 Mus musculus cDNA, 3 end
100374_at	1146.2	P	NC	~2.0	Cluster Incl AL133300:Mus musculus Doublesex related transcript (Dmrt1)
104031_at	753.4	A	NC	~1.7	Cluster Incl U46155:Patched homolog
104848_at	3579.7	P	NC	1.6	Cluster Incl AI506668:vn58f01.x1 Mus musculus cDNA, 3 end

101765_at	-537.3	A	NC	~1.4	Cluster Incl X63240:Anti-Mullerian hormone
108855_at	287	A	NC	~1.2	Cluster Incl AI647526:uk40g07.x1 Mus musculus cDNA, 3 end
99861_at	305.7	A	NC	1.1	Cluster Incl X76292:Desert hedgehog homolog
92183_at	-49.7	A	NC	~1.1	Cluster Incl Z79787:Dystrobrevin
102713_at	1539.6	P	NC	1	Cluster Incl M98339:GATA-binding protein 4
100272_at	-87.4	A	NC	~-1.2	Cluster Incl Z29652:Selected mouse cDNA on the Y
105119_i_at	-1028	A	NC	~-1.2	Cluster Incl AI550379:vx13a04.x1 Mus musculus cDNA, 3 end
100006_at	845.2	P	NC	-1.2	Cluster Incl D21253:Cadherin 11
92184_at	-318.4	A	NC	~-1.3	Cluster Incl X95226:Dystrobrevin
93856_at	9436.8	P	NC	-1.4	Cluster Incl M55512:Wilms tumor homolog
93643_at	1734	P	NC	-1.8	Cluster Incl AJ243851:Mus musculus mRNA for putative LIM-homeodomain protein alpha isoform
103238_at	-197.4	A	D	~-3.2	Cluster Incl M89797:Wingless-related MMTV integration site 4

Table 3.1 List of Sex Differentiating Genes that were Examined within our Data Set

Average Difference (Avg Diff) and Absolute Call (Abs Call) are the male values obtained. Difference Call (Diff Call) and fold change are values obtained when comparing male against female. Female has been set as the baseline, therefore all positive fold changes are male expressing and negative fold changes are female expressing.

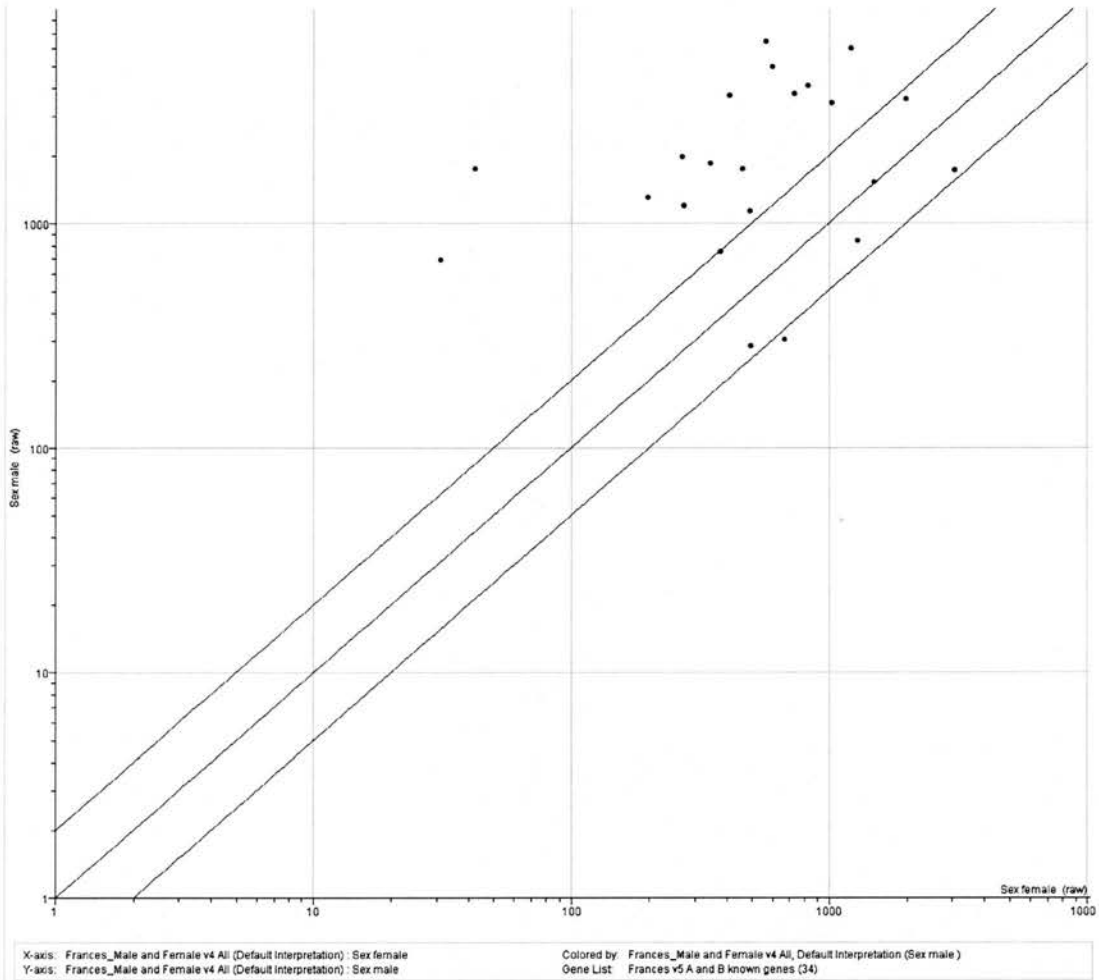


Figure 3.6 Scatter Plot Analysis of Sex Differentially Expressed Genes that were Examined in our Data Set

Each point represents the normalised expression level of an individual transcript within male and female RNA samples. Genes that are more highly expressed in male lie closer to the Y axis and genes that are more highly expressed in female lie closer to the X axis.

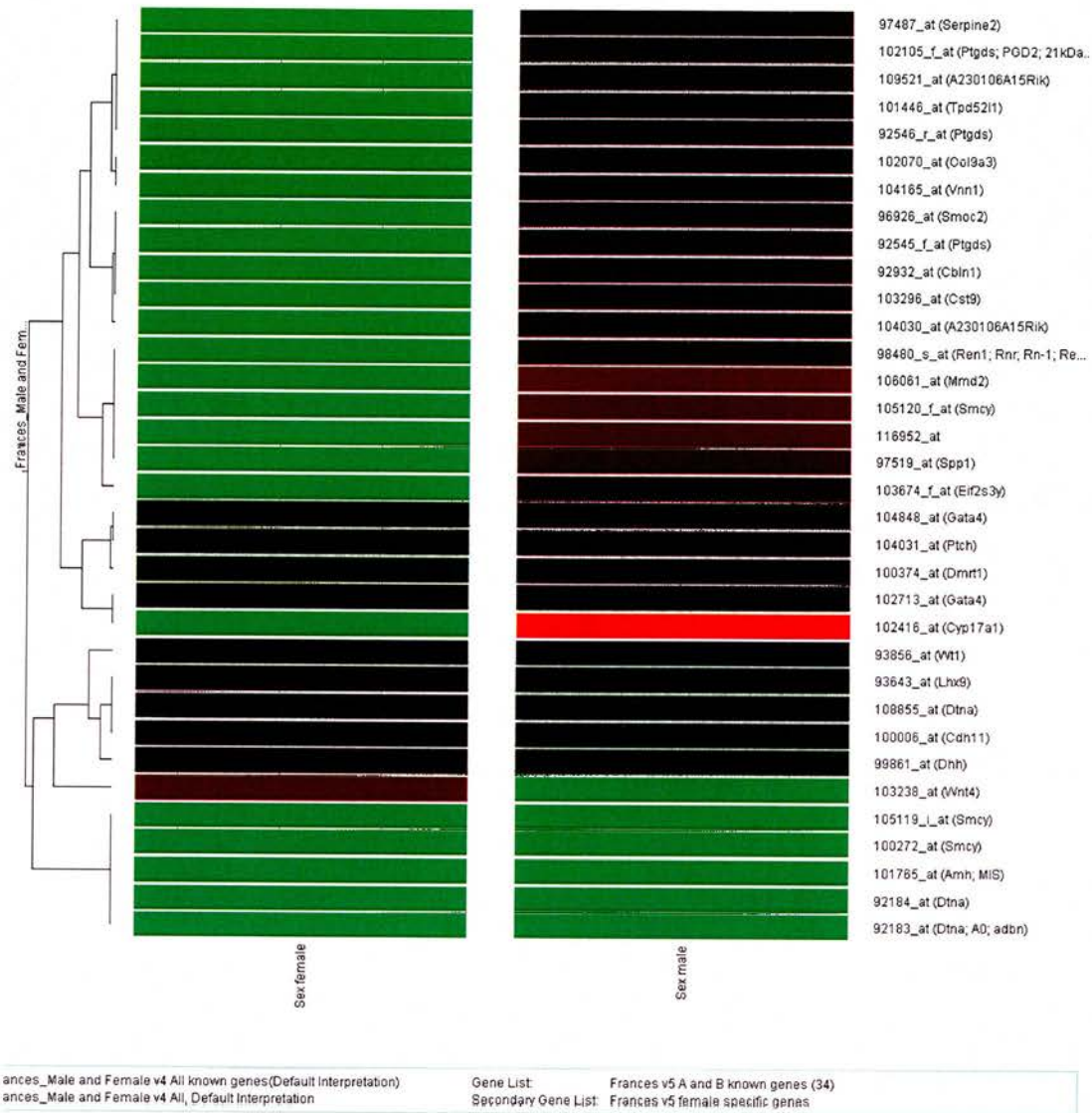


Figure 3.7 Heat Map of Sex Differentially Expressed Genes that were Examined in our Data Set

Female signal intensities (left column) and male signal intensities (right column) of genes listed on right. All signal intensities are calculated against the median. Green indicates a decrease, red indicates an increase and black indicates no deviation from the median.

independent from those on the heat maps. Heat maps are a colour-coded representation based on signal intensity, whereby a change in signal between male and female are shown by a change in red and green. Signal intensities for heat maps are calculated against the median and any deviation from the median results in colour change of green or red depending upon the sex specificity of the gene. For example, genes that are coloured green in female and red in male indicate a change up and therefore are more highly expressed in males than females. Conversely, genes that are coloured red in female and green in male indicate a change down and therefore are more highly expressed in females than males. Genes that are black in both male and female represent little or no change in signal intensities between the two sexes. Genes that are either green or red in both male and female indicate a deviation from the median, but no difference between the expression levels which means such genes are increased or decreased in both sexes. Many genes from this list were taken from a PCR-based cDNA subtraction screen, whereby E12.5 XY gonadal cDNA was compared with E12.5 XX gonadal cDNA (Menke and Page, 2002). A proportion of these genes were in agreement with our data set and the others gave a no change value. This may be due to technical differences between the two systems, such as genes on our arrays are only represented by 25 nucleotides and the RNA samples may not hybridise well to these short sequences. Thus, low signals may be detected in both male and female samples resulting in no change in expression level being measured in our system.

Not all genes from this list were expected to be differentially expressed at E12.5. For example, *Wt1* was expected to be 'present' as it first appears in the urogenital ridge at E9.0 and continues to be expressed after formation of the gonads (Shimamura et al., 1997). More noticeably, the expression of *Wt1* is not sexually differentiated at E12.5 (Morrish and Sinclair, 2002). Our results were in agreement with this as *Wt1* was found to be 'present' and not changed. This also helped to verify our data set to be stage specific at E12.5 and that the genes and ESTs represent expression levels during the sex differentiation process.

Well-known sex differentially expressed genes at E12.5 include *Sry* and *Sox9* in the male gonad and *Wnt4* in the female gonad. Unfortunately, *Sox9* was amongst the incorrectly submitted sequences and so it's expression within our data set could not be checked. *Sry* was expected to be 'present' in male but not greatly increased as it's expression peaks at E11.5 and is completely switched off by E13.0 (Hacker, Capel et al. 1995). Our data set revealed both Sry probe sets to be unchanged but one probe set was registered to be 'present' and the other was registered as 'absent'. This difference may be due to hybridisation of the RNA samples to the individual probe sets on the microarray. The expression of *Wnt4* is down-regulated in the developing male gonad after E11.5, but it persists in the developing female gonad (Vainio et al., 1999). Therefore, *Wnt4* was expected to be 'absent' in the male gonad and 'present' in the female gonad. Consequently, the difference call should be a decrease. Our data set was in agreement with this and showed a fold change value of 18.4.

Another gene predicted to be more highly expressed in male gonads than female gonads is Podocyte Expressed 1 (*Pod-1*). This is also in agreement with our findings, where our data showed a 'present call' in the male sample and an 'absent call' in the female sample. More specifically, the expression of *Pod-1* was found to be increased in male gonads by a fold change of 3.0. The expression pattern of *Pod-1* has been shown to be complex as it is dependent not only on the sex of the embryos but also on their stage of development (Tamura et al., 2001). For example, at E11.5, the expression of *Pod-1* was found to be at the same level in male and female gonads, but it became stronger in the male than the female at E13.5 (Tamura et al., 2001). This high level of *Pod-1* expression was found to persist throughout the foetal stages of development, but it slowly declined during the neonatal stages until 20 days after birth (Tamura et al., 2001), which coincides with progression of the first wave of meiosis and spermatid formation in the testis. In contrast, the expression level of *Pod-1* in the ovary was increased after birth and became stronger than the testis by three weeks after birth (Tamura et al., 2001). *Pod-1* was found to be specifically expressed in the coelomic epithelial cells, peritubular myoid and epithelial-like cells (Tamura et al., 2001). Interestingly, the expression pattern of *Pod-1* in the foetal

testis is closely related and for the most part complementary to that of *Sf1* (Tamura et al., 2001), which suggests that these two genes may interact with each other in the sex determination cascade.

As previously mentioned, two testis-specific genes discovered by a cDNA microarray screen for genes displaying sexually dimorphic expression during murine gonad development are Protease nexin-1 (*Pn-1*) and Vanin-1 (*Vnn1*) (Grimmond et al., 2000). The male-specific expression of *Pn-1* is reported to begin as early as E11.25 and found to be expressed in Sertoli cells at E12.5 (Grimmond et al., 2000). The expression of *Vnn1* has been found in pre-Sertoli cells at E11.5 and its expression pattern continues to be testis specific during embryonic development, whereas no expression is detected in female gonads (Grimmond et al., 2000). These genes were checked against our screen. *Pn-1* was found to be 'present' in the male and increased with a fold change of 4.0. *Vnn1* was found to be 'present' in the male and increased with a fold change of 3.6. Hence, our results corresponded with those from Grimmond et al (Grimmond et al., 2000).

In our data set, the *double-sex* related gene (*DMRT1*) was found to be 'present' in male gonads but 'absent' in female gonads. However, the difference (fold change) between them was only 1.9, which is considered to be below the significant threshold and the difference call was deemed to be unchanged. This may be due to the fact that, in the mouse, *Dmrt1* is still expressed in the genital ridges of both sexes at E12.5 and only becomes sexually dimorphic at E14.5, where its expression remains strong in the male sex cords but decreases to a low level in the female ovary (De Grandi et al., 2000). This specific expression pattern has led to the hypothesis that *Dmrt1* plays a role in sex determination in either sex, or with a later role in testis differentiation, or possibly both. Indeed, knockout analysis in mice demonstrated the important role of *Dmrt1* in testicular differentiation, whereby homozygous *Dmrt1*$^{-/-}$ mutants have severely hypoplastic testes with disorganised seminiferous tubules and missing germ cells (Raymond et al., 2000).

Another gene found to be 'present' in male gonads but 'absent' in female gonads, with a low detectable fold change of 1.0, is *Gata-4* (a member of the GATA family of transcription factors). The expression of Gata-4 has been detected in E11.5 gonads of both XX and XY mouse embryos, specifically in the developing somatic cell lineages (Sertoli in testis and granulosa in ovary) but not primordial germ cells (Viger et al., 1998). In addition, Gata-4 expression has been shown to be continuous in Sertoli cells throughout embryonic development but was notably down-regulated in the ovary at E13.5 (Viger et al., 1998). Hence, *Gata-4* may be involved in early gonadal development and possibly sexual differentiation. The fact that Gata-4 is expressed in both sexes and only becomes sexually dimorphic at a stage later than the gonads used in our screen may account for the small fold change value that was detected.

However, there were a few genes which are known to be differentially expressed in male and female gonads at E12.5 which were not reported in our CHIP screening results, the most notable being *Mis*. Our data set revealed *Mis* to be 'absent' in both male and female samples with a difference call of 'no change'. Although a fold change value of 1.8 was calculated, it is deemed insignificant due to the expression of the gene not being 'present' in at least one sample. The undetectable differences in some genes may be due to the oligonucleotide sequences used by Affymetrix being taken from a less specific region of that particular gene. This would result in poor hybridisation of the RNA samples to the probes on the CHIP and thus producing an overall weak signal in the data set.

3.3.6 Potential Sex Determination Candidates

Although our screen did not identify all the sex differentiated genes that were expected, the results still produced a sex specific data set. Therefore, genes and ESTs with significant fold changes (greater than 2 and –2) were still considered to be valid. In order to select the most probable male and female developmental genes and

MALE Affy Probe Set ID	A Chip F_Avg Diff	F_Abs Call	F_Diff Call	F_Fold Change	M_Avg Diff	M_Abs Call	M_Diff Call	M_Fold Change	Description
104637_at	159.9	A	D	-12.2	4804.3	P	I	12.2	Cluster Incl AI326397:mm20b10.x1 Mus musculus cDNA, 3 end
92932_at	331.8	A	D	-6.7	2991.9	P	I	6.7	Cluster Incl X61448:M.musculus mRNA for D3 clone
101446_at	985.1	A	D	-4.3	4846.3	P	I	4.3	Cluster Incl AF004428:Tumor protein D52-like 1
96336_at	1057.7	A	D	-4.3	4546.4	P	I	4.3	Cluster Incl AI844626:UI-M-AL1-ahr-a-04-0-UI.s1 Mus musculus cDNA, 3 end
96926_at	216.2	A	D	-4.2	1599.2	P	I	4.2	Cluster Incl AA980164:ua31a05.r1 Mus musculus cDNA, 5 end
96771_at	418.9	A	D	-3.9	1697.6	P	I	3.9	Cluster Incl AI006228:ua88d09.r1 Mus musculus cDNA, 5 end
93187_at	407.4	A	D	-3.6	1500.1	P	I	3.6	Cluster Incl AW048347:UI-M-BH1-akk-e-03-0-UI.s1 Mus musculus cDNA, 3 end
94208_at	253.8	A	D	-3.2	1199.9	P	I	3.2	Cluster Incl AW045202:UI-M-BH1-alh-c-08-0-UI.s1 Mus musculus cDNA, 3 end
104135_at	252.7	A	D	-3.1	1093.8	P	I	3.1	Cluster Incl AW045474:UI-M-BH1-akr-c-02-0-UI.s1 Mus musculus cDNA, 3 end
103891_i_at	399.5	A	D	-3.1	1335.2	P	I	3.1	Cluster Incl AI197161:ue51h10.r1 Mus musculus cDNA, 5 end
103451_at	498.9	A	D	-3.1	1635.2	P	I	3.1	Cluster Incl AI835159:UI-M-AQ0-aaf-e-10-0-UI.s1 Mus musculus cDNA, 3 end
93294_at	206.4	A	D	-3.1	1070.1	P	I	3.1	Cluster Incl M70642:Fibroblast inducible secreted protein
98109_at	-130.4	A	D	-2.7	567.1	P	I	2.7	Cluster Incl AI839528:UI-M-AN0-acn-e-06-0-UI.s1 Mus musculus cDNA, 3 end
96658_at	233.3	A	D	-2.7	870	P	I	2.7	Cluster Incl AI841906:UI-M-AO0-acd-e-10-0-UI.s1 Mus musculus cDNA, 3 end
103222_at	-24.1	A	D	-2.7	716.9	P	I	2.7	Cluster Incl L21671:Mouse Eps8 mRNA sequence, complete cds
97482_at	408.5	A	D	-2.5	484.8	P	I	2.5	Cluster Incl AA832565:vw43b11.r1 Mus musculus cDNA, 5 end

103009_at	-381.2	A	D	-2.5	562.6	P	I	2.5	Cluster Incl X92590:Histone cell cycle regulation defective homolog A (S. cerevisiae)
97446_at	281.3	A	D	-2.4	843.6	P	I	2.4	Cluster Incl AW125010:UI-M-BH2.1-apv-g-03-0-UI.s1 Mus musculus cDNA, 3 end
96221_at	128.7	A	D	-2.4	729.7	P	I	2.4	Cluster Incl AI606300:vm55g05.y1 Mus musculus cDNA, 5 end
92927_at	443.2	A	D	-2.4	957.8	P	I	2.4	Cluster Incl L10426:Mus musculus ets-related protein 81 (ER81) mRNA, complete cds
93327_at	202.2	A	D	-2.3	589	P	I	2.3	Cluster Incl AI842665:UI-M-AO1-aeo-a-02-0-UI.s1 Mus musculus cDNA, 3 end
100403_at	590.4	A	D	-2.3	1360.4	P	I	2.3	Cluster Incl AA839903:vw65f04.r1 Mus musculus cDNA, 5 end
97310_at	135.8	A	D	-2.3	698.7	P	I	2.3	Cluster Incl AW124318:UI-M-BH2.1-apq-d-11-0-UI.s1 Mus musculus cDNA, 3 end
98125_at	102.6	A	D	-2.2	611.8	P	I	2.2	Cluster Incl AI849193:UI-M-AJ1-agy-g-11-0-UI.s1 Mus musculus cDNA, 3 end
96357_at	189.9	A	D	-2.2	692	P	I	2.2	Cluster Incl AW212775:uo66e09.x1 Mus musculus cDNA, 3 end
104432_at	-90.1	A	D	-2.2	393	P	I	2.2	Cluster Incl AF016482:Mus musculus rho7 (rho7) mRNA, partial cds
103842_at	95.8	A	D	-2.2	628.1	P	I	2.2	Cluster Incl AJ007376:Mus musculus mRNA for DBY RNA helicase
101992_at	1236.7	A	D	-2.1	2554.1	P	I	2.1	Cluster Incl U13393:Mus musculus delta proteasome subunit mRNA, complete cds
92768_s_at	826.9	A	D	-2	1719.7	P	I	2	Cluster Incl M15268:Aminolevulinic acid synthase 2, erythroid

MALE B Chip Affy Probe Set ID	F_Avg Diff	F_Abs Call	F_Diff Call	F_Fold Change	M_Avg Diff	M_Abs Call	M_Diff Call	M_Fold Change	Description
113089_at	-931.8	A	D	-4.5	4444.3	P	I	4.5	Cluster Incl AA008996:mg99h01.r1 Mus musculus cDNA, 5 end
107566_at	860.6	A	D	-4	5692.5	P	I	4	Cluster Incl AI849532:UI-M-AH1-agt-g-12-

						0-UI.s1 Mus musculus cDNA, 3 end
111903_at	1084.1 A	D	-3.9	5050 P	I	3.9 Cluster Incl AI852203:UI-M-BH0-ajc-c-08-0-UI.s1 Mus musculus cDNA, 3 end
112900_at	351.3 A	D	-3.4	3270.2 P	I	3.4 Cluster Incl AA682034:vu13a03.s1 Mus musculus cDNA, 3 end
109521_at	825.4 A	D	-3.4	4125.2 P	I	3.4 Cluster Incl AW123386:UI-M-BH2.1-apf-h-01-0-UI.s1 Mus musculus cDNA, 3 end
112784_at	313.3 A	D	-3.1	3699.1 P	I	3.1 Cluster Incl AI839993:UI-M-AH0-acs-f-09-0-UI.s1 Mus musculus cDNA, 3 end
105120_f_at	-477.1 A	D	-3.1	2053.8 P	I	3.1 Cluster Incl AI550379:vx13a04.x1 Mus musculus cDNA, 3 end
109296_at	-506.1 A	D	-3	1792 P	I	3 Cluster Incl AU044228:AU044228 Mus musculus cDNA, 3 end
107046_at	1437.1 A	D	-2.7	3825.5 P	I	2.7 Cluster Incl AI838194:UI-M-AL0-abx-a-10-0-UI.s1 Mus musculus cDNA, 3 end
114058_at	1173.4 A	D	-2.6	3441.8 P	I	2.6 Cluster Incl AA152809:mq45d08.r1 Mus musculus cDNA, 5 end
116162_g_at	2695.5 A	D	-2.5	4405.5 P	I	2.5 Cluster Incl AI430099:md74e01.y1 Mus musculus cDNA, 5 end
115947_at	-702.8 A	D	-2.5	1285 P	I	2.5 Cluster Incl AI481075:vf93b08.x1 Mus musculus cDNA, 3 end
108008_at	589.2 A	D	-2.5	2431.4 P	I	2.5 Cluster Incl AI551848:vo92h10.x1 Mus musculus cDNA, 3 end
116952_at	-301.5 A	D	-2.4	1582 P	I	2.4 Cluster Incl AW061185:UI-M-BH1-anc-e-06-0-UI.s1 Mus musculus cDNA, 3 end
115102_at	1464.3 A	D	-2.4	3575.5 P	I	2.4 Cluster Incl AW060550:UI-M-BH1-ann-g-08-0-UI.s1 Mus musculus cDNA, 3 end
110713_at	814 A	D	-2.3	2113.8 P	I	2.3 Cluster Incl AW124288:UI-M-BH2.1-apq-b-05-0-UI.s1 Mus musculus cDNA, 3 end
109172_at	1825.1 A	D	-2.3	4237.3 P	I	2.3 Cluster Incl AI854623:UI-M-BH0-ake-e-05-0-UI.s1 Mus musculus cDNA, 3 end
114872_at	94.8 A	D	-2.2	1714.7 P	I	2.2 Cluster Incl AI467050:vb94a12.x1 Mus musculus cDNA, 3 end
114784_at	250.5 A	D	-2.2	1480.8 P	I	2.2 Cluster Incl AW124221:UI-M-BH2.1-aph-d-05-0-UI.s1 Mus musculus cDNA, 3 end

104954_at	117.3	A	D	-2.2	1776.2	P	I	2.2	Cluster Incl AI508971:vc44c02.y1 Mus musculus cDNA, 5 end
110129_at	-452	A	D	-2.2	1109	P	I	2.2	Cluster Incl AW211368:uo80b10.y1 Mus musculus cDNA, 5 end
115096_at	274.2	A	D	-2.1	4754	P	I	2.1	Cluster Incl AI844484:UI-M-AJ1-ahe-e-09-0-UI.s1 Mus musculus cDNA, 3 end
110400_f_at	-675	A	D	-2.1	845.8	P	I	2.1	Cluster Incl AI604974:vf20h09.x1 Mus musculus cDNA, 3 end
108745_at	24.8	A	D	-2.1	1585.7	P	I	2.1	Cluster Incl AI466491:vx41d08.y1 Mus musculus cDNA, 5 end
105667_at	117.5	A	D	-2	1526.1	P	I	2	Cluster Incl AI323290:mh18c05.x1 Mus musculus cDNA, 3 end

FEMALE	**A Chip**								
Affy Probe Set ID	**F_Avg Diff**	**F_Abs Call**	**F_Diff Call**	**F_Fold Change**	**M_Avg Diff**	**M_Abs Call**	**M_Diff Call**	**M_Fold Change**	**Description**
98131_at	1650.5	P	I	4.2	-97	A	D	-4.2	Cluster Incl D78646:Mouse mRNA for zeta-crystallin/quinone reductase, partial cds
104548_at	1569.3	P	I	4	347.1	A	D	-4	Cluster Incl Y15443:Mus musculus mRNA for 50C15 protein
95135_at	2241.2	P	I	4	593.5	A	D	-4	Cluster Incl AI844396:UI-M-AL1-ahp-e-05-0-UI.s1 Mus musculus cDNA, 3 end
103293_at	669	P	I	3.6	-889.8	A	D	-3.6	Cluster Incl AF079901:Mus musculus 28 kDa cis-Golgi SNARE (GS28) mRNA, complete cds
94881_at	995.3	P	I	3.5	10.2	A	D	-3.5	Cluster Incl AW048937:UI-M-BH1-amo-d-08-0-UI.s1 Mus musculus cDNA, 3 end
96605_at	806	P	I	3.2	251.1	A	D	-3.2	Cluster Incl AI787183:ui86a11.y1 Mus musculus cDNA, 5 end
98454_at	867.2	P	I	3	4.8	A	D	-3	Cluster Incl Y14771:Mus musculus mRNA for paralemmin
102058_at	1017.3	P	I	2.8	113.6	A	D	-2.8	Cluster Incl AI845667:UI-M-AQ1-ady-c-11-0-UI.s1 Mus musculus cDNA, 3 end

102925_at	1377.3	P	I	2.8	494.8	A	D	-2.8	Cluster Incl AA285446:vb82a12.r1 Mus musculus cDNA, 5 end
104023_at	739.6	P	I	2.8	215	A	D	-2.8	Cluster Incl AW060457:UI-M-BH1-anj-f-12-0-UI.s1 Mus musculus cDNA, 3 end
104648_at	927.4	P	I	2.8	187	A	D	-2.8	Cluster Incl AI844597:UI-M-AJ1-ahh-c-06-0-UI.s2 Mus musculus cDNA, 3 end
100051_at	2021.6	P	I	2.7	746.6	A	D	-2.7	Cluster Incl U17297:Erythrocyte protein band 7.2
92256_at	1009.8	P	I	2.7	254.9	A	D	-2.7	Cluster Incl AI853714:UI-M-BH0-ajq-d-07-0-UI.s1 Mus musculus cDNA, 3 end
99701_f_at	685.6	P	I	2.7	-246.1	A	D	-2.7	Cluster Incl AJ005560:Small proline-rich protein 2B
102215_at	1192.4	P	I	2.6	252.1	A	D	-2.6	Cluster Incl Y10495:Mus musculus mRNA for CDV-1R protein
92987_at	1116.1	P	I	2.6	21.6	A	D	-2.6	Cluster Incl M28383:Solute carrier family 4 (anion exchanger), member 3
93428_at	538.4	P	I	2.5	-138.6	A	D	-2.5	Cluster Incl AW121806:UI-M-BH2.3-anx-b-07-0-UI.s2 Mus musculus cDNA, 3 end
95621_at	563.4	P	I	2.5	-102.9	A	D	-2.5	Cluster Incl AA606367:vo47h02.r1 Mus musculus cDNA, 5 end

FEMALE	B Chip								
Affy Probe Set ID	F_Avg Diff	F_Abs Call	F_Diff Call	F_Fold Change	M_Avg Diff	M_Abs Call	M_Diff Call	M_Fold Change	Description
115255_at	8646.9	P	I	8.3	-1214.5	A	D	-8.3	Cluster Incl AA920710:vy49b09.r1 Mus musculus cDNA, 5 end
105754_at	10180.3	P	I	8.2	-209.3	A	D	-8.2	Cluster Incl AI846736:UI-M-AO1-ael-b-02-0-UI.s1 Mus musculus cDNA, 3 end
113914_at	22051.3	P	I	4.9	3527.8	A	D	-4.9	Cluster Incl AW121680:UI-M-BH2.3-aoc-e-04-0-UI.s1 Mus musculus cDNA, 3 end
116600_at	4016.2	P	I	3.6	513.6	A	D	-3.6	Cluster Incl AI848448:UI-M-AH1-ags-f-05-0-UI.s1 Mus musculus cDNA, 3 end
113066_at	2679	P	I	3.5	-695.6	A	D	-3.5	Cluster Incl AA638515:vo54b11.r1 Mus musculus cDNA, 5 end
113425_at	5823.2	P	I	3.5	1651.5	A	D	-3.5	Cluster Incl AA986447:ue13d02.x1 Mus

						musculus cDNA, 3 end
114658_at	6449.5 P	I	3.4	1894.4 A	D	-3.4 Cluster Incl AW123164:UI-M-BH2.1-apd-d-06-0-UI.s1 Mus musculus cDNA, 3 end
112307_at	3229.5 P	I	3.1	-97.3 A	D	-3.1 Cluster Incl AI851057:UI-M-BH0-ajv-g-07-0-UI.s1 Mus musculus cDNA, 3 end
106290_at	4659.6 P	I	2.9	1346.9 A	D	-2.9 Cluster Incl AI839436:UI-M-AN0-aco-f-08-0-UI.s1 Mus musculus cDNA, 3 end
110715_at	2974.3 P	I	2.9	443.5 A	D	-2.9 Cluster Incl AA214997:mu83g08.r1 Mus musculus cDNA, 5 end
115354_at	2953.4 P	I	2.8	572.8 A	D	-2.8 Cluster Incl AA895031:vx65f07.r1 Mus musculus cDNA, 5 end
114963_at	1834.9 P	I	2.7	-745 A	D	-2.7 Cluster Incl AI327212:mg81e01.x1 Mus musculus cDNA, 3 end
105229_at	1441.5 P	I	2.6	-521.9 A	D	-2.6 Cluster Incl AW048072:UI-M-BH1-alq-a-02-0-UI.s1 Mus musculus cDNA, 3 end
114138_at	2504.2 P	I	2.6	354.9 A	D	-2.6 Cluster Incl AI643851:vu09h05.x1 Mus musculus cDNA, 3 end
115993_at	2706.5 P	I	2.6	507.6 A	D	-2.6 Cluster Incl AI604578:vm27b08.y1 Mus musculus cDNA, 5 end
106949_at	3037.4 P	I	2.5	2774 A	D	-2.5 Cluster Incl AI850212:UI-M-BG1-aie-c-02-0-UI.s1 Mus musculus cDNA, 3 end
111398_at	17441.5 P	I	2.5	7759.1 A	D	-2.5 Cluster Incl AI605905:mu39b02.y1 Mus musculus cDNA, 5 end

Table 3.2 List of Potential Sex Determination Candidates that were Selected from our Data Set

Male and female candidates were chosen from A and B chips. F and M denote female and male Average Difference (Avg Diff) and Absolute Call (Abs Call) values respectively. Difference Call (Diff Call) and fold change are values obtained when comparing female against male or male against female. In these cases, if F value, then female is the target and male has been set as the baseline. Thus, female values are positive and male values are negative. Conversely, if M value, then male is the target and female has been set as the baseline. Therefore, male values are positive and female values are negative.

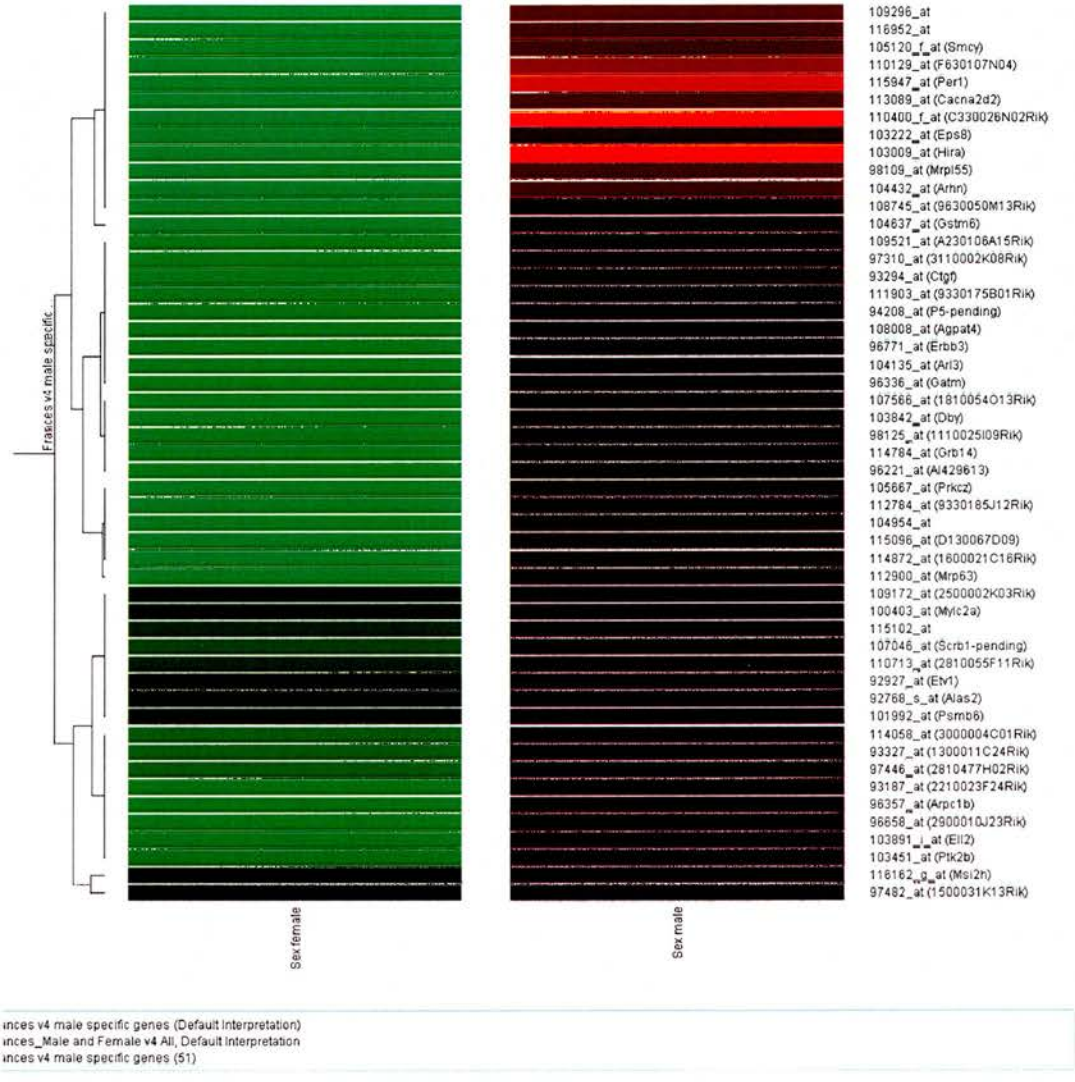


Figure 3.8 Male Potential Sex Determination Candidates that were Selected from our Data Set

Female signal intensities (left column) and male signal intensities (right column) of genes listed on right. All signal intensities are calculated against the median. Green indicates a decrease, red indicates an increase and black indicates no deviation from the median.

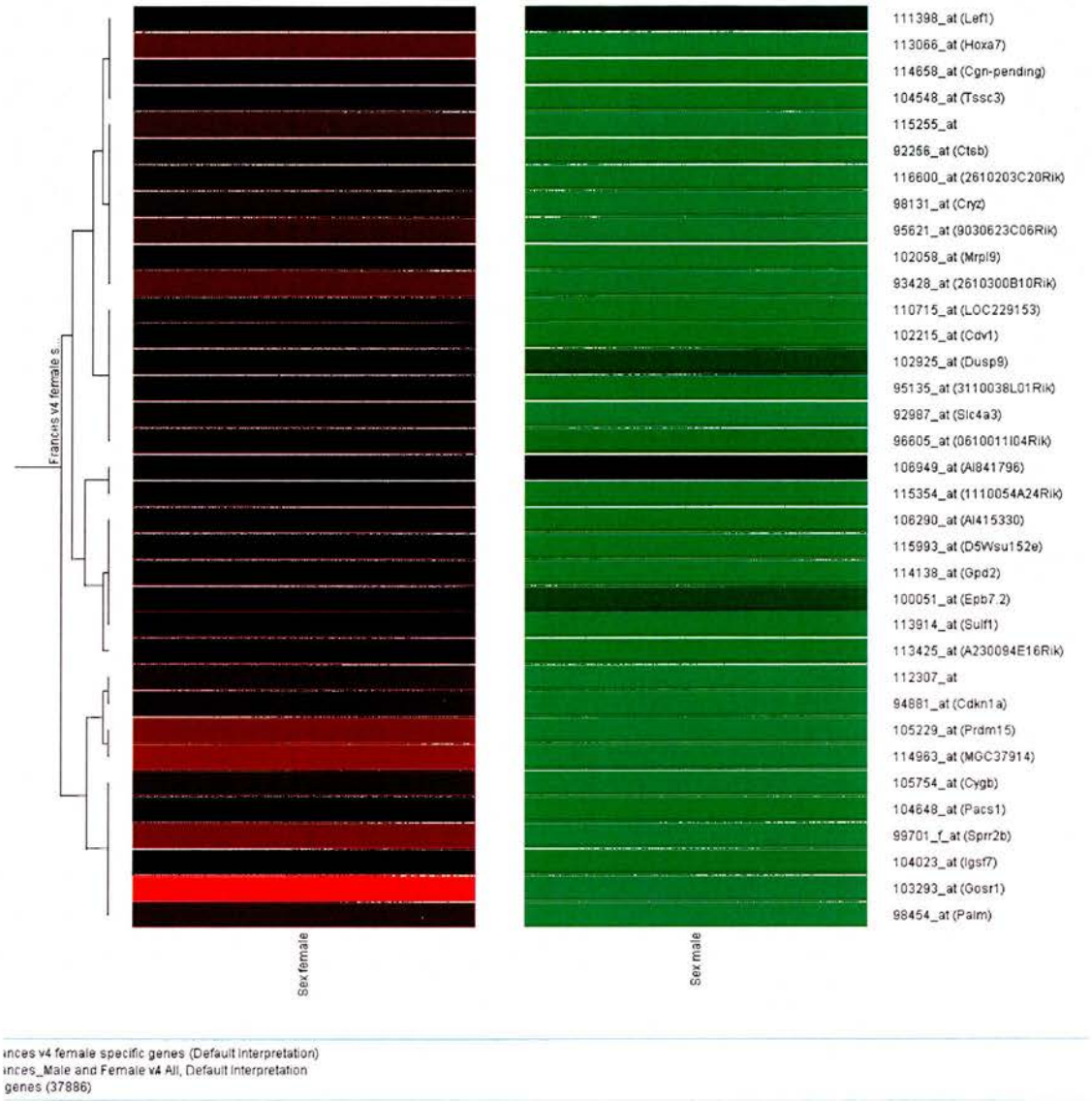


Figure 3.9 Female Potential Sex Determination Candidates that were Selected from our Data Set

Female signal intensities (left column) and male signal intensities (right column) of genes listed on right. All signal intensities are calculated against the median. Green indicates a decrease, red indicates an increase and black indicates no deviation from the median.

ESTs, genes and ESTs with the highest fold changes and which were 'present' in one RNA sample and 'absent' in the other were regarded as potential candidates which may be involved in the sex determination process. Ranked from the highest fold changes, 54 increased (male specific) and 35 decreased (female specific) genes and ESTs were selected for further investigation and analysis. Table 3.2 shows a list of the selected genes and ESTs along with the fold changes. Figures 3.8 and 3.9 show heat maps of the selected male and female candidates (GeneSpring 6.1, Silicon Genetics). In general, the potential male candidates are green in female and red in male whereas the potential female candidates are green in male and red in female. One or two candidates show black in both samples due to the signal intensities in male and female being fairly close. For example, Affymetrix probe set ID 106949_at is a potential female candidate as it is 'present' in female and 'absent' in male with a significant fold change of 2.5. But the raw signal gave similar values in male and female (2774 and 3037.4 respectively), therefore a difference in colour was not observed on the heat map as there was no large difference in signal intensity.

It is estimated that in humans, chromosomal locations of approximately 50% of mutations in male-to-female sex reversal patients are unidentified. Thus, the human equivalent of the selected genes and ESTs were checked to assess whether they may be related to complete or partial sex reversal in affected patients. Hence, their human chromosomal locations were obtained using Ensembl EnsMart (Appendix 7). The list of possible male and female sex determination candidates was passed onto our collaborators (Prof. Giovanna Camerino, Univ. of Pavia) for further investigations into the human sex developmental pathways.

3.3.7 Using a Statistical Algorithm

The screen 1 data set was initially analysed using Affymetrix MAS 4.0.1 as this was the current version available at that time. Affymetrix MAS 4.0.1 generates expression values by empirical algorithms. This version has now been updated to

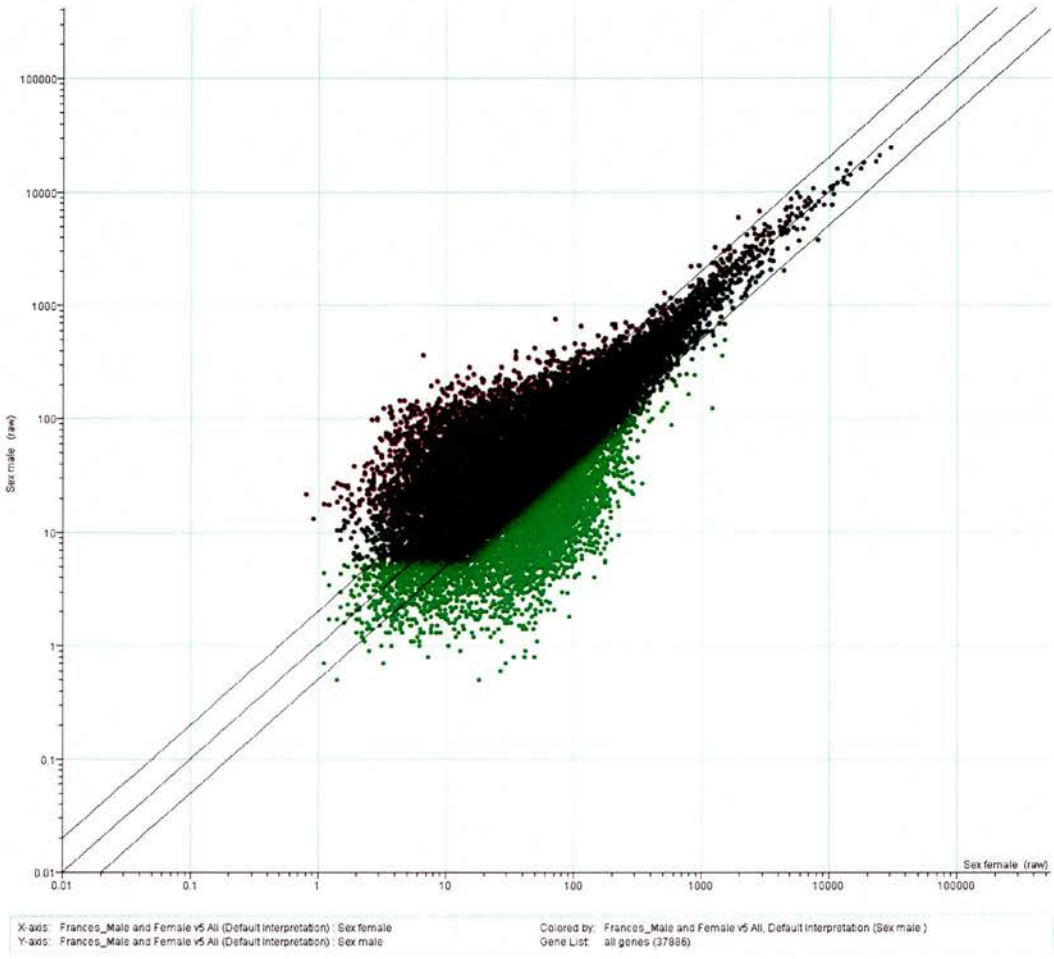


Figure 3.10 Scatter Plot Analysis of the Gene Expression Profile of Wild-type E12.5 Gonads Using a Statistical Algorithm

Each point represents the normalised expression level of an individual transcript within male and female RNA samples. Genes that are more highly expressed in male lie closer to the Y axis and genes that are more highly expressed in female lie closer to the X axis.

Affymetrix MAS 5.0 and generates expression values by statistical algorithms instead. Affymetrix claim this newer version to be more accurate as it provides the additional value of statistical significance (p values) and confidence limits. Affymetrix MAS 5.0 has also eliminated negative expression values (signals) and only positive values are generated. Taking all this into consideration, the screen 1 data set was re-analysed with Affymetrix MAS 5.0 to ascertain whether a statistical algorithm will alter the previously calculated values of genes and ESTs within the data set.

Firstly, a scatter plot analysis (GeneSpring 6.1, Silicon Genetics) of the female gonadal pool of RNA versus the male gonadal pool of RNA was generated to obtain an overall distribution of the genes and ESTs calculated by the statistical algorithm (Figure 3.10). This was compared to the scatter plot calculated using an empirical algorithm in Affymetrix MAS 4.0.1 (Figure 3.3). It is immediately obvious that the global distribution between the two forms of algorithms produces a different dispersal of genes and ESTs while the raw data is the same. The scatter plot produced from Affymetrix MAS 5.0 (Figure 3.10) shows a larger clustering of genes and ESTs. There are also no genes or ESTs positioned on the X axis or Y axis. These noticeable differences are due to the statistical algorithm not producing any negative values.

To further investigate the effect of the statistical algorithm, the same criteria used in Affymetrix MAS 4.0.1 (fold change greater than 2 and present in at least one sample) were applied to the re-analysis with Affymetrix MAS 5.0. This yielded 638 genes and ESTs to be differentially expressed which is a substantial decrease compared to the 1096 genes and ESTs found previously. Of the 638 genes and ESTs differentially regulated, 296 of these were up-regulated (more highly expressed in male) and 342 were down-regulated (more highly expressed in female). This is represented in a scatter plot in figure 3.11 (GeneSpring 6.1, Silicon Genetics). A comparison of the Affymetrix MAS 4.0.1 data set with the Affymetrix MAS 5.0 data set, using the above mentioned criteria, resulted in a match of 289 genes and ESTs.

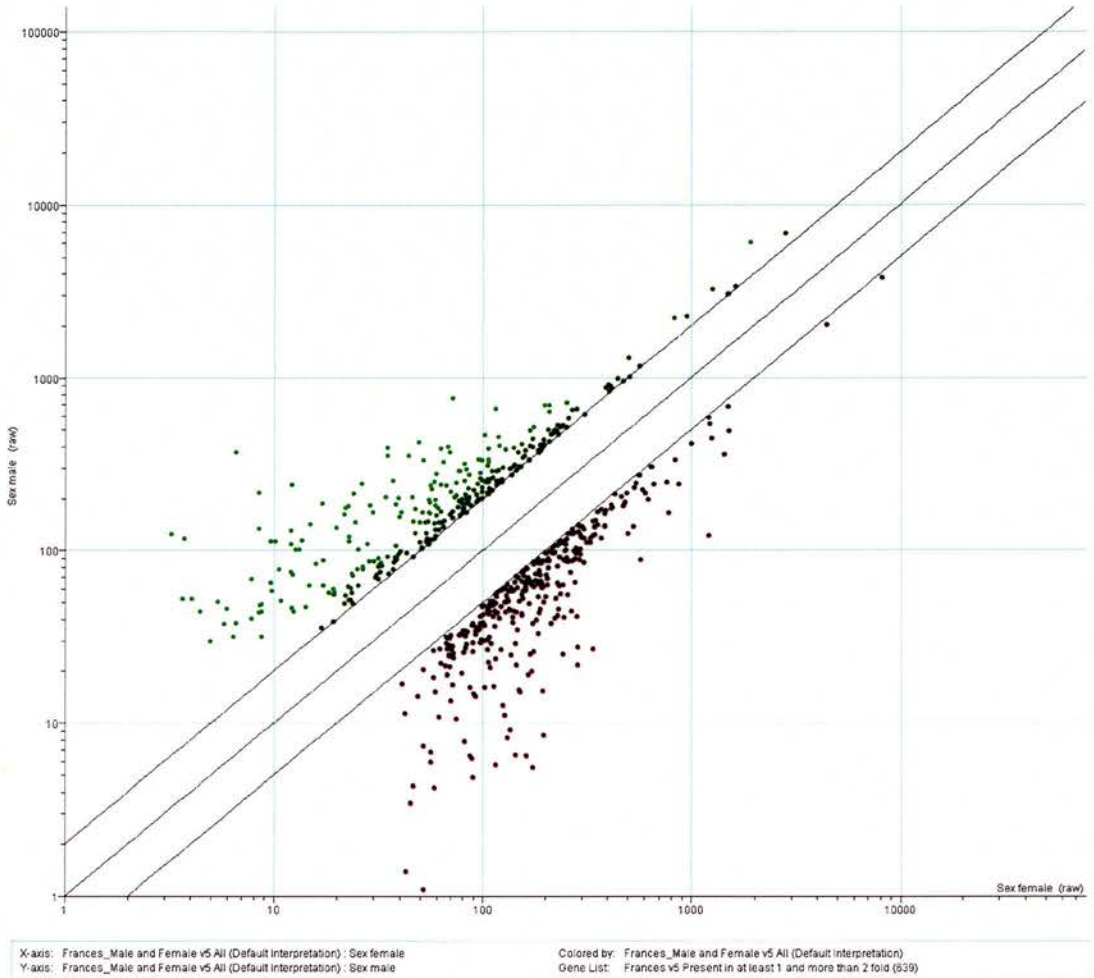


Figure 3.11 Scatter Plot Analysis of Genes Expressed Greater than 2 or –2 and Present in at Least One Sample of Wild-type E12.5 Gonads Using a Statistical Algorithm

Each point represents the normalised expression level of an individual transcript within male and female RNA samples. Genes that are more highly expressed in male lie closer to the Y axis and genes that are more highly expressed in female lie closer to the X axis.

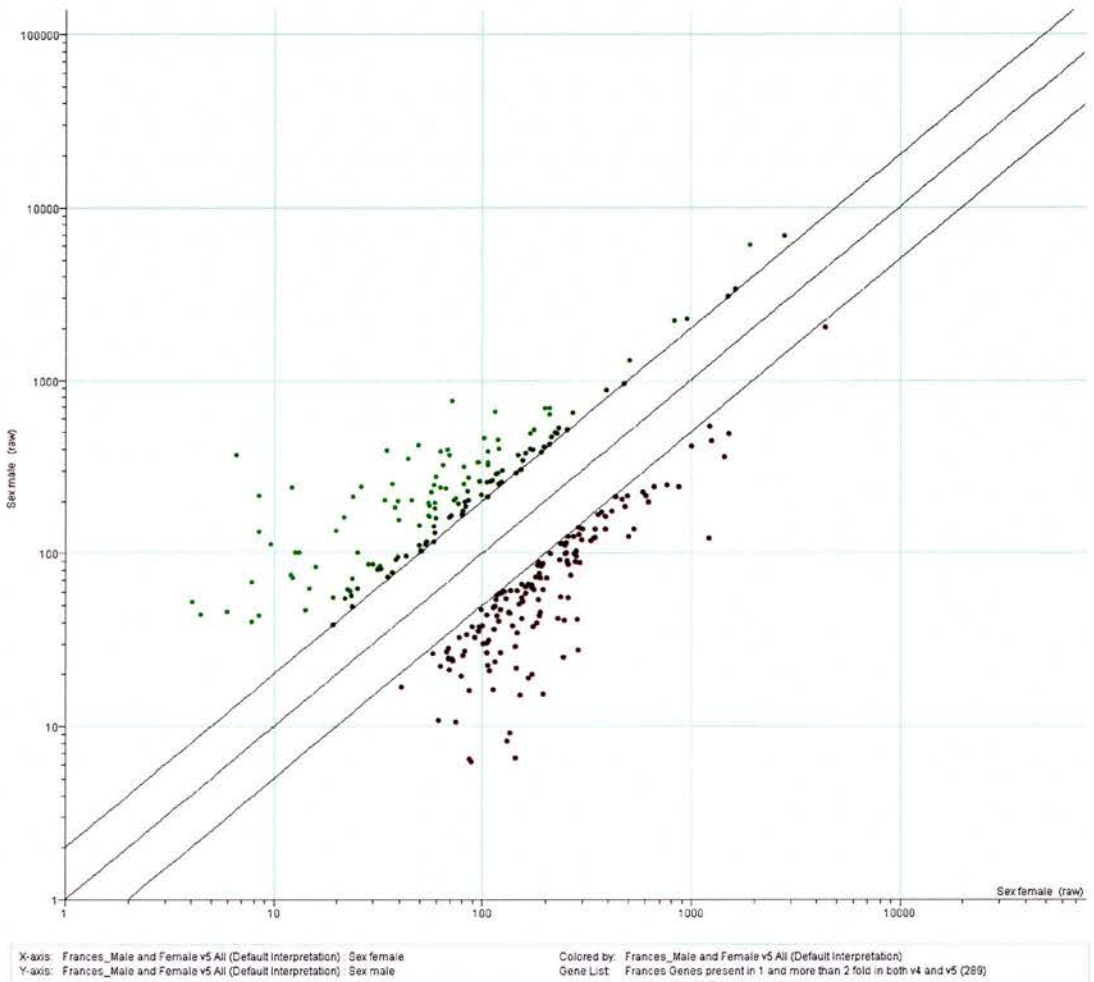


Figure 3.12 Scatter Plot Analysis Showing Genes Greater than 2 or –2 and Present in at Least One Sample in Both Empirical and Statistical Algorithms

Each point represents the normalised expression level of an individual transcript within male and female RNA samples. Genes that are more highly expressed in male lie closer to the Y axis and genes that are more highly expressed in female lie closer to the X axis.

Figure 3.12 shows a scatter plot (GeneSpring 6.1, Silicon Genetics) of these genes and ESTs and they are listed in Appendix 6. Although the expression levels of various genes and ESTs were expected to be different based on the different distribution from the scatter plots (Figures 3.3 and 3.10), this less than 50% overlap was surprising.

A closer inspection of how certain genes were expressed in both data sets revealed that the trend of genes and ESTs were the same. The difference was mainly in the fold change. In general, Affymetrix MAS 5.0 produced lower fold change values that has resulted in fewer differentiated genes and ESTs. For demonstration purposes, the list of potential candidates selected for further analysis in Affymetrix MAS 4.0.1 was converted to its statistical algorithm equivalent and heat maps were generated in GeneSpring 6.1 (Silicon Genetics). The potential male candidates are green in female and red in male as before. Concurrently, the potential female candidates are green in male and red in female as expected. Therefore, the trend within the selected potential candidates is the same.

In conclusion, the statistical algorithm has affected the ranking of genes and ESTs but has not altered the sex specificity. Although many of the previously selected potential candidates would have been selected again, if Affymetrix MAS 5.0 were available at the beginning, it's possible that other genes and ESTs may have been selected for further analysis as well. For example, Affymetrix probe set ID 92420_at (Neurotrophin 3; *Nt3*) had 'present' and 'increased' calls but a fold change of 1.8 in Affymetrix MAS 4.0.1 which is below our specified criteria and was therefore excluded. However, Affymetrix MAS 5.0 calculated the same 'present' and 'increased' calls but with a fold change of 2.1 and thus would have been included as a potential male candidate. NT3 is derived from the family of neurotrophins which are growth factors involved in tissue morphogenesis. Recently, it has been reported that NT3 expression is elevated in the embryonic rat testis during the time of seminiferous cord formation and inhibition of NT3 caused a reduction in the expression of Sox9 (Cupp et al., 2003). The authors hypothesised that when male

sex determination is initiated, the developing Sertoli cells express NT3 as a chemotactic agent for migrating mesonephros cells, which are essential to promote embryonic testis cord formation and influence downstream male sex differentiation.

3.4 Screen 2

3.4.1 Experimental Design

Due to the flawed design of the Affymetrix murine genome U74 microarrays, a complete data set was not obtained in screen 1. Therefore, a plan for the experiment to be repeated using the corrected microarrays (Affymetrix murine genome U74 version 2) was intended. Plus, it would be more viable to corroborate the genes identified in the first data set. However, due to time restrictions, it was impossible to collect the same amount of total RNA used in screen 1 for screen 2. Consequently, it was decided to try the new Affymetrix small sample labelling protocol version II. This protocol is designed to allow small quantities of total RNA to be amplified and labelled by two cycles of cDNA synthesis and *in vitro* transcription reactions for target amplification prior to hybridisation to microarrays for expression profiling. Unfortunately, there was not enough time to test the reproducibility of the protocol and the consistency of the RNA amplification. Thus, the protocol was followed strictly according to manufacturer's instructions (section 2.4.3).

As for screen 1, wild-type male and female gonads were quantitatively staged (as described in section 2.3.3) and dissected from mouse embryos at E12.5 (28 to 30 tail somites). Again, the adjacent mesonephros was removed from each gonad and equal amounts of RNA were hybridised to each GeneChip. Figure 3.13 shows an aliquot of the amplified male and female fragmented cRNA samples run on an agarose gel to ensure the cRNA was of good quality before hybridisation to the microarrays. The Affymetrix GeneChip arrays were scanned and analysed with Affymetrix MAS 5.0 software. Normalisation procedures used for screen 1 were applied to screen 2.

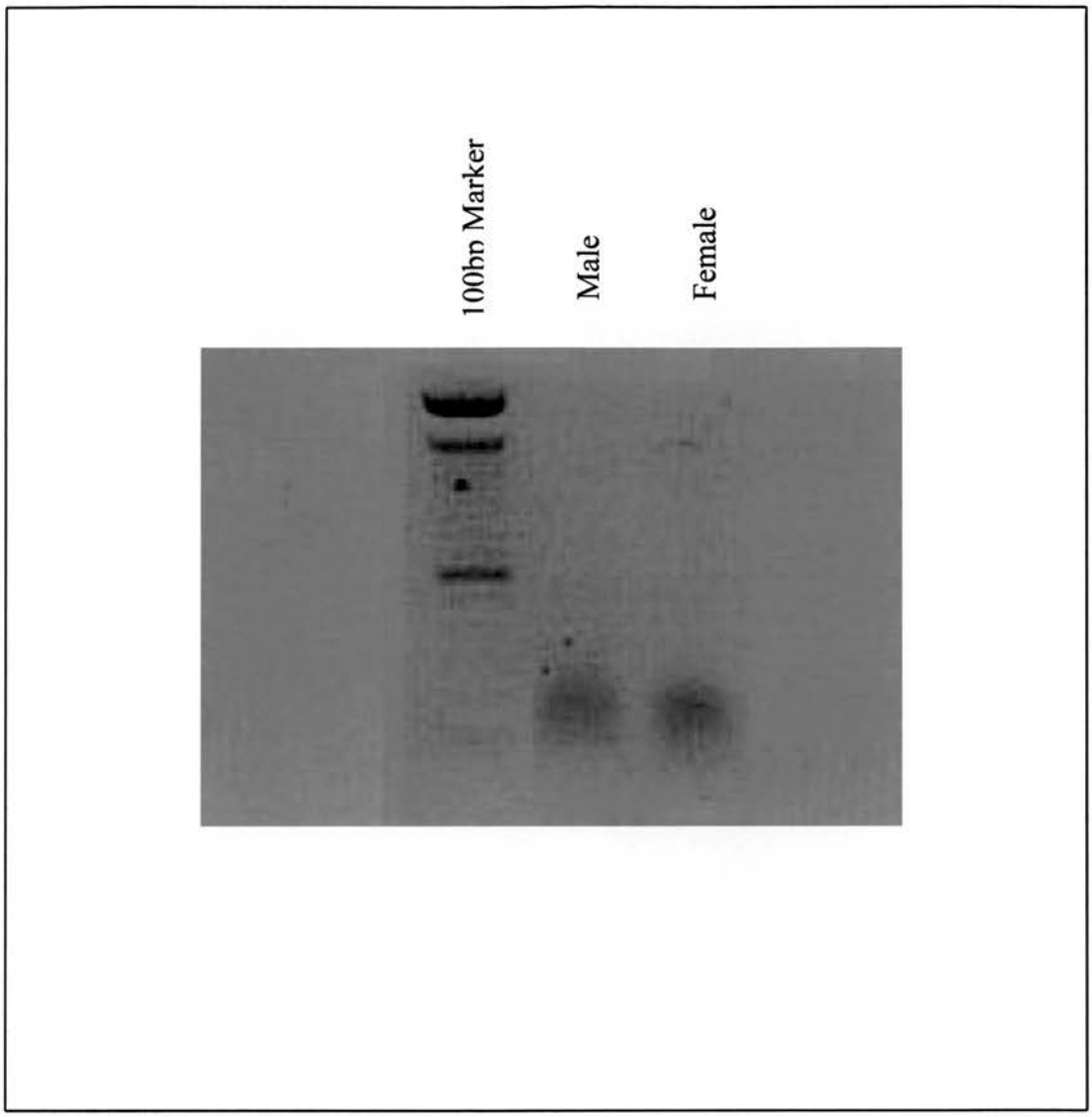


Figure 3.13 Fragmented cRNA Samples for Screen 2

16µl of labelled male and female cRNA samples were fragmented. 1µl of each fragmented cRNA sample was loaded onto an agarose gel prior to hybridisation onto microarrays. Results show both cRNA samples were fragmented and of significant quality.

However, it must be noted that screen 2 is not a replication of screen 1 as different protocols were used so it is not feasible to compare the two data sets at a quantitative gene expression level. But the two data sets may be compared at a biological level where the differential expression of genes and ESTs in both screens should be similar.

3.4.2 Results

Before the analysis of the second data set commenced, the statistics of the arrays were checked by the generation of their report files. Appendices 4e and 4f show the report files for male and female respectively. As in screen 1, the percentage of probe sets scored as 'present' and the 3'-5' GAPDH ratio were examined. The MG_U74Av2 chip hybridised with the male sample scored as 39.3% probe sets 'present' and the 3'-5' GAPDH ratio was 60.38 (Appendix 4e). The percentage of probe sets 'present' was an expected value. However, the 3'-5' GAPDH ratio is far too high. This reveals that the cRNA transcripts were too short and poor hybridisation may have occurred. This is most likely caused by the second *in vitro* transcription reaction not working well, thus not producing full-length transcripts. The MG_U74Av2 chip hybridised with the female sample scored as 12.5% probe sets 'present' and the 3'-5' GAPDH ratio was 1.29 (Appendix 4f). Although the 3'-5' GAPDH ratio fell with the manufacturer's guidelines of less than 3, the absolute calls for GAPDH and β-actin control probes were 'absent'. This indicates that the quality of this RNA sample was not sufficient. Degradation may have occurred at same point during the amplification procedure. In addition, the percentage of probe sets 'present' is too low. Again, this is due to the poor quality of the RNA resulting in poor hybridisation to the microarray. In summary, both RNA samples did not amplify well and the arrays for screen 2 were unsuccessful. If time had permitted, the reproducibility and consistency of the RNA amplification for E12.5 gonadal tissue would have been tested and optimisation of the protocol would have been carried out if required.

3.5 Discussion

The use of microarrays has proven to be a helpful tool in fishing for novel sex determining genes. Although all the potential candidates will have to be verified by other methods, such as *in-situ* hybridisation, microarrays have provided a good starting point. There are different approaches that can be applied when using microarrays. We chose to take the simple and economical approach where only one specific time point was used with pooled RNA samples and thus many arrays were not necessary. This approach was sufficient for our needs. A different but more costly and time-consuming approach would have been to repeat the screens (Affymetrix recommends at least two replicates) for a more concrete data set, where replicates would be cross-matched and so the number of false positives would be reduced. Replicating screens will also allow genes to be statistically analysed, for example hierarchical clustering could group genes which have the same biological function. This would identify patterns within the data sets and distinguish biologically relevant groups of genes. However, in a biological system with pooled RNA samples, it is difficult to reproduce such a solid data set as the system has too many variables. Alternatively, more time points could have been selected (for example, E11.5 and E13.5) to allow for a wider gene expression profile. From this, the sex differentially expressed genes would be identified more easily. Of course, the latter two approaches would be more feasible if smaller amounts of total RNA could be used. Therefore, time would need to be spent on acquiring a well-optimised, consistent and reproducible RNA amplification protocol if a more array-based approach is to be taken.

Following the start of our Affymetrix sex determination screen, other laboratories have performed two other sex determination screens using E12.5 gonads. However, these were done by a suppression subtraction hybridisation method. The first one, as mentioned above, was performed by Menke and Page (Menke and Page, 2002). In their screen, they identified 19 genes that were expressed at significantly high levels in male gonads. From these 19 genes, 6 of these were also found to be more highly

expressed at a significant level in our screen. These 6 genes were: Cerebellin 1 precursor protein (*Cbln1*), Procollagen type IX α3 (*Col9a3*), Prostaglandin D2 synthase (*Ptdgs*), Renin (*Ren1*), Secreted phosphoprotein 1 (*Spp1*) and Protease nexin-1 (*Pn-1*). More recently, a second screen performed by McClive et al (McClive et al., 2003) identified 54 genes with high expression in male gonads versus low expression in female gonads. From these 54 genes, 8 of these were also found to be more highly expressed in male gonads than female gonads in our screen. These 8 genes were: Transcription factor 21 (*Tcf21*), Alcohol dehydrogenase 5 (*Adh5*), Heterogeneous nuclear riboproteins methyltransferase-like 2 (*Hrmt1l2*), Vanin 1 (*Vnn1*) and Cerebellin 1 precursor protein (*Cbln1*), Procollagen type IX α3 (*Col9a3*), Prostaglandin D2 synthase (*Ptdgs*) and Protease nexin-1 (*Pn-1*) again. Thus, in our Affymetrix microarray approach, we have identified common genes with the above mentioned suppression subtraction hybridisation approaches. Although in each screen the number of common genes is not large, it should be noted that there were only 4 genes that are common when comparing the two suppression subtraction hybridisation screens against each other. Hence, most approaches will reveal common genes but also genes that are only found in that particular study. Interestingly, both screens did not identify *Sox9* either and only the screen by Menke and Page identified *Mis*. Thus, results of each screen will depend upon the methods used and the time points analysed. Combined approaches are required to generate a more complete gene expression profile of the developing gonad.

The preliminary attempt to identify novel sex determining candidates was relatively successful in screen 1. As the exact screen was not replicated, it was difficult to rule out the false positives within the data set. Therefore, a wide scale verification method was chosen for the next step. Hence, 89 putative candidates in total were selected to be further analysed. It is expected that some of the selected candidates represent false positives and hence will not be differentially expressed in male or female gonads. Thus RNA *in-situ* hybridisation of E12.5 gonads will be performed to identify the true sex differentiating genes and ESTs (Chapter 4). Moreover, quantitative real-time PCR will be used to determine the temporal gene expression

pattern of each candidate throughout gonadal differentiation (Chapter 5). To summarise, the aim of using Affymetrix GeneChips to identify novel sex determining candidates has been accomplished. In addition, the approach we took can be used for other biological systems.

Chapter 4 Verification of Microarray Data

Chapter 4 Verification of Microarray Data

4.1 Introduction

When performing microarray screens, it is important to confirm the results obtained. This can be either done by repeating the screens several times to obtain a statistically significant set of data, or to use alternative methods to confirm differential expression patterns. This chapter examines the sex specific expression of potential candidates by performing PCR amplification and whole-mount *in-situ* hybridisation.

4.2 Analysis of Potential Sex Determination Candidates by PCR Amplification

To investigate the potential sex determination candidates (as selected in chapter 3, Table 3.2), a total of nine candidate genes from the A chip were chosen for a qualitative RT-PCR analysis using RNA isolated from E12.5 male and female gonads. These included the five highest ranking candidates (Affymetrix probe set ID: 104637_at (fold change of 12.3), 92932_at (fold change of 6.7), 101446_at (fold change of 4.3), 96336_at (fold change of 4.3) and 96926_at (fold change of 4.2). A further four candidates were selected to assess whether lower fold change values would also show a differential male and female expression pattern by PCR amplification. These candidates were Affymetrix probe set ID: 104135_at (fold change of 3.1), 98109_at (fold change of 2.7), 97446_at (fold change of 2.4) and 100403_at (fold change of 2.3). In addition, GAPDH primers were used as a positive control to ensure that male and female pools of cDNA were of equal quantity.

Figure 4.1 shows the results of the PCR amplification. As all the candidates chosen for this assay were previously identified to be enriched in male gonads, it was expected that each candidate would be more highly expressed in the male sample than in the female sample. Affymetrix probe set ID 92932_at (fold change of 6.7)

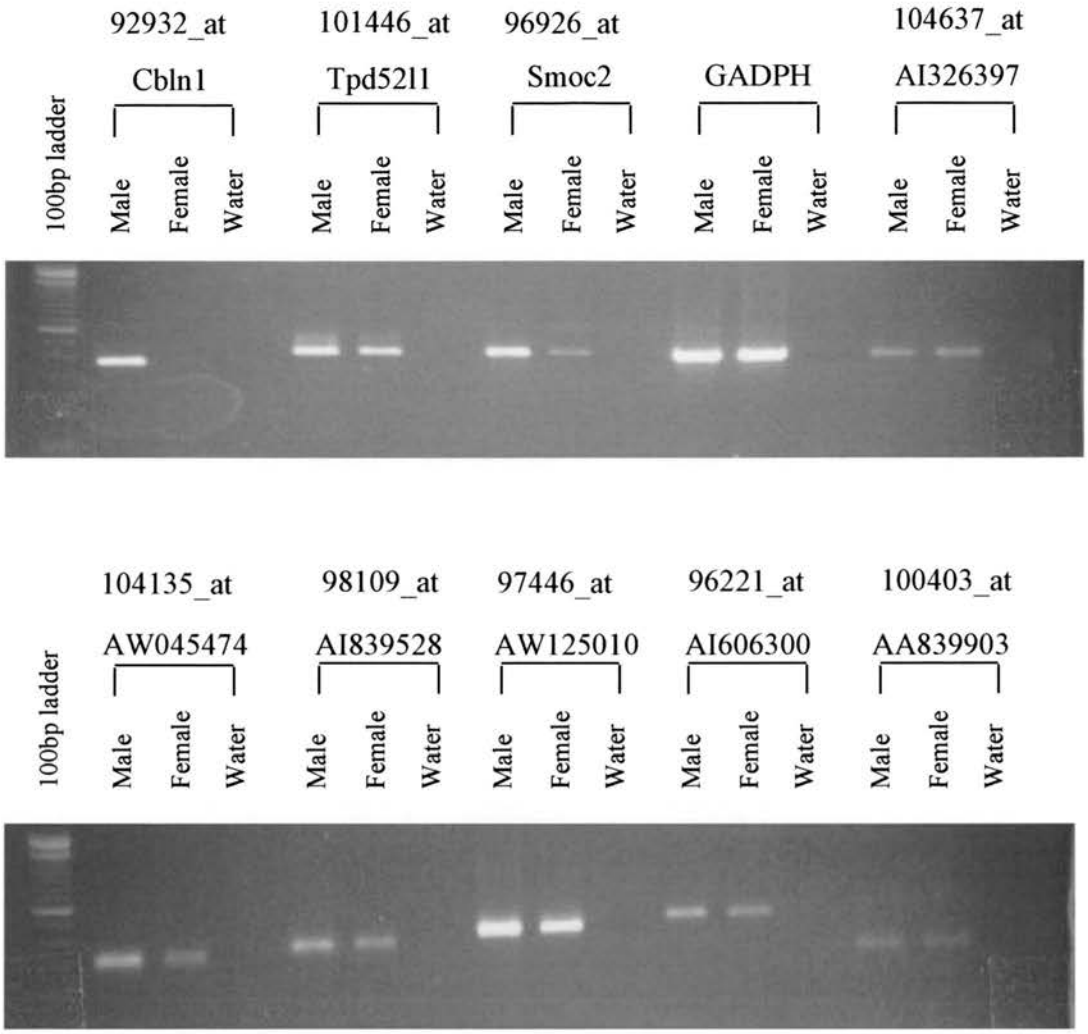


Figure 4.1 Expression Analysis of Potential Candidates by PCR Amplification

Expression analysis of 9 potential candidates by PCR amplification (35 cycles) using male and female gonadal samples. GAPDH was used as a positive control to ensure that both samples were of equal concentration. Water was used as a negative control. Samples are labeled by Affymetrix probe set ID and GenBank descriptions corresponding to microarray data represented in Table 3.2.

was striking as strong expression in the male sample can be seen, whereas no expression in the female sample was detected. Affymetrix probe set ID 96926_at (fold change of 4.2) also showed a clear difference between the two samples, where a dramatic decrease in the female sample was observed compared to the strong expression in the male sample. A high level of expression in the male sample and a low level of expression in the female sample could also be seen in Affymetrix probe set ID 101446_at (fold change of 4.3) and Affymetrix probe set ID 104135_at (fold change of 3.1). It should also be noted that the GAPDH positive control shows no difference between male and female samples, confirming that both cDNA samples are of equal concentration and differences in expression levels between male and female were not due to a difference in DNA concentration. These results gave us the confidence to proceed with RNA *in-situ* hybridisation for these candidates and for an additional 80 previously selected candidates.

4.3 Further Analysis of Potential Sex Determination Candidates by Whole-mount In-situ Hybridisation

RNA *in-situ* hybridisation probes were designed for all of the selected candidates. Anti-sense probes were produced by PCR amplification (as described in section 2.5.1 and as listed in Appendix 1). Subsequently, whole-mount *in-situ* hybridisation was performed on eviscerated E12.5 male and female embryos for each probe. Results were separated into the following four categories: increased expression in testis, increased expression in ovary, equal expression in both sexes and undetectable expression in either sex. It should be noted that from the 89 selected potential sex-determining candidates in screen 1 (chapter 3), 54 candidates were identified to be more highly expressed in male gonads and 35 candidates were identified to be more highly expressed in female gonads (Table 3.2). Subsequently, of the 89 probes that were assayed by whole-mount *in-situ* hybridisation, 3 were found to be confirmed with increased expression in testis , none were found to be increased in the ovary , 41

were found to be expressed in both testis and ovary and 45 were found to have undetectable expression in either testis or ovary .

The 3 candidates with confirmed high expression in testis were Affymetrix probe set ID 92932_at (fold change of 6.7), Affymetrix probe set ID 101446_at (fold change of 4.3) and Affymetrix probe set ID 96926_at (fold change of 4.2). These 3 candidates were novel and of unknown function. Their GenBank descriptions were cerebellin 1 precursor protein (*Cbln1*) for Affymetrix probe set ID 92932_at (fold change of 6.7), tumour protein D52-like 1 (*Tpd52l1*) for Affymetrix probe set ID 101446_at (fold change of 4.3) and secreted modular calcium-binding protein 2 (*Smoc2*) for Affymetrix probe set ID 96926_at (fold change of 4.2). It should be noted that the GenBank descriptions for these candidates will be used hereafter.

Figures 4.2, 4.3 and 4.4 show the whole-mount *in-situ* hybridisation on E12.5 gonads of *Cbln1*, *Tpd52l1* and *Smoc2* respectively. It can be seen that the expression of each candidate is only within the testis as staining can only be detected in male gonads, while female gonads remain unstained. Thus, all 3 potential candidates initially identified from screen 1 have been verified to be sex differentially expressed at E12.5.

4.4 Discussion

It is important to verify microarray data by other methods to rule out false positives within the data set. Using PCR amplification as an initial indicator for several potential candidates showed that differential expression levels could be obtained between male and female gonadal samples. This led to the wide scale, whole-mount *in-situ* hybridisation approach of assaying 89 potential candidates. Unfortunately, this method only yielded confirmation of 3 candidate genes to be male specifically expressed at E12.5. Moreover, none of the female potential candidates were positive by *in-situ* hybridisation. This may be due to the RNA *in-situ* hybridisation probes

MALE

FEMALE

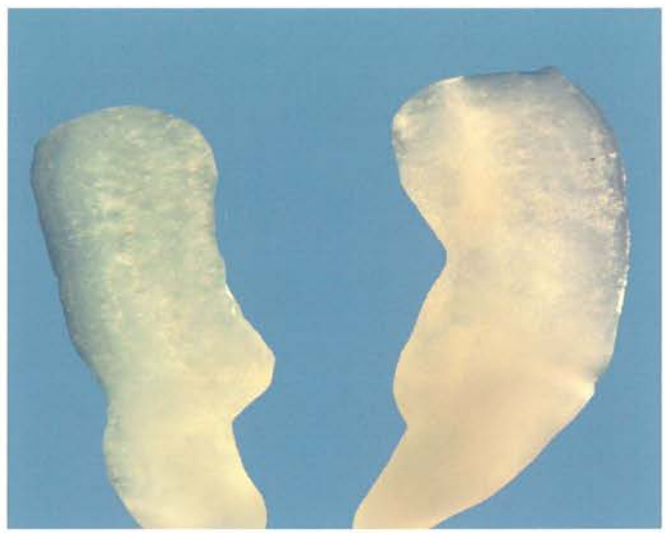
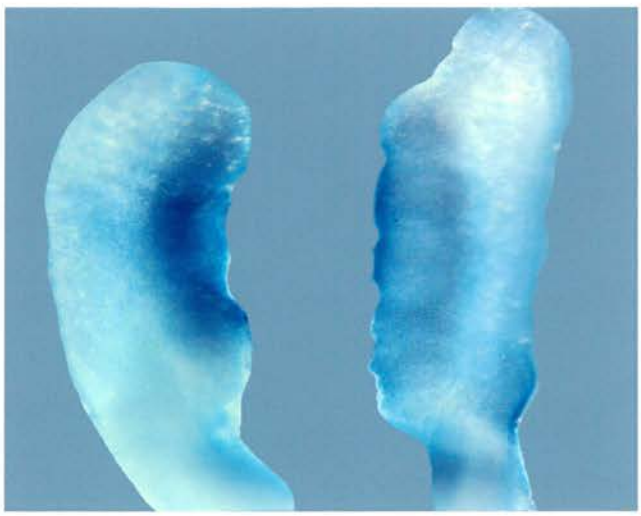


Figure 4.2 *Cbln1 In-situ* Hybridisation of E12.5 Gonads

Whole-mount *in-situ* hybridisation of E12.5 gonads using anti-sense *Cbln1* probe. Top panel shows *Cbln1* staining in male gonads, whereas no staining is observed in female gonads (lower panel). Sense probes were negative for both male and female gonads (data not shown).

MALE

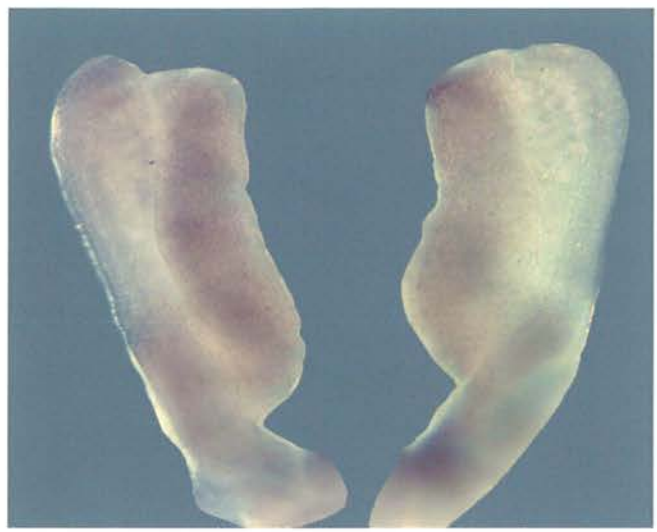


FEMALE

Figure 4.3 *Tpd52l1 In-situ* Hybridisation of E12.5 Gonads

Whole-mount *in-situ* hybridisation of E12.5 gonads using anti-sense *Tpd52l1* probe. Top panel shows *Tpd52l1* staining in male gonads, whereas no staining is observed in female gonads (lower panel). Sense probes were negative for both male and female gonads (data not shown).

MALE



FEMALE

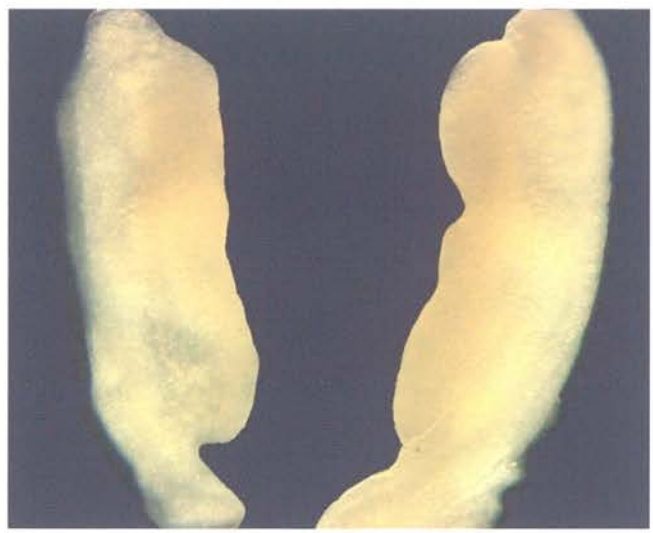


Figure 4.4 *Smoc2 In-situ* Hybridisation of E12.5 Gonads

Whole-mount *in-situ* hybridisation of E12.5 gonads using anti-sense *Smoc2* probe. Top panel shows *Smoc2* staining in male gonads, whereas no staining is observed in female gonads (lower panel). Sense probes were negative for both male and female gonads (data not shown).

not being specific enough or optimisation of individual probes may be required. Alternatively, the transcripts are lowly expressed and are therefore not detected by *in-situ* hybridisation. However, it should be noted that screen 1 identified follistatin (*Fst*) as a female specific gene with a fold change value of 9 (high-lighted in green, Appendix 5). This gene was also later identified in the subtraction hybridisation screen by Menke and Page (Menke and Page, 2002) as more highly expressed in female gonads than male gonads. Therefore, the most likely explanation is that there were technical difficulties with the *in-situ* hybridisations. These technical difficulties can range from insufficient proteinase K treatment to poor hybridisation of RNA probes. Future optimisation of *in-situ* hybridisation conditions should resolve these problems.

In summary, although the whole-mount *in-situ* hybridisation results helped to confirm male specific expression in the 3 candidates that were more highly expressed in the male gonads, the differential expression between male and female gonads of the remaining 86 candidates cannot be ruled out. Hence, there may be more male and female specific genes within the selected potential candidates that were assayed. Due to the clear differential expression of the 3 male specific candidates, further analysis of these 3 putative genes will be carried out in an attempt to determine their expression pattern and gene expression profile. This will be done by further whole-mount *in-situ* hybridisation at E11.5, E13.5 and E14.5 (chapter 5). In addition, real-time PCR of these candidates will be conducted to gain a more quantitative measure of differential expression between male and female gonads.

Chapter 5 Expression Analysis of Cbln1, Tpd52l1 and Smoc2

Chapter 5 Expression Analysis of Cbln1, Tpd52l1 and Smoc2

5.1 Introduction

To gain further insights into the sexually dimorphic expression patterns and gene expression profiles of *Cbln1*, *Tpd52l1* and *Smoc2*, more detailed analyses were required. This chapter will investigate the other embryonic tissues within which these genes may be expressed by PCR amplification. It also extends the initial verification of these candidate genes (described in chapter 4) and shows further whole-mount *in-situ* hybridisation of male and female gonads at E11.5, E13.5 and E14.5. In addition, RNA *in-situ* hybridisation of E14.5 gonadal sections will be performed in an attempt to identify the specific cell type within which each gene is expressed. Moreover, quantitative real-time PCR of these candidates will be carried out to verify the fold changes from the microarray analysis and to determine their temporal expression patterns, ranging from 17 to 36 tail somites.

5.2 Expression Pattern of Cbln1, Tpd52l1 and Smoc2 by PCR Amplification

We have shown that *Cbln1*, *Tpd52l1* and *Smoc2* have a sexually dimorphic expression pattern at E12.5. However, their expression patterns within other tissues were unknown and were therefore investigated. PCR amplification of cDNA from E12.5 heart, lungs, kidneys, liver, forelimbs, hindlimbs and the head was carried out. Results are shown in figure 5.1. For *Cbln1*, strong expression in the forelimbs, hindlimbs and the head, moderate expression in the liver, yet no expression in the heart, lungs or kidneys is detected. *Tpd52l1* showed strong expression in the forelimbs, hindlimbs and the head, along with moderate expression in the heart, lungs and liver. However, expression could not be detected in the kidneys. Finally, for *Smoc2*, expression seems to be ubiquitous as it is detected in all of the embryonic organs that were assayed.

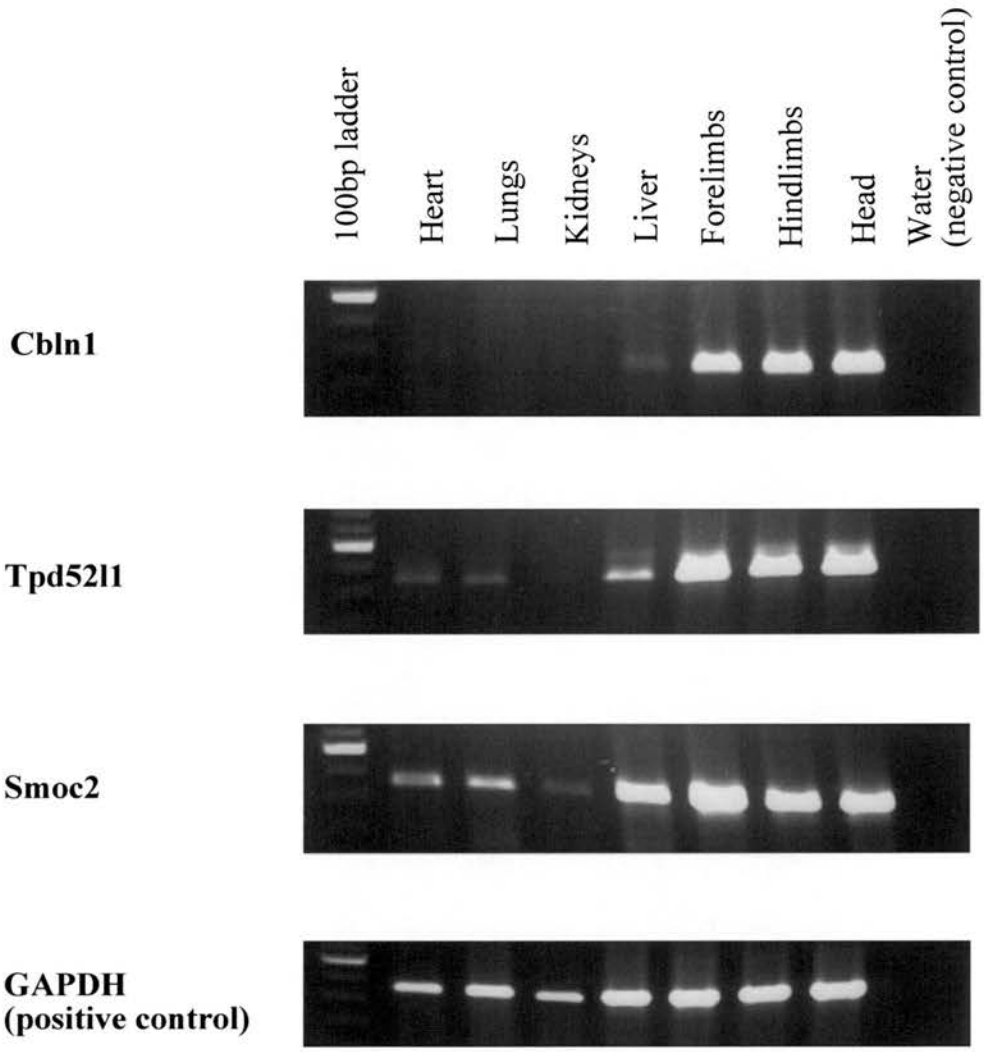


Figure 5.1 Expression Pattern of *Cbln1*, *Tpd52l1* and *Smoc2* by PCR Amplification

RNA was isolated from the above E12.5 organs and cDNA was synthesised accordingly. Expression analysis of each sample was determined by PCR amplification (35 cycles) and assayed by gel electrophoresis. GAPDH was used as a positive control to observe any differences in cDNA concentration. Water was used as a negative control for each gene.

5.3 Expression Pattern of Cbln1, Tpd52l1 and Smoc2 in Gonads

5.3.1 Whole-mount *In-situ* Hybridisation of Gonads

Extending the initial sex differential verification of *Cbln1*, *Tpd52l1* and *Smoc2* with E12.5 gonads, whole-mount *in-situ* hybridisation was also carried out for E11.5, E13.5 and E14.5 gonads. *Cbln1* results (Figure 5.2) indicate stronger expression in male gonads than female gonads at each stage assayed. However, the staining pattern seems to be diffuse throughout the gonad, particularly in the E13.5 and E14.5 male gonads. Although staining is not detected in E11.5 and E13.5 female gonads, there may be a low level of expression in E14.5 female gonads.

Figure 5.3 shows the *Tpd52l1* results. Marginal staining can be observed in E11.5 male gonads, whereas the E11.5 female gonad shows no staining. At E13.5, *Tpd52l1* expression can be detected in male gonads, whereas female gonads are not staining. In fact, *Tpd52l1* expression seems to be expressed within the testicular cords. This expression pattern is also shown in E14.5 male gonads. However, there is marginal staining in female gonads also at this stage.

Smoc2 results (Figure 5.4) also show staining in E11.5 male gonads, when compared to E11.5 female gonads where no staining is detected. Staining is also observed in E13.5 and E14.5 gonads. Although the expression is stronger in male gonads, there is slight expression in the female gonads. It should also be noted that the mesonephroi of male and female E13.5 and E14.5 urogenital ridges are stained, whereby expression may be stronger in the female mesonephroi at these stages.

5.3.2 RNA *In-situ* Hybridisation on Gonadal Sections

RNA *in-situ* hybridisation was conducted on E14.5 transverse gonadal sections (as described in section 2.5.3). Figures 5.5, 5.6 and 5.7 shows the expression patterns of

Figure 5.2 *Cbln1 In-situ* Hybridisation of E11.5, E13.5 and E14.5 Gonads

Whole-mount *in-situ* hybridisation of E11.5, E13.5 and E14.5 gonads using anti-sense *Cbln1* probe. Staining detected in male gonads, no staining observed in female gonads. Sense probe was negative for both male and female gonads (data not shown).

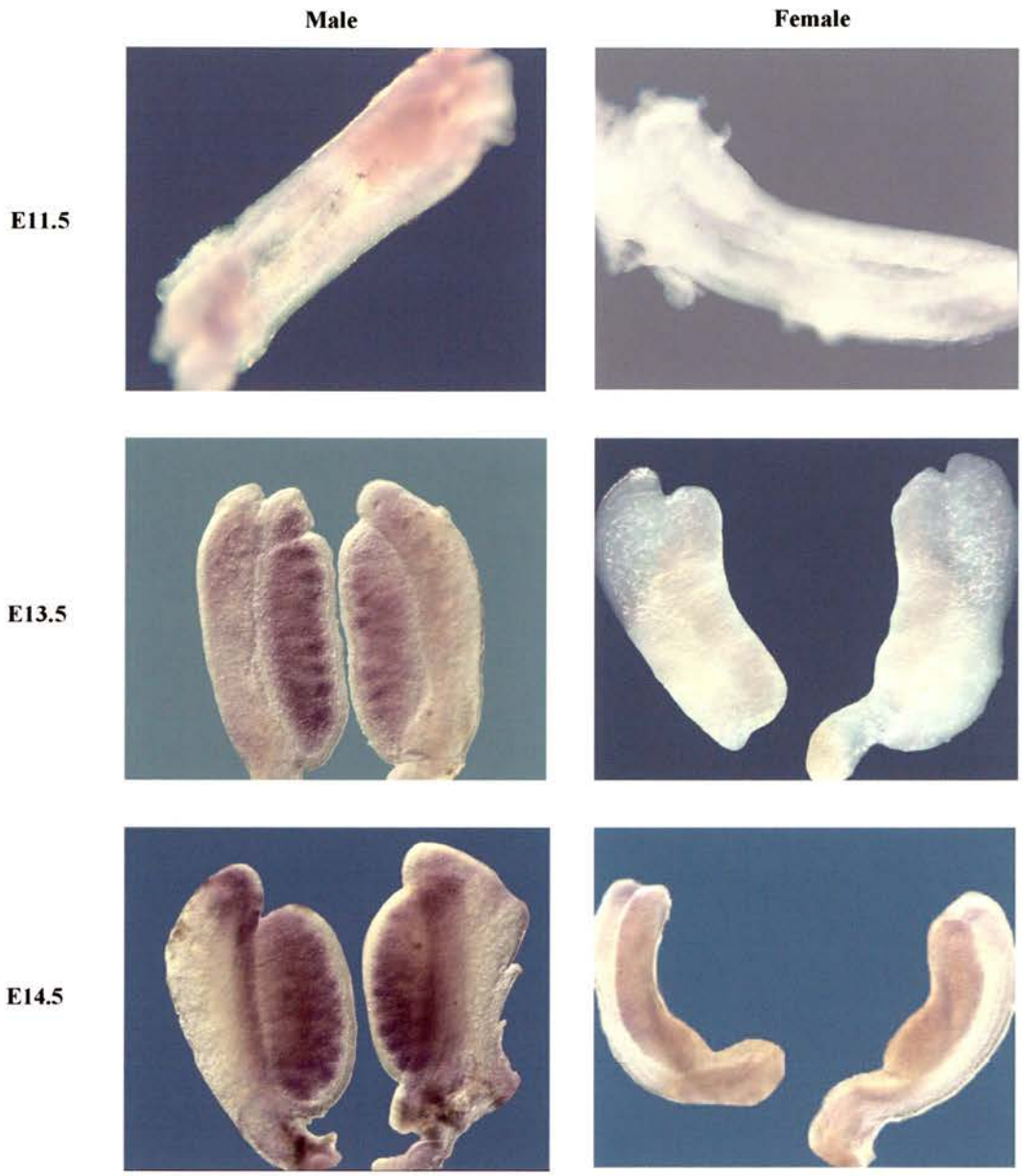


Figure 5.3 *Tpd52l1 In-situ* Hybridisation of E11.5, E13.5 and E14.5 Gonads

Whole-mount *in-situ* hybridisation of E11.5, E13.5 and E14.5 gonads using anti-sense *Tpd52l1* probe. Staining detected in male gonads and marginal staining found in E14.5 female gonads, no staining observed in E11.5 and E13.5 female gonads. Sense probe was negative for both male and female gonads (data not shown).

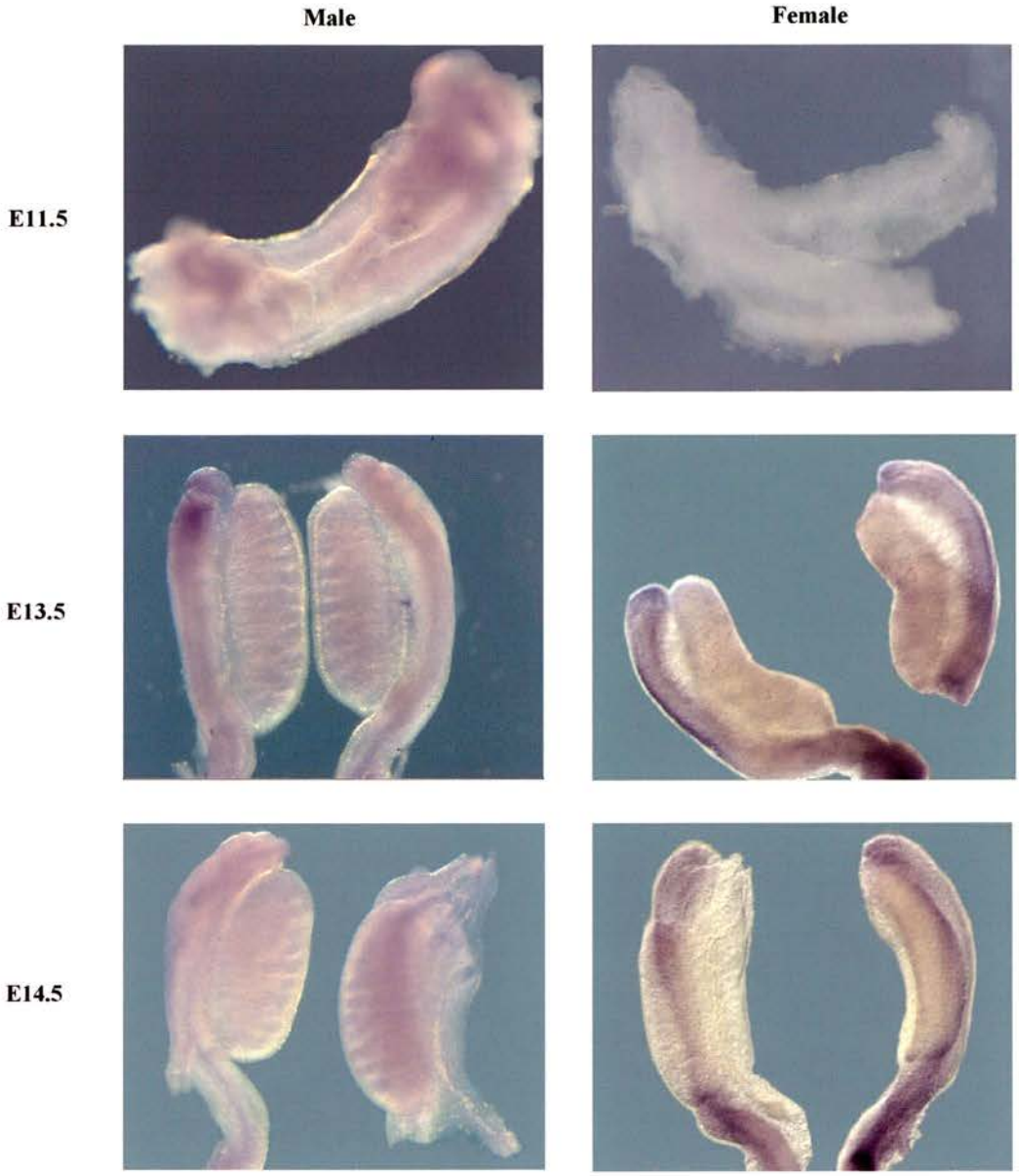


Figure 5.4 *Smoc2 In-situ* Hybridisation of E11.5, E13.5 and E14.5 Gonads

Whole-mount *in-situ* hybridisation of E11.5, E13.5 and E14.5 gonads using anti-sense *Smoc2* probe. Staining detected in male gonads and marginal staining found in E13.5 and E14.5 female gonads. Staining also detected in the mesonephroi at E13.5 and E14.5. No staining observed in E11.5 female gonads. Sense probe was negative for both male and female gonads (data not shown).

Cbln1, *Tpd52l1* and *Smoc2* respectively. The *Cbln1* anti-sense probe revealed the expression of *Cbln1* to be mainly localised within the central region of the male gonad. However, this expression does not seem to be restricted to the sex cords. Moreover, cortical regions seem to be moderately stained. It should be noted that the sense control showed no staining within the male gonad and no expression was detected in the female gonad when using both sense and anti-sense probes. The *Tpd52l1* anti-sense probe showed strong staining within the sex cords of the male gonad, but also slight staining in the region surrounding the sex cords. As for *Cbln1*, no *Tpd52l1* expression is detected in the female gonad or with the sense probe in male and female gonads. The *Smoc2* anti-sense probe revealed *Smoc2* expression throughout the male gonad. In contrast, no expression could be detected in the female gonad and the sense probe also showed no staining in the male and female gonads. The Sertoli cell marker, *Mis*, was used as a positive control (Figure 5.8) and clear expression can be detected within the sex cords of the testis.

5.4 Gene Expression Profiles of Cbln1, Tpd52l1 and Smoc2 in Gonads

5.4.1 Experimental Design

In order to assess the sex-specific expression levels of *Cbln1*, *Tpd52l1* and *Smoc2*, a quantitative analysis approach using real-time PCR was taken. Wild-type male and female gonad pairs, ranging from early E11.5 to late E12.5, were used to obtain their temporal expression patterns. These gonad samples were accurately staged by tail somite counting (as described in section 2.3.3) and ranged from 17 tail somites (early E11.5) to 36 tail somites (late E12.5). Each tail somite stage was treated as individual samples to generate a specific time scale for assaying gene expression profiles in relation to the developmental stages of the gonad. It was of particular interest to identify whether the potential sex determination candidates (*Cbln1*, *Tpd52l1* and *Smoc2*) were involved in the initial stages of sex development, as they appeared to be sexually dimorphic in their expression at E11.5 (shown by *in-situ*

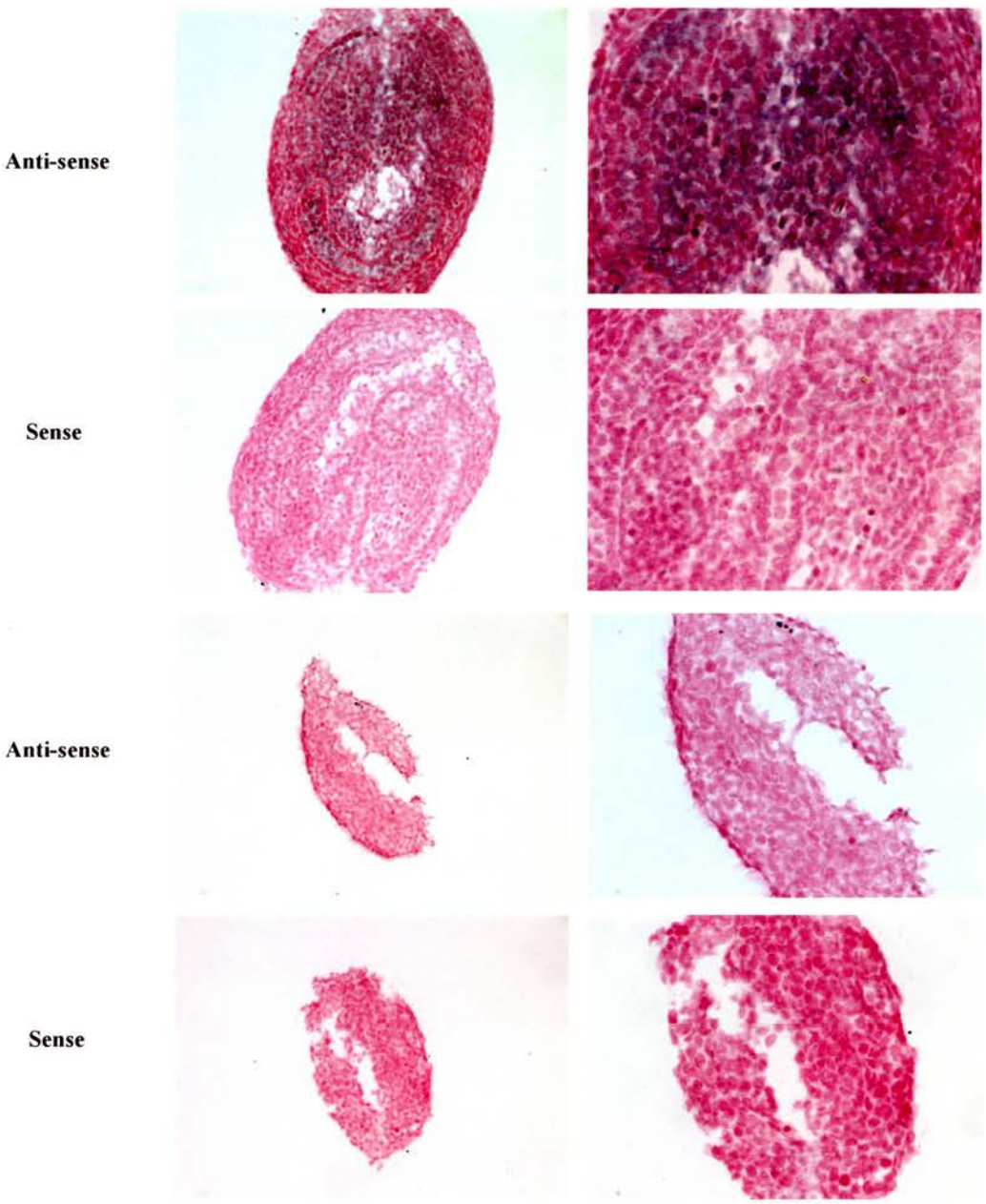


Figure 5.5 *Cbln1 In-situ* Hybridisation of E14.5 Gonadal Sections

RNA *in-situ* hybridisation of E14.5 gonadal sections using *Cbln1* sense and anti-sense probes. Top 4 panels show testis, bottom 4 panels show ovary. Staining only detected in testis using anti-sense probe.

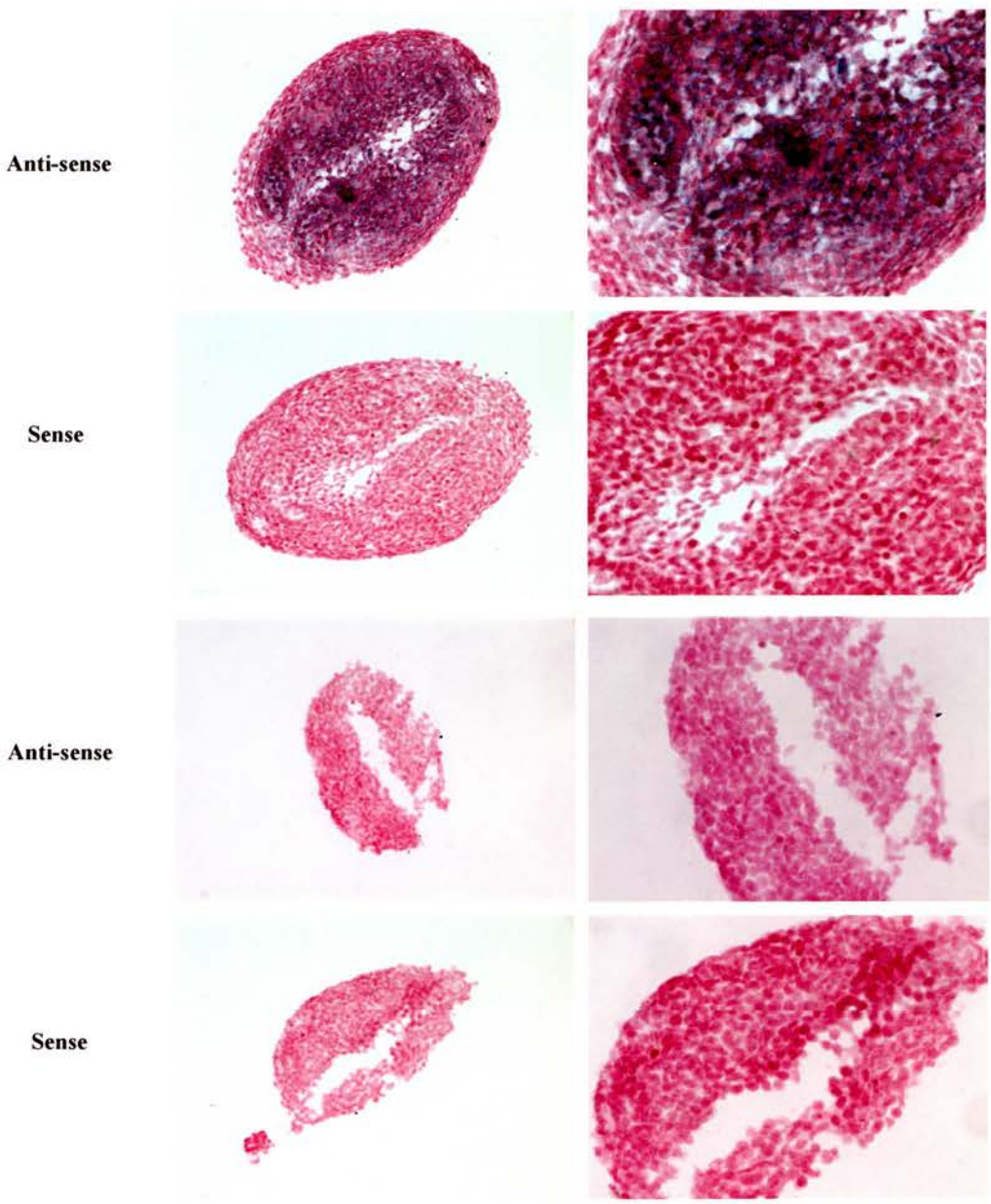


Figure 5.6 *Tpd52l1 In-situ* Hybridisation of E14.5 Gonadal Sections

RNA *in-situ* hybridisation of E14.5 gonadal sections using *Tpd52l1* sense and anti-sense probes. Top 4 panels show testis, bottom 4 panels show ovary. Staining only detected in testis using anti-sense probe.

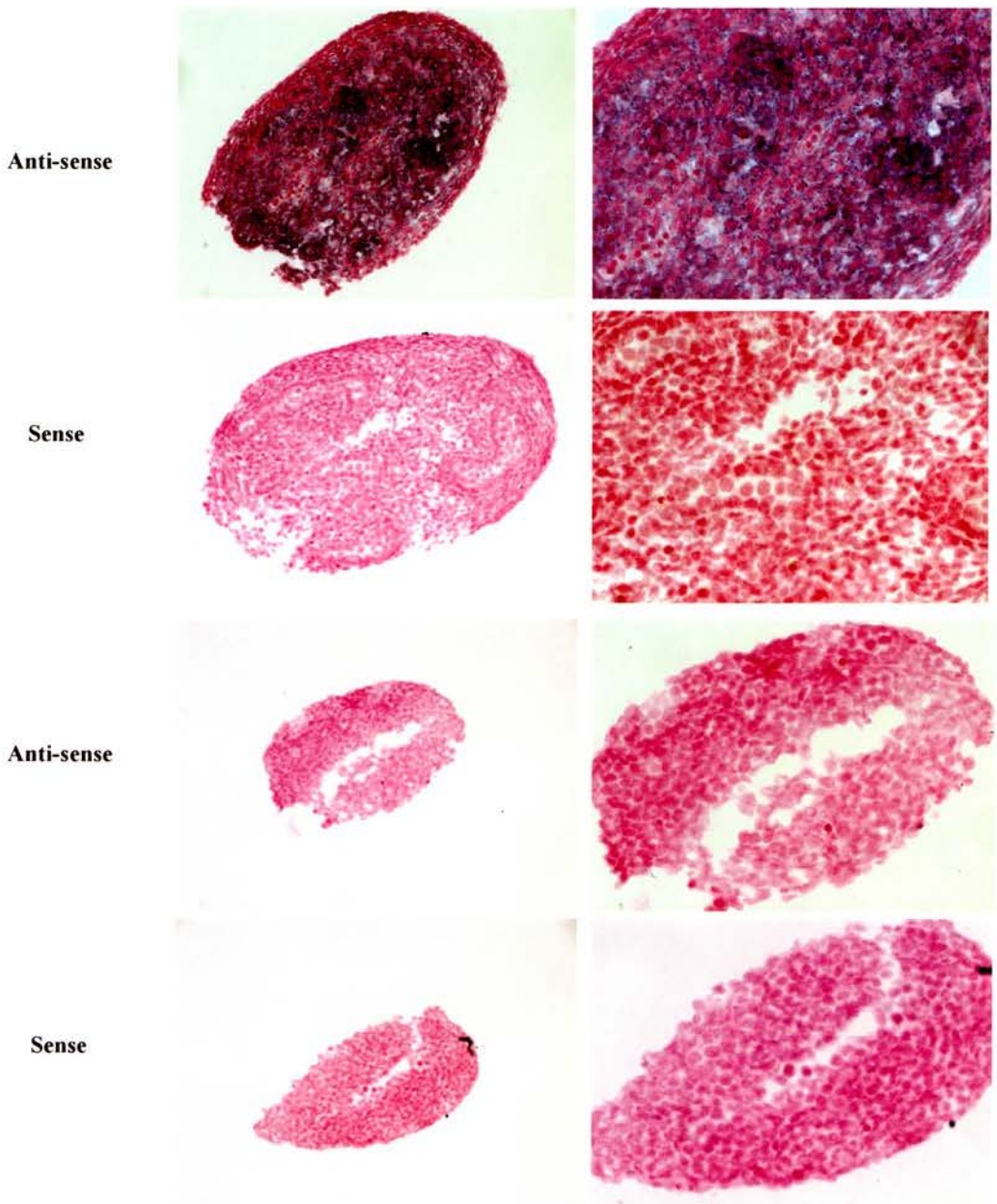


Figure 5.7 *Smoc2 In-situ* Hybridisation of E14.5 Gonadal Sections

RNA *in-situ* hybridisation of E14.5 gonadal sections using *Smoc2* sense and anti-sense probes. Top 4 panels show testis, bottom 4 panels show ovary. Staining only detected in testis using anti-sense probe.

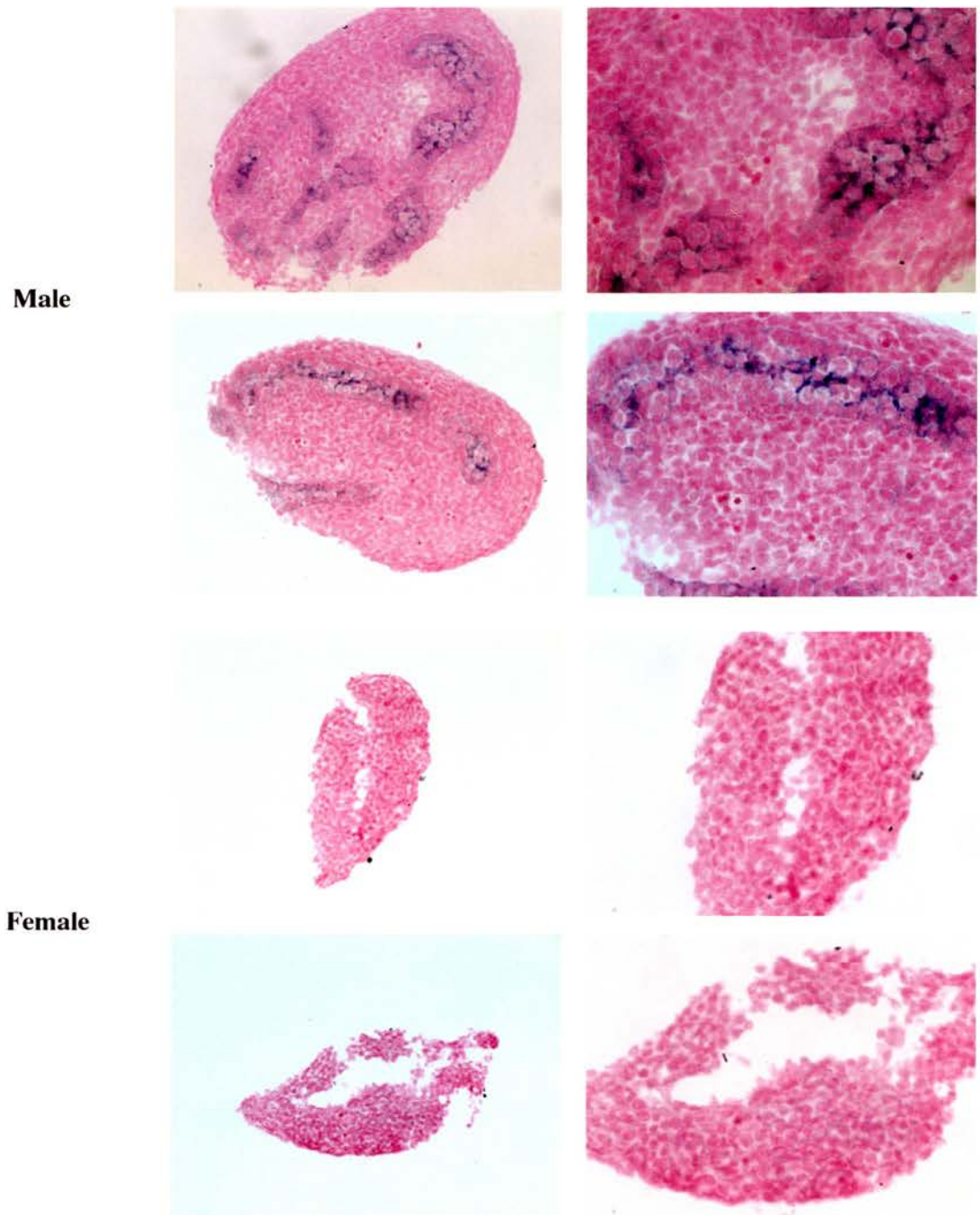


Figure 5.8 *Mis In-situ* Hybridisation of E14.5 Gonadal Sections

RNA *in-situ* hybridisation of E14.5 gonadal sections using *Mis* anti-sense probe (positive control). Top 4 panels show testis, bottom 4 panels show ovary. Staining only detected in sex cords of testis.

hybridisation, Figures 5.2, 5.3 and 5.4). Therefore, real-time PCR analyses of *Sry*, *Sox9* and *Mis* were also performed to relate the expression of the newly identified genes to that of known players in the sex determination pathway.

5.4.2 Real-time PCR Standardisation

In every real-time PCR run, a dilution series of cDNA is included to generate a standard curve for quantification of the cDNA samples. To ensure accurate quantification of each cDNA sample, cDNA amplification curves should fall within the standard curves of the dilution series. Figure 5.9 presents an example of the real-time amplification curves obtained from a real-time PCR run using the Lightcycler (Roche). Fluorescence (y axis) is plotted against the cycle number (x axis). Fluorescence is measured once per sample at the end of the annealing step for specific hybridisation probes and once per sample at the end of the elongation step for SYBR green probes. In both cases, the software calculates a linear regression line through the data points of the standard curves, which enables the respective concentration of each sample to be calculated. The slope of the regression line is also calculated by the software and represents the overall reaction efficiency. Manufacturer's state to achieve standard curve efficiency between 1.5 and 2.2, the slope has to be between –5.7 and –2.9 (Roche). The software also calculates an error value that indicates the variation between each sample, such as pipetting errors. An error value of up to 0.6 corresponds to a x value (concentration) deviation of up to 50% (Roche).

For all SYBR green reactions, a melting curve analysis should be added to the end of the real-time PCR run. A melting curve analysis shows the specificity of the reaction by demonstrating the specific amplicons (PCR products) generated in a graphical representation (Figure 5.10). Hence, any unspecific products such as primer dimers that are formed can be detected.

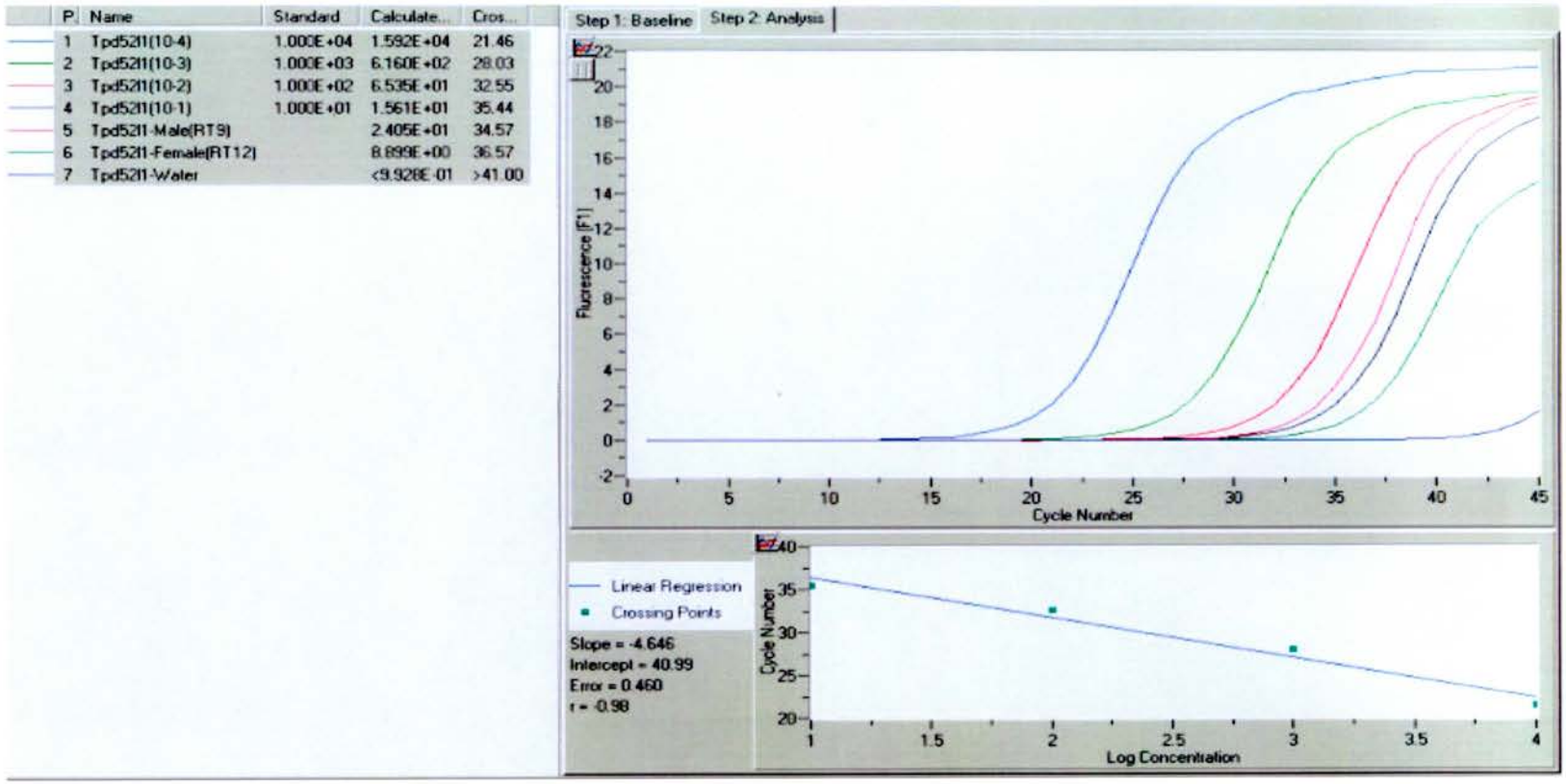


Figure 5.9 Real-time PCR Amplification Curves

An example of real-time PCR amplification curves (SYBR green system) obtained from a real-time PCR run using the Lightcycler (Roche). Amplification products of dilution series occur at regular intervals. cDNA samples are amplified within curves of dilution series. Insignificant amplification of water at end of run.

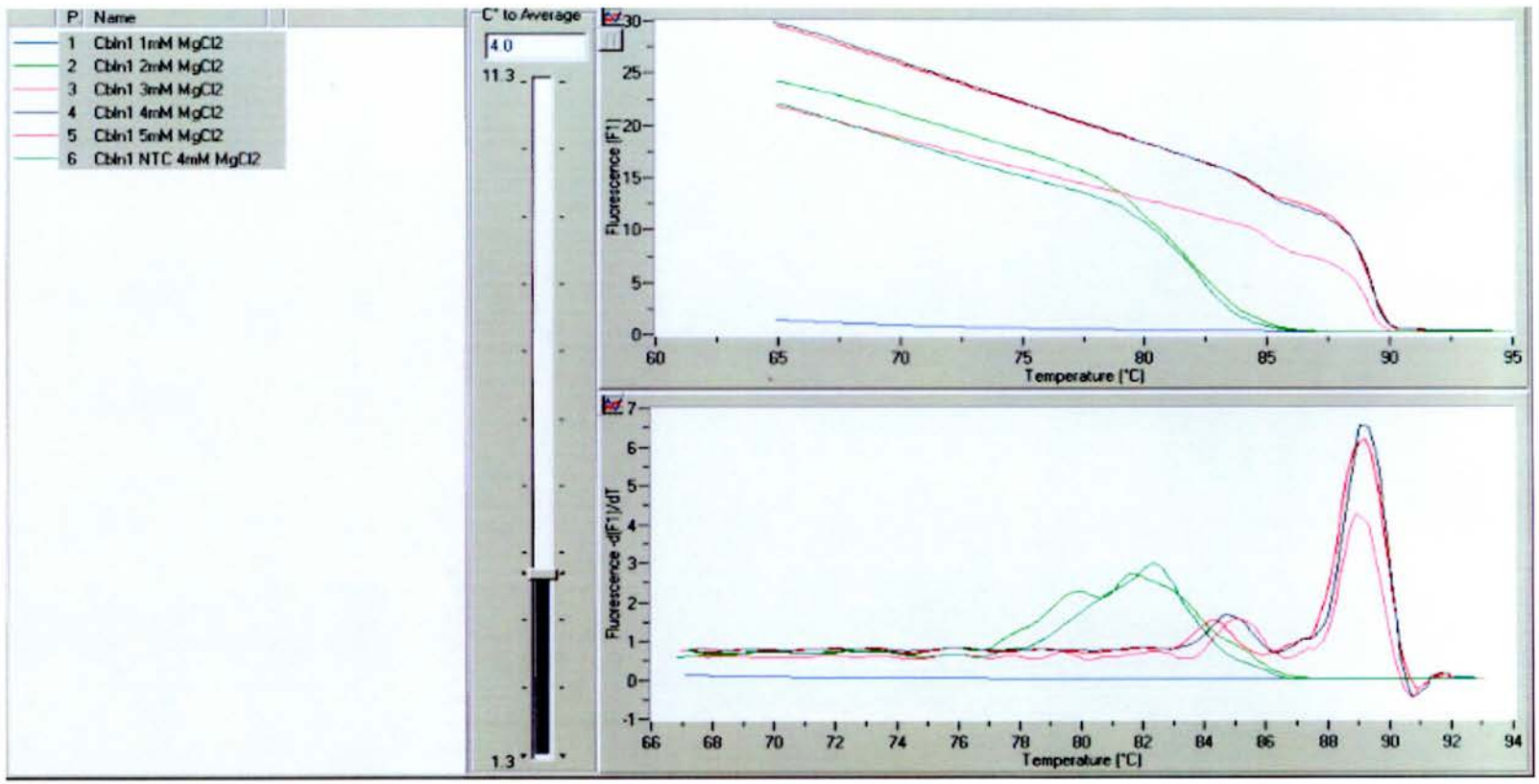


Figure 5.10 Real-time PCR Melting Curves

An example of melting curves obtained when using the SYBR green system. Specific and non-specific PCR products can be observed.

5.4.3 Real-time PCR Results – Quantification of Male and Female Expression Levels

The SYBR green system was used to quantify the expression levels of the potential sex determination candidates (*Cbln1*, *Tpd52l1* and *Smoc2*). Target sequences previously used for PCR (Figure 5.1) were also used for this real-time PCR analysis and expression levels were normalised against *Gapdh* (SYBR green system). Figure 5.11 shows the differential expression levels of *Cbln1*. The y axis represents the expression of *Cbln1* over the expression of *Gapdh* (normalised expression) against individual E12.5 male and female gonadal cDNA (x axis). Microarray data and *in-situ* hybridisation results have shown *Cbln1* to be increased in male gonads, therefore the expression of *Cbln1* was expected to be more highly expressed in the male sample and expressed lowly in the female sample. Indeed, the real-time PCR results showed expression levels to be high in male and low in female. More specifically, the difference in expression levels (fold change) between the two samples was found to be 5.75. This difference value is in good agreement with the fold change value of 6.7 obtained from the microarray analysis (chapter 3).

Figure 5.12 shows the differential expression levels of *Tpd52l1*. The E12.5 male and female gonadal cDNA used in the *Cbln1* real-time PCR analysis (as mentioned above) was also used for this analysis. Microarray data and *in-situ* hybridisation results have also shown *Tpd52l1* to be increased in male gonads, therefore the expression of *Tpd52l1* was expected to be more highly expressed in the male sample and less expressed in the female sample. Indeed, the real-time PCR results showed a high level of expression in the male sample and a low level of expression in the female sample. Furthermore, the difference in expression levels (fold change) between the two samples was found to be 3.05. This fold change value is comparable to the fold change value of 4.2 obtained from the micoarray analysis (chapter 3).

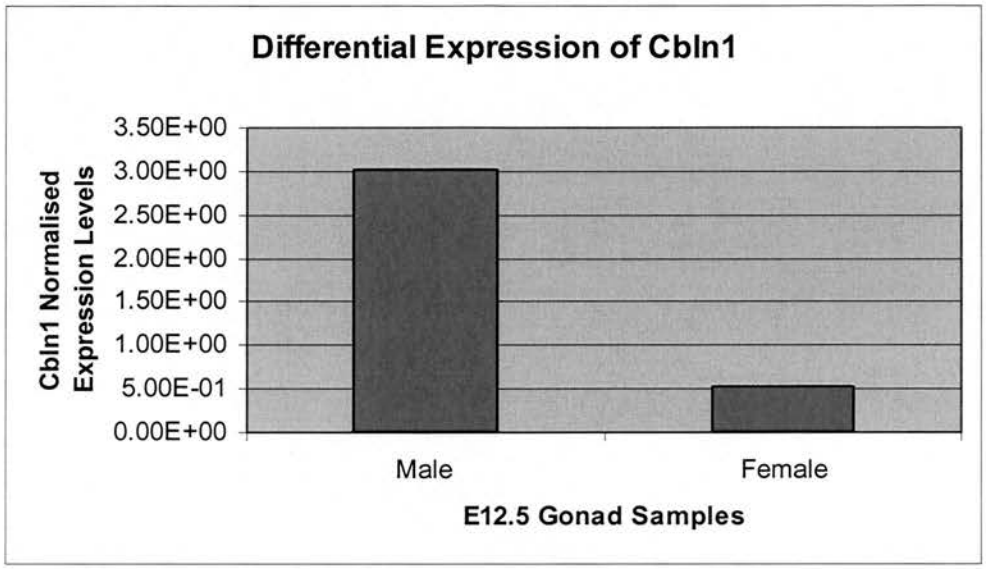


Figure 5.11 Differential Expression Levels of *Cbln1*

Total RNA was extracted from E12.5 male gonads and E12.5 female gonads. 5μg of each RNA sample was reverse transcribed to produce cDNA. Expression levels of *Cbln1* were measured by real-time PCR. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH.

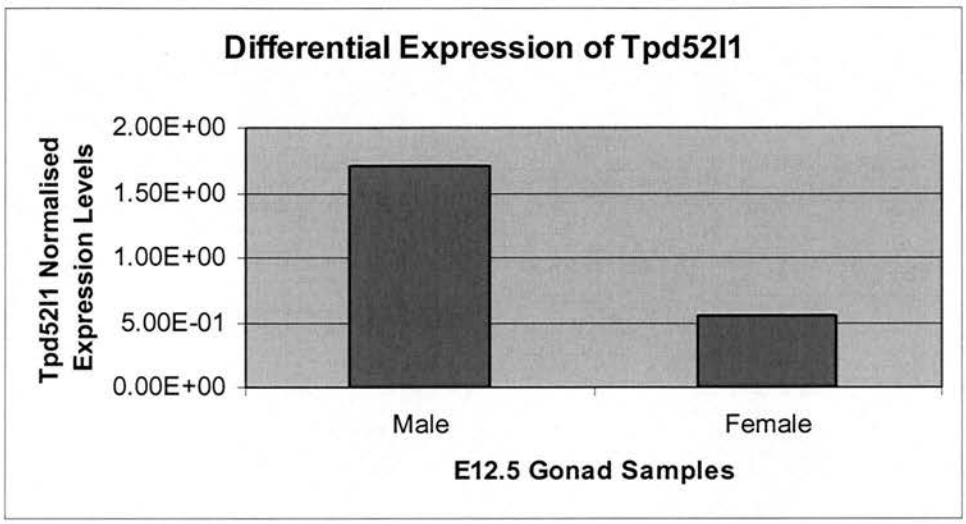


Figure 5.12 Differential Expression Levels of *Tpd52l1*

Total RNA was extracted from E12.5 male gonads and E12.5 female gonads. 5μg of each RNA sample was reverse transcribed to produce cDNA. Expression levels of *Tpd52l1* were measured by real-time PCR. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH.

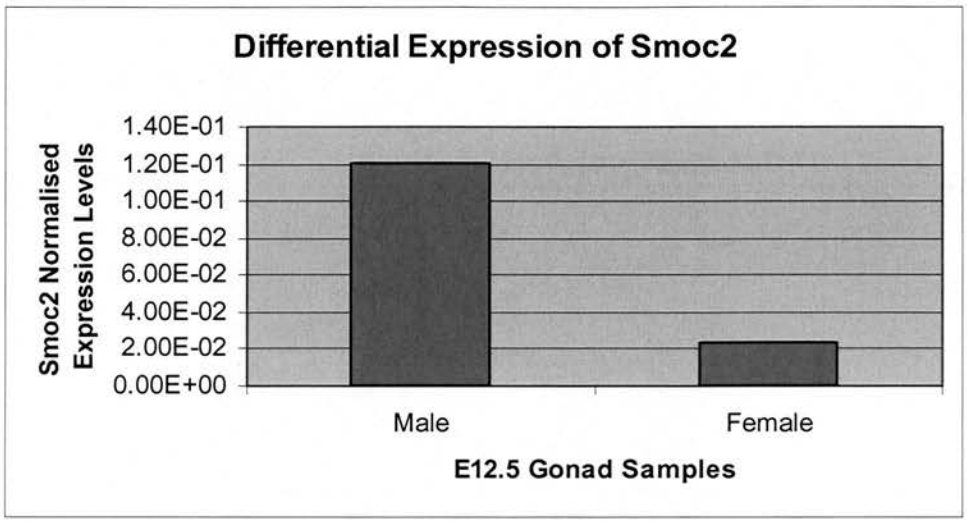


Figure 5.13 Differential Expression Levels of *Smoc2*

Total RNA was extracted from E12.5 male gonads and E12.5 female gonads. 5μg of each RNA sample was reverse transcribed to produce cDNA. Expression levels of *Smoc2* were measured by real-time PCR. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH.

Using the same E12.5 male and female gonadal cDNA samples as were used for the *Cbln1* and *Tpd52l1* real-time PCR analyses (as mentioned above), Figure 5.13 shows the differential expression levels of *Smoc2*. Similar to *Cbln1* and *Tpd52l1*, the real-time PCR results of *Smoc2* showed a high level of expression in male and a low level of expression in female. Furthermore, the difference in expression levels (fold change) between the two samples was found to be 5.20. This factor difference is relatively close to the fold change value of 4.1 obtained from the microarray analysis (chapter 3).

5.4.4 Real-time PCR Results – Temporal Expression Patterns

To begin the temporal expression pattern studies, genes with distinctive and known expression profiles in gonadal development (such as *Sry*, *Sox9* and *Mis*) would first be used to ensure that the use of this system was viable and that gene expression profiles could be obtained. Subsequently, temporal expression patterns of *Cbln1*, *Tpd52l1* and *Smoc2* could be related to these known male specific genes. Sequence-specific hybridisation probes for *Sry*, *Sox9* and *Mis* were used to quantify the expression level of the respective gene in each sample. All samples were normalised against a sequence-specific hybridisation probe for *Gapdh*. Gene expression profiles of each gene were plotted on separate line graphs, whereby the normalised expression levels (y axis) were plotted against the tail somite stage of each gonadal sample (x axis).

In males, the expression of *Sry* has been shown to start at E10.5, which then sharply increases to a peak at E11.5 and becomes switched off before E12.5 (Hacker et al., 1995; Jeske et al., 1995). In females, the *Sry* gene is absent and, as expected, *Sry* expression is not detected. Figure 5.14 represents the temporal expression profile of *Sry* within the gonadal samples assayed. Indeed, the temporal expression pattern obtained is as expected. More specifically, *Sry* is expressed at a significant level in the male sample at 17 tail somites (early E11.5) compared to the female sample.

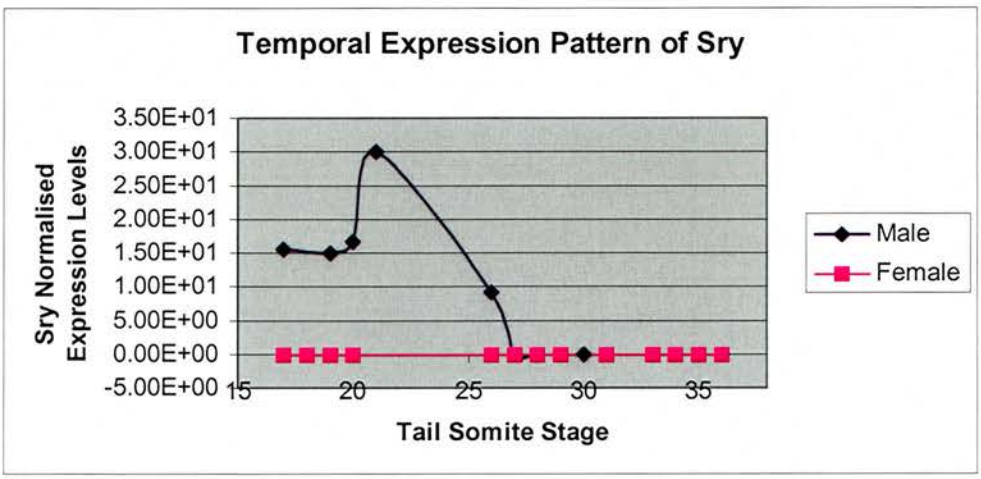


Figure 5.14 Temporal Expression Pattern of *Sry*

Total RNA was extracted from male and female gonad pairs, ranging from 17 to 36 tail somites. 10μg of each RNA sample was reverse transcribed to produce cDNA. Expression levels of *Sry* were measured by real-time PCR. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH. The temporal expression pattern of the male gonads is represented in blue and the temporal expression pattern of the female gonads is represented in pink.

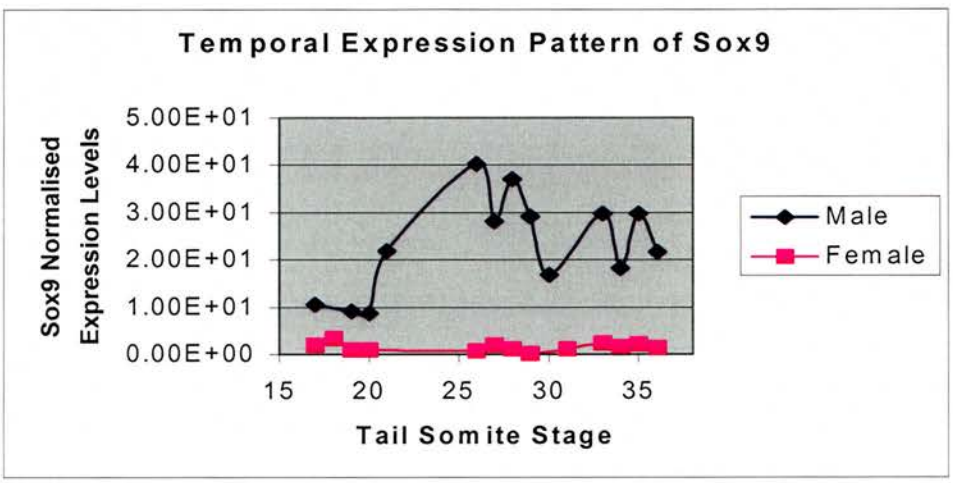


Figure 5.15 Temporal Expression Pattern of *Sox9*

Total RNA was extracted from male and female gonad pairs, ranging from 17 to 36 tail somites. 10μg of each RNA sample was reverse transcribed to produce cDNA. Expression levels of *Sox9* were measured by real-time PCR. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH. The temporal expression pattern of the male gonads is represented in blue and the temporal expression pattern of the female gonads is represented in pink.

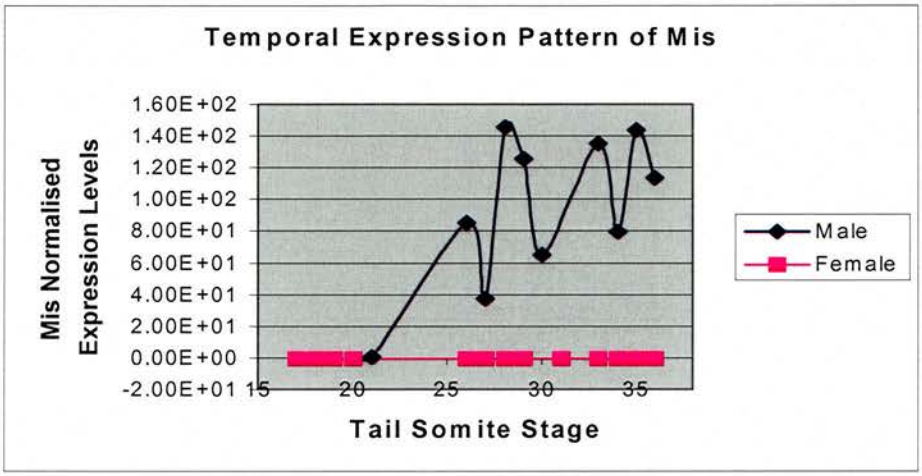


Figure 5.16 Temporal Expresssion Pattern of *Mis*

Total RNA was extracted from male and female gonad pairs, ranging from 17 to 36 tail somites. 10μg of each RNA sample was reverse transcribed to produce cDNA. Expression levels of *Mis* were measured by real-time PCR. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH. The temporal expression pattern of the male gonads is represented in blue and the temporal expression pattern of the female gonads is represented in pink.

This expression continues at the same level until 20 tail somites (E11.5), where it then sharply increases to a peak at 21 tail somites (E11.5). It then gradually decreases and is not expressed from 27 tail somites (early E12.5) to 36 tail somites (late E12.5). The expression of *Sry* was not detected in the female samples (ranging from 17 to 36 tail somites).

Small amounts of cytoplasmic Sox9 can be found in both male and female gonads at E10.5. However, the expression of *Sox9* is activated through the expression of *Sry* in XY gonads and so a high level of nuclear Sox9 is not detected until E11.5 in male gonads (de Santa Barbara et al., 2000; Morais da Silva et al., 1996). *Sox9* expression continues to be highly expressed at E12.5 to E13.5 as its expression becomes localised to the sex cords in the testis (Kent et al., 1996). As the expression of *Sox9* is increased in males, it is decreased in females. Figure 5.15 represents the temporal expression pattern of *Sox9* in the gonadal samples assayed. The temporal expression pattern obtained was found to be similar to the published data. Low levels of *Sox9* are detected in males between 17 to 20 tail somites (E11.5), levels then increased at 21 tail somites (E11.5) and continue to rise towards E12.5. Between early E12.5 to late E12.5 (26 to 36 tail somites) in the male samples, the expression of *Sox9* oscillates, indicating a degree of variability in *Sox9* expression at these time points. However, *Sox9* expression levels remain high in the male samples and low in the female samples throughout this time scale at each time point. In fact, the expression of *Sox9* in the female samples remains at a low level at all stages.

Mis is first detected at E11.5 in the male gonad (Hacker et al., 1995). From E12.5, *Mis* is highly detected in males and cannot be detected in females (Munsterberg and Lovell-Badge, 1991). *Mis* is then only down-regulated in males after birth and, in contrast, can be detected in females 6 days after birth (Munsterberg and Lovell-Badge, 1991). Figure 5.16 represents the temporal expression pattern of *Mis* in the gonadal samples assayed. The temporal expression pattern obtained showed a similar trend as determined in other studies. Whilst no expression of *Mis* could be detected at E11.5 (17 to 21 tail somites) in male or female samples, at 26 tail somites

(early E12.5), the expression of *Mis* becomes sexually dimorphic, where it increases dramatically in male samples and remains at undetectable levels in female samples. Although the expression deviates from the expected smooth curve between early E12.5 to late E12.5 (26 to 36 tail somites) in the male samples, the trend of the expression can be seen to be increasing towards the later tail somite stages. No *Mis* expression was detected in female samples throughout the time course of the expression profile (early E11.5 to late E12.5, 17 to 36 tail somites).

Taking all three gene expression profiles (*Sry*, *Sox9* and *Mis*) together, the results demonstrated that correct gene expression profiles could be obtained by real-time PCR analysis. Therefore, the same gonadal male and female cDNA samples were used for analysis of *Cbln1*, *Tpd52l1* and *Smoc2* gene expression profiles. The SYBR green system was used to quantify the expression level of the respective gene in each sample. All samples were normalised against *Gapdh* primers also using the SYBR green system.

Figure 5.17 shows the temporal expression pattern of *Cbln1*. The expression levels of *Cbln1* seem to be low in both male and female samples to begin with (at E11.5, from 17 to 21 tail somites). Then from E12.5 (26 to 36 tail somites), the expression of *Cbln1* becomes increased in male samples and decreased in female samples. Although, in the male profile, there is a slight dip in expression level at 29 and 30 tail somites, *Cbln1* is still more highly expressed in the male samples than in the female samples at those specific stages. It can also be seen that the expression of *Cbln1* is increasing in the male as the tail somite stage increases. In contrast, the expression of *Cbln1* in the female remains low throughout E12.5 (26 to 36 tail somites).

Figure 5.18 shows the temporal expression pattern of *Tpd52l1*. The expression levels of *Tpd52l1* seem to be quite dynamic and variable in both male and female samples. However, it should be noted that the *Tpd52l1* is still more highly expressed in male than in female from 21 tail somites (E11.5) onwards. There seem to be particular large differences in expression levels between male and female samples at

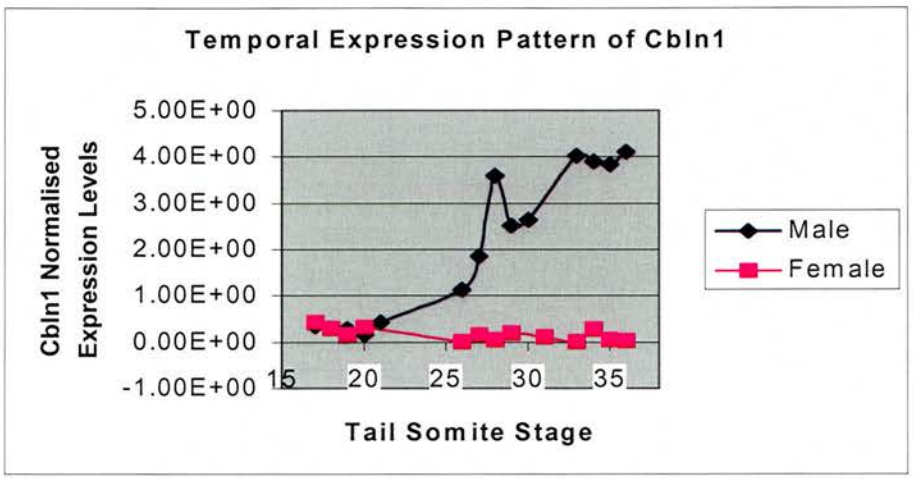


Figure 5.17 Temporal Expression Pattern of *Cbln1*

Total RNA was extracted from male and female gonad pairs, ranging from 17 to 36 tail somites. 10μg of each RNA sample was reverse transcribed to produce cDNA. Expression levels of *Cbln1* were measured by real-time PCR. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH. The temporal expression pattern of the male gonads is represented in blue and the temporal expression pattern of the female gonads is represented in pink.

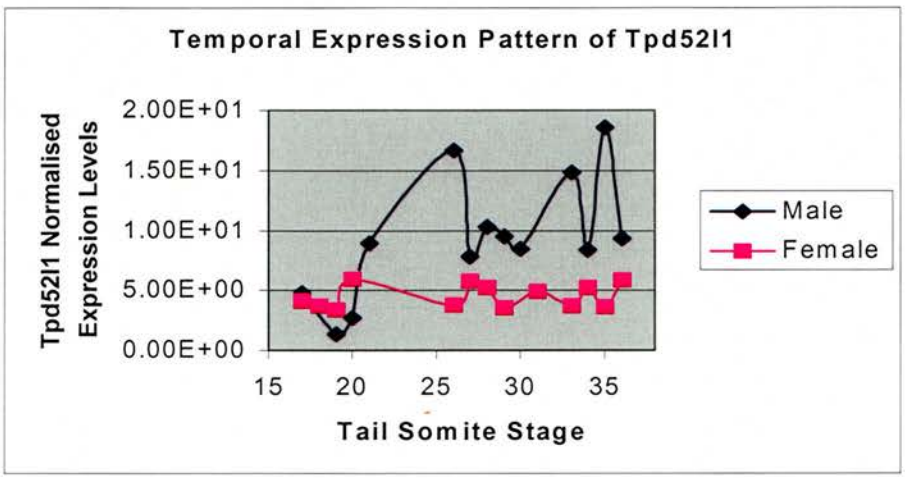


Figure 5.18 Temporal Expression Pattern of *Tpd52l1*

Total RNA was extracted from male and female gonad pairs, ranging from 17 to 36 tail somites. 10μg of each RNA sample was reverse transcribed to produce cDNA. Expression levels of *Tpd52l1* were measured by real-time PCR. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH. The temporal expression pattern of the male gonads is represented in blue and the temporal expression pattern of the female gonads is represented in pink.

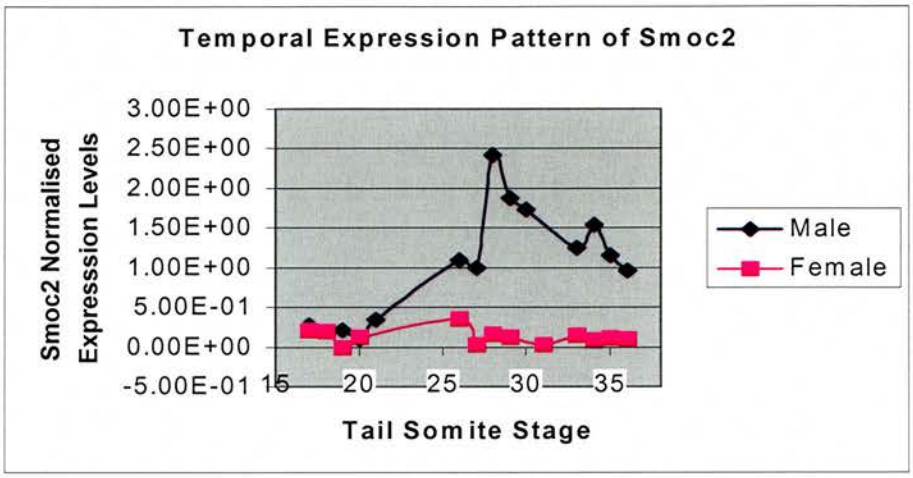


Figure 5.19 Temporal Expression Pattern of *Smoc2*

Total RNA was extracted from male and female gonad pairs, ranging from 17 to 36 tail somites. 10μg of each RNA sample was reverse transcribed to produce cDNA. Expression levels of *Smoc2* were measured by real-time PCR. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH. The temporal expression pattern of the male gonads is represented in blue and the temporal expression pattern of the female gonads is represented in pink.

25 tail somites (early E12.5) and 35 tail somites (late E12.5). This degree of variability is also observed at a lesser extent in female samples as the expression of *Tpd52l1* remains at a constant low level throughout the time scale of 17 to 36 tail somites.

Figure 5.19 shows the temporal expression pattern of *Smoc2*. The expression levels of *Smoc2* seem to be similarly expressed at low levels in both male and female samples to begin with (at E11.5, from 17 to 21 tail somites). Then at early E12.5 (26 tail somites), *Smoc2* expression levels are increased in male while they remain at low levels in female. In particular, there seems to be a peak of expression at E12.5 (28 tail somites) in the male sample which gradually decreases from that point onwards towards late E12.5 (36 tail somites). At 36 tail somites, *Smoc2* expression is still at a significantly high level in male. Therefore, *Smoc2* is still sexually dimorphic at this time point. Moreover, the expression levels of *Smoc2* in female remain low throughout the time scale of 17 to 36 tail somites.

5.5 Discussion

PCR amplification and gel electrophoresis gave an indication of the expression patterns of *Cbln1*, *Tpd52l1* and *Smoc2* within various E12.5 embryonic organs. Unfortunately, RNA *in-situ* hybridisation of whole embryos (E10.5, E11.5 and E12.5) was unsuccessful and did not confirm or provide further insight into their expression patterns. Interpretation of the data was difficult due to technical difficulties with detection as background staining occurred. This may be a result of the probe design or optimisation of the whole-mount *in-situ* hybridisation protocol is required. Both these factors are also true for the whole-mount *in-situ* hybridisation of gonads. Although each probe is more highly expressed in male gonads than female gonads, the expression patterns are diffuse and could be more specific. This is particularly seen in *Cbln1*, as Menke and Page showed *Cbln1* to be expressed in a

more specific pattern within the interstitial cells of an E13.5 testis (Menke and Page, 2002).

Although RNA *in-situ* hybridisation of E14.5 gonadal sections showed staining in the testis and no staining in the ovary, the specific cell type within which each gene is expressed could not be determined. This is again possibly due to technical difficulties or the probe itself. Whilst all sense probes were negative (no staining), anti-sense probes showed a diffuse staining in the testis and staining within an individual cell type could not be distinguished. Different sets of probes should be designed and optimisation of the protocol should be carried out to establish a more specific expression pattern of each gene.

Quantitative real-time PCR analysis has further verified the sexually dimorphic patterns, showing *Cbln1*, *Tpd52l1* and *Smoc2* to be more highly expressed in male gonads. Interestingly, the fold changes measured by real-time PCR analysis are comparable to the fold changes obtained from the Affymetrix microarray screen 1, thus confirming the quantitative expression levels between male and female gonads.

In addition, the real-time PCR analyses of *Cbln1*, *Tpd52l1* and *Smoc2* have allowed these individual temporal expression patterns at the initial stages of sex differentiation (early E11.5 to late E12.5) to be generated. These gene expression profiles have given an indication of their expression level and expression pattern throughout the specific time course assayed. It should be noted that the temporal expression patterns of *Cbln1*, *Tpd52l1* and *Smoc2* did not display large differences between male and female samples throughout E11.5 (17 to 21 tail somites). Therefore, the marginal staining detected within E11.5 male gonads by whole-mount *in-situ* hybridisation of these genes may be due to weaker expression levels of these genes at this time point.

Various degrees of variability between each consecutive sample, as observed within the *Tpd52l1* temporal expression pattern in particular, may be caused by dissection

artefacts and impurities within individual samples. Nevertheless, a clear trend can be observed within each time-course of gene expression. Although variability may occur due to the use of individual samples, one would have expected a smooth curve, whereas the data displayed a more scattered pattern. It should be noted that the temporal expression pattern of *Sox9* and *Mis* also showed this scatter, yet it is known that this is not the case *in vivo*. Thus, it is important to observe the trend of each temporal expression pattern, rather than examining the expression levels at individual time points.

Comparing the temporal expression patterns of *Cbln1*, *Tpd52l1* and *Smoc2* to those of *Sry*, *Sox9* and *Mis*, similar trends can be identified. For example, the time-course of *Cbln1* gene expression (Figure 5.17) resembles the temporal expression pattern of *Mis* (Figure 5.16), from the respect that both genes become sexually dimorphic at early E12.5 (26 tail somites) and continues to increase in the male. However, little or no expression can be detected in females. As *Cbln1* is expressed at approximately the same time and has a similar expression pattern to *Mis*, this suggests that there may be a correlation between these two genes. Further investigations are needed to certify and clarify this preliminary finding.

In addition, the advantage of examining the initial stages of sex differentiation allows early targets of sex determination to be investigated. In this case, it was interesting to explore whether *Cbln1*, *Tpd52l1* and *Smoc2* are early targets within the male development pathway. As expression levels of all three genes seem to increase when the expression of *Sry* is almost switched off, it is unlikely that they are direct downstream targets of *Sry*. However, they may be downstream targets of *Sox9*. Promoter-binding studies need to be conducted to provide concrete evidence of such regulation (further discussed in chapter 7).

Chapter 6 Gene-specific Knockdown in Gonadal Organ Culture by siRNA

Chapter 6 Gene-specific Knockdown in Gonadal Organ Culture by siRNA

6.1 Introduction

In this chapter, genes known to be involved in sex determination (such as *Sry* and *Sox9*) will be targeted by duplex siRNAs. These siRNAs will first be tested on TM4 cells grown in monolayer culture and then tried on gonadal organ cultures. The outcome of these experiments may establish the use of sex-specific siRNAs on gonad cultures and may provide a faster method of investigating sex-determining genes. This will enable many sex-determining candidates to be assayed and the downstream effects of silenced genes can also be analysed, leading to a better understanding of the sex developmental pathways.

6.2 Testing of siRNAs on TM4 Cells

6.2.1 TM4 Cell Line

The TM4 cell line is a mouse sertoli cell line (purchased from ECACC). Firstly, these TM4 cells were tested to assess whether they expressed the male specific genes, *Sry* and *Sox9*. The RNA of a confluent plate of cells was extracted and cDNA was synthesised (as described in section 2.2) for PCR amplification. An aliquot of the RNA was saved to run alongside the cDNA sample to test for DNA contamination. This is particularly important as the *Sry* gene is intron-less. Known male samples were used as a positive control to ensure both sets of primers were functioning. Figure 6.1 shows the results of the PCR on a DNA agarose gel. Lanes 1-4 represent *Sry* and lanes 5-8 represent *Sox9*. All controls showed the expected results. In the TM4 cells cDNA samples, lane 5 was positive which indicates that the TM4 cells express *Sox9*. However, lane 1 was negative indicating that the TM4 cells did not express *Sry*. Therefore, only Sox9 siRNAs could be used on the TM4 cells.

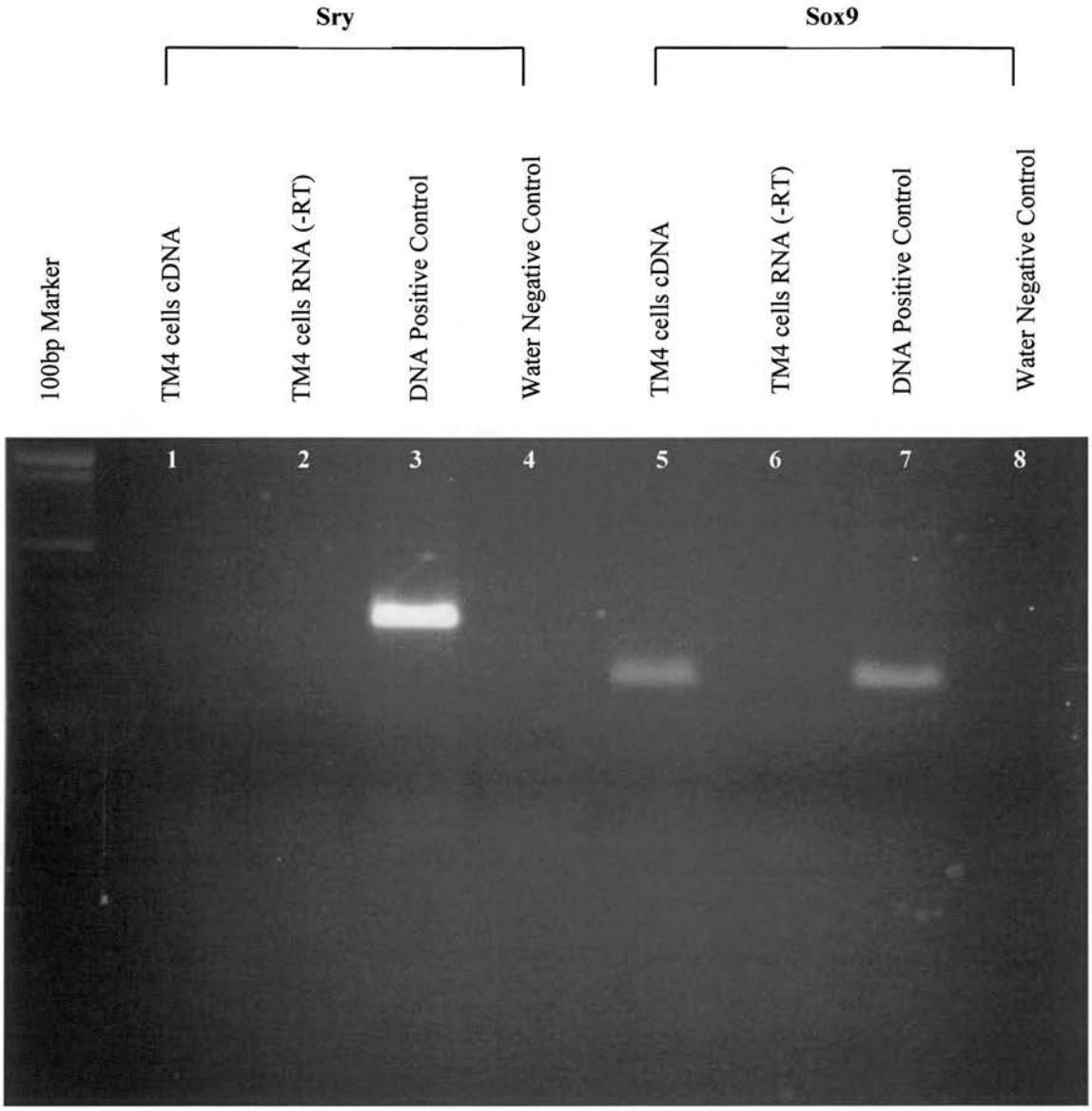


Figure 6.1 *Sry* and *Sox9* Expression Analysis of the TM4 Cell Line

RT-PCR analysis of the TM4 cell line using *Sry* and *Sox9* primers. Total RNA was collected from a confluent plate of cultured TM4 cells. 5μg of RNA was reverse transcribed to produce cDNA. The same aliquot of DNA/RNA template was used for both sets of primers. Results show the TM4 cell line to express *Sox9*, but not *Sry* (lane numbers correspond to text).

6.2.2 Experimental Design and Results

Four different siRNA oligonucleotide templates were designed and produced for *Sox9* (as described in section 2.7.1). These have different target sequences for *Sox9* to test for the most effective siRNA. They are referred to as Sox9(10), Sox9(13), Sox9(33) and Sox9(44) (individual target sequences can be found in Appendix 3). Cells were cultured and plated at an equal density as described in section 2.6.1 and the four siRNAs were individually applied to separate plates of cells. A negative control of the transfectant (oligofectamine) alone was included to assess if it is toxic to the cells and to ascertain that any reduction in expression levels was due to the siRNA and not the transfectant causing apoptosis in the cells. In addition, a plate of untreated cells (no siRNA or transfectant) was also included to account for the standard level of expression in each experiment. After 24 hours post-transfection with siRNA, cells were prepared for RNA extraction and cDNA synthesis (as described in section 2.2). Using sequence-specific hybridisation probes for *Sox9* (Appendix 2), levels of each siRNA knockdown were quantitatively measured by real-time PCR (as described in section 2.9.1). Expression levels of each sample was determined in the same run and normalised by measuring the amount of Gapdh, (Appendix 2).

The above experiment was repeated before analysing the results. Variable expression levels were observed between the two sets of data (data not shown). It was unclear whether this variability was due to differences in cell density, cell death or another factor within the individual experiments. Therefore, replicates of each sample were performed in one experiment. Figure 6.2 shows a bar chart of the Sox9 siRNA knockdown observed. The replicates of each sample run alongside its corresponding sample. All expression levels have been normalised against Gapdh to account for any differences in cDNA concentration between each sample. All four Sox9 siRNAs showed a reduction in *Sox9* expression levels compared to the untreated cells and transfectant (oligofectamine) alone cells. Knockdown in expression levels between replicates were comparable, however all four Sox9

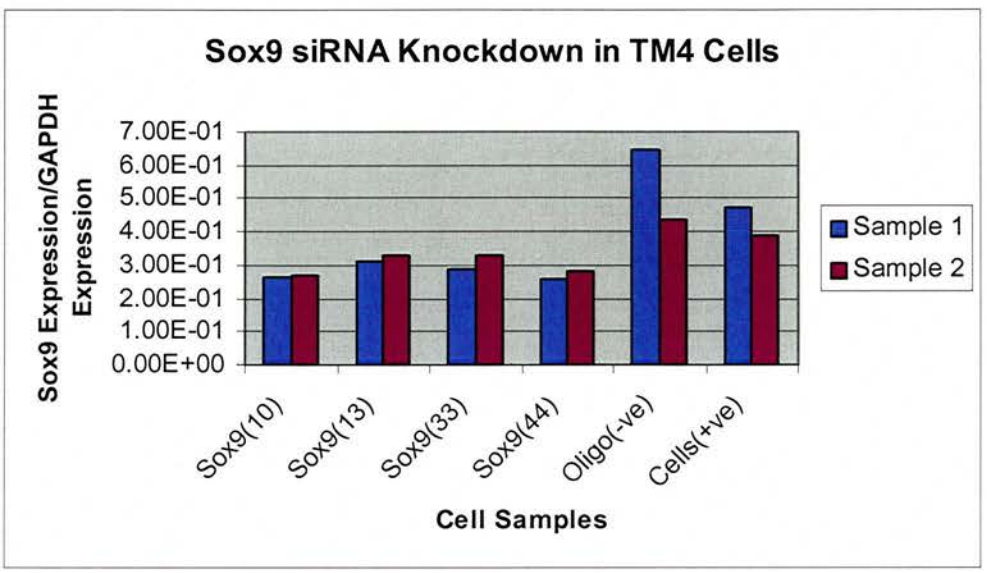


Figure 6.2 Real-time PCR Analysis of Sox9 Expression in TM4 Cells Following siRNA Treatment

Total RNA was collected from Sox9(10), Sox9(13), Sox9(33) and Sox9(44) siRNA treated cells, along with transfectant (Oligofectamine; Oligo(-ve)) only and untreated cells (cells(+ve)). 5μg of RNA was reverse transcribed to produce cDNA. Real-time PCR analysis for *Sox9* was performed. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH.

siRNAs seemed to suppress *Sox9* expression to a similar extent. Therefore, from the four Sox9 siRNAs, no particular siRNA was found to be more effective at repressing *Sox9* expression in the TM4 cell system. Although the siRNAs did not completely repress *Sox9* expression levels, preliminary results indicated that the Sox9 siRNAs can individually knockdown the expression of *Sox9* by 15 - 45%. This encouraged proceedings with the intended siRNA assays on gonadal organ cultures.

6.3 Repression of Sox9 Expression by siRNA in Gonad Cultures

6.3.1 Experimental Design

As all four Sox9 siRNAs seemed to have similar knockdown effects on the cell system, but none seemed to repress the expression of *Sox9* completely, all four Sox9 siRNAs were used in the same cocktail with the hope to cause a more efficient knockdown effect on the gonad culture system. Wild-type E12.5 urogenital ridges were dissected for cultures which would be quantitatively analysed by real-time PCR (as described in section 2.9.1). As for the cell system, sequence-specific hybridisation probes for *Sox9* (Appendix 2) were used and each sample was normalised by measuring the amount of Gapdh (Appendix 2). *Sox8*$^{+/+}$ E12.5 urogenital ridges were dissected for cultures which would be phenotypically analysed by X-gal staining. E12.5 was the chosen time-point as the expression of Sox9 is at a high level at this stage and the effect of siRNA is believed to be most effective when the targeted gene is highly expressed, allowing the siRNA to produce a more efficient knockdown effect. Gonad cultures were treated as described in section 2.6.2. The gonad plus mesonephros (urogenital ridge) was used for all gonadal organ cultures. Each pair of gonads and mesonephroi was separated, where one gonad plus its mesonephros was treated with the Sox9 siRNA cocktail mix and the corresponding gonad and mesonephros from the same embryo was treated with the Pax2 control siRNA. This ruled out embryonic stage differences and allowed gonad cultures to be directly compared. Pax2 siRNA (kind gift of Dr Peter

Hohenstein) was used as a control since gonads do not express *Pax2* and consequently are unaffected in *Pax2* knockout mice (Torres et al., 1995). Therefore, the Pax2 siRNA treated gonad culture can be considered as the control sample in each embryo to compare the standard level of Sox9 expression against the corresponding Sox9 siRNA treated gonad cultures.

6.3.2 Real-time PCR Results

Wild-type E12.5 gonad cultures were incubated for 24 hours after siRNA treatment and should therefore be approximately equivalent to E13.5 gonads. Figure 6.3 shows a bar chart of the expression levels of *Sox9* produced by gonad cultures treated with the Sox9 siRNA cocktail and the Pax2 control siRNA. All ten embryos were obtained from the same litter. Genotyping of each embryo revealed eight of these to be male and two to be female. Although *Sox9* is a male sex determination gene and is not expressed in female gonads at E12.5, female littermates were included in this primary assay to compare the differences in expression levels between the two sexes and to assess the validity of the results against previously established data. Indeed, in the two female gonad cultures (C97 and C102) expression of *Sox9* was undetectable (Figure 6.3), whereas the eight male gonad cultures showed positive levels of *Sox9* expression. From these eight male gonad cultures, six showed a reduction in *Sox9* expression levels following the Sox9 siRNA cocktail treatment when compared to their contra-lateral gonad cultures treated with the Pax2 siRNA, which showed higher levels of *Sox9* expression. However, the knockdown levels of *Sox9* varied amongst these six samples. Samples C94, C96 and C100 showed the greatest reduction in *Sox9* expression, whereby Sox9 siRNA treated samples were approximately 50% lower than the corresponding Pax2 control treated sample. Samples C98 and C99 showed a 38% and 39% reduction respectively. Whereas the Sox9 siRNA sample in C103 only showed a 20% knockdown effect. Before repeating these preliminary results, the phenotypic effect of the Sox9 siRNA cocktail was investigated by X-gal staining of the gonad cultures. This will allow more

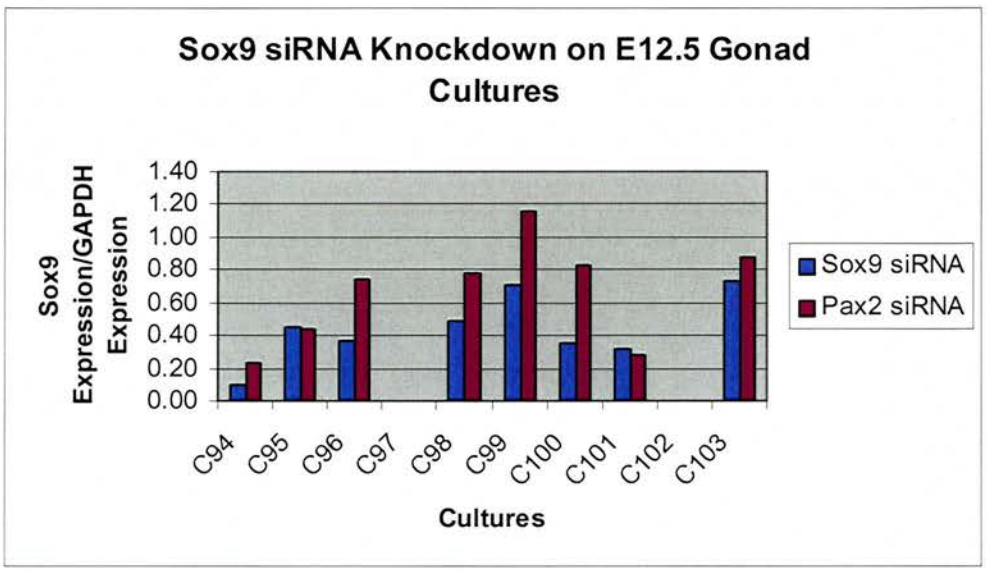


Figure 6.3 Real-time PCR Analysis of Sox9 Expression in E12.5 Gonad Cultures Following siRNA Treatment

Total RNA was collected from wild-type E12.5 gonads cultured for 24 hours following Sox9 and Pax2 siRNA treatment (corresponding gonads were used). 5µg of RNA was reverse transcribed to produce cDNA. Real-time PCR analysis for *Sox9* was performed. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH. Samples C97 and C102 are female gonad cultures. All other samples are male.

concrete conclusions to be drawn and aid the decision on the approach for the next step.

6.3.3 X-Gal Staining Results

Culturing of gonads commenced at E12.5 (stage of dissection) and was incubated for 24 hours following siRNA treatment. Therefore, the stage of the gonad culture should have reached approximately E13.5 when culturing stopped. Normally, *in vivo*, phenotypic differences between testis and ovary can be observed in E13.5 gonads. However, developmental delays are expected to occur with *in vitro* systems. Thus, *Sox8* mice (from Dr Michael Wegner) were used as they contain a *lacZ* marker fused with the *Sox8* start codon. Since *Sox8* is only expressed in the developing testis from E12 (Sock et al., 2001), following the activation of *Sox9*, all male gonad cultures should be *LacZ* positive.

To confirm this hypothesis and to observe the growth of the gonadal organ cultures phenotypically, E11.5 and E12.5 male and female gonads (with mesonephroi attached) were taken for X-gal staining (as described in section 2.8). At the same time, from the same litter, E11.5 and E12.5 male and female gonads were placed into culture. After three days (E12.5) and four days (E11.5) culturing, these explants were also taken for X-gal staining. Figure 6.4 shows the results for the E12.5 X-gal staining of both the gonads and the cultures. The difference in male and female samples is clearly observed by the *lacZ* positive staining in male samples and the *lacZ* negative staining in female samples. Although the sex cords are not formed in an organised manner compared to normal E15.5 gonads, the structure of the sex cords can be clearly distinguished. Figure 6.5 shows the results for the E11.5 X-gal staining of both the gonads and the cultures. There was no staining in either male or female gonads. This was expected as the expression of *Sox8* is not switched on until E12 (Sock et al., 2001). However, after 3 days in culture, *lacZ* positive cells were observed in XY gonad cultures but not in XX gonad cultures. This change in E11.5 male cultures indicated that the gonadal organ culture system is working well.

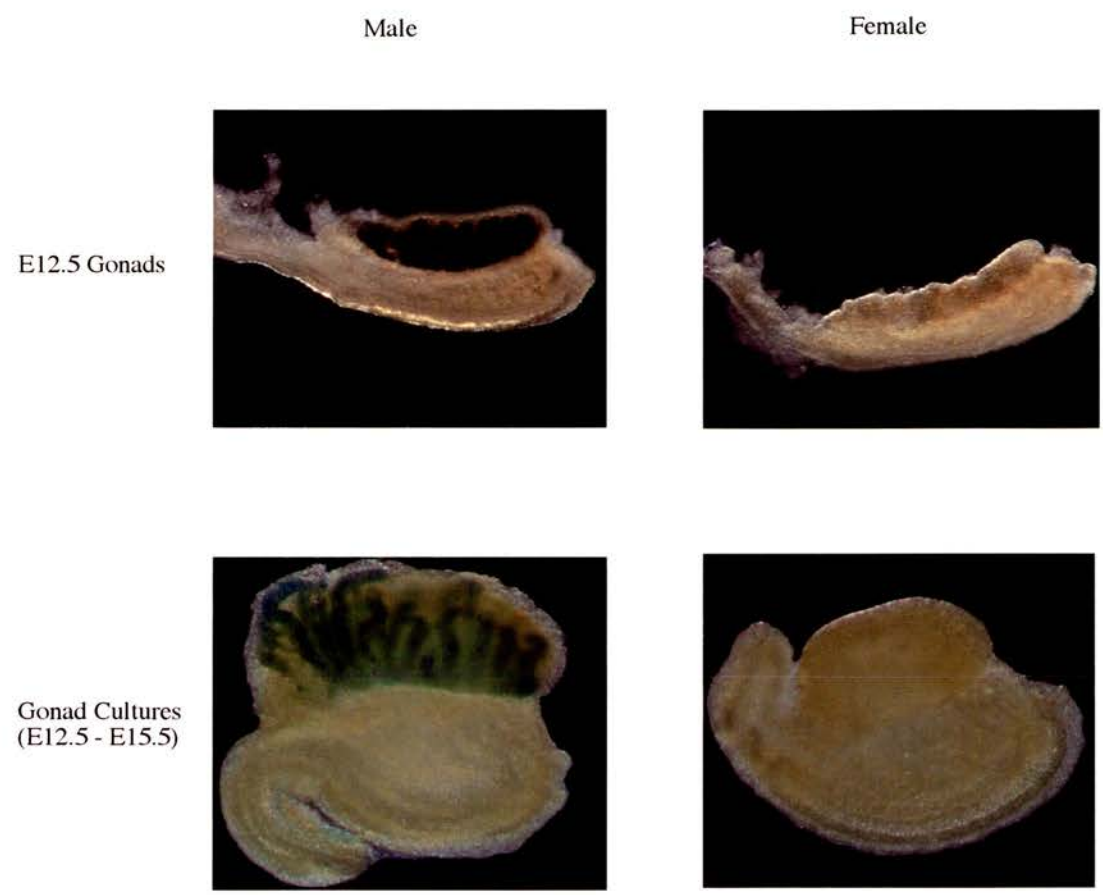


Figure 6.4 X-Gal Staining of E12.5 Gonads and Gonad Cultures

Heterozygous *Sox8* embryos were dissected at E12.5. The top two panels show X-gal staining of a male gonad (left) and female gonad (right). The bottom two panels show X-gal staining of male (left) and female (right) gonads cultured from E12.5 to E15.5 (3 days).

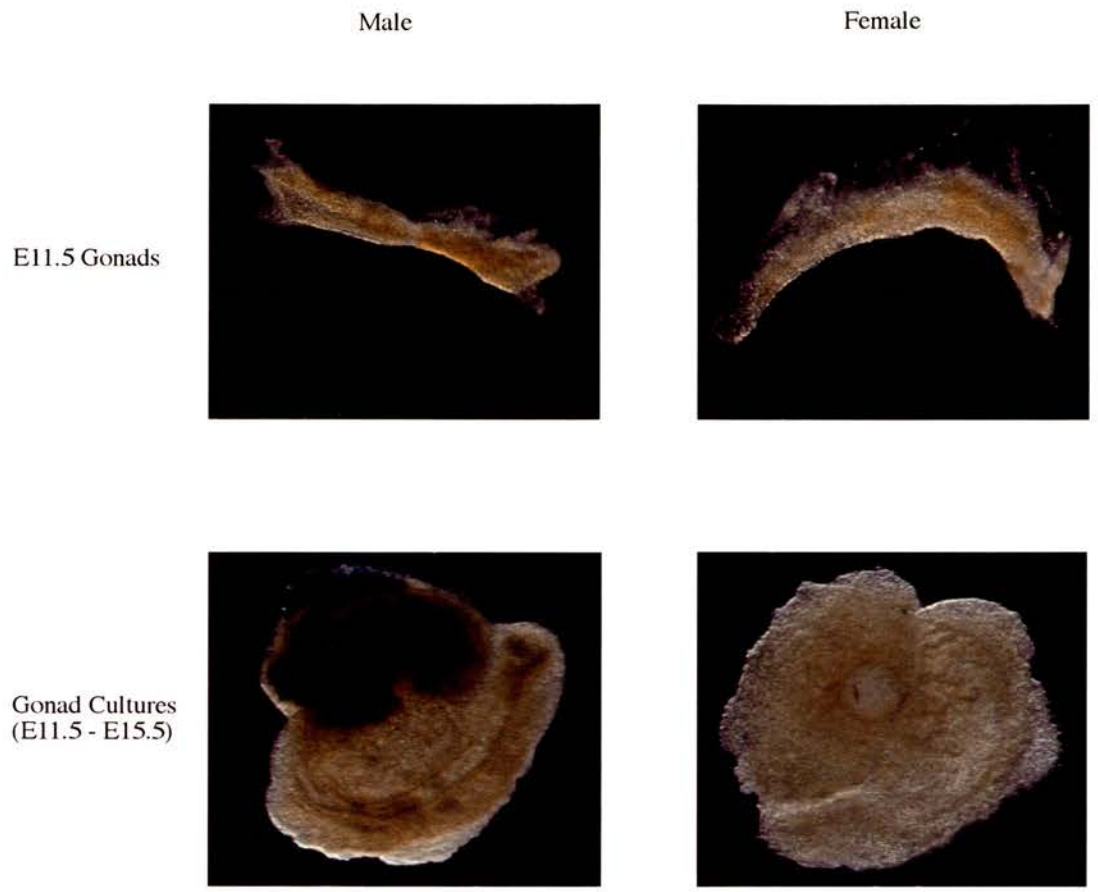


Figure 6.5 X-Gal Staining of E11.5 Gonads and Gonad Cultures

Heterozygous *Sox8* embryos were dissected at E11.5. The top two panels show X-gal staining of a male gonad (left) and female gonad (right). The bottom two panels show X-gal staining of male (left) and female (right) gonads cultured from E11.5 to E15.5 (4 days).

Although distinct sex cords could not be observed in the E11.5 male cultures compared to the E12.5 male cultures (as mentioned above), a clear *lacZ* positive stain is found. Therefore, the male gonad cultures can be differentiated from the female gonad cultures on the basis of *lacZ* staining.

To test whether Sox9 siRNA interfered with the expression of *Sox8*, E12.5 gonad cultures were treated with the Sox9 siRNA cocktail (Figure 6.6). All gonad cultures were *lacZ* positive. This indicated that the Sox9 siRNA cocktail was not sufficient or could not effectively suppress the expression of *Sox8* within the male E12.5 gonad cultures. A second independent experiment gave the same results. All gonad cultures were *lacZ* positive in a further three Sox9 siRNA cocktail treated male gonad cultures (results not shown).

Taking these results and the quantitative real-time PCR result together, it was concluded that the Sox9 siRNA cocktail does not effectively knockdown the expression of *Sox9* in gonadal organ cultures. This may be due to the fact that by E12.5 *Sox9* has already activated *Sox8* and subsequently the expression levels of *Sox8* may be maintained despite the reduction of *Sox9*. To address this issue, we performed the same experiments using gonad cultures from E11.5.

6.3.4 Sox9 siRNA on E11.5 Gonad Cultures

E11.5 gonad cultures were incubated for 24 hours after siRNA treatment and should therefore be approximately equivalent to E12.5 gonads. The X-gal staining procedure revealed *lacZ* positive staining in Sox9 siRNA treated cultures as well as Pax2 siRNA treated cultures (Figure 6.7). This indicated that the Sox9 siRNA cocktail did not repress the expression of *Sox9* to a sufficient extent in order to prevent its activation of *Sox8* and consequently the expression of *lacZ* in the gonad cultures.

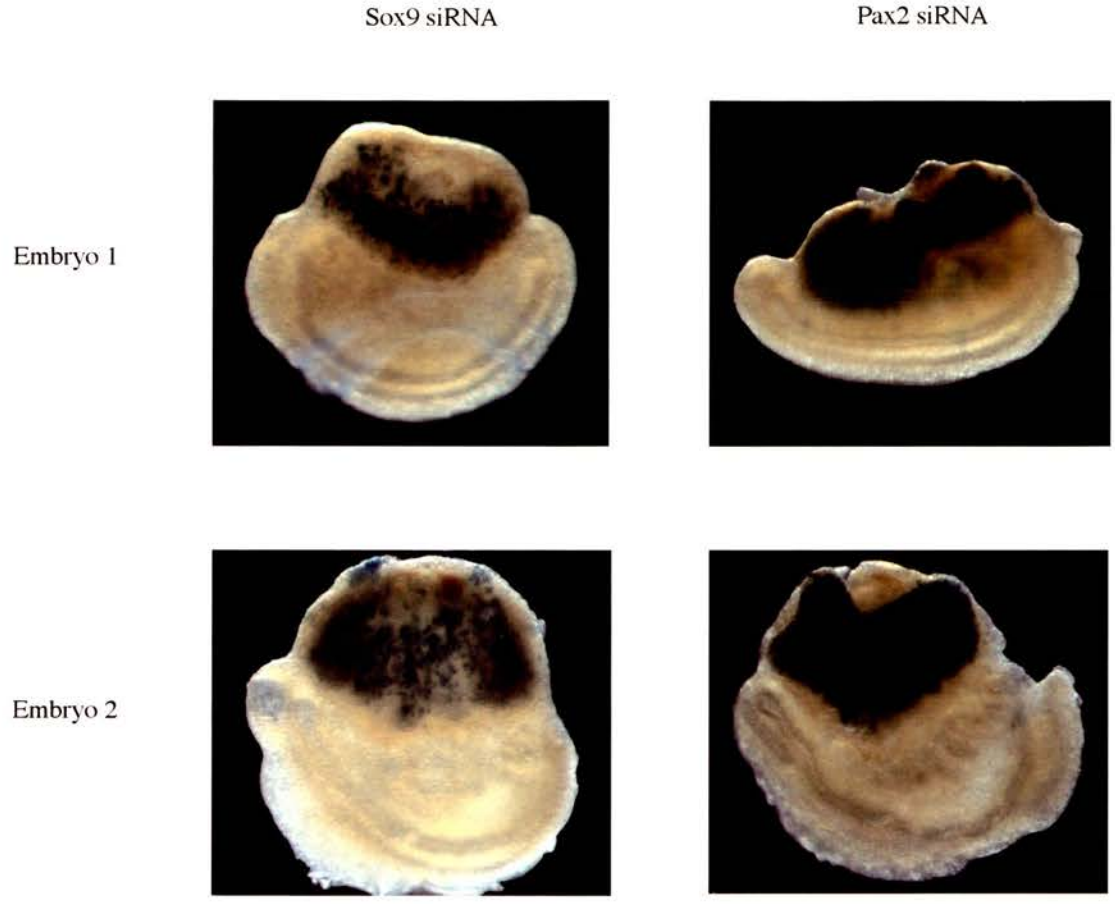


Figure 6.6 X-Gal Staining of E12.5 Gonad Cultures Treated with Sox9 siRNA

Heterozygous *Sox8* embryos were dissected at E12.5 and gonads were cultured to E13.5 (24 hours). The top two panels represent a pair of gonads from embryo 1 and the bottom two panels represent a pair of gonads from embryo 2. Both embryos are male. Each pair of gonads was separated, one treated with the Sox9 siRNA cocktail mix and the other with the Pax2 control siRNA. All gonad cultures were X-gal stained under the same conditions.

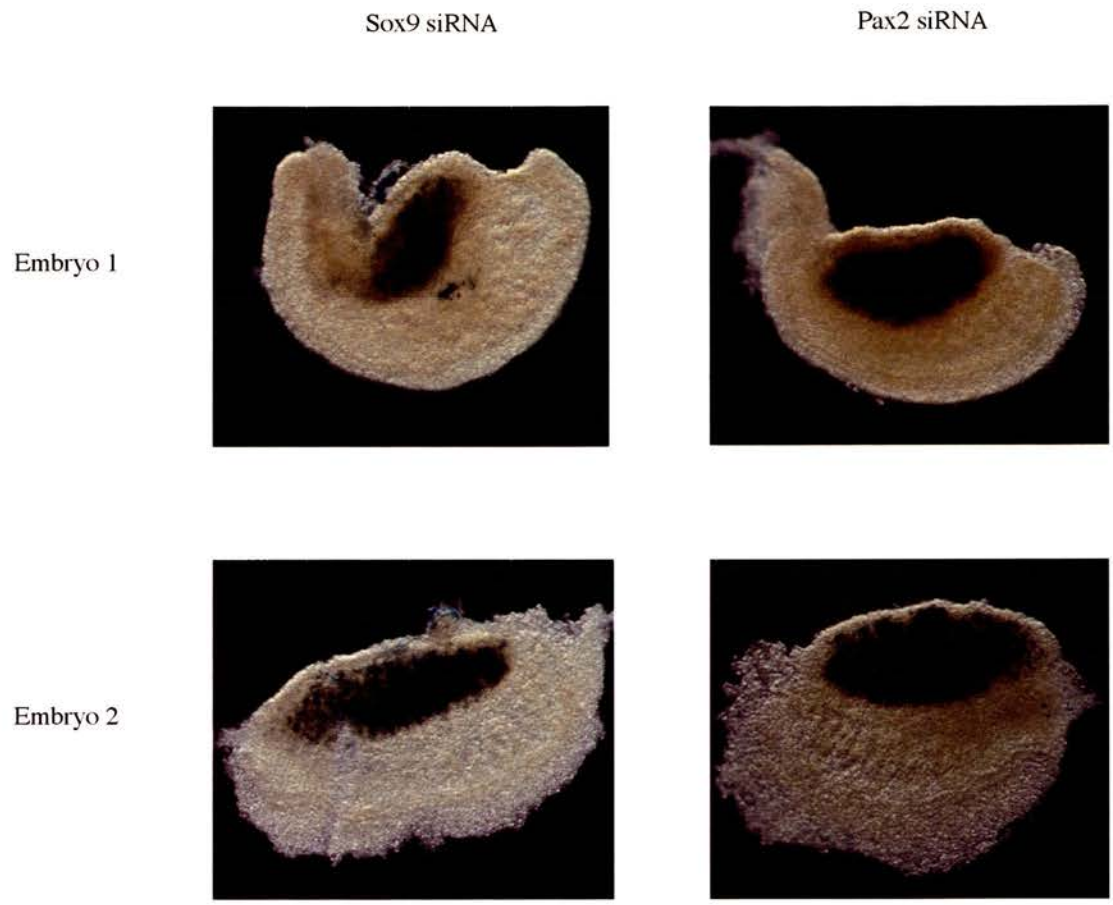


Figure 6.7 X-Gal Staining of E11.5 Gonad Cultures Treated with Sox9 siRNA

Heterozygous *Sox8* embryos were dissected at E11.5 and gonads were cultured to E12.5 (24 hours). The top two panels represent a pair of gonads from embryo 1 and the bottom two panels represent a pair of gonads from embryo 2. Both embryos are male. Each pair of gonads was separated, one treated with the Sox9 siRNA cocktail mix and the other with the Pax2 control siRNA. All gonad cultures were X-gal stained under the same conditions.

Quantitative real-time PCR results of E11.5 gonads cultured for 24 hours to E12.5 revealed three out of the four male gonad cultures to show a decrease in *Sox9* expression levels when treated with the Sox9 siRNA cocktail (Figure 6.8). The levels of *Sox9* knockdown seemed to be more successful in this case compared to the E12.5 attempt. In the three samples that showed a reduction in *Sox9* expression, the knockdown effect was much more effective in E11.5 than E12.5. More specifically, sample C121 showed a 44% knockdown effect, sample C123 showed a 72% knockdown effect and sample C122 showed a 77% knockdown effect. Knockdown levels of E12.5 gonad cultures only ranged from 20% to approximately 50% (Figure 6.3). Therefore, in gonadal organ cultures, the siRNA system is more effective when the expression levels of the target gene has not reached its peak. In addition, the siRNAs should be applied at a time point shortly after the target gene has been switched on, before the expression levels are too high to be repressed or knocked down.

Given the quite promising knockdown levels of *Sox9* in these E11.5 gonad cultures, it would be interesting to investigate the downstream effects of this *Sox9* repression in the gonadal organ culture system. Therefore, the *Mis* expression levels of these E11.5 gonad cultures were also measured by real-time PCR using sequence-specific hybridisation probes for *Mis* (Appendix 2). The expression levels of *Mis* were expected to be reduced in the samples where *Sox9* had been knocked down as *Mis* has been suggested to be a direct downstream target of *Sox9* (Arango et al., 1999; De Santa Barbara et al., 1998). Indeed, the three samples (C121, C122 and C123) that showed a reduction in *Sox9* expression also showed a reduction in *Mis* expression (Figure 6.9). In particular, there was a correlation between the down regulation of *Sox9* and the down regulation of *Mis*. Sample C121 showed a 44% knockdown in *Sox9* expression and a 38% knockdown in *Mis* expression, sample C123 showed a 72% knockdown in *Sox9* expression and a 81% knockdown in *Mis* expression and sample C122 showed a 77% knockdown in *Sox9* expression and a 92% knockdown in *Mis* expression. It should also be noted that there was no reduction of *Mis* expression in the male sample C120 in which *Sox9* levels were also unaffected. This

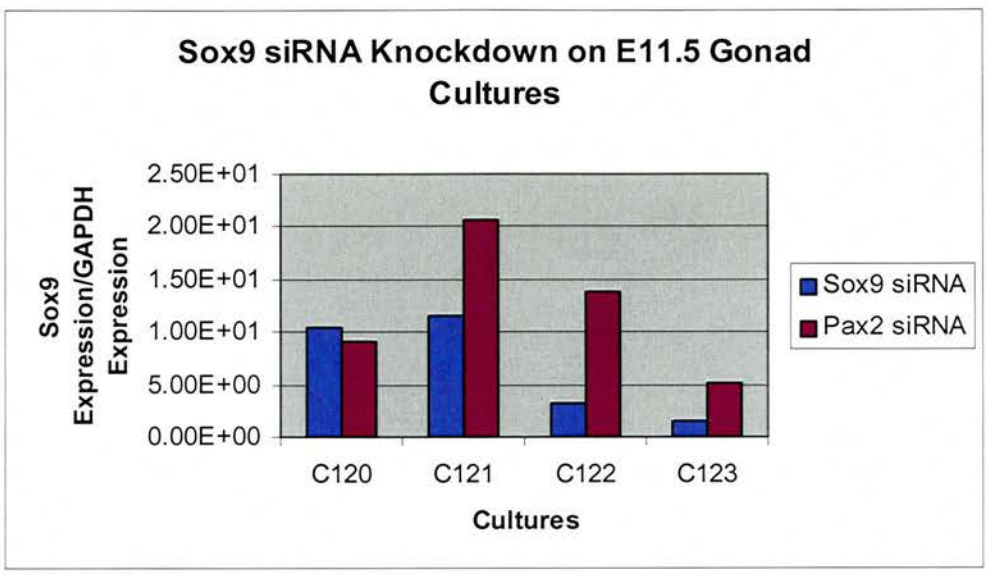


Figure 6.8 Real-time PCR Analysis of Sox9 Expression in E11.5 Gonad Cultures Following siRNA Treatment

Total RNA was collected from wild-type E11.5 gonads cultured for 24 hours following Sox9 and Pax2 siRNA treatment (corresponding gonads were used). 5μg of RNA was reverse transcribed to produce cDNA. Real-time PCR analysis for *Sox9* was performed. To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH. All samples are male.

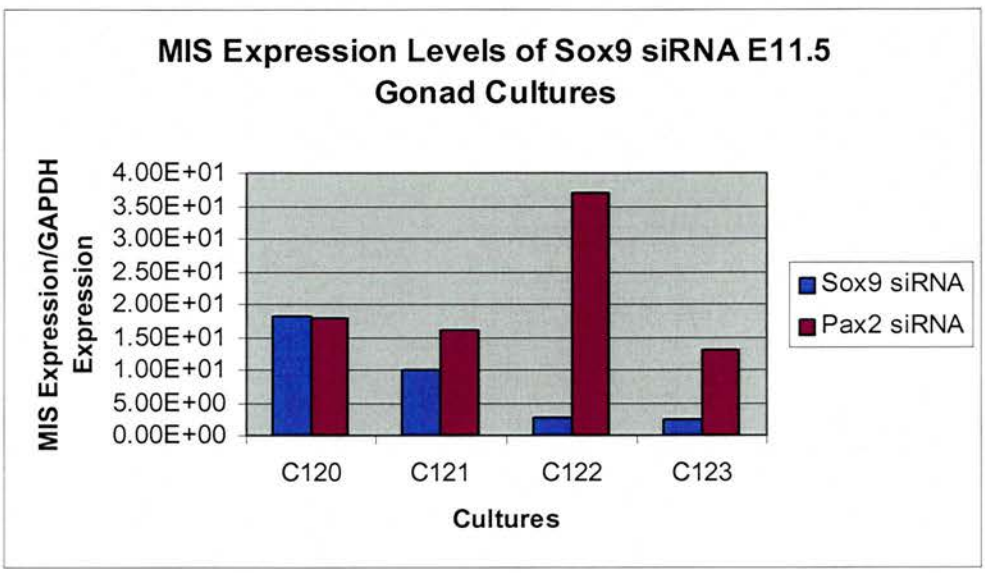


Figure 6.9 Real-time PCR Analysis of MIS Expression in E11.5 Gonad Cultures Following Sox9 siRNA Treatment

Total RNA was collected from wild-type E11.5 gonads cultured for 24 hours following Sox9 and Pax2 siRNA treatment (corresponding gonads were used). 5µg of RNA was reverse transcribed to produce cDNA. Real-time PCR analysis for *MIS* was performed (same cDNA samples as previously used in Figure 6.9). To control for the levels of cDNA, samples were normalised with respect to the expression levels of GAPDH. All samples are male.

illustrates that the reduction of *Mis* is a result of the reduced levels of *Sox9* and not caused directly by the presence of the transfectant (oligofectamine) or the siRNA oligonucleotides in the culturing medium. This further illustrates that the siRNA system works to a significant extent on gonadal organ cultures as the knockdown of the target gene reflects the *in vivo* results of downstream targets. More importantly, these data demonstrate that the approach of applying siRNAs in an organ culture system can be used to study molecular pathways.

6.4 Repression of Sry Expression by siRNA in Gonad Cultures

6.4.1 Experimental Design

Four different siRNA oligonucleotide templates were designed and produced for *Sry* (as described in section 2.7.1). They are referred to as Sry(6), Sry(7), Sry(9) and Sry(11) (individual target sequences can be found in Appendix 3). Since *Sry* is not expressed in TM4 cells (Figure 6.1), there was no system available to test for the efficiency of each Sry siRNA. Therefore, the technique of combining all four siRNAs in the one cocktail mix, as for Sox9, was also used for Sry. The Pax2 siRNA was again used as the control. $Sox8^{+/+}$ E10.5 urogenital ridges were used for gonad cultures as the expression of *Sry* is switched on at E10.5, where low levels of the *Sry* transcripts can be detected in XY gonads (Hacker et al., 1995; Jeske et al., 1995). These gonad cultures were X-gal stained to analyse whether the Sry siRNA cocktail mix would cause a phenotypic effect.

6.4.2 X-Gal Staining Results

Gonads were cultured for 48 hours (from E10.5 to approximately E12.5) before they were removed for X-gal staining. Figure 6.10 shows the results of two pairs of male gonad cultures from the same litter of embryos. The Pax2 control siRNA cultures

Figure 6.10 X-Gal Staining of E10.5 Gonad Cultures (1) Treated with Sry siRNA

Heterozygous *Sox8* embryos were dissected at E10.5 and gonads were cultured to E12.5 (48 hours). The top two panels represent a pair of gonads from embryo 1 and the bottom two panels represent a pair of gonads from embryo 2. Both embryos are male. Each pair of gonads was separated, one treated with the Sry siRNA cocktail mix and the other with the Pax2 control siRNA. All gonad cultures were X-gal stained under the same conditions.

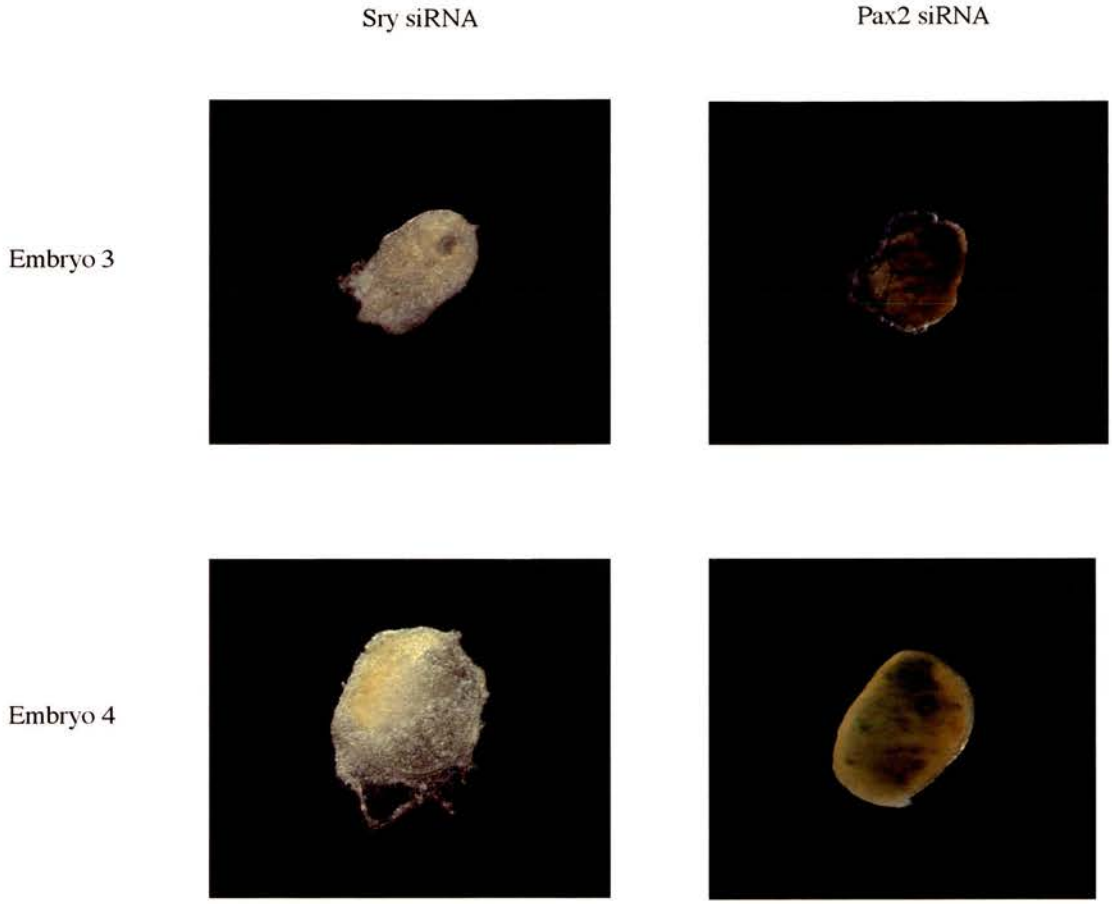


Figure 6.11 X-Gal Staining of E10.5 Gonad Cultures (2) Treated with Sry siRNA

Heterozygous *Sox8* embryos were dissected at E10.5 and gonads were cultured to E12.5 (48 hours). The top two panels represent a pair of gonads from embryo 3 and the bottom two panels represent a pair of gonads from embryo 4. Both embryos are male. Each pair of gonads was separated, one treated with the Sry siRNA cocktail mix and the other with the Pax2 control siRNA. All gonad cultures were X-gal stained under the same conditions.

clearly showed *lacZ* positive cells (blue staining observed), indicating that the siRNA treatment did not significantly delay the development of these cultures and did not interfere with the male sex determination process. In contrast, cultures treated with Sry siRNA were *lacZ* negative (no blue staining observed). These data show that the Sry siRNA cocktail mix had successfully repressed the expression of *Sry* in this gonadal organ culture system, whereby no staining was observed. The fact that the Pax2 siRNA treated corresponding cultures showed staining signifies that the Sry siRNA treated, *lacZ* negative cultures were not negative due to developmental delays in this organ culturing system where *lacZ* was not yet expressed. This experiment was repeated to confirm and to assess the reproducibility of these results. Again, only the Pax2 control siRNA cultures were stained and the Sry siRNA cocktail mix cultures were not stained (Figure 6.11). Hypothetically, all these Sry siRNA treated cultures could be considered to be sex-reversed as they are non-male and phenotypically resembled female gonad cultures rather than male. In conclusion, the Sry siRNA cocktail mix worked efficiently in this gonadal organ culture system and successfully altered the phenotype of the male gonad cultures.

6.5 Discussion

The ability of siRNAs to silence specific genes within a biological system was only discovered in the last few years (Fire et al., 1998). However, it has become one of the most promising research tools in biology. At present, when using siRNAs, it is common to select several sequences of the target gene to test for the most effective siRNA. As it is not efficient to conduct this in organ cultures where an excessive number of embryos would be required, the TM4 murine sertoli cell line (purchased from ECACC) was used. This cell line expresses the male sex determination gene *Sox9*, but did not express the sex-determining gene *Sry*. Therefore, only the Sox9 siRNAs could be assayed on these TM4 cells. In order to produce reliable results, replicates of each siRNA were conducted. Although the expression of *Sox9* was not completely reduced by any of the individual siRNAs, the 15 - 45% knockdown of

each siRNA was encouraging. Further experiments using other conditions, for example different transfection substitutes, may allow for further optimisation.

Since none of the Sox9 siRNAs were clearly more efficient, it was decided to combine all four siRNAs in one mix. Taking this approach may cause a greater knockdown effect or it may prevent a particular siRNA acting more efficient. However, following the decision to attempt this combined approach, the use of two different siRNAs targeted against two different genes was used by Voorhoeve and Agami (Voorhoeve and Agami, 2003). They showed that, combining the use of two siRNA-retroviral vectors, simultaneous suppression of both pRb and p53 was obtained in cells (Voorhoeve and Agami, 2003). Thus, it is feasible to produce a knockdown effect by combining different siRNAs in one reaction.

In addition, recently developed dicer-based systems can be used to produce many different siRNAs in the same reaction. This dicer system consists of a population of several siRNAs generated by digesting long dsRNA with ribonuclease III. Ribonuclease III cleaves dsRNA into 12-30 nucleotide dsRNA fragments with 2-3 nucleotide 3' overhangs and 5' phosphate and 3' hydroxyl termini (Byrom et al., 2003). These overhangs and termini are similar to those produced by "dicer" in the *in vivo* RNAi pathway. The success of this dicer system has been illustrated in mouse embryos whereby dicer-siRNAs were injected into the lumen of the neural tube at specific regions and delivered into neuroepithelial cells by directed electroporation (Calegari et al., 2002). These dicer-siRNAs were directed against β-gal and coelectroporated into neuroepithelial cells together with reporter plasmids expressing GFP and β-gal. Results only showed the presence of GFP, no β-gal expression was detected (Calegari et al., 2002). This demonstrates the specificity that can be achieved when using a number of siRNAs in the same reaction.

In gonadal organ cultures, the time point at which the siRNA is applied is important to the efficiency of the siRNA. As measured by quantitative real-time PCR, the Sox9 siRNA mix is almost twice more effective on E11.5 gonad cultures than E12.5

gonad cultures. This may be due to the presence of higher amounts of *Sox9* at E12.5. Alternatively, the amount of siRNA transfected may not have been sufficient to suppress *Sox9* efficiently. Moreover, E12.5 gonadal tissue contains more cell layers which may have caused the siRNA not to reach deeper layers of the tissue, leading to the siRNA not being delivered to all targeted cells. For example, the siRNA may only affect the top layer of cells and cannot reach the cells within the central layer of the gonad. Interestingly, although X-gal staining of Sox9 siRNA treated gonad cultures were *lacZ* positive, a slight effect of the Sox9 siRNA cocktail mix may be observed as there are fewer *lacZ* positive cells in the Sox9 siRNA treated samples when compared to the corresponding Pax2 control siRNA samples. However, it should be noted that the X-gal staining assays are a qualitative and not quantitative method of analysis. Taking the E11.5 and E12.5 real-time PCR results separately, the variation in *Sox9* expression levels within individual experiments may be due to the specific developmental stage of each embryo as tail somites were not counted for each pair of gonad cultures. Hence, the gonadal stage and the onset of genes may be slightly different which may cause a variation in siRNA knockdown levels. Of course, it may also be due to experimental differences as this is a sophisticated system with many parameters involved. Therefore, siRNA work on gonadal organ cultures must be repeated several times before conclusions can be drawn.

In contrast to *Sox9*, the siRNA knockdown experiments appeared to be more efficient for *Sry*. All cultures treated with Sry siRNA showed a sex-reversed phenotype. Silencing *Sry* may be easier than *Sox9* from the aspect that it is the initial switch for male development (Hacker et al., 1995), thus once activated, it determines the rest of the male developmental cascade. Therefore, if repression of *Sry* is successful, it should theoretically hinder further development of the male gonad cultures. In addition, changes in sex may be due to the testis being more sensitive to a decrease in *Sry* levels compared to a decrease in *Sox9* levels. This is based on the fact that male gonadal development seems to be particularly sensitive to the expression of *Sry* as it has been shown that a reduction of *Sry* expression to less than 30% of wild-type levels can cause male-to-female sex reversal in mice (Nagamine et al., 1999).

Consequently, lack of or too low levels of *Sry* could inhibit the activation of *Sox9* which would also prevent the upregulation of *Sox8* and thus the expression of *lacZ* in the $Sox8^{+/+}$ cultures. On the other hand, the expression of *Sry* is highly dynamic: low levels of *Sry* transcripts can be detected at E10.5 (when the formation of the bipotential gonad begins), it then reaches a peak at E11.5 (prior to overt differentiation), before being completely switched off by E13.0 (Hacker et al., 1995; Jeske et al., 1995). Therefore, it is critical to choose an early time point in order to allow the siRNA to cause an effect. In addition, as male specific expression commences at E11.5 and the expression of *Sox9* is sufficient to cause male development (Vidal et al., 2001), E10.5 gonad cultures had to be used for the Sry siRNA assays.

X-gal staining results of the Sry siRNA treated E10.5 gonad cultures have shown a repression of *Sry* as all these cultures were *lacZ* negative compared to the *lacZ* positive staining of their Pax2 control siRNA treated corresponding cultures. The successful knockdown of *Sry* expression levels in the Sry siRNA treated cultures prevented the downsteam activation of *Sox9*, which lead to the absence of *Sox8* expression and consequently no *lacZ* staining. These non-male gonad cultures demonstrate that the Sry siRNA cocktail mix worked effectively in the gonadal organ culture system and has even altered the phenotype of these male gonad cultures. Moreover, this shows that using more than one siRNA in one transfection mix can be efficient and successful.

However, further investigations into these non-male gonad cultures are required to determine whether they are truly sex-reversed. More specifically, it will be important to establish whether these non-male gonad cultures express female specific genes. This could be examined by *in-situ* hybridisation using female specific markers, such as *Bmp2* (Ross et al., 2003) and *Fst* (Yao et al., 2004). In addition, a Leydig cell-specific marker, such as *P450scc*, should be used to explore the presence of Leydig cells within our non-male cultures. It is conceivable that the Sry siRNA cocktail has only effected Sertoli cells within the cultures as *Sry*, *Sox9* and *Sox8* are

all expressed in Sertoli cells. Investigations into Leydig cells will reveal the extent of the Sry siRNA effect and will clarify whether these cultures have a truly non-male phenotype.

In summary, the task of setting up a siRNA system to investigate the function of novel sex determination genes has been achieved. The successful Sry siRNA repression can be used to further investigate the male developmental pathway. For example, the downstream targets of *Sry* could be deduced and confirmed by real-time PCR analysis of their expression levels following Sry siRNA knockdown. Subsequently, specific siRNAs for these genes can be produced and assayed on gonadal organ cultures themselves (further discussed in chapter 7). Taking both approaches into consideration, this siRNA gonad culture set up can be adapted and used for many genes involved in sex determination. Using the *lacZ* system as an indictor in the *Sox8* mice has been helpful in setting up the technology to use siRNAs on gonadal organ cultures. However, it should be noted that this system has to be adjusted and adapted for other gonadal development analyses. For example, *Sox8* activation occurs at early E12.5, whereas other genes act at a later stage. Consequently, investigations into genes which are involved in sex determination of male gonads at later stages will require a different system as *Sox8* will already have been activated, thus *lacZ* cannot be used as an indicator in these cases. Additionally, although the use of this *lacZ* system as an indicator in the *Sox8* mice is male specific, analysis of female specific genes is also possible by using wild-type gonad cultures and measuring their expression levels by real-time PCR. In fact, since specific gene repression by siRNA is unlikely to be 100% complete, quantification of male and female specific gene expression may be preferential, as it does not rely on an absolute inactivation of the target gene. In conclusion, the siRNA gonad culture system can be used for investigations into male and female pathways, which may lead to a better understanding of the sex development cascades. Furthermore, RNA isolated from siRNA-treated gonad cultures could be used on gonad-specific microarrays, which would detect changes in gene expression. Thus, the downstream

targets of the repressed gene would be identified and this may allow additional elements of the sex developmental pathways to be elucidated.

Chapter 7 General Discussion and Future Directions

Chapter 7 General Discussion and Future Directions

7.1 Overview of Project

The work presented in this thesis aims to identify new genes involved in the development of the gonad, which may lead to a better understanding of the male and female developmental pathways. Affymetrix microarrays were used to screen for novel candidates involved in sex determination. Differential expression patterns were verified by whole-mount *in-situ* hybridisation and this thesis focused on the 3 male specific genes (*Cbln1*, *Tpd52l1* and *Smoc2*) that were identified. Differences in expression levels of these transcripts were confirmed by real-time PCR analysis. Temporal expression patterns of these potential candidates were also examined by real-time PCR, whereby the overall trend for each gene could be monitored during the initial stages of gonadal differentiation (early E11.5 to late E12.5). These temporal expression patterns allowed speculation about the factors which may regulate these novel candidates at the early stages of male sexual development. Extending this work, the establishment of gene-specific knockdown by siRNA in gonadal organ cultures will allow functional analysis of candidate genes isolated from the sex determination screen.

7.2 Future Prospects of Microarrays in Sex Determination

There were many potential sex differentially expressed genes within our microarray data set. Several more potential candidates may be verified to have sexually dimorphic expression patterns following optimisation of the whole-mount *in-situ* hybridisation protocol and this may lead to identification of further novel candidates.

Advances in microarray technology now permits small amounts of total RNA to be used. However, a well-optimised, consistent and reproducible RNA amplification protocol is required for individual RNA samples. Once this has been achieved for

embryonic gonadal tissue, it will allow more sex determination screens to be performed. In particular, genes that are important in the sex development pathways could be identified by applying RNA isolated from sex-reversed gonads to microarrays. For example, XY homozygous knockout mice for the *Sox9* gene are completely sex reversed (Chaboissier et al., 2004). Thus, applying RNA isolated from male, female and mutant sex-reversed gonads to microarrays will allow genes that are involved in this male to female sex-reversed phenotype to be distinguished. This screen may also allow a gene to be pinpointed within the sex determination cascade and may demonstrate additional genes regulated by *Sry*.

7.3 Future Analyses of Cbln1, Tpd52l1 and Smoc2

From our sex differential screen, the three genes that were analysed in more detail were *Cbln1*, *Tpd52l1* and *Smoc2*, as they were confirmed to have a sexually dimorphic expression pattern at E12.5 by whole-mount *in-situ* hybridisation (Figures 4.2, 4.3 and 4.4). Nevertheless, optimisation of the *in-situ* hybridisation protocol and designing different sets of probes may improve the *in-situ* hybridisation results for these genes. This may lead to a more defined staining pattern of the gonads. Moreover, once optimisation has been achieved, analysis of sections following whole-mount *in-situ* hybridisation can be conducted. These results may then reveal the specific cell type within which each gene is expressed.

The preliminary temporal expression patterns of *Cbln1*, *Tpd52l1* and *Smoc2* suggested that these genes may be regulated by an early male specifically expressed gene, such as *Sox9*. To further investigate this hypothesis, the expression pattern of these genes could be examined in *Sox9* knockout mice. If these genes are not expressed, this implies that they may be downstream targets of *Sox9*. Subsequently, promoter-binding studies could be conducted to confirm the regulation of these genes by *Sox9*. Consequently, the promoter regions of these genes would be checked for the SOX9-binding site (CCCTTTGAGA) (Arango et al., 1999). If this motif is

found within a gene, electrophoretic mobility shift assays (EMSA) and co-transfection assays could be performed to investigate if this motif is actively involved in the transcription of the gene. If results reveal that SOX9 can bind and activate transcription of the gene, this indicates that *Sox9* regulates the gene. In case absolute proof is required, then this binding site could be mutated in ES cells, as conducted for *Mis* (Arango et al., 1999). Hence, mice with a targeted mutation of the SOX9-binding site within the gene's promoter would be generated. Analysis of the testes and ovaries of these mice will reveal whether mutation of the SOX9-binding site affects the expression levels of the gene of interest. Thus, the regulation of this gene by *Sox9* will be elucidated.

7.4 Current Developments of Cbln1, Tpd52l1 and Smoc2

7.4.1 Biological Profile of *Cbln1*

Precerebellin-1 or cerebellin 1 precursor protein (*Cbln1*) belongs to a larger superfamily of several atypical collagens (Urade et al., 1991). It shares significant amino acid homology with the globular (non-collagen) domains of the complement components C1qA, B and C (Urade et al., 1991; Wada and Ohtani, 1991), implying some common characteristics of structure and/or function. As the C1q complex consists of heterotrimers of C1qA, B and C that bind to surface immunoglobulin, it has been hypothesised that multiple precerebellins may assemble into homo- or heteromeric complexes that bind to brain-specific members of the immunoglobulin superfamily (Urade et al., 1991).

Molecular genetic analysis of *Cbln1* has determined that additional *Cbln1* transcripts are derived either from alternative splicing of the 3'-untranslated region (3'-UTR) of *Cbln1* or the use of different polyadenylation sites (Kavety and Morgan, 1998). To date, four independent transcripts have been cloned from an adult mouse cerebellum cDNA library (Kavety and Morgan, 1998). With one possible exception, these

various processing events do not affect the amino acid sequence of precerebellin (Kavety and Morgan, 1998).

The expression pattern of *Cbln1* within embryonic organs is presently unknown. This thesis describes *Cbln1* expression in E12.5 forelimbs, hindlimbs, head and liver. However, no expression was detected in the heart, lungs or kidneys (Figure 5.1). By sex differential screens, *Cbln1* has been identified to be sexually dimorphically expressed in E12.5 gonads (this thesis, Figure 4.2 and Menke and Page, 2002), which implies that *Cbln1* may play a role in sex determination and gonadal development.

Although homozygous *Cbln1* knockout mice do not show a male-to-female sex-reversed phenotype (personal correspondence with Dr James Morgan, St. Jude Children's Research Hospital, Memphis), the involvement of *Cbln1* within the sex development cascade should not be ruled out. For example, functional redundancy with another gene may have occurred and thus no gonadal phenotype is observed. Therefore, it is possible that *Cbln1* may still be important in the activation of male specific genes within the testicular development pathway. Further analysis of the homozygous *Cbln1* knockout gonads are required to establish a better understanding of *Cbln1* in sex determination and gonadal development.

Interestingly, a novel member of the cerebellin family, Cerebellin 4 precursor protein (*Cbln4*), has also been identified to be expressed in a sexually dimorphic pattern at E12.5 ((Menke and Page, 2002). Since *Cbln1* and *Cbln4* are both highly expressed in E12.5 testis and given that members of this family can form ternary complexes, it is possible that these two genes associate with each other to activate other genes within the male development pathway. However, *Cbln1* seems to be localised within the interstitial cells (this thesis, Figure 5.5 and (Menke and Page, 2002), whilst *Cbln4* seems to be expressed in the testicular cords (Menke and Page, 2002). This suggests the possibility of mutual induction signals mediated by these two molecules in the specification of different cell types/compartments of the testes. Alternatively, the distinct expression patterns of these two molecules correlate with distinct functions

in these different cell types/compartments. Further investigations of these two molecules are required to deduce their mutual or separate involvement in testicular development.

7.4.2 Biological Profile of *Tpd52l1*

The tumour protein D52 family represents a conserved group of small hydrophilic proteins, consisting of D52 (Byrne et al., 1995), D53 (TPD52L1) (Byrne et al., 1996) and D54 (TPD52L2) (Nourse et al., 1998). These three D52-like proteins are coiled-coil motif bearing proteins, indicative of protein-protein interactions, and may be capable of hetero- and/or homodimer formation (Byrne et al., 1996). D52-like gene transcripts are subject to alternative splicing, with sequences encoding a region termed insert 3 being affected in all three D52-like genes (Nourse et al., 1998). A 14-3-3 binding motif within one of two alternatively spliced exons encoding insert 3 has recently been identified, leading to the discovery of 14-3-3 proteins as novel heterologous partners for human D53 (Boutros et al., 2003).

In the mouse, *Tpd52l1* has been mapped to chromosome 10A4-10B2 (Byrne et al., 1998). However, little is known of its expression pattern. This thesis has shown *Tpd52l1* to be strongly expressed in E12.5 forelimbs, hindlimbs and head. Plus, moderately expressed in the heart, lungs and liver. However, no expression was detected in the kidneys (Figure 5.1). In the gonads, *Tpd52l1* was found to be sex differentially expressed by E12.5 – E14.5 (Figures 4.3 and 5.3). This suggests that *Tpd52l1* may be involved in the sex determination and gonadal development pathways. Furthermore, RNA *in-situ* hybridisation of E14.5 gonadal sections indicated *Tpd52l1* to be expressed within Sertoli cells (Figure 5.6). The known male specific genes *Sry*, *Sox9* and *Mis* are Sertoli cell markers, thus *Tpd52l1* may act within this developmental cascade. Real-time PCR analysis revealed *Tpd52l1* expression to increase in male gonads between 21 tail somites (E11.5) and 25 tail somites (early E12.5) (Figure 5.18). More specifically, this increase occurs shortly

after the onset of *Sox9* (Figure 5.15). It would be interesting to further analyse whether *Tpd52l1* is a direct downstream target of *Sox9*.

7.4.3 Biological Profile of *Smoc2*

The secreted modular calcium-binding protein 2 (*Smoc2*) contains an EF-hand calcium-binding domain homologous to that in BM-40 (Vannahme et al., 2003). In addition, it contains two thyroglobulin-like domains, a follistatin-like domain and a novel domain found only in the homologous SMOC1 (Vannahme et al., 2002). Phylogenetic analysis of the calcium-binding domain sequences showed that SMOC1 and SMOC2 form a separate group within the BM-40 family (Vannahme et al., 2003). Circular dichroism (CD) spectroscopy revealed SMOC2 to be a glycoprotein with a calcium-dependent conformation (Vannahme et al., 2003).

Northern blot analysis results from different tissues of adult mice showed a broad expression pattern of *Smoc2*, whereby a prominent signal in the ovary and clear signals in the heart, muscle and spleen were detected (Vannahme et al., 2003). RT-PCR revealed wide-spread expression in many adult tissues, such as the brain, thymus, lung, heart, liver, kidney, spleen, testis, ovary and skeletal muscle (Vannahme et al., 2003). Embryonic tissue distribution of *Smoc2* also exhibited expression in E12.5 heart, lungs, kidneys, liver, forelimbs, hindlimbs and the head (Figure 5.1). In embryonic gonads, *Smoc2* has been found to be expressed in a sexually dimorphic pattern by E12.5 – E14.5 (Figures 4.4 and 5.4). RNA *in-situ* hybridisation of E14.5 gonadal sections indicated *Smoc2* expression throughout the male gonad (Figure 5.7), whereas no specific expression within a particular cell type could be identified. As *Smoc2* has now been revealed to localise in basement membranes and is also present in the extracellular environment when calcium-bound (Vannahme et al., 2003), it is not surprising that it is widely expressed throughout the testis. These findings imply that the up-regulation of *Smoc2* in male gonads may only represent the formation of testicular cords. Hence, *Smoc2* appears to be

involved in the structure of the gonad rather than in the sex differentiation process. Thus, *Smoc2* is unlikely to play a role in sex determination.

7.5 Future Prospects of siRNA on Gonadal Organ Cultures

7.5.1 Quantitative Analysis of *Sry* siRNA-treated Gonad Cultures

The work in this thesis describes the application of siRNA in gonadal organ cultures. More specifically, the successful knockdown of *Sry* on gonadal organ cultures by siRNA treatment resulted in $Sox8^{+/+}$ male explants to be *lacZ* negative, representing the lack of expression of this male specific marker (Figures 6.10 and 6.11). It would be interesting to perform real-time PCR analysis of these Sry siRNA-treated gonad cultures on *Sry* itself but also on other genes. Both male and female genes should be analysed to measure the extent of sex reversal produced by the application of siRNA. Real-time PCR analysis of *Sry* will quantitatively establish the repressive effects of siRNA against *Sry*. Subsequently, other genes that may be a downstream target of *Sry* or act within the *Sry* cascade can also be examined by measurement of their expression levels.

It would also be interesting to measure the expression levels of genes which act within the female developmental pathway, such as *Fst* (Yao et al., 2004). Increased levels of Fst within Sry siRNA-treated gonad cultures would indicate a male-to-female sex reversed phenotype. Thus, the true extent of siRNA application on gonadal organ cultures will also be revealed.

7.5.2 Use of Other siRNAs on Gonad Cultures

The experimental siRNA system described in this thesis is a promising method of analysing genes involved in sex determination. Obviously, siRNA against *Sry* was

conducted as a proof of principle, where the successful knockdown effects have been shown (Figures 6.10 and 6.11). The technology can now be used to study putative sex-determining genes identified in our screen.

Investigations into potential female specific candidates are also possible by designing siRNAs against these genes and applying them onto wild-type gonad cultures. Their expression levels would then be quantified by real-time PCR analysis, whereby any reductions would be measured. Of course, this system would have to be tested with known female specific genes, such as *Bmp2* (Ross et al., 2003) and *Fst* (Yao et al., 2004), before assaying potential female candidates identified from the Affymetrix sex differential microarray analysis.

7.5.3 Generation of *LacZ* Mice for Analysis of Later-staged Male Specific Genes

This thesis described the use of a *lacZ* marker fused with the *Sox8* start codon (Sock et al., 2001) as an indicator of male specific expression. This system has aided the set up of siRNA technology on gonadal organ cultures and allowed for the analysis of early genes in testicular development. To investigate genes that are involved in later stages of male gonad development, there will be a requirement for the generation of transgenic mice carrying the *lacZ* marker fused with the start codon of a male specific gene expressed at a later developmental stage.

Recently, promyelocytic leukemia zinc-finger (*Plzf*) has been identified as a spermatogonia-specific transcription factor in the testis (Costoya et al., 2004). Mice lacking *Zfp145*, encoding the transcriptional repressor Plzf, progressively lose spermatogonia with age due to increases in apoptosis and subsequent loss of tubule structure (Costoya et al., 2004). However, overt differentiation defects did not occur and supporting Sertoli cells were not lost (Costoya et al., 2004). In wild-type mice, expression of *Plzf* is first detected in the germ cells at E17.5 during development of the testis and was found to peak in the postnatal testis at around 1 week (Costoya et

al., 2004). Thus, a *lacZ* marker could be fused with the *Plzf* start codon to produce transgenic mice that would serve as an indicator of phenotypically male gonad cultures for later staged testicular specific genes. Although the onset of expression at E17.5 may be too late and may cause problems, this system could be particularly useful in identifying genes that are involved in germ cell maturation of the testis.

Unfortunately, there are many genes that are not suitable as they are expressed in both male and female gonads before becoming sexually dimorphic at a later stage. For example, *Dmrt1* is expressed in both sexes from E10.5 – E13.5 and only becomes sex differentially expressed at E14.5 (De Grandi et al., 2000). Ideally, the gene should be male specifically expressed throughout the development of the gonad to provide a clear indication of being phenotypically male. Moreover, the onset of the gene's expression would preferably be from E14.5 to allow later staged male specific genes to be analysed, whilst minimising technical difficulties of culturing gonads for prolonged periods before the *lacZ* marker is switched on. Currently, there are no genes that match these criteria. However, as more male specific genes become identified, a suitable gene may become available to allow the generation of transgenic mice for this later staged system.

7.5.4 Alternative RNAi Approach

Although the application of siRNAs on gonadal organ cultures has been a successful *in vitro* approach to analyse the function of sex-determining genes, advances in RNAi technology can now be used in whole animal systems. A new vector, named pDECAP (Deletion of Cap structure and poly(A)), which expresses long double-strand RNA (ds-RNA) from an RNA polymerase II (Pol II) promoter has been developed (Shinagawa and Ishii, 2003). Transcripts from pDECAP lack both the 5'-cap structure and the 3'-poly(A) tail that facilitate ds-RNA export to the cytoplasm (Shinagawa and Ishii, 2003), hence long ds-RNA from pDECAP does not induce the interferon response. In general, Pol II transcripts are transferred to the cytosol

immediately after transcription, where they induce interferon synthesis (Stark et al., 1998), which leads to a block in translation and sequence-nonspecific mRNA degradation (Elbashir et al., 2001; Paddison et al., 2002). However, the design of pDECAP efficiently blocks the export of long ds-RNA to the cytosol (Shinagawa and Ishii, 2003). Since ribonuclease(s) such as Dicer, which cleave long ds-RNA, are localised in both the nucleus and the cytoplasm (Lee et al., 2002), the long ds-RNA expressed from the pDECAP vector is processed into siRNA in the nucleus and then moves to the cytosol to induce the degradation of target mRNA (Shinagawa and Ishii, 2003). Transgenic mice embryos expressing long ds-RNA for the transcriptional corepressor Ski from the pDECAP vector displayed similar phenotypes to those of *Ski*-deficient embryos (Shinagawa and Ishii, 2003).

Adopting this novel approach, the pDECAP vector could also be used to generate transgenic mice for the analysis of genes involved in sex determination and gonadal development. This innovative technique would initially be assayed on known sex determination genes, such as *Sox9*, to assess if similar phenotypes can be obtained. This process would involve investigating the effect of long ds-RNA for the *Sox9* gene on Sox9 protein levels. The pDECAP vector expressing ds-RNA for a fragment of Sox9 (pDECAP-Sox9) could be cotransfected into TM4 cells (murine Sertoli cell line expressing Sox9) together with a Sox9 expression vector. Subsequently, Sox9 protein levels could be examined by Western blot analysis. If levels of Sox9 protein are reduced in cotransfection of pDECAP-Sox9 and the pDECAP-Sox9 vector alone, this indicates that the pDECAP-Sox9 vector can cause a knockdown of Sox9 expression levels. Further investigations would be conducted to ensure that *Sox9* ds-RNA expressed from pDECAP-Sox9 does not induce the interferon response and that endogenous levels of *Sox9* mRNA are also reduced by transfection of pDECAP-Sox9. If results obtained reveal the desired outcome, the pDECAP-Sox9 vector can be prepared and injected into fertilised mouse oocytes for the generation of transgenic mice. Thus, RNAi may be used to study the function of sex determination genes by gonad specific knockdown in whole animal systems. This *in vivo* RNAi method may complement and allow a more in depth functional

analysis of sex determination genes when compared to the application of siRNA on gonadal organ cultures. However, it does not replace the generation of mutant mice by homologous recombination in embryonic stem (ES) cells (Capecchi, 1989) as a complete knockdown of the gene's expression level may not be achieved by pDECAP.

7.6 Conclusion

Following the discovery of *SRY*, over a decade of study on genes involved in sex determination and gonadal development has been conducted. Whilst several other genes, such as *WT1*, *SF1*, *DAX1* and *SOX9*, have been shown to play an important part in the sex developmental process, our knowledge and understanding remains incomplete. The utilisation of microarrays to screen for novel sex determination candidates has been an advancement towards the elucidation of this process, yet many other genes remain unidentified. Moreover, the functional roles of these potential candidates are unknown. Although further work is required, the identification of three novel sexually dimorphic expressed genes in this thesis has been an important step forward towards unravelling the sex development cascade. In addition, the establishment of the siRNA system on gonadal organ cultures described in this thesis will allow the functional properties of many potential candidates to be analysed, prior to the more time-consuming approaches of generating transgenic mice. Thus, taking the microarray analysis and siRNA system together, these approaches may provide an increased understanding of the male and female developmental pathways.

References

References

Achermann, J. C., Ito, M., Hindmarsh, P. C., and Jameson, J. L. (1999). A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal and adrenal failure in humans. Nat Genet *22*, 125-126.

Achermann, J. C., Meeks, J. J., and Jameson, J. L. (2001). Phenotypic spectrum of mutations in DAX-1 and SF-1. Mol Cell Endocrinol *185*, 17-25.

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

Albrecht, K. H., Capel, B., Washburn, L. L., and Eicher, E. M. (2000). Defective mesonephric cell migration is associated with abnormal testis cord development in C57BL/6J XY(Mus domesticus) mice. Dev Biol *225*, 26-36.

Albrecht, K. H., and Eicher, E. M. (2001). Evidence that Sry is expressed in pre-Sertoli cells and Sertoli and granulosa cells have a common precursor. Dev Biol *240*, 92-107.

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.

Baarends, W. M., van Helmond, M. J., Post, M., van der Schoot, P. J., Hoogerbrugge, J. W., de Winter, J. P., Uilenbroek, J. T., Karels, B., Wilming, L. G., Meijers, J. H., and et al. (1994). A novel member of the transmembrane serine/threonine kinase receptor family is specifically expressed in the gonads and in mesenchymal cells adjacent to the mullerian duct. Development *120*, 189-197.

Bardoni, B., Zanaria, E., Guioli, S., Floridia, G., Worley, K. C., Tonini, G., Ferrante, E., Chiumello, G., McCabe, E. R., Fraccaro, M., and et al. (1994). A dosage sensitive locus at chromosome Xp21 is involved in male to female sex reversal. Nat Genet *7*, 497-501.

Behringer, R. R., Cate, R. L., Froelick, G. J., Palmiter, R. D., and Brinster, R. L. (1990). Abnormal sexual development in transgenic mice chronically expressing mullerian inhibiting substance. Nature *345*, 167-170.

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

Bennington, J. a. B., J. (1975). Tumors of the kidney, renal pelvis, and ureter, Washington, D.C).

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

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.

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.

Bishop, C. E., Whitworth, D. J., Qin, Y., Agoulnik, A. I., Agoulnik, I. U., Harrison, W. R., Behringer, R. R., and Overbeek, P. A. (2000). A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. Nat Genet *26*, 490-494.

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.

Bland, M. L., Jamieson, C. A., Akana, S. F., Bornstein, S. R., Eisenhofer, G., Dallman, M. F., and Ingraham, H. A. (2000). Haploinsufficiency of steroidogenic factor-1 in mice disrupts adrenal development leading to an impaired stress response. Proc Natl Acad Sci U S A *97*, 14488-14493.

Borello, U., Buffa, V., Sonnino, C., Melchionna, R., Vivarelli, E., and Cossu, G. (1999). Differential expression of the Wnt putative receptors Frizzled during mouse somitogenesis. Mech Dev *89*, 173-177.

Boutros, R., Bailey, A. M., Wilson, S. H., and Byrne, J. A. (2003). Alternative splicing as a mechanism for regulating 14-3-3 binding: interactions between hD53 (TPD52L1) and 14-3-3 proteins. J Mol Biol *332*, 675-687.

Bowles, J., Schepers, G., and Koopman, P. (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. Dev Biol *227*, 239-255.

Brennan, J., Karl, J., and Capel, B. (2002). Divergent vascular mechanisms downstream of Sry establish the arterial system in the XY gonad. Dev Biol *244*, 418-428.

Buehr, M., Gu, S., and McLaren, A. (1993). Mesonephric contribution to testis differentiation in the fetal mouse. Development *117*, 273-281.

Bullejos, M., Bowles, J., and Koopman, P. (2002). Extensive vascularization of developing mouse ovaries revealed by caveolin-1 expression. Dev Dyn *225*, 95-99.

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

Byrne, J. A., Mattei, M. G., and Basset, P. (1996). Definition of the tumor protein D52 (TPD52) gene family through cloning of D52 homologues in human (hD53) and mouse (mD52). Genomics *35*, 523-532.

Byrne, J. A., Mattei, M. G., Basset, P., and Gunning, P. (1998). Identification and in situ hybridization mapping of a mouse Tpd52l1 (D53) orthologue to chromosome 10A4-B2. Cytogenet Cell Genet *81*, 199-201.

Byrne, J. A., Tomasetto, C., Garnier, J. M., Rouyer, N., Mattei, M. G., Bellocq, J. P., Rio, M. C., and Basset, P. (1995). A screening method to identify genes commonly overexpressed in carcinomas and the identification of a novel complementary DNA sequence. Cancer Res *55*, 2896-2903.

Byrom, M. W., Cheng, A. M., and Ford, L. P. (2003). Characterizing RNAi induced with siRNA cocktails generated by RNase III. Ambion TechNotes *10*, 4–6.

Calegari, F., Haubensak, W., Yang, D., Huttner, W. B., and Buchholz, F. (2002). Tissue-specific RNA interference in postimplantation mouse embryos with endoribonuclease-prepared short interfering RNA. Proc Natl Acad Sci U S A *99*, 14236-14240.

Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., and et al. (1990). Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. Cell *60*, 509-520.

Capecchi, M. R. (1989). Altering the genome by homologous recombination. Science *244*, 1288-1292.

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.

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. Biol Reprod *63*, 1825-1838.

Clarkson, M. J., and Harley, V. R. (2002). Sex with two SOX on: SRY and SOX9 in testis development. Trends Endocrinol Metab *13*, 106-111.

Colvin, J. S., Feldman, B., Nadeau, J. H., Goldfarb, M., and Ornitz, D. M. (1999). Genomic organization and embryonic expression of the mouse fibroblast growth factor 9 gene. Dev Dyn *216*, 72-88.

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

Colvin, J. S., White, A. C., Pratt, S. J., and Ornitz, D. M. (2001b). Lung hypoplasia and neonatal death in Fgf9-null mice identify this gene as an essential regulator of lung mesenchyme. Development *128*, 2095-2106.

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

Cupp, A. S., Uzumcu, M., and Skinner, M. K. (2003). Chemotactic role of neurotropin 3 in the embryonic testis that facilitates male sex determination. Biol Reprod *68*, 2033-2037.

Davies, J. A., Ladomery, M., Hohenstein, P., Michael, L., Shafe, A., Spraggon, L., and Hastie, N. (2004). Development of an siRNA-based method for repressing specific genes in renal organ culture and its use to show that the Wt1 tumour suppressor is required for nephron differentiation. Hum Mol Genet *13*, 235-246.

De Grandi, A., Calvari, V., Bertini, V., Bulfone, A., Peverali, G., Camerino, G., Borsani, G., and Guioli, S. (2000). The expression pattern of a mouse doublesex-related gene is consistent with a role in gonadal differentiation. Mech Dev *90*, 323-326.

De Santa Barbara, P., Bonneaud, N., Boizet, B., Desclozeaux, M., Moniot, B., Sudbeck, P., Scherer, G., Poulat, F., and Berta, P. (1998). Direct interaction of SRY-related protein SOX9 and steroidogenic factor 1 regulates transcription of the human anti-Mullerian hormone gene. Mol Cell Biol *18*, 6653-6665.

de Santa Barbara, P., Moniot, B., Poulat, F., and Berta, P. (2000). Expression and subcellular localization of SF-1, SOX9, WT1, and AMH proteins during early human testicular development. Dev Dyn *217*, 293-298.

Dehbi, M., Ghahremani, M., Lechner, M., Dressler, G., and Pelletier, J. (1996). The paired-box transcription factor, PAX2, positively modulates expression of the Wilms' tumor suppressor gene (WT1). Oncogene *13*, 447-453.

di Clemente, N., Wilson, C., Faure, E., Boussin, L., Carmillo, P., Tizard, R., Picard, J. Y., Vigier, B., Josso, N., and Cate, R. (1994). Cloning, expression, and alternative splicing of the receptor for anti-Mullerian hormone. Mol Endocrinol *8*, 1006-1020.

Durlinger, A. L., Kramer, P., Karels, B., de Jong, F. H., Uilenbroek, J. T., Grootegoed, J. A., and Themmen, A. P. (1999). Control of primordial follicle recruitment by anti-Mullerian hormone in the mouse ovary. Endocrinology *140*, 5789-5796.

Dyche, W. J. (1979). A comparative study of the differentiation and involution of the Mullerian duct and Wolffian duct in the male and female fetal mouse. J Morphol *162*, 175-209.

Dykxhoorn, D. M., Novina, C. D., and Sharp, P. A. (2003). Killing the messenger: short RNAs that silence gene expression. Nat Rev Mol Cell Biol *4*, 457-467.

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001a). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001b). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev *15*, 188-200.

Eriksson, A., Tohonen, V., Wedell, A., and Nordqvist, K. (2002). Isolation of the human testatin gene and analysis in patients with abnormal gonadal development. Mol Hum Reprod *8*, 8-15.

Failli, V., Rogard, M., Mattei, M. G., Vernier, P., and Retaux, S. (2000). Lhx9 and Lhx9alpha LIM-homeodomain factors: genomic structure, expression patterns, chromosomal localization, and phylogenetic analysis. Genomics *64*, 307-317.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Foster, J. W. (1996). Mutations in SOX9 cause both autosomal sex reversal and campomelic dysplasia. Acta Paediatr Jpn *38*, 405-411.

Francavilla, S., Cordeschi, G., Properzi, G., Concordia, N., Cappa, F., and Pozzi, V. (1990). Ultrastructure of fetal human gonad before sexual differentiation and during early testicular and ovarian development. J Submicrosc Cytol Pathol *22*, 389-400.

Gessler, M., Poustka, A., Cavenee, W., Neve, R. L., Orkin, S. H., and Bruns, G. A. (1990). Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosome jumping. Nature *343*, 774-778.

Gilbert, S. F. (2000). Developmental Biology, Sixth edn, Sinauer Associates, Inc.

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

Giuili, G., Shen, W. H., and Ingraham, H. A. (1997). The nuclear receptor SF-1 mediates sexually dimorphic expression of Mullerian Inhibiting Substance, in vivo. Development *124*, 1799-1807.

Gomperts, M., Wylie, C., and Heasman, J. (1994). Primordial germ cell migration. Ciba Found Symp *182*, 121-134; discussion 134-129.

Grimmond, S., Van Hateren, N., Siggers, P., Arkell, R., Larder, R., Soares, M. B., de Fatima Bonaldo, M., Smith, L., Tymowska-Lalanne, Z., Wells, C., and Greenfield, A. (2000). Sexually dimorphic expression of protease nexin-1 and vanin-1 in the developing mouse gonad prior to overt differentiation suggests a role in mammalian sexual development. Hum Mol Genet *9*, 1553-1560.

Gu, H., Marth, J. D., Orban, P. C., Mossmann, H., and Rajewsky, K. (1994). Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. Science *265*, 103-106.

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.

Gubbay, J., Vivian, N., Economou, A., Jackson, D., Goodfellow, P., and Lovell-Badge, R. (1992). Inverted repeat structure of the Sry locus in mice. Proc Natl Acad Sci U S A *89*, 7953-7957.

Guo, J. K., Hammes, A., Chaboissier, M. C., Vidal, V., Xing, Y., Wong, F., and Schedl, A. (2002). Early gonadal development: exploring Wt1 and Sox9 function. Novartis Found Symp *244*, 23-31; discussion 31-42, 253-257.

Guo, W., Burris, T. P., and McCabe, E. R. (1995). Expression of DAX-1, the gene responsible for X-linked adrenal hypoplasia congenita and hypogonadotropic hypogonadism, in the hypothalamic-pituitary-adrenal/gonadal axis. Biochem Mol Med *56*, 8-13.

Gura, T. (2000). A silence that speaks volumes. Nature *404*, 804-808.

Habert, R., Lejeune, H., and Saez, J. M. (2001). Origin, differentiation and regulation of fetal and adult Leydig cells. Mol Cell Endocrinol *179*, 47-74.

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.

Hawkins, J. R., Taylor, A., Berta, P., Levilliers, J., Van der Auwera, B., and Goodfellow, P. N. (1992). Mutational analysis of SRY: nonsense and missense mutations in XY sex reversal. Hum Genet *88*, 471-474.

Hirobe, S., He, W. W., Lee, M. M., and Donahoe, P. K. (1992). Mullerian inhibiting substance messenger ribonucleic acid expression in granulosa and Sertoli cells coincides with their mitotic activity. Endocrinology *131*, 854-862.

Hoyle, C., Narvaez, V., Alldus, G., Lovell-Badge, R., and Swain, A. (2002). Dax1 expression is dependent on steroidogenic factor 1 in the developing gonad. Mol Endocrinol *16*, 747-756.

Ikeda, Y., Shen, W. H., Ingraham, H. A., and Parker, K. L. (1994). Developmental expression of mouse steroidogenic factor-1, an essential regulator of the steroid hydroxylases. Mol Endocrinol *8*, 654-662.

Ikeda, Y., Swain, A., Weber, T. J., Hentges, K. E., Zanaria, E., Lalli, E., Tamai, K. T., Sassone-Corsi, P., Lovell-Badge, R., Camerino, G., and Parker, K. L. (1996). Steroidogenic factor 1 and Dax-1 colocalize in multiple cell lineages: potential links in endocrine development. Mol Endocrinol *10*, 1261-1272.

Jamin, S. P., Arango, N. A., Mishina, Y., Hanks, M. C., and Behringer, R. R. (2002). Requirement of Bmpr1a for Mullerian duct regression during male sexual development. Nat Genet *32*, 408-410.

Jeays-Ward, K., Hoyle, C., Brennan, J., Dandonneau, M., Alldus, G., Capel, B., and Swain, A. (2003). Endothelial and steroidogenic cell migration are regulated by WNT4 in the developing mammalian gonad. Development *130*, 3663-3670.

Jeske, Y. W., Bowles, J., Greenfield, A., and Koopman, P. (1995). Expression of a linear Sry transcript in the mouse genital ridge. Nat Genet *10*, 480-482.

Jeyasuria, P., Ikeda, Y., Jamin, S. P., Zhao, L., De Rooij, D. G., Themmen, A. P., Behringer, R. R., and Parker, K. L. (2004). Cell-specific Knockout of Steroidogenic Factor 1 Reveals Its Essential Roles in Gonadal Function. Mol Endocrinol.

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. Am J Hum Genet *68*, 1102-1109.

Jordan, B. K., Shen, J. H., Olaso, R., Ingraham, H. A., and Vilain, E. (2003). Wnt4 overexpression disrupts normal testicular vasculature and inhibits testosterone synthesis by repressing steroidogenic factor 1/beta-catenin synergy. Proc Natl Acad Sci U S A *100*, 10866-10871.

Jost, A. (1953). Problems of fetal endocrinology: the gonadal and hypophyseal hormones. Recent Prog Horm Res *8*, 379-418.

Jost, A., Vigier, B., Prepin, J., and Perchellet, J. P. (1973). Studies on sex differentiation in mammals. Recent Prog Horm Res *29*, 1-41.

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

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. H. a. B., J.B.L. (1999). The anatomical basis of mouse development, Academic Press).

Kavety, B., and Morgan, J. I. (1998). Characterization of transcript processing of the gene encoding precerebellin-1. Brain Res Mol Brain Res *63*, 98-104.

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.

Klamt, B., Koziell, A., Poulat, F., Wieacker, P., Scambler, P., Berta, P., and Gessler, M. (1998). Frasier syndrome is caused by defective alternative splicing of WT1 leading to an altered ratio of WT1 +/-KTS splice isoforms. Hum Mol Genet *7*, 709-714.

Koopman, P. (1999). Sry and Sox9: mammalian testis-determining genes. Cell Mol Life Sci *55*, 839-856.

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.

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.

Kuhl, M., Sheldahl, L. C., Malbon, C. C., and Moon, R. T. (2000). Ca(2+)/calmodulin-dependent protein kinase II is stimulated by Wnt and Frizzled homologs and promotes ventral cell fates in Xenopus. J Biol Chem *275*, 12701-12711.

Kuhn, R., Schwenk, F., Aguet, M., and Rajewsky, K. (1995). Inducible gene targeting in mice. Science *269*, 1427-1429.

Kuo, C. T., Morrisey, E. E., Anandappa, R., Sigrist, K., Lu, M. M., Parmacek, M. S., Soudais, C., and Leiden, J. M. (1997). GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. Genes Dev *11*, 1048-1060.

Larsson, S. H., Charlieu, J. P., Miyagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., van Heyningen, V., and Hastie, N. D. (1995). Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. Cell *81*, 391-401.

Lee, Y., Jeon, K., Lee, J. T., Kim, S., and Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. Embo J *21*, 4663-4670.

Li, X., Nokkala, E., Yan, W., Streng, T., Saarinen, N., Warri, A., Huhtaniemi, I., Santti, R., Makela, S., and Poutanen, M. (2001). Altered structure and function of reproductive organs in transgenic male mice overexpressing human aromatase. Endocrinology *142*, 2435-2442.

Little, M., Holmes, G., and Walsh, P. (1999). WT1: what has the last decade told us? Bioessays *21*, 191-202.

Lovell-Badge, R., Canning, C., and Sekido, R. (2002). Sex-determining genes in mice: building pathways. Novartis Found Symp *244*, 4-18; discussion 18-22, 35-42, 253-257.

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.

Lyet, L., Louis, F., Forest, M. G., Josso, N., Behringer, R. R., and Vigier, B. (1995). Ontogeny of reproductive abnormalities induced by deregulation of anti-mullerian hormone expression in transgenic mice. Biol Reprod *52*, 444-454.

Martineau, J., Nordqvist, K., Tilmann, C., Lovell-Badge, R., and Capel, B. (1997). Male-specific cell migration into the developing gonad. Curr Biol *7*, 958-968.

McClive, P. J., Hurley, T. M., Sarraj, M. A., van den Bergen, J. A., and Sinclair, A. H. (2003). Subtractive hybridisation screen identifies sexually dimorphic gene expression in the embryonic mouse gonad. Genesis *37*, 84-90.

McConnell, M. J., Cunliffe, H. E., Chua, L. J., Ward, T. A., and Eccles, M. R. (1997). Differential regulation of the human Wilms tumour suppressor gene (WT1) promoter by two isoforms of PAX2. Oncogene *14*, 2689-2700.

McLaren, A. (1988). Somatic and germ-cell sex in mammals. Philos Trans R Soc Lond B Biol Sci *322*, 3-9.

McLaren, A. (1991). Development of the mammalian gonad: the fate of the supporting cell lineage. Bioessays *13*, 151-156.

McLaren, A., and Southee, D. (1997). Entry of mouse embryonic germ cells into meiosis. Dev Biol *187*, 107-113.

Meeks, J. J., Crawford, S. E., Russell, T. A., Morohashi, K., Weiss, J., and Jameson, J. L. (2003a). Dax1 regulates testis cord organization during gonadal differentiation. Development *130*, 1029-1036.

Meeks, J. J., Russell, T. A., Jeffs, B., Huhtaniemi, I., Weiss, J., and Jameson, J. L. (2003b). Leydig cell-specific expression of DAX1 improves fertility of the Dax1-deficient mouse. Biol Reprod *69*, 154-160.

Meeks, J. J., Weiss, J., and Jameson, J. L. (2003c). Dax1 is required for testis determination. Nat Genet *34*, 32-33.

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

Miller, J. R. (2002). The Wnts. Genome Biol *3*, REVIEWS3001.

Mishina, Y., Whitworth, D. J., Racine, C., and Behringer, R. R. (1999). High specificity of Mullerian-inhibiting substance signaling in vivo. Endocrinology *140*, 2084-2088.

Mittwoch, U. (1989). Sex differentiation in mammals and tempo of growth: probabilities vs. switches. J Theor Biol *137*, 445-455.

Miyamoto, N., Yoshida, M., Kuratani, S., Matsuo, I., and Aizawa, S. (1997). Defects of urogenital development in mice lacking Emx2. Development *124*, 1653-1664.

Molkentin, J. D., Lin, Q., Duncan, S. A., and Olson, E. N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. Genes Dev *11*, 1061-1072.

Moore, A. W., McInnes, L., Kreidberg, J., Hastie, N. D., and Schedl, A. (1999). YAC complementation shows a requirement for Wt1 in the development of epicardium, adrenal gland and throughout nephrogenesis. Development *126*, 1845-1857.

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. Nat Genet *14*, 62-68.

Morrish, B. C., and Sinclair, A. H. (2002). Vertebrate sex determination: many means to an end. Reproduction *124*, 447-457.

Munsterberg, A., and Lovell-Badge, R. (1991). Expression of the mouse anti-mullerian hormone gene suggests a role in both male and female sexual differentiation. Development *113*, 613-624.

Nachtigal, M. W., Hirokawa, Y., Enyeart-VanHouten, D. L., Flanagan, J. N., Hammer, G. D., and Ingraham, H. A. (1998). Wilms' tumor 1 and Dax-1 modulate the orphan nuclear receptor SF-1 in sex-specific gene expression. Cell *93*, 445-454.

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.

Natoli, T. A., Alberta, J. A., Bortvin, A., Taglienti, M. E., Menke, D. B., Loring, J., Jaenisch, R., Page, D. C., Housman, D. E., and Kreidberg, J. A. (2004). Wt1 functions in the development of germ cells in addition to somatic cell lineages of the testis. Dev Biol *268*, 429-440.

Nourse, C. R., Mattei, M. G., Gunning, P., and Byrne, J. A. (1998). Cloning of a third member of the D52 gene family indicates alternative coding sequence usage in D52-like transcripts. Biochim Biophys Acta *1443*, 155-168.

Nusse, R., Brown, A., Papkoff, J., Scambler, P., Shackleford, G., McMahon, A., Moon, R., and Varmus, H. (1991). A new nomenclature for int-1 and related genes: the Wnt gene family. Cell *64*, 231.

Nykanen, A., Haley, B., and Zamore, P. D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell *107*, 309-321.

Ornitz, D. M., and Itoh, N. (2001). Fibroblast growth factors. Genome Biol *2*, REVIEWS3005.

Paddison, P. J., Caudy, A. A., and Hannon, G. J. (2002). Stable suppression of gene expression by RNAi in mammalian cells. Proc Natl Acad Sci U S A *99*, 1443-1448.

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.

Pandur, P., Maurus, D., and Kuhl, M. (2002). Increasingly complex: new players enter the Wnt signaling network. Bioessays *24*, 881-884.

Parker, K. L., and Schimmer, B. P. (1997). Steroidogenic factor 1: a key determinant of endocrine development and function. Endocr Rev *18*, 361-377.

Parker, K. L., Schimmer, B. P., and Schedl, A. (2001). Genes essential for early events in gonadal development. Exs, 11-24.

Pearce, J. J., Singh, P. B., and Gaunt, S. J. (1992). The mouse has a Polycomb-like chromobox gene. Development *114*, 921-929.

Picard, J. Y., Tran, D., and Josso, N. (1978). Biosynthesis of labelled anti-mullerian hormone by fetal testes: evidence for the glycoprotein nature of the hormone and for its disulfide-bonded structure. Mol Cell Endocrinol *12*, 17-30.

Racine, C., Rey, R., Forest, M. G., Louis, F., Ferre, A., Huhtaniemi, I., Josso, N., and di Clemente, N. (1998). Receptors for anti-mullerian hormone on Leydig cells are responsible for its effects on steroidogenesis and cell differentiation. Proc Natl Acad Sci U S A *95*, 594-599.

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.

Reijnen, M. J., Hamer, K. M., den Blaauwen, J. L., Lambrechts, C., Schoneveld, I., van Driel, R., and Otte, A. P. (1995). Polycomb and bmi-1 homologs are expressed in overlapping patterns in Xenopus embryos and are able to interact with each other. Mech Dev *53*, 35-46.

Rijsewijk, F., Schuermann, M., Wagenaar, E., Parren, P., Weigel, D., and Nusse, R. (1987). The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. Cell *50*, 649-657.

Ross, A. J., Tilman, C., Yao, H., MacLaughlin, D., and Capel, B. (2003). AMH induces mesonephric cell migration in XX gonads. Mol Cell Endocrinol *211*, 1-7.

Sadler, T. W. (1995). Langman's medical embryology, Seventh edition edn, Williams & Wilkins).

Schepers, G., Wilson, M., Wilhelm, D., and Koopman, P. (2003). SOX8 is expressed during testis differentiation in mice and synergizes with SF1 to activate the Amh promoter in vitro. J Biol Chem *278*, 28101-28108.

Schmahl, J., and Capel, B. (2003). Cell proliferation is necessary for the determination of male fate in the gonad. Dev Biol *258*, 264-276.

Schmahl, J., Eicher, E. M., Washburn, L. L., and Capel, B. (2000). Sry induces cell proliferation in the mouse gonad. Development *127*, 65-73.

Schmahl, J., Yao, H. H., Pierucci-Alves, F., and Capel, B. (2003). Colocalization of WT1 and cell proliferation reveals conserved mechanisms in temperature-dependent sex determination. Genesis *35*, 193-201.

Schwarz, D. S., Hutvagner, G., Haley, B., and Zamore, P. D. (2002). Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. Mol Cell *10*, 537-548.

Sharp, P. A. (2001). RNA interference--2001. Genes Dev *15*, 485-490.

Shimamura, R., Fraizer, G. C., Trapman, J., Lau Yf, C., and Saunders, G. F. (1997). The Wilms' tumor gene WT1 can regulate genes involved in sex determination and differentiation: SRY, Mullerian-inhibiting substance, and the androgen receptor. Clin Cancer Res *3*, 2571-2580.

Shinagawa, T., and Ishii, S. (2003). Generation of Ski-knockdown mice by expressing a long double-strand RNA from an RNA polymerase II promoter. Genes Dev *17*, 1340-1345.

Sim, E. U., Smith, A., Szilagi, E., Rae, F., Ioannou, P., Lindsay, M. H., and Little, M. H. (2002). Wnt-4 regulation by the Wilms' tumour suppressor gene, WT1. Oncogene *21*, 2948-2960.

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.

Slusarski, D. C., Corces, V. G., and Moon, R. T. (1997). Interaction of Wnt and a Frizzled homologue triggers G-protein-linked phosphatidylinositol signalling. Nature *390*, 410-413.

Sock, E., Schmidt, K., Hermanns-Borgmeyer, I., Bosl, M. R., and Wegner, M. (2001). Idiopathic weight reduction in mice deficient in the high-mobility-group transcription factor Sox8. Mol Cell Biol *21*, 6951-6959.

Stark, G. R., Kerr, I. M., Williams, B. R., Silverman, R. H., and Schreiber, R. D. (1998). How cells respond to interferons. Annu Rev Biochem *67*, 227-264.

Su, H., and Lau, Y. F. (1993). Identification of the transcriptional unit, structural organization, and promoter sequence of the human sex-determining region Y (SRY) gene, using a reverse genetic approach. Am J Hum Genet *52*, 24-38.

Svensson, E. C., Huggins, G. S., Lin, H., Clendenin, C., Jiang, F., Tufts, R., Dardik, F. B., and Leiden, J. M. (2000). A syndrome of tricuspid atresia in mice with a targeted mutation of the gene encoding Fog-2. Nat Genet *25*, 353-356.

Svensson, E. C., Tufts, R. L., Polk, C. E., and Leiden, J. M. (1999). Molecular cloning of FOG-2: a modulator of transcription factor GATA-4 in cardiomyocytes. Proc Natl Acad Sci U S A *96*, 956-961.

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

Swain, A., Narvaez, V., Burgoyne, P., Camerino, G., and Lovell-Badge, R. (1998). Dax1 antagonizes Sry action in mammalian sex determination. Nature *391*, 761-767.

Swain, A., Zanaria, E., Hacker, A., Lovell-Badge, R., and Camerino, G. (1996). Mouse Dax1 expression is consistent with a role in sex determination as well as in adrenal and hypothalamus function. Nat Genet *12*, 404-409.

Tamura, M., Kanno, Y., Chuma, S., Saito, T., and Nakatsuji, N. (2001). Pod-1/Capsulin shows a sex- and stage-dependent expression pattern in the mouse gonad development and represses expression of Ad4BP/SF-1. Mech Dev *102*, 135-144.

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.

Tilmann, C., and Capel, B. (1999). Mesonephric cell migration induces testis cord formation and Sertoli cell differentiation in the mammalian gonad. Development *126*, 2883-2890.

Tohonen, V., Osterlund, C., and Nordqvist, K. (1998). Testatin: a cystatin-related gene expressed during early testis development. Proc Natl Acad Sci U S A *95*, 14208-14213.

Torres, M., Gomez-Pardo, E., Dressler, G. R., and Gruss, P. (1995). Pax-2 controls multiple steps of urogenital development. Development *121*, 4057-4065.

Tran, D., Muesy-Dessole, N., and Josso, N. (1977). Anti-Mullerian hormone is a functional marker of foetal Sertoli cells. Nature *269*, 411-412.

Tung, P. S., Skinner, M. K., and Fritz, I. B. (1984). Cooperativity between Sertoli cells and peritubular myoid cells in the formation of the basal lamina in the seminiferous tubule. Ann N Y Acad Sci *438*, 435-446.

Urade, Y., Oberdick, J., Molinar-Rode, R., and Morgan, J. I. (1991). Precerebellin is a cerebellum-specific protein with similarity to the globular domain of complement C1q B chain. Proc Natl Acad Sci U S A *88*, 1069-1073.

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.

Vannahme, C., Gosling, S., Paulsson, M., Maurer, P., and Hartmann, U. (2003). Characterization of SMOC-2, a modular extracellular calcium-binding protein. Biochem J *373*, 805-814.

Vannahme, C., Smyth, N., Miosge, N., Gosling, S., Frie, C., Paulsson, M., Maurer, P., and Hartmann, U. (2002). Characterization of SMOC-1, a novel modular calcium-binding protein in basement membranes. J Biol Chem *277*, 37977-37986.

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

Viger, R. S., Mertineit, C., Trasler, J. M., and Nemer, M. (1998). Transcription factor GATA-4 is expressed in a sexually dimorphic pattern during mouse gonadal development and is a potent activator of the Mullerian inhibiting substance promoter. Development *125*, 2665-2675.

Voorhoeve, P. M., and Agami, R. (2003). The tumor-suppressive functions of the human INK4A locus. Cancer Cell *4*, 311-319.

Wada, C., and Ohtani, H. (1991). Molecular cloning of rat cerebellin-like protein cDNA which encodes a novel membrane-associated glycoprotein. Brain Res Mol Brain Res *9*, 71-77.

Wang, Z. J., Jeffs, B., Ito, M., Achermann, J. C., Yu, R. N., Hales, D. B., and Jameson, J. L. (2001). Aromatase (Cyp19) expression is up-regulated by targeted disruption of Dax1. Proc Natl Acad Sci U S A *98*, 7988-7993.

Watanabe, K., Clarke, T. R., Lane, A. H., Wang, X., and Donahoe, P. K. (2000). Endogenous expression of Mullerian inhibiting substance in early postnatal rat sertoli cells requires multiple steroidogenic factor-1 and GATA-4-binding sites. Proc Natl Acad Sci U S A *97*, 1624-1629.

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

Wilson, M., and Koopman, P. (2002). Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators. Curr Opin Genet Dev *12*, 441-446.

Yamanaka, H., Moriguchi, T., Masuyama, N., Kusakabe, M., Hanafusa, H., Takada, R., Takada, S., and Nishida, E. (2002). JNK functions in the non-canonical Wnt pathway to regulate convergent extension movements in vertebrates. EMBO Rep *3*, 69-75.

Yao, H. H., DiNapoli, L., and Capel, B. (2003). Meiotic germ cells antagonize mesonephric cell migration and testis cord formation in mouse gonads. Development *130*, 5895-5902.

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.

Yoshida, M., Suda, Y., Matsuo, I., Miyamoto, N., Takeda, N., Kuratani, S., and Aizawa, S. (1997). Emx1 and Emx2 functions in development of dorsal telencephalon. Development *124*, 101-111.

Yu, R. N., Ito, M., Saunders, T. L., Camper, S. A., and Jameson, J. L. (1998). Role of Ahch in gonadal development and gametogenesis. Nat Genet *20*, 353-357.

Zamore, P. D., Tuschl, T., Sharp, P. A., and Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.

Zanaria, E., Muscatelli, F., Bardoni, B., Strom, T. M., Guioli, S., Guo, W., Lalli, E., Moser, C., Walker, A. P., McCabe, E. R., and et al. (1994). An unusual member of the nuclear hormone receptor superfamily responsible for X-linked adrenal hypoplasia congenita. Nature *372*, 635-641.

Appendices

Appendix 1

Sequences of primers used in PCR and production of riboprobes

For riboprobes, T3 and T7 promoter sequences were placed in front of 3' and 5' primer sequences. For example, T3 promoter–3' primer and T7 promoter–5'primer.

T3 promoter sequence: cgcaattaaccctcactaaagggaacatctaga

T7 promoter sequence: gcgtaatacgactcactatagggcgactcgag

GenBank Description	Primer Sequence
AI326397_5'	ggagatggagcctgcatact
AA008996_3'	agctaggtcaggtgggtaat
AA008996_5'	tcagtgtgtgtatggtccat
AA152809_3'	gagctaagccggcatttgtca
AA152809_5'	aactgggtgttctgaccacga
AA214997_3'	agcttctttgggagctgtct
AA214997_5'	gatgtcagtgggaatccagt
AA285446_3'	ctgagtatgtgacgtctgct
AA285446_5'	gacaggatcactgagactct
AA606367_3'	agacgaagacacattcccgt
AA606367_5'	attgtctccttgcctgcctt
AA638515_3'	aaaacacaggtagctgggct
AA638515_5'	gggtgactttgtagcatggt
AA682034_3'	atcctgcagtgtagtcggct
AA682034_5'	acagcggtgaggaacatggt
AA832565_3'	ctgcgtgtccagcatccata
AA832565_5'	tgagctgccaactcagctga
AA839903_3'	tgtgaggaagacggtgaagtt
AA839903_5'	ggcacaacgtgctcttctaat
AA895031_3'	ctgcagtgcctacagtcata
AA895031_5'	gcaggcgaatgtacaagaga
AA920710_3'	tggtcaagtgtctctggtga
AA920710_5'	gtgtacggttctcagcatga
AA986447_3'	gcactgtgtgttctcgtgta
AA986447_5'	gttcatggctgagtgcttca

AF016482_3'	tcagaagatcgggaggaaca
AF016482_5'	aactacactgccagcttcga
AF079901_3'	ttcgaaggtggtcatgctct
AF079901_5'	tgacgttgggcagcaaagat
AI006228_3'	cctgtggtcaactgtcatct
AI006228_5'	acttgcctttggcgcattct
AI197161_3'	caaaagcttttcaccttgca
AI197161_5'	ttcggcacaggtgaatttga
AI323290_3'	actcactgtcttctccagga
AI323290_5'	atcctgattaccagcgtgga
AI326397_3'	atgaagaccagccgcttcct
AI327212_3'	gtggccagtccttcattgtt
AI327212_5'	gcaggtggactcacatcttt
AI430099_3'	aacttcgcccagtcttgcct
AI430099_5'	ggctgccatacaccatggat
AI466491_3'	ttcactcctcagcttctcca
AI466491_5'	gatgacgatgacgatgacga
AI467050_3'	ttggtactgtcaaacacctt
AI467050_5'	tgcagttcaatggctcagtt
AI481075_3'	tgcccacataggacatagca
AI481075_5'	aaaggagccagagaagggca
AI508971_3'	ttactgaccagagcagtgga
AI508971_5'	atcagtgttcatcacgggca
AI550379_3'	cctaccttagtcttttccat
AI550379_5'	tatgctagtcccaggaatat
AI551848_3'	tgcgttcacaggcagacatca
AI551848_5'	gttagagaagacgcagcacca
AI604578_3'	accatgccaaggagactagt
AI604578_5'	aaagtcgcccagatgtttgtt
AI604974_3'	cattagcctggacgtacact
AI604974_5'	tcttgaccatcagaggcagt
AI605905_3'	ttgcaaacgggtttggtcca
AI605905_5'	gtcctgagagatcttgtcca
AI606300_3'	gaggatagcgctccatccat
AI606300_5'	gagacctgcacagagagact
AI643851_3'	tcgttctagactggaactgt
AI643851_5'	gctttcagctcttactcagt
AI787183_3'	agcggatgctggaatcctta
AI787183_5'	gaggaatcagacgatgtcca
AI835159_3'	tccccctttcacaactgct
AI835159_5'	cttcgatttttgggtggggt
AI838194_3'	tccccagacttagggtggga
AI838194_5'	gttacagtctcgccacgaga
AI839436_3'	gagaagacagctaaacagct
AI839436_5'	cattagtggcttctttccct

AI839528_3'	tgcgccaccaagcctatact
AI839528_5'	gtggaactccctgacagact
AI839993_3'	tgtggccatcatccttctca
AI839993_5'	ccaggtcaagaagcacaaca
AI841906_3'	ctggagaagcatatctccct
AI841906_5'	gacaaccacacacacctact
AI842665_3'	catcttctcagtgcctctcca
AI842665_5'	aaagctagacccacaggtgca
AI844396_3'	aagcgcacaaacagcactgt
AI844396_5'	acaacagtggagcccttcat
AI844484_3'	accaaacacgtggtcccaca
AI844484_5'	tcttctccaaggtgtgacca
AI844597_3'	ttcatgagaggctgaggaca
AI844597_5'	ccacttattctcaggcccaa
AI844626_3'	caatcacatccttaacagggt
AI844626_5'	agtgtttgcatagtcgaggtt
AI845667_3'	aaggtgtcttgagccatgga
AI845667_5'	gtacactgtagctgtctcca
AI846736_3'	ggatttggctcttgctctga
AI846736_5'	ttgctgatgcctccaagaca
AI848448_3'	cattcttggctttcagccact
AI848448_5'	atcacactttgtgcttgggct
AI849193_3'	tatttcgccaactgcaccgt
AI849193_5'	gtcgtggttcagtgttcgat
AI849532_3'	aaagttgcaggcctcagaga
AI849532_5'	ggcagtagaggatggagtta
AI850212_3'	acatagtgtccctgttgcct
AI850212_5'	aacctctgcacataacttggt
AI851057_3'	aagggaggttaggaaaggta
AI851057_5'	aatgcagctggtagaatttca
AI852203_3'	ttctgcacacgagcagaactt
AI852203_5'	gcagcatgattctcatgcact
AI853714_3'	ctgcatgtcctctctctact
AI853714_5'	ctacattgtcatggtggcct
AI854623_3'	tgcctgagacctgagatccta
AI854623_5'	tagcacagagagtcaggtgga
AJ005560_3'	gacaaggctcacagcatact
AJ005560_5'	gtcttactaccagcagcagt
AJ007376_3'	cctcctctcgatctttctgt
AJ007376_5'	tctagagtccgtccttgtgt
AU044228_3'	aatagcccaagtgaccaaga
AU044228_5'	gtttgttctttcgcatggca
AW045202_3'	tcagtatggtgtccagggat
AW045202_5'	gccagaacctgattcatggt
AW045474_3'	cgtccaggatggcatgaact
AW045474_5'	cgataccgagcctctgagat

AW048072_3'	ggatggtgttagccagtgat
AW048072_5'	aacacaggtcacagcacctt
AW048347_3'	ctcttcaaccacaggctcaa
AW048347_5'	ctgtgtgccaggtagttcta
AW048937_3'	agcagttgtacaaggagcca
AW048937_5'	cttctcgtgagacgcttaca
AW060457_3'	gagttgtctgtggttccaagt
AW060457_5'	ggccttgatgtcaacagactt
AW060550_3'	ggtgggctgcatgaaattcaa
AW060550_5'	caatcactgggcgtaatgaca
AW061185_3'	agatctgacagtagctgcca
AW061185_5'	agagggatggagtcatacca
AW121680_3'	atgtgcacagcagacactca
AW121680_5'	cactaccaaccaagcagaca
AW121806_3'	atctgtggcttcagcacaca
AW121806_5'	aggacacagccacagtgata
AW123164_3'	ggctttgatacgcctggtaa
AW123164_5'	tggagataagaggcgaacca
AW123386_3'	tgtgactaccgtgtgtgtgt
AW123386_5'	gacagacagagacacacact
AW124221_3'	catactctggatgatggccat
AW124221_5'	tgtaacccgcaggtgaaatct
AW124288_3'	tttgtcctggcctcaggtaa
AW124288_5'	attaatcgcaagggcaggca
AW124318_3'	agcttccagcctcccaccat
AW124318_5'	gctttatgagtagcaggtggt
AW125010_3'	tgttcgagattcctcccaggt
AW125010_5'	aggctcagtcatcagctgtct
AW211368_3'	tcctgatataaaaggaacagact
AW211368_5'	ggtgggaatacttggagtact
AW212775_3'	ctaagggctgctttgctaaa
AW212775_5'	ggccttgtttttggtttgca
Cbln1_3'	gtagctacgaacacgtagga
Cbln1_5'	attgagagggtgacttggga
D78646_3'	gtgcagagacttgaggaaga
D78646_5'	ggcgaaagcgttttggttca
L10426_3'	gacatctggcgttggtacat
L10426_5'	gaggagcagaatggatggat
L21671_3'	gatttgcacatccctgtcga
L21671_5'	ctgaagttgctggatgccaa
M15268_3'	gagctgagtctttctggtgt
M15268_5'	atggtggcagcagctatgtt
M28383_3'	tcgtgatagtcagtgctggt
M28383_5'	aatccgtcaagctccagtgt
M70642_3'	acagttgtaatggcaggcaca
M70642_5'	ccaaagcagctgcaaatacca

Smoc2_3’	tagggtagcgtgaggcaata
Smoc2_5’	caaagatccacagctggaga
Tpd52l1_3’	cttatggaatggcggatgga
Tpd52l1_5’	aaggcttgttggagacggaa
U13393_3'	tgcagacattcgtccttggt
U13393_5'	acgatcatggctgtgcagtt
U17297_3'	gactatggcaggagtgactt
U17297_5'	ccatgaatgagcaggcaagt
X92590_3'	cacaaccacctgtctgtgaa
X92590_5'	gaaagagcagaacctcgtca
Y10495_3'	ctcgtcaacctcactcttca
Y10495_5'	agaaggtggtttctgagcca
Y14771_3'	ttgtccttctcgaccgtgat
Y14771_5'	gtaggaagcaggcagagatt
Y15443_3'	cagcagcaacgggaatatct
Y15443_5'	ggtatggaagaagaagcgct

Appendix 2

Sequence-specific hybridisation probes for real-time PCR

***Gapdh* cDNA:** 5'-attcaacggcacagtcaagg-3' and 5'-tggatgcagggatgatgttc-3', hybridisation probes 5'-ccagaagactgtggatggcccct-x and LC-Red640-tggaaagctgtggcgtgatggc-p.

***Mis*:** 5'-gtcctacatctggctgaagtgat-3' and 5'-ccgagtagggcagaggttct-3', hybridisation probes 5'-ggccctgttagtgctatactctggacc-x and LC-Red640-gcccccaggtcacagtcacagg-p.

***Sox9*:** 5'-gacaagcggaggccgaa-3' and 5'-ccagcttgcacgtcggtt-3', hybridisation probes 5'-cttgcagcgccttgaagatagcattag-x and LC-Red640-gagatgtgagtctgttccgtggcctc-p.

***Sry*:** 5'-agcctatgtgtagttccttggtc-3' and 5'-tgcataaggagtcacattttgct-3', hybridisation probes 5'-caatctggcagttgagttaatgtgcagat-x and LC-Red640-ccattcattcatcccacatatacttgccc-p.

Appendix 3

SiRNA sequences

- **Pax2**
  Target sequence: **aatgtgtcaggcacacagacg**
  Anti-sense strand siRNA: **aatgtgtcaggcacacagacg**cctgtctc
  Sense strand siRNA: **aacgtctgtgtgcctgacaca**cctgtctc

- **Sox 9**

  Target sequence 10: **aacacccggccccaggagaac**
  Anti-sense strand siRNA: **aacacccggccccaggagaac**cctgtctc
  Sense strand siRNA: **aagttctcctggggccgggtg**cctgtctc

  Target sequence 13: **aagaaggagagcgaggaagat**
  Anti-sense strand siRNA: **aagaaggagagcgaggaagat**cctgtctc
  Sense strand siRNA: **aaatcttcctcgctctccttc**cctgtctc

  Target sequence 33: **aagtcggtgaagaacggacaa**
  Anti-sense strand siRNA: **aagtcggtgaagaacggacaa**cctgtctc
  Sense strand siRNA: **aattgtccgttcttcaccgac**cctgtctc

  Target sequence 44: **aaccgacgtgcaagctggcaa**
  Anti-sense strand siRNA: **aaccgacgtgcaagctggcaa**cctgtctc
  Sense strand siRNA: **aattgccagcttgcacgtcgg**cctgtctc

- **Sry**

  Target sequence 6: **aatacagagatcagcaagcag**
  Anti-sense strand siRNA: **aatacagagatcagcaagcag**cctgtctc
  Sense strand siRNA: **aactgcttgctgatctctgta**cctgtctc

  Target sequence 7: **aagcagctgggatgcaggtgg**
  Anti-sense strand siRNA: **aagcagctgggatgcaggtgg**cctgtctc
  Sense strand siRNA: **aaccacctgcatcccagctgc**cctgtctc

  Target sequence 9: **aagccttacagaagccgaaaa**
  Anti-sense strand siRNA: **aagccttacagaagccgaaaa**cctgtctc
  Sense strand siRNA: **aattttcggcttctgtaaggc**cctgtctc

  Target sequence 11: **aaaaaaggccctttttccagg**
  Anti-sense strand siRNA: **aaaaaaggccctttttccagg**cctgtctc
  Sense strand siRNA: **aacctggaaaaagggcctttt**cctgtctc

Appendix 4

Affymetrix report files

(a) Male MG_U74A chip report file
(b) Female MG_U74A chip report file
(c) Male MG_U74B chip report file
(d) Female MG_U74B chip report file
(e) Male MG_U74Av2 chip report file
(f) Female MG_U74Av2 chip report file

Data for control probes β-actin and GAPDH are highlighted in bold. Total probe sets present on array is highlighted in red. The 3'-5' GAPDH ratio is highlighted in blue.

(a) Male MG_U74A chip report file

```
Filename:          Old-A_Male.CHP
Probe Array Type:  MG_U74A
Probe Pair Thr:    8
Controls:          Antisense
```

```
Alpha1:            0.04
Alpha2:            0.06
Tau:         0.015
Noise (RawQ):      16.920
Scale Factor (SF):       0.830
TGT Value:         100
Norm Factor (NF): 1.000
Probe Mask File: C:\PROGRAM FILES\GENECHIP\LIBRARY\MG_U74A.MSK
```

```
Background:
      Avg: 439.93 Std: 11.85  Min: 418.90 Max: 467.60
Noise:
      Avg: 12.91  Std: 0.59   Min: 11.50  Max: 15.00
Corner+
      Avg: 183          Count: 32
Corner-
      Avg: 2937         Count: 32
Central-
      Avg: 4073         Count: 9
```

```
The following data represents probe sets that exceed the probe pair
threshold and are not called "No Call".

Total Probe Sets: 10028
Number Present:   3162  31.5%
Number Absent:    6616  66.0%
Number Marginal:  250   2.5%

Average Signal (P):      508.8
Average Signal (A):      31.4
Average Signal (M):      90.4
Average Signal (All):    183.4
```

```
Housekeeping Controls:
Probe Set          Sig(5')     Det(5')     Sig(M')     Det(M')
      Sig(3')      Det(3')     Sig(all)    Sig(3'/5')
B-ACTINMUR/M12481 821.1 P      1054.1      P     2762.5      P
      1545.93      3.36
GAPDHMUR/M32599   2407.6       P     2378.4      P     4256.7      P
      3014.24      1.77
TRANSRECMUR/X57349      3.7   A     23.5  A     22.3  A     16.47
      6.11
PYRUCARBMUR/L09192      46.0  A     10.6  A     91.3  A     49.30
      1.98
18SRNAMUR         166.3 P     1337.4      P     214.4 P     572.70
      1.29
```

(b) Female MG_U74A chip report file

```
Filename:          Old-A_Female.CHP
Probe Array Type: MG_U74A
Probe Pair Thr:    8
Controls:          Antisense
```

```
Alpha1:            0.04
Alpha2:            0.06
Tau:         0.015
Noise (RawQ):      17.790
Scale Factor (SF):       1.027
TGT Value:         100
Norm Factor (NF): 1.000
Probe Mask File: C:\PROGRAM FILES\GENECHIP\LIBRARY\MG_U74A.MSK
```

```
Background:
      Avg: 441.19 Std: 15.70  Min: 412.20 Max: 480.10
Noise:
      Avg: 14.24  Std: 0.71   Min: 12.40  Max: 16.00
Corner+
      Avg: 202          Count: 32
Corner-
      Avg: 3358         Count: 32
Central-
      Avg: 3816         Count: 9
```

The following data represents probe sets that exceed the probe pair threshold and are not called "No Call".

```
Total Probe Sets: 10028
Number Present:   2855  28.5%
Number Absent:    6915  69.0%
Number Marginal:  258   2.6%

Average Signal (P):      506.0
Average Signal (A):      38.9
Average Signal (M):      112.2
Average Signal (All):    173.8
```

Housekeeping Controls:

Probe Set	Sig(5')	Det(5')	Sig(M')	Det(M')	
Sig(3')	Det(3')	Sig(all)	Sig(3'/5')		
B-ACTINMUR/M12481	**459.9 P**	**677.4 P**	**2262.7**	**P**	**1133.31**
4.92					
GAPDHMUR/M32599	**1269.0**	**P**	**1952.3**	**P**	**3107.1** **P**
2109.46	**2.45**				
TRANSRECMUR/X57349	5.9	A	31.3	A	22.3 A 19.82
3.77					
PYRUCARBMUR/L09192	51.3	A	16.1	A	117.6 M 61.67
2.29					
18SRNAMUR	109.9 P	576.3 P	89.6	P	258.60
0.82					

(c) Male MG_U74B chip report file

```
Filename:          Old-B_Male.CHP
Probe Array Type: MG_U74B
Probe Pair Thr:    8
Controls:          Antisense
```

```
Alpha1:            0.04
Alpha2:            0.06
Tau:         0.015
Noise (RawQ):      12.880
Scale Factor (SF):       3.318
TGT Value:         100
Norm Factor (NF): 1.000
Probe Mask File: C:\PROGRAM FILES\GENECHIP\LIBRARY\MG_U74B.MSK
```

```
Background:
      Avg: 432.24 Std: 17.09  Min: 391.30 Max: 493.30
Noise:
      Avg: 10.32  Std: 0.94   Min: 8.80   Max: 13.90
Corner+
      Avg: 583          Count: 32
Corner-
      Avg: 45748        Count: 32
Central-
      Avg: 45616        Count: 9
```

```
The following data represents probe sets that exceed the probe pair
threshold and are not called "No Call".

Total Probe Sets: 9580
Number Present:   1159  12.1%
Number Absent:    8236  86.0%
Number Marginal:  185   1.9%

Average Signal (P):      466.0
Average Signal (A):      73.1
Average Signal (M):      209.8
Average Signal (All):    123.2
```

```
Housekeeping Controls:
Probe Set          Sig(5')      Det(5')      Sig(M')      Det(M')
      Sig(3')      Det(3')      Sig(all)     Sig(3'/5')
B-ACTINMUR/M12481 1690.6       P      2122.0       P      7233.5       P
      3682.06      4.28
GAPDHMUR/M32599   3925.6       P      6383.0       P      8697.6       P
      6335.41      2.22
TRANSRECMUR/X57349       90.4  A      66.2  A      186.3 P      114.32
      2.06
PYRUCARBMUR/L09192       15.1  A      42.7  A      181.0 A      79.61
      11.98
18SRNAMUR          412.4 P      4829.2       P      513.0 P      1918.18
      1.24
```

(d) Female MG_U74B chip report file

```
Filename:          Old-B_Female.CHP
Probe Array Type: MG_U74B
Probe Pair Thr:    8
Controls:          Antisense
```

```
Alpha1:            0.04
Alpha2:            0.06
Tau:         0.015
Noise (RawQ):      11.700
Scale Factor (SF):       3.389
TGT Value:         100
Norm Factor (NF): 1.000
Probe Mask File: C:\PROGRAM FILES\GENECHIP\LIBRARY\MG_U74B.MSK
```

```
Background:
      Avg: 348.53 Std: 5.74   Min: 336.20 Max: 369.30
Noise:
      Avg: 9.33   Std: 0.36   Min: 8.30   Max: 10.30
Corner+
      Avg: 1193          Count: 32
Corner-
      Avg: 45825         Count: 32
Central-
      Avg: 45823         Count: 9
```

```
The following data represents probe sets that exceed the probe pair
threshold and are not called "No Call".

Total Probe Sets: 9580
Number Present:    1407   14.7%
Number Absent:     7973   83.2%
Number Marginal:   200    2.1%

Average Signal (P):      441.8
Average Signal (A):      67.3
Average Signal (M):      212.1
Average Signal (All):    125.3
```

```
Housekeeping Controls:
Probe Set          Sig(5')      Det(5')      Sig(M')      Det(M')
      Sig(3')      Det(3')      Sig(all)     Sig(3'/5')
B-ACTINMUR/M12481  1193.2       P      1799.9       P      5927.9       P
      2973.67      4.97
GAPDHMUR/M32599    1996.4       P      5157.8       P      7041.5       P
      4731.91      3.53
TRANSRECMUR/X57349       71.2   A      7.5    A      53.5   A      44.07
      0.75
PYRUCARBMUR/L09192       21.7   A      17.2   A      157.0  A      65.32
      7.22
18SRNAMUR          278.8 M      2401.3       P      472.9  P      1050.99
      1.70
```

(e) Male MG_U74Av2 chip report file

```
Filename:          Male.CHP
Probe Array Type:  MG_U74Av2
Algorithm:         Statistical
Probe Pair Thr:    8
Controls:          Antisense
```

```
Alpha1:            0.04
Alpha2:            0.06
Tau:         0.015
Noise (RawQ):      3.180
Scale Factor (SF):       2.099
TGT Value:         200
Norm Factor (NF): 1.000
```

```
Background:
      Avg: 74.44   Std: 1.64    Min: 71.60   Max: 80.10
Noise:
      Avg: 3.29    Std: 0.15    Min: 2.80    Max: 3.60
Corner+
      Avg: 137           Count: 32
Corner-
      Avg: 11730         Count: 32
Central-
      Avg: 9620          Count: 9
```

The following data represents probe sets that exceed the probe pair threshold and are not called "No Call".

```
Total Probe Sets: 12473
Number Present:   4900   39.3%
Number Absent:    7221   57.9%
Number Marginal:  352    2.8%

Average Signal (P):      671.4
Average Signal (A):      50.4
Average Signal (M):      141.7
Average Signal (All):    297.0
```

```
Housekeeping Controls:
Probe Set          Sig(5')      Det(5')      Sig(M')      Det(M')
      Sig(3')      Det(3')      Sig(all)     Sig(3'/5')
B-ACTINMUR/M12481 305.7 P      1193.8       P     8942.5       P
      3480.68      29.25
GAPDHMUR/M32599    110.4 P      502.0 P      6669.2       P     2427.22
      60.38
TRANSRECMUR/X57349       20.5  A      1.9   A      47.5  P      23.31
      2.31
PYRUCARBMUR/L09192       76.6  A      55.8  A      43.4  A      58.59
      0.57
18SRNAMUR          424.7 P      1641.2       P     944.1 P      1003.32
      2.22
```

Spike Controls:

Probe Set	Sig(5')	Det(5')	Sig(M')	Det(M')
Sig(3')	Det(3')	Sig(all)	Sig(3'/5')	
BIOB	153.1 M	257.4 P	191.9 P	200.81 1.25
BIOC	537.9 P		415.5 P	476.70 0.77
BIODN	479.9 P		2761.8 P	1620.86 5.76
CREX	5501.5 P		9130.9 P	7316.23 1.66
DAPX	16.4 A	38.0 A	35.9 A	30.09 2.19
LYSX	5.8 A	3.9 A	2.7 A	4.15 0.46
PHEX	2.9 A	3.5 A	88.4 A	31.60 30.15
THRX	2.5 A	5.3 A	11.4 A	6.42 4.50
TRPNX	6.2 A	3.0 A	4.2 A	4.47 0.68

(f) Female MG_U74Av2 chip report file

```
Filename:          Female.CHP
Probe Array Type:  MG_U74Av2
Algorithm:         Statistical
Probe Pair Thr:    8
Controls:          Antisense
```

```
Alpha1:            0.04
Alpha2:            0.06
Tau:         0.015
Noise (RawQ):      2.140
Scale Factor (SF):       18.518
TGT Value:         200
Norm Factor (NF): 1.000
```

```
Background:
      Avg: 57.13   Std: 0.73    Min: 55.60   Max: 58.70
Noise:
      Avg: 1.81    Std: 0.05    Min: 1.70    Max: 1.90
Corner+
      Avg: 79            Count: 32
Corner-
      Avg: 11204         Count: 32
Central-
      Avg: 10722         Count: 9
```

```
The following data represents probe sets that exceed the probe pair
threshold and are not called "No Call".

Total Probe Sets: 12473
Number Present:   1564   12.5%
Number Absent:    10667  85.5%
Number Marginal:  242    1.9%

Average Signal (P):      1981.9
Average Signal (A):      140.7
Average Signal (M):      462.3
Average Signal (All):    377.8
```

```
Housekeeping Controls:
Probe Set          Sig(5')      Det(5')      Sig(M')      Det(M')
      Sig(3')      Det(3')      Sig(all)     Sig(3'/5')
B-ACTINMUR/M12481  233.7 A      50.8  A      2445.0       P      909.83
      10.46
GAPDHMUR/M32599    238.4 A      363.0 A      308.3 A      303.23
      1.29
TRANSRECMUR/X57349       63.3   A      14.1   A      139.1 A      72.17
      2.20
PYRUCARBMUR/L09192       37.3   A      21.8   A      88.7  A      49.24
      2.38
18SRNAMUR          3061.7       P      10553.8      P      4031.5       P
      5882.33      1.32
```

Spike Controls:

Probe Set	Sig(5')		Det(5')		Sig(M')		Det(M')	
	Sig(3')		Det(3')		Sig(all)		Sig(3'/5')	
BIOB	2092.6	P	4471.8	P	2559.1	P	3041.14	1.22
BIOC	6170.2	P			4709.2	P	5439.68	0.76
BIODN	6125.4	P			26887.7	P	16506.58	4.39
CREX	56520.2	P			93378.5	P	74949.34	1.65
DAPX	75.6	A	113.8	A	23.0	A	70.82	0.30
LYSX	94.5	A	93.7	A	121.3	A	103.16	1.28
PHEX	21.6	A	25.8	A	24.6	A	23.98	1.14
THRX	46.8	A	97.5	A	30.9	A	58.39	0.66
TRPNX	46.2	A	32.4	A	67.4	A	48.70	1.46

Appendix 5 List of sex differentiating genes from screen 1 (MG_U74A)

	M_F-Av1 Avg Diff	M_F-Av1 Abs Call	M_F-Av1 Diff Call	M_F-Av1 Fold Change	Description
104637_at	5990.8	P	I	12.3	Cluster Incl AI326397:mm20b10.x1 Mus musculus cDNA, 3 end
104030_at	**6499**	**P**	**I**	**11.4**	**Cluster Incl AI848841:UI-M-AJ1-ahb-c-08-0-UI.s1 Mus musculus cDNA, 3 end**
100068_at	12070.2	P	I	8.5	Cluster Incl M74570:Mouse aldehyde dehydrogenase II mRNA, complete cds
103296_at	**5003**	**P**	**I**	**7.3**	**Cluster Incl Y18243:Mus musculus mRNA for testatin**
101882_s_	2268.5	P	I	7.2	Cluster Incl U03715:Procollagen, type XVIII, alpha 1
92932_at	3730.8	P	I	6.7	Cluster Incl X61448:M.musculus mRNA for D3 clone
103050_at	1622.9	P	I	6.5	Cluster Incl AF035717:Transcription factor 21
98469_at	2666.9	A	NC	5.9	Cluster Incl D21099:Serine/threonine kinase 5
100601_at	2675.2	P	I	5.3	Cluster Incl AF022110:Mus musculus integrin beta-5 mRNA, complete cds
102105_f_	3801.9	P	I	5.2	Cluster Incl AI840733:UI-M-AM0-adn-b-04-0-UI.s1 Mus musculus cDNA, 3 end
98423_at	5793.3	P	I	4.9	Cluster Incl M81445:Gap junction membrane channel protein beta 2
101082_at	3163.4	P	I	4.6	Cluster Incl J02652:Malic enzyme, supernatant
101446_at	6043.1	P	I	4.3	Cluster Incl AF004428:Tumor protein D52-like 1
96336_at	5669.2	P	I	4.3	Cluster Incl AI844626:UI-M-AL1-ahr-a-04-0-UI.s1 Mus musculus cDNA, 3 end
96926_at	1994.1	P	I	4.2	Cluster Incl AA980164:ua31a05.r1 Mus musculus cDNA, 5 end
98480_s_a	1763.7	P	I	4.2	Cluster Incl M32352:Mouse renin (Ren-1-d) gene, complete cds
100928_at	2121.7	P	I	4.1	Cluster Incl X75285:Fibulin 2
93009_at	6420	P	MI	4.1	Cluster Incl J04696:Mouse glutathione S-transferase class mu (GST5-5) mRNA, complete cds
92738_at	1910.4	P	I	4	Cluster Incl D49921:Mouse mRNA for glial cell line-derived neurotrophic factor(GDNF), complete cds
97487_at	1857.4	P	I	4	Cluster Incl X70296:Serine protease inhibitor 4
94185_at	399.6	A	I	3.9	Cluster Incl AF056492:Mus musculus TRAAK K+ channel subunit mRNA, complete cds
94555_at	1544.4	A	NC	3.9	Cluster Incl AI845579:UI-M-AQ1-adx-b-12-0-UI.s1 Mus musculus cDNA, 3 end
96771_at	2116.9	P	I	3.9	Cluster Incl AI006228:ua88d09.r1 Mus musculus cDNA, 5 end
97519_at	1396.8	P	I	3.9	Cluster Incl X13986:Mouse mRNA for minopontin
101027_s_	2132.9	P	NC	3.6	Cluster Incl AF069051:Mus musculus pituitary tumor transforming gene protein (PTTG) mRNA, comple
104165_at	1762.3	P	I	3.6	Cluster Incl AJ132098:Mus musculus mRNA for Vanin-1
93187_at	1870.5	P	I	3.6	Cluster Incl AW048347:UI-M-BH1-akk-e-03-0-UI.s1 Mus musculus cDNA, 3 end
99694_r_a	-574	A	NC	3.6	Cluster Incl AV326295:AV326295 Mus musculus cDNA, 3 end
94778_at	3592.3	P	MI	3.5	Cluster Incl U96401:Mus musculus aldehyde dehydrogenase Ahd-2-like mRNA, complete cds
97660_s_a	-1197.3	A	NC	3.5	Cluster Incl AJ223070:Mus musculus mRNA for TCF-4 protein

101531_at	3705.7	P	I	3.4	Cluster Incl AI527354:uj51b11.x1 Mus musculus cDNA, 3 end
102070_at	3457.5	P	I	3.4	Cluster Incl AW212495:uo89g01.x1 Mus musculus cDNA, 3 end
92442_at	1350	A	MI	3.4	Cluster Incl AI840554:UI-M-AM0-adp-f-02-0-UI.s1 Mus musculus cDNA, 3 end
99010_at	1699.8	P	I	3.4	Cluster Incl AB024538:Mus musculus Islr(immunoglobulin superfamily containing leucine-rich repeat) m
101365_at	451.6	A	MI	3.3	Cluster Incl M81342:Fibroblast growth factor receptor 3
96876_at	2996.3	P	I	3.3	Cluster Incl U34259:Mus musculus Golgi 4-transmembrane spanning transporter MTP mRNA, complet
100613_at	1602.3	P	NC	3.2	Cluster Incl AF030635:Mus musculus FK-506 binding protein homolog (SAM11) mRNA, complete cds
101345_at	-1156	A	NC	3.2	Cluster Incl AF030001:Mus musculus major histocompatibility locus class III region-butyrophilin-like pro
103451_at	2039	P	I	3.2	Cluster Incl AI835159:UI-M-AQ0-aaf-e-10-0-UI.s1 Mus musculus cDNA, 3 end
94208_at	1496.2	P	I	3.2	Cluster Incl AW045202:UI-M-BH1-alh-c-08-0-UI.s1 Mus musculus cDNA, 3 end
96994_at	-276.3	A	NC	3.2	Cluster Incl AJ132691:Mus musculus (strain 129) partial HIC-1 gene, exon 1
100279_at	-24	A	NC	3.1	Cluster Incl X76989:M.musculus mRNA for gal beta 1,3 galNAc alpha 2,3-sialyltransferase
100328_s_	-15.5	A	NC	3.1	Cluster Incl U96684:Paired-Ig-like receptor A3
102094_f_	70247.2	P	I	3.1	Cluster Incl AI841270:UI-M-AM0-adu-d-10-0-UI.s1 Mus musculus cDNA, 3 end
102770_at	2008.8	P	NC	3.1	Cluster Incl M26385:Glycophorin A
103891_i_	1665	P	I	3.1	Cluster Incl AI197161:ue51h10.r1 Mus musculus cDNA, 5 end
104135_at	1363.9	P	I	3.1	Cluster Incl AW045474:UI-M-BH1-akr-c-02-0-UI.s1 Mus musculus cDNA, 3 end
92197_r_a	-5578.1	A	NC	3.1	Cluster Incl X83733:M.musculus SAP62-AMH gene
93294_at	1334.3	P	I	3.1	Cluster Incl M70642:Fibroblast inducible secreted protein
94301_at	2033.8	P	I	3.1	Cluster Incl AI843269:UI-M-AO1-aei-c-08-0-UI.s1 Mus musculus cDNA, 3 end
99475_at	7576.1	P	I	3.1	Cluster Incl U88327:Mus musculus suppressor of cytokine signalling-2 (SOCS-2) mRNA, complete cds
101971_at	4508	P	MI	3	Cluster Incl AW061255:UI-M-BH1-amj-e-08-0-UI.s2 Mus musculus cDNA, 3 end
103580_at	961.1	A	NC	3	Cluster Incl AI845588:UI-M-AQ1-adx-c-11-0-UI.s1 Mus musculus cDNA, 3 end
92473_at	-19.3	A	NC	3	Cluster Incl Y11092:M.musculus mRNA for map kinase interacting kinase, Mnk2
92545_f_at	1306.9	P	I	3	Cluster Incl AB006361:Mus musculus mRNA for prostaglandin D synthetase, complete cds
94831_at	449.3	P	NC	3	Cluster Incl M65270:Mouse cathepsin B gene
97542_at	-238.1	A	NC	3	Cluster Incl U04443:Mus musculus non-muscle myosin light chain 3 (MLC3nm) mRNA, partial cds
104625_at	413.3	A	D	-3	Cluster Incl AA874130:vw87g02.r1 Mus musculus cDNA, 5 end
104675_at	-623.7	A	NC	-3	Cluster Incl D01093:Proprotein convertase subtilisin/kexin type 4
97317_at	229.5	P	D	-3	Cluster Incl AW122933:UI-M-BH2.1-apa-h-09-0-UI.s1 Mus musculus cDNA, 3 end
98627_at	1249.3	P	MD	-3	Cluster Incl X81580:Insulin-like growth factor binding protein 2
102789_at	371.3	A	NC	-3.1	Cluster Incl AB000096:GATA-binding protein 2
103163_f_	79.8	A	D	-3.1	Cluster Incl AV092014:AV092014 Mus musculus cDNA
103175_f_	-1728.7	A	NC	-3.1	Cluster Incl AV120644:AV120644 Mus musculus cDNA
103655_at	-23.4	A	NC	-3.1	Cluster Incl AF074882:Mus musculus strain C57BL/6J histone deacetylase 3 (Hdac3) mRNA, alternativ
100385_at	561.1	A	NC	-3.2	Cluster Incl X75832:M.musculus Glra1 gene (exon1)

103238_at	-197.4	A	D	-3.2	Cluster Incl M89797:Wingless-related MMTV integration site 4
92778_i_at	466.3	A	D	-3.2	Cluster Incl Z22552:M.musculus membrane glycoprotein gene
94987_at	75.2	A	D	-3.2	Cluster Incl U25051:Phosphatidylethanolamine N-methyltransferase
95397_at	668.2	A	MD	-3.2	Cluster Incl AI852661:UI-M-BH0-aji-a-10-0-UI.s1 Mus musculus cDNA, 3 end
96605_at	313.1	A	D	-3.2	Cluster Incl AI787183:ui86a11.y1 Mus musculus cDNA, 5 end
99230_f_at	-170.5	A	NC	-3.2	Cluster Incl AV377350:AV377350 Mus musculus cDNA, 3 end
103628_at	645.8	A	NC	-3.3	Cluster Incl D16503:Lymphoid enhancer binding factor 1
93027_r_at	-682.9	A	NC	-3.3	Cluster Incl AI836214:UI-M-AP0-abg-a-02-0-UI.s1 Mus musculus cDNA, 3 end
97245_at	189.4	A	MD	-3.3	Cluster Incl AW123191:UI-M-BH2.1-apg-b-09-0-UI.s1 Mus musculus cDNA, 3 end
98714_at	-316.8	A	NC	-3.3	Cluster Incl AV038316:AV038316 Mus musculus cDNA
93769_at	-317.9	A	NC	-3.4	Cluster Incl AI838708:UI-M-AO0-acf-a-02-0-UI.s1 Mus musculus cDNA, 3 end
98133_at	210.6	A	NC	-3.4	Cluster Incl D26352:Calbindin-28K
103921_i_at	926.2	M	D	-3.5	Cluster Incl AI839690:UI-M-AN0-acp-e-04-0-UI.s1 Mus musculus cDNA, 3 end
94881_at	12.7	A	D	-3.5	Cluster Incl AW048937:UI-M-BH1-amo-d-08-0-UI.s1 Mus musculus cDNA, 3 end
100263_r_at	-586.9	A	NC	-3.6	Cluster Incl AV336781:AV336781 Mus musculus cDNA, 3 end
103293_at	-1109.6	A	D	-3.6	Cluster Incl AF079901:Mus musculus 28 kDa cis-Golgi SNARE (GS28) mRNA, complete cds
103729_at	671.3	A	D	-3.6	Cluster Incl M36775:Laminin, alpha 1
93594_r_at	785.2	A	MD	-3.8	Cluster Incl U87948:Epithelial membrane protein 3
100483_at	1521	P	D	-3.9	Cluster Incl M89641:Interferon (alpha and beta) receptor
93090_at	2617.4	P	D	-3.9	Cluster Incl M23362:Fibroblast growth factor receptor 2
100904_at	-683.7	A	NC	-4	Cluster Incl AI852257:UI-M-BH0-aje-b-03-0-UI.s1 Mus musculus cDNA, 3 end
104548_at	432.8	A	D	-4	Cluster Incl Y15443:Mus musculus mRNA for 50C15 protein
95135_at	740.1	A	D	-4	Cluster Incl AI844396:UI-M-AL1-ahp-e-05-0-UI.s1 Mus musculus cDNA, 3 end
95793_at	341	A	D	-4.1	Cluster Incl AJ005562:Small proline-rich protein 2D
98131_at	-120.9	A	D	-4.2	Cluster Incl D78646:Mouse mRNA for zeta-crystallin/quinone reductase, partial cds
101441_i_at	797.5	P	D	-4.5	Cluster Incl AF031127:Mus musculus inositol trisphosphate receptor type 2 (Itpr2) gene, partial cds
99034_at	547.3	P	NC	-4.5	Cluster Incl Y15001:Mus musculus mRNA for iroquois homeobox protein 3
103902_at	571.1	A	D	-5.3	Cluster Incl AI835436:UI-M-AQ0-aaj-h-09-0-UI.s1 Mus musculus cDNA, 3 end
95234_f_at	-1202.8	A	NC	-5.3	Cluster Incl AV350152:AV350152 Mus musculus cDNA, 3 end
101526_at	520.1	A	MD	-6.4	Cluster Incl X14759:Homeo box, msh-like 1
98817_at	**117.4**	**A**	**D**	**-9**	**Cluster Incl Z29532:M.musculus mRNA for follistatin**

Appendix 6 Genes more than 2 fold and present in at least 1 in both MAS v4 and v5

	female		male			
Systematic	Flags	Raw	Flags	Raw	Genbank	Description
104637_at	A	6.5	P	372.3	AI326397	glutathione S-transferase, mu 6
111976_at	A	8.4	P	215.5	AA960347	serine protease inhibitor, Kunitz type 1
114872_at	A	12.1	P	241.7	AI467050	RIKEN cDNA 1600021C16 gene
98459_at	A	8.4	P	134	AA913994	serine hydroxymethyl transferase 1 (soluble)
103674_f_at	A	4	P	52.5	AJ006584	eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked
93327_at	A	9.6	P	113.6	AI842665	RIKEN cDNA 1300011C24 gene
111209_at	A	34.7	P	395	AW122433	down-regulated by Ctnnb1, a
100068_at	P	71	P	768.6	M74570	aldehyde dehydrogenase family 1, subfamily A1
98405_at	A	4.4	P	44.3	U96700	serine protease inhibitor 6
112900_at	A	26.1	P	244	AA682034	mitochondrial ribosomal protein 63
113193_at	A	23.8	P	214	AI841117	guanine nucleotide releasing factor 2
102646_at	A	7.7	P	68.6	AW122284	Mus musculus, clone IMAGE:4972955, mRNA
103296_at	P	49	P	426.5	Y18243	cystatin 9
104044_at	A	12.5	P	101.5	AI853663	RIKEN cDNA 1300006N24 gene
114131_at	A	43.8	P	353.5	AI846599	ESTs, Moderately similar to NRTR_MOUSE Neurturin receptor alpha precursor (NTNR-alpha)
97104_g_at	A	5.9	P	46.2	AF031380	hypothetical protein; Mus musculus KOI-4 gene, partial cds.
99535_at	A	13.1	P	101.8	AW047630	CCR4 carbon catabolite repression 4-like (S. cerevisiae)
98480_s_at	A	21.5	P	161.7	M32352	renin (Ren-1-d); Mouse renin (Ren-1-d) gene, complete cds.
94208_at	A	19.6	P	136.2	AW045202	protein disulfide isomerase-related protein
109365_at	A	36.7	P	253.6	AI838035	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130083H14
99098_at	A	11.9	P	75.7	AW045533	farnesyl diphosphate synthetase
104030_at	A	62	P	390	AI848841	UI-M-AJ1-ahb-c-08-0-UI.s1 NIH_BMAP_MOB_N Mus musculus cDNA clone
109411_at	A	34	P	204.9	AA656082	optic atrophy 1 homolog (human)
97310_at	A	12.2	P	72.6	AW124318	RIKEN cDNA 3110002K08 gene
96336_at	P	67.5	P	399.5	AI844626	glycine amidinotransferase (L-arginine:glycine amidinotransferase)
93009_at	P	113.6	P	665.9	J04696	glutathione S-transferase, mu 2
117044_at	A	68.7	P	371.2	AI850402	RIKEN cDNA C130078N17 gene
103842_at	A	15.7	P	84.2	AJ007376	DEAD (aspartate-glutamate-alanine-aspartate) box polypeptide, Y chromosome
96325_at	A	8.4	P	44	AW124874	RIKEN cDNA 2510039O18 gene
101513_at	A	7.7	P	40.3	AI840454	prefoldin 5
115260_at	A	39	P	200.6	AW121522	RIKEN cDNA 1200015K23 gene

92932_at	A	64.4	P	325.2	X61448	cerebellin 1 precursor protein
AFFX-Trans	A	37.9	P	186.3	X57349	transferrin receptor
96876_at	P	59.3	P	278.4	U34259	lysosomal-associated protein transmembrane 4A
101446_at	A	100.7	P	467.3	AF004428	tumor protein D52-like 1
93187_at	A	45.5	P	203.5	AW048347	RIKEN cDNA 2210023F24 gene
110494_at	A	58	P	250.7	AA239040	RIKEN cDNA 2810411G20 gene
92841_f_at	A	14.6	P	63.1	X51429	chromogranin B
104158_at	A	25.1	P	102.1	AW046671	UI-M-BH1-alm-a-09-0-UI.s1 NIH_BMAP_M_S2 Mus musculus cDNA clone
113079_at	A	56.4	P	226.5	AA238413	Mak3p homolog (S. cerevisiae)
108537_at	A	39.6	P	156.1	AI846354	RIKEN cDNA 6330565B14 gene
101082_at	M	81	P	317	J02652	malic enzyme, supernatant
113759_at	A	62.4	P	241.9	AI847140	phosphodiesterase 4B, cAMP specific
109172_at	A	118.2	P	456	AI854623	RIKEN cDNA 2500002K03 gene
102070_at	P	105.3	P	392.4	AW212495	procollagen, type IX, alpha 3
100928_at	A	65.8	P	237.8	X75285	fibulin 2
97504_at	A	54.5	P	196.2	M83749	cyclin D2
114820_at	A	94.2	P	339	AI553622	RIKEN cDNA A230102O09 gene
107566_at	A	94.7	P	337.9	AI849532	RIKEN cDNA 1810054O13 gene
95599_at	P	197.8	P	694.1	D28941	sialyltransferase 4C (beta-galactosidase alpha-2,3-sialytransferase)
96771_at	P	54.9	P	190.8	AI006228	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)
101971_at	P	118.7	P	405.5	AW061255	RIKEN cDNA 2500002L14 gene
109976_at	A	58.6	P	198	AI159735	hypothetical protein 6430576E21
95634_at	A	14.1	P	47.2	AI848107	RIKEN cDNA 0610010K14 gene
99475_at	P	208.3	P	694.1	U88327	suppressor of cytokine signaling 2
112784_at	A	85.2	P	275.9	AI839993	RIKEN cDNA 9330185J12 gene
109521_at	A	105.5	P	338.1	AW123386	RIKEN cDNA A230106A15 gene
101869_s_a	P	1901	P	6073.3	J00413	Mouse beta-globin major gene.
96135_at	P	58.6	P	183.8	AA833425	RIKEN cDNA 3110003A17 gene
109361_at	A	81.1	P	252.1	AA170632	RIKEN cDNA 1200009B18 gene
114058_at	A	105.2	P	324.6	AA152809	RIKEN cDNA 3000004C01 gene
108018_at	A	208.2	P	639.4	AA959436	RIKEN cDNA 2810002E22 gene
92738_at	A	28.3	P	86.9	D49921	glial cell line derived neurotrophic factor
103891_i_at	A	54.4	P	166.8	AI197161	ELL-related RNA polymerase II, elongation factor
103220_at	A	23.6	P	71.7	AI851350	RIKEN cDNA 5730494G16 gene
114461_at	A	54.9	P	164.2	AI874945	ESTs, Weakly similar to I48668 zinc finger protein 51 - mouse [M.musculus]
112860_at	P	174.3	P	520.3	AI837433	RIKEN cDNA 2310066E14 gene

97487_at	A	49.1	P	146.2	X70296	serine (or cysteine) proteinase inhibitor, clade E, member 2
115879_at	P	167.5	P	495.9	AI837146	RIKEN cDNA 1700012G19 gene
94778_at	A	29.6	P	87.5	U96401	aldehyde dehydrogenase family 1, subfamily A7
92294_at	A	19	P	56	AW060793	ESTs, Weakly similar to JC4945 ADP-ribosylation factor 1 - mouse [M.musculus]
114328_f_at	A	74.2	P	209.6	AI227486	Mus musculus, clone IMAGE:4015738, mRNA
103451_at	P	72.7	P	203	AI835159	UI-M-AQ0-aaf-e-10-0-UI.s1 NIH_BMAP_MHI Mus musculus cDNA clone UI-M-AQ0-aaf-e-10-0-UI 3'
93294_at	A	22.4	P	62.4	M70642	connective tissue growth factor
108789_at	A	96.4	P	263.3	AA759531	Mus musculus 12 days embryo embryonic body between diaphragm region and neck cDNA
110713_at	A	59.1	P	161.3	AW124288	RIKEN cDNA 2810055F11 gene
106281_f_at	P	821.9	P	2236.7	AI851600	Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library
115029_at	A	84.5	P	229.6	AI626964	DNA segment, Chr 10, ERATO Doi 739, expressed
99078_at	A	32.3	P	85.2	AI839522	RIKEN cDNA 1110033C18 gene
96658_at	A	23.1	P	60.9	AI841906	RIKEN cDNA 2900010J23 gene
100528_at	A	31.5	P	82.6	AI852278	ubiquitin-conjugating enzyme E2H
112674_g_a	P	502	P	1304.8	AI414938	RIKEN cDNA 6030408B16 gene
100333_at	A	31.2	P	80.7	M13521	serum amyloid A-1; Murine serum amyloid A-1 (SAA-1) gene.
113902_at	A	76	P	195.3	AA759863	RIKEN cDNA 2610001M19 gene
93596_i_at	P	146.8	P	373.5	AA624586	RIKEN cDNA 2410043G19 gene
92564_at	A	21.8	P	55.3	AI891475	leucine rich repeat (in FLII) interacting protein 1
104562_at	A	24.9	P	63.1	AW049632	RIKEN cDNA 5730403M16 gene
96221_at	A	32.5	P	81.8	AI606300	expressed sequence AI429613
112452_at	A	115.4	P	290.2	AA967482	hypothetical protein LOC234699
107074_at	M	117.5	P	293.5	AI838083	guanosine monophosphate reductase
112824_at	A	104.1	P	258.8	AI847975	RIKEN cDNA 1110015K06 gene
92545_f_at	A	58	P	144	AB006361	prostaglandin D2 synthase (brain)
103065_at	A	39	P	96.1	M73696	solute carrier family 20, member 1
94781_at	P	2807	P	6883.2	V00714	Mouse gene for alpha-globin.
103901_at	A	23.3	P	56.9	AW213777	RIKEN cDNA 4930451A13 gene
106195_at	A	81.8	P	199.4	AI851948	RIKEN cDNA C130018M11 gene
101992_at	M	124.4	P	302.9	U13393	proteasome (prosome, macropain) subunit, beta type 6
93285_at	P	267.4	P	650.9	AI845584	dual specificity phosphatase 6
111849_at	A	108.8	P	263.5	AW050288	leucine-zipper-like transcriptional regulator, 1
102094_f_at	P	944.1	P	2277.5	AI841270	UI-M-AM0-adu-d-10-0-UI.s1 NIH_BMAP_MAM Mus musculus cDNA clone
104333_at	A	38.1	P	91.9	U69488	Mus musculus viral envelope like protein (G7e) gene, complete cds.
97468_at	P	168.4	P	404.9	AB025409	CDC28 protein kinase 1
93482_at	P	159.5	P	382.8	AI117835	RIKEN cDNA 9530072E15 gene

108780_at	A	110.3	P	264.7	AI845395	UI-M-AO1-aej-a-05-0-UI.s1 NIH_BMAP_MPG_N Mus musculus cDNA clone
116198_at	A	85.3	P	203.5	AI152021	anaphase-promoting complex subunit 7
97446_at	A	68.6	P	161.8	AW125010	RIKEN cDNA 2810477H02 gene
109533_at	A	70.4	P	165.8	AW121643	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130021L12
96900_at	P	230.4	P	531.8	AW125480	RIKEN cDNA 1620401E04 gene
111903_at	A	173.1	P	397.6	AI852203	RIKEN cDNA 9330175B01 gene
110958_at	P	219.4	P	503.4	AA959919	RIKEN cDNA 6720458D17 gene
116165_at	A	82.9	P	188.7	AA755396	RIKEN cDNA 2010011I20 gene
104165_at	A	58.6	P	133.2	AJ132098	
103953_at	A	49.4	P	112.1	U91538	SEC22 vesicle trafficking protein-like 1 (S. cerevisiae)
97259_at	A	42.9	P	97	AJ005983	cyclic AMP phosphoprotein
95406_at	P	389.7	P	879.4	AW125347	RIKEN cDNA 1810037I17 gene
97312_at	P	97.9	P	219.3	AB014464	CD164 antigen
115514_at	P	211.9	P	474.3	AA387402	RIKEN cDNA 8030402F09 gene
109161_at	A	153.9	P	344.1	AW121069	phosphatidic acid phosphatase type 2B
110531_at	A	79.7	P	177.7	AA466584	ESTs
104477_at	P	225.1	P	496.6	AW047643	ESTs
96926_at	A	53.5	P	117.4	AA980164	SPARC related modular calcium binding 2
104135_at	A	52.9	P	113.4	AW045474	ADP-ribosylation-like 3
96629_at	P	119.1	P	254.2	X04097	DNA segment, Chr 7, Roswell Park 2 complex, expressed
96275_f_at	A	36.8	P	78.3	AI843327	Mus musculus 10 day old male pancreas cDNA, RIKEN full-length enriched library, clone
114663_at	A	122.9	P	260.4	AI835464	RIKEN cDNA B230208H17 gene
113001_at	A	79.5	P	168.3	AI643492	RIKEN cDNA 3110054G10 gene
95944_at	P	34.8	P	73.6	AV299153	RIKEN cDNA 2810407E23 gene
110743_at	P	195.5	P	413.3	AW121291	RIKEN cDNA E130203B14 gene
98610_at	A	23.6	P	49.4	AW047926	mitochondrial ribosomal protein S28
93543_f_at	P	1627	P	3387.3	J03952	glutathione S-transferase, mu 1
101109_at	P	251.4	P	522.9	U43512	dystroglycan 1
92927_at	A	50.2	P	104.1	L10426	ets variant gene 1
AFFX-18SR	P	1489	P	3083.3	X00686	Mouse gene for 18S rRNA.
106557_at	P	206.6	P	427.7	AI132668	RIKEN cDNA D130038B21 gene
116614_at	A	189	P	387.1	AA647405	RIKEN cDNA 1700037B15 gene
100912_at	A	19	P	38.9	AI182009	RIKEN cDNA 2410012M04 gene
115043_at	P	471.1	P	961.5	AI842996	ATPase, Na+/K+ transporting, beta 2 polypeptide
101531_at	A	58.2	P	118	AI527354	aldolase 2, B isoform
96696_at	P	144.7	P	293.1	AI837110	UI-M-AK0-adc-e-02-0-UI.s1 NIH_BMAP_MHY Mus musculus cDNA clone UI-M-AK0-adc-e-02-0-UI

92768_s_at	A	105.7	P	214	M15268	aminolevulinic acid synthase 2, erythroid
116008_at	P	152.3	P	305.3	AA673849	ESTs
98915_at	P	254	P	126.8	AI849082	RIKEN cDNA 1600023E10 gene
93729_at	P	116.8	A	58.2	Y15197	microtubule-associated protein 7
113914_at	P	288	M	143	AW121680	sulfatase 1
103845_at	P	233.2	P	115.1	AI839005	solute carrier family 31, member 1
109402_at	P	431.8	M	213.1	AI020360	RIKEN cDNA 2310028N02 gene
97924_at	P	120.8	A	59	AJ132236	UDP-N-acetylglucosamine-2-epimerase/N-acetylmannosamine kinase
106580_at	P	184	A	89.7	AI834858	melanoma antigen, family E, 1
114991_at	P	236.7	A	115	AI593437	Mus musculus mRNA for mszf55-2, partial cds.
103352_at	P	123.7	A	60	X65603	dolichyl-phosphate (UDP-N-acetylglucosamine) acetylglucosaminephosphotransferase 1
98154_at	P	98.4	A	47.7	AW050133	RIKEN cDNA 1300004C11 gene
93240_f_at	P	208.5	P	101	AI838094	RIKEN cDNA 5730594O13 gene
114271_at	P	126.7	A	61	AA794755	SEC22 vesicle trafficking protein-like 2 (S. cerevisiae)
116816_at	P	355	P	169.2	AI874742	ESTs
95620_at	P	367.8	P	174.8	AW120882	RIKEN cDNA 2310016E22 gene
100464_at	P	115.4	M	54.8	AI840585	RIKEN cDNA 3110043O21 gene
109057_at	P	182.2	A	86.3	AA237109	RIKEN cDNA 2310058O09 gene
97844_at	P	269.8	P	126.1	U67187	regulator of G-protein signaling 2
93503_at	P	249.6	P	116.5	U88567	secreted frizzled-related sequence protein 2
94379_at	P	57.1	A	26.6	D17577	kinesin family member 1B
113155_at	P	299.9	A	139.2	AI854828	RIKEN cDNA 3222401G21 gene
95471_at	P	4402	P	2036.7	U22399	cyclin-dependent kinase inhibitor 1C (P57)
115993_at	P	247.1	A	112.5	AI604578	vm27b08.y1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA clone IMAGE:991383 5'
106290_at	P	246.3	A	112.1	AI839436	UI-M-AN0-aco-f-08-0-UI.s1 NIH_BMAP_MBG Mus musculus cDNA clone UI-M-AN0-aco-f-08-0-UI
97951_s_at	P	135.4	A	61.5	U37775	
105552_at	P	194.2	A	87.8	AI464754	ESTs
112880_at	P	288.1	A	130	AA144420	matrix metalloproteinase 23
116085_at	P	1221	P	546.5	AI225392	sideroflexin 3
112947_at	P	189.3	A	84.6	AA693298	RIKEN cDNA 2310022A10 gene
114748_at	P	461.5	P	206	AI853387	death associated protein kinase 1
109963_f_at	P	490.2	A	216.3	AI852122	solute carrier family 25 (mitochondrial deoxynucleotide carrier), member 19
98594_at	P	152.6	P	66.9	AW125453	UI-M-BH2.2-aqm-a-02-0-UI.s1 NIH_BMAP_M_S3.2 Mus musculus cDNA clone
94356_at	P	113.9	A	49.9	AI593074	transformation related protein 53 binding protein 1
97105_at	P	111.7	A	48.8	AI642677	Mus musculus adult male hypothalamus cDNA, RIKEN full-length enriched library, clone
99847_at	P	88.3	A	38.1	X73523	sialyltransferase 4A (beta-galactosidase alpha-2,3-sialytransferase)

116872_at	P	411.4	A	177.5	AI845668	expressed sequence AI845668
95033_at	P	384.6	P	164.1	AW049513	RIKEN cDNA C230043E16 gene
98802_at	P	77.2	A	32.9	D30782	epiregulin
101079_at	P	129.1	A	55	AF093140	nuclear RNA export factor 1 homolog (S. cerevisiae)
93896_at	P	104.4	A	44.4	D13903	protein tyrosine phosphatase, receptor type, D
102771_at	P	145.5	A	61.8	AF091628	SET domain, bifurcated 1
106583_at	P	997.9	P	420.8	AI851530	pleckstrin homology domain-containing, family A (phosphoinositide binding specific) member 2
101357_at	P	68.1	A	28.5	X14971	adaptor protein complex AP-2, alpha 1 subunit
105162_at	P	184.7	A	77.2	AW122704	RIKEN cDNA A830080H07 gene
100400_at	P	40.8	A	17	AW124329	RIKEN cDNA 4921531G14 gene
114322_at	P	177.5	A	73.6	AI843036	RIKEN cDNA C330021A05 gene
94390_at	P	83.3	A	34.2	AI646981	A kinase (PRKA) anchor protein 8
109554_at	P	248.1	A	101.8	AA611111	Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched library, clone
113293_at	P	246.6	A	100.6	AW124722	Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732435K03
97958_at	P	83.4	A	34	AW122355	RIKEN cDNA 2010005I16 gene
92581_at	P	158.7	P	64.5	U07159	acetyl-Coenzyme A dehydrogenase, medium chain
106493_at	P	164.4	A	66.8	AI956751	ESTs
92653_at	P	67	A	27.2	AI482432	RIKEN cDNA D530037H12 gene
97386_at	P	345.1	P	139.8	AI853294	RIKEN cDNA 1110032O19 gene
107536_at	P	296.6	A	120.1	AI851964	RIKEN cDNA 3110052D19 gene
112797_at	P	233	A	92.8	AI838560	inhibitor of growth family, member 3
98922_at	P	95.7	M	37.8	L34260	intergral membrane protein 1
102859_at	P	475.1	P	187.6	AW121304	DNA segment, Chr 6, ERATO Doi 253, expressed
101526_at	P	186.6	P	73.1	X14759	homeo box, msh-like 1
97154_f_at	P	121.7	P	47.5	AA413015	RIKEN cDNA C030034I22 gene
110140_f_at	P	581.7	A	226.5	AW060592	pantothenate kinase 1
96605_at	P	170.4	A	66.3	AI787183	RIKEN cDNA 0610011I04 gene
96580_at	P	99.5	A	38.4	AF020199	pre B-cell leukemia transcription factor 3
114248_at	P	116.5	A	44.1	AI505771	RIKEN cDNA A730004F22 gene
113443_r_at	P	95.3	A	36	AI429737	RIKEN cDNA 4933426L22 gene
116688_at	P	277.5	A	104.6	AW049492	hypothetical protein LOC269878
112920_at	P	657	M	244.8	AW125065	UI-M-BH2.1-apx-d-02-0-UI.s1 NIH_BMAP_M_S3.1 Mus musculus cDNA clone
113940_at	P	271.3	A	100.8	AA754816	RIKEN cDNA 1110019O13 gene
100902_at	P	159.1	A	59.1	AI846549	RIKEN cDNA 2610019F03 gene
92233_at	P	171.2	P	63.4	AI098965	RIKEN cDNA 1810007M14 gene
110344_at	P	336.6	A	123.1	AW123792	RIKEN cDNA 2210404D11 gene

92502_at	P	68.3	A	24.9	X95504	pleiomorphic adenoma gene-like 1
106043_g_a	P	251.2	A	91.5	AI317347	RIKEN cDNA D630014A15 gene
116019_at	P	343.8	P	124.9	AA646779	sulfatase 1
92618_at	P	62	A	22.5	AW125253	small EDRK-rich factor 2
105644_at	P	153.4	A	55.6	AI852681	ESTs
108569_at	P	326.8	A	118.4	AI851400	RIKEN cDNA 1110014J05 gene
101923_at	P	67.9	A	24.6	U34277	phospholipase A2, group VII (platelet-activating factor acetylhydrolase, plasma)
114025_at	P	200.3	A	72.4	AI643935	runt related transcription factor 1
93145_at	P	91.6	A	33.1	AI840824	ESTs
106153_g_a	P	602.1	A	217.5	AI845991	RIKEN cDNA 2610019M19 gene
115084_at	P	384.6	A	138.7	AA863989	ESTs
100571_at	P	1250	P	450.3	AW123934	UI-M-BH2.3-aqc-b-11-0-UI.s1 NIH_BMAP_M_S3.3 Mus musculus cDNA clone
95135_at	P	250.3	A	89.7	AI844396	RIKEN cDNA 3110038L01 gene
103729_at	P	174.3	A	62.4	M36775	laminin, alpha 1
109989_at	P	277.5	A	96.7	AW124029	ESTs
93887_at	P	133	P	46.2	AI854351	multiple PDZ domain protein
101139_r_at	P	71.2	A	24.7	U37531	salivary apomucin; Mus musculus mucin apoprotein mRNA, complete cds.
114842_g_a	P	252.9	A	87	AA855356	neurofibromatosis 1
102215_at	P	119	A	40.9	Y10495	carnitine deficiency-associated gene expressed in ventricle 1
102925_at	P	155.4	P	53.3	AA285446	dual specificity phosphatase 9
111057_at	P	149.1	A	51.1	AI323892	RIKEN cDNA 5033421K01 gene
113797_at	P	133.8	A	45.5	AA968101	RIKEN cDNA 2810407B07 gene
92974_at	P	71.7	A	24	X52533	zinc finger protein 37
102368_at	P	82.1	A	27.4	AW121186	RIKEN cDNA 2610208E05 gene
93090_at	P	755.7	P	250.5	M23362	fibroblast growth factor receptor 2
112828_at	P	1505	P	497.2	AW121214	podocalyxin-like
98627_at	P	276.7	P	90.6	X81580	insulin-like growth factor binding protein 2
102093_f_at	P	112.7	A	36.7	AI461979	splicing factor, arginine/serine-rich 4 (SRp75)
103700_r_at	P	79.9	A	25.8	AI646638	mt86b06.y1 Soares mouse lymph node NbMLN Mus musculus cDNA clone IMAGE:636755 5'
114658_at	P	616.9	A	198.3	AW123164	cingulin
99034_at	P	193.5	A	62.1	Y15001	Iroquois related homeobox 3 (Drosophila)
93574_at	P	68.8	A	21.3	AF036164	serine (or cysteine) proteinase inhibitor, clade F), member 1
101441_i_at	P	289.5	P	89	AF031127	Mus musculus inositol trisphosphate receptor type 2 (Itpr2) gene, partial cds.
102642_at	P	100	A	30.1	L33415	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2
92399_at	P	106.8	A	31.9	D26532	runt related transcription factor 1
104548_at	P	183.7	A	54.3	Y15443	tumor-suppressing subchromosomal transferable fragment 3

96596_at	P	103.9	A	30.5	U52073	N-myc downstream regulated-like
107019_at	P	262.2	A	75.7	AW050053	Mus musculus adult male olfactory brain cDNA, RIKEN full-length enriched library, clone
116983_at	P	865.5	P	244	AW123760	RIKEN cDNA C230098O21 gene
94056_at	P	138.3	A	38.5	M21285	stearoyl-CoA desaturase; Mouse stearoyl-CoA desaturase gene, exon 6.
114766_at	P	152.8	A	42.4	AI853837	UI-M-BH0-ajr-c-06-0-UI.s1 NIH_BMAP_M_S1 Mus musculus cDNA clone UI-M-BH0-ajr-c-06-0-UI 3'
97994_at	P	528.5	P	139.8	AI019193	transcription factor 7, T-cell specific
92314_at	P	104.1	A	26.8	U25652	procollagen, type XII, alpha 1
100483_at	P	495.5	P	126.7	M89641	interferon (alpha and beta) receptor 1
111398_at	P	1432	A	365.1	AI605905	lymphoid enhancer binding factor 1
95537_at	P	78.2	A	19.6	AB019577	Unc-51 like kinase 2 (C. elegans)
109935_at	P	187.1	A	46.3	AA710824	RIKEN cDNA 4930544L10 gene
114207_at	P	235.4	A	56.7	AW125250	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:C130025B04
112889_at	P	145	A	34.8	AI852697	hypothetical protein, MNCb-4193
110428_at	P	184.7	A	43.7	AA797538	EST AA792894
111229_at	P	172.2	A	38.1	AW123416	RIKEN cDNA 5033402L14 gene
107144_at	P	254.4	A	56.1	AA271544	RIKEN cDNA 2410012A13 gene
116067_at	P	180.6	A	39.8	AI847821	RIKEN cDNA 9030607L17 gene
102734_at	P	121.7	A	26.8	U88909	baculoviral IAP repeat-containing 3
92256_at	P	105.2	A	22.6	AI853714	cathepsin B
97317_at	P	113.9	A	23.9	AW122933	ectonucleotide pyrophosphatase/phosphodiesterase 2
110194_at	P	142	A	29.3	AW060408	ESTs
102058_at	P	107.2	A	21.2	AI845667	mitochondrial ribosomal protein L9
98853_at	P	86.1	A	16.3	D30779	phospholipase A2, group IB, pancreas, receptor
106812_at	P	226.1	A	42.3	AI839608	UI-M-AN0-acr-g-06-0-UI.s1 NIH_BMAP_MBG Mus musculus cDNA clone
102168_at	P	61	A	10.9	AL078630	Mouse DNA sequence from clone CT7-BM573K1 on chromosome 17, complete sequence.
112320_at	P	244	A	41.2	AI852394	RIKEN cDNA B930096L08 gene
115001_at	P	144.2	A	21.9	AW121891	cDNA sequence BC032200
111726_at	P	281.9	A	41.8	AA797617	pecanex homolog (Drosophila)
104451_at	P	112.4	A	16.5	AI852578	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2
92696_at	P	74.1	A	10.6	U09563	nuclear receptor subfamily 6, group A, member 1
110983_at	P	171.8	A	20	AA789558	RIKEN cDNA 2610318I15 gene
112414_at	P	164.3	A	19.1	AA673475	pantothenate kinase 3
108279_at	P	242	A	25.2	AI646532	RIKEN cDNA E130014J05 gene
113517_at	P	1206	P	123.9	AI197468	ESTs
111046_r_at	P	149.6	A	15.2	AI957367	ul86f05.x1 Sugano mouse kidney mkia Mus musculus cDNA clone IMAGE:2158977 3'
98817_at	P	284.6	A	27.8	Z29532	follistatin

110248_at	P	193.2	A	15.4	AI020111	cell division cycle 37 homolog (S. cerevisiae)-like
103462_at	P	86.3	A	6.5	AW122239	dedicator of cyto-kinesis 2
100499_at	P	87.3	A	6.3	D29797	syntaxin 3
104965_at	P	134.8	A	9.2	AI448604	RIKEN cDNA 2610016C23 gene
94881_at	P	130.7	A	8.3	AW048937	cyclin-dependent kinase inhibitor 1A (P21)
111425_at	P	142.8	A	6.6	AA608224	ets variant gene 3

Appendix 7 List of Potential Sex Determination Candidates and Their Homologous Human Genes

List of potential male and female candidates from Table 3.2. Their human equivalent along with the chromosome location was obtained using Ensembl EnsMart

AFFY MG U74A	Description	Chromosome Name	Ensembl Gene ID	Human Chromosome	Human Ensembl Gene ID
103009_at	HIRA PROTEIN (TUP1 LIKE ENHANCER OF SPLIT PROTEIN 1). [Source:SWISSPROT;Acc:Q61666]	16	ENSMUSG00000022702.2	22	ENSG00000100084.3
95134_at		X	ENSMUSG00000008035.1	X	ENSG00000165175.4
95135_at		X	ENSMUSG00000008035.1	X	ENSG00000165175.4
102925_at	DUAL SPECIFICITY PHOSPHATASE 9; DUAL-SPECIFICITY MAP KINASE PHOSPHATASE 4. [Source:RefSeq;Acc:NM_029352]	X	ENSMUSG00000031383.1	X	ENSG00000130829.4
92768_s_at	5-AMINOLEVULINIC ACID SYNTHASE, ERYTHROID-SPECIFIC, MITOCHONDRIAL PRECURSOR (EC 2.3.1.37) (DELTA-AMINOLEVULINATE SYNTHASE) (DELTA-ALA SYNTHETASE) (ALAS-E). [Source:SWISSPROT;Acc:P08680]	X	ENSMUSG00000025270.2	X	ENSG00000158578.4
92768_s_at	5-AMINOLEVULINIC ACID SYNTHASE, ERYTHROID-SPECIFIC, MITOCHONDRIAL PRECURSOR (EC 2.3.1.37) (DELTA-AMINOLEVULINATE SYNTHASE) (DELTA-ALA SYNTHETASE) (ALAS-E). [Source:SWISSPROT;Acc:P08680]	X	ENSMUSG00000025270.2	X	ENSG00000158578.4
92927_at		13	ENSMUSG00000047643.1	7	ENSG00000006468.1
92927_at	ER81 PROTEIN (ETS TRANSLOCATION	12	ENSMUSG00000004151.2	7	ENSG00000006468.1

	VARIANT 1). [Source:SWISSPROT;Acc:P41164]				
92927_at	ER81 PROTEIN (ETS TRANSLOCATION VARIANT 1). [Source:SWISSPROT;Acc:P41164]	12	ENSMUSG00000004151.2	7	ENSG00000006468.1
92927_at	ER81 PROTEIN (ETS TRANSLOCATION VARIANT 1). [Source:SWISSPROT;Acc:P41164]	12	ENSMUSG00000004151.2	7	ENSG00000006468.1
97309_at	HSC70-INTERACTING PROTEIN (HIP) (PUTATIVE TUMOR SUPPRESSOR ST13). [Source:SWISSPROT;Acc:Q99L47]	15	ENSMUSG00000022403.1	22	ENSG00000100380.3
97310_at	HSC70-INTERACTING PROTEIN (HIP) (PUTATIVE TUMOR SUPPRESSOR ST13). [Source:SWISSPROT;Acc:Q99L47]	15	ENSMUSG00000022403.1	22	ENSG00000100380.3
94207_at	PROTEIN DISULFIDE ISOMERASE A6 PRECURSOR. [Source:SPTREMBL;Acc:Q8BK54]	12	ENSMUSG00000020571.2	2	ENSG00000143870.1
94208_at	PROTEIN DISULFIDE ISOMERASE A6 PRECURSOR. [Source:SPTREMBL;Acc:Q8BK54]	12	ENSMUSG00000020571.2	2	ENSG00000143870.1
94209_g_at	PROTEIN DISULFIDE ISOMERASE A6 PRECURSOR. [Source:SPTREMBL;Acc:Q8BK54]	12	ENSMUSG00000020571.2	2	ENSG00000143870.1
103451_at	PROTEIN TYROSINE KINASE 2 BETA (EC 2.7.1.112) (FOCAL ADHESION KINASE 2) (FADK 2) (PROLINE-RICH TYROSINE KINASE 2) (CELL ADHESION KINASE BETA) (CAK BETA) (CALCIUM-DEPENDENT TYROSINE KINASE) (CADTK) (RELATED ADHESION FOCAL TYROSINE KINASE). [Source:SWISSPRO	14	ENSMUSG00000022042.2	8	ENSG00000120899.3
97482_at	MO25-LIKE PROTEIN. [Source:SWISSPROT;Acc:Q9DB16]	14	ENSMUSG00000021981.2	13	ENSG00000102547.4
96221_at		10	ENSMUSG00000019842.2	6	ENSG00000056972.3

98454_at	PARALEMMIN. [Source:SWISSPROT;Acc:Q9Z0P4]	10	ENSMUSG00000035863.2	19	ENSG00000099864.2
93294_at	CONNECTIVE TISSUE GROWTH FACTOR PRECURSOR (FISP-12 PROTEIN) (HYPERTROPHIC CHONDROCYTE-SPECIFIC PROTEIN 24). [Source:SWISSPROT;Acc:P29268]	10	ENSMUSG00000019997.1	6	ENSG00000118523.1
98131_at	QUINONE OXIDOREDUCTASE (EC 1.6.5.5) (NADPH:QUINONE REDUCTASE) (ZETA- CRYSTALLIN). [Source:SWISSPROT;Acc:P47199]	3	ENSMUSG00000028199.1	1	ENSG00000116791.1
99701_f_at	SMALL PROLINE-RICH PROTEIN 2B. [Source:RefSeq;Acc:NM_011469]	3	ENSMUSG00000050092.2	1	ENSG00000169457.1
102058_at	60S RIBOSOMAL PROTEIN L9, MITOCHONDRIAL PRECURSOR (L9MT) (FRAGMENT). [Source:SWISSPROT;Acc:Q99N94]	3	ENSMUSG00000028140.2	1	ENSG00000143436.1
94881_at	CYCLIN-DEPENDENT KINASE INHIBITOR 1 (P21) (CDK-INTERACTING PROTEIN 1) (MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN). [Source:SWISSPROT;Acc:P39689]	17	ENSMUSG00000023067.1	6	ENSG00000124762.1
98067_at	CYCLIN-DEPENDENT KINASE INHIBITOR 1 (P21) (CDK-INTERACTING PROTEIN 1) (MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN). [Source:SWISSPROT;Acc:P39689]	17	ENSMUSG00000023067.1	6	ENSG00000124762.1
96605_at		6	ENSMUSG00000023367.2	7	ENSG00000002933.1
97446_at	HELG. [Source:RefSeq;Acc:NM_133347]	9	ENSMUSG00000032480.2	3	ENSG00000132153.2
97446_at	HELG. [Source:RefSeq;Acc:NM_133347]	9	ENSMUSG00000032480.2	3	ENSG00000132153.2
97446_at	HELG. [Source:RefSeq;Acc:NM_133347]	9	ENSMUSG00000032480.2	3	ENSG00000132153.2
102215_at	CARNITINE DEFICIENCY-ASSOCIATED	5	ENSMUSG00000029469.2	12	ENSG00000122970.3

	PROTEIN EXPRESSED IN VENTRICLE 1 (CDV-1 PROTEIN). [Source:SWISSPROT;Acc:O35594]				
104548_at	PLECKSTRIN HOMOLOGY-LIKE DOMAIN, FAMILY A, MEMBER 2; TUMOR-SUPPRESSING SUBCHROMOSOMAL TRANSFERABLE FRAGMENT 3. [Source:RefSeq;Acc:NM_009434]	7	ENSMUSG00000010760.2	11	ENSG00000181649.2
98125_at	BBP-LIKE PROTEIN 2. [Source:RefSeq;Acc:NM_026795]	7	ENSMUSG00000030515.2	15	ENSG00000185418.3
98125_at	BBP-LIKE PROTEIN 2. [Source:RefSeq;Acc:NM_026795]	7	ENSMUSG00000030515.2	15	ENSG00000185418.3
92987_at	ANION EXCHANGE PROTEIN 3 (NEURONAL BAND 3-LIKE PROTEIN). [Source:SWISSPROT;Acc:P16283]	1	ENSMUSG00000006576.2	2	ENSG00000114923.2
92987_at	ANION EXCHANGE PROTEIN 3 (NEURONAL BAND 3-LIKE PROTEIN). [Source:SWISSPROT;Acc:P16283]	1	ENSMUSG00000006576.2	2	ENSG00000114923.2
92987_at	ANION EXCHANGE PROTEIN 3 (NEURONAL BAND 3-LIKE PROTEIN). [Source:SWISSPROT;Acc:P16283]	1	ENSMUSG00000006576.2	2	ENSG00000114923.2
100403_at	MYOSIN LIGHT CHAIN, REGULATORY A. [Source:RefSeq;Acc:NM_022879]	11	ENSMUSG00000020469.1	7	ENSG00000106631.1
98109_at	MITOCHONDRIAL RIBOSOMAL PROTEIN L55. [Source:RefSeq;Acc:NM_026035]	11	ENSMUSG00000036860.1	1	ENSG00000162910.4
101992_at	PROTEASOME SUBUNIT BETA TYPE 6 PRECURSOR (EC 3.4.25.1) (PROTEASOME DELTA CHAIN) (MACROPAIN DELTA CHAIN) (MULTICATALYTIC ENDOPEPTIDASE COMPLEX DELTA CHAIN) (PROTEASOME SUBUNIT Y). [Source:SWISSPROT;Acc:Q60692]	11	ENSMUSG00000018286.1	17	ENSG00000142507.1
93327_at		11	ENSMUSG00000040158.1	17	ENSG00000108392.1

103293_at	28 KDA GOLGI SNARE PROTEIN (GOLGI SNAP RECEPTOR COMPLEX MEMBER 1) (28 KDA CIS-GOLGI SNARE P28) (GOS-28). [Source:SWISSPROT;Acc:O88630]	11	ENSMUSG00000010392.1	17	ENSG00000108587.2
95621_at		11	ENSMUSG00000035775.1	17	ENSG00000171431.1
104432_at	RHO-RELATED GTP-BINDING PROTEIN RHON (RHO7). [Source:SWISSPROT;Acc:Q9QYM5]	11	ENSMUSG00000001313.1	17	ENSG00000108830.1
93187_at		11	ENSMUSG00000020773.2	17	ENSG00000132481.1
104023_at	IMMUNOGLOBULIN SUPERFAMILY, MEMBER 7; DENDRITIC CELL-DERIVED IMMUNOGLOBULIN RECEPTOR 1; LEUKOCYTE MONO-IG-LIKE RECEPTOR 2. [Source:RefSeq;Acc:NM_134158]	11	ENSMUSG00000044811.2	17	ENSG00000167850.3
100051_at	ERYTHROCYTE BAND 7 INTEGRAL MEMBRANE PROTEIN (STOMATIN) (PROTEIN 7.2B). [Source:SWISSPROT;Acc:P54116]	2	ENSMUSG00000026880.2	9	ENSG00000148175.1
96336_at	GLYCINE AMIDINOTRANSFERASE, MITOCHONDRIAL PRECURSOR (EC 2.1.4.1) (L- ARGININE:GLYCINE AMIDINOTRANSFERASE) (TRANSAMIDINASE) (AT). [Source:SWISSPROT;Acc:Q9D964]	2	ENSMUSG00000027199.1	15	ENSG00000171766.2
93327_at		2	ENSMUSG00000050752.1	17	ENSG00000108392.1
96658_at		2	ENSMUSG00000044627.1	9	ENSG00000175854.1
96771_AT (Not found)					
104637_AT (Not found)					
103222_AT (Not found)					
93428_AT (Not found)					
96357_AT (Not found)					
92256_AT (Not found)					
103891_I_AT (Not found)					

104648_AT (Not found) 103842_AT (Not found)					

AFFY MG U74B	Description	Chromosome Name	Ensembl Gene ID	Human Chromosome	Human Ensembl Gene ID
114963_at		19	ENSMUSG00000023350.1	11	ENSG00000084207.3
111722_at		15	ENSMUSG00000036649.2	8	ENSG00000022567.1
111903_at		15	ENSMUSG00000036649.2	8	ENSG00000022567.1
111722_at		15	ENSMUSG00000036649.2	8	ENSG00000022567.1
111903_at		15	ENSMUSG00000036649.2	8	ENSG00000022567.1
111722_at		15	ENSMUSG00000036649.2	8	ENSG00000022567.1
111903_at		15	ENSMUSG00000036649.2	8	ENSG00000022567.1
107046_at	PDZ-DOMAIN PROTEIN SCRIBBLE; SCRIBBLE. [Source:RefSeq;Acc:NM_134089]	15	ENSMUSG00000022568.2	8	ENSG00000180900.4
107046_at	PDZ-DOMAIN PROTEIN SCRIBBLE; SCRIBBLE. [Source:RefSeq;Acc:NM_134089]	15	ENSMUSG00000022568.2	8	ENSG00000180900.4
107046_at	PDZ-DOMAIN PROTEIN SCRIBBLE; SCRIBBLE. [Source:RefSeq;Acc:NM_134089]	15	ENSMUSG00000022568.2	8	ENSG00000180900.4
106290_at	AMPHOTERIN INDUCED GENE 2; ALIVIN 1. [Source:RefSeq;Acc:NM_178114]	15	ENSMUSG00000048218.2	12	ENSG00000139211.2
110713_at		12	ENSMUSG00000019718.1	14	ENSG00000126790.2
110715_at	PHOSPHOLIPASE D1 (EC 3.1.4.4) (PLD 1) (CHOLINE PHOSPHATASE 1) (PHOSPHATIDYLCHOLINE-HYDROLYZING PHOSPHOLIPASE D1) (MPLD1). [Source:SWISSPROT;Acc:Q9Z280]	3	ENSMUSG00000027695.2	3	ENSG00000075651.2
110715_at	PHOSPHOLIPASE D1 (EC 3.1.4.4) (PLD 1) (CHOLINE PHOSPHATASE 1) (PHOSPHATIDYLCHOLINE-HYDROLYZING PHOSPHOLIPASE D1) (MPLD1).	3	ENSMUSG00000027695.2	3	ENSG00000075651.2

	[Source:SWISSPROT;Acc:Q9Z280]				
110715_at	PHOSPHOLIPASE D1 (EC 3.1.4.4) (PLD 1) (CHOLINE PHOSPHATASE 1) (PHOSPHATIDYLCHOLINE-HYDROLYZING PHOSPHOLIPASE D1) (MPLD1).	3	ENSMUSG00000027695.2	3	ENSG00000075651.2
	[Source:SWISSPROT;Acc:Q9Z280]				
108008_at	1-ACYLGLYCEROL-3-PHOSPHATE O-ACYLTRANSFERASE 1 (LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE, DELTA).	17	ENSMUSG00000023827.1	6	ENSG00000026652.2
	[Source:RefSeq;Acc:NM_026644]				
111398_at	LYMPHOID ENHANCER BINDING FACTOR 1 (LEF-1).	3	ENSMUSG00000027985.2	4	ENSG00000138795.1
	[Source:SWISSPROT;Acc:P27782]				
111398_at	LYMPHOID ENHANCER BINDING FACTOR 1 (LEF-1).	3	ENSMUSG00000027985.2	4	ENSG00000138795.1
	[Source:SWISSPROT;Acc:P27782]				
106949_at		6	ENSMUSG00000037172.2	7	ENSG00000127359.1
113425_at	PROTEIN PHOSPHATASE 1, REGULATORY (INHIBITOR) SUBUNIT 9A; NEURABIN; NEURAL TISSUE-SPECIFIC F-ACTIN BINDING PROTEIN.	6	ENSMUSG00000032827.2	7	ENSG00000158528.1
	[Source:RefSeq;Acc:NM_181595]				
109172_at	PUTATIVE DEOXYRIBOSE-PHOSPHATE ALDOLASE (EC 4.1.2.4) (PHOSPHODEOXYRIBOALDOLASE) (DEOXYRIBOALDOLASE) (DERA).	6	ENSMUSG00000030225.2	12	ENSG00000023697.2
	[Source:SWISSPROT;Acc:Q91YP3]				
113066_at	HOMEOBOX PROTEIN HOX-A7 (HOX-1.1) (M6-12) (M6).	6	ENSMUSG00000038236.1	7	ENSG00000122592.1
	[Source:SWISSPROT;Acc:P02830]				
113089_at	CALCIUM CHANNEL, VOLTAGE-DEPENDENT, ALPHA 2/DELTA SUBUNIT 2; ALPHA 2 DELTA CALCIUM CHANNEL SUBUNIT; DUCKY.	9	ENSMUSG00000010066.2	3	ENSG00000007402.2

	[Source:RefSeq;Acc:NM_020263]				
115354_at	ACID SPHINGOMYELINASE-LIKE PHOSPHODIESTERASE 3B PRECURSOR (EC 3.1.4.-) (ASM-LIKE PHOSPHODIESTERASE 3B). [Source:SWISSPROT;Acc:P58242]	4	ENSMUSG00000028885.1	1	ENSG00000130768.1
107566_at		7	ENSMUSG00000010307.1	11	ENSG00000151117.1
113914_at	EXTRACELLULAR SULFATASE SULF-1 PRECURSOR (EC 3.1.6.-) (MSULF-1). [Source:SWISSPROT;Acc:Q8K007]	1	ENSMUSG00000016918.2	8	ENSG00000137573.3
116019_at	EXTRACELLULAR SULFATASE SULF-1 PRECURSOR (EC 3.1.6.-) (MSULF-1). [Source:SWISSPROT;Acc:Q8K007]	1	ENSMUSG00000016918.2	8	ENSG00000137573.3
113914_at	EXTRACELLULAR SULFATASE SULF-1 PRECURSOR (EC 3.1.6.-) (MSULF-1). [Source:SWISSPROT;Acc:Q8K007]	1	ENSMUSG00000016918.2	8	ENSG00000137573.3
116019_at	EXTRACELLULAR SULFATASE SULF-1 PRECURSOR (EC 3.1.6.-) (MSULF-1). [Source:SWISSPROT;Acc:Q8K007]	1	ENSMUSG00000016918.2	8	ENSG00000137573.3
113914_at	EXTRACELLULAR SULFATASE SULF-1 PRECURSOR (EC 3.1.6.-) (MSULF-1). [Source:SWISSPROT;Acc:Q8K007]	1	ENSMUSG00000016918.2	8	ENSG00000137573.3
116019_at	EXTRACELLULAR SULFATASE SULF-1 PRECURSOR (EC 3.1.6.-) (MSULF-1). [Source:SWISSPROT;Acc:Q8K007]	1	ENSMUSG00000016918.2	8	ENSG00000137573.3
111158_at		11	ENSMUSG00000020472.2	1	ENSG00000162714.2
115096_at		11	ENSMUSG00000020472.2	1	ENSG00000162714.2
114058_at		11	ENSMUSG00000051378.2	17	ENSG00000186185.1
114058_at		11	ENSMUSG00000051378.2	17	ENSG00000186185.1
113012_at	GLYCEROL-3-PHOSPHATE DEHYDROGENASE, MITOCHONDRIAL PRECURSOR (EC 1.1.99.5) (GPD-M) (GPDH-M). [Source:SWISSPROT;Acc:Q64521]	2	ENSMUSG00000026827.1	2	ENSG00000115159.2

114138_at	GLYCEROL-3-PHOSPHATE DEHYDROGENASE, MITOCHONDRIAL PRECURSOR (EC 1.1.99.5) (GPD-M) (GPDH-M). [Source:SWISSPROT;Acc:Q64521]	2	ENSMUSG00000026827.1	2	ENSG00000115159.2
114139_at	GLYCEROL-3-PHOSPHATE DEHYDROGENASE, MITOCHONDRIAL PRECURSOR (EC 1.1.99.5) (GPD-M) (GPDH-M). [Source:SWISSPROT;Acc:Q64521]	2	ENSMUSG00000026827.1	2	ENSG00000115159.2
117316_at	GLYCEROL-3-PHOSPHATE DEHYDROGENASE, MITOCHONDRIAL PRECURSOR (EC 1.1.99.5) (GPD-M) (GPDH-M). [Source:SWISSPROT;Acc:Q64521]	2	ENSMUSG00000026827.1	2	ENSG00000115159.2
114784_at	GROWTH FACTOR RECEPTOR-BOUND PROTEIN 14 (GRB14 ADAPTER PROTEIN). [Source:SWISSPROT;Acc:Q9JLM9]	2	ENSMUSG00000026888.1	2	ENSG00000115290.1
112307_at		2	ENSMUSG00000040007.2	15	ENSG00000140320.1
108745_at		2	ENSMUSG00000050530.2	10	ENSG00000148468.2
108745_at		2	ENSMUSG00000050530.2	10	ENSG00000148468.2
108745_at		2	ENSMUSG00000050530.2	10	ENSG00000148468.2
115255_at		7_random_NT_081257	ENSMUSG00000040914.1	19	ENSG00000104983.1
104954_AT (Not found)					
105120_F_AT (Not found)					
112784_AT (Not found)					
110400_F_AT (Not found)					
112900_AT (Not found)					
116952_AT (Not found)					
109296_AT (Not found)					
114658_AT (Not found)					

116162_G_AT (Not found) 105667_AT (Not found) 115947_AT (Not found) 115993_AT (Not found) 114872_AT (Not found)				
105754_AT (Not found) 110129_AT (Not found) 105229_AT (Not found) 109521_AT (Not found) 116600_AT (Not found)				