NORTHWESTERN UNIVERSTIY

Uncovering Chromatin-Mediated Regulatory Mechanisms of Mammalian Sex Determination

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULLFILMENT OF THE
REQUIREMENTS

for the degree

DOCTOR IN PHILOSOPHY

The Field of Biology

by

Sara Alexandra García Moreno Duel

CHICAGO, ILLINOIS

December 2018

ABSTRACT

A fundamental goal in biology is to understand how distinct cell types containing the same genetic information arise from a single stem cell throughout development. Sex determination is a key developmental process that requires a unidirectional commitment of an initially bipotential gonad towards either the male or female fate. This makes sex determination a unique model to study cell fate commitment and differentiation *in vivo.* Based on the accumulating evidence that dynamic chromatin structure regulates cell fate decisions, our overarching goal was to identify epigenetic mechanisms and cis-regulatory elements that contribute to the bipotential state of the fetal gonad and to the establishment and maintenance of the male or female fate during sex determination in mice.

In this study, we developed a quantitative genome-wide profile of the active H3K4me3 and the repressive H3K27me3 histone modifications in isolated XY and XX gonadal supporting cells before and after sex determination. We show that male- and female-promoting genes are bivalent before sex determination, providing insight into how the bipotential state of the gonad is established at the epigenetic level. Surprisingly, after sex determination, many genes that promote the alternate pathway remain bivalent, possibly contributing to the ability of these cells to transdifferentiate even in adults. The finding that bivalency is retained at female-promoting genes whose expression declines after commitment to the male fate led us to question what could be silencing these bivalent genes throughout sex determination. It is known that the Polycomb group of proteins (PcG) mediate repression of bivalent genes in ESCs, and that loss of CBX2*,* the PcG subunit that binds H3K27me3 and mediates silencing, leads to upregulation of the female pathway and ovary development in XY individuals. We found that many genes in the Wnt signaling pathway were targeted for H3K27me3-mediated repression in Sertoli (XY) cells, leading us to test whether deletion of the ovarian-promoting gene *Wnt4* could rescue male

development in *Cbx2* mutants. In accordance with our hypothesis, we show that expression of the male-determining gene *Sry* and testis development were rescued in XY *Cbx2*^{-/-}*;Wnt4*^{-/-} mice. We also show that CBX2 binds the downstream Wnt signaler *Lef1*. Our findings suggest that *Cbx2* induces *Sry* indirectly by repressing the key ovarian-promoting/anti-testis Wnt pathway.

In addition to histone modifications that regulate chromatin accessibility at promoter regions, the precise spatiotemporal regulation of genes during development is orchestrated by cis-regulatory genomic elements (such as enhancers, silencers and insulators) that function at a distance from the transcription start site. However, identifying functional regulatory sites that drive cell differentiation *in vivo* has been complicated by the high numbers of cells required for whole-genome epigenetic assays. Our inability to pinpoint the location of these sites has limited our capacity to study the mechanisms that regulate the expression of sex-determining genes. To address this limitation, we performed an Assay for Transposase-Accessibility followed by next-generation sequencing (ATAC-seq), a low-input assay that identifies regulatory elements, in purified XX and XY gonadal supporting cells before and after sex determination in mice. To distinguish enhancers from silencers and insulators, we also performed Chromatin Immunoprecipitation followed by next-generation sequencing (ChIP-seq) for H3K27ac, a histone modification that marks active enhancers. We show that XX and XY supporting cells initiate sex determination with similar chromatin landscapes in accordance to their bipotential nature, and acquire sex-specific regulatory elements as they commit to the male or female fate.

To validate our approach, we identified two functional gonad-specific enhancers: one downstream of *Bmp2,* an ovary-promoting gene, and one upstream of *Sox9*, a testis-promoting gene. Deletion of the distal enhancer upstream of *Sox9,* named *Enh13,* led to development of phenotypically-normal XY females. This remarkable finding suggests that *Enh13* alone is sufficient to upregulate *Sox9* and is necessary for testis development.

The work presented in this dissertation greatly increases our understanding of the complex regulatory network underlying mammalian sex determination. In addition to providing insight into the role that chromatin dynamics plays in regulating cell fate decisions during a key developmental event *in vivo*, our datasets provide a powerful resource for identifying non-coding regulatory elements that could lead to Disorders of Sexual Development when disrupted.

ACKNOWLEDGEMENTS

I'd first like to thank my parents, my constant cheerleaders, for their never-ending support and unwavering enthusiasm for all my accomplishments, no matter how small. Their genuine interest and excitement for everything I've ever done has kept me motivated and has allowed me to accomplish every goal I have set for myself. Thank you for everything you've done for me.

*

A huge thanks to my sister, especially during these last five years- it's been fun being a student at the same time as you. Thanks for the many Chicago visits and for hosting me in DC, for the phone calls, the races we've run and all the yellow gifts you've sent me to cheer me up, you've made getting to the finish line of my PhD pretty enjoyable.

*

To my brilliant mentor and role model, the late Dr. Danielle Maatouk- in the short time that I was lucky enough to know her, she left an indelible mark in my life and development as a scientist. I will forever be grateful for the time she spent teaching me and inspiring me to reach higher and do better. She is greatly missed.

*

Thanks to the Maatouk Lab, Chris and Bella, what can I say other than gonad or go home? Working with you always made lab fun and I will miss having you guys around, especially now that we've proven that we can publish on our own. Thank you for being my lab family.

*

I am eternally grateful to Dr. Blanche Capel for her generosity, and the Capel lab for being so warm and welcoming. I could not have finished without their support.

I'd like to thank my committee members Dr. Debu Chakravarti, Dr.Jhumku Khotz and Dr. Ann

Harris, and Dr. Teresa Woodruff for their time and guidance throughout this process.

*

*

I'm also grateful for the wonderful people at the core facilities at Northwestern University without

whom this thesis would not exist.

*

I am especially grateful for my "old" friends who have been there from the beginning and

continue to love and support me no matter the distance. It's amazing how little things seem to

have changed when we're together. Special mention to Ana, Berta, Sam, and Paola whom I

miss every day.

*

And for my "new" friends that have made Chicago feel like home and are making it very hard to

leave. Thank you Kelly and Leah, Stephanie and Nico, Tania, Andrew, Matt and everyone else

who has made this PhD and the Chicago winters bearable… and even fun! I'd also like to thank

the Plebanek family for welcoming me into their family and home. I'll miss you all!

*

I'd like to thank all of my Mexican and American extended family, my grandmothers, Gamma

and Coquis, my aunts and uncles, and all of my cousins. Thank you pretending to be interested

in my work despite not understanding it!

I am immensely grateful to all the mice that have given their life to science- this part of my career has never been easy for me, but it could not be done without them. They are not taken for granted.

*

And last, but most certainly not least, thank you Mike, for sharing this crazy journey with me, for the many late-night talks of life and science, for the life we've created together and the future we've come to dream of. Thank you for being my rock in times of need and assuring me that I could (and would!) surpass any obstacle thrown my way. I especially want to thank you for always considering me in your decisions and making sure that we build our future together, not in spite of each other. Thank you for being my friend, my family, my love. I can't wait for the next chapter in our lives. Duke, here we come!

*

To Danielle

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER FOUR

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

1.1. Opening

Sexual reproduction requires the development of dimorphic yet compatible organisms within the same species. Vertebrates have evolved a common template to achieve development of two distinct sexes. Initially male and female embryos are indistinguishable. During development, the embryonic gonad forms with the unique ability to differentiate into two alternative organs: testes (males) or ovaries (females) (Fig. 1.1). Gonadal differentiation diverges based on a genetic or environmental switch that activates one pathway and represses the other. This system presents an interesting model to explore the various levels of regulation involved in fate commitment of single cells and the coordination of the entire cell community that makes up the bipotential gonad.

One interesting characteristic of vertebrate sex determination is the plastic ability to switch between a male and female fate. For example, in most reptilian species, the incubation temperature of the egg controls the sex of the offspring, and switching eggs between temperatures during a critical window of development results in a switch to the opposite sex. Many fish can undergo sex reversal as adults, and although mammals do not undergo full sex reversal, the removal of certain key transcription factors can cause a switch in the identity of female cells to male cells (Ottolenghi et al., 2007, Uhlenhaut et al., 2009) or male cells to female cells (Matson et al., 2011). Importantly, gonadal cells do not switch randomly to another fate, but specifically to their developmental alternative fate.

Although multiple mechanisms are interwoven in the process of fate commitment and canalization as one sex or the other, epigenetic regulation is emerging as an important component. Epigenetics, the study of changes to gene function without changes in the DNA sequence, is capable of imposing a stable differentiation state throughout cell division. Between 1940 and 1956, Conrad Waddington proposed the concept of an epigenetic landscape to

describe the process of cell fate commitment during development. He envisioned that a cell progresses towards a differentiated state through a series of fate decisions that are stabilized by changes to its epigenome. These changes maintain cell-type-specific gene expression patterns that channel the cell along certain pathways while eliminating other possible pathways the cell might take (Waddington, 1940, Waddington, 1942, Waddington, 1956).

Although there are several lines of evidence emerging in the field of vertebrate sex determination that suggest that epigenetic regulation is a key element in the process of commitment to and canalization of the male or female fate, the epigenetic mechanisms underlying sex determination remain greatly unexplored. In this chapter, I will summarize the current literature on the epigenetic and chromatin-mediated regulation of sex determination in mammals and, in some instances, other vertebrates, and introduce the gaps in knowledge that our lab intends to address.
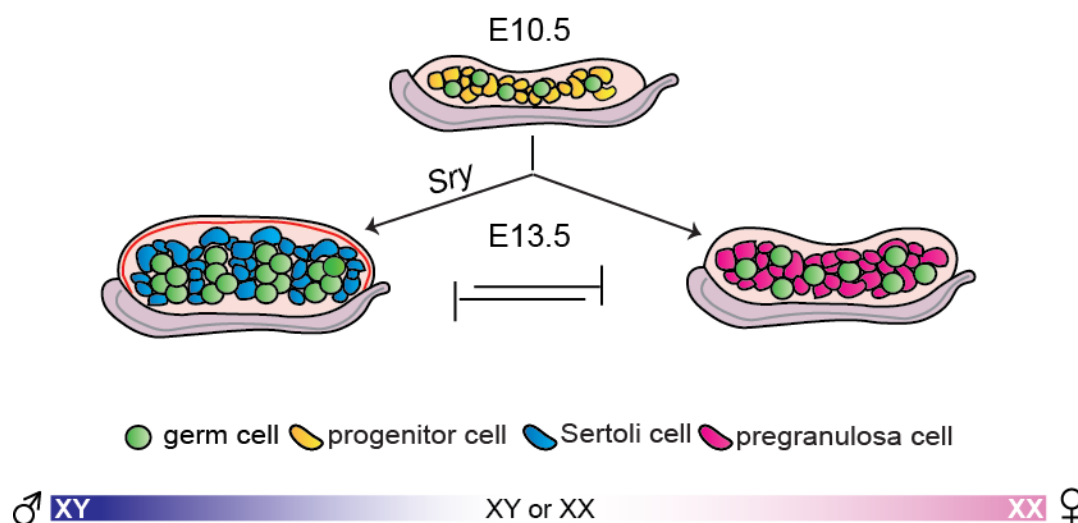
Figure 1.1. Overview of mammalian sex determination. In mice, gonads form at E10.5 and are indistinguishable between XY and XX individuals. The somatic progenitor cells that compose the fetal gonad (top) are initially bipotential. Following expression of the Y-encoded *Sry* gene at E11.5, progenitor cells differentiate into Sertoli cells and gonads develop as testes (left). In absence of *Sry*, progenitor cells differentiate into granulosa cells and gonads develop as ovaries (right). Commitment to the male or female fate requires repression of the alternate pathway.

1.2. Overview of mammalian sex determination

For many years the term "primary sex determination" has been used in vertebrates to refer to the fate decision in the bipotential gonad that leads to the commitment to testis or ovary fate. This is because (1) the first microscopic differences between male and female development occur in the gonad, and (2) once the gonad is committed to testis or ovary fate, the hormonal secretions of the testis or ovary are the dominant influence on male and female secondary sex characteristics. In mammals, sex determination is determined by the genetic sex of an individual, where XY individuals typically develop testes, and XX individuals typically develop ovaries. Below, I will give a brief overview of the genetic mechanisms that regulate the male vs. female fate decision from an initially bipotential system.

1.2.1. Development of the bipotential gonad

Gonadal development initiates with the formation of a bipotential gonadal ridge at around mid-gestation (embryonic day 10 (E10)) in the mouse (Figure 1.1). It forms by the thickening of the coelomic epithelium that covers the middle part of the mesoderm, the mesonephros, on either side of the hindgut. Loss-of-function studies in mice have identified *Gata4* (Page, 2013) as a key factor for initial formation of the ridge, and genes such as *Sf1* (Luo et al., 1994), *Wt1* (Hammes et al., 2001), *Lhx9* (Birk; et al., 2000), *Emx2* (Kusaka et al., 2010)*, Cbx2 (M33)* (Katoh-Fukui et al., 1998) and *Six1/4* (Fujimoto et al., 2013) for its growth and maintenance. The coelomic epithelium gives rise to precursors of the gonadal somatic cell lineages (Karl and Capel, 1998), while germ cells develop extra-gonadally as primordial germ cells (PGCs) (Chiquoine, 1954, Ginsburg et al., 1990) . PGCs proliferate as they migrate down the hindgut endoderm (Tam and Snow, 1981) and colonize the gonads at E10.5. At this early stage in development, the gonad is a unique system in which both PGCs and somatic cells that

compose it are bipotential and have the full ability to differentiate into two opposing organs: the testis or the ovary (Figure 1.1) (McLaren, 1988). The initial decision to follow the male or female pathway occurs within a subset of the somatic cells named the supporting cell lineage due to their role in nurturing and supporting the germ cells throughout development and adult life. If the bipotential supporting lineage engages the male fate, they differentiate into Sertoli cells in the testis; if the engage the female fate, they differentiate into granulosa cells in the ovaries (Figure 1.1) (Albrecht and Eicher, 2001). Sertoli and granulosa cells then signal to germ cells to either enter spermatogenesis or develop as oocytes, respectively. The bipotential gonad is not only morphologically indistinguishable between XX and XY individuals, but transcriptome analysis on whole gonads (Munger et al., 2013) and on purified supporting cells prior to sex determination (Nef et al., 2005) show that XX and XY cells co-express a number of genes that will not become sex-specific until after their male or female fate is specified. The primordial gonad exists in a bipotential state for ~1.5 days, however, the mechanisms that establish and maintain its bipotential nature are not well understood.

1.2.2. Genetic regulation of male fate commitment in mammals

The discovery of the testis-determining gene *Sry* on the Y chromosome was crucial in the field of sex determination as it represented the top of the male-determination cascade of most mammals. In humans, analysis of XX males carrying a translocation of a fragment of the Y chromosome, and of XY females carrying point mutations in the same locus, led to the discovery of the *SRY* gene (Sinclair et al., 1990). Soon after, *Sry,* the mouse homolog of *SRY*, was identified (Gubbay et al., 1990). The sufficiency of this single gene to initiate testis development was proven by Koopman et al. in 1991 (Koopman et al., 1991) by transgenically expressing *Sry* in XX mice, inducing complete female-to-male sex reversal (females developed as normal males). The spatiotemporal expression of *Sry* is very tightly regulated: its expression

is restricted to the supporting cell lineage, and is only transiently activated starting at E10.5, peaking at E11.5 and disappearing by E12.5 (Hacker et al., 1995). Relatively little is known about what regulates *Sry* expression, however *Wt1* (Bradford et al., 2009), *Sf1* [3], *Gata4/Fog2* (Manuylov et al., 2007), *Cbx2* (Katoh-Fukui et al., 2012) and MAP3K4 (Bogani et al., 2009) have all been implicated.

Until now, the sole function of *Sry* seems to be the upregulation of the transcription factor SOX9, which is initially expressed at low levels both in male and female supporting cells (Albrecht and Eicher, 2001). *Sox9* is essential for testis development, as conditional deletion of *Sox9* in XY gonads results in complete sex reversal (Chaboissier et al., 2004). *Sox9* upregulation is achieved by binding of SRY and SF1 at the core element of the testis-specific enhancer (TESCO) approximately 14kb upstream of *Sox9* (Sekido and Lovell-Badge, 2008)*. Sox9* upregulation in itself is not sufficient for testis-determination; rather, its high expression must be maintained after *Sry* expression has been extinguished. This occurs in a cell-autonomous manner through binding of SF1 and SOX9 itself to TESCO*,* and through a non-cell-autonomous manner by secretion of paracrine factors such as PGD2 (Moniot et al., 2009), and FGF9 that acts via its receptor FGFR2 (Kim et al., 2007b). Together, these positive feed-forward loops ensure that *Sox9* reaches a threshold level within the nucleus to direct Sertoli cell differentiation, as well as a threshold of Sertoli cells in the gonad to direct testis development (Schmahl and Capel, 2003).

Importantly, as the male pathway is engaged, the female pathway must be actively repressed. The reverse is also true during ovarian development in females. Although not much is known about how the opposing pathways are repressed, it has become evident that FGF9 and FGFR2 also play an important role during testis development by antagonizing the female pathway through repression of *Wnt4 (Kim et al., 2006b).* This is evidenced by the fact that *Wnt4*

is overexpressed in sex-reversed *Fgf9*[-/-] and *Fgfr2*[-/-] XY mice, and that sex-reversal is rescued in *Fgf9;Wnt4* and *Fgfr2;Wnt4* double mutants (Jameson et al., 2012a). Furthermore, fate commitment must be maintained throughout the adult life, as evidenced by the postnatal transdifferentiation of Sertoli cells into granulosa cells following conditional deletion of *Dmrt1* in the postnatal testis (Matson et al., 2011).

1.2.3. Genetic regulation of female fate commitment in mammals

In the absence of *Sry*, *Sox9* expression is not upregulated and the male-female balance of the bipotential gonad is tipped in favor of the female pathway. Historically, the female pathway had been considered the "default pathway", largely based on experiments performed by Alfred Jost in 1947 in which the removal of rabbit fetal gonads *in utero* led to development of XY rabbits with female external genitalia (Jost, 1952). Though this may be true for sex differentiation, in which hormonal secretions pattern the internal and external reproductive system, it has since become clear that there is a female pathway that must become activated and maintained while simultaneously repressing the male pathway in order for the gonads themselves to develop as ovaries. The search for an ovary-determining gene, an analog to *Sry*, has not identified a single gene. However, there is clear evidence that the *Wnt4/Rspo1* pathway is a key driver of ovary development (Chassot et al., 2008a, Vainio et al., 1999). *Wnt4* and *Rspo1* expression become female-specific at ~E12.5. *Wnt4* acts downstream of *Rspo1* and stabilizes β-catenin for nuclear internalization (Tevosian and Manuylov, 2008, Tomaselli et al., 2011, Chassot et al., 2008b, Liu et al., 2009). Although deletion of either *Wnt4* or *Rspo1* in the mouse only leads to partial female-to-male sex reversal with development of ovotestis (Chassot et al., 2008b, Vainio et al., 1999), the ectopic expression of a stabilized form of β-catenin in XY gonadal somatic cells leads to complete male-to-female sex reversal (Maatouk et al., 2008).

This suggests a model in which SOX9 and β-catenin compete in the nucleus at the bipotential stage: the accumulation of SOX9 following *Sry* expression outcompetes β-catenin engaging the male fate, whereas the failure of SOX9 to reach a certain threshold in a critical time-period allows β-catenin to antagonize the male pathway and promote female development.

Just as the female fate is actively antagonized in the testis, the male fate is suppressed in the ovary. This commitment to the female fate is maintained in the adult ovary by FOXL2 and ESR1/2. FOXL2 antagonizes the male pathway by binding TESCO and repressing *Sox9*. As expected, loss of *Foxl2* upregulates *Sox9* expression, resulting in postnatal transdifferentiation of granulosa cells into Sertoli cells in XX mice (Uhlenhaut et al., 2009).

The remarkable plasticity that underlies mammalian sex determination led us to posit that supporting cells must retain a certain memory, possibly epigenetic in nature, of their bipotential state and their ability to follow the opposing pathway. In the following sections, we will review the emerging evidence that the bipotential nature of supporting cells, as well as the canalization of the male or female pathway, are epigenetically regulated in mammals and other vertebrates.

1.3. Are XX and XY chromatin landscapes created equal?

For all heterogametic species, genetic sex (the presence of XX/XY or ZZ/ZW chromosomes) is established at fertilization long before the gonad forms, differentiates, and produces hormones. The presence of heteromorphic sex chromosomes leads to disparities in gene dosage between males and females. In flies and worms, the mechanisms that control dosage compensation are directly linked to sex determination (Cline and Meyer, 1996). In mammals, an XX/XY heterogametic species, the process of dosage compensation is not obviously linked to primary sex determination. To balance the difference in dosage of genes on the X between males and females, one X chromosome in females is inactivated in the late

blastocyst stage through heterochromatin formation and spreading (Zhao et al., 2008, Pinter et al., 2012). At one time it was believed that X-inactivation silenced all genes on the inactive X. However, in recent years it has become clear that around 15% of X-linked genes in humans (Carrel and Willard, 2005) and 3.3% of X-linked genes in mice (Yang et al., 2010a) escape inactivation and are expressed at higher levels in females than in males. Conversely, males express Y-linked spermatogenesis genes that are not shared with females. Therefore, XX and XY cells harbor differences attributable to their sex chromosome complement. It is clear that if a testis is induced in an XX embryo, the embryo will develop as a phenotypic male despite the fact that the genetic sex of each cell is female (Koopman et al., 1991, Vidal et al., 2001, Qin et al., 2004, Polanco et al., 2010). This means that gonadal sex is dominant in mammalian systems. Nonetheless, the sex chromosome complement of vertebrate cells could influence their differentiation before or after steroid-producing gonads develop. In fact there is clear evidence in birds that this is the case (McQueen and Clinton, 2009).

Interestingly, a number of X escapees and Y-linked genes play crucial roles in chromatin regulation (Table 1). Chromatin is regulated at many levels by epigenetic mechanisms, such as DNA methylation and histone modifications that can alter the 3D organization of chromatin within the nucleus. Changes to chromatin structure ultimately modulate transcription factor accessibility to DNA binding sites, thereby regulating gene expression. Examples of Y-linked and X escapee genes in this category will be reviewed below.

1.3.1. Y-linked chromatin modifiers

The Y chromosome is the smallest chromosome and Y-encoded genes mostly function in male sex determination, testis development and fertility. Amongst the Y-encoded genes are a few chromatin modifiers: the histone demethylase *Uty* (ubiquitously-transcribed TPR gene on

the Y chromosome), the H3K4-specific histone demethylase *Jarid1d* (Jumonji, AT-Rich Interactive Domain 1d)*,* and in humans, the histone acetyltransferase *CDY* (Chromodomain protein Y-linked) (Dhanoa et al., 2016)*.* Of these, *Uty* and *Jarid1d* have homologues on the X chromosome, and are reviewed in the following section.

      CDY is exclusively expressed in the testes. The CDY family of proteins contains two conserved domains: a chromodomain, which interacts with the repressive H3K9me histone modification (Kim et al., 2006a), and a catalytic histone acetyltransferase domain, which has high affinity for H4 (Lahn et al., 2002). However, this catalytic domain has also been found to recruit histone deacetylases (HDACs) (Caron et al., 2003). As histone acetylation correlates with transcriptional activity, *CDY* may function as both a transcriptional activator and a repressor. It is speculated that *CDY*-mediated hyperacetylation of H4 may play an important role in facilitating the transition from histones to protamines, nuclear proteins that replace histones in sperm (Lahn et al., 2002). However, whether *CDY* can regulate transcription at a genome-wide level through its interaction with H3K9me and HDACs remains to be elucidated.

Table 1.1. Y-linked and X escapee chromatin modifiers.

Y-linked chromatin modifiers		
Gene	**Molecular Function**	**References**
Uty	H3K27me3-specific histone demethylase	Lee et al., 2007; Wang et al., 2013; Walport et al., 2014
Jarid1d	H3K4-specific demethylase	Lee et al., 2007; Xu et al., 2008a
CDY (human)	H4-specific hyperacetylation, interacts with methylated H3K9 through its chromodomain, and recruits HDACs	Lahn et al., 2002; Kim et al., 2006a; Caron et al., 2003
X escapee chromatin modifiers		
Gene	**Molecular Function**	**References**
Utx	H3K27me3-specific histone demethylase	Hong et al., 2007; Xu et al., 2008b; Shpargel et al., 2012; Welstead et al., 2012; Wang et al., 2013
Jarid1c	H3K4-specific demethylase	(Iwase et al., 2007); Xu et al., 2008a
MeCP2	Binds methylated CpGs, recruits HDACs, HP1, H3K9 methyltransferases, the DNA-demethylase MBD2, and the chromatin remodeler ATRX, and promotes CTCF binding and formation of higher-order chromatin loops	Hendrich and Bird, 1998; Nan et al., 1998; Agarwal et al., 2007; Fuks et al., 2003; Becker et al., 2013; Nan et al., 2007; Kernohan et al., 2014

1.3.2. X escapee chromatin modifiers

Utx (ubiquitously-transcribed TPR gene on the X chromosome) is an X-linked gene that escapes X-inactivation in both mice and humans (Greenfield et al., 1998). This gene encodes a histone demethylase, which specifically removes the repressive histone modification H3K27me3 (Hong et al., 2007). *Utx* is critical for embryonic development, as *Utx*-null females die at midgestation (Shpargel et al., 2012, Welstead et al., 2012, Wang et al., 2013). *Utx* has a homolog on the Y chromosome, known as *Uty.* Although early *in vitro* studies showed that *Uty* was catalytically inactive (Lan et al., 2007), *Utx*-null males are viable and fertile (Shpargel et al., 2012, Welstead et al., 2012, Wang et al., 2013). Furthermore, *Uty*-null males phenocopy *Utx*-null males, suggesting that *Utx* and *Uty* are functionally redundant (Wang et al., 2013). Consistent with this interpretation, a more recent study found that human *Uty* is an active demethylase, but has reduced activity due to point mutations that affect substrate binding (Walport et al., 2014). However, even though *Utx* and *Uty* may have similar functions, their expression levels and patterns differ in XX and XY embryos (Xu et al., 2008b), suggesting that they may regulate chromatin in a sex-specific manner.

Jarid1c (Jumonji, AT-Rich Interactive Domain 1c) is an X escapee that encodes an H3K4-specific histone demethylase that plays an important role in brain development and function. *Jarid1c* is more highly expressed in XX than XY mice, independent of whether an ovary or a testis is present (Xu et al., 2008a). Importantly, expression of its Y-linked homolog *Jarid1d* is unable to compensate for loss of *Jarid1c* despite its similar function, suggesting that sex-specific expression of this histone demethylase may contribute to differences in brain development and function such as neurite length and aggressive behavior (Xu et al., 2008a, Lee et al., 2007).

Another X escapee, *MeCP2* (methyl CpG-binding protein 2), plays a crucial role in brain development. Although *MECP2* is not itself a chromatin modifier, it has a high binding affinity to methylated CpGs and ability to interact with epigenetic machinery (Hendrich and Bird, 1998). Once bound, *MECP2* can regulate chromatin structure by recruiting HDACS, heterochromatin protein 1 (HP1), H3K9 histone methyltransferases, the DNA demethylase MBP1, and the chromatin remodeler ATRX amongst others, and in one case was able to alter higher-order chromatin structure by inducing chromatin looping (Nan et al., 1998, Agarwal et al., 2007, Fuks et al., 2003, Becker et al., 2013, Nan et al., 2007, Kernohan et al., 2014).

Although there is currently no phenotypic evidence that these sex-linked chromatin modifiers directly regulate primary sex determination, sexually dimorphic expression of chromatin modifiers could directly affect the activity of many genes leading to sex-specific transcriptional programs. Whether XX and XY cells initiate development with different chromatin landscapes, and whether these differences have significant phenotypic consequences in fetal or adult life is an area of active research using models that can disentangle genetic and hormonal effects (Burgoyne et al., 1995, Arnold, 2009).

1.4. The Chromatin Landscape at the Bipotential Stage

Despite differences in sex-chromosome complements, XX and XY fetal gonads are initially bipotential, and the somatic cells that compose the gonad have the full ability to differentiate into either Sertoli cells (males) or granulosa cells (females) (Albrecht and Eicher, 2001). Consistent with their bipotential nature, both XX and XY progenitor cells co-express genes later associated with Sertoli- and granulosa-cell development (Nef et al., 2005, Munger et al., 2013). In other pluripotent systems, such as embryonic stem cells (ESCs), chromatin has an open configuration, which is associated with active transcription due to accessibility of

transcription factor binding sites and transcription start sites. This open configuration confers pluripotent cells with a unique plasticity that enables differentiation into a wide range of lineages (Atlasi and Stunnenberg, 2017). In ESCs, promoters of developmental genes are often marked by both the active H3K4me3 and the repressive H3K27me3 histone modifications (Box 1) (Bernstein et al., 2006a, Azuara et al., 2006). Despite their antagonistic functions, this "bivalency" plays a key role in pluripotency by maintaining genes in a poised state for rapid activation or repression upon differentiation. As pluripotent cells differentiate, chromatin becomes more restricted by epigenetic deposition of DNA methylation and histone modifications that establish heritable cell-type-specific gene expression patterns (Atlasi and Stunnenberg, 2017).

Based on this model, one hypothesis is that the chromatin landscape in XX and XY bipotential progenitor cells is similar, enabling equal access to promoters of sex-determining genes and regulatory elements of both sexes, and that after sex-determination, the chromatin landscape becomes more restricted to canalize either the male or female fate and repress the alternative pathway. Investigation of this hypothesis (discussed in Chapter 3) revealed that, similar to ESCs, key sex-determining genes that are co-expressed in XX and XY mouse progenitor cells are bivalent, possibly contributing to the bipotential nature of these cells (García-Moreno et al., unpublished). The upregulation of the male or female pathway is accompanied by mutually antagonistic mechanisms that repress the alternate cell fate (Kim et al., 2006b, Ottolenghi et al., 2007, Chassot et al., 2008b, Uhlenhaut et al., 2009, Matson et al., 2011). At the chromatin level, this transition is reflected in a reorganization of histone marks around sex determining genes. Genes associated with the female pathway lose their repressive mark when the ovarian pathway is activated, and genes associated with the male pathway lose their repressive mark when the testis pathway is activated. However, ovary-determining genes

that become transcriptionally repressed when the gonad commits to a male fate remain bivalent, similar to the fate of testis-determining genes in the ovary (García-Moreno et al., unpublished). Maintaining the alternate cell fate in a state poised for activation possibly contributes to the unique capacity of supporting cells to transdifferentiate from Sertoli cells to granulosa cells and vice-versa, even during adulthood (Uhlenhaut et al., 2009, Matson et al., 2011). As bivalency is a conserved epigenetic mechanism (Lesch et al., 2016), it will be interesting to determine whether bivalent sex-determining genes confer plasticity to non-mammalian systems that can switch between the male and female fate in response to environmental stimuli.

Another study performed DNaseI hypersensitive site sequencing (DNaseI-seq) and chromatin immunoprecipitation followed by sequencing (ChIP-seq) for H3K27ac, a histone modification indicative of active enhancers, to generate a profile of the chromatin landscape in purified mouse Sertoli cells at embryonic day 13.5 (E13.5), soon after commitment to the male pathway (Maatouk et al., 2017). DNaseI-seq identified 28,133 nucleosome-depleted regions (NDRs) (corresponding to regulatory elements) that were unique to Sertoli cells when compared to six other tissues. As expected, these NDRs were commonly located near Sertoli-specific genes, implying that they represent regulatory elements that drive Sertoli cell differentiation. Interestingly, granulosa-specific genes had a similar enrichment of Sertoli-unique NDRs in their neighborhood in Sertoli cells, suggesting that the chromatin architecture between Sertoli and granulosa cells is similar and may reflect their shared progenitor state. However, H3K27ac-positive NDRs were significantly enriched only around Sertoli-specific genes. Together these findings indicate that as sex-determining pathways diverge, genes in Sertoli cells that are associated with the male pathway lose their repressive H3K27me3 mark and acquire H3K27ac. Simultaneously, genes associated with the female pathway remain bivalent in Sertoli cells, and

retain NDRs in their enhancer/promoter regions. It is likely that these H3K27ac-negative NDRs (inactive enhancers) associated with granulosa-specific genes are bound by repressors that block upregulation of the female-determining pathway.

1.5. Epigenetic Regulation of the Primary Male Switch

The balance between male and female fate in the early gonad is disrupted by master regulators of sex-determination that act as a "switch" to activate the male pathway and repress the female pathway. There is increasing evidence that these primary switches are epigenetically regulated. Below, we will review how the primary mammalian switch *Sry* is epigenetically regulated during sex determination, and provide evidence that the epigenetic regulation of master regulators of male sex-determination is conserved in non-mammalian systems.

1.5.1. The mammalian switch, Sry

In mammals, it is the presence of the Y-encoded transcription factor *Sex–determining Region Y (Sry)* that directs the bipotential fetal gonad towards the male fate (Sinclair et al., 1990, Gubbay et al., 1990, Koopman et al., 1991). *Sry* expression breaks the initial male-female balance in bipotential progenitors and directs differentiation of Sertoli cells through a downstream signaling cascade (Albrecht and Eicher, 2001). In XY mice, *Sry* is transiently expressed between E10.5 and E12.5 (Hacker et al., 1995). Although there are still unanswered questions about the events leading to the activation of *Sry*, a number of epigenetic modifiers have been identified.

1.5.1.1. Regulation of *Sry* mediated by DNA methylation

DNA methylation, which in mammals occurs primarily at cytosines located next to

guanines (CpG), plays a crucial role during development by epigenetically controlling gene silencing and establishing cell-type-specific gene expression patterns during differentiation (Razin and Cedar, 1984). The methylome of the Y chromosome was first explored by Nishino et al. (Nishino et al., 2004). This group generated an *in vivo* methylation profile of XY mouse embryos throughout development using bisulfite sequencing, a standard technique that identifies methylated CpGs. There are 16 CpGs in the 4.5kb region upstream of *Sry* that cluster into two regions. Region I contains four CpGs and overlaps the TSS of a circular transcript of *Sry,* which is believed to be untranslated (Dolci et al., 1997), whereas Region II contains seven CpGs and overlaps the TSS of the linear transcript of *Sry,* which is translated into a functional protein from E10.5-E12.5 in XY fetal gonads (Jeske et al., 1995). Bisulfite sequencing revealed that both regions were hypomethylated at the blastocyst stage and became hypermethylated in E8.5 embryos in which *Sry* is not yet expressed. At the time of *Sry* activation (E11.5), both regions lost methylation in the gonad, but not in other tissues such as the liver. Finally, at E15.5, Region II became re-methylated, whereas Region I remained hypomethylated. It is unlikely that hypomethylation at Region I is of functional significance, as this region did not induce activity in an *in vitro* promoter assay (Nishino et al., 2004). However, the possible three-dimensional structure of the inverted repeats around this locus may interfere with *in vitro* promoter assays (Gubbay et al., 1990, Capel et al., 1993). Nonetheless, this work established a clear anti-correlation between promoter CpG methylation and *Sry* expression in the gonad throughout sex determination (Fig. 1.2).

To investigate the mechanistic link between methylation and *Sry* expression, Nishino et al. established an *in vitro* system by culturing primary cells isolated from E11.5 XY gonads (Nishino et al., 2004). Introduction of previously *in vitro*-methylated constructs containing either the circular or linear promoter regions strongly suppressed the activity of ectopically introduced

Sry. These results clearly showed that *Sry* gene expression can be epigenetically regulated through DNA methylation-mediated silencing.

In mammals, DNA methylation occurs primarily at the CpG site. Non-CpG methylation, which occurs at CNG sites in which N can be any nucleotide, is more common in plants (Gruenbaum et al., 1981), bacteria (Yoder et al., 1997, Casadesus and Low, 2006) and yeast (Pinarbasi et al., 1996), but was observed at CC(A/T)GG motifs in various mammalian *systems* (Malone et al., 2001, Franchina and Kay, 2000, Agirre et al., 2003, Imamura et al., 2005). Interestingly, the promoter for both the circular and linear transcripts of *Sry* (regions I and II) contains a CCTGG motif (Nishino et al., 2011). In the gonad, this motif has the opposite methylation dynamic than its neighboring CpG sites, and coincides with *Sry's* temporal expression pattern. While CpG sites transiently lose methylation at E11.5, CCTGG sites are transiently methylated (Fig. 1.2). An *in vitro* promoter assay further showed that CCTGG methylation can induce *Sry* activity, suggesting that non-CpG methylation is associated with active transcription and may play an important role in *Sry* activation. However, although the CCTGG site at Region II was conserved in six mouse strains, two of these strains (C57BL/6 and 129S1) did not exhibit CCTGG methylation at E11.5, suggesting that non-CpG methylation may be background-dependent and raises the possibility of strain-specific differences in epigenetic regulation of *Sry* (Nishino et al., 2011).

Although it is well established that DNA methylation strongly correlates with *Sry's* activity, the mechanisms by which its promoter CpG sites are dynamically methylated during such a precise window of development have not been described. This is partly due to the fact that *in vivo* disruption of DNA methyltransferases (DNMTs) leads to early embryonic lethality (Li et al., 1992, Okano et al., 1999). Warr and colleagues hypothesized that GADD45g*,* a protein that can induce gene expression by recruiting DNA repair factors to demethylate target genes

(Barreto et al., 2007, Niehrs and Schafer, 2012), was involved in *Sry's* activation based on the observation that loss of *Gadd45g* caused complete male-to-female sex reversal in mice due to failure to activate *Sry* (Warr et al., 2012)*.* However, loss of *Gadd45g* did not alter the normal methylation status of *Sry* during sex determination, suggesting that it does not activate *Sry* through promoter demethylation, but must function through an alternative mechanism (Gierl et al., 2012). Despite this, the same study found that demethylation at *Sry's* promoter begins as early as 5-6 tail somites, and is therefore one of the earliest male-specific molecular events preceding *Sry* expression.

Interesting data concerning the epigenetic regulation of sex determination came from cloned animals that were generated by somatic cell nuclear transfer (SCNT). Cloned animals exhibit anomalies of sex differentiation despite having a normal karyotype (Inoue et al., 2009). For example, canine SCNT results in male-to-female sex reversal in almost 23% of cases (Jeong et al., 2016). A genetic analysis of *SRY+* XY DSD dogs found no genetic differences that could account for sex reversal. However, the *SRY* locus in XY DSD dogs was hypermethylated compared to control XY dogs (Jeong et al., 2016). This hypermethylation corresponded with reduced expression of *SRY* (in addition to other downstream male-determining genes), and an upregulation of the ovarian developmental pathway. Although this case is specific to an artificial system, it strongly supports the idea that epigenetic regulation can control the sex-determination network.

Not much is known about the role that DNA methylation plays in regulating the human *SRY* gene. In one study, an epigenetic profile of the Y chromosome was generated from blood cells of two individuals from distinct populations (Singh et al., 2011). A CTCF transcription factor binding site was found in the promoter region of *SRY*, which overlapped an unmethylated CpG site. CTCF binds chromatin insulators and acts as a transcriptional repressor. Interestingly, this

site was enriched for repressive histone modifications, pointing to the possibility that the human *SRY* gene may be silenced through recruitment of repressors and not through DNA methylation. However, caution should be taken when drawing conclusions based on analysis of a non-gonadal cell type that does not usually express *SRY.*

Active genes	Silenced genes
H3K4me2/3 H3K27ac H3K9ac H4K16ac	H3K27me3 H3K9me2/3

Table 1.2. Histone modifications mentioned in this work that mark active promoters or enhancers (green), or silenced genes (red).

1.5.1.2. Regulation of *Sry* mediated by histone modifications

　　Evidence that histone modifiers are critical regulators of mammalian sex determination came from the observation that loss of the chromobox protein homologue 2 (*Cbx2*) leads to complete male-to-female sex reversal and development of hypoplastic gonads in mice (Katoh-Fukui et al., 1998) and humans (Biason-Lauber et al., 2009). CBX2 is part of the Polycomb-group (PcG) of proteins that typically function as transcriptional repressors and play a critical role in regulating gene silencing during development (Simon and Kingston, 2009, Margueron and Reinberg, 2011). The PcG proteins assemble into two complexes: the Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2). Canonically, PRC2 first catalyzes H3K27me3 at the promoter of its target genes. PRC1 then binds H3K27me3 through its CBX2 subunit, and maintains gene repression through chromatin compaction (Bernstein et al., 2006b, Lau et al., 2017). Sex-reversal in XY *Cbx2* mutants is characterized by a failure to upregulate *Sry* and its

downstream target *SRY*-box 9 (*Sox9*) (Katoh-Fukui et al., 2012). As sex-reversal was rescued by forced expression of *Sry,* CBX2 was proposed to act as an early activator of *Sry* expression (Katoh-Fukui et al., 2012). Although some subunits of the PcG proteins have been reported to drive active transcription of certain target genes (Jacob et al., 2011, Mousavi et al., 2012, Herz et al., 2012), CBX2 has not been shown to act directly as an activator. Therefore, a role for CBX2 as an activator was unexpected. Recent evidence suggests that PRC2 may be involved in "transcriptional focusing" by spatially confining H3K27me3+ silent domains (Lu, 2017). It is possible that disruption of the PcG complex, such as loss of CBX2, may lead to aberrant spread of these domains resulting in silencing of underlying genes, possibly including *Sry*. Another possibility is that CBX2 induces *Sry* expression indirectly by repressing a repressor. Because the male and female pathways are mutually antagonistic, stabilization of the male fate requires active repression of female-determining genes, raising the possibility that CBX2 has a widespread role in the repression of the female-pathway, which may otherwise block male fate commitment.

Further evidence for the epigenetic regulation of *Sry* mediated through histone modifiers came from studies of the histone demethylase *Jmjd1a*. *Jmjd1a*-deficient mice exhibit male-to-female sex reversal (Kuroki et al., 2013). *Jmjd1a* encodes a histone demethylase that acts specifically on the repressive H3K9me2 histone modification. ChIP for H3K9me2 and the active histone modification mark H3K4me2 was performed in purified progenitor cells from gonads of XY mice at E11.5 when *Sry* expression is at its peak. As opposed to control XY mice, the *Sry* locus of *Jmjd1a*-deficient mice retained high enrichment of the repressive H3K9me2 mark, and had low enrichment of the active H3K4me2 mark. Consistently, *Sry* expression levels were reduced in mutant mice. This suggests that *Jmjd1a*-dependent demethylation of H3K9me2 is required for subsequent H3K4me2 accumulation at *Sry's* promoter, leading to an open

chromatin conformation and *Sry* activation (Fig. 1.2). Consistent with this observation, a separate study found that the *Sry* promoter has low levels of H3K9me3 and is enriched for the active H3K4me3 modification in E12.5 mouse testes. In contrast, adult testes have high H3K9me3 and low H3K4me3 enrichment, suggesting that the *Sry* locus regains H3K9me3 after it is silenced, and that this mark maintains long-term repression of *Sry* throughout adulthood in mice (Sinha et al., 2017). Recently, sex-reversal of *Jmjd1a*-deficient mice was rescued by disrupting the GLP/G9a H3K9 methyltransferase complex, pointing towards this complex as the one that catalyzes H3K9me2 at the *Sry* locus (Fig. 1.2) (Kuroki et al., 2017). This study highlights how various epigenetic factors interact to fine-tune the expression of *Sry* by balancing the levels of H3K9me2 at its promoter.
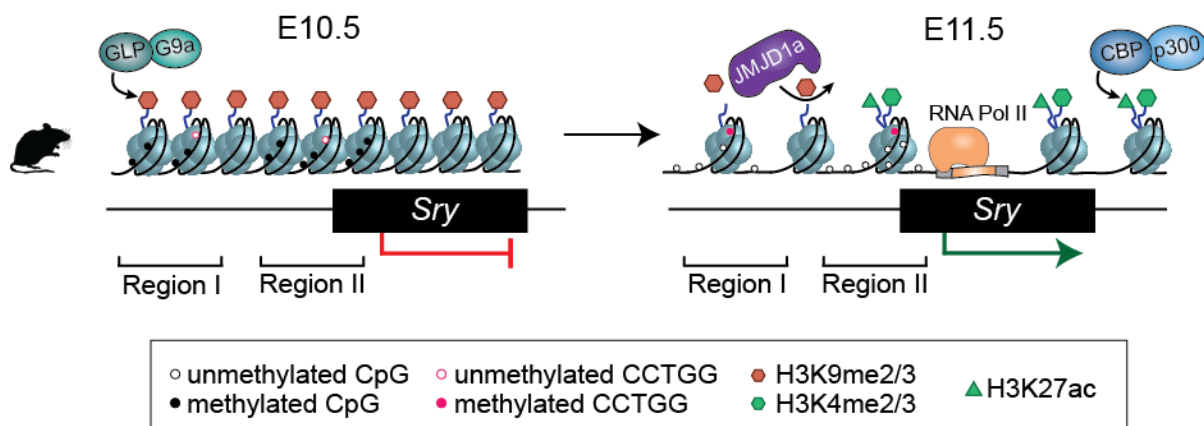
Figure 1.2. Epigenetic regulation of *Sry* in mammals. Prior to sex determination (left) *Sry* is repressed and enriched for the repressive GLP/G9a-mediated H3K9me2 histone modification. Region I overlaps the TSS of an untranscribed circular *Sry* transcript, Region II overlaps the TSS of a functional linear *Sry* transcript. At E10.5, regions I and II contain methylated CpGs, and unmethylated CCTGG sites. At E11.5, *Sry* activation requires *Jmjd1a*-mediated removal of H3K9me2, CBP/p300-mediated deposition of H3K27ac, and accumulation of the active H3K4me2 at its promoter. CpGs at region I and II become demethylated, and CCTGG sites become methylated.

In addition to accumulation of H3K4me2 at *Sry's* promoter, activation of *Sry* requires deposition of H3K27ac by the CREB-binding protein (CBP) and its paralogue p300 (Fig. 1.2) (Carre et al., 2017). A recent study found that deletion of either gene in gonadal somatic cells results in partial male-to-female sex reversal, whereas loss of three of the four *p300/Cbp* functional alleles leads to complete male-to-female sex reversal. ChIP showed reduced levels of H3K27ac at the promoter of *Sry* in *p300/Cbp* mutants at E11.5, corresponding to reduced *Sry* and *Sox9* expression levels. This study indicates that this histone/lysine acetyltransferase complex is a key component in the activation of *Sry* during testis development.

Interestingly, epigenetic profiling of human blood cells showed that the human *SRY* gene promoter was enriched for the active H3K9ac mark and the repressive H3K27me3 mark, but H3K9me3 was only enriched over the gene body (Singh et al., 2011). The conflicting H3K9ac and H3K9me3 modifications may be a reflection of mixed cell populations, as *SRY* is expressed in the B-cell lineage but not in other blood lineages. Studies of human gonadal lineages are required to understand which chromatin modifiers are involved in *SRY* activation during sex determination in humans. However, there is evidence that disruption of epigenetic regulation can lead to disorders of sexual development (DSDs). Consistent with the mouse studies mentioned above, one XY phenotypically normal girl with hypoplastic ovaries was found to have inherited two loss-of-function mutations in the *CBX2* gene (Biason-Lauber et al., 2009). In a separate case, an *SRY+* XY woman with streak gonads had a hyperacetylated *SRY* locus relative to control XY males (Mitsuhashi et al., 2010). This hyperacetylation coincided with prolonged *SRY* activity, which the authors suggested might destabilize the downstream male-determining network, although further experiments are required to fully explain this case.

1.5.2. Evidence for Epigenetic Regulation of the Key Switch Gene, Dmrt1, in other Vertebrates

Several new lines of evidence suggest that epigenetic regulation of key switch genes may be a conserved element of sex determination pathways in different species. *Dmrt1* (doublesex and mab3-related transcription factor 1) is a highly conserved transcription factor that plays a critical role in vertebrate sex determination (Matson and Zarkower, 2012) and has been shown to act as the master sex-determining gene in many non-mammalian systems (Masuyama et al., 2012, Yoshimoto et al., 2008, Smith et al., 2009). In mice, *Dmrt1* is not required for testis-development; however, it is essential for maintaining the Sertoli cell phenotype in adult testes, as loss of *Dmrt1* causes transdifferentiation of Sertoli cells into granulosa cells (Matson et al., 2011). Although there are currently no studies investigating the epigenetic regulation of *Dmrt1* in mammals, its epigenetic regulation has been investigated in the half-smooth tongue sole, chickens, and the red-eared slider turtle.

1.5.2.1. Epigenetic regulation of *dmrt1* in fish

dmrt1 is highly conserved in fish and plays an important role in male sex determination (Marchand et al., 2000, Guo et al., 2005). The half-smooth tongue sole *Cynoglossus semilaevis* is a flatfish that is a fascinating model to study the epigenetic regulation of sex determination, as it has both genetic- and temperature-dependent sex-determining mechanisms. The genetic sex of this species is determined by its Z and W chromosome composition: males carry two Z chromosomes, whereas females carry a Z and a W chromosome. In this species, dosage of the Z chromosome is the sex-determining factor. Despite the inheritance of sex chromosomes, ~14% of ZW females develop as males under normal conditions (22°C), and sex reversal can increase to ~73% when fish are reared at high temperatures (28°C) (Chen et al., 2014). These

sex-reversed fish are "pseudomales" that can reproduce with normal females. Importantly, the offspring of pseudomales and normal females maintain high sex-reversal rates under normal conditions (22°C), suggesting that the ability to sex reverse can be inherited.

Shao et al. generated a DNA methylation profile of female, male and pseudomale fish to investigate the role of DNA methylation in the transition from genetic to environmental sex determination (Shao et al., 2014). This study identified *dmrt1,* which is Z-encoded in the tongue sole (Chen et al., 2014), as a critical gene that responds to temperature change through DNA methylation. Consistent with its expression pattern, *dmrt1* maintains low methylation levels throughout life in male gonads, but is increasingly methylated in female gonads starting at the time of sex determination and throughout adulthood. In ZW pseudomales, the *dmrt1* gene has low methylation levels similar to ZZ males, and is therefore active and able to drive testis-development. This male-specific methylation is then inherited in the ZW offspring of ZW pseudomales and ZW females, accounting for a higher rate of sex-reversal compared to ZW offspring between normal ZZ males and ZW females. This finding suggests that temperature can override genetic sex-determination by directly altering the epigenetic marks at *dmrt1,* and that in this species where methylation is not stripped from the genome in the early embryo, this effect is transgenerational.

1.5.2.2. Epigenetic regulation of *DMRT1* by MHM in chickens

As in the half-smooth tongue sole, chickens have a ZZ (male) and ZW (female) sex chromosome composition. *DMRT1* is also a Z-linked gene in chickens that is more highly expressed in males than in females. Overexpression of *DMRT1* in chick ZW gonads induced the male pathway (Lambeth et al., 2014), whereas knockdown of *DMRT1* led to feminization of male gonads (Smith et al., 2009).
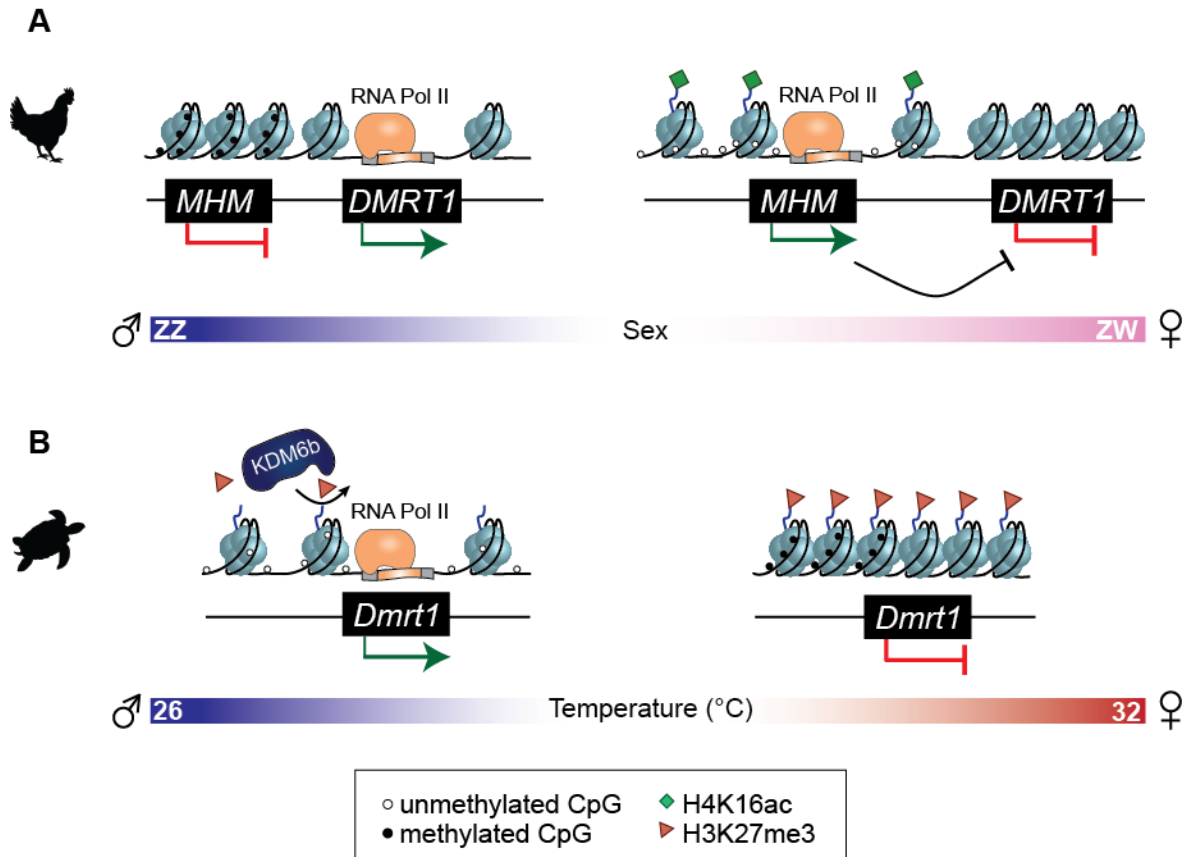
Figure 1.3. Epigenetic regulation of *Dmrt1*. A) Epigenetic regulation of *DMRT1* in chickens. In males (ZZ, left), a region neighboring the male-determining gene *DMRT1* is hypermethylated (MHM) and repressed. In females (ZW, right), MHM is hypomethylated and enriched for the active H4K16ac histone modification. The open chromatin conformation of MHM in females enables transcription of a lncRNA that inhibits expression of *DMRT1.* B) Epigenetic regulation of *Dmrt1* in red-eared slider turtles. During embryogenesis, turtles develop as males if eggs are incubated at lower temperatures (blue, left) and as females at higher temperatures (red, right). At male-producing temperatures, *Kdm6b* is upregulated. KDM6b-mediated removal of the repressive H3K27me3 and demethylation of CpGs at the promoter of *Dmrt1* is required for its activation and subsequent testis development. At female-producing temperatures, enrichment of H3K27me3 and CpG methylation at the promoter of *Dmrt1* inhibits its expression and leads to ovary development.

A region neighboring *DMRT1* is differentially methylated between males and females (Teranishi et al., 2001). In males, this region is hypermethylated (known as the male hypermethylated region or MHM) and is transcriptionally inactive, whereas in females, this region is hypomethylated and transcribed into a long non-coding RNA. In addition to being hypomethylated, the female MHM locus is enriched for the active H4K16ac mark (Bisoni et al., 2005) and has an open chromatin configuration (Itoh et al., 2011). MHM is hypothesized to inhibit *DMRT1* expression in females, as the MHM transcript accumulates at the site of transcription adjacent to *DMRT1* (Teranishi et al., 2001) (Fig. 1.3A). Consistent with this hypothesis, injection of exogenous MHM into rooster testicles led to a strong downregulation of *DMRT1* and a paling of the comb color, indicative of a change in sex hormones. Importantly, transcriptional levels of other key sex-determining genes such as *AMH* and *SOX9* were unaffected, suggesting that MHM specifically regulates *DMRT1* (Yang et al., 2010b)*.*

1.5.2.3. Epigenetic regulation of *Dmrt1* by KDM6B in turtles

The red-eared slider turtle *Trachemys scripta* lacks heteromorphic sex chromosomes. Instead, the temperature of incubation acts as a sex determinant. During embryogenesis, turtles develop as males if eggs are exposed to lower temperatures (26°C), and as females if they are exposed to higher temperatures (32°C). Prior to sex determination, *Dmrt1* expression in the embryonic gonads is sexually dimorphic, with higher levels at MPT (male-producing temperature) than at FPT (female-producing temperature) (Kettlewell et al., 2000, Czerwinski et al., 2016). Recently, *Dmrt1* was shown to be both necessary and sufficient to initiate male sex determination in *T. scripta* (Ge et al., 2017). Moreover, *Dmrt1* responds rapidly to temperature shifts, which establishes it as a master regulator of temperature-dependent sex determination*.* Interestingly, DNA methylation of the *Dmrt1* promoter region is inversely correlated with its

expression and fluctuates in response to temperature changes. For example, the methylation level of the *Dmrt1* promoter greatly increased when eggs were shifted from MPT to FPT. In contrast, shifting eggs from FPT to MPT resulted in decreased promoter DNA methylation and gene activation (Ge et al., 2017) (Fig. 1.3B). This suggests that, similar to the tongue sole flatfish, temperature can regulate both sex-determining gene expression and promoter methylation. However, although the previous studies show an inverse correlation between CpG promoter methylation and *Dmrt1* expression, whether DNA methylation is a cause or a consequence of expression changes in temperature-dependent systems has not been established.

Transcriptome sequencing of *T. scripta* embryonic gonads at MPT and FPT revealed that in addition to *Dmrt1, Kdm6b* (or *Jmjd3*) is also upregulated at MPT and its expression levels fluctuate in response to temperature changes (Czerwinski et al., 2016). *Kdm6b* is an H3K27me3-specific demethylase that can activate target genes through the removal of this repressive histone modification. Loss of *Kdm6b* in MPT gonads results in a complete shift from the male to the female sexual trajectory, suggesting that it also plays a key role in male sex-determination (Ge et al., 2017, submitted). Knockdown of *Kdm6b* resulted in a significant downregulation of *Dmrt1* expression in MPT gonads, pointing towards *Dmrt1* as a downstream target of *Kdm6b.* Consistent with this hypothesis, ChIP-qPCR revealed that the *Dmrt1* promoter had low levels of H3K27me3 in MPT gonads in which *Kdm6b* and *Dmrt1* expression levels are high. In contrast, the *Dmrt1* promoter was enriched for H3K27me3 in FPT gonads in which *Kdm6b* and *Dmrt1* expression levels were low. Importantly, H3K27me3 levels at *Dmrt1* increased upon loss of *Kdm6b,* suggesting that *Kdm6b* activates *Dmrt1* by removing H3K27me3 at MPT leading to activation of this switch gene that induces male sexual development (Fig.

1.3B). This study provides the first direct mechanistic link between epigenetic and temperature regulation of sex determination.

1.6. Chromatin-mediated Regulation of *Sry's* Downstream Target *Sox9*

In the developing testis, *Sry's* primary function is to upregulate *Sox9* expression (Sekido and Lovell-Badge, 2008). *Sox9* is part of the *Sry*-type HMG box (Sox) family, a group of proteins with over 60% similarity to the SRY high-mobility group (HMG) box region (Denny et al., 1992a). In mammals, several Sox genes, including *Sox9,* can replace *Sry* (Zhao and Koopman, 2012). For example, directed expression of *Sox9* in XX mice induces testis development, demonstrating that *Sox9* is sufficient for male sex determination (Vidal et al., 2001). Consistent with this, duplication of *SOX9* in XX humans can lead to female-to-male sex reversal (Cox et al., 2011). On the other hand, *Sox9*-null XY gonads develop as ovaries (Barrionuevo et al., 2006), and loss-of-function mutations of the human *SOX9* gene can lead to female development (Kwok et al., 1995, Stoeva et al., 2011). These studies highlight the role of *Sox9* as a critical regulator of male sex determination and Sertoli cell development.

1.6.1. Enhancer-mediated regulation of Sox9 in mice

Sox9 lies downstream of a 2Mb gene desert, known to contain a number of different regulatory elements that control the spatiotemporal expression of *Sox9* during development (Bagheri-Fam et al., 2006). The size and complexity of this regulatory region is not surprising, as *Sox9* is expressed in a number of different tissues such as cartilage, brain, pituitary, lung, pancreas, and testes, and is probably driven by tissue-specific enhancers. Bangheri-Fam et al. performed the first characterization of tissue-specific regulatory elements within this region (Bagheri-Fam et al., 2006). This group used transgenic mice expressing *lacZ* as a reporter gene driven by a *Sox9* minimal promoter to assess the regulatory potential of seven conserved

regions (between pufferfish and humans) lying upstream and downstream of *Sox9*. This assay identified 3 elements that drove tissue-specific *lacZ* expression, however none of these elements drove *lacZ* expression in the testes. Two years later, the first testis-specific *Sox9* enhancer (TES) was identified (Sekido and Lovell-Badge, 2008) using a similar approach. TES and its core element TESCO lie -13 to -10kb upstream of the *Sox9* TSS, and are bound by SF1, SRY and SOX9 in the fetal gonad. This supports a model for the regulation of *Sox9* in which the transcription factor SF1 initiates, SRY upregulates and SOX9 maintains expression in the developing gonad by binding to a testis-specific enhancer. Noteworthy, the key ovarian-promoting transcription factor FOXL2 was found to bind TES and repress *Sox9* (Uhlenhaut et al., 2009)*.* Therefore, this regulatory element plays a major role in directing both tissue- and sex-specific expression of *Sox9* by integrating positive and negative signals*.*

Until recently (2017), TES/TESCO remained the only *Sox9* enhancer identified that mimics endogenous *Sox9* expression in the gonads. It was not until the advent of CRISPR/Cas9 as a gene-editing tool that TES and TESCO were deleted in mice with the goal of determining the specific contribution of this enhancer to sex determination. Surprisingly, TES/TESCO null XY mice developed normally, despite a ~55% reduction of normal *Sox9* levels (Gonen et al., 2017). Perhaps the most important take-away from this study is the revelation that TES does not act alone, and that there must be one or more additional enhancers that coordinate to account for the full levels of *Sox9.* This is a reasonable assumption as multiple enhancers within *Sox9's* regulatory region have been shown to regulate its expression in cartilage (Yao et al., 2015), and because no mutations within the human homolog of TESCO have been identified in DSD patients (Georg et al., 2010).

The methods described above used to identify enhancers are labor-intensive, time consuming, and biased. They require one to know where to look in the genome, to take

fragments that span the whole region of interest, and to further break those fragments down into progressively smaller portions until a core element is identified. It is perhaps because of this that TES remained the only identified testis-specific enhancer for almost a decade. Luckily, technological advances such as high-throughput sequencing led to the development of genome-wide and unbiased assays that rendered the identification of regulatory elements easy and affordable.

Regulatory elements lie within regions of open chromatin that are typically devoid of nucleosomes, known as nucleosome-depleted regions (NDRs). The open chromatin conformation allows DNA binding motifs to be accessed by transcription factors, insulators or repressors, or in the case of promoters, by RNA Polymerase II. Treating chromatin with a DNA-specific enzyme, DNase I, selectively digests exposed DNA at NDRs, whereas DNA tightly wrapped around histone proteins is protected against its activity. The digested DNA can then be amplified, sequenced, and mapped back to the genome, generating a map of NDRs. This method is known as DNaseI hypersensitive site sequencing (DnaseI-seq). Furthermore, regulatory regions have specific chromatin signatures depending on their function. For example, active enhancers are marked by H3K27ac and H3K4me1, primed enhancers are marked only by H3K4me1, and poised enhancers by H3K4me1 and H3K27me3 (Shlyueva et al., 2014). Therefore, by combining DNaseI-seq with ChIP-seq for various histone modifications, one can not only identify regulatory elements, but also classify them based on their function.

Maatouk et al. developed a profile of the chromatin landscape in FACS-purified Sertoli cells from E13.5 mouse testes by performing DNaseI-seq and ChIP-seq for H3K27ac. DNaseI-seq revealed 33 putative enhancers upstream of *Sox9* (Maatouk et al., 2017), 16 of which were assayed for regulatory potential by transient transgenics in the mouse (discussed in Chapter 4) (Gonen et al., 2018b). Four of these enhancers drove gonad-specific *lacZ* expression, validating

DNaseI-seq as a tool for identifying functional enhancers. Of these four enhancers, one overlaps a region in humans known as XYSR, which is known to be associated to DSDs (see below). Amazingly, deletion of this enhancer, termed Enh13, caused complete male-to-female sex reversal (Gonen et al., 2018b) resulting from reduced *Sox9* expression. Therefore, Enh13 is necessary for normal testis development in the mouse by driving testis-specific *Sox9* expression at the time of sex determination. Whether it is necessary for multiple testis-specific enhancers to function together to regulate both the timing and expression levels of *Sox9,* or whether most enhancers are redundant remains to be elucidated. Regardless, it is remarkable that deletion of 557bp of a non-coding sequence is able to alter the fate of a cell lineage and the sex of an organism.

1.6.2. Disruption of the regulatory region upstream of SOX9 causes DSDs in humans

The experimental challenges presented when working with human subjects makes it difficult to identify enhancers that regulate the human Sox9 gene. However, insight into how this gene is regulated can be gained from studying cases of human disease caused by *SOX9* disruption. SOX9 acts as a master transcription factor in chondrocytes; therefore *SOX9* mutations can cause a skeletal malformation syndrome known as campomelic dysplasia (CD). Breakpoints in the regulatory region upstream of the human Sox9 gene had been previously described in CD individuals, defining a cis-regulatory region driving chondrocyte-specific expression of *SOX9* (Leipoldt et al., 2007). Years later, a systematic screening of this region revealed multiple chondrocyte-specific enhancers (Yao et al., 2015). Interestingly, up to two-thirds of individuals affected by CD display 46,XY male-to-female sex reversal, suggesting that this region contains enhancers that drive *SOX9* expression both in the cartilage and in the testes. However, there are cases of DSDs with no signs of skeletal malformation, suggesting

that that there are tissue-specific regions within the Sox9 locus that can lead to isolated forms of disease when disrupted.

For example, a 46,XX DSD newborn was characterized as a developmentally normal male with no signs of CD, suggesting that no cartilage-specific regulatory element was disrupted (Refai et al., 2010). Typically, 46,XX DSD males are *SRY*-positive due to a translocation of the *SRY* locus on to an autosome. However, this individual was *SRY*-negative, and instead presented a translocation of chromosome 12 to chromosome 17, with the breakpoint on chromosome 17 mapping ~776Kb upstream of *SOX9.* Rather, the 12:17 translocation inserted a regulatory element for the gene Deynar or pseudogene LOC204010 on chromosome 12 into the *SOX9* locus, potentially leading to upregulation of *SOX9* and subsequent testis development in absence of *SRY* (Refai et al., 2010). This case is reminiscent of the *Odsex* mutation that led to normal male development in XX *Sry*-negative mice due to the transgenic insertion of the *Dct* promoter ~1Mb upstream of *Sox9*, driving its expression in the gonads (Bishop et al., 2000, Qin et al., 2004). This was the first study in mice showing that a long-range cis-regulatory element could drive *Sox9* expression in the gonads.

To identify testis-specific regulatory regions in humans, Benko et al. performed copy number variation mapping at the *SOX9* locus in three *SRY*-negative 46,XX patients with female-to-male sex reversal, and two *SRY*-positive 46,XY patients with male-to-female sex reversal (Benko et al., 2011). No cases presented any form of skeletal malformations or abnormalities reminiscent of CD. All three 46,XX male patients shared a duplication upstream of *SOX9,* whereas both 46,XY females shared a deletion in this same region. The overlap between the duplications and deletions revealed a minimal non-coding 78kb region located 517-595kb upstream of the *SOX9* promoter, which was named *Reversal of Sex* (*RevSex*). *RevSex* is hypothesized to contain at least one critical testis-specific enhancer, supporting a model in

which gain-of-function mutations (resulting from duplications or triplications of *RevSex*) cause gonadal-specific *SOX9* upregulation leading to testis development, and loss-of-function mutations (resulting from deletions of *RevSex*) cause gonad-specific loss of *SOX9* and ovarian development. A separate study then identified a second region critical for gonad-specific *SOX9* expression by using the same approach in four families with heterozygous deletions upstream of *SOX9* associated to 46,XY DSDs with no skeletal malformations (Kim et al., 2015). This region was termed XY sex reversal region (XYSR), as its deletion causes male-to-female sex reversal and lies 607-640 kb upstream of the *SOX9* promoter.

1.6.3. Higher-order chromatin conformation at the SOX9 locus

Cis-regulatory elements are able to exert their function over long distances by chromatin looping, which brings the regulatory element in contact with the promoter of its target gene. Several technologies have been developed that analyze the 3D chromatin organization within the nucleus and can identify enhancer-promoter interactions. These methods are known as chromatin conformation capture technologies and depend on chemical cross-linkers that "freeze" interactions between distal loci. To identify these interactions, the cross-linked chromatin is fragmented and the loose ends (that correspond to the distinct interacting loci) are ligated to each other. These ligated fragments are then amplified and mapped back to the genome, with half of the fragment mapping to promoter region, and the other half mapping to an interacting element.

Smyk and Szafranski et al. performed chromosome conformation capture-on-chip (4C) using the *SOX9* promoter as bait to identify long-range interactions in a Sertoli cell line that might regulate sex determination in humans (Smyk et al., 2013). Surprisingly, although several interacting regions upstream and downstream of *SOX9* were identified, these regions did not

encompass *RevSex*. It is possible that this is because *RevSex* does not interact with the SOX9 promoter but rather with other regulatory elements. Alternatively, *RevSex* may interact only transiently with *SOX9* during the initial stages of sex determination leading to its upregulation*,* whereas other enhancers may be required for *SOX9* maintenance during adulthood. Consistent with this hypothesize, ChIP-seq in human fibroblasts from a 46,XX DSD patient harboring a *RevSex* duplication found that *RevSex* is enriched for the repressive H3K27me3 and H3K9me3 histone marks. These marks are indicative of a closed chromatin conformation making it unlikely that this region is active at a differentiated state. However, the *RevSex* duplication was probably active early on, leading to *SOX9* upregulation in absence of *SRY* and female-to-male sex reversal (Lybaek et al., 2014). In contrast to *RevSex*, the region spanning from TESCO down to the *SOX9* promoter had higher levels of the active H3K4me3 histone mark in XY than in XX individuals, suggesting that TESCO might be important for maintaining normal levels of *SOX9* during adulthood (Lybaek et al., 2014).

Enhancers and their target genes are contained into subchromosomal structures of higher-order chromatin called topologically associated domains (TADs). Binding of CTCF insulators function as boundaries that delimit each TAD, creating discrete genomic units that restrict enhancer-promoter interactions (Nora et al., 2012, Dixon et al., 2012). A more novel chromosome conformation capture technique known as Hi-C is able to identify genome-wide interactions and TADs. A recent study used Hi-C to identify that the mouse and human *Sox9/SOX9* locus is compartmentalized into two major TADs, one of which contains the Sox9 gene and most of the 2Mb gene desert including *RevSex*, and another upstream TAD containing two potassium channel genes *KCNJ2* and *KCNJ16* (Franke et al., 2016). Astoundingly, although duplications overlapping *RevSex* can cause female-to-male sex reversal, if these duplications span further upstream towards the neighboring genes *KCNJ2* and

KCNJ16, sexual development is not affected but limb development is. This is due to the fact that DSD-associated duplications remain within the *SOX9* TAD, increasing interactions between gonad-specific enhancers and the *SOX9* promoter. In contrast, larger duplications that span farther upstream disrupt the TAD boundary and create novel interactions between *KCNJ2* and *SOX9.* This leads to *KCNJ2* mis-expression in the limbs, probably through regulatory elements originally belonging to *SOX9*. This finding demonstrates that TADs are stable genomic units that determine the regulation of a gene and the pathogenic effect when disrupted.

Interestingly, four of the regions identified by 4C that interact with the *SOX9* promoter overlap long non-coding RNAs (lncRNAs) (Smyk et al., 2013). LncRNAs are transcriptional regulators that are defined as non-protein coding transcripts that are longer than 200 nucleotides. Additionally, *RevSex* itself encodes a testis-specific lncRNA that has a higher enrichment of H3K9me3 in XX than in XY individuals (Lybaek et al., 2014). Together, these observations point to lncRNAs as possible epigenetic regulators of *SOX9.*

1.7. *Sry* and other Sox genes as epigenetic regulators

Both SRY and SOX9 are transcription factors (TFs) that recognize and bind a specific DNA sequence (Denny et al., 1992b). Although TFs typically act as direct activators when binding to DNA motifs of target genes, below we will review several studies that support a model in which SRY and SOX9 can themselves function as epigenetic regulators. Despite *Sry* acting as the primary male-determining switch in mice and humans, sequence conservation between the *mSry* and *hSRY* genes is low and restricted to 79 amino acids that encode the HMG box, suggesting that this region is of high functional significance (Sinclair et al., 1990, Ferrari et al., 1992). The HMG box of SRY has an L-shaped structure formed by an extended segment and three α-helices (Read et al., 1993). Contrary to most TFs that bind the major groove of DNA, the

SRY HMG box binds the minor groove of DNA inducing a 60°-70° bend in the double helix (Fig. 1.4A). In addition to DNA bending activity, HMG1 proteins can also bind to DNA cruciform structures, suggesting that SRY could also mediate attachment sites in chromatin loops (Fig. 1.4B) (Ferrari et al., 1992). This ability to directly modulate the 3D architecture of DNA leads to the speculation that SRY and other HMG-box containing TFs, such as the Sox genes, are core determinants of cell-type-specific chromatin landscapes. In support of this, it has been suggested that SRY might mediate effects at a distance by mechanically displacing factors associated with chromatin near the point of flexure. Alternatively, DNA bending might bring remote sites in close contact with each other to facilitate the interaction of transcription factors, or to promote the establishment of chromatin loops (Fig. 1.4A) (Bianchi and Beltrame, 1998). Importantly, the majority of sex-reversing point mutations in the *SRY* gene fall within the HMG box-encoding region (Berta et al., 1990). These mutations cause alterations to the SRY protein that reduce or ablate DNA binding affinity, suggesting that DNA binding and bending is of critical functional importance (Jager et al., 1992, Harley et al., 1992, Pontiggia et al., 1994, Murphy et al., 2001, Assumpcao et al., 2002).

In addition to directly bending DNA and altering the 3D conformation of chromatin, SRY and SOX9 can act as epigenetic regulators by interacting with chromatin-modifying complexes. In E11.5 fetal gonads, SRY interacts with Krüppel-associated box only (KRAB-O) protein and its obligatory co-repressor Krab-associating protein 1 (KAP1) (Oh et al., 2005, Peng et al., 2009). KAP1 then recruits heterochromatin protein 1 (Hp1), HDACs and the SETDB1 methyltransferase, which function as gene silencers by creating a repressive chromatin environment. Therefore, *Sry* may have a dual function in early sex determination by activating male-determining genes through direct binding to regulatory elements (such as binding to the testis-specific enhancer TESCO upstream of *Sox9* (Sekido and Lovell-Badge, 2008)), and

A

B
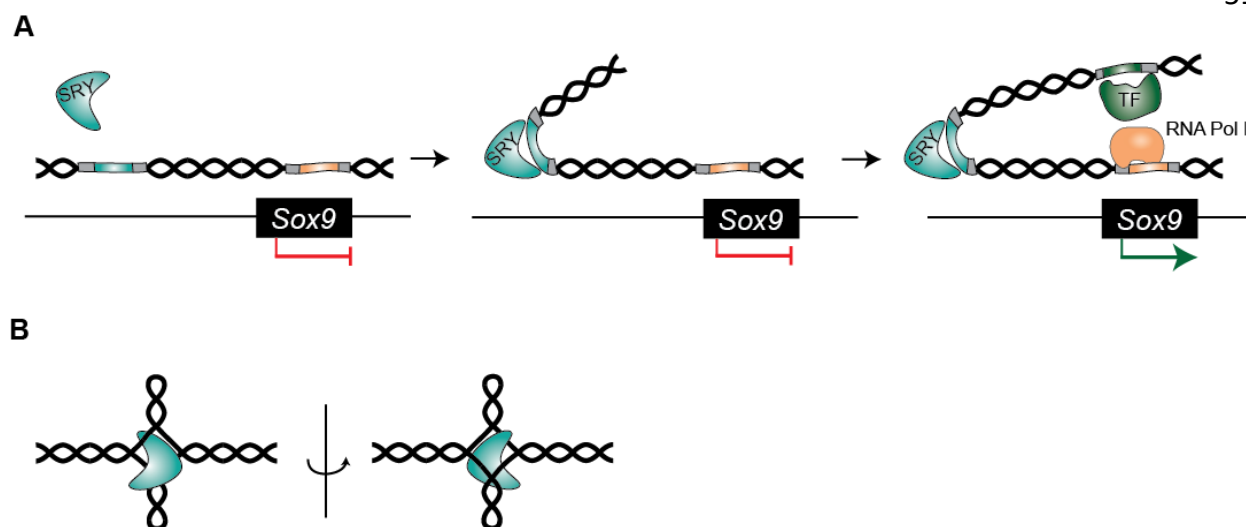
Figure 1.4. Overview of SRY mechanism in mammals. A) *Sry* encodes a transcription factor that recognizes a specific DNA-binding motif (blue box). SRY contains an L-shaped HMG box that binds the minor groove of DNA and induces a 60-70 degree bend in the double helix (middle panel). It is speculated that DNA bending by SRY may mediate effects at a distance by promoting chromatin looping, which brings transcription factors in close proximity to TSS. Depicted is a putative example of SRY-mediated activation of *Sox9* by binding to an upstream enhancer (right panel). B) The HMG box of SRY can bind to DNA cruciform structures regardless of DNA sequence.

repressing the female pathway through recruitment of KRAB-O/KAP1 chromatin-mediated repression machinery. Interestingly, KRAB-O knockdown in *Sry*-expressing cultured cells reduced levels of *Sox9,* suggesting that KRAB-O may also mediate *Sry* function. However, KRAB-O knockdown mice exhibited normal testis development, possibly due to redundancy from over 100 KRAB genes that are expressed in the mouse fetal gonad (Polanco et al., 2009).

In addition to p300/CBP-mediated activation of *Sry* through H3K27ac deposition, SOX9 can itself epigenetically regulate target genes by interacting with this histone acetyltransferase complex (Tsuda et al., 2003)*.* In human chondrocytes, p300/CBP significantly enhanced *Sox9*-mediated transcription of target genes through p300-mediated histone acetylation (Tsuda et al., 2003, Furumatsu et al., 2005). Although these studies were not performed in the developing gonad, it is possible that *Sox9* functions during sex determination by acetylating SOX9 binding

sites and recruiting additional activators to Sertoli-specific genes and regulatory regions. In this way, SRY and SOX9 may play critical roles during sex determination by epigenetically regulating downstream targets that further promote the divergence of the male pathway from the bipotential state.

Although I have limited this introduction to studies of the initial steps in the sex-determination cascade, further studies, including the ones described in chapters 3 and 4, are sure to uncover additional levels of epigenetic regulation that further promote the canalization and maintenance of the male and female pathways. In chapter 3, I explore the role that chromatin plays in maintaining sex-determining genes poised for activation following expression, or lack thereof, of the primary male switch *Sry.* Furthermore, I discuss our finding that silenced genes that promote the alternate pathway remain poised after sex determination, possibly contributing to the remarkable plasticity that enables transdifferentiation of male to female cells and vice versa. Finally, I present a novel function of the chromatin modifier CBX2 in repressing poised female-pathway genes during testis development. In chapter 4, I explore how the chromatin landscapes of male and female progenitor cells are similar at the bipotential stage but become sex-specific as cells acquire accessibility at regulatory elements that enable binding of transcription factors and silencers that promote differentiation of Sertoli or granulosa cells. Finally, I will discuss how our datasets led to the identification of novel enhancers that are critical for the regulation of sex-determining genes during mammalian sex determination.

CHAPTER TWO

MATERIALS AND METHODS

2.1. Mice

The *Sox9-CFP*, *TESm-CFP* and *Sf1-eGFP* reporter mouse lines, and the *Cbx2, Fgf9* and *Wnt4* KO mouse lines previously generated (Kim et al., 2007b, Nef et al., 2005, Katoh-Fukui et al., 1998, Colvin et al., 2001, Vainio et al., 1999) were maintained on a C57BL/6 (B6) background. Timed matings were established between reporter males and ICR females for their larger litter sizes. The morning of a vaginal plug was considered E0.5. Embryos were collected at E10.5/E13.5. and genetic sex was determined using PCR for the presence/absence of a region of the Y chromosome (Table 2.1). The mutant mice were genotyped using the following primers:

Table 2.1. Genotyping primers.

Genotype	Forward Primer	Primer mutation	Reverse Primer
Cbx2	GTAGCCAAGCCAGA GCTGAA	CCGCTTCCATTGCTCAGCGGT	ACCACAGGCCTCTTTGG TGT
Fgf9	GCAAGGGAGGGGAG TTGGATATACC	GAAATCCAGTCCTGCAGTACA GCTGC	CAGCCCAAGCTTTCGCG AGCTCG
Wnt4	CAACAACGAGGCTG GCAGG	CGCATTGTCTGAGTAGGTGTC ATTC	CCCGCATGTGTGTCAAG ATGG
XY	TGAAGCTTTTGGCTT TGAG	N/A	CCGCTGCCAAATTCTTT GG
LacZ	GGCGTTAACTCGGC GTTTCAT	N/A	CGGTTAACGCCTCGAAT CAGC

2.2. Cell Isolation

Gonads were microdissected from E10.5 or E13.5 embryos and removed from the mesonephros. Isolated E10.5 gonads and E13.5 XX gonads were then incubated in 450μl 0.05% trypsin, and E13.5 XY gonads were incubated in 450μl 0.05% trypsin plus 50μl of 2.5% collagenase for 5-7 minutes at 37°C. Trypsin was quenched by slowly adding 100μl of PBS/3%

BSA. Supernatant was removed, and gonads were carefully rinsed with 300μl PBS/3% BSA.

Gonads were then mechanically dissociated in 300μl PBS/3% BSA and the cell suspension was

strained through a 35μm filter top. The tube and filter top were rinsed with an additional 200 μl

PBS/3% BSA. The supporting cells were collected using fluorescence-activated cell sorting

(FACS) into 1.5ml tubes containing 100μl of PBS/3% BSA.

2.3. Adult Sertoli Cell Purification

Sertoli cells were isolated as in (Anway et al., 2003). Testes of >2-month-old B6 males

were dissected and de-capsulated, and placed in a 50ml conical tube with PBS. Testes were

swirled and rinsed twice in PBS. The PBS was decanted and replaced by 25ml of a collagenase

solution (0.5mg/ml collagenase in HBSS). Testes were incubated for 15 minutes in a shaker at

100 oscillations/minute at 34°C to disperse tubules without fragmenting them. After incubation,

the supernatant (containing interstitial cells) was decanted. Tubules were washed 3x in 20ml

HBSS by gentle swirling and allowing tubules to settle, and then incubated in 25ml of a trypsin

solution (0.5mg/ml trypsin in HBSS) for 10 minutes at 37°C in a water bath without shaking.

After two washes in HBSS, tubules were rinsed in PBS/3%BSA to quench the trypsin.

To separate the Sertoli cells from the germ cells, the tubules were incubated in 25ml of

an enzyme solution (0.1% collagenase, 3%BSA, 1ul DNase in HBSS) for 40 minutes in a shaker

at 100 oscillations/minute at 34°C. After incubation, tubules were further homogenized by

pipetting up and down with a p1000. Disaggregated cells were recovered by centrifuging at 500

rpm for 5 minutes. The pellet was subsequently washed 3x in PBS.

To lyse germ cells and increase the purity of the Sertoli cell population, cells were

submitted to a hypotonic shock in a solution of 1:2 HBSS to H_2O (30ml total). The solution was

mixed by inverting the tube twice, and immediately centrifuged at 500 rpm for 5 minutes at 4°C

to recover the Sertoli cells. The cell pellet was then washed twice in 1ml of PBS and resuspended in 500μl of PBS/3% BSA. Cells were then FACS-purified to collect the large-cell population to eliminate contaminating germ cells, nuclei and other debris (Fig. 2.1D).
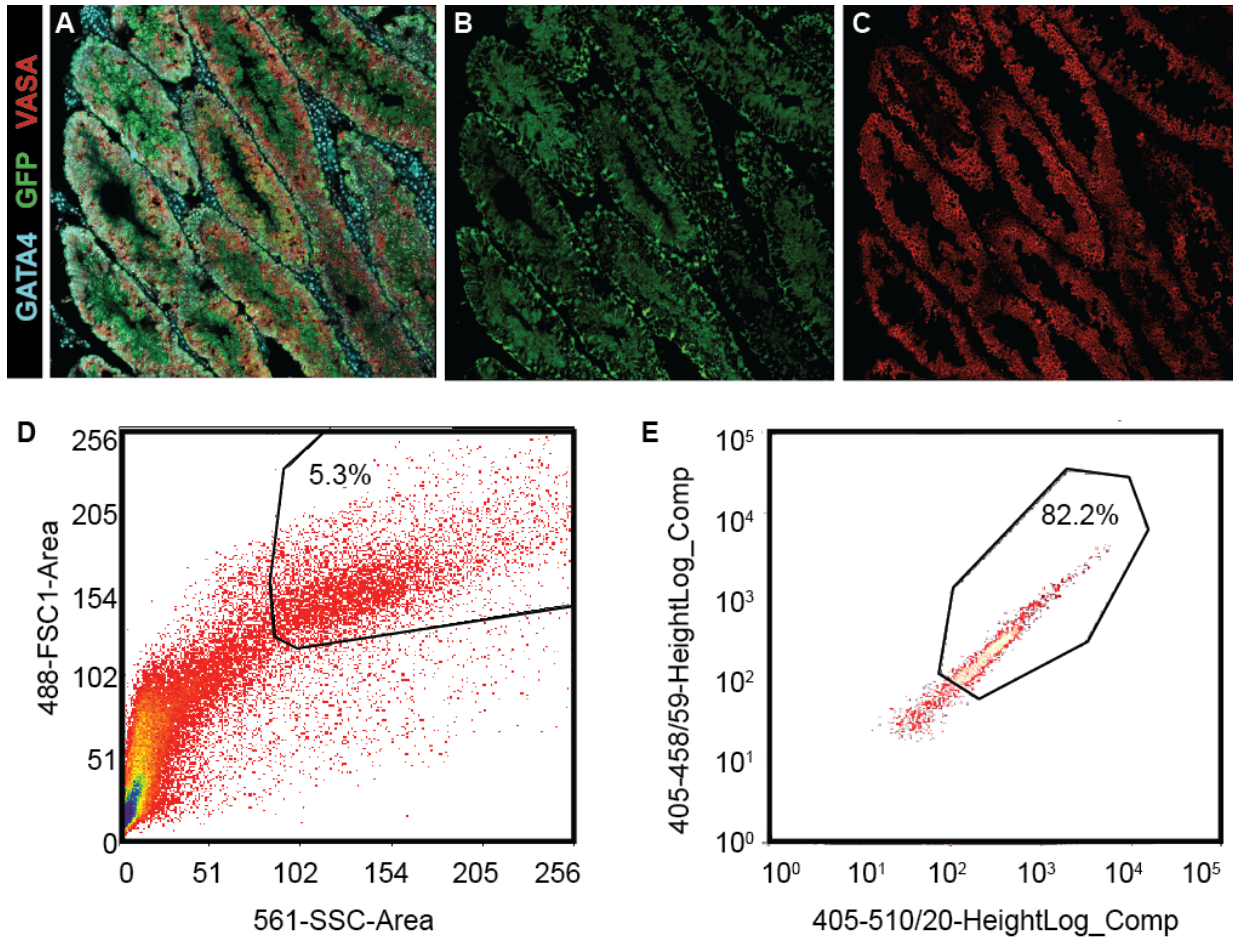
Figure 2.1. Adult Sertoli cell purification. A-C) Immunofluorescent analysis of an adult (>2 month-old) *Sox9-CFP* mouse stained with the supporting cell marker (GATA4, blue), the CFP reporter that marks Sertoli cells (GFP, green) and the germ cell marker (VASA, red). D) Flow cytometry scatterplot of the population of large cells (boxed). E) Flow cytometry scatterplot of *Sox9-CFP* cell correlation to the large-cell population in D. 80% of *Sox9-CFP+* cells fall within the large cell population.

To validate our approach, we purified Sertoli cells as described above from a *Sox9-CFP* adult mouse, which has CFP-expressing Sertoli cells. Immunofluorescent analysis showed that

CFP is still expressed in adult Sertoli cells (Fig. 2.1A and B), and that this marker is not co-expressed in interstitial cells (Fig. 2.1A) nor in germ cells (Fig. 2.1E). Furthermore, ~80% of *Sox9-CFP+* cells fell within the large-cell population collected by FACS after the hypotonic shock solution, suggesting that this population is enriched for Sertoli cells. ~400K Sertoli cells were routinely recovered from 1 adult mouse and used for native ChIP-qPCR as described below.

2.4. Chromatin Immunoprecipitation (ChIP)

2.4.1. ChIP-seq

Multiplexed indexed T7 chromatin immunoprecipitation (Mint-ChIP) was performed with no modifications as in Van Galen *et al.* 2016 on ~30K–100K FACS-purified supporting cells on two biological replicates, each replicate contained cells from pooled gonads of multiple embryos (Fig. 2.2). All buffers and reagents were taken directly from the Mint-ChIP protocol. The protocol is briefly described below.

Day 1: Adapter ligation and ChIP

Timed matings were set up so that E10.5 and E13.5 XY and XX supporting cells could be FACS-purified and processed for ChIP on the same day. FACS-purified cells were resuspended in 20μl of PBS/2xPIC (PIC, Thermo Fisher 1862209). Cells were lysed by adding 20μl of *2x Lysis Buffer* with 150u on MNase (NEB #M0247S). and incubating on ice for 20 minutes (the optimal units of MNase was determined by an MNase titration on FACS-purified cells shown in Fig. 2.2). Digest was done at 37°C for 10 minutes. In parallel, 400K *Drosophila* S2 cells were also resuspended in 20μl of PBS/2xPIC, lysed and digested to use as carrier chromatin. After digest, 37μl of *Enzyme Mix* and 4μl of each adapter barcode (Ad_BC##) were

quickly added to each sample while on ice (Table 2.2). At this stage, each unique barcode

represents one sample: XX E10.5, XY E10.5, XX E13.5 and XY E13.5. Adapter ligation was

performed by incubating at room temperature for 2 hours. Reaction was ended by adding 80μl

of *Lysis Dilution Buffer.*

All barcoded samples and un-indexed carrier chromatin were then pooled and split for

ChIP. Volume for each ChIP was brought up to 200μl and incubated overnight at 4°C on tube

rotator with either 3μl of H3 antibody, 3μl of H3K4me3 antibody, 5μl H3K27me3 antibody or 2μl

of H3K27ac (Table 2.3).
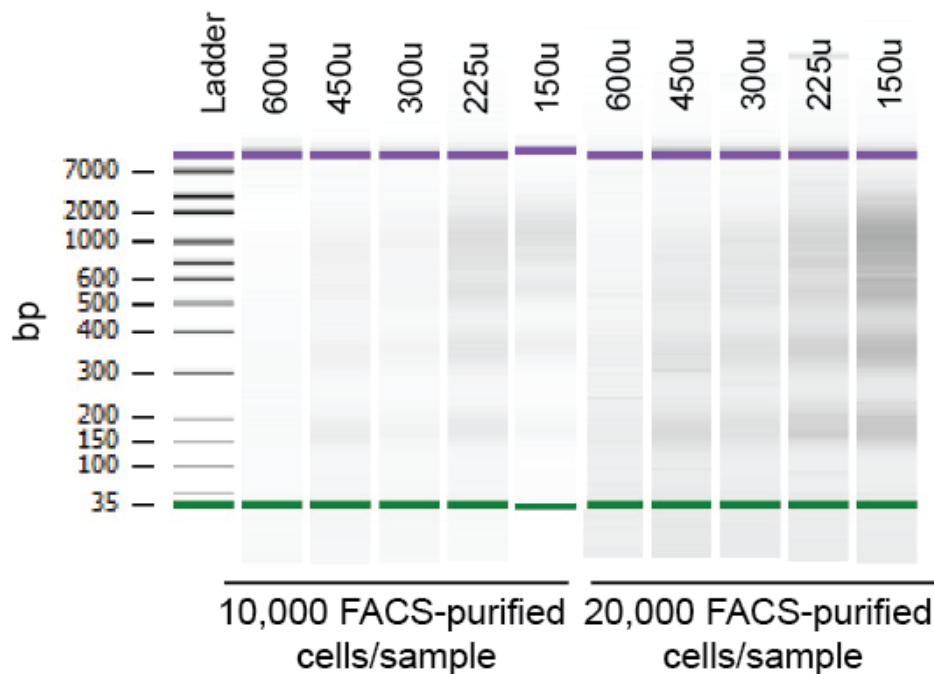
Figure 2.2. MNase titration on FACS-purified cells. MNase units for chromatin digest was optimized on 10K (left 5 lanes) and 20K (right 5 lanes) FACS-purified Sertoli cells. Fragment size was analyzed on a biolanalyzer. 150u of MNase was selected as the optimal concentration due to fragment size distribution and higher final chromatin concentration.

Day 2: In vitro transcription

25µl of protein G Dynabeads (Thermo Fisher 10004D) per sample were magnetized and washed 1x in 25µl of *Lysis Dilution Buffer*. 25µl of washed beads were then added to each ChIP and incubated for 3 hours at 4°C on a tube rotator. After binding, beads were washed 2x in 200µl of ice cold *RIPA*, 1x in 200µl of ice cold *RIPA/High Salt Buffer*, 1x in 200µl of ice cold *LiCl Wash Buffer,* and 1x in ice cold TE. Beads were resuspended in 97.5µl of *ChIP Elution Buffer* and 2.5µl of 10mg/ml protK. DNA was eluted from beads by incubating at 63°C for 1 hour.

Using a 96-well plate magnet, beads were removed and eluted DNA was transferred to new wells and subjected to a SPRI bead cleanup to isolate DNA. 100µl of SPRI beads were added to 100µl of eluate, mixed and incubated for 10 minutes at room temperature. SPRI beads were then magnetized for 5 minutes and washed 2x in 70% EtOH. Beads were air-dried for 5 minutes and resuspended in 8µl EB. 1µl was used to quantify chromatin using a Qbit. 10µl of T7 RNA polymerase mix (NEB HiScribe T7 kit) were added to the 8µl of purified DNA, mixed, and incubated at 37°C for 16 hours for *in vitro* transcription from DNA to RNA.

Day 3: Library preparation

1µl of DNase was added to each sample and incubated at 37°C for 15 minutes. 15µ/sample of Silane beads (binds up to 3µg of RNA) were washed in 60µl of *RLT Buffer,* resuspended in 60µl of *RLT Buffer*, and added to each sample. Beads were incubated for 1 minute at room temperature. 48µl of 100% EtOH was mixed into each sample, and incubated for an additional 5 minutes. Beads were then magnetized and washed 2x in 70% EtOH. Beads were air-dried at room temperature for 2 minutes and RNA was eluted by adding 9µl of *TE Buffer* to beads. 1µl was used for RNA quantification on a Qbit.

For reverse transcription (RNA to cDNA), 3µl of the primer PvG748 and 1.5µl of annealing buffer were added to 8µl of RNA. The adapter was annealed to the RNA by incubating at 65°C for 5 minutes and immediately placed on ice for 1 minute. Then, 15µl of 2x First-strand Reaction Mix and 3µl of SSIII/RNaseOUT Enzyme Mix were added to each sample and placed in a thermocycler using the following program: 25°C for 10 minutes, 50°C for 50 minutes, 85°C for 5 minutes and held at 4°C. 1µl of RNase Cocktail (Invitrogen AM2286) was added to each sample and incubated for 15 minutes at 37°C to remove remaining RNA. cDNA was purified by first adding 68µl of H_2O to bring volume up to 100µl and then adding 100µl of SPRI beads. cDNA was purified as described above (day 2). 1µl was used to quantify single-stranded cDNA on a Qbit.

1-200ng of cDNA was brought up to 22.5µl with H_2O, and was added to 25µl of PFU Ultra II HS 2x Master mix and 2.5 of a PCR barcode (PCR_BC##). At this stage, each unique PCR barcode represents a histone modification used for ChIP: H3, H3K4me3, H3K27me3 or H3K27ac. Adapter modified fragments were then amplified by PCR using the program: 95°C for 3 minutes, 4 cycles of 95 °C for 15 seconds, 65 °C for 30 seconds and 72 °C for 30 seconds, 10-12 cycles of 95 °C for 15 seconds and 72 °C for 60 minutes, and finishing with 72 °C for 5 minutes and held at 4 °C. Samples were quantified and subjected to 75bp, single-end Next-Generation Sequencing on an Illumina sequencer. After sequencing, samples were demultiplexed by unique barcode combination. The adapter barcodes and PCR barcodes used for Chapter 3 are presented in Table 2.2.

2.4.2. Native ChIP-qPCR for histone modifications

ChIP-qPCR was modified from the Mint-ChIP protocol described above using the same buffers and reagents (Van Galen *et al.* 2016) (Fig. 2.2). Briefly, FACS-purified cells or 400K

Table 2.2. Barcodes and adapters used for Mint-ChIP.

PCR Barcode		Adapter Barcode		Sample name	Pooled Sample Name
PCR_06	AGGTGCGA	Ad_BC02	CATGCTTA	H3_SF1M	
		Ad_BC03	GCACATCT	H3_SF1F	H3
		Ad_BC04	TGCTCGAC	H3_Sox9	Rep 1
		Ad_BC05	AGCAATTC	H3_TESm	
PCR_07	CCAGTTAG	Ad_BC02	CATGCTTA	H3K4me3_SF1M	
		Ad_BC03	GCACATCT	H3K4me3_SF1F	H3K4me3
		Ad_BC04	TGCTCGAC	H3K4me3_Sox9	Rep 1
		Ad_BC05	AGCAATTC	H3K4me3_TESm	
PCR_08	TTGAGCCT	Ad_BC02	CATGCTTA	H3K27me3_SF1M	
		Ad_BC03	GCACATCT	H3K27me3_SF1F	H3K27me3
		Ad_BC04	TGCTCGAC	H3K27me3_Sox9	Rep 1
		Ad_BC05	AGCAATTC	H3K27me3_TESm	
PCR_02	CATGCTTA	Ad_BC02	CATGCTTA	H3_SF1M	
		Ad_BC03	GCACATCT	H3_SF1F	H3
		Ad_BC06	AGTTGCTT	H3_Sox9	Rep 2
		Ad_BC08	TTGAGCCT	H3_TESm	
PCR_08	TTGAGCCT	Ad_BC02	CATGCTTA	H3K4me3_SF1M	
		Ad_BC03	GCACATCT	H3K4me3_SF1F	H3K4me3
		Ad_BC06	AGTTGCTT	H3K4me3_Sox9	Rep 2
		Ad_BC08	TTGAGCCT	H3K4me3_TESm	
PCR_05	AGCAATTC	Ad_BC02	CATGCTTA	H3K27me3_SF1M	
		Ad_BC03	GCACATCT	H3K27me3_SF1F	H3K27me3
		Ad_BC06	AGTTGCTT	H3K27me3_Sox9	Rep2
		Ad_BC08	TTGAGCCT	H3K27me3_TESm	

Drosophila S2 cells were resuspended in 20µl 2xPIC/PBS and 20µl 2x *Lysis Buffer* containing 150u of MNase. Cells were lysed on ice for 20 minutes and chromatin was digested for 10 minutes at 37°C. The digest was stopped by adding 40µl *Dilution Buffer + EGTA* to a final concentration of 25mM and incubated on ice 10 minutes. The FACS-purified cells and the S2

cells were pooled, and the sample was brought up to 200µl/IP with *Dilution Buffer + PIC*.

Samples were then split and incubated with 5µl of H3K4me3 antibody, 5µl H3K27me3 antibody

and 5µl IgG (Table 2.3) as a control overnight at 4°C on a tube rotator.

The next day, 25µl of protG Dynabeads were washed 2x in *Dilution Buffer + PIC,* added

to each IP, and incubated for an additional 3hrs at 4°C on a rotator. Samples were magnetized,

and the unbound portion was recovered. Beads were washed 2x in 200µl of ice-cold *RIPA*, 1x in

200µl ice-cold *RIPA/High Salt Buffer*, 1x in 200µl ice-cold *LiCl Wash Buffer* and 1x in 100µl of

TE by moving tubes from the front of the magnet to the back twice. Washed beads were

resuspended in 100µl *ChIP Elution Buffer*. Beads and unbound samples were then incubated

with 0.5µl of RNase Cocktail and 0.5µl of protK (10mg/ml) at 37°C for 30 minutes and 63°C for 1

hour. DNA was purified using a MinElute Qiagen Kit (28004) and used for qPCR. qPCR was

analyzed as Ratio Bound/Unbound=1/2^(Bound CT–Unbound CT), and IgG values were

subtracted. Each ChIP-qPCR was performed three times, on pooled cells from multiple

embryos. Primers used for ChIP-qPCR are listed in Table 2.4. qPCR was run at 58°C instead of

60°C.

Table 2.3. ChIP antibodies.

Target	Catalog Number	µl/IP
H3	Active Motif 39763	3
H3K27me3	Cell Signaling Technologies 9733S	5
H3K4me3	Active Motif 39159	3
H3K27ac	Active Motif 39133	2
Cbx2	Bethyl A302-524A	5
IgG	Cell Signaling Technologies 2729S	3-5

2.4.3. Cross-linked ChIP-qPCR for CBX2

Buffers

- *Lysis Buffer (LB):* 0.5% IGEPAL, 150mM NaCl, 10mM Tris-HCL pH 7.5, 1mM EDTA.

- *RIPA:* 0.1% DOC, 0.1% SDS, 1% Triton X-100, 10mM Tris-HCL pH 8, 1mM EDTA, 140mM NaCL

- *RIPA/High NaCl:* 0.1% DOC, 0.1% SDS, 1% Triton X-100, 10mM Tris-HCL pH 8, 1mM EDTA, 360mM NaCl.

- *LiCl Buffer:* 250mM LiCl, 0.5% IGEPAL, 0.5% DOC, 1mM EDTA, 10mM Tris-HCL pH 8.

- *TE:* 10mM Tris-HCL pH 8, 1mM EDTA.

- *Elution Buffer (EB):* TE with 300mM NaCl and 0.1% SDS.

All buffers were supplemented with a final concentration of 1xPIC and 0.01M PMSF to inhibit proteases on day of use. All reagents are the same ones as used above.

Day 1- Chromatin shearing

~50 E13.5 and E14.5 gonads (previously fixed in 1% PFA for 10 minutes and stored at -80°C) were collected in single tube by resuspending in 200μl *Lysis Buffer/PIC*. Gonads were homogenized with plastic pestle that fits 1.5ml tubes, and pestle was rinsed with 100μl *Lysis Buffer/PMSF/PIC*. Suspension was pipetted up and down with a p200 to further break the tissue apart. Gonads were then incubated on ice for 30 minutes to lyse cells. Every 10 minutes, the suspension was pipetted with a p200 to further homogenize the tissue and help cell lysis.

After cell lysis, nuclei were collected by centrifuging at 500rpm for 5 minutes at 4°C. The supernatant was removed, and the nuclei pellet was resuspended in 300μl *RIPA/PMSF/PIC*. Nuclei were then lysed on ice for 20min. Chromatin was then sheared in a Biorupter for three rounds of 15 minutes at High with cycles of 30s on/30s off (45 minutes total). Sheared

chromatin was centrifuged at 13,000rpm for 20 minutes at 4°C to pellet debris. The supernatant containing the sheared chromatin was recovered and transferred to a fresh 1.5ml tube. 30μl (10%) was taken as Input and the rest was snap-frozen and stored at -80°C.

Input was brought up to 100μl with *ChIP Elution Buffer* and transferred to a 96 well plate. 0.5μl of RNase Cocktail and 0.5μl of protk 10ug/μl was added to the input, covered with optical tape and incubate in thermocycler at 37°C for 30 minutes for RNase digest and 4 hours at 63°C for protK digest and reverse crosslinking. Input DNA was purified using a Qiagen Minelute PCR purification Kit and eluted in two rounds of 10μl to recover as much DNA as possible (final 20μl). Chromatin was quantified using a nano-drop, and ~0.5μg were run on a ~1.8% agarose gel + Ethidium Bromide next to 5μl 100bp ladder to check for shearing efficiency. Shearing should result in enrichment of 100-400bp fragments before proceeding with ChIP (Fig. 2.1). Input was stored at -20°C for later use.

Day 2- ChIP

ProtG Dynabeads were vortexed for 30 seconds until fully resuspended. 25μl of beads were added to two separate 1.5ml tubes. Beads were magnetized and the supernatant was removed, then washed 2x with *RIPA.* Washed beads were resuspended in 200μl RIPA and 5μl CBX2 antibody (Table 2.3) or 5μl of IgG antibody (Table 2.3) as control for unspecific binding. Beads were incubated on a tube rotator at 4°C for 3 hours prior to ChIP.

After incubating, the bead-antibody complex was magnetized and the supernatant containing excess antibody was removed. The bead-antibody complex was washed once with 200μl *RIPA*, and resuspended in 50μl *RIPA/PMSF/PIC* by pipetting gently. In the meantime, sheared chromatin from the previous day was thawed and split so that each tube contained ~20μg of chromatin (~150μl). The 25μl of bead-CBX2 complex were added to one tube of

sheared chromatin (200μl total) and 50μl bead-IgG complex was added to the other tube of sheared chromatin from the same sample (200μl total). Bead-antibody-chromatin complex was then incubated overnight at 4°C on a tube rotator.
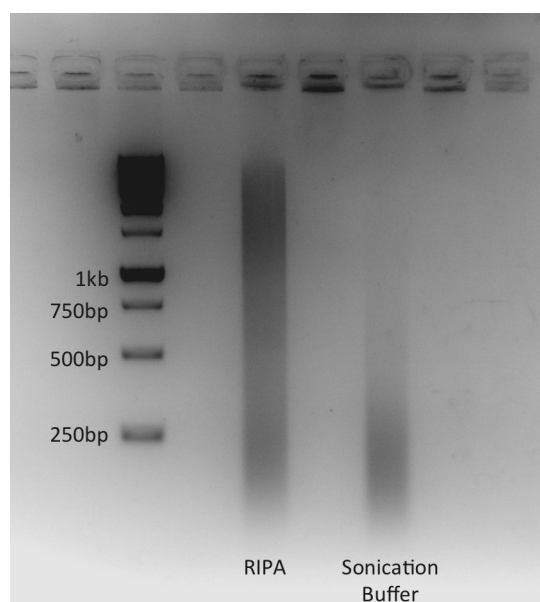
Figure 2.3. Chromatin shearing optimization. One whole adult testis was divided in 2 and processed for sonication. Half was sheared in 300μl RIPA and the other half in 300μl Sonication Buffer. Both were sonicated at the same time in a Biorupter Sonicator for 3 rounds of 15min on High, 30s on/ 30s off. Samples were then incubated with 1μl RNase Cocktail for 30min at 37°C, then with 1μl ProtK for 2hrs at 55°C and reverse crosslinked ON at 65°C. Chromatin was quantified on Nanodrop and 1μg from each condition was loaded on an agarose gel. The rest was used for ChIP. ChIP was unsuccessful on chromatin sheared in Sonication Buffer, probably due to the high percentage of SDS (1%SDS). ChIP in chromatin sheared in RIPA buffer (0.1% SDS) was successful.

Day 3- Washing and reverse crosslinking

The beads-antibody-chromatin complex was magnetized, and the supernatant was removed. Beads were washed 2x in 200μl ice-cold *RIPA*, 1x in 200μl ice-cold *RIPA/High Salt Buffer*, 1x in 200μl ice-cold *LiCl Wash Buffer*, and 1x in 200μl ice-cold *TE buffer* by moving tube from front-to-back twice for each wash. Washed beads were resuspended in 100μl *ChIP Elution*

Buffer and transferred to a 96 well plate. 0.5µl of RNase Cocktail and 0.5µl of protk (10µg/µl) were added to each sample and incubated in thermocycler at 37°C for 30 minutes for RNase digest and 4 hours at 63°C for protK digest and reverse crosslinking. Beads were gently pipetted before incubating. After incubation, beads were gently pipetted to resuspend them, then magnetized on 96 well plate magnet. The supernatant now containing ChIPed DNA was recovered and placed in a fresh tube for the purification step.

Day 4- DNA purification and qPCR

Input DNA was purified using the Qiagen Minelute PCR purification Kit and eluted in two rounds of 10µl to recover as much DNA as possible (final 20µl). 70µl nuclease-free H2O were added to the CBX2 ChIP and IgG ChIP (each in 20µl of EB) to bring up to a final volume of 90µl. 35µl of nuclease-free H2O was added to 10µl of Input (thawed on ice) to bring up to a final volume of 45µl. At this point, all samples and inputs are at the same dilution. All samples for briefly vortexed and centrifuged down before RT-PCR.

RT-PCR was run with input at a 1:100 dilution in nuclease-free H_2O, and SYBR Green was used as a reporter. qPCR primers are listed in Table 2.4. qPCR was run at 58°C instead of 60°C. ChIP was analyzed as % by first adjusting mean Input CT values by subtracting the log2 of the dilution factor used (Log2 of 100 = 6.64). Adjusted Input CT=Input CT- 6.64. % Input=100*2^(Adjusted Input CT – IP CT).

2.5. Assay for Transposase Accessible Chromatin (ATAC-seq)

ATAC-Seq libraries where prepared by Christopher Futtner, Ph.D., from approximately 60,000 FACS-purified cells as previously described with no modifications (Buenrostro et al.,

2013) (Fig. 2.2). ATAC-seq was performed on two biological replicates, with each replicate containing pooled FACS-sorted cells from gonads of multiple embryos.

Table 2.4. ChIP-qPCR primers.

Gene	Forward Primer	Reverse Primer
Axin2	ACCTACTCACTTCCATTCCCC	TCCTTTCCACATTCTCCCCAA
Bmp2	GACTTCTTGAACTTGCCG	TGTTGCTTTTCTTCGCCTCC
Dmrt1	TGCCAGTCTCTGTTAGCCAA	CTACAAACCTCAGCCGTGTG
Fgf9	ATCTGACTCACACCCAACT	AGCAGCAGCCCGAAGACATT
Foxl2	ATACGAATCAGAACGGAGCG	TCGGTGGGTTTTCTTGGC
Fst	AGAGAGAGAGAGGAGTCG	TCCACAAGTCAGAAGCAA
Fzd1	ACACGCACATACACATACACCT	TTCCTCGCCAGCCACTGA
Gapdh	TCCTATCCTGGGAACCATCACC	TCTTTGGACCCGCCTCATTT
Hoxd13	TGGGCTATGGCTACCACTTC	GACACTTCCTTGGCTCTTGC
Lef1	GCTCCCAGGTTCTACAGATGGC	AAAATTCTCCGGTTCCCACTGTC
Lgr5	CCTCTGGACCACAGGAAGT	AGACCATGACTGCGCTCTG
Oct4	TGGCTGAGTGGGCTGTAAGG	CAAACCAGTTGCTCGGATGC
Rspo1	GAAACTGGTCAGATGCTC	ACCCAGGATGCTTGCTAA
Sox9	TCCTCCCTTTAGCCAACC	AGGCGTCTGGACTTAG
Wnt2b	GCAGTCAGCTAAATGGAGCAGA	AGCTGCTGGTTCCACTTGCTTC
Wnt4	CGTGGGAGAAGTAATAAAAGAA	CTTAGGAACTGGAAGGCTGTG
Wnt5a	GCACTAACCCATGAGACATTG	TGTGAAGTTTAGGTGAACCG
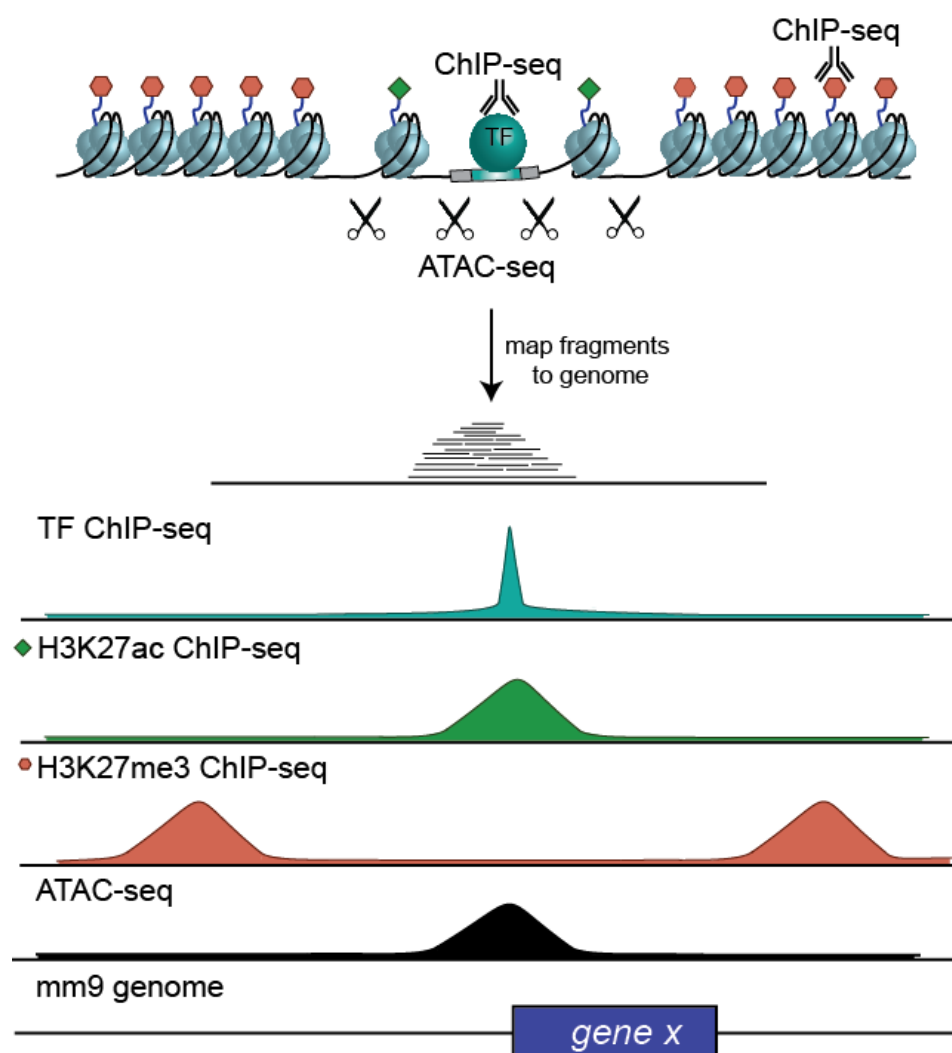
Figure 2.4. Diagram of ChIP-seq and ATAC-seq. Briefly, ATAC-seq identifies regions of nucleosome-depletion when the hyperactive transposase Tn5 fragments and ligates adapters into naked DNA in a single step. Adapter-ligated fragments are then amplified and sequenced, and mapped back to the genome (here, sequencing was mapped back to the mouse mm9 genome, dark blue). Regions of enrichment are shown as "peaks" in ATAC-seq tracks and represent nucleosome-depleted regions (black). ChIP-seq identifies regions of the genome that are either bound by a TF or protein, or identifies histone modifications by using an antibody that specifically binds to the target of interest. ChIP-seq for TFs produces tracks with narrow peaks representing the location of TF binding (blue), whereas ChIP-seq for histone modifications typically produces tracks with broad peaks representing large regions of chromatin decorated by the histone modification of interest (for example H3K27ac in green, or H3K27me3 in red). For TFs and other proteins, chromatin must by cross-linked prior to immunoprecipitation to preserve the DNA binding. However, ChIP-seq for histone modifications can be done without chemical crosslinking as DNA is naturally tightly bound to histones.

2.6. Bioinformatics

2.6.1. ChIP-seq

Single-end, 75bp sequencing was used for ChIP samples. For visualization on the UCSC genome browser and reducing error due to low reads (caused by low cell numbers), replicates were concatenated to maximize number of reads. Demultiplexed .fastq.gz files were first analyzed for quality by uploading each file to the FASTQ program, The number of bp that required trimming from the 5' end of each read was determined based on the quality of the sequencing (typically ~7bp) Reads were trimmed using: trim_galore –stringency 5 –clip_R1 7 <file name>. Trimmed files were then unzipped with gunzip <trimmed file>. Once unzipped, the trimmed files were aligned to the mm9 mouse genome was performed using Bowtie with the following function: bowtie –p 4 –S ../index/mm9 <trimmed file> <new file.sam>, where –p is the number of nodes on the server that are used, and -S indicates that the output is a sam file. Samtools was used to convert sam files to bam files with the following function: samotools view –b <file.sam> -o <file.bam>. Files were then ordered by chromosome position using samtools sort <file.bam. <file_sorted>. Next, unmapped reads were eliminated by using the function samtools view <file_sorted.bam> -F 4 –b –h –o <new file_aligned>, were –F 4 stands for unaligned reads.

Significant regions of enrichment (peaks) were identified using the program MACS2/2.1.0. After loading the program (module load MACS2/2.1.0), peaks were called with the command macs2 callpeak –g 1.87e9 –bdg –broad –t <ChIP_for_modification_of_interest_aligned> -c <H3_aligned> -n <new file_peaks>, where -g represents the size of the genome (mm9), --bdg outputs a bedgraph file, the --broad setting is used for histone modifications which have broader enrichment regions than TFs, -t stands for "treatment" and in this case represents H3K27me3, H3K4me3 or H3K27ac, and -c stands for

"control" and represents H3. Finally, peak files were sorted with the function LC_COLLATE=C sort –k1,1 –k2,2n <ChIP_peaks_treat_pileup.bdg > > <new file_peaks_sorted.bdg>, and bigwig files were created using the function bedGraphToBigWig <bdg file> ../chrom.sizes.txt <output.bw>, where the chrom.sizes.txt file for the mm9 genome was directly downloaded from the UCSC genome browser. To upload biwig files to the UCSC genome browser, bigwig files were first uploaded to vault and the url link copied into the custom tracks. To make bed files to overlay the bigwig files, .bw files were transferred to the computer, opened in excel and modified to contain three columns: the first is the name of the chromosome, the second is the position start and the third is the end position. The file was saved as a .txt file and uploaded to custom tracks in the UCSC genome browser.

Although MACS is important to create bigwig files that allow us to visualize the ChIP-seq tracks on the genome browser, I've found that HOMER is better at calling peaks and is easier to manipulate the settings. Therefore, peaks were also called using the findPeaks function with the settings "--style histone", and once again using H3 as a control as described above. The size was set at 5000 for H3K27me3 and 1000 for H3K4me3 and H3K27ac which have narrower peaks. The setting "-C 0" was used in MNase ChIP to disable fold enrichment limit of expected unique tag positions.

To limit analyses only to gene promoters, Bedtools Intersect was used to intersect the ChIP-seq called peak files with a file containing all TSSs annotated in the mm9 genome downloaded from the UCSC genome browser, which was expanded to span 2kb upstream and 2kb downstream of the TSS. Tag counts for each 2kb promoter region were obtained using the HOMER annotatePeaks.pl function and were normalized to the sample's corresponding H3 tag counts. For example, the command annotatePeaks.pl mm9_2kbTSS.txt mm9 –noadj –fragLength 0 –d H3K27me3_TagDir/ H3K4me3_TagDir/ H3_TagDir > Tagcounts.txt was used,

where –noadj turns off normalization to total reads since in this case sequencing was normalized to total H3, and –fragLength 0 eliminates normalization to fragment length and provides raw reads for the 4kb window. Then, "1" was added to all tag counts to eliminate ceros that affect divisions, and H3K27me3 or H3K4me3 tags were divided by the corresponding number of H3 tags in each region. Scatterplots of promoter tag counts normalized to H3 were generated in Excel.

Histograms of average enrichment around the TSS were generated using HOMER's annotatePeaks.pl function with "-size 20000" and "-hist 1000". By using tss mm9, the histogram centers around all TSSs in the genome. To restrict it to a set of TSS (such as Sertoli- or granulosa-specific TSS), the option –list <specific_TSS> was added to the function. All quantitative comparisons between sexes and timepoints were only performed between replicates that were processed together (i.e. between samples in replicate 1 or replicate 2, as these were FACS-purified and processed for ChIP on the same day).

2.6.2. ATAC-seq

ATAC-seq was processed as described as described above, except using input chromatin as a control. Briefly, single-end ATAC-Seq reads were trimmed for quality and to remove adapters using Trimgalore with a stringency setting of 5. Trimmed reads were aligned to the mm9 genome using Bowtie with the parameters -m 1 and --best. MACS2 was used to call peaks using the --nomodel and --broad settings. Peaks were considered to be true positives if they were present in both replicates.

2.6.3. NDR analyses

All NDR analyses were performed with HOMER. Categorization of ubiquitous NDRs was determined by identifying overlapping NDRs using the mergePeaks function between our datasets and previously published DNaseI-seq data performed in mouse fibroblasts, ESCs, heart, kidney, brain and liver (Maatouk et al., 2017). The remaining non-overlapping NDRs were considered gonad-specific NDRs. MergePeaks was then used between gonad-specific NDRs and H3K27ac ChIP-seq to identify gonad-specific H3K27ac+ NDRs. MergePeaks was also used to identify resolved and *de novo* NDRs within our datasets. The annotatePeaks.pl function was used to generate a list of genes associated to each NDR (each NDR is associated to its nearest TSS), as well as their genomic allocation and distance to the TSS. NDRs were categorized into intergenic, intronic, exonic (exon, 5' UTR and 3' UTR) or promoter (TSS and promoter). To determine the average number of NDRs associated to Sertoli-specific or granulosa-specific genes, we first identified Sertoli-specific genes (genes expressed >1.5 fold higher in E13.5 Sertoli cells compared to E13.5 granulosa cells), and granulosa-specific genes (genes expressed >1.5 fold higher in E13.5 granulosa cells compared to E13.5 Sertoli cells) from a previous microarray study (Jameson et al., 2012b), then counted the number of NDRs associated to each Sertoli- or granulosa-specific TSS. TF motif analyses were performed using HOMER with default settings and the -mask option.

2.7. *In vivo* transient transgenic assays

2.7.1. Cloning the enhancer-reporter plasmids

The putative enhancer was amplified by PCR and cloned into the NotI site of the pSfi-Hsp68–LacZ reporter vector (Addgene #33351). The primers used for the cloning were TgBmp2_DNG/F_inf: <u>ACCGCGGTGGCGGCCGC</u>ATAGAAGATTGCCAGACTCC and

TgBmp2_DNG/R_inf: <u>TCCACTAGTTCTAGA</u>ACAGGGATTCTCTGTATAGC. Cloning was carried out using In-Fusion HD (Clonetech #639646). All plasmids were sequenced using Sanger sequencing to verify the correct sequence was inserted.

2.7.2. Microinjection of the enhancer-reporter plasmids

To prepare DNA for zygote injection, 50 μg of the Enhancer-Hsp68-LacZ plasmid was linearized with NotI-HF and XhoI and gel purified by electroelution. The DNA was phenol-chloroform extracted, ethanol precipitated and resuspended in TE buffer (10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0). The DNA was further purified on a DNA-cleanup column (Qiagen PCR Purification Kit) and eluted again in TE Buffer. 5ng/μl of the purified plasmid was injected into the pronucleus of each C57BL/6Jx CBA hybrid zygote. Embryos were transferred on the same day into oviducts pseudopregnant CD1 recipient mother. Injections and embryo transfer were done by the Crick Genetic Manipulation Service. Embryos were then harvested at E13.5 and stained for β-galactosidase to identify embryos positive for enhancer activity.

2.7.3. X-Galactosidase staining

Embryos were dissected at E13.5 and genotyped for the *LacZ* gene via PCR (Table 2.1). The embryo bodies and dissected gonads were fixed in fresh cold fixative (0.2% Glutaraldehyde, 2% Formahdehyde) in buffer L_0 (100 mM Phosphate buffer pH7.2, 2 mM $MgCl_2$ and 5 mM EGTA) for 30 minutes at 4°C. Embryos were washed three times in L_0 buffer, each wash for 5 minutes. Embryos were then moved to a freshly made and filtered X-Gal staining solution containing 100 mM Phosphate buffer pH7.2, 2 mM $MgCl_2$, 5 mM EGTA, 0.02% NP40, 0.01% Sodium-Deoxycholate, 50 mM $K_3Fe(CN)_6$, 50 mM $K_4Fe(CN)_6·3H_2O$ and 1 mg/ml X-Gal. Incubation in X-Gal staining solution was done at 37°C and varied from a few hours up to

overnight, depending on the strength of the β-Gal expression. The next day, embryos were washed three times in PBS and images of the gonads and bodies were taken.

2.8. Immunofluorescence

Embryonic gonads were dissected and fixed for 30 minutes-1 hour at room temperature in 4% paraformaldehyde. Gonads were washed 3 times in PBS and dehydrated by incubating them in 25% MeOH/PBS, 50% MeOH/PBS, 75% MeOH/PBS and 100% MeOH/PBS for 15 minutes each (in order) at room temperature. Gonads were stored at -20°C in 100% MeOH until staining.

For staining, gonads were rehydrated by incubating in 100% MeOH/PBS, 75% MeOH/PBS, 50% MeOH/PBS and 25% MeOH/PBS for 15 minutes each (in order) at room temperature. Gonads were washed 3 times in PBS at room temperature, then permeabilized for 1 hour in PBS/1% Triton X-100 on a rocker. Gonads were then incubated in *Blocking Buffer* (PBS, 1% Triton X-100, 10% FBS, 3% BSA) for 2 hours at room temperature on a rocker. Finally, gonads were incubated overnight at 4°C in *Blocking Buffer* and the following primary antibodies: SRY, SOX9, FOXL2, GATA4, and PECAM1 (Table 2.5).

The next day, gonad were washed 3 times in *Wash Buffer* (PBS, 0.1% Triton X-100) for 30 minutes at room temperature on a rocker. After washing, gonads were incubated overnight at 4°C in blocking buffer and the following secondary antibodies: Alexa Fluor488-labeled anti-rat or anti-mouse, Alexa Fluor647-labeled anti-goat and Cy3-labeled anti-rabbit (Table 2.5). Incubation was done on a rocker in tin foil to protect from the light. Nuclei were stained with DAPI at 1:500. Whole samples were mounted in DABCO (Sigma-Aldrich) in 90% glycerol and imaged with confocal scanning microscopy using a Leica SP2.

Table 2.5. Antibodies for immunofluorescent analyses.

Target	Catalog Number	Concentration
	Primary antibodies	
FOXL2	Novus Biologicals NB100-1277	1:250
GATA4	Santa Cruz SC-1237	1:100
PECAM1	BD BioSciences	1:250
SOX9	EMD Millipore AB5535	1:3000
SRY	Gift from Dr. Dagmar Wilhelm	1:100
	Secondary Antibodies	
Alexa Fluor 488 anti-rat	Life Technologies	1:500
Alexa Fluor 488 anti-mouse	Life Technologies	1:500
Alexa Fluor 647 anti-goat	Life Technologies	1:500
Cy3 anti-rabbit	Jackson ImmunoResearch 711-165-152	1:500

CHAPTER THREE

CBX2 IS REQUIRED TO REPRESS FEMALE FATE AT BIVALENT LOCI

The work in this chapter was done with contributions from Lin Yi-Tzu, Christopher Futtner, Isabella Salamone, Blanche Capel and Danielle Maatouk.

3.1. Background and Significance

Gonadal sex determination, the process by which the bipotential fetal gonad initiates development as either a testis or an ovary, is the first critical step in the development of sexually dimorphic internal and external reproductive organs. Sex determination initiates with a binary cell fate decision within a single somatic cell lineage of the gonad, known as the supporting cell lineage (Albrecht and Eicher, 2001, Burgoyne et al., 1995). Cells of this lineage are initially held in a bipotential state in the early fetal gonad, poised between male and female fate. In most mammalian species, including mice and humans, expression of the Y-encoded transcription factor *Sex-determining Region Y* (*Sry*) is required to direct the male fate of supporting cells (Sinclair et al., 1990, Gubbay et al., 1990, Koopman et al., 1991). In mice, *Sry* is transiently expressed in XY progenitor cells from ~E10.5-E12.5, soon after the gonad is first formed (Bullejos and Koopman, 2001, Hacker et al., 1995). *Sry's* primary function is to upregulate its downstream target *Sox9* (Sekido et al., 2004). *Sox9* upregulation and subsequent maintenance through *Fgf9* leads to Sertoli cell differentiation and establishment of the testis pathway (Kim et al., 2006b). In the absence of a Y chromosome, the canonical Wnt signaling molecules, *Wnt4* and *Rspo1,* become upregulated in XX progenitor cells (Vainio, 1999, Chassot, 2008). The subsequent downstream stabilization of β-catenin (Maatouk et al., 2008) together with upregulation of other transcription factors such as *Foxl2* (Ottolenghi et al., 2007)*,* leads to the differentiation of granulosa cells in the ovary (Fig. 3.1).

Importantly, upregulation of either Sertoli- or granulosa-promoting pathways is accompanied by mutually antagonistic mechanisms, which are critical for repressing the alternate pathway at the time of sex determination (Kim et al., 2006b, Jameson et al., 2012a). Mapping the up- or down-regulation of each gene in the XX and XY gonad between E11.0 and E12.0, revealed that many genes associated with the female pathway became female-specific

by down-regulation in the XY gonad, while a smaller group of male genes became male-specific by down-regulation in the XX gonad (Munger et al., 2013, Jameson et al., 2012b). This data established the importance of gene repression in the initiation of male (or female) development.

Repression of the alternative fate is actively maintained throughout adulthood. Evidence for this came from a conditional deletion of *Foxl2* in post-natal granulosa cells, which led to their transdifferentiation towards a Sertoli-like state and reorganization of the ovary into testicular tissue (Uhlenhaut et al., 2009). Conversely, conditional deletion of the XY-determining gene *Dmrt1* in post-natal testes, led to transdifferentiation of Sertoli cells into granulosa cells accompanied by ovarian reorganization (Matson et al., 2011). The remarkable plasticity that enables *in vivo* transdifferentiation towards the alternate cell fate led us to posit that gonadal supporting cells retain an epigenetic memory of their bipotential state. Despite increasing evidence that cell fate decisions are epigenetically regulated in many systems (Yellajoshyula et al., 2012, Ikawa et al., 2016, Minoux et al., 2017, Bracken et al., 2006, Pasini et al., 2007), whether epigenetic mechanisms drive the unilateral fate commitment of gonadal bipotential supporting cells towards the Sertoli or granulosa cell fate, and maintain that fate after sex determination, is not well understood.

Loss of two chromatin modifiers, the lysine-demethylase *Jmjd1a* (Kuroki et al., 2013) and the chromobox homolog *Cbx2* (Katoh-Fukui et al., 1998, Katoh-Fukui et al., 2012, Biason-Lauber et al., 2009), cause complete male-to-female sex reversal due to loss of *Sry* expression at E11.5. JMJD1A is a lysine-demethylase that has been shown to directly control *Sry* expression by removing the repressive H3K9me2 histone modification at its promoter during the activation step (Kuroki et al., 2013). Originally, CBX2 was also proposed to be required for activation of *Sry* (Katoh-Fukui et al, 1998; Katoh-Fukui et al, 2012). However CBX2 is a subunit

of the Polycomb group (PcG) of proteins, a large complex of chromatin modifiers (Hashimoto et al., 1998) classically shown to maintain transcriptional repression by catalyzing H3K27me3 at the promoter of its target genes (Kirmizis et al., 2004). Subsequent binding of CBX2 (or other CBX proteins) to H3K27me3 leads to chromatin compaction (Grau et al., 2011, Lau et al., 2017), which limits transcription factor and RNA Pol II accessibility. Whether CBX2 acts as a direct or indirect mediator of *Sry* expression is unknown.

The PcG proteins often work alongside the Trithorax group of proteins (TrxG), which have the opposite role of maintaining transcriptional expression through the deposition of H3K4me3 at active promoters (Santos-Rosa et al., 2002). During development, PcG and TrxG can modify the same target genes. The co-occurrence of an active and a repressive histone modification at the promoter of developmental genes maintains these so-called "bivalent" genes in a poised state competent for rapid activation or repression (Bernstein et al., 2006a, Azuara et al., 2006, Pan et al., 2007). Hence, the balance of the PcG and TrxG proteins acts to fine-tune the timing of gene expression and stabilize cell fate commitment upon differentiation.

With the goal of understanding the epigenetic mechanisms that regulate sex-specific expression of sex-determining genes while retaining a memory of the alternate cell fate, we compared the genome-wide profile of H3K4me3 and H3K27me3 in the supporting cell lineage of the gonad at the bipotential stage (E10.5) and after sex determination has occurred (E13.5). We found that key sex-determining genes are bivalent prior to sex determination, providing insight into how the bipotential state of gonadal progenitor cells is established. Surprisingly, repressed sex-determining genes remain bivalent after sex-determination, and at least in males, bivalency is retained even in adulthood, possibly contributing to the highly plastic nature of the gonad long after the initial fate commitment. We found that many elements of the Wnt signaling pathway are targeted for H3K27me3-mediated repression during testis-development. As Wnt signaling is

known to act as a repressor of male fate (Kim et al., 2006b, Maatouk et al., 2008, Jameson et al., 2012a), we tested the possibility that loss of Wnt signaling could rescue male development in *Cbx2* mutants. We found that *Sry* expression and testis development were rescued in XY *Cbx2*$^{-/-}$;*Wnt4*$^{-/-}$ double knockout mice. Furthermore, we show that CBX2 targets the promoter of the downstream Wnt signaler *Lef1* in testes, suggesting that *Cbx2* induces *Sry* expression indirectly by repressing female-promoting genes. This work suggests a widespread role for the PcG proteins in repressing the alternative female-determining pathway to establish male fate in supporting cell progenitors.

3.2. Results

3.2.1. In vivo *chromatin profiling of gonadal supporting cells*
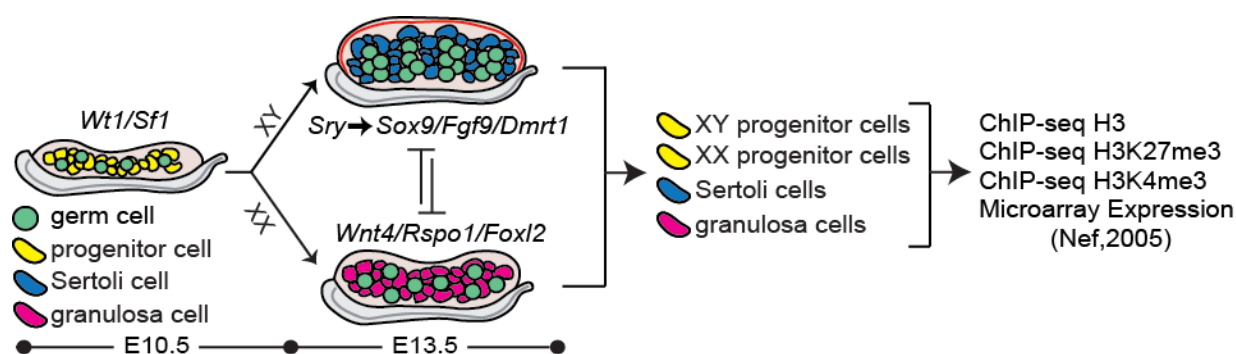
Figure 3.1. Overview of sex determination and workflow. Briefly, supporting progenitor cells (yellow) are bipotential and indistinguishable between XX and XY gonads. Expression of *Sry* directs Sertoli cell differentiation in the testis through upregulation of male-determining genes such as *Sox9, Fgf9* and *Dmrt1*. In the absence of *Sry*, upregulation of female-determining genes *Wnt4, Rspo1,* and *Foxl2,* direct differentiation of granulosa cells in the ovary. XY and XX progenitor cells, Sertoli cells, and granulosa cells were FACS-purified and submitted to ChIP-seq for H3, H3K4me3 and H3K27me3. Further analysis made use of microarray expression data from E10.5 and E13.5 purified supporting cells (Nef, S. et al, 2005).

To investigate the chromatin state of gonadal supporting cells at the time of gonad formation and after their differentiation into either Sertoli cells or granulosa cells, we performed ChIP-seq for H3K4me3, H3K27me3, and for Histone 3 (H3) as a means of normalizing across populations (van Galen et al., 2016). To circumvent the limitation of low numbers of cells obtained from fetal gonads, we used a highly sensitive protocol (van Galen et al., 2016) capable of assaying as few as 500 purified cells per biological replicate. ChIP-seq was performed in 50-100K FACS (fluorescence activated cell sorted) purified XY and XX progenitor cells from E10.5 gonads of *Sf1-GFP* transgenic mice (Beverdam and Koopman, 2006), Sertoli cells from E13.5 gonads of *Sox9-CFP* transgenic mice (Kim et al., 2007b), and granulosa cells from gonads of E13.5 *TESm-CFP* transgenic mice (Fig. 3.1). ChIP-seq was performed on two biological replicates (each replicate containing pooled cells from multiple embryos) and was further validated by ChIP-qPCR (Fig. 3.2 and 3.3).

ChIP-seq revealed that the largest subset of promoters in all cell types were marked only by H3K4me3 (42-45%), whereas a smaller subset was marked by H3K27me3, the majority of which also overlapped with H3K4me3 (Fig. 3.4A). The percentage of bivalent promoters decreased from E10.5 to E13.5 in both XX and XY cells, whereas the percentage of H3K4me3-only and H3K27me3-only promoters increased (Fig. 3.4A and Fig. 3.5 A-B'), suggesting that the histone modifications at many bivalent genes resolve into one mark or the other as they become active or repressed during differentiation.

To investigate the association between these histone modifications and transcriptional activity in the supporting cell lineage, we made use of the microarray gene expression dataset performed by Nef et. al. (Nef et al., 2005) in FACS purified XY and XX supporting cells from *Sf1-GFP* transgenic mice at E10.5 and E13.5 (Beverdam and Koopman, 2006).
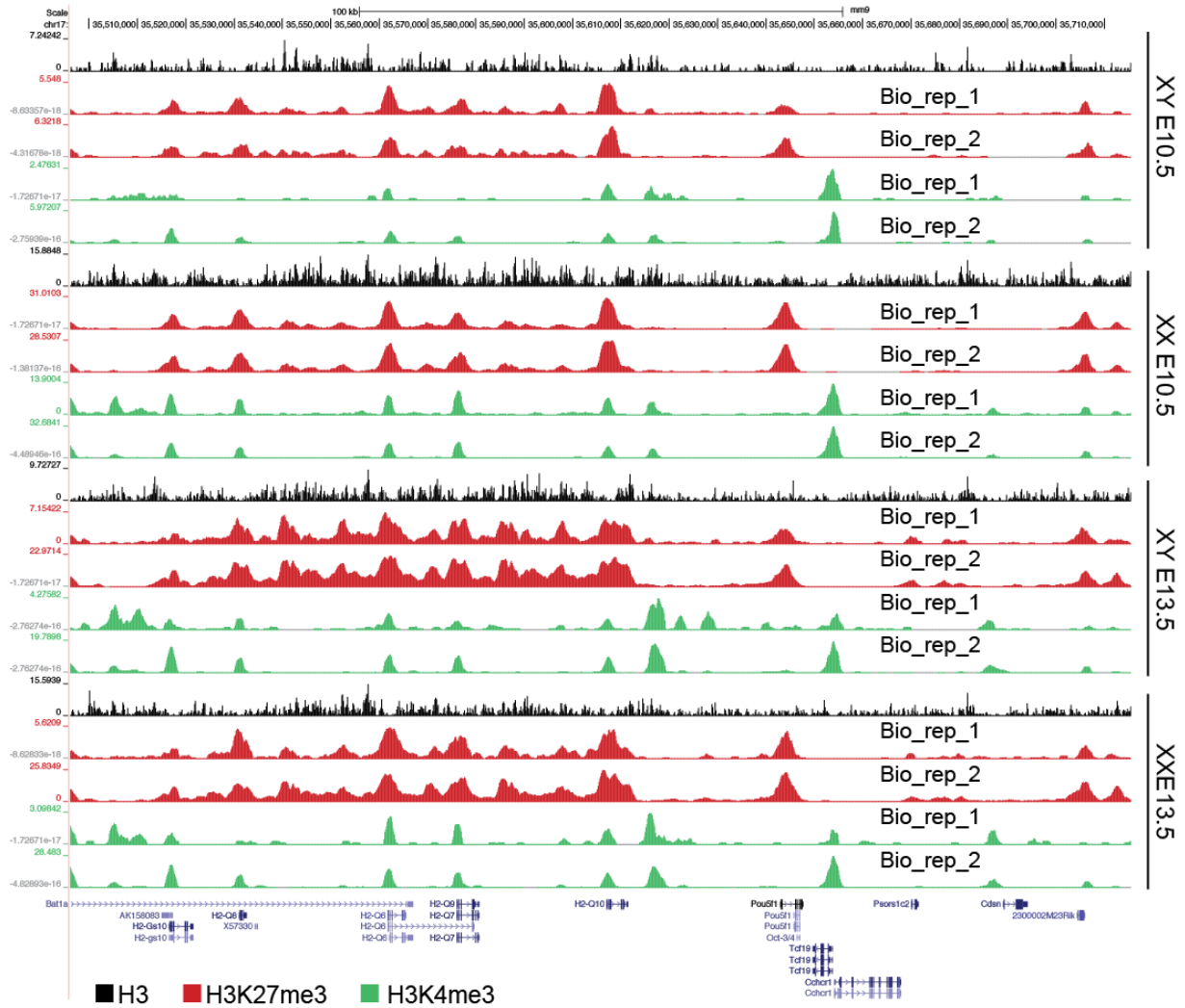
Figure 3.2. ChIP-seq biological replicates are similar. ChIP-seq tracks for H3 (black), H3K27me3 (red) and H3K4me3 (green) for both biological replicates are shown side by side in XY and XX, E10.5 and E13.5 purified supporting cells. Each replicate contains pooled cells from multiple embryos.
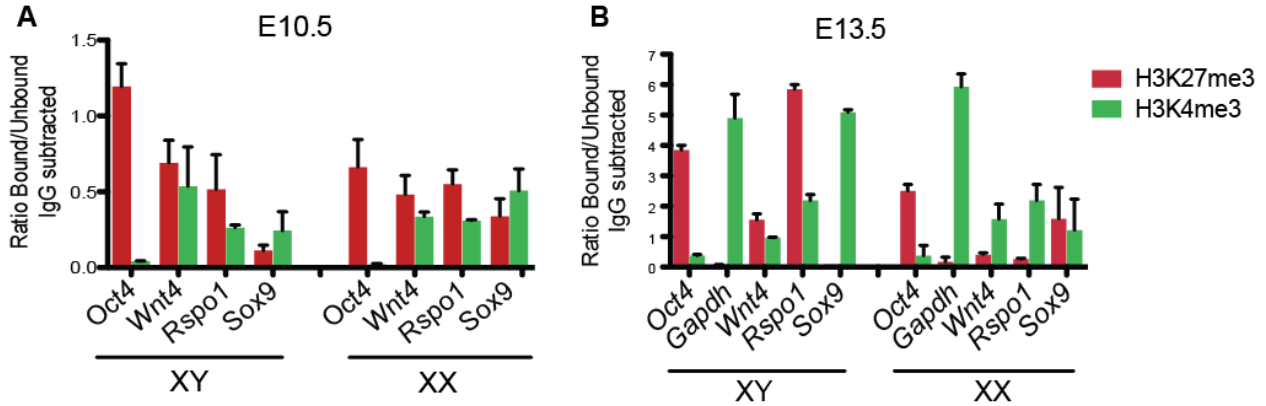
Figure 3.3. ChIP-qPCR validates ChIP-seq. ChIP-qPCR validation of ChIP-seq for H3K27me3 (red) and H3K4me3 (green) in FACS-purified E10.5 XY and XX cells (A) and E13.5 XY and XX (B). Each ChIP-qPCR was performed on 3 biological replicates, each replicate contained pooled cells from multiple embryos.

Previous transcriptional profiling showed that E10.5-E11.0 XY and XX progenitor cells were nearly indistinguishable (Fig. 3.5C), with only a handful of X- or Y-linked genes differentially expressed at this stage (Nef et al., 2005, Munger et al., 2013). Consistent with this, we found that the vast majority of bivalent genes are shared between XY and XX progenitor cells (3545/3788 (XY) and 3545/4561 (XX)) (Fig. 3.5D), suggesting that these cells are not only morphologically and transcriptionally indistinguishable, but also epigenetically similar. The larger number of female-specific bivalent genes is most likely due to the lower number of sequencing reads obtained from our XY-progenitor samples. Differentiation into either Sertoli or granulosa cells is accompanied by the upregulation of a number of Sertoli-promoting or granulosa-promoting genes (Fig. 3.5C'). Consistent with this, there is a reduction in the number of overlapping bivalent promoters between XY and XX cells, from 3545 genes at E10.5 to 2963 genes at E13.5 (Fig. 3.5D'). This suggests that Sertoli and granulosa cells acquire unique subsets of bivalent genes that promote divergence as they differentiate into distinct cell types.
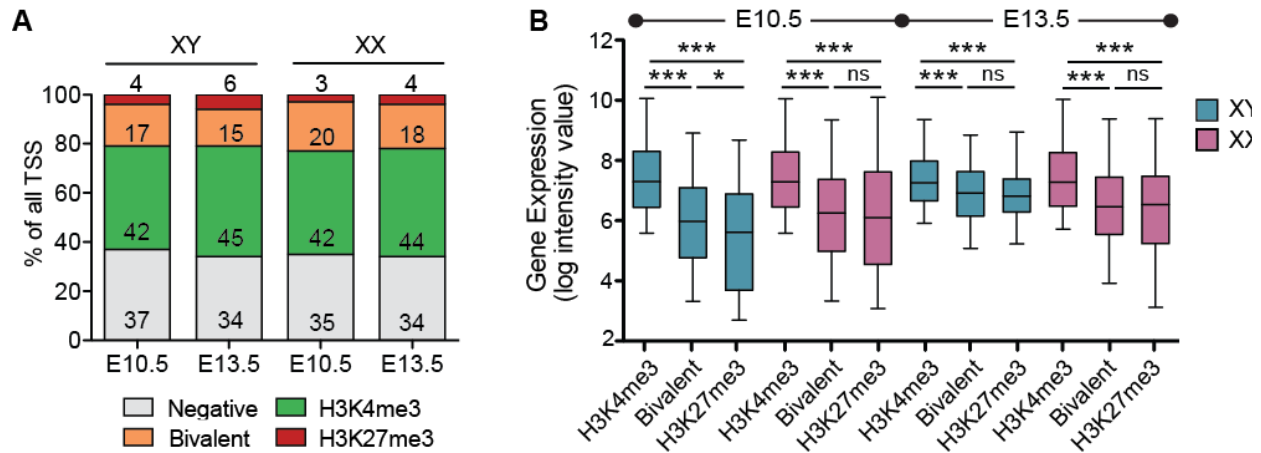
Figure 3.4. Epigenetic profiling of supporting cells during sex determination. A) Percentage of total number of promoters annotated in the mm9 genome marked by H3K27me3-only (red), both H3K27me3 and H3K4me3 (orange), H3K4me3-only (green), or neither (grey). B) Boxplots of gene expression values (log intensity value from Nef et al., 2005) for H3K4me3-only promoters, H3K4me3 and H3K27me3 promoters (bivalent) or H3K27me3-only promoters, in XY (blue) or XX (pink) supporting cells at E10.5 (left) and E13.5 (right) (outliers excluded). *** represents p<0.0001 as determined by student's t test.

3.2.2. Key sex-determining genes are poised at the bipotential stage

Bivalent genes play a crucial role in maintaining pluripotency by fine-tuning the timing of gene expression and ensuring the correct lineage commitment. Embryonic stem cells that lack *Eed* activity, a core component of the PcG proteins, initiate premature differentiation and can differentiate into the wrong lineage (Azuara et al., 2006). The supporting progenitor cells that compose the XY and XX E10.5 fetal gonad remain in a bipotential state until peak expression of *Sry* around E11.5 in males (Sinclair et al., 1990, Gubbay et al., 1990, Koopman et al., 1991), or, in absence of *Sry* in females, accumulation of *Wnt4/Rspo1* and the downstream target β-catenin over a critical threshold (Vainio et al., 1999, Chassot et al., 2008b, Maatouk et al., 2008). Despite extensive transcriptional profiling of progenitor cells at this stage (Nef et al., 2005, Munger et al., 2013), the mechanisms that establish and maintain the bipotential state for ~24hrs are not well understood.
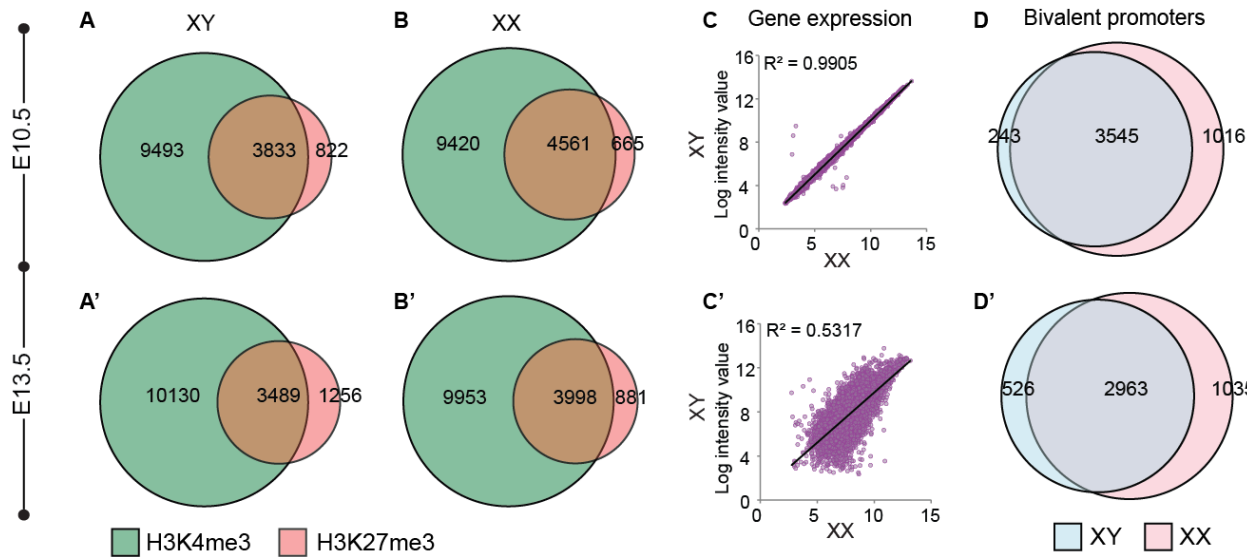
Figure 3.5. Sertoli and granulosa cells acquire sex-specific H3K27me3, H3K4me3 and transcriptional profiles during sex determination. A-B') Venn diagrams depicting number of promoters marked by H3K27me3 (red) and H3K4me3 (green) in XX and XY supporting cells at E10.5 (top) and E13.5 (bottom). B&B') Linear correlation between expression profiles of XY (y axis) and XX (x axis) progenitor cells at E10.5 (top), and between Sertoli and granulosa cells at E13.5 (bottom). D&D') Venn diagrams depicting number of overlapping bivalent promoters between males (blue) and females (pink) at E10.5 (top) and E13.5 (bottom).

We hypothesized that key sex-determining genes that drive either Sertoli or granulosa differentiation are bivalent at E10.5, held in a poised state for either activation or repression following the appropriate sex-determining signal. To investigate this hypothesis, we identified the set of genes that become either Sertoli- or granulosa-specific at E13.5 (Jameson et al., 2012b), surveyed H3K4me3 and H3K27me3 deposition at E10.5, and compared the patterns to control genes whose expression is not specifically associated with the Sertoli or granulosa pathway (Fig. 3.6A and B, and Fig. 3.7A and B). Promoters were differentially enriched for H3K4me3 and H3K27me3 consistent with their expression patterns. For example, the promoter of the constitutively expressed *TATA-Box Binding Protein (Tbp)* was marked by high H3K4me3 and lacked H3K27me3 in both XY and XX supporting cells (Fig. 3.6A and B, and Fig. 3.7A and B). In contrast, the promoter of the germ-cell specific gene *Oct4,* which is repressed in the

supporting cell lineage, is marked by high H3K27me3 and lacks H3K4me3 (Fig. 3.6A and B, and Fig. 3.7A and B). Less than 5% of both Sertoli- and granulosa-specific gene promoters were marked by H3K27me3-only (Fig. 3.6C). In fact, the vast majority harbored some degree of both H3K4me3 and H3K27me3 enrichment, and approximately 30% of Sertoli- and granulosa-specific genes had "high" enrichment (>2.5 log2 enrichment normalized to total H3) of both histone modification marks (Fig. 3.6A-C), suggesting a role for bivalent promoters in maintaining plasticity of supporting progenitor cells to follow either of two fates.

In accordance with our hypothesis, promoters of key XY-determining genes such as *Sox9, Fgf9* and *Dmrt1,* and key female-determining genes *Wnt4, Rspo1* and *Foxl2* were bivalent prior to sex determination in both sexes (Fig. 3.6A and B, and Fig. 3.7A and B). Genes such as *Wt1,* which are crucial for gonadal development in both sexe*s,* were also bivalent at E10.5 (Fig. 3.7A and B). The observed bivalency is consistent with the similar levels of expression between XY and XX progenitor cells (Fig. 3.7C). The enrichment level of H3K27me3 and H3K4me3 modifications of bivalent genes very closely mirrors their gene expression levels. For example, *Foxl2* and *Dmrt1* are repressed at this stage with levels similar to the negative control *Oct4* (Fig. 3.7C), and have a higher H3K27me3/H3K4me3 ratio than other sex-determining genes (Fig. 3.6A and B). In contrast, genes such as *Sox9, Fgf9, Wnt4* and *Rspo1* have higher levels of expression (Fig. 3.7C) and, accordingly, a higher H3K4me3/H3K27me3 ratio (Fig. 3.6A and B). However, despite the differences in expression levels, sex determination genes have high enrichment levels of both H3K27me3 and H3K4me3, suggesting that there is transcriptional variability amongst bivalent genes as has been reported (Bernstein et al., 2006a).

With the exception of a few genes linked to the sex chromosomes (Jameson et al., 2012b, Munger et al., 2013), XY and XX progenitor cells are transcriptionally indistinguishable at E10.5-E11.0 (Nef et al., 2005). At this stage, genes that later play a crucial role during sex

determination are expressed at similar levels (Fig. 3.5C and Fig. 3.7C). Interestingly, in ventral foregut endoderm cells, a population of bipotential cells that give rise to either liver or pancreatic cell types, certain histone modifications can predetermine cell fate (Xu et al., 2011). We therefore asked whether gonadal progenitor cells are epigenetically predisposed to their male or female fate. However, no significant differences were observed between XY and XX progenitor cells for mean enrichment levels of H3K4me3 and H3K27me3 at sex-determining genes (Fig. 3.7D), suggesting that promoter deposition of H3K4me3 and H3K27me3 does not direct differentiation of XY and XX progenitor cells, but rather contributes to their bipotential state. Interestingly, levels of H3K27me3 are slightly higher in XY than in XX supporting cells, possibly due to the higher transcriptional levels of the H3K27me3-demethylse *Utx* in females, a gene which escapes X-inactivation (Greenfield et al., 1998, Jameson et al., 2012b).

3.2.3. Repressed key sex-determining genes remain poised after sex determination

Having established that key sex-determining genes are bivalent prior to sex determination, we next asked whether their promoters resolve into sex-specific patterns of H3K4me3 and H3K27me3 after sex determination, consistent with the upregulation of either the male or female pathway and repression of the alternate fate. To investigate this, we compared the genome-wide profile of H3K4me3 and H3K27me3 in purified Sertoli and granulosa cells from E13.5 XY and XX gonads. We show that upregulation of sex-determining genes is accompanied by a strong reduction of the repressive H3K27me3 mark at their promoter by E13.5 (Fig. 3.8A and B and Fig. 3.9). For example, the key XY-determining genes *Sox9, Fgf9* and *Dmrt1*, which become upregulated in E13.5 XY supporting cells following expression of *Sry* (Fig. 3.8C)*,* show a >2-fold reduction of H3K27me3 enrichment in Sertoli cells (Fig. 3.8A and Fig. 3.9A). In total, 260 of the 382 (68%) XY-determining genes that were bivalent at E10.5 lose
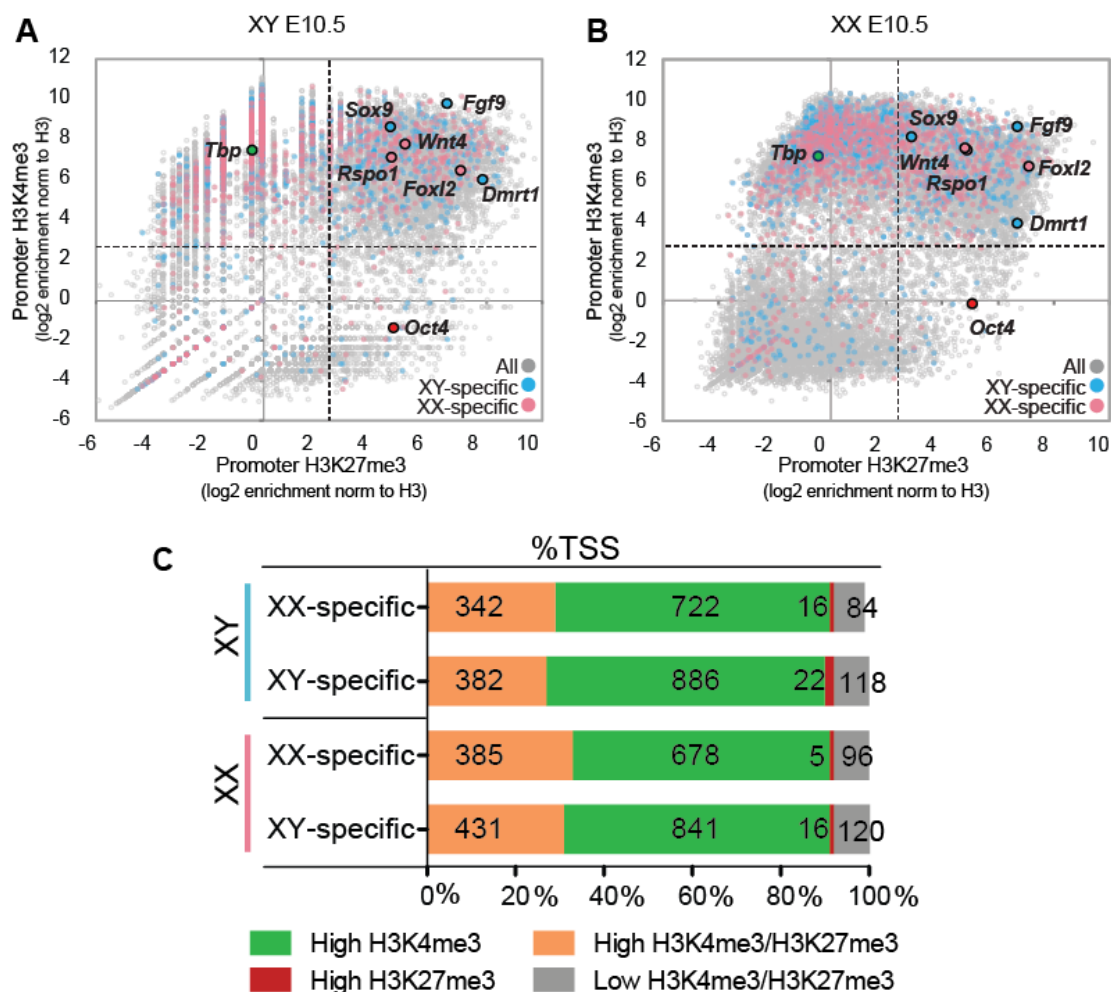
Figure 3.6. Promoters of many sex-determining genes are marked by high H3K27me3 and high H3K4me3 prior to sex determination. A&B) Scatterplots depicting enrichment of H3K27me3 (y axis) and H3K4me3 (x axis) normalized to total H3 within a 2kb promoter region around all TSS (mm9) (grey), in XY (left) and XX (right) progenitor cells. Genes that become either Sertoli- or granulosa-specific at E13.5 (Jameson et al., 2012b) are blue or pink respectively. *Oct4* (red) is a germ-cell specific gene that is not expressed in supporting cells; *Tbp* (green) is a constitutively active gene. Note that promoters of many sex determination genes (*Sox9, Fgf9, Dmrt1, Wnt4, Rspo1, Foxl2*) have high (log2>2.5) H3K27me3 and H3K4me3 enrichment C) Percentage of total promoters marked by high enrichment levels (>2.5 log2 enrichment normalized to H3) of H3K27me3 (red), H3K4me3 (green), both (orange), or neither (grey) throughout sex determination. Numbers of genes in each category are shown.
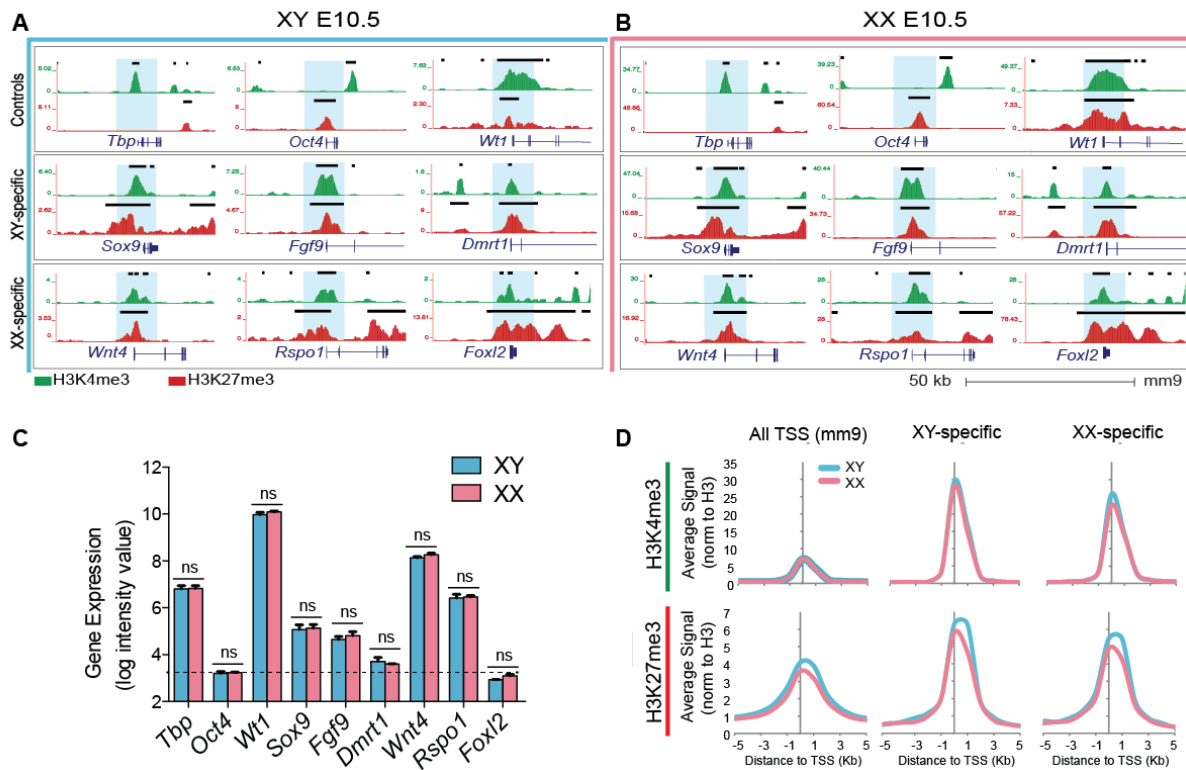
Figure 3.7. **Promoters of key sex-determining genes are bivalent prior to sex determination.** A&B) Genome browser tracks showing ChIP-seq profiles for H3K4me3 (green) and H3K27me3 (red) in XY (blue, left) and XX (pink, right) progenitor cells. Black boxes represent significant enrichment when compared to flanking regions as determined by HOMER. Note that SD gene promoters are marked by both histone modifications. C) Bar graphs denoting gene expression log intensity values from Nef et al, 2005, for select genes in XY (blue) and XX (pink) progenitor cells. Differences between XY and XX for each gene are not significant as determined by student's t test. D) Average H3K4me3 (top) and H3K27me3 (bottom) signal normalized to H3 around the TSS (5kb upstream to 5kb downstream) (x axis) at all TSS (mm9), XY-specific promoters, and XX-specific promoters in XY (blue) and XX (pink) progenitor cells.

H3K27me3 enrichment and shift towards an H3K4me3-only state in Sertoli cells, consistent with an upregulation of Sertoli-specific gene transcription (Fig. 3.8A and D). Conversely, the key female-determining genes *Wnt4, Rspo1* and *Foxl2,* which become upregulated in E13.5 XX supporting cells in the absence of *Sry* (Fig. 3.8C)*,* show a >2-fold reduction of H3K27me3 enrichment in granulosa cells (Fig. 3.8B and Fig. 3.9B). Accordingly, 173 of the 385 (45%) female-determining genes that were bivalent at E10.5 shift towards an H3K4me3-only state in granulosa cells by E13.5 (Fig. 3.8B and D). Promoter H3K4me3 enrichment at these genes does not significantly increase during sex determination, but rather is retained at similar levels (Fig. 3.9C). It is important to note that *Foxl2,* a small single-exon gene, is embedded within an H3K27me3-dense locus and appears to retain H3K27me3 in granulosa cells. However, a closer look at the TSS and gene body itself reveals that these regions do in fact lose H3K27me3 upon differentiation (Fig. 3.9B), although this is not evident in our scatterplot analyses that used a 2kb window around the TSS for calculating enrichment levels.

Remarkably, we found that sex-determining genes that promote the alternate cell fate and become transcriptionally repressed during sex determination did not lose their active H3K4me3 mark and resolve into H3K27me3-only promoters. Instead, repressed sex-determining genes remained bivalent even after progenitor cells had departed from their bipotential state and differentiated into either Sertoli or granulosa cells. For example, *Wnt4, Rspo1* and *Foxl2,* female-determining genes that are actively repressed in Sertoli cells (Fig. 3.8C), retain high enrichment of both H3K4me3 and H3K27me3 (Fig. 3.8A and Fig. 3.9A). Similarly, *Sox9, Fgf9,* and *Dmrt1,* male-determining genes that are actively repressed in granulosa cells (Fig. 3.8C), retain high enrichment of both H3K4me3 and H3K27me3 (Fig. 3.8B and Fig. 3.9B). To determine whether this bivalency at repressed genes is retained in adult stages, we performed ChIP-qPCR in purified adult Sertoli cells (> 2 months old males) and

found that the female-specific genes *Wnt4, Rspo1, Foxl2, Bmp2* and *Fst* still retained enrichment of both H3K27me3 and H3K4me3 (Fig. 3.10).

These results are not exclusive to these key sex-determining genes. In fact, 242 of the 342 (71%) bivalent female-determining genes at E10.5 remain bivalent at E13.5 in Sertoli cells (Fig. 3.6C and Fig. 3.8D), and 301 of the 431 (70%) of bivalent male-determining genes at E10.5 remain bivalent at E13.5 in granulosa cells (Fig. 3.6C and Fig. 3.8D). These results are reflected in the analyses of the mean enrichment level of H3K4me3 and H3K27me3 at the promoters of male- and female-specific genes in Sertoli and granulosa cells (Fig. 3.9D). While average H3K4me3 enrichment is higher at male-specific genes in Sertoli cells and at female-specific genes in granulosa cells, genes promoting the alternate cell fate have not completely lost H3K4me3, suggesting that even repressed genes retain this active mark. However, the average H3K27me3 enrichment level is higher at female-specific genes in Sertoli cells and at male-specific genes in granulosa cells, consistent with its role in repression (Fig. 3.9D). Interestingly, there is a significant difference between the mean H3K27me3 enrichment at female-determining genes in Sertoli cells and granulosa cells, whereas this difference is not as marked at male-determining genes. This may be a reflection of the total percentage of bivalent genes that have not resolved at E13.5 (Fig. 3.9D): while 260 of the initial 382 (68%) of XY-specific bivalent genes have resolved in E13.5 Sertoli cells, only 173 of the initial 385 (45%) of XX-specific bivalent genes have resolved in E13.5 granulosa cells, possibly because the female developmental pathway is not yet fully established at E13.5 (Jameson et al., 2012b).

Our results suggest that key sex-determining genes that promote the alternate cell fate remain in a poised state for activation even after sex determination has occurred, and in the case of males, even in adult Sertoli cells, possibly providing supporting cells with an epigenetic
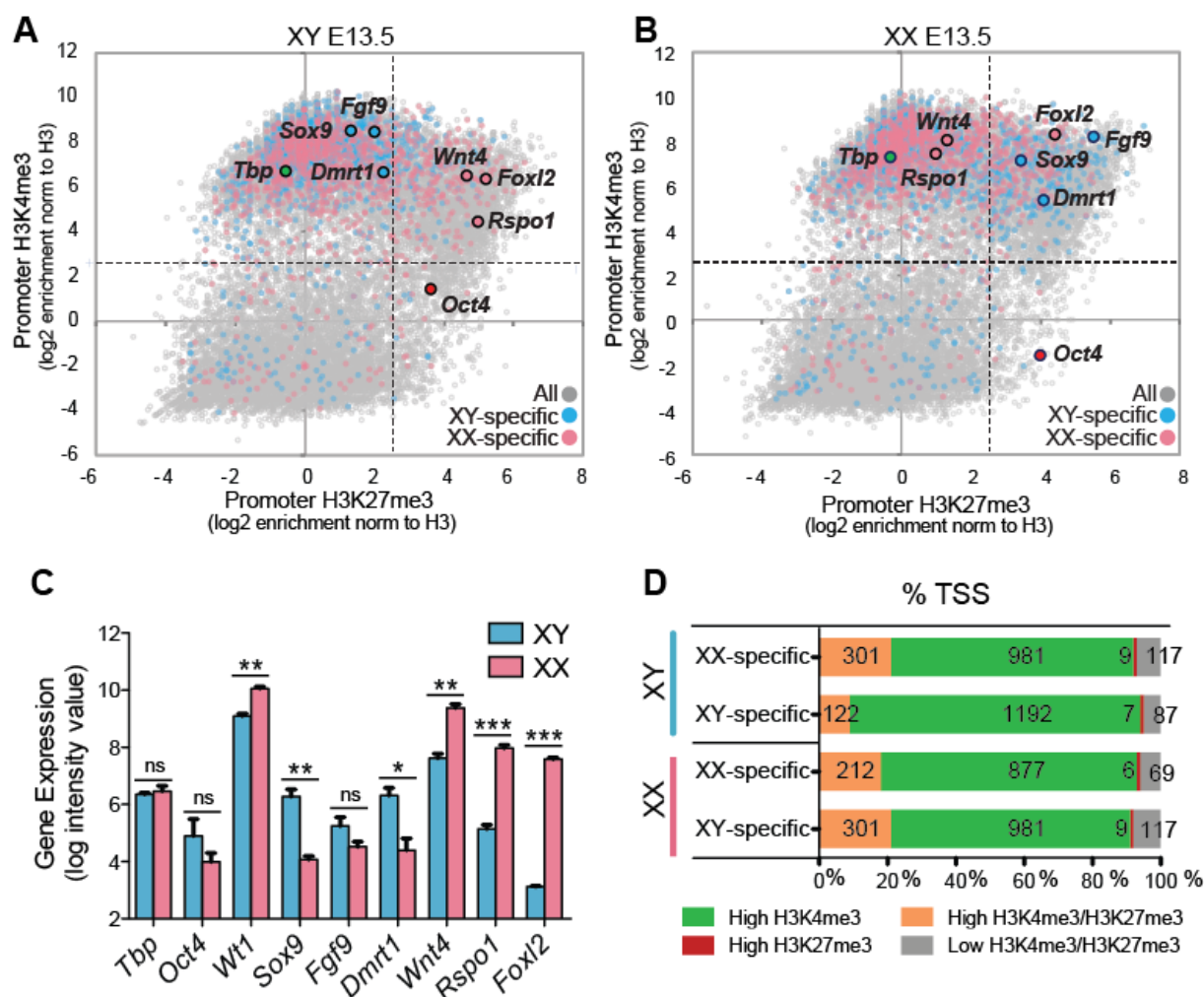
Figure 3.8. Upregulation of sex-determining genes is accompanied by loss of promoter H3K27me3. A&B) Scatterplots depicting enrichment of H3K27me3 (y axis) and H3K4me3 (x axis) normalized to total H3 at a 2kb promoter region around the TSS of all annotated mm9 genes (grey) in Sertoli cells (left) and granulosa cells (right) at E13.5. Genes that become either Sertoli- or granulosa-specific at E13.5 as determined by Jameson et al., 2012 are in blue and pink respectively. Note that promoters of SD genes that promote the alternate sex have high (log2>2.5) H3K27me3 and H3K4me3 enrichment even after sex determination. C) Bar graphs of gene expression log intensity values from Nef et al, 2005, for select genes in XY (blue) and XX (pink) progenitor cells. *** represents p<0.0001 as determined by student's t test compared to the negative control *Oct4*. D) Percentage of total promoters marked by high enrichment levels (>2.5 log2 enrichment normalized to H3) of H3K27me3 (red), H3K4me3 (green), both (orange), or neither (grey) throughout sex determination. Numbers of genes in each category are shown.
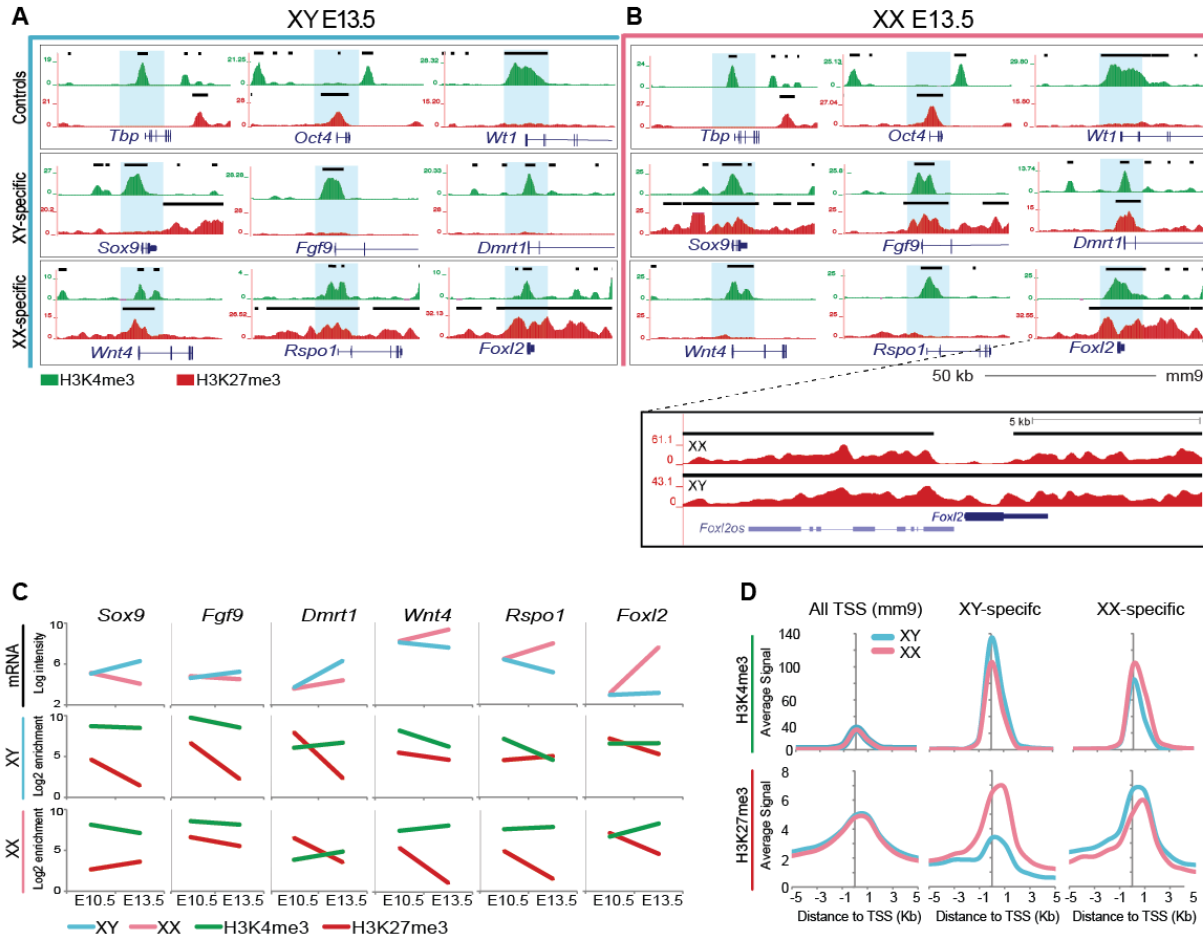
Figure 3.9. Repressed key sex-determining genes remain bivalent after sex determination. A&B) Genome browser tracks showing ChIP-seq profiles for H3K4me3 (green) and H3K27me3 (red) in Sertoli cells (blue, left) and granulosa cells (pink, right) at E13.5. Black boxes represent significant enrichment when compared to flanking regions as determined by HOMER. Note that repressed key SD genes are marked by both histone modifications. A closer look at H3K27me3 ChIP-seq tracks at *Foxl2* (boxed) in granulosa cells (top) and Sertoli cells (bottom) shows loss of H3K27me3 at the promoter of *Foxl2* XX but not XY cells. C) Time course of gene expression (log intensity values from Nef et al., 2005) from E10.5 to E13.5 in XY (blue) and XX (pink) supporting cells (top row), and of H3K4me3 (green) and H3K27me3 (red) promoter enrichment value normalized to H3 in XY (middle row) and XX (bottom row) supporting cells. D) Average H3K4me3 (top) and H3K27me3 (bottom) signal normalized to H3 around the TSS (5kb upstream to 5kb downstream) (x axis) at all TSS (mm9), XY-specific promoters, and XX-specific promoters in Sertoli (blue) and granulosa (pink) cells.

Adult Sertoli Cells
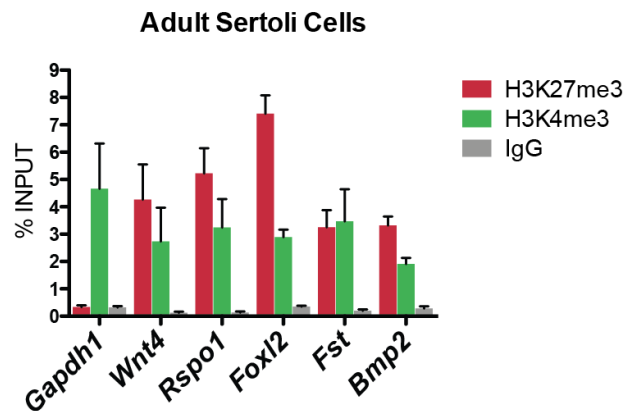
Figure 3.10. Female pathway genes remain bivalent in adult Sertoli cells. ChIP-qPCR for H3K27me3 (red), H3K4me3 (green) and IgG (grey) at the promoter of several female-specific genes in purified Sertoli cells from adult (>2m/o) males. Each qPCR was performed on 3 biological replicates, each replicate contained purified Sertoli cells from 1-2 adult males.

memory of their bipotential state and contributing to their ability to transdifferentiate long after sex determination.

3.2.4. Sex reversal is rescued in Cbx2;Wnt4 double knock-out mice

Based on previous findings, we became interested in understanding what mechanism could be protecting genes that drive the alternate pathway from loss of H3K27me3 and subsequent upregulation. Classically, the PcG complexes, Polycomb Repressive Complex 1 and 2 (PRC1 & PRC2) mediate repression of target genes. Canonically, PRC2 recruits PRC1 by catalyzing H3K27me3 at the promoter of its target genes. Binding of PRC1 to H3K27me3 through its CBX subunit maintains transcriptional repression through chromatin compaction (Lau et al., 2017). It was shown previously that loss of CBX2 leads to complete male-to-female sex reversal in mice and humans, which is believed to result from loss of *Sry* and *Sox9* expression (Katoh-Fukui et al., 1998, Biason-Lauber et al., 2009, Katoh-Fukui et al., 2012). In accord with previous results, we found that loss of *Cbx2* leads to loss of *Sry* and *Sox9* expression in E13.5 XY gonads, inducing male-to-female sex reversal in C57BL/6J mice (Fig. 3.11A-C"). Testis
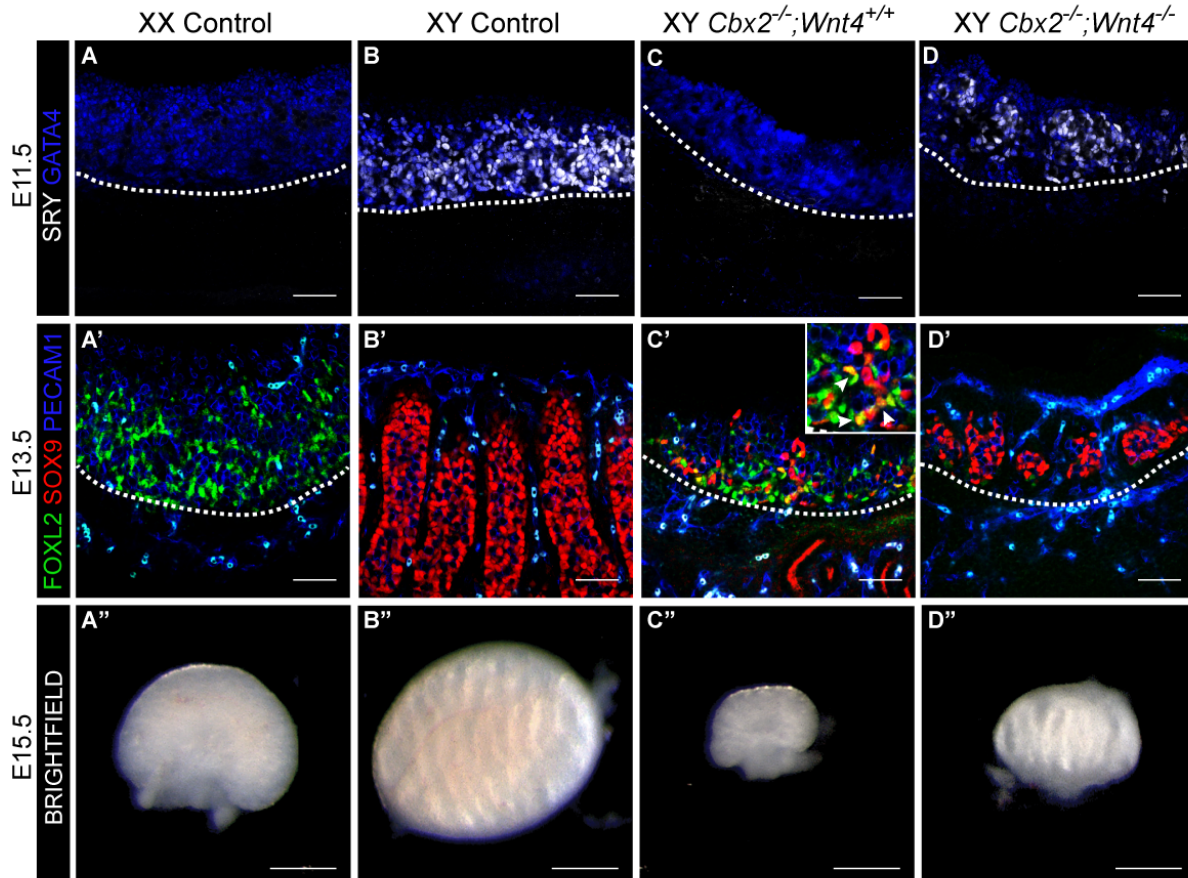
Figure 3.11. ***Sry* and *Sox9* expression is rescued in *Cbx2;Wnt4* double knockout XY gonads.** *Sry* expression and testis development are rescued in XY gonads of *Cbx2;Wnt4* double knockout (dKO) mice. Immunofluorescent analysis of gonads from E11.5 (A-D) and E13.5 (A'-D') gonads, and brightfield images of E15.5 (A"-D") gonads. E11.5 gonads (A-D) are stained with somatic cell marker GATA4 (blue) and SRY (white). E13.5 gonads (A'-D') are stained with the granulosa cell marker FOXL2 (green) and the Sertoli cell marker SOX9 (red). WT XX gonads have FOXL2-expressing granulosa cells (A'), whereas WT XY gonads have SOX9-expressing Sertoli cells (red), which are organized around germ cells forming testis cords (B'). Loss of *Cbx2* in E13.5 XY gonads leads to reduction of SOX9+ Sertoli cells (red) and gain of FOXL2+ granulosa cells (green). Some individual cells express both markers (yellow, arrowheads in inset). Testis cords are lost and the morphology resembles WT XX (A'). SRY, SOX9, and testis cords are rescued in *Cbx2;Wnt4* dKO gonads (D'). Bright field images of E15.5 gonads show the morphology of WT ovaries (A") and WT testis (B"). Loss of *Cbx2* causes male-to-female sex reversal leading to formation of functional but hypoplastic ovaries (C"). Loss of *Wnt4* on a *Cbx2* KO background recues testis morphology, but not testis size. Images correspond to one of n=3. Scale bars represent 50um (A-D') and 500um(A"-D").

cords that typically characterize this stage of male development (Fig. 3.11B'&B"), were completely absent (Fig. 3.11C'&C"), and *Cbx2*[-/-] gonads develop as functional, albeit hypoplastic ovaries (Fig. 3.11C-C"). However, at E13.5, XY *Cbx2*[-/-] mutants have both FOXL2- and SOX9-expressing cells (Fig. 3.11C'), including some individual cells that express both markers simultaneously (inset in Fig. 3.11C'). This phenotype suggested a failure to repress the female pathway in XY cells.

To investigate a repressive role for PcG during testis development, we performed functional annotation of all granulosa-specific promoters and surrounding loci marked by H3K27me3 in XY cells. The highest represented developmental process is the female reproductive system, confirming the role the PcG complex plays in repressing the female pathway (Fig. 3.12B). Specifically, the Wnt signaling pathway is highly represented (Fig. 3.12A). A closer look revealed that in addition to *Wnt4* and *Rspo1* (Fig. 3.9), several other members of the Wnt signaling pathway are marked by H3K27me3, such as *Wnt2b, Wnt5a, Axin2, Fzd1, Lef1,* and *Lgr5* (Fig. 3.12C), all of which become female-specific upon sex-determination (Jameson et al., 2012b). Our results suggest that the PcG proteins play a widespread role in repressing the female pathway by targeting the Wnt signaling pathway for H3K27me3-mediated repression in Sertoli cells.

Several previous lines of work showed that activation of Wnt signaling in the XY gonad could suppress male development (Maatouk et al., 2008), and that the predominant role of *Fgf9* in the male pathway is to repress *Wnt4* (Kim et al., 2006b, Jameson et al., 2012a). Based on the high representation of Wnt pathway genes bearing H3K27me3 marks in XY cells, we predicted that CBX2 normally acts as a repressor of the Wnt pathway during male development. We therefore hypothesized that loss of CBX2 results in an inability to repress Wnt signaling in the XY gonad, which has the indirect consequence of silencing *Sry*, *Sox9*, and the male pathway.

To test this prediction, we created a *Cbx2^-/-;Wnt4^-/-* double knockout mouse. In accord with

our hypothesis, immunofluorescent analysis of E11.5 XY *Cbx2^-/-;Wnt4^-/-* gonads showed that

SRY expression was rescued (Fig. 3.11D), similar to wild type XY mice (Fig. 3.11B).

Furthermore, SOX9 expression and testis cord formation were also rescued in E13.5 *Cbx2^-/-*

;Wnt4^-/- gonads (Fig. 3.11D'&D"), which develop as testes (Fig. 3.11D"). These results clearly

indicate that *Cbx2* is not required for SRY or SOX9 expression.
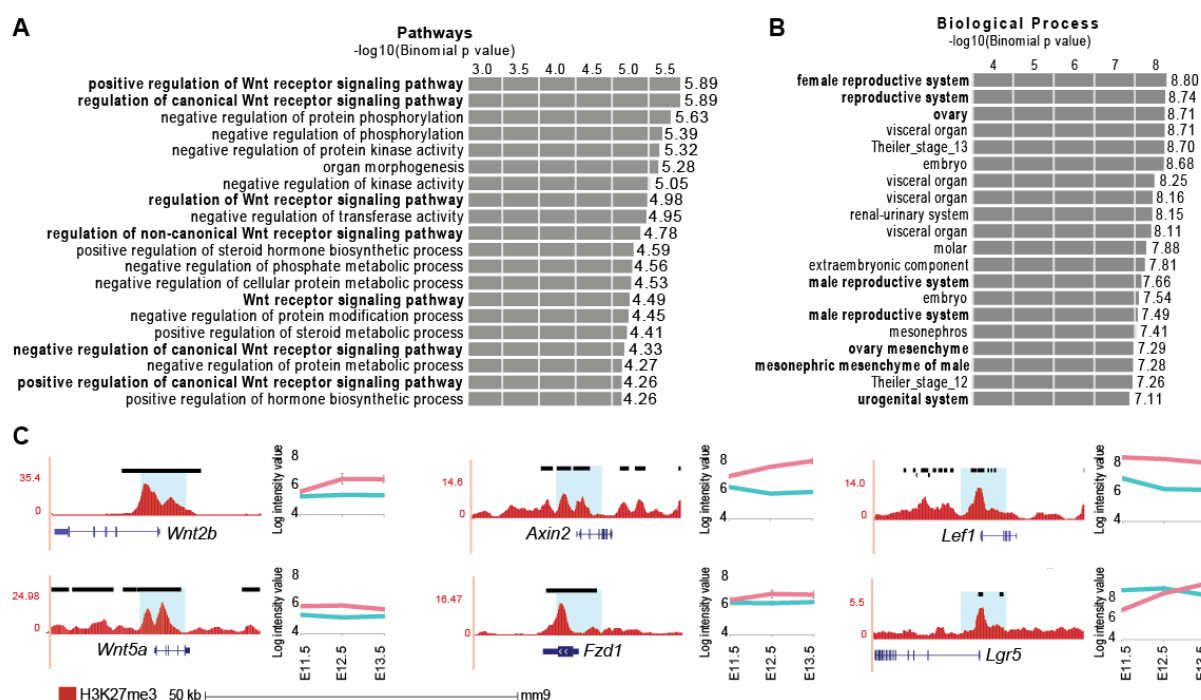
Figure 3.12. The Wnt pathway is targeted for H3K27me3-mediated repression in Sertoli cells. A) Gene Ontology functional analysis of granulosa-specific promoters and flanking regions marked by H3K27me3 shows that the Wnt signaling pathway is significantly targeted for repression in Sertoli cells. B) Gene Ontology functional analysis of granulosa-specific promoters and flanking regions marked by H3K27me3 shows that the developmental process most highly represented are those associated to the formation of the reproductive system, in particular the female reproductive and urogenital system. C) Genome browser tracks showing ChIP-seq profiles for H3K27me3 (red) at the promoter of several members of the Wnt signaling pathway. Black boxes represent regions significantly enriched for H3K27me3 compared to flanking regions as determined by HOMER. Gene expression profiles for each of these during sex determination (E11.5-E13.5) genes are to the right, in XY supporting cells (blue) and XX supporting cells (pink) (obtained from Jameson et al, 2012b).

It was originally reported that $Cbx2^{-/-}$ gonads were hypoplastic due to a somatic cell proliferation defect (Katoh-Fukui et al., 2012). Despite the fact that XY $Cbx2^{-/-};Wnt4^{-/-}$ gonads express testis markers and are morphologically similar to XY wild type gonads (Fig.3.11D-D"), gonad size is not rescued, consistent with the idea that *Cbx2* regulates sex determination and gonad size through independent pathways (Katoh-Fukui et al., 2012). Our results suggest that CBX2 positively regulates *Sry* and *Sox9* expression by directly repressing the granulosa-promoting Wnt signaling pathway.

Fgf9 promotes male development by repressing *Wnt4* during testis development (Jameson et al., 2012a). To investigate whether *Fgf9* synergizes with *Cbx2* to repress the female pathway, we crossed *Fgf9* and *Cbx2* mutants. Although homozygous loss of either *Cbx2* or *Fgf9* leads to complete male to female sex reversal (Katoh-Fukui et al., 1998, Colvin et al., 2001), gonads with heterozygous loss of either gene develop as wild type gonads (Fig. 3.13A-C). However, gonads of $Cbx2^{+/-};Fgf9^{+/-}$ double heterozygotes show partial male-to-female sex reversal at the gonadal poles (Fig. 3.13D), similar to other models of partial sex reversal (Jameson et al., 2012a, Bagheri-Fam et al., 2008). These experiments are consistent with the idea that *Cbx2* and *Fgf9* act synergistically to repress the female fate.
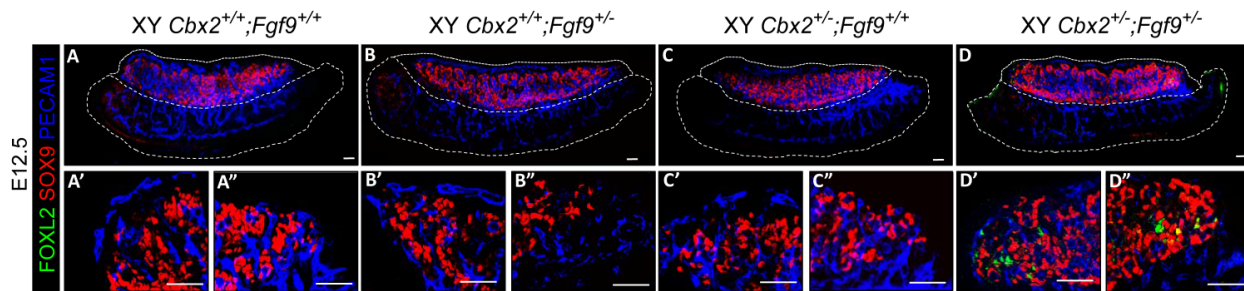
Figure 3.13. *Fgf9* and *Cbx2* act synergistically to repress the female pathway. (A-D") Immunofluorescent analysis of XY E12.5 gonads stained with granulosa marker FOXL2 (green) and Sertoli marker SOX9 (red). XY gonads of *Cbx2;Fgf9* double heterozygous mice have FOXL2+ granulosa cells at the gonadal poles. The anterior (left, eg. A') and posterior (right, eg. A") poles of each gonad (A-D) are enlarged in the bottom row. Images are representative of n>3. Scale bars represent 50um.
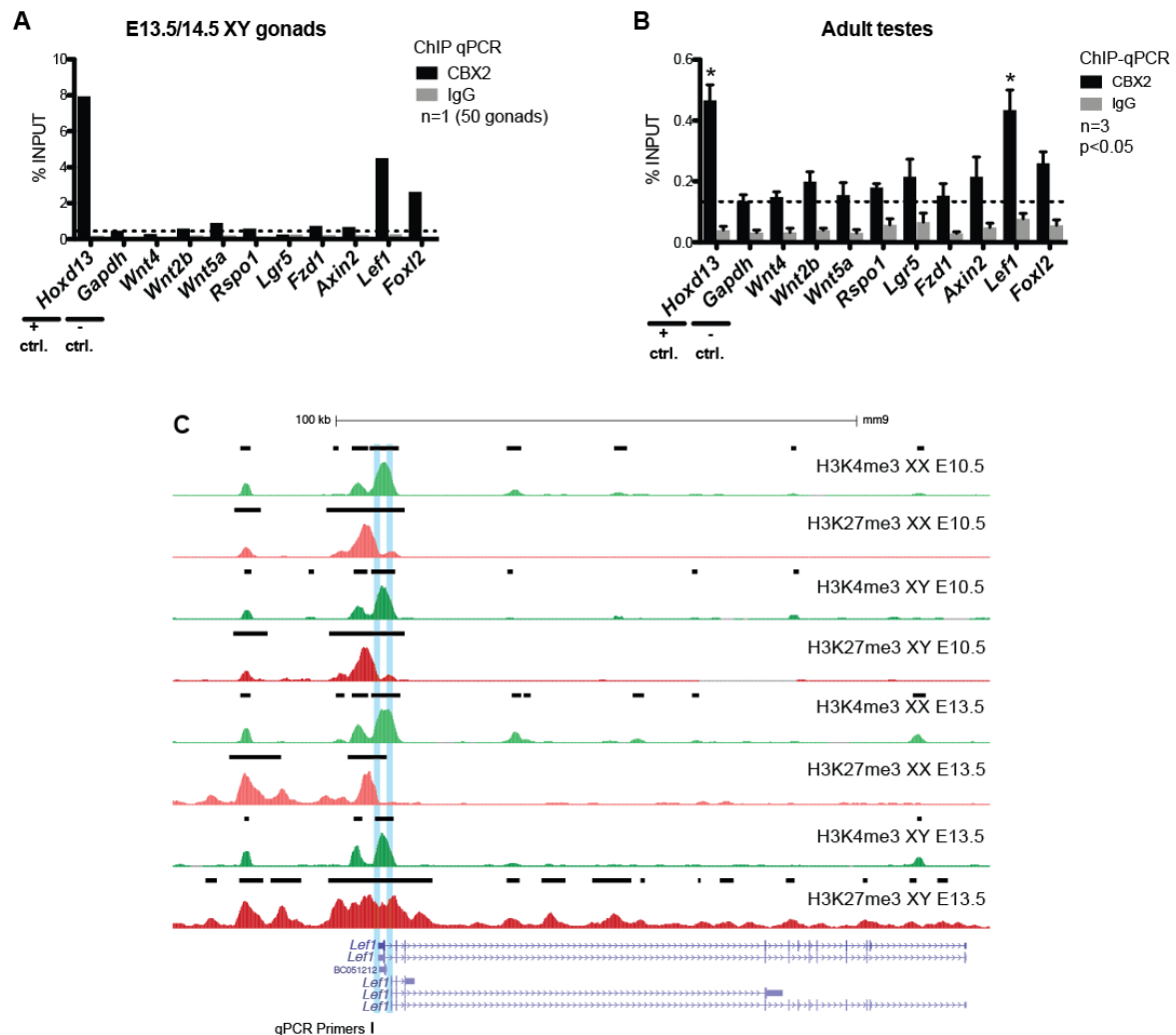
Figure 3.14. CBX2 targets *Lef1* in testes. ChIP-qPCR for CBX2 on embryonic gonads (50 pooled E13.5 and E14.5 XY gonads) (A) and on testes from 3 adult males (B). * represents p<0.01 as determined by student's t test when compared to the negative control *Gapdh.* C) UCSC genome browser tracks showing ChIP-seq profiles for H3K4me3 (green) and H3K27me3 (red) at the promoter of *Lef1*. Black boxes above tracks represent regions significantly enriched for H3K4me3 or H3K27me3 compared to H3 as determined by HOMER. Two possible TSS for *Lef1* annotated in the mm9 genome are highlighted in blue. The black box at the bottom represents the region targeted for CBX2 ChIP-qPCR above.

3.2.5. The PcG subunit CBX2 targets the downstream Wnt signaler Lef1 during testis development

To determine whether CBX2 directly targets *Wnt4* and/or other female-specific Wnt signalers, we performed ChIP-qPCR in pooled E13.5 and E14.5 XY gonads, and in adult (>2m/o) testes (Fig. 3.14A and B). *Hoxd13*, a known CBX2 target gene (Lau et al., 2017), was used as a positive control. In contrast, the promoter region of the constitutively active gene *Gapdh* was used as a negative control. Surprisingly, CBX2 did not bind *Wnt4,* or other Wnt signalers such as *Wnt2b, Wnt5a, Rspo1, Axin2, Fzd1,* and *Lgr5.* However, *Lef1* was significantly targeted by CBX2 in both embryonic and adult testes, with levels similar to the positive control *Hoxd13. Lef1* is a female-specific gene (Fig 3.12C and Fig. 3.15) (Nef et al., 2005, Jameson et al., 2012b) known to interact with β-catenin in the nucleus to drive expression of target genes (Behrens et al., 1996). Furthermore, *Lef1* is bivalent at the bipotential stage (Fig. 3.14C). While the region marked by H3K27me3 decreases in XX cells at E13.5, it increases to engulf the promoter region, a region downstream of the TSS, and most of the gene body in XY E13.5 cells (Fig. 3.14C), in accordance with its repression. Therefore, our results suggest that CBX2 inhibits upregulation of the female pathway during testis development by directly repressing *Lef1.*

3.3. Discussion

The early progenitors in the bipotential mouse gonad are a valuable model for the investigation of how cells resolve their fate and commit to one of two differentiation pathways. To provide insight into the epigenetic mechanisms that regulate this cell-fate decision, we developed an *in vivo* quantitative profile of H3K4me3 and H3K27me3 enrichment in XY and XX gonadal supporting cells at time points before (E10.5) and after (E13.5) sex determination. We

show that genes essential to establish both male and female fate are initially poised in both the XX and XY gonad, suggesting that bivalency defines the bipotential state. Once male or female sex determination occurs, the male- or female-expressed bivalent genes resolve into H3K4me3-only promoters. However, the genes associated with the alternate sexual fate retain their bivalent marks, suggesting that they remain poised at E13.5 and, at least in males, in adult Sertoli cells. We show that male to female sex-reversal caused by loss of CBX2, the PRC1 protein that recognizes the H3K27me3 mark, can be rescued by simultaneous loss of *Wnt4*. Furthermore, we show that CBX2 directly binds *Wnt4's* downstream target *Lef1*, a female-specific co-factor of -catenin. These findings suggest that loss of *Sry* expression and the male pathway in *Cbx2* mutants may be the indirect consequence of failure to repress the female pathway.

The co-occurrence of the active H3K4me3 and the repressive H3K27me3 marks was first reported in ESCs, in which a cohort of developmental genes is "bivalent", poised for future repression or activation (Bernstein et al., 2006a, Azuara et al., 2006, Pan et al., 2007). Our findings are consistent with the idea that epigenetic regulation plays a critical role in maintaining pluripotency and facilitating divergence into multiple differentiated states. We show that most genes that become either male- or female-specific after sex determination initially harbor both the H3K4me3 and H3K27me3 modifications at their promoters. More specifically, key sex-determining genes (genes known to cause sex disorders when disrupted) such as the male-determining genes *Sox9, Fgf9* and *Dmrt1*, and the female-determining genes *Wnt4, Rspo1* and *Foxl2*, have high enrichment levels of both chromatin marks at the bipotential stage (>log 2.5 enrichment normalized to total H3). To correlate chromatin states and transcriptional regulation, we made use of the microarray gene expression data collected by Nef et al. (Nef et al., 2005) in XY and XX gonadal somatic cells. Interestingly, we found transcriptional variability even

amongst genes with high enrichment levels of H3K27me3 and H3K4me3 marks. This has been previously reported (Bernstein et al., 2006a), and could correlate with the presence of additional histone variants such as H2A.X, with the acquisition of transcription factors that act as repressors or activators, or with recruitment of PRC1.

Typically, upregulation of bivalent genes is accompanied by loss of H3K27me3 at the promoter, whereas repression is accompanied by a loss of H3K4me3. This binary resolution maintains the transcriptional state throughout cell divisions even in absence of the initial signal. We therefore predicted that sex-determining genes would resolve into sex-specific patterns of histone modifications, stably transmitting the cell fate decision throughout adulthood. Instead, we show that although sex-determining genes that become upregulated do lose their repressive mark, sex-determining genes that promote the alternate pathway and become transcriptionally repressed persist in a poised state even after differentiation into either Sertoli or granulosa cells.

Based on our observations, we propose a model in which sex-determining genes that are initially expressed at similar levels between males and females are bivalent in progenitor cells, and are held in a chromatin-accessible conformation. This state maintains the balance between the male and female fate yet renders genes responsive to the sex-determining signal triggered by *Sry* in males or presence of *Wnt4/Rspo1* in females. H3K4me3 and H3K27me3 then serve a dual role during sex determination: on the one hand, sex-determining genes that become upregulated upon differentiation resolve into H3K4me3-only promoters, stably transmitting that commitment throughout cell division. On the other hand, sex-determining genes of the alternate pathway remain poised, competent for activation when they receive an appropriate signal. The finding that female-determining genes retain enrichment of both H3K24me3 and H3K27me3 in Sertoli cells purified from adult testes led us to speculate that transmission of poised sex-determining genes from the bipotential to the differentiated state

could confer the remarkable plasticity that enables *in vivo* transdifferentiation of the supporting cell lineage to the alternate cell fate, even long after the fate-decision has been made. Our findings may also explain why repression of the alternate cell fate is an active process throughout adulthood. In the absence of active repressors, poised promoters may be vulnerable to differentiation toward the alternate cell fate. Our data suggest that maintenance of a bivalent but silent state requires the activity of PRC1.

Loss of the PRC1 subunit CBX2 leads to upregulation of the female pathway and complete male-to-female sex reversal. This was originally thought to be due to a failure to express *Sry*, implying an activating role for PRC1 (Katoh-Fukui et al, 1998; Katoh-Fukui et al, 2012). Although some PcG proteins have been reported to promote activation of certain targets (Jacob et al., 2011, Herz et al., 2012, Mousavi et al., 2012), we hypothesized that downregulation of the male pathway is the indirect consequence of failure to repress the female program in XY *Cbx2* mutants. Our present results are consistent with this hypothesis and clearly show that 1) *Cbx2* is not required for *Sry* or *Sox9* expression, as *Cbx2;Wnt4* double mutant mice express both of these genes, and 2) CBX2 regulation of *Sry* is indirect through repression
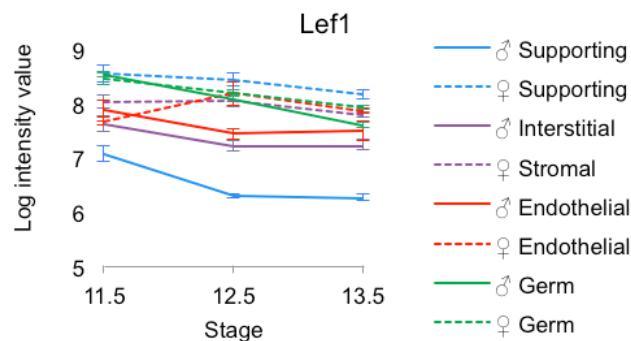
Figure 3.15. *Lef1* gene expression profile in XX and XY gonads during sex determination. Gene expression profile from Jameson et al., 2012 of *Lef1* during sex determination in purified germ and somatic cells. Note that *Lef1* is upregulated in female supporting cells and repressed in male supporting cells.

of *Wnt4's* downstream target *Lef1,* which inhibits transcription of downstream granulosa-determining genes with anti-*Sry* activity*.*

Interestingly, we found that CBX2 does not target *Wnt4* itself, but rather it's downstream target *Lef1.* Accordingly, *Lef1* expression is high in granulosa cells and repressed in Sertoli cells (Fig. 3.15) (Jameson et al., 2012b). LEF1 (also known as TCF1-α) interacts with β-catenin in the nucleus to drive expression of Wnt target genes (Behrens et al., 1996). We therefore propose a model in which presence of WNT4 in early XY fetal gonads (E10.5-E11.5) (Jameson et al., 2012b, Nef et al., 2005), is unable to drive expression of female-promoting target genes despite nuclear internalization of β-catenin, due to CBX2-mediated repression of *Lef1..* Upon loss of *Cbx2, Lef1* is expressed and LEF1 interacts with nuclear β-catenin, driving expression of the female pathway and leading to ovary development in XY individuals. However, in *Cbx2;Wnt4* double KO XY gonads, the female pathway cannot be upregulated despite presence of LEF1, as β-catenin is degraded in the absence of WNT4*,* resulting in testis development.

Multiple genes associated with male development are repressed in XX cells, thus it is still unclear why loss of *Cbx2* does not disrupt ovary development, despite its similar expression levels in XY and XX supporting cells. It is possible that maintenance of H3K27me3 at Sertoli-specific genes in XX cells is mediated by CBX subunits other than CBX2. Alternatively, PRC1 may achieve specificity by interacting with sex-specific transcription factors or non-coding RNAs during sex determination. Further insight into the mechanistic functions of CBX2 will benefit from more advanced techniques that enable ChIP for non-histone proteins using very small amounts of starting material. Our results highlight the importance of epigenetic mechanisms in the establishment and resolution of the bipotential state of gonadal precursors. We are the first to develop an *in vivo* genome-wide epigenetic profile of the supporting cell lineage throughout sex determination. We provide insight into how the bipotential state is established, and an

explanation for how differentiated supporting cells retain an epigenetic memory of their

bipotential state (Fig. 3.16). Furthermore, we describe a widespread role for the PcG proteins in

repressing the female pathway during testis development and show that its subunit CBX2 is

required to directly repress *Wnt4's* downstream target *Lef1* in Sertoli cells (Fig. 31.6). Expanding

our knowledge of the regulatory mechanisms underlying sex determination may increase our

ability to diagnose disorders of sex development with unknown etiologies, over 50% of which

remain undiagnosed.
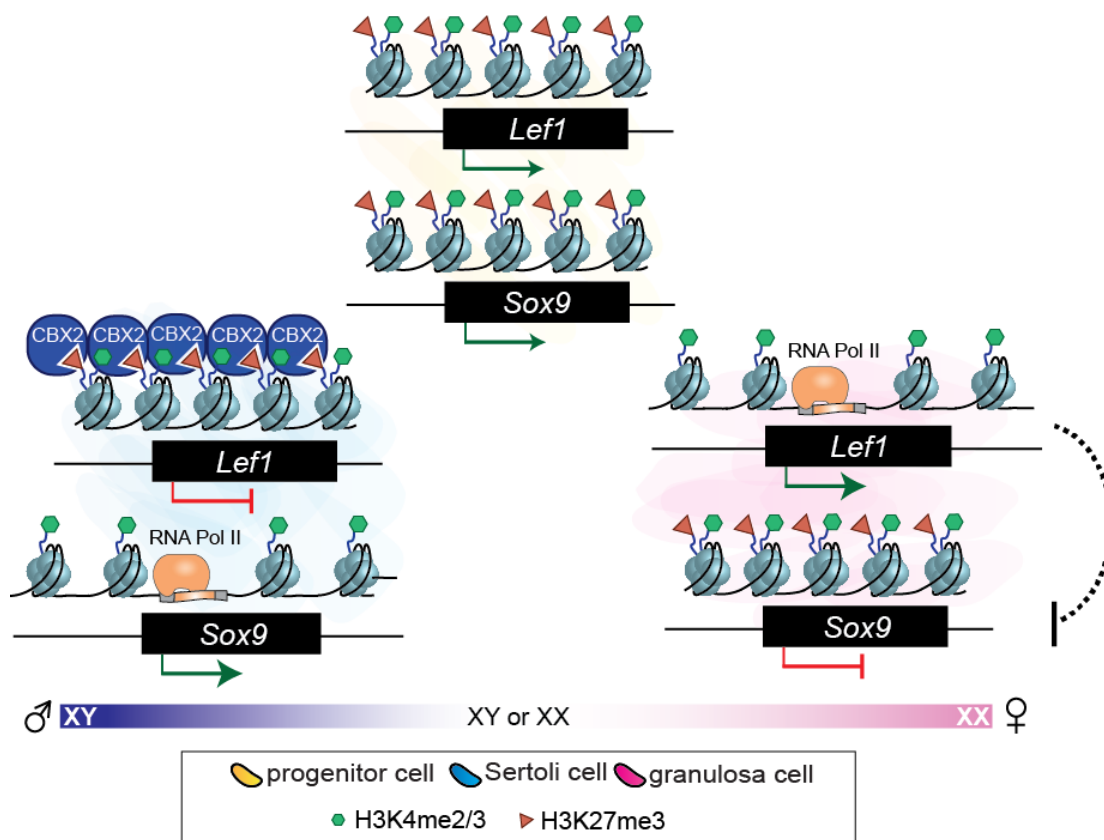
Figure 3.16. Model for CBX2 repression of the female pathway during male fate commitment. Initially, male (i. e. *Sox9*) and female (i. e. *Lef1*) pathway genes are bivalent at the bipotential stage. Upregulation of the male pathway in Sertoli cells is accompanied by loss of H3K27me3 at the promoters of male pathway genes. Conversely, upregulation of female pathway genes in granulosa cells is accompanied by loss of H3K27me3 at promoters of female pathway genes. CBX2 binds *Lef1* in males and maintains silencing of the female pathway. In females, expression of *Lef1* leads to upregulation of Wnt pathway genes that suppress the male fate.

CHAPTER FOUR

GONADAL SUPPORTING CELLS ACQUIRE SEX-SPECIFIC CHROMATIN LANDSCAPES DURING MAMMALIAN SEX DETERMINATION

4.1. Background and Significance

Sex determination is a powerful system to study cell differentiation. Although the chromosomal sex of a mammal is established at fertilization by inheritance of an X or Y chromosome from the father, XX and XY embryos are initially indistinguishable. Even the early gonad has no obvious sex-specific features, and initiates development as a bipotential organ with the ability to differentiate into either a testis (typically in XY individuals) or an ovary (typically in XX individuals) (Fig. 4.1A) (Burgoyne et al., 1995, Albrecht and Eicher, 2001). A subset of the somatic cells that compose the bipotential gonad, known as the supporting cell lineage, initiate gonadal sex determination by engaging either the male or the female pathway and promoting divergence from the progenitor state. Therefore, gonadal sex determination hinges on the ability of a population of bipotential cells to commit to one of two cell fates at a precise moment in development (Albrecht and Eicher, 2001).

In eutharian and metatherian mammals, gonadal sex determination is triggered by expression of the Y-encoded gene *Sry* around mid-gestation (Gubbay et al., 1990, Sinclair et al., 1990, Koopman et al., 1991). *Sry* upregulates its downstream target *Sox9,* a transcription factor (TF) which then directs differentiation of Sertoli cells (Hacker et al., 1995, Bullejos and Koopman, 2001, Sekido et al., 2004). In XX gonads that lack *Sry*, the *Wnt4/Rspo1/Ctnnb1* pathway is involved in directing the supporting progenitor cells to differentiate as granulosa cells (Fig. 4.1A) (Vainio et al., 1999, Parma et al., 2006, Maatouk et al., 2008). Although it may appear that gonadal sex determination is simply defined by the presence or absence of *Sry,* a complex network of male- or female-promoting signaling pathways coexist at the bipotential stage that require tight regulation (Jameson et al., 2012b, Munger et al., 2013). Evidence that gene dosage must be tightly regulated comes from studies of humans with Disorders of Sex Developments (DSDs) that have duplications or deletions in the region upstream of the *Sox9*

locus, a region devoid of coding genes but enriched for regulatory elements. While duplications in XX individuals lead to increased levels of *Sox9* causing female-to-male sex reversal, deletions in XY individuals lead to decreased *Sox9* levels causing male-female sex reversal (Wagner et al., 1994, Benko et al., 2011, Lybaek et al., 2014, Kim et al., 2015). This highlights how a slight disruption to this network can be enough to send the system towards the opposite pathway. However, our inability to pinpoint the location of cis-regulatory elements limits our capacity to study the mechanisms that regulate the precise spatiotemporal expression of sex-determining genes. Additionally, only ~20% of individuals with DSDs will receive a genetic diagnosis (Lux et al., 2009), likely due to mutations residing in non-coding regions that cannot be identified by standard diagnostic techniques such as karyotyping, sequencing of individual genes or even whole-exome sequencing.

To identify genomic elements that regulate sex determination, we developed a map of the chromatin accessibility landscape of the supporting cell lineage before and after commitment to the male or female fate. We purified XX and XY supporting cells before (E10.5) and after (E13.5) sex determination in mice, and performed Assay for Transposase-Accessible Chromatin (ATAC-seq) and Chromatin Immunoprecipitation followed by sequencing (ChIP-seq) for H3K27ac, a histone modification indicative of active enhancers (Creyghton et al., 2010). We show that XX and XY progenitor cells from E10.5 bipotential gonads have similar chromatin accessibility landscapes, and that these resolve into sex-specific patterns after differentiation into either granulosa (XX) or Sertoli (XY) cells. H3K27ac+ gonad-specific nucleosome-depleted regions (NDRs) are enriched around granulosa-promoting genes in granulosa cells, and Sertoli-promoting genes in Sertoli cells. Furthermore, these NDRs are enriched for binding motifs of transcription factors (TFs) linked to supporting cell differentiation, suggesting that gonad-specific NDRs are cis-regulatory elements that establish and/or maintain sex-specific transcriptional

programs throughout sex determination. Finally, we demonstrate the power of our dataset to identify novel enhancers by validating the *in vivo* activity of an enhancer downstream of *Bmp2,* a female-specific gene (Nef et al., 2005, Kashimada et al., 2011). This work will increase our understanding of the complex regulatory network that drives mammalian sex determination and is an invaluable resource for identifying functional elements that could harbor non-coding mutations that cause DSDs and escape routine diagnostic techniques.

4.2. Results

4.2.1. Gonadal supporting cells acquire sex-specific chromatin landscapes during sex determination

Regulatory elements lie within regions of open chromatin that are typically devoid of nucleosomes, known as nucleosome-depleted regions (NDRs) (Lee et al., 2004). The open chromatin conformation allows DNA-binding motifs to be accessed by transcription factors, transcriptional repressors, DNA-binding proteins with insulator activity, or in the case of promoters, by RNA Polymerase II (Song et al., 2011). To investigate the chromatin accessibility landscape of gonadal supporting cells at the time of gonad formation and after their differentiation into either Sertoli cells or granulosa cells, we performed ATAC-seq, a highly sensitive protocol capable of identifying NDRs in as few as 50K cells (Buenrostro et al., 2013). As NDRs correlate with a number of different genomic elements, such as enhancers, silencers, insulators and promoters, we also performed low-cell ChIP-seq for H3K27ac, a histone modification indicative of active enhancers (Creyghton et al., 2010, Rada-Iglesias et al., 2011, van Galen et al., 2016). ATAC-seq and ChIP-seq were performed in FACS-purified XX and XY bipotential progenitor cells from E10.5 gonads of *Sf1-GFP* (Beverdam and Koopman, 2006), differentiated Sertoli cells from E13.5 gonads of *TESCO*-CFP (Sekido and Lovell-Badge, 2008),
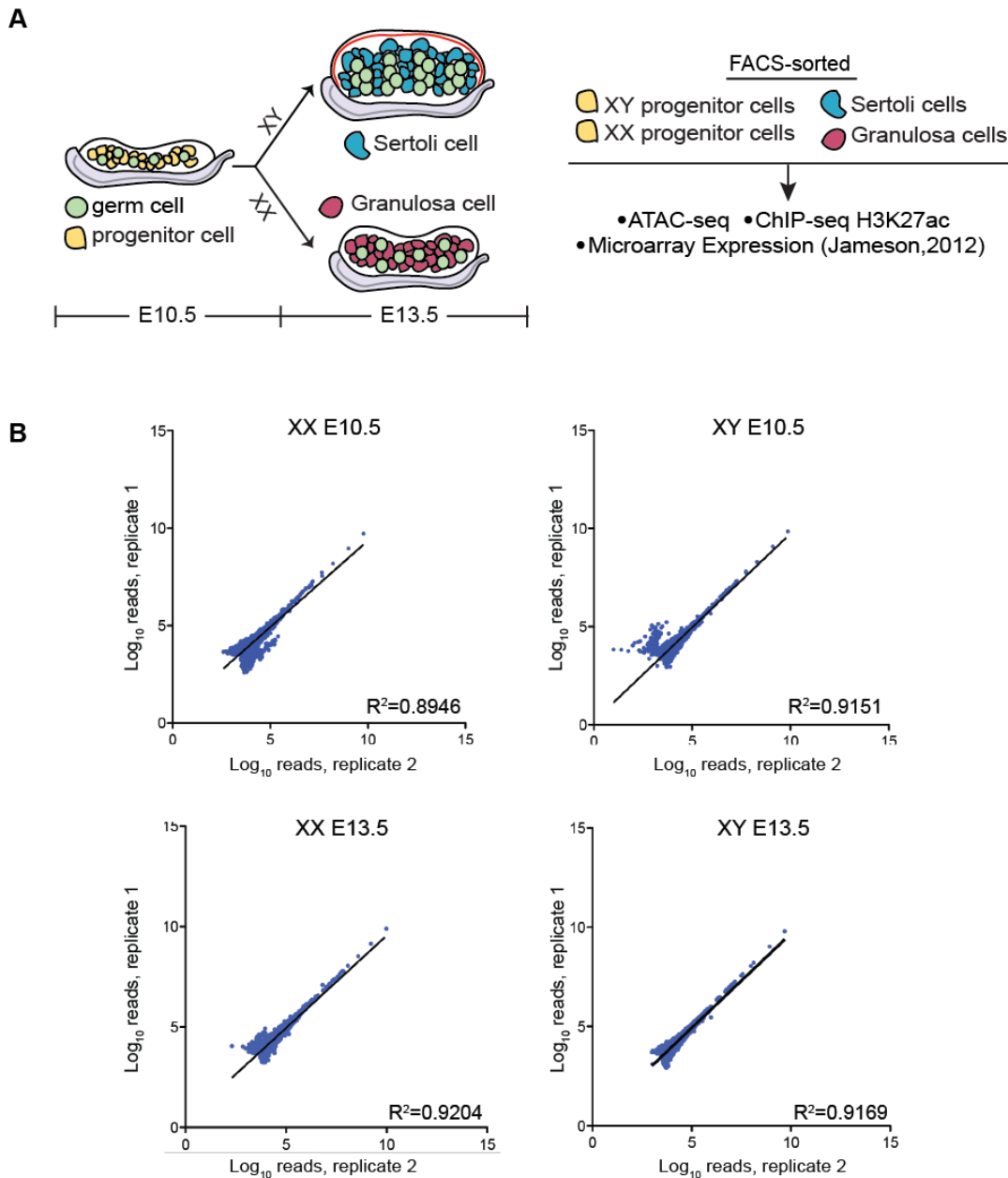
Figure 4.1. Overview of workflow and replicates. A) Briefly, supporting progenitor cells (yellow) are bipotential and indistinguishable between XX and XY gonads at E10.5. Expression of the Y-linked *Sry* gene (at E11.5) directs Sertoli cell differentiation in the testis (XY, blue). Absence of *Sry* directs differentiation of granulosa cells in the ovary (XX, pink). XX and XY progenitor cells (E10.5), Sertoli cells (E13.5), and granulosa cells (E13.5) were FACS-purified and used for ATAC-seq and ChIP-seq for H3K27ac. Further analysis made use of microarray expression data from purified supporting cells (Jameson et al., 2012b). B) The Log_{10} number of ATAC-seq reads for each biological replicate are plotted with replicate 1 on the Y axis and replicate 2 on the X axis. The plots show high correlation amongst the replicates (R squared value shown on bottom right).
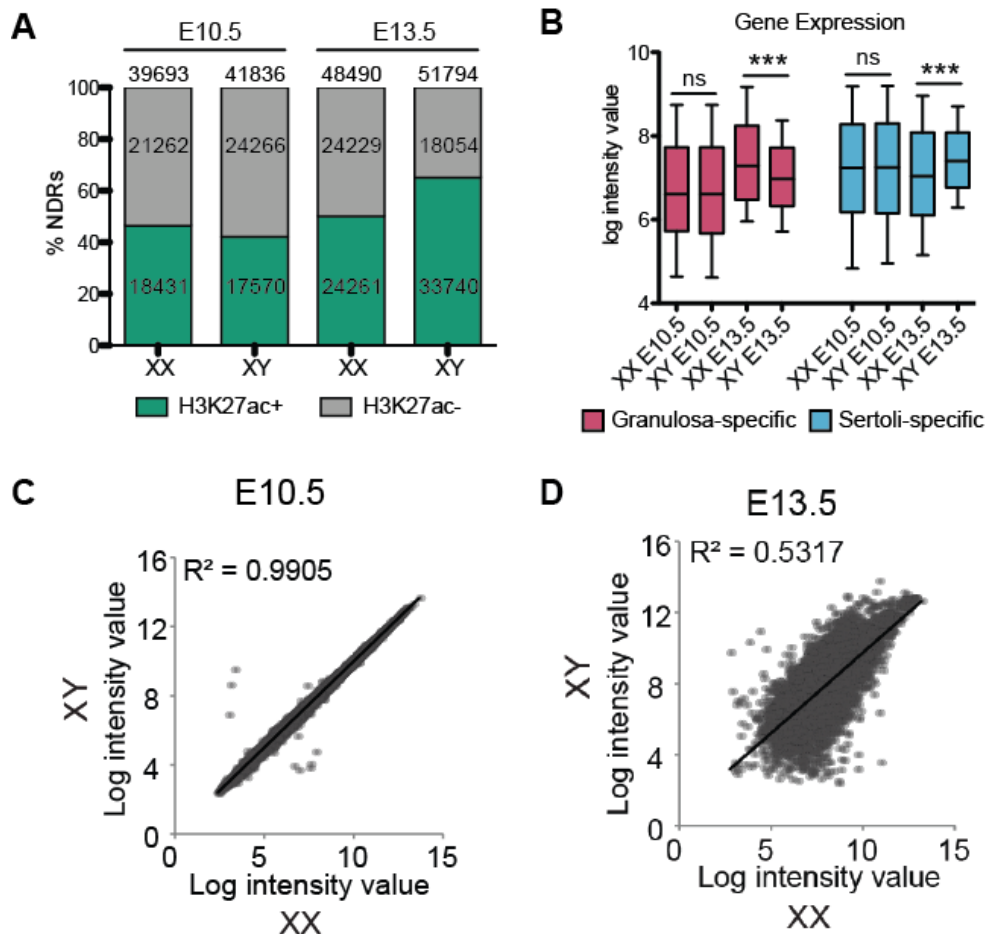
Figure 4.2. Chromatin accessibility and transcriptional profiling of XX and XY supporting cells before and after sex determination. A) Percent (and number) of H3K27ac-negative (grey) and H3K27ac-postive (green) NDRs in XX and XY cells at E10.5 (left) and E13.5 (right). B) Log intensity values from Nef et al 2005. of genes that become granulosa-specific (pink) or Sertoli-specific (blue) in XX and XY cells at E10.5 and E13.5. C-D) Log intensity gene expression values from Nef et al 2005 from XY cells (Y axis) and XX axis (X axis) at E10.5 (C) and E13.5 (D), showing similarity between the transcriptional profiles of XX and XY cells prior to sex determination and acquisition of sex-specific transcriptional profiles after sex determination.

and differentiated granulosa cells from gonads of E13.5 *TESMS-CFP* transgenic mice (Gonen et al., 2018a) (Fig. 4.1A). ATAC-seq and ChIP-seq were performed on 2 biological replicates (each replicate contained pooled cells from multiple embryos) and only NDRs present in both were used in our analyses (Fig. 4.1B).

ATAC-seq identified ~40,000 NDRs in XX and XY E10.5 progenitor cells (Fig. 4.2A). After sex determination, the number of NDRs increased from 39,693 NDRs in XX progenitor cells to 48,490 in differentiated granulosa cells, and from 41,836 NDRs in XY progenitor cells to 52,794 NDRs in Sertoli cells (Fig. 4.2A). In addition, ChIP-seq revealed that ~40% of NDRs were also marked by H3K27ac in XX and XY progenitor cells, and after sex determination, this number increased to 50% in granulosa cells, and 65% in Sertoli cells (Fig. 4.2A).

Previous transcriptional profiling showed that E10.5-E11.0 XY and XX progenitor cells were mostly indistinguishable at this stage (Fig. 4.2B and C) (Nef et al., 2005, Jameson et al., 2012b, Munger et al., 2013). Accordingly, we found that the vast majority of NDRs were shared between XX and XY progenitor cells (90% of XX NDRs (36,020/39,693) and 86% of XY NDRs (36,020/41,836)) (Fig. 4.3A and B), suggesting that these cells are not only morphologically and transcriptionally indistinguishable, but also have similar chromatin landscapes. Differentiation into either Sertoli or granulosa cells is accompanied by the upregulation of a number of Sertoli- or granulosa-promoting genes (Fig. 4.2B and D) (Jameson et al., 2012b). Consistently, there was an increase in total numbers of NDRs, from 39,393 at E10.5 to 48,490 at E13.5 in XX cells, and from 41,836 at E10.5 to 51,794 at E13.5 in XY cells (Fig. 4.2A). However, the number of overlapping NDRs between XX and XY cells decreased from 36,020 at E10.5 to 26,572 at E13.5 (55% of XX and XY NDRs) (Fig. 4.3A). Only ~40% of NDRs that were present at E10.5 were retained at E13.5; therefore, the majority (~60%) of NDRs at E13.5 were newly acquired (Fig. 4.3B). Of these newly acquired NDRs, one third were shared between Sertoli and
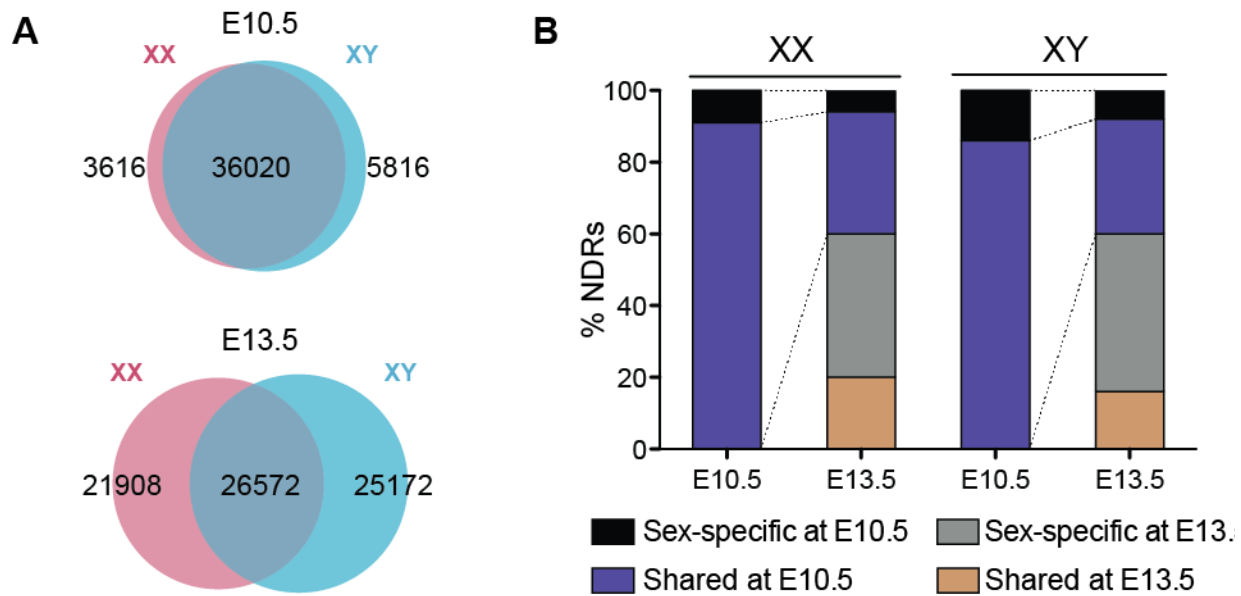
Figure 4.3. Chromatin architecture is remodeled during mammalian sex determination. A) Venn diagrams of all NDRs in XX (pink) and XY (blue) supporting cells at E10.5 (left) and E13.5 (right). B) Percent of NDRs that are shared between XX and XY cells at E10.5 (purple) and at E13.5 (orange), or specific to either XX or XY cells at E10.5 (black) or at E13.5 (grey). The purple and black bars at E13.5 represent NDRs that were retained from E10.5, while the orange and grey represent newly acquired NDRs.

granulosa cells, and two thirds were specific to either Sertoli or granulosa cells (Fig. 4.3B). Together, these data suggest that differentiation of Sertoli and granulosa cells is accompanied by an increase in Sertoli- and granulosa-specific NDRs that may promote sex-specific transcriptional programs as the cells diverge from their progenitor state into distinct cell types.

4.2.2. Gonad-specific NDRs are distal regulatory elements that neighbor Sertoli- and granulosa-promoting genes

To identify which NDRs are specific to the supporting cell lineage and more likely to regulate sex-determining genes, we removed NDRs that were also present in six other mouse tissues (primary skin fibroblasts, embryonic stem cells and adult kidney, liver, heart and brain) (Maatouk et al., 2017) from all NDRs identified in both sexes and timepoints. NDRs shared

between all cell types were termed "ubiquitous NDRs," whereas those found only in the supporting cell lineage were termed "gonad-specific NDRs." We predicted that gonad-specific NDRs would be enriched around Sertoli-promoting genes in Sertoli cells and around granulosa-promoting genes in granulosa cells. To test this hypothesis, we utilized the set of genes that becomes specific in either Sertoli or granulosa cells at E13.5 (Jameson et al., 2012b) and determined the number of NDRs associated with each gene by annotating them to the closest TSS using Hypergeometric Optimization of Motif Enrichment (HOMER). Although regulatory elements can regulate TSSs at a distance, to date the only way of properly assigning a regulatory element to its target TSS is by chromatin-conformation capture analyses. However, these same assays have shown that looping largely occurs between genes and proximal enhancers within Topologically Associated Domains (TADs) (Dekker et al., 2002). On average, the TSSs of both Sertoli- or granulosa-promoting genes neighbored one ubiquitous NDR (possibly corresponding to the promoter) in both time points and sexes, suggesting that ubiquitous NDRs are not preferentially enriched around Sertoli- or granulosa-promoting genes (Fig. 4.4A). Interestingly, gonad-specific and H3K27ac+ gonad-specific NDRs were enriched around granulosa-promoting genes as in both XX and XY cells at E10.5, suggesting that progenitor cells are predisposed towards the granulosa cells fate (Fig. 4.4B and 4.4C).

In accordance with our hypothesis, differentiation of granulosa cells was accompanied by an increase in gonad-specific NDRs neighboring granulosa-promoting genes (~2.5 NDRs/TSS), but not Sertoli-promoting genes (~1 NDR/TSS) (Fig. 4.4B). H3K27ac+ NDRs were similarly enriched around granulosa- (~1.5 NDRs/TSS) but not Sertoli-promoting genes (<1 NDR/TSS) (Fig. 4.4C). In contrast, differentiation of Sertoli cells was accompanied by an increase in gonad-specific NDRs neighboring both granulosa- and Sertoli-promoting genes (~2.5 NDRs/TSS) (Fig. 4.4B). However, H3K27ac+ NDRs, indicative of active enhancers, were
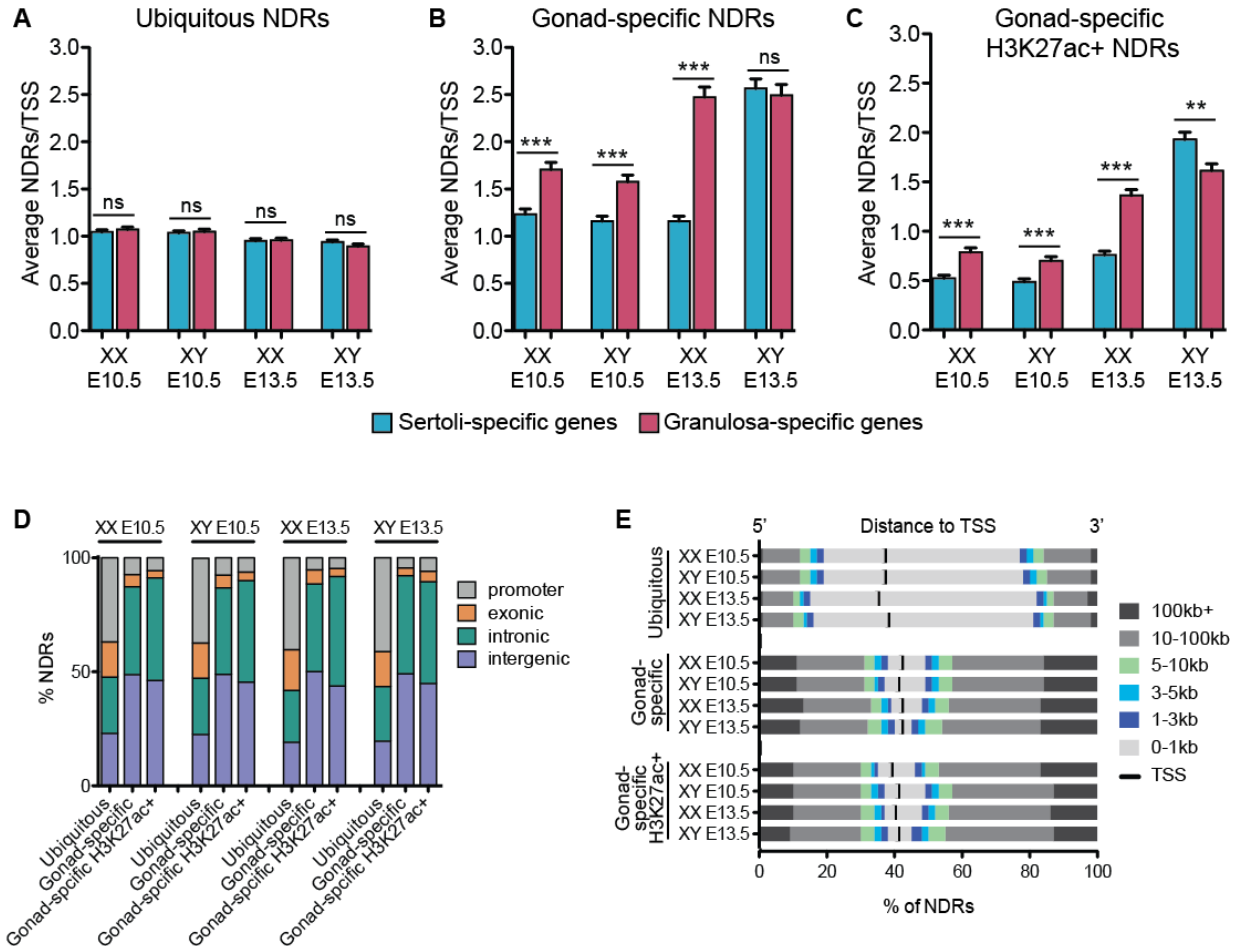
Figure 4.4. Gonad-specific NDRs are distal enhancers that neighbor Sertoli- and granulosa-promoting genes. A-C) Graphs showing the average number of NDRs associated to the nearest TSS as determined by HOMER. Based on their cell-type specificity, NDRs were classified as ubiquitous (A), gonad-specific (B) or H3K27ac+ gonad-specific (C). NDRs were associated to the nearest TSS of genes that become granulosa cell-specific (pink, 823 total genes) and Sertoli cell-specific (blue, 1026 total genes). Granulosa-specific and Sertoli-specific genes were obtained from (Jameson et al., 2012b). *** represents p<0.0001 as determined by a student's t test. D) Ubiquitous, gonad-specific and gonad-specific H3K27ac-positive NDRs were subcategorized based on their genomic location (promoter, grey; exonic, orange; intronic, green; intergenic, purple). E) Ubiquitous, gonad-specific and gonad-specific H3K27ac-positive NDRs were subcategorized based on their 5' (left) and 3' (right) distance to the associated TSS (TSS, black middle line; 0-1kb, light grey; 1-3kb, purple; 3-5kb, blue; 5-10kb, green; 10-100kb, medium grey; >100kb, dark grey).

significantly enriched mostly around Sertoli-promoting genes in Sertoli cells (~2 NDRs/TSS

vs 1.6 NDRs/TSS around granulosa-promoting genes) (Fig. 4.4C). Therefore, H3K27ac-

negative NDRs neighboring granulosa-promoting genes in Sertoli cells may be bound by

repressors that block upregulation of the female-determining pathway (Li et al., 2014).

ATAC-seq revealed that the genomic distribution of ubiquitous NDRs differed from

gonad-specific NDRs in both time points and sexes. Around 40% of ubiquitous NDRs

overlapped promoters, ~25% overlapped an intron and ~25% fell were intergenic (Fig. 4.4D).

Accordingly, the majority (~60%) of ubiquitous NDRs fell within 1kb of the TSS (Fig. 4.4E). In

contrast, only ~10% of all gonad-specific and H3K27ac+ gonad-specific NDRs overlapped a

promoter, whereas ~40% fell within an intron and ~50% were intergenic (Fig. 4.4D).

Accordingly, ~10% of all gonad-specific and H3K27ac+ gonad-specific NDRs fell within 1kb of

the TSS, while the majority of gonad-specific NDRs were >10kb away (Fig. 4.4E). Our data

suggests that NDRs that are common amongst different cell types largely represent promoters,

whereas NDRs that are specific to the supporting cell lineage are mostly intergenic,

representative of distal regulatory elements, and primarily neighbor granulosa- or Sertoli-

promoting genes.

4.2.3. Gonad-specific NDRs are enriched for transcription factor binding motifs that promote
supporting cell development

To identify transcription factor binding sites that potentially bind putative granulosa and

Sertoli regulatory elements, we conducted a TF binding motif analysis using HOMER. The TF

binding motifs identified within ubiquitous NDRs were very similar between both time points and

sexes, with CTCF and BORIS being the top represented motifs in all categories (Table 4.1 and

Table S1). BORIS is a paralog of CTCF and shares a nearly identical binding domain (Sleutels

et al., 2012). The prevalence of CTCF binding sites in ubiquitous NDRs is consistent with previous literature that shows high conservation of CTCF binding sites amongst different cell types (Kim et al., 2007a, Xi et al., 2007). Other motifs identified in ubiquitous NDRs were for ETS, ELF1, ELK1, ELK4 and FLI1 (Table 4.1 and Table S1), which are all members of the Ets (E26 transformation-specific) family of transcription factors. These ubiquitous transcription factors have a wide range of functions during development including cell differentiation, cell cycle control, proliferation and apoptosis (Sharrocks, 2001). In addition to ETS factors, binding motifs for SP1, YY1, GABPA and NRF1 identified in ubiquitous NDRs are commonly found in promoters, further supporting our finding that ubiquitous NDRs largely represent promoters (Fig. 4.4D and E) (Lin et al., 2007).

CTCF and BORIS binding sites were also enriched amongst gonad-specific NDRs, but not amongst H3K27ac+ gonad-specific NDRs (Table 4.1 and Table S1), reinforcing the idea that H3K27ac+ NDRs function as enhancers and not as other regulatory elements such as insulators (Creyghton et al., 2010, Rada-Iglesias et al., 2011).

At E10.5, other TF motifs enriched in gonad-specific NDRs in both sexes were binding motifs for the Homeobox family of genes (HOX) and its subfamilies Lim homeobox genes (LHX) and Pancreatic and Duodenal Homeobox 1 (PDX1), GATA transcription factors, and the TEA domain family (TEAD) (Table 4.1 and Table S1). While the HOX family of transcription factors are important for morphogenesis in all tissues, the LHX and GATA motifs are likely more specific to gonad morphogenesis, as *Lhx9* and *Gata4* are required for early gonad development in both sexes (Birk et al., 2000, Hu et al., 2013). Although TEAD motifs were highly represented, their role during gonad development is still unknown. Interestingly, at E10.5, SOX motifs, which likely represent the binding sites for the key Sertoli-promoting TFs SRY, SOX9 and SOX8

(Chaboissier et al., 2004, Barrionuevo et al., 2006), were only represented in the H3K27ac+ gonad-specific category in both XX and XY cells (Table 4.1 and Table 1).

At E13.5, LHX, GATA, PDX1 and TEAD motifs were still present in gonad-specific and H3K27ac+ gonad-specific NDRs in both sexes, with the addition of binding motifs for Neurofibromin 1 (*Nf1*), Estrogen-related receptor beta (*Esrrb*), and Nuclear receptor subfamily 5 group A member 2 (*Nr5a2*), the latter as the top represented motif (Table 4.1 and Table 1). The Nr5a2 motif likely represents its sibling nuclear receptor Steroidogenic Factor 1 (*Sf1* or *Nr5a1)*, which is required for early gonad formation in both sexes and is important for the upregulation of *Sox9* in Sertoli cells (Sekido and Lovell-Badge, 2008).

TCF, LHX and FOX motifs were prevalent in gonad-specific NDRs in granulosa cells (Table 4.1 and Table 1). Wnt signaling, which regulates target genes through downstream interactions between β-catenin and Tcf/Lef factors, is essential for ovarian development (Behrens et al., 1996, Vainio et al., 1999). Likewise, the forkhead box member *Foxl2* encodes a transcription factor that is required for the long-term maintenance of granulosa cell identity, as well as the repression of the male fate (Uhlenhaut et al., 2009). Finally, although *Lhx9* is required for proper gonad formation of both sexes, *Lhx9* becomes highly expressed in granulosa cells during sex determination (Birk et al., 2000).

In contrast to granulosa cells, gonad-specific NDRs in Sertoli cells were highly enriched for SOX and DMRT motifs (Table 4.1 and Table S1). As mentioned previously, *Sox9* and *Sox8* play a crucial role in driving differentiation of progenitor cells towards the Sertoli cell lineage (Chaboissier et al., 2004, Barrionuevo et al., 2006), and *Dmrt1* is required to prevent postnatal transdifferentiation of Sertoli cells into granulosa cells by repressing the female pathway (Matson et al., 2011).

		E10.5 XX		E10.5 XY		E13.5 XX		E13.5 XY	
		Motif	% Target	Motif	% Target	Motif	% Target	Motif	% Target
Ubiquitous NDRs		CTCF	19.85%	CTCF	19.02%	CTCF	11.93%	CTCF	13.69%
		BORIS	21.63%	BORIS	20.98%	BORIS	14.17%	BORIS	15.44%
		CTCF*	1.06%	CTCF*	0.98%	ELK4	16.31%	ELK4	16.23%
		ELK4	15.31%	ELK4	15.34%	ELK1	15.93%	ELK1	15.79%
		ELK1	14.96%	ELK1	14.97%	Fli1	20.55%	NFY	10.97%
		Fli1	19.68%	Fli1	19.60%	ETS	9.50%	ETS	9.67%
		ELF1	13.35%	NFY	10.61%	ELF1	14.27%	Fli1	20.22%
		ETS	8.92%	ELF1	13.39%	NFY	10.56%	Sp1	14.54%
		Sp1	13.07%	ETS	8.96%	GABPA	15.59%	ELF1	14.27%
		NFY	10.32%	Sp1	13.50%	Sp1	14.15%	GABPA	15.65%
		GABPA	14.99%	GABPA	15.00%	ETV1	20.18%	KLF14	38.75%
		ETV1	19.70%	ETV1	19.64%	CTCF*	0.57%	KLF5	29.84%
		ETS1	14.99%	YY1	2.55%	ETS1	15.14%	ETV1	20.22%
		EWS:FLI1	8.57%	ETS1	14.69%	GFY-STAF	1.65%	CTCF*	0.56%
		KLF14	36.91%	Neurod1	8.87%	EWS:FLI1	8.89%	ETS1	15.03%
Gonad-specific NDRs		CTCF	13.21%	CTCF	12.76%	Nr5a2	18.56%	Nr5a2	24.85%
		BORIS	13.21%	BORIS	12.92%	Nr5a2†	15.30%	Nr5a2†	20.38%
		Hoxc9	12.82%	Hoxc9	13.35%	NF1	10.00%	Esrrβ	18.02%
		Hoxa9	13.86%	Hoxa9	14.17%	CTCF	5.47%	CTCF	7.22%
		Pdx1	15.84%	Pdx1	16.03%	Esrrβ	13.35%	BORIS	6.99%
		REST-NRSF	2.05%	REST	2.13%	Gata3	24.86%	Errα	32.64%
		Hoxb4	4.66%	Hoxb4	4.79%	Gata4	18.36%	NF1	8.93%
		Lhx3	19.49%	Hoxa2	2.88%	Lhx2	19.07%	Gata3	22.62%
		Tead4	12.47%	Tead4	12.65%	GATA2	13.13%	Dmrt6	6.16%
		GATA2	9.33%	TEAD	10.01%	Lhx1	19.27%	Sox10	25.58%
		Hoxa2	2.59%	Nkx6.1	28.42%	GATA1	11.82%	DMRT1	6.86%
		Nkx6.1	28.21%	Lhx3	19.64%	Lhx3	25.27%	Gata4	16.10%
		GATA1	8.39%	Lhx2	13.24%	Nkx6.1	35.23%	Sox9	16.16%
		Gata3	17.79%	Tead2	7.95%	BORIS	5.54%	Sox3	26.84%
		TEAD	9.57%	RFX	2.56%	TLX	8.35%	Sox6	23.87%
H3K27ac+ gonad-specific NDRs		Hoxc9	8.43%	Hoxc9	7.98%	Nr5a2	15.15%	Nr5a2	17.65%
		Hoxa9	10.06%	Hoxa9	9.67%	Nr5a2†	12.34%	Nr5a2†	14.28%
		Pdx1	12.50%	Pdx1	12.38%	NF1	6.98%	Esrrβ	12.58%
		Sox4	11.92%	Tead4	10.56%	Gata3	22.48%	Gata4	15.20%
		Tead4	10.48%	Sox4	11.28%	Gata4	15.65%	GATA2	10.38%
		Sox10	20.39%	TEAD	8.62%	GATA2	11.11%	GATA1	9.07%
		Sox3	21.52%	Hoxb4	3.13%	Esrrβ	11.10%	Gata3	21.55%
		Tead2	6.73%	Gata4	11.93%	GATA1	9.79%	Sox10	22.88%
		TEAD	8.30%	Sox3	20.88%	Lhx2	15.78%	Sox3	24.61%
		Sox2	11.40%	Gata3	17.07%	Lhx1	16.33%	NF1	5.40%
		Lhx3	18.40%	Sox10	19.49%	NF1‡	21.83%	Sox4	12.56%
		Sox6	18.66%	Sox2	11.12%	TLX	6.44%	Errα	28.00%
		Gata4	12.18%	Tead2	6.78%	Nkx6.1	32.31%	DMRT1	4.67%
		Nkx6.1	27.40%	GATA2	8.17%	Lhx3	22.17%	Dmrt6	4.05%
		SOX15	13.35%	GATA1	7.40%	TCF21	11.82%	Sox2	13.03%

Table 4.1. Transcription factor motifs in ubiquitous, gonad-specific and H3K27ac+ gonad-specific NDRs throughout sex determination. TF motif analysis performed in HOMER for ubiquitous (top row), all gonad-specific (middle row), and H3K27ac+ gonad-specific (bottom row) NDRs in XX (grey) and XY (white) supporting cells at E10.5 (left two columns) and E13.5 (right two columns). The top 15 motifs are shown in order of significance (p<e10), along with the percent of NDRs in which these motifs are found. *CTCF satellite elements, †Nr5a2 motifs are from different ChIP experiments but have similar consensus sequences, ‡NF1 half-site. More motifs in Table S1.

Taken together, our data shows that NDRs that are specific to gonads are initially enriched for binding motifs of TFs required for early gonad formation (likely *Lhx9*, *Gata4* and *Sf1*), and they acquire NDRs enriched for binding motifs of TFs that drive either granulosa (likely *Tcf1*, *Lhx9* and *Foxl2*) or Sertoli (likely *Sry*, *Sox9, Sox8* and *Dmrt1*) cell development.

4.2.4. Most NDRs arise de novo in differentiated Sertoli or granulosa cells

We next asked whether NDRs that are sex-specific at E13.5 are primarily inherited from the progenitor state or if chromatin is opening at regulatory elements during differentiation. To do this, we identified four classes of sex-specific NDRs: 1) NDRs that were present in both XX and XY E10.5 progenitor cells and remained open in granulosa cells but closed in Sertoli cells (resolved granulosa), 2) NDRs that were present in both XX and XY E10.5 progenitor cells and remained open in Sertoli cells but closed in granulosa cells (resolved Sertoli), 3) NDRs that were not present at E10.5 in either sex but appeared *de novo* in granulosa cells (*de novo* granulosa), and 4) NDRs that were not present at E10.5 in either sex but appeared *de novo* in Sertoli cells (*de novo* Sertoli). For example, both a granulosa-specific resolved and a granulosa-specific *de novo* NDR lie downstream of *Axin2,* a gene that becomes upregulated in granulosa cells (Fig. 4.5A and B). Similarly, a Sertoli-specific resolved and a Sertoli-specific *de novo* NDR lie downstream of *Sox8*, a Sertoli-promoting gene (Fig. 4.6A and B). Our data shows that *de novo* NDRs were ~4 times more prevalent than resolved NDRs in both granulosa and Sertoli cells (Fig. 4.5C and Fig. 4.6C).

To understand whether resolved NDRs and *de novo* NDRs could have different biological functions, we performed TF motif analyses for each category using HOMER. Interestingly, we found that resolved NDRs were enriched for different TF motifs than *de novo* NDRs in both male and females. Furthermore, NDRs which resolved as either Sertoli- or

granulosa-specific also differed from each other (with the exception of CTCF and BORIS

motifs) despite being present in both sexes at E10.5 (Fig. 4.5D and E, Fig. 4.6D and E, and

Table S1).
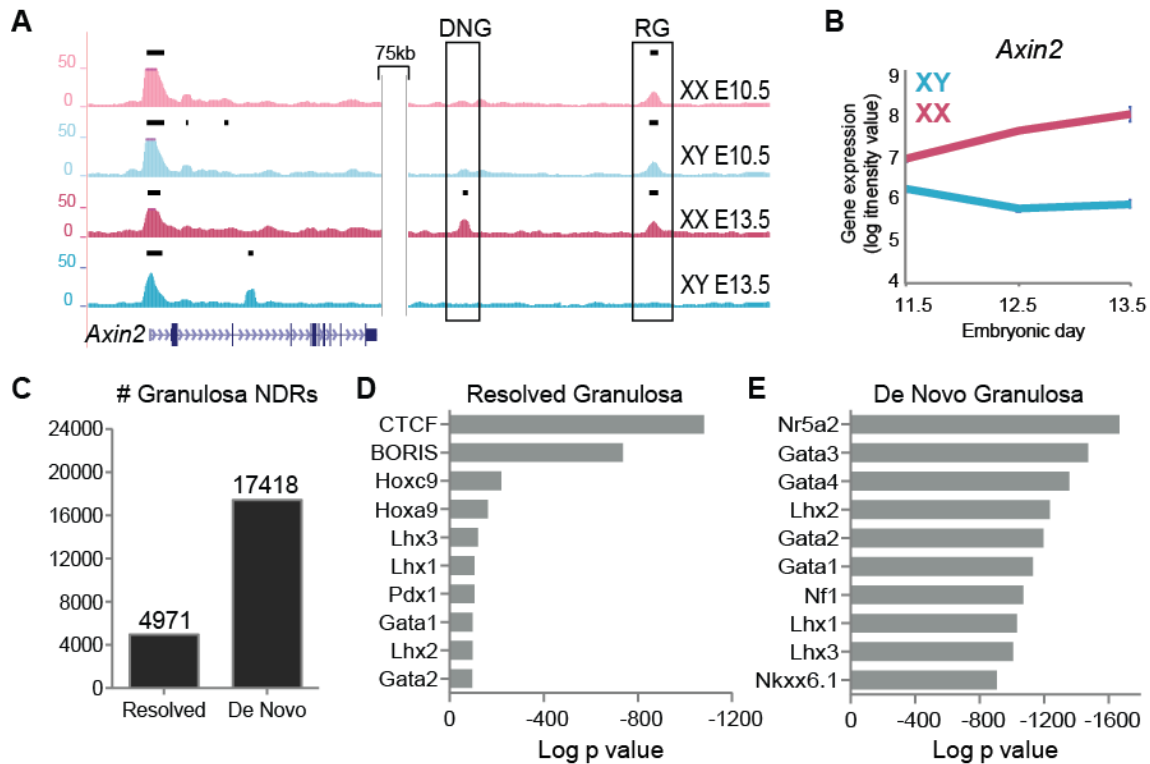
Figure 4.5. Gonad-specific NDRs are either retained from the progenitor state or arise *de novo* in granulosa cells. A) Genome Browser tracks of ATAC-seq in XX E10.5 (light pink), XY E10.5 (light blue), granulosa (dark pink) and Sertoli (dark blue) purified cells of the genomic location surrounding *Axin2*. Black boxes above tracks represent significant enrichment compared to flanking regions as determined by HOMER. A resolved granulosa (RG) and a *de novo* granulosa (DNG) NDR (boxed) lie downstream of *Axin2*. B) *Axin2* gene expression profile in purified XX (pink) and XY (blue) supporting cells throughout sex determination (E11.5-E13.5) (Jameson et al., 2012b). C) Number of resolved and *de novo* NDRs in granulosa cells. D and E) Top 10 TF binding motifs identified in resolved and *de novo* NDRs in granulosa cells ranked in order of significance (additional motifs in Table S1).
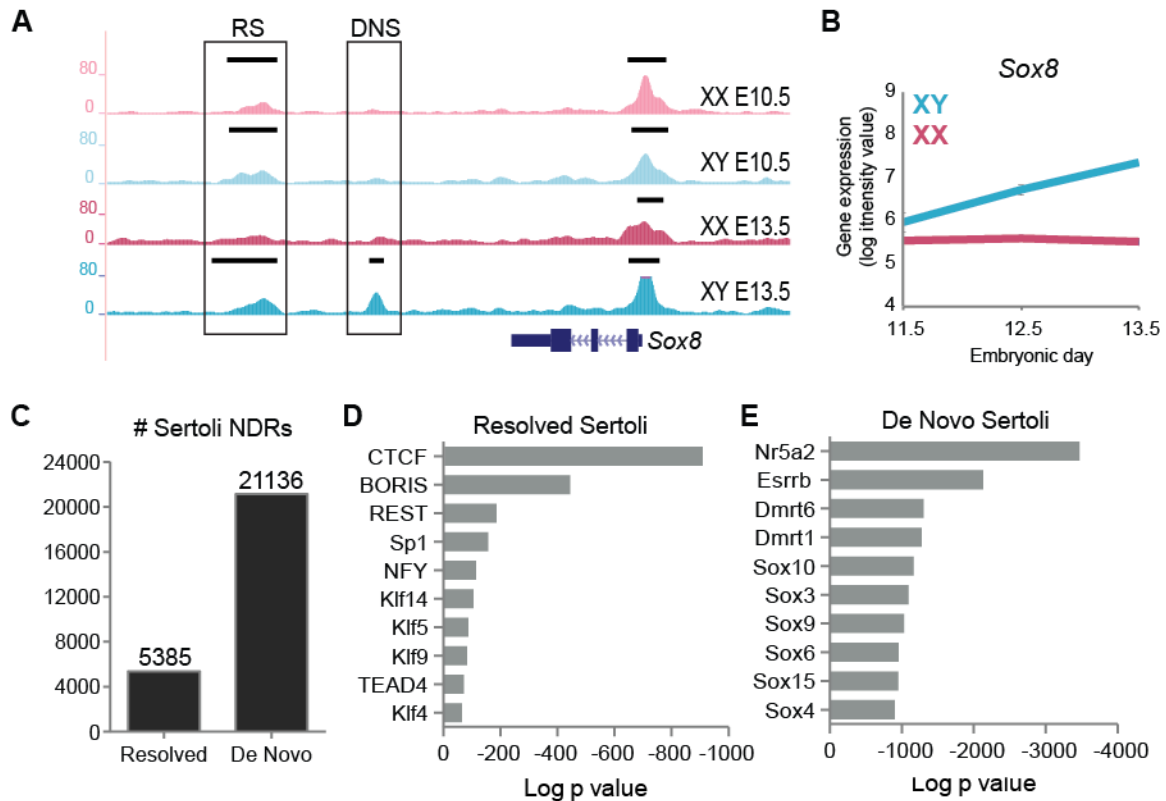
Figure 4.6. Gonad-specific NDRs are either retained from the progenitor state or arise *de novo* in Sertoli cells. A) Genome Browser tracks of ATAC-seq in XX E10.5 (light pink), XY E10.5 (light blue), granulosa (dark pink) and Sertoli (dark blue) purified cells of the genomic location surrounding *Sox8*. Black boxes above tracks represent significant enrichment compared to flanking regions as determined by HOMER. A resolved Sertoli (RS) and a *de novo* Sertoli (DNS) NDR (boxed) lie downstream of *Sox8*. B) *Sox8* gene expression profile in purified XX (pink) and XY (blue) supporting cells throughout sex determination (E11.5-E13.5) (Jameson et al., 2012b). C) Number of resolved and *de novo* NDRs in Sertoli cells. D and E) Top 10 TF binding motifs identified in resolved and *de novo* NDRs in Sertoli cells ranked in order of significance (additional motifs in Table S1).

In granulosa cells, resolved NDRs were enriched for HOX, LHX and GATA motifs, similar to motifs identified in gonad-specific NDRs at E10.5, that play an important role in early gonad development (Fig. 4.5D, Table 4.1 and Table S1). Interestingly, resolved Sertoli NDRs were enriched for different TFs, such as RE1 Silencing Transcription Factor (*Rest*), Specificity Protein 1 (*Sp1*), Nuclear Transcription Factor Y (*Nfy*), TEAD factors, and several Krüppel-like Factors (*Klf*), which likely represent *Klf4*, a transcription factor involved in Sertoli cell differentiation (Godmann et al., 2008) (Fig. 4.6D and Table S1).

In *de novo* granulosa NDRs, HOX motifs were no longer enriched, but GATA and LHX binding sites remained highly represented along with motifs for *Nr5a2, Nf1*, and NK6 homeobox 1 (*Nkx6.1)*, a transcription factor with an unknown function in gonadal development (Fig. 4.5E). TCF motifs were also identified amongst the top 25 represented motifs, but surprisingly, no FOX motifs were present in either resolved nor *de novo* granulosa NDRs (Table S1). In *de novo* Sertoli NDRs, binding sites for *Nr5a2*, SOX genes and DMRT genes dominated (Fig. 4.6E). Amongst the top 25 motifs were also members of the Ets family of proteins (Table S1), which can function downstream of the Fgf signaling cascade, a pathway important for repressing the female fate and promoting Sertoli cell development (Kim et al., 2006b).

Our data show that the majority of granulosa- and Sertoli-specific NDRs arise *de novo* during sex determination and are distinct from each other as they contain binding sites of TFs that promote either granulosa or Sertoli cell differentiation. A smaller proportion (around 25%) of NDRs were present in both sexes prior to sex determination and retained only in one sex after sex determination. These NDRs were also enriched for different TF binding motifs between granulosa and Sertoli cells. Together, our data suggests that NDRs that become sex-specific at E13.5 promote the divergence of granulosa and Sertoli cells from their common progenitor, whether or not they were present prior to differentiation.

4.2.5. ATAC-seq identifies a novel enhancer downstream of Bmp2

Until now, only a few distal regulatory elements have been identified for sex-determining genes (Sekido and Lovell-Badge, 2008, Maatouk et al., 2017, Gonen et al., 2018a). Our data set provides an excellent resource to identify distal regulatory elements, including functional enhancers that regulate Sertoli or granulosa cell development. To demonstrate this, we performed a transient transgenic assay with a putative enhancer downstream of *Bmp2.* We identified a 922bp-long, H3K27ac+ gonad-specific NDR that arises *de novo* in granulosa cells and lies 139kb downstream of the *Bmp2* TSS (its closest TSS) (Fig. 4.7A). Previous studies have shown that *Bmp2* functions downstream of *Wnt4* during ovarian development and synergizes with FOXL2 to upregulate *Fst,* a female-specific gene which causes sex reversal when disrupted (Kashimada et al., 2011, Yao et al., 2004)*.* As *Bmp2* becomes upregulated in granulosa cells during sex determination (Fig. 4.7B) (Jameson et al., 2012b), we speculated that this putative enhancer might be regulating female-specific expression of *Bmp2.*

To test the activity of this putative enhancer *in vivo*, we cloned it upstream of an *Hsp68-LacZ* reporter plasmid and used this to generate transient transgenic mice. Our reporter assay showed that this putative *Bmp2* enhancer was active in XX E13.5 gonads (5/6 transgenic embryos examined), but not in the mesonephros, consistent with its cell-type specificity (Fig. 4.7C and Fig. 4.8). Interestingly, *LacZ* expression was consistently stronger in the anterior pole of the gonad (Fig. 4.7C and Fig. 4.8). Despite this NDR being present only in granulosa cells, our transgenic reporter was also active in 2/7 XY gonads of transgenic embryos (Fig. 4.9). We identified 6 SOX, 3 TCF and 8 LHX consensus binding motifs within the putative *Bmp2* enhancer (Fig. 4.10), possibly explaining how this enhancer-reporter construct can function in both granulosa and Sertoli cells depending on which TF it interacts with during early gonadal development. The ability to drive *LacZ* expression in gonads of both sexes despite the putative

enhancer being present in only one sex is consistent with previous findings (Gonen et al., 2018a) and further highlights the bipotential nature of early supporting progenitors. Finally, the putative *Bmp2* enhancer shows sequence conservation with humans, suggesting that it may have an important role during human sex determination (Fig. 4.7A). These results demonstrate that our dataset can be used to identify functional gonad-specific enhancers that may be playing a critical role during gonadal sex determination by establishing granulosa- and Sertoli-specific transcriptional programs.

The following sections 4.2.8. and 4.2.9. are largely based on the publication: Gonen, N., Futtner, C. R., Wood, S., **García-Moreno, S. A**., Salamone, I. M., Samson, S. C., Sekido, R., Poulat, F., Maatouk D. M., Lovell-Badge, R., 2018. Sex reversal following deletion of a single distal enhancer of Sox9. *Science*, vol 29, issue 6396, pp. 1469-1473.

4.2.6. ATAC-seq identifies a novel enhancer upstream of Sox9

Sox9 functions in many embryonic and adult cell types (Jo et al., 2014), and genetic and molecular evidence suggests that its regulatory region is spread over a gene desert of at least 2 Mb 5′ to the coding sequence (Symon and Harley, 2017). The only enhancer known to be relevant for expression in Sertoli cells was TES, a 3.2-kb element mapping 13 kb 5′ from the transcriptional start site, and its 1.4-kb core, TESCO (Sekido and Lovell-Badge, 2008). Targeted deletion of TES or TESCO reduced *Sox9* expression levels in the early and postnatal mouse testis to about 45% of normal but did not result in XY female development (Gonen et al., 2017). We therefore used several unbiased approaches to systematically screen for additional gonadenhancers upstream of mouse *Sox9*, including ATAC-Seq and ChIP-seq for H3K27ac in XX and XY E10.5 and E13.5 purified supporting cells. We also used DNaseI-seq data previously obtained from E13.5 and E15.5 sorted Sertoli cells (Maatouk et al., 2017). ATAC-Seq and DNaseI-seq identified 33 putative enhancers, out of which we chose only those present at both E13.5 and E15.5 (14 enhancers) for in vivo validation by transgenic assay (Fig. 4.11A).
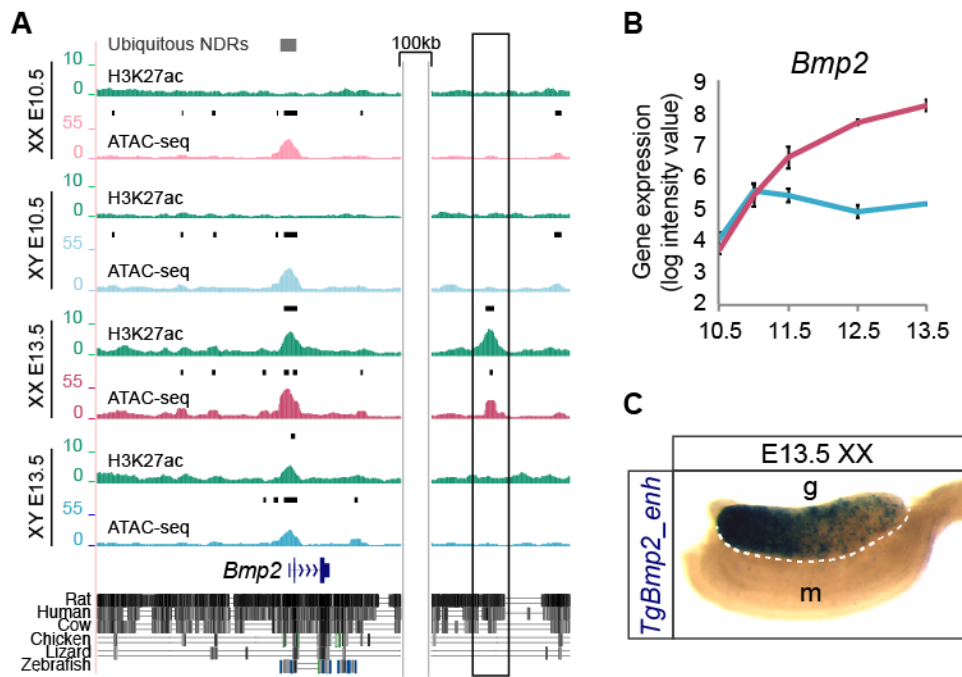
Figure 4.7. Transient transgenic analysis of a granulosa-specific NDR downstream of *Bmp2*. A) Genome Browser tracks of ATAC-seq in XX E10.5 (light pink), XY E10.5 (light blue), granulosa (dark pink) and Sertoli (dark blue) purified cells and corresponding H3K27ac ChIP-seq profiles (green) at the genomic location surrounding *Bmp2*. Black boxes above tracks represent significant enrichment compared to flanking regions as determined by HOMER. Grey boxes represent ubiquitous NDRs (top). Sequence conservation between rat, human, cow, chicken, lizard, and zebrafish are shown below. A *de novo* gonad-specific H3K27ac+ granulosa NDR downstream of *Bmp2* (boxed) was selected for transient transgenic analysis and cloned into an *hsp68-LacZ* reporter cassette. B) Gene expression profile in purified XX (pink) and XY (blue) supporting cells throughout sex determination (E10.5-E13.5) showing that *Bmp2* becomes granulosa-specific (Nef et al., 2005). C) An E13.5 XX gonad from a transgenic embryo shows β-galactosidase expression (*TgBmp2_enh,* dark blue) in the gonad (g), but not the mesonephros (m). *TgBmp2_enh* expression is stronger in the anterior pole of the gonad (left).
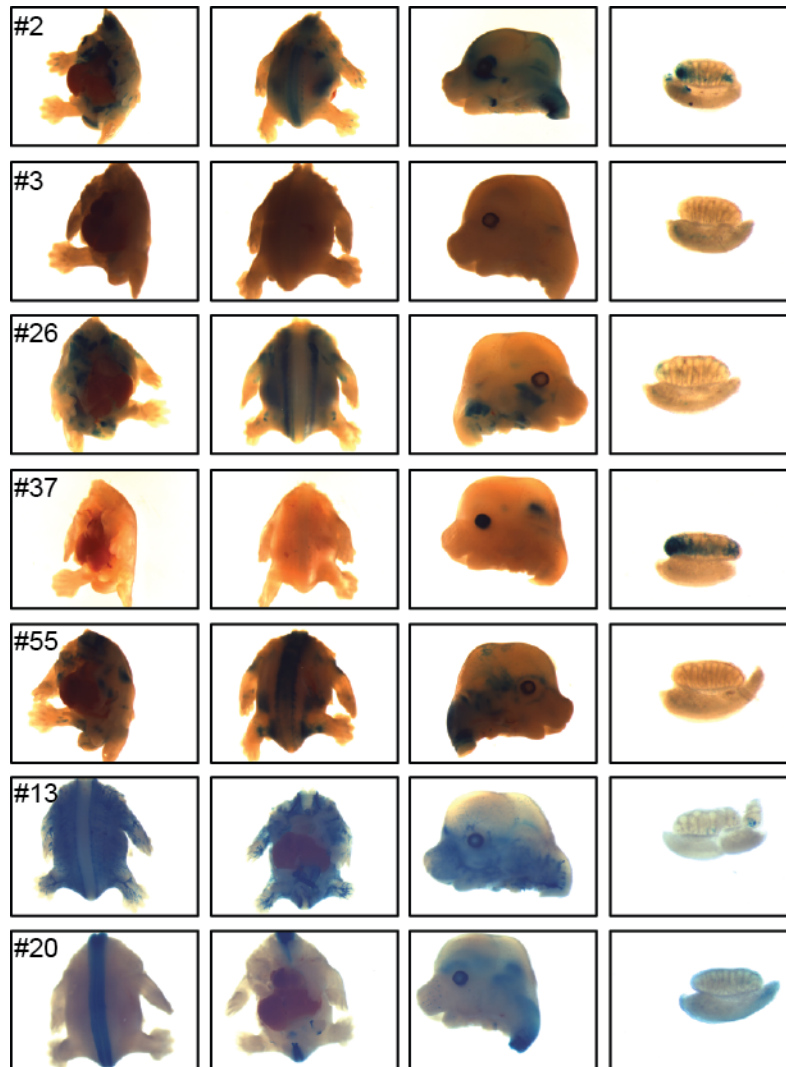
Figure 4.8. Transient transgenic analysis of a putative enhancer downstream of *Bmp2* in female embryos. β-Gal expression pattern (blue) in the body (back, left; front, right), head, and gonads of the female E13.5 transient transgenic reporter embryos that were positive for β-Gal. 5/6 positive embryos had β-Gal expression in the gonads (#11, 67, 78, 42 and 63).
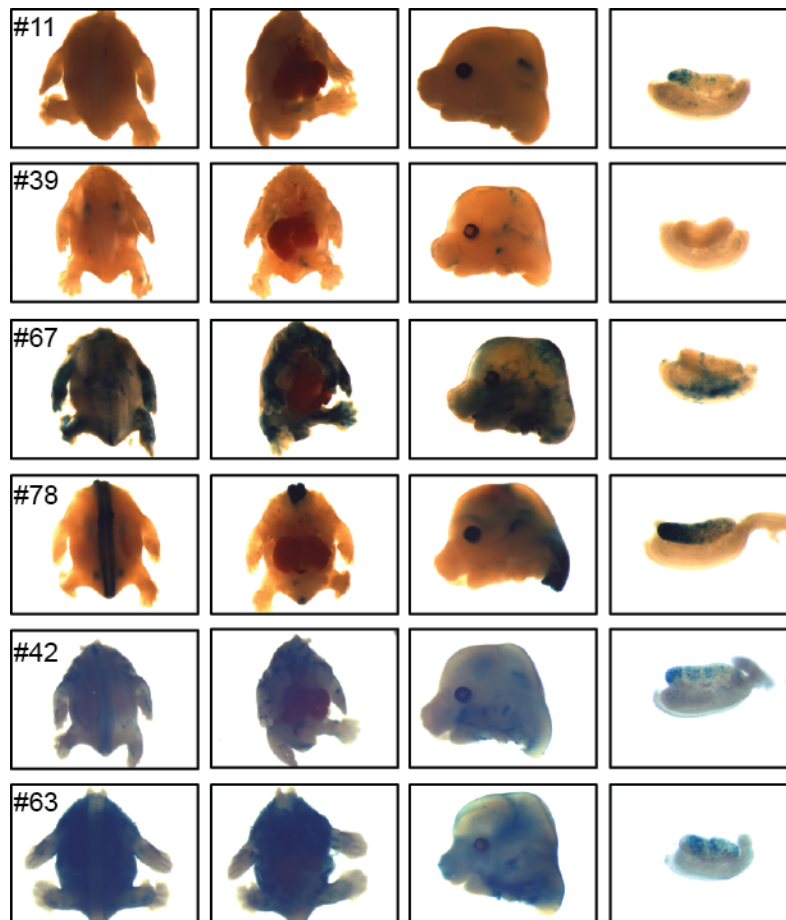
Figure 4.9. Transient transgenic analysis of a putative enhancer downstream of *Bmp2* in male embryos. β-Gal expression pattern (blue) in the body (back, left; front, right), head, and gonads of the male E13.5 transient transgenic reporter embryos that were positive for β-Gal. 2/7 positive embryos had β-Gal expression in the gonads (#2 and 37).

> [mm9, chr2:133518475–133519397 – 923 bp]
CTCCAAATATCATAGTTAACCTATTGATAATGCTCTGGTTATTACTTCATAGAAGATTGCCAGACTCCCACCAAATG
CCATTAAGAGCCCTGGTTCCTGGTTTCCTGTGAGTCACCAATCCTCCTGAATAATCTTTGGAATCAGATGTCTTTAT
GTGTCAGGTC**ATTGT**ATGCCAGCTTGCTTTAT**TCAAA**AGTAGTAC**TTGT**TCTATTTCTTTGCC**TTGT**TTTTAAACAT
GGACCCTGCTGTTTCCACACAGGAAGGAATAATGAATACTTCTGAGTTTCTATGAGACTCTTTCTCTTCTACTAT**TA
ATT**GCTTAGCTGACC**TTGT**AACTAGACTCCTCTTTGATCTGG**TAATT**TCTCATTTCATGCTTGCAGATGTTT**TCAAA**
TGTTCCAATATTTTCCCTCCCTGGTGAAGTGGACTCC**TTGT**GGTGCTTTCATACATATTTATCTTGGCCACCACGAC
TGCATTTGA**TAATT**TTTTTCTAGGA**ACAA**AAGTGATTAAGGTTGGTGATGAGAAGGAATTCCAGAGTGTGCTGATAT
CAGTTTGAGACGTTTGCATTGGGGTATCTCCTTCTATACCACTAACTAGCAGAAAGAGGTTAATATGTTTACCCATT
GACTTTATAAGTATATATTAAACATATGTTAAAAAGTAAGCCCACTAAATGCAACTTAAAAGAACTGATTTAAAAAA
A**AATTA**ATGTTTTA**TAATT**GGTGGTGGCACATGCCTT**TAATT**CCAGGAGG**AATTA**AACTTGGGAGGCAGAGGCAATT
GGATCTCAGAGTTCTAGGCCACTCTGATCTACAGAGAGAGTACTAGGATAGCCAGGGCTATACAGAGAATCCCTGTC
TCGAAAAACTAACTAACTAA**AATTA**TATTTTTAACGTTAACAGGAAGGAGCTGTATGTTGGCTGTGCATGCCTGTA

SOX9 binding motif
SOX consensus binding motif
LHX consensus binding motif
TCF consensus binding motif

Figure 4.10. SOX, LHX and TCF consensus binding motifs within the putative *Bmp2* enhancer. The DNA sequence of the granulosa-specific *Bmp2* downstream enhancer used for our *in vivo* reporter assay has 6 SOX binding motifs (blue), 8 LHX motifs (pink) and 2 TCF motifs (purple).

Of the enhancers chosen for *in vivo* reporter analysis, twelve failed to give any gonadal *LacZ* activity, although many showed staining in other tissues in which *Sox9* is normally expressed, such as chondrocytes, brain, and spinal cord (Gonen et al., 2018a). The remaining four showed gonad expression, and these constructs were reinjected to generate stable lines in order to better study their activity in both males and females during development (Gonen et al., 2018a). Of these four, "Enh13" was the strongest Sertoli cell-specific peak within this region in both the DNaseI-seq and ATAC-seq data. Enh13 (557 bp long, 565 kb 5′) is highly conserved among mammals and is located toward the distal 5′ end of a 25.7-kb region in mice that shows conserved synteny with a 32.5-kb region upstream of human *SOX9* termed XY SR, the deletion of which is associated with sex reversal (Kim et al., 2015) (Fig. 4.11A and B). Furthermore, H3K27ac ChIP-seq data marked Enh13 as active in both Sertoli and granulosa cells. ATAC-seq data from E10.5 genital ridges show that Enh13 is not open at this stage, irrespective of chromosomal sex (Fig. 4.11B), suggesting that it opens coincident with specification of the

supporting cell lineage from SF1-positive cells of the coelomic epithelium (Jakob and Lovell-Badge, 2011). Transient transgenic analysis showed that Enh13 drove *LacZ* expression in E13.5 testes, and in some cases, also in E13.5 ovaries (Fig. 4.11C), suggesting that it is an active enhancer that potentially regulates *Sox9* expression in Sertoli cells during sex determination.
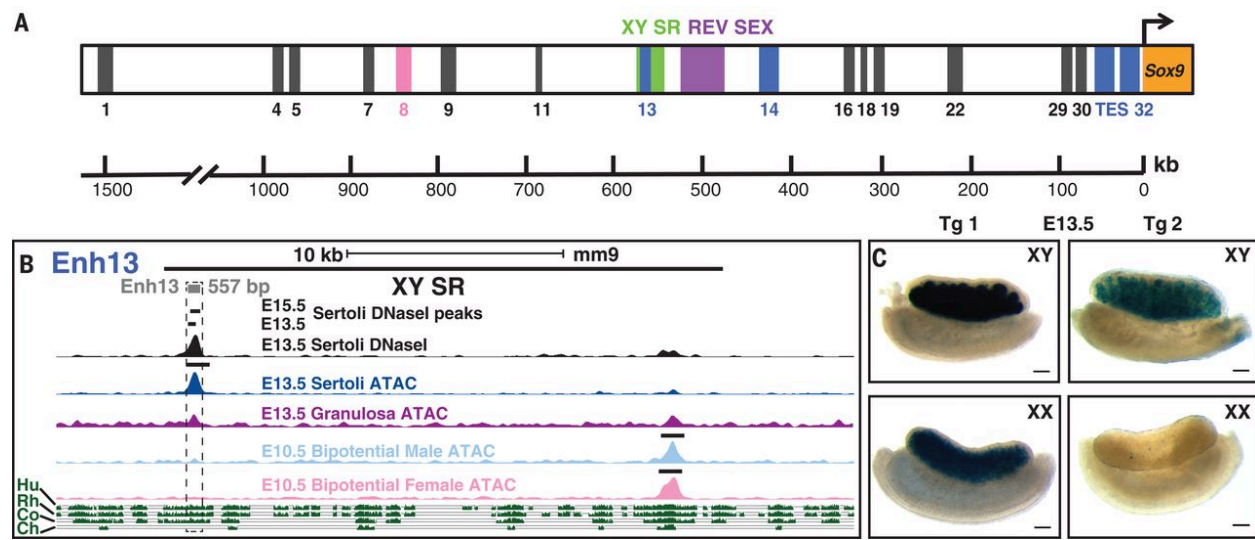
Figure 4.11. Enh13 is a testis-positive enhancer of *Sox9* located within the XY SR region. A) A schematic representation of the gene desert upstream of the mouse *Sox9* gene and the locations of the putative enhancers identified by ATAC-seq and DNaseI-seq that were screened in vivo with transgenic reporter mice. Enhancers that did not drive gonad expression of LacZ are shown in gray. Enhancers that drove testis-specific and ovary-specific LacZ expression are shown in blue and pink, respectively. The mouse regions that show conserved synteny with the human XY SR and REV SEX are depicted in green and purple boxes, respectively. B) Enh13 (gray box) is located at the 5′ side of the 25.7-kb mouse equivalent XY SR locus (heavy black line). Data from DNaseI-seq (black) on E15.5 and E13.5 XY sorted Sertoli cells and ATAC-seq on E13.5 sorted Sertoli cells (blue) and granulosa cells (purple), as well as E10.5 sorted somatic cells, at the Enh13 genomic region are presented. Peaks correspond to nucleosome-depleted regions and are marked by a black horizontal line if they are significantly enriched compared to flanking regions, as determined by model-based analysis of ChIP-seq, and present in at least two biological replicates. The gray box overlaying the peak indicates the cloned fragment. Green areas represent sequence conservation among mice, humans (Hu), rhesus monkeys (Rh), cows (Co), and chickens (Ch) (sequence conservation tracks were obtained from the University of California–Santa Cruz). C) β-Gal staining (blue) of E13.5 testes and ovaries from two representative independent stable Enh13 transgenic (Tg) lines. Scale bars, 100 μm.

4.2.7. Deletion of the distal upstream enhancer of Sox9, Enh13, causes male-to-female sex reversal

To determine whether Enh13 is a critical enhancer of *Sox9* expression and testis development, we used CRISPR/Cas9 genome editing to derive mice homozygous for deletions of Enh13. Homozygous deletion always led to XY female development on a wild-type background (Fig. 4.12A). This result was surprising, because if TES accounts for 55% of *Sox9* expression in early Sertoli cells, any additional enhancer(s) should not account for more than 45% and, when this enhancer is deleted, levels of *Sox9* should remain higher than the threshold of ~25% below which sex reversal might be expected (Gonen et al., 2017). Nevertheless, whereas XY Enh13$^{+/-}$ embryos still underwent normal testis development, XY Enh13$^{-/-}$ embryos produced ovaries indistinguishable from those of XX wild-type embryos, with no signs of testis cords or a coelomic vessel (Fig. 4.12B and C). Immunofluorescence analysis of E13.5 XY Enh13$^{+/-}$ and Enh13$^{-/-}$ gonads for SOX9 and FOXL2 showed that the former are still testes whereas the latter are fully sex-reversed ovaries (Fig. 4.12C). Similar analysis of XX Enh13-deleted gonads did not show any obvious phenotype (Gonen et al., 2018a). Quantitative PCR analysis revealed that while *Sry* mRNA levels were unchanged in E11.5 XY Enh13$^{-/-}$ gonads, *Sox9* levels decreased to 21% compared to WT gonads, and were similar to levels in XX gonads at this same stage, suggesting that Enh13 is required to upregulate *Sox9* in the early XY gonad (Gonen et al., 2018a).

Enh13 contains a single consensus SRY binding site in mice as well as a SOX9 site to which SRY can also bind (Gonen et al., 2018a). ChIP-qPCR on E11.5 gonads dissected from *Sry-Myc*transgenic embryos was performed using a specific antibody against the MYC tag (Sekido et al., 2004). There was an 11-fold enrichment in SRY-MYC–positive versus negative gonads with primers spanning the SOX9 consensus site and a six-fold enrichment with primers

spanning the SRY site whereas primers against the strongest SRY binding site in TESCO

(Sekido et al., 2004)showed only a five-fold enrichment (Gonen et al., 2018a). ChIP assays also

revealed SOX9 to be bound at similar levels to both Enh13 and TESCO in cells from E13.5

testes (Gonen et al., 2018a).

Our data supports the conclusion that Enh13 is a critical enhancer of *Sox9* in Sertoli

cells and is required for testis development. Furthermore, Enh13 plays a more significant role

than TES during early gonadal development.
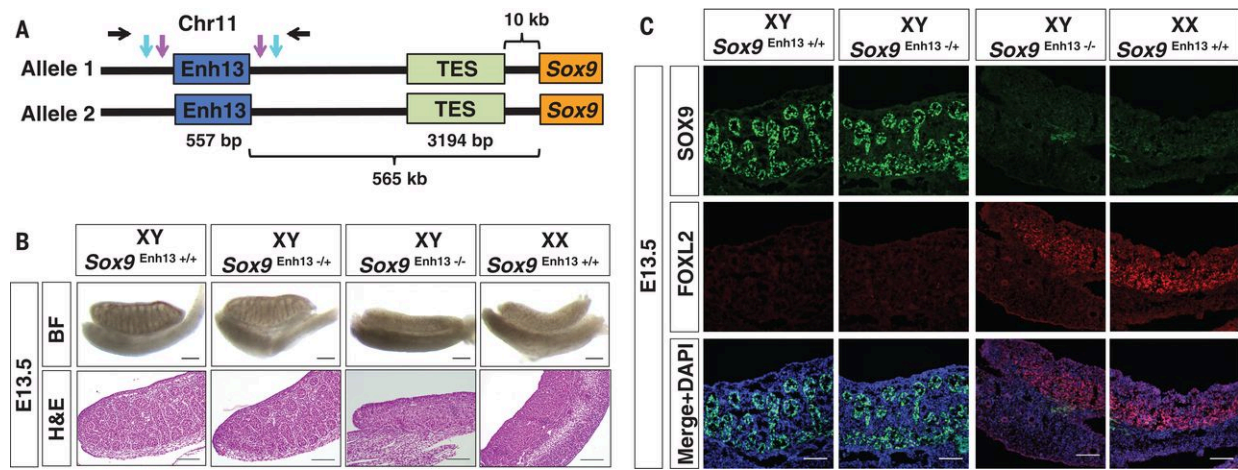
Figure 4.12. Deletion of Enh13 leads to complete XY male-to-female sex reversal. A) A schematic representation of the locations of Enh13 and TES upstream of *Sox9*. Blue and purple arrows represent the external and internal single guide RNAs, respectively, used to delete Enh13. Black arrows represent the PCR primers used to genotype embryos and mice with Enh13 deletion. Chr11, chromosome 11. B) Bright-field (BF) pictures and hematoxylin and eosin (H&E)–stained sections of E13.5 XY Enh13$^{+/+}$, Enh13$^{+/-}$, and Enh13$^{-/-}$ and XX Enh13$^{+/+}$ gonads. C) Immunostaining of E13.5 XY wild-type, Enh13$^{+/-}$, and Enh13$^{-/-}$ and XX wild-type gonads. Gonads were stained for Sertoli marker SOX9 (green), granulosa marker FOXL2 (red), and 4′,6-diamidino-2-phenylindole (DAPI) (blue). Sex-reversed gonads are indistinguishable from wild-type XX gonads, whereas the heterozygous deletion does not appear to alter testis morphogenesis. Scale bars in B) and C), 100 μm.

4.3. Discussion

Cis-regulatory elements, such as enhancers, play a critical role during development by coordinating the precise spatiotemporal expression of each gene within a transcriptional network. Consequently, disruption of these elements has the potential to cause congenital disorders. To increase our understanding of the complex and antagonistic developmental networks underlying mammalian sex determination, we identified *in vivo* cis-regulatory elements by employing whole-genome approaches.

We used ATAC-seq to assay for regions of open chromatin coupled with ChIP-seq for H3K27ac, a marker of active enhancers and promoters (Creyghton et al., 2010), in purified XX and XY supporting cells before (E10.5) and after (E13.5) gonadal sex determination. We show that prior to sex determination, XX and XY progenitor cells have similar chromatin accessibility patterns that resolve into sex-specific patterns as they differentiate into either granulosa or Sertoli cells. To distinguish NDRs common to many cell types (ubiquitous) from those found only in the supporting cell lineage (gonad-specific), we made use of previously published DNaseI-seq data from mouse ESCs, fibroblasts, heart, liver, kidney, and brain (Maatouk et al., 2017). Although these NDRs were identified by DNaseI-seq, ATAC-seq was previously shown to have a high correlation with DNaseI-seq, with the advantage of requiring very few numbers of cells (Buenrostro et al., 2013). We found that, in contrast to ubiquitous NDRs, gonad-specific NDRs as well as H3K27ac+ gonad-specific NDRs were enriched around granulosa-promoting genes in granulosa cells and Sertoli-promoting genes in Sertoli cells. Furthermore, we show that binding motifs for TFs that promote granulosa- or Sertoli-development are enriched in gonad-specific NDRs that neighbor granulosa- or Sertoli-promoting genes, respectively. Finally, we demonstrate the power of our dataset to identify enhancers by characterizing a novel enhancer

downstream of *Bmp2* and a novel enhancer upstream of *Sox9* that function in gonads of transgenic mice as predicted.

The progenitor cells of the gonad exist in a bipotential state with the ability to differentiate into either Sertoli (XY) or granulosa (XX) cells depending on the developmental signal they receive (Albrecht and Eicher, 2001). Extensive transcriptional profiling of these cells at E10.5 has shown that less than 1% of genes are transcribed in a sexually dimorphic way and are likely Y-linked genes or X-linked genes that escape X inactivation (Nef et al., 2005, Jameson et al., 2012b, Munger et al., 2013). Therefore, autosomal genes that promote Sertoli or granulosa cell development (such as SOX and Wnt signaling genes) are initially transcribed at similar levels in both XX and XY progenitor cells. However, despite increasing evidence that cell differentiation is epigenetically regulated, it remains unclear whether the chromatin state of progenitor cells plays a role in maintaining their bipotential nature and establishing their male or female fate. We predicted that the chromatin accessibility landscape of XX and XY progenitor cells would also be similar. Accordingly, we found that the vast majority (~90%) of NDRs are shared between males and females prior to gonadal sex determination, and the enrichment patterns of TF binding motifs within these NDRs are nearly identical in XX and XY supporting cells at this stage. Interestingly, gonad-specific NDRs were enriched around granulosa-promoting genes in both sexes at E10.5, suggesting that both XX and XY progenitor cells are predisposed towards the granulosa cell lineage. This is consistent with previous findings that show that there are more granulosa-promoting genes expressed at the bipotential stage than there are Sertoli-promoting genes, and initiation of the male pathway requires upregulation of the testis pathway and simultaneous repression of the ovarian pathway (Jameson et al., 2012b).

An intriguing finding was an enrichment of SOX binding motifs in H3K27ac+ gonad-specific NDRs at E10.5 in both sexes. This suggests that regions for SOX binding sites are not

only accessible but active at E10.5. As *Sry* and SOX genes share at least 50% homology in their DNA binding domains, the presence of SOX motifs at E10.5 could also represent binding sites for SRY itself (Denny et al., 1992a). This finding indicates that the bipotential supporting lineage is poised to rapidly engage the male pathway upon *Sry*'s transient activation and downstream *Sox9* upregulation, regardless of its chromosomal complement. In support of this, transgenic expression of *Sry* in early XX gonads is sufficient to drive Sertoli cell development (Koopman et al., 1991, Hiramatsu et al., 2009). In absence of *Sry* or if *Sry* expression is delayed outside of its functional window (~12 to 15 tail somites)*,* the pro-ovarian Wnt signaling pathway is not repressed and the female pathway becomes engaged (Hiramatsu et al., 2009). Therefore, the absence of SOX motifs in granulosa cell NDRs likely represents a closure of early SOX motif-containing NDRs that were not bound by *Sry* nor *Sox9* (Vainio et al., 1999, Chassot et al., 2008b)*.* The similarity of the enhancer distribution and binding sites within these enhancers between XX and XY cells prior to sex determination is consistent with, and possibly contributes to, their similar transcriptional profiles and ability to engage either pathway.

As enhancers play a critical role in establishing tissue-specific gene expression patterns during development, we predicted that differentiation of Sertoli and granulosa cells would require enhancers that upregulate Sertoli- and granulosa-promoting genes, and that these enhancers would be gonad-specific. In accordance with our hypothesis, we found only 55% of NDRs were shared between Sertoli and granulosa cells at E13.5 in contrast to ~90% at E10.5. As expected, H3K27ac+ gonad-specific NDRs were more likely to neighbor granulosa-promoting genes in granulosa cells, and Sertoli-promoting genes in Sertoli cells, suggesting that these enhancers establish granulosa- and Sertoli-specific transcriptional profiles. In support of this idea, our data shows that granulosa-specific NDRs are enriched with binding sites of TFs important for granulosa cell differentiation such as TCF and LHX factors, whereas Sertoli-

specific NDRs are enriched for SOX and DMRT motifs, TFs required for Sertoli cell development and maintenance. The sibling nuclear receptors SF1 (NR5A1) and NR5A2 share a binding motif, identified in the HOMER database as NR5A2, which was the most significantly represented motif in both XX and XY gonad-specific and H3K27ac+ gonad-specific NDRs. *Sf1* is crucial for early gonadogenesis in both sexes as well as for the initiation of the male pathway through *Sox9* signaling (Sekido and Lovell-Badge, 2008). Importantly, FOXL2, which is required for ovarian development and maintenance, has also been shown to bind this motif (Benayoun et al., 2008). Since the precise binding motif for FOXL2 is not included in HOMER's TF motif database, we speculate that the NR5A2 motifs identified in Sertoli and granulosa cell NDRs in our analyses represent both SF1 and FOXL2; however, further analyses using CHIP is needed to determine which TFs are binding these sites.

Sertoli cells can transdifferentiate into granulosa cells upon loss of *Dmrt1* even long after sex determination has occurred, emphasizing the need to actively repress the female-fate even during adulthood (Matson et al., 2011). In accordance with previous studies, we found that all gonad-specific NDRs in Sertoli cells were equally enriched around Sertoli- and granulosa-promoting genes, whereas H3K27ac+ gonad-specific NDRs were mostly enriched only around Sertoli-promoting genes (Maatouk et al., 2017). It was originally speculated that NDRs neighboring granulosa-promoting genes in Sertoli cells could be due to a delay in chromatin closure at these sites (Maatouk et al., 2017). However, our data clearly show an increase in the average number of NDRs neighboring granulosa-promoting genes from E10.5 to E13.5 in XY cells, suggesting that these NDRs are newly acquired. Therefore, we favor the interpretation that these NDRs, which are likely H3K27ac-negative, are bound by TFs that maintain the Sertoli fate by blocking upregulation of granulosa-promoting genes (Maatouk et al., 2017).

Interestingly, our data also found that the majority (~75%) of sex-specific NDRs were not inherited from a progenitor state, but rather appeared *de novo* following sex determination. When we analyzed the TF motifs enriched in NDRs that resolved from a progenitor state into sex-specific NDRs versus those that arise *de novo*, we found that each category was enriched for different TFs. *De novo* NDRs were enriched for TFs that drive supporting cell development, similar to the gonad-specific NDRs discussed above. Of interest, binding sites for CTCF were present in both resolved granulosa and resolved Sertoli NDRs, suggesting that different CTCF sites are maintained during sex determination depending on whether bipotential cells engage the male or female pathway. CTCF factors are important for the formation of TADs, higher-order chromatin structures that regulate promoter-enhancer interactions by confining them into specific regions (Ghirlando and Felsenfeld, 2016). Taken together, our data suggests that chromatin undergoes extensive remodeling within the supporting cell lineage throughout sex determination, including loss and gain of sex-specific NDRs that interact with different TFs and promoters.

Apart from increasing our understanding of the role that chromatin remodeling plays *in vivo* during a key developmental event, this dataset will provide a useful tool for the identification of functional gonad-specific enhancers and associated TFs that drive mammalian sex determination. As proof of concept, we identified a *de novo* granulosa enhancer 139kb downstream of the granulosa-promoting gene *Bmp2* and showed that it was functional in ovaries of transgenic mice by performing an *in vivo* reporter assay. Our finding that the majority of gonad-specific NDRs fall >10kb away from a TSS highlights the importance of using unbiased, genome-wide techniques for the identification of critical regulatory elements.
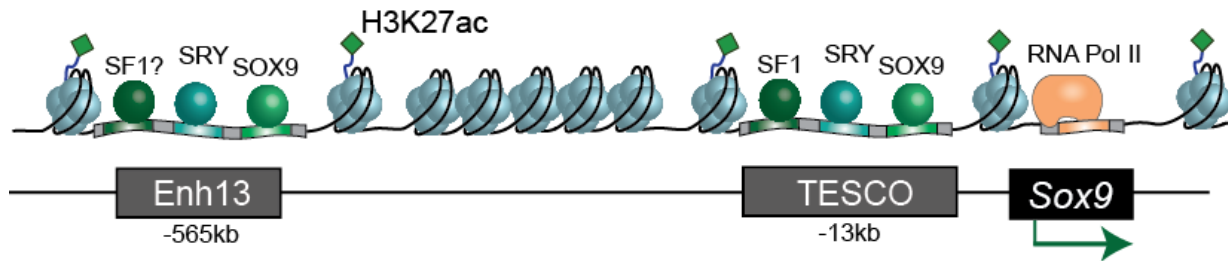
Figure 4.13. Model for Enh13 and TESCO regulation of *Sox9* in mice. Enh13 lies 565kb upstream of *Sox9,* and TESCO lies 13kb upstream of *Sox9,* and both enhancers are marked by H3K27ac in E13.5 Sertoli cells. Enh13 is bound by SRY and SOX9 in E11.5 XY gonads, whereas TESCO is bound by SF1, SRY and SOX9. Preferential binding of SRY to Enh13 over TESCO supports the hypothesis that Enh13 is critical for specification of the male fate by upregulating *Sox9* immediately after *Sry* activation, whereas TESCO acts at a later timepoint and is partially redundant to Enh13. Binding of SOX9 to Enh13 and TESCO suggests that once *Sox9* is upregulated, it autoregulates its expression to maintain its upregulation.

In addition to this putative *Bmp2* enhancer, our dataset led to the identification of an enhancer 565kb upstream of *Sox9* (named Enhancer 13) that, when deleted, caused complete male-to-female sex reversal (Gonen et al., 2018a). The notion of redundancy or "shadow enhancers" within a regulatory region is well established (Lagha et al., 2012, Visel et al., 2009), and recent data suggest that deletion of single, even "ultraconserved" enhancers from developmentally important genes can have at most subtle if not undetectable effects (Dickel et al., 2018, Osterwalder et al., 2018). It was therefore remarkable to see that deleting Enh13 alone phenocopies the loss of *Sox9* itself within the supporting cell lineage (Lopes et al., 2016, Barrionuevo et al., 2006).

Substantial evidence points to the time-dependent action of SRY on *Sox9*. Preferential binding of SRY to Enh13 over TESCO at E11.5 supports the hypothesis that the former is more critical because it initiates up-regulation of *Sox9* while the latter is secondary (Fig. 4.13). If *Sox9* fails to reach a critical threshold within a few hours, then ovary-determining and/or anti-testis factors, such as Wnt signaling, accumulate to a sufficient level to repress *Sox9* and make it refractory to male-promoting factors, including SRY, even though expression of the latter

persists in XY gonads when Sertoli cells fail to differentiate (Hiramatsu et al., 2009). The finding that SOX9 itself binds Enh13 suggests that, like TES, Enh13 is used by SOX9 to autoregulate its expression (Fig. 4.13). We suggest that Enh13 is an early-acting enhancer, such that, without it, *Sox9* transcription fails to increase to a level where the other enhancers can act before the gene is silenced. It is only later that TES, and perhaps other enhancers, begins to act in a more redundant fashion, although it is conceivable that each enhancer has a major role to play during distinct phases of Sertoli cell development from the fetal to the adult testis. It will be of interest to determine how Enh13 activity cascades into the recruitment of the other enhancers.

Although identified in mice, Enh13 maps to a region of the human genome known as XY SR, deletions and duplications of which have resulted in human DSDs (Kim et al., 2015). Unlike several other gonadal enhancers, Enh13 appears well conserved, has a clear role in mice to initiate up-regulation of *Sox9* expression, in response to SRY activity, and it may contribute to its maintenance. (Sekido et al., 2004). This makes it very likely to play a similar role in humans given its location within the XY SR region (Kim et al., 2015, Ohnesorg et al., 2016). If so, heterozygosity for Enh13 deletions in humans should mimic heterozygosity for null mutations in SOX9, with XY female sex reversal occurring in about 70% of cases but perhaps without other CD phenotypes. Thus, our dataset is not only predictive of enhancers driving expression of sex-determining genes in mice but has the potential to increase the number of diagnoses in human DSDs with unknown etiologies.

CHAPTER FIVE

CONCLUSIONS AND FUTURE DIRECTIONS

In vertebrates, the formation of a bipotential gonad is the first and critical step for the development of sexually dimorphic organisms. The unique ability of the gonad to engage one of two possible developmental pathways during sex determination provides a valuable model to study cell fate decisions and transitions. Accumulating evidence suggests that various epigenetic mechanisms and cis-regulatory elements underlie the resolution of cell fate in the early gonad by fine-tuning the timing and expression levels of sex-determining genes. However, until now it was unclear whether the chromatin state of progenitor cells plays a role in maintaining their bipotential nature and establishing their male or female fate.

Our findings suggest that the plasticity of the bipotential gonad is reflected in (and perhaps fostered by) a poised chromatin landscape that resolves through the activation of genes associated with one fate, and the maintenance of bivalency at genes associated with the opposite fate. We show that XX and XY progenitor cells approach sex determination with similar chromatin landscapes. The finding that, before sex determination, both male- and female-promoting genes in XX and XY progenitor cells are poised for activation, and that active enhancers are enriched for SOX binding motifs, supports the idea that these cells are competent to rapidly engage the male pathway following *Sry* expression, but can follow the female pathway in absence of this signal. We are the first to show that XX and XY progenitor cells are not only morphologically and transcriptionally indistinguishable, but also have indistinguishable chromatin landscapes, possibly contributing to their bipotential nature. However, it is important to note that our studies are restricted to three histone marks (H3K27ac, H3K27me3 and H3K4me3) out of hundreds. In one study, the histone mark H3K9acK14ac predisposed a bipotential lineage towards either the pancreatic or liver cell fate (Xu et al., 2011). Other studies have shown that in pluripotent lineages, H3K4me1 primes enhancers for future activation (Rada-Iglesias et al., 2011). Therefore, moving forward it will be interesting to

investigate whether other histone modifications may be predisposing gonadal progenitor cells towards their male or female fate.

The activation of genes associated with the male or female fate is initiated by transcription factors (i.e. SRY in males). However, although TFs initiate transcriptional changes, their ability to bind DNA is modulated by the chromatin landscape. The finding that H3K27me3 is removed from the promoter of bivalent genes that become upregulated further suggests that TFs and chromatin modifiers must act in tandem to regulate target genes. The resolution of bivalency towards an active H3K4me3-only state during sex determination is likely mediated by an H3K27me3-specific demethylase. In mammals, there are two H3K27me3 demethylases: *Kdm6b* and *Utx.* Importantly, *Kdm6b* has been shown to be critical for the activation of the male switch *Dmrt1* in turtles by removing H3K27me3 during sex determination (Ge et al., 2017). Accordingly, loss of *Kdm6b* in turtles leads to male-to-female sex reversal (Ge et al., 2018). Future studies will determine whether *Kdm6b* and/or *Utx* are critical for the upregulation of male- or female-determining genes during sex determination in mammals.

At present, commitment to one fate or the other does not appear to be associated with an epigenetic lock-out of the alternative pathway. Instead, our data show that the alternate pathway remains bivalent even after sex determination has occurred. Although we show that the chromatin modifier CBX2 plays an important role in stabilizing the male fate by silencing bivalent female-promoting genes, the observation that bivalency is retained even at silenced genes suggest that maintenance of cell fate depends on TFs rather than more stable epigenetic repressive mechanisms such as formation of heterochromatin. This is supported by the finding that sex-specific NDRs in differentiated Sertoli and granulosa cells are enriched for TFs that promote their differentiation (SOX and DMRT motifs in Sertoli cells, LHX, TCF and FOXL2 in granulosa cells). Dependence on TFs rather than more stable repressive mechanisms may not

be unusual between lineages that originate from a common precursor, especially when there is an evolutionary advantage to maintaining plasticity. In the case of the supporting cell lineage in the gonad, one possibility is that this is an evolutionary holdover from species that undergo sex reversal, but this idea has not been investigated. The interlacing of TFs and the chromatin landscape during cell fate commitment and maintenance is an important area of future study, which will benefit from the unique characteristics of the bipotential gonad.

In addition to advancing our understanding of the basic biological principles underlying molecular and cellular fate commitment and development, increasing our understanding of the epigenetic landscape during sex determination has implications for our ability to diagnose human DSDs with unknown etiologies. Currently ~80% of patients do not receive a genetic diagnosis. In many of these undiagnosed cases, the cause may be due to mutations residing in non-coding regions that escape conventional diagnostic screens such as karyotyping, sequencing of individual genes or even whole-exome sequencing. In this thesis, we have generated datasets that define regulatory regions beyond the coding genome in mice, and have validated these datasets by identifying an enhancer critical for testis development. In the future, determining the extent to which these regulatory regions are conserved in humans will increase our ability to identify disease-causing mutations and increase diagnoses in individuals with DSDs.

REFERENCES

AGARWAL, N., HARDT, T., BRERO, A., NOWAK, D., ROTHBAUER, U., BECKER, A., LEONHARDT, H. & CARDOSO, M. C. 2007. MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. *Nucleic Acids Res,* 35**,** 5402-8.

AGGER, K., CLOOS, P. A., CHRISTENSEN, J., PASINI, D., ROSE, S., RAPPSILBER, J., ISSAEVA, I., CANAANI, E., SALCINI, A. E. & HELIN, K. 2007. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature,* 449**,** 731-4.

AGIRRE, X., VIZMANOS, J. L., CALASANZ, M. J., GARCIA-DELGADO, M., LARRAYOZ, M. J. & NOVO, F. J. 2003. Methylation of CpG dinucleotides and/or CCWGG motifs at the promoter of TP53 correlates with decreased gene expression in a subset of acute lymphoblastic leukemia patients. *Oncogene,* 22**,** 1070-2.

ALBRECHT, K. H. & 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.

ANWAY, M. D., FOLMER, J., WRIGHT, W. W. & ZIRKIN, B. R. 2003. Isolation of sertoli cells from adult rat testes: an approach to ex vivo studies of Sertoli cell function. *Biol Reprod,* 68**,** 996-1002.

ARNOLD, A. P. 2009. Mouse models for evaluating sex chromosome effects that cause sex differences in non-gonadal tissues. *J Neuroendocrinol,* 21**,** 377-86.

ASSUMPCAO, J. G., BENEDETTI, C. E., MACIEL-GUERRA, A. T., GUERRA, G., JR., BAPTISTA, M. T., SCOLFARO, M. R. & DE MELLO, M. P. 2002. Novel mutations affecting SRY DNA-binding activity: the HMG box N65H associated with 46,XY pure gonadal dysgenesis and the familial non-HMG box R30I associated with variable phenotypes. *J Mol Med (Berl),* 80**,** 782-90.

ATLASI, Y. & STUNNENBERG, H. G. 2017. The interplay of epigenetic marks during stem cell differentiation and development. *Nat Rev Genet*.

AZUARA, V., PERRY, P., SAUER, S., SPIVAKOV, M., JORGENSEN, H. F., JOHN, R. M., GOUTI, M., CASANOVA, M., WARNES, G., MERKENSCHLAGER, M. & FISHER, A. G. 2006. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol,* 8**,** 532-8.

BAGHERI-FAM, S., BARRIONUEVO, F., DOHRMANN, U., GUNTHER, T., SCHULE, R., KEMLER, R., MALLO, M., KANZLER, B. & SCHERER, G. 2006. Long-range upstream and downstream enhancers control distinct subsets of the complex spatiotemporal Sox9 expression pattern. *Dev Biol,* 291**,** 382-97.

BAGHERI-FAM, S., SIM, H., BERNARD, P., JAYAKODY, I., TAKETO, M. M., SCHERER, G. & HARLEY, V. R. 2008. Loss of Fgfr2 leads to partial XY sex reversal. *Dev Biol,* 314**,** 71-83.

BARRETO, G., SCHAFER, A., MARHOLD, J., STACH, D., SWAMINATHAN, S. K., HANDA, V., DODERLEIN, G., MALTRY, N., WU, W., LYKO, F. & NIEHRS, C. 2007. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature,* 445**,** 671-5.

BARRIONUEVO, F., BAGHERI-FAM, S., KLATTIG, J., KIST, R., TAKETO, M. M., ENGLERT, C. & SCHERER, G. 2006. Homozygous inactivation of Sox9 causes complete XY sex reversal in mice. *Biol Reprod,* 74**,** 195-201.

BECKER, A., ALLMANN, L., HOFSTATTER, M., CASA, V., WEBER, P., LEHMKUHL, A., HERCE, H. D. & CARDOSO, M. C. 2013. Direct homo- and hetero-interactions of MeCP2 and MBD2. *PLoS One,* 8**,** e53730.

BEHRENS, J., VON KRIES, J. P., KUHL, M., BRUHN, L., WEDLICH, D., GROSSCHEDL, R. & BIRCHMEIER, W. 1996. Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature,* 382**,** 638-42.

BENAYOUN, B. A., CABURET, S., DIPIETROMARIA, A., BAILLY-BECHET, M., BATISTA, F., FELLOUS, M., VAIMAN, D. & VEITIA, R. A. 2008. The identification and characterization of a FOXL2 response element provides insights into the pathogenesis of mutant alleles. *Hum Mol Genet,* 17**,** 3118-27.

BENKO, S., GORDON, C. T., MALLET, D., SREENIVASAN, R., THAUVIN-ROBINET, C., BRENDEHAUG, A., THOMAS, S., BRULAND, O., DAVID, M., NICOLINO, M., LABALME, A., SANLAVILLE, D., CALLIER, P., MALAN, V., HUET, F., MOLVEN, A., DIJOUD, F., MUNNICH, A., FAIVRE, L., AMIEL, J., HARLEY, V., HOUGE, G., MOREL, Y. & LYONNET, S. 2011. Disruption of a long distance regulatory region upstream of SOX9 in isolated disorders of sex development. *J Med Genet,* 48**,** 825-30.

BERNSTEIN, B. E., MIKKELSEN, T. S., XIE, X., KAMAL, M., HUEBERT, D. J., CUFF, J., FRY, B., MEISSNER, A., WERNIG, M., PLATH, K., JAENISCH, R., WAGSCHAL, A., FEIL, R., SCHREIBER, S. L. & LANDER, E. S. 2006a. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell,* 125**,** 315-26.

BERNSTEIN, E., DUNCAN, E. M., MASUI, O., GIL, J., HEARD, E. & ALLIS, C. D. 2006b. Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol Cell Biol,* 26**,** 2560-9.

BERTA, P., HAWKINS, J. R., SINCLAIR, A. H., TAYLOR, A., GRIFFITHS, B. L., GOODFELLOW, P. N. & FELLOUS, M. 1990. Genetic evidence equating SRY and the testis-determining factor. *Nature,* 348**,** 448-50.

BEVERDAM, A. & KOOPMAN, P. 2006. Expression profiling of purified mouse gonadal somatic cells during the critical time window of sex determination reveals novel candidate genes for human sexual dysgenesis syndromes. *Hum Mol Genet,* 15**,** 417-31.

BIANCHI, M. E. & BELTRAME, M. 1998. Flexing DNA: HMG-box proteins and their partners. *Am J Hum Genet,* 63**,** 1573-7.

BIASON-LAUBER, A., KONRAD, D., MEYER, M., DEBEAUFORT, C. & SCHOENLE, E. J. 2009. Ovaries and female phenotype in a girl with 46,XY karyotype and mutations in the CBX2 gene. *Am J Hum Genet,* 84**,** 658-63.

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., PORTER, F. D. & WESTPHAL, H. 2000. The LIM homeobox gene Lhx9 is essential for mouse gonad formation. *Nature,* 403**,** 909-13.

BIRK;, CASIANO;, WASSIF;, COGLIATI;, ZHAO;, ZHAO;, GRINBERG;, HUANG;, KREIDBERGK;, PARKER;, PORTER; & WESTPHAL 2000. The LIM homeobox gene Lhx9 is essential for 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. & OVERBEEK, P. A. 2000. A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. *Nat Genet,* 26**,** 490-4.

BISONI, L., BATLLE-MORERA, L., BIRD, A. P., SUZUKI, M. & MCQUEEN, H. A. 2005. Female-specific hyperacetylation of histone H4 in the chicken Z chromosome. *Chromosome Res,* 13**,** 205-14.

BOGANI, D., SIGGERS, P., BRIXEY, R., WARR, N., BEDDOW, S., EDWARDS, J., WILLIAMS, D., WILHELM, D., KOOPMAN, P., FLAVELL, R. A., CHI, H., OSTRER, H., WELLS, S., CHEESEMAN, M. & GREENFIELD, A. 2009. Loss of mitogen-activated protein kinase

kinase kinase 4 (MAP3K4) reveals a requirement for MAPK signalling in mouse sex determination. *PLoS Biol,* 7**,** e1000196.

BRACKEN, A. P., DIETRICH, N., PASINI, D., HANSEN, K. H. & HELIN, K. 2006. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev,* 20**,** 1123-36.

BRADFORD, S. T., WILHELM, D., BANDIERA, R., VIDAL, V., SCHEDL, A. & KOOPMAN, P. 2009. A cell-autonomous role for WT1 in regulating Sry in vivo. *Hum Mol Genet,* 18**,** 3429-38.

BUENROSTRO, J. D., GIRESI, P. G., ZABA, L. C., CHANG, H. Y. & GREENLEAF, W. J. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods,* 10**,** 1213-8.

BULLEJOS, M. & KOOPMAN, P. 2001. Spatially dynamic expression of Sry in mouse genital ridges. *Dev Dyn,* 221**,** 201-5.

BURGOYNE, P. S., THORNHILL, A. R., BOUDREAN, S. K., DARLING, S. M., BISHOP, C. E. & EVANS, E. P. 1995. The genetic basis of XX-XY differences present before gonadal sex differentiation in the mouse. *Philos Trans R Soc Lond B Biol Sci,* 350**,** 253-60 discussion 260-1.

CAPEL, B., SWAIN, A., NICOLIS, S., HACKER, A., WALTER, M., KOOPMAN, P., GOODFELLOW, P. & LOVELL-BADGE, R. 1993. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell,* 73**,** 1019-30.

CARON, C., PIVOT-PAJOT, C., VAN GRUNSVEN, L. A., COL, E., LESTRAT, C., ROUSSEAUX, S. & KHOCHBIN, S. 2003. Cdyl: a new transcriptional co-repressor. *EMBO Rep,* 4**,** 877-82.

CARRE, G. A., SIGGERS, P., XIPOLITA, M., BRINDLE, P., LUTZ, B., WELLS, S. & GREENFIELD, A. 2017. Loss of p300 and CBP disrupts histone acetylation at the mouse Sry promoter and causes XY gonadal sex reversal. *Hum Mol Genet*.

CARREL, L. & WILLARD, H. F. 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature,* 434**,** 400-4.

CASADESUS, J. & LOW, D. 2006. Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev,* 70**,** 830-56.

CHABOISSIER, M. C., KOBAYASHI, A., VIDAL, V. I., LUTZKENDORF, S., VAN DE KANT, H. J., WEGNER, M., DE ROOIJ, D. G., BEHRINGER, R. R. & SCHEDL, A. 2004. Functional analysis of Sox8 and Sox9 during sex determination in the mouse. *Development,* 131**,** 1891-901.

CHASSOT, A. A., GREGOIRE, E. P., MAGLIANO, M., LAVERY, R. & CHABOISSIER, M. C. 2008a. Genetics of ovarian differentiation: Rspo1, a major player. *Sex Dev,* 2**,** 219-27.

CHASSOT, A. A., RANC, F., GREGOIRE, E. P., ROEPERS-GAJADIEN, H. L., TAKETO, M. M., CAMERINO, G., DE ROOIJ, D. G., SCHEDL, A. & CHABOISSIER, M. C. 2008b. Activation of beta-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. *Hum Mol Genet,* 17**,** 1264-77.

CHEN, S., ZHANG, G., SHAO, C., HUANG, Q., LIU, G., ZHANG, P., SONG, W., AN, N., CHALOPIN, D., VOLFF, J. N., HONG, Y., LI, Q., SHA, Z., ZHOU, H., XIE, M., YU, Q., LIU, Y., XIANG, H., WANG, N., WU, K., YANG, C., ZHOU, Q., LIAO, X., YANG, L., HU, Q., ZHANG, J., MENG, L., JIN, L., TIAN, Y., LIAN, J., YANG, J., MIAO, G., LIU, S., LIANG, Z., YAN, F., LI, Y., SUN, B., ZHANG, H., ZHANG, J., ZHU, Y., DU, M., ZHAO, Y., SCHARTL, M., TANG, Q. & WANG, J. 2014. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. *Nat Genet,* 46**,** 253-60.

CHIQUOINE, A. D. 1954. The identification, origin, and migration of the primordial germ cells in the mouse embryo. *Anat Rec,* 118**,** 135-46.

CLINE, T. W. & MEYER, B. J. 1996. Vive la difference: males vs females in flies vs worms. *Annu Rev Genet,* 30**,** 637-702.

COLVIN, J. S., GREEN, R. P., SCHMAHL, J., CAPEL, B. & ORNITZ, D. M. 2001. Male-to-female sex reversal in mice lacking fibroblast growth factor 9. *Cell,* 104**,** 875-89.

COX, J. J., WILLATT, L., HOMFRAY, T. & WOODS, C. G. 2011. A SOX9 duplication and familial 46,XX developmental testicular disorder. *N Engl J Med,* 364**,** 91-3.

CREYGHTON, M. P., CHENG, A. W., WELSTEAD, G. G., KOOISTRA, T., CAREY, B. W., STEINE, E. J., HANNA, J., LODATO, M. A., FRAMPTON, G. M., SHARP, P. A., BOYER, L. A., YOUNG, R. A. & JAENISCH, R. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A,* 107**,** 21931-6.

CZERWINSKI, M., NATARAJAN, A., BARSKE, L., LOOGER, L. L. & CAPEL, B. 2016. A timecourse analysis of systemic and gonadal effects of temperature on sexual development of the red-eared slider turtle Trachemys scripta elegans. *Dev Biol,* 420**,** 166-177.

DEKKER, J., RIPPE, K., DEKKER, M. & KLECKNER, N. 2002. Capturing chromosome conformation. *Science,* 295**,** 1306-11.

DENNY, P., SWIFT, S., BRAND, N., DABHADE, N., BARTON, P. & ASHWORTH, A. 1992a. A conserved family of genes related to the testis determining gene, SRY. *Nucleic Acids Res,* 20**,** 2887.

DENNY, P., SWIFT, S., CONNOR, F. & ASHWORTH, A. 1992b. An SRY-related gene expressed during spermatogenesis in the mouse encodes a sequence-specific DNA-binding protein. *EMBO J,* 11**,** 3705-12.

DHANOA, J. K., MUKHOPADHYAY, C. S. & ARORA, J. S. 2016. Y-chromosomal genes affecting male fertility: A review. *Vet World,* 9**,** 783-91.

DICKEL, D. E., YPSILANTI, A. R., PLA, R., ZHU, Y., BAROZZI, I., MANNION, B. J., KHIN, Y. S., FUKUDA-YUZAWA, Y., PLAJZER-FRICK, I., PICKLE, C. S., LEE, E. A., HARRINGTON, A. N., PHAM, Q. T., GARVIN, T. H., KATO, M., OSTERWALDER, M., AKIYAMA, J. A., AFZAL, V., RUBENSTEIN, J. L. R., PENNACCHIO, L. A. & VISEL, A. 2018. Ultraconserved Enhancers Are Required for Normal Development. *Cell,* 172**,** 491-499 e15.

DIXON, J. R., SELVARAJ, S., YUE, F., KIM, A., LI, Y., SHEN, Y., HU, M., LIU, J. S. & REN, B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature,* 485**,** 376-80.

DOLCI, S., GRIMALDI, P., GEREMIA, R., PESCE, M. & ROSSI, P. 1997. Identification of a promoter region generating Sry circular transcripts both in germ cells from male adult mice and in male mouse embryonal gonads. *Biol Reprod,* 57**,** 1128-35.

FERRARI, S., HARLEY, V. R., PONTIGGIA, A., GOODFELLOW, P. N., LOVELL-BADGE, R. & BIANCHI, M. E. 1992. SRY, like HMG1, recognizes sharp angles in DNA. *EMBO J,* 11**,** 4497-506.

FRANCHINA, M. & KAY, P. H. 2000. Evidence that cytosine residues within 5'-CCTGG-3' pentanucleotides can be methylated in human DNA independently of the methylating system that modifies 5'-CG-3' dinucleotides. *DNA Cell Biol,* 19**,** 521-6.

FRANKE, M., IBRAHIM, D. M., ANDREY, G., SCHWARZER, W., HEINRICH, V., SCHOPFLIN, R., KRAFT, K., KEMPFER, R., JERKOVIC, I., CHAN, W. L., SPIELMANN, M., TIMMERMANN, B., WITTLER, L., KURTH, I., CAMBIASO, P., ZUFFARDI, O., HOUGE,

G., LAMBIE, L., BRANCATI, F., POMBO, A., VINGRON, M., SPITZ, F. & MUNDLOS, S. 2016. Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature,* 538**,** 265-269.

FUJIMOTO, Y., TANAKA, S. S., YAMAGUCHI, Y. L., KOBAYASHI, H., KUROKI, S., TACHIBANA, M., SHINOMURA, M., KANAI, Y., MOROHASHI, K., KAWAKAMI, K. & NISHINAKAMURA, R. 2013. Homeoproteins Six1 and Six4 regulate male sex determination and mouse gonadal development. *Dev Cell,* 26**,** 416-30.

FUKS, F., HURD, P. J., WOLF, D., NAN, X., BIRD, A. P. & KOUZARIDES, T. 2003. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem,* 278**,** 4035-40.

FURUMATSU, T., TSUDA, M., YOSHIDA, K., TANIGUCHI, N., ITO, T., HASHIMOTO, M., ITO, T. & ASAHARA, H. 2005. Sox9 and p300 cooperatively regulate chromatin-mediated transcription. *J Biol Chem,* 280**,** 35203-8.

GE, C., YE, J., WEBER, C., SUN, W., ZHANG, H., ZHOU, Y., CAI, C., QIAN, G. & CAPEL, B. 2018. The histone demethylase KDM6B regulates temperature-dependent sex determination in a turtle species. *Science,* 360**,** 645-648.

GE, C., YE, J., ZHANG, H., ZHANG, Y., SUN, W., SANG, Y., CAPEL, B. & QIAN, G. 2017. Dmrt1 induces the male pathway in a turtle species with temperature-dependent sex determination. *Development,* 144**,** 2222-2233.

GEORG, I., BAGHERI-FAM, S., KNOWER, K. C., WIEACKER, P., SCHERER, G. & HARLEY, V. R. 2010. Mutations of the SRY-responsive enhancer of SOX9 are uncommon in XY gonadal dysgenesis. *Sex Dev,* 4**,** 321-5.

GHIRLANDO, R. & FELSENFELD, G. 2016. CTCF: making the right connections. *Genes Dev,* 30**,** 881-91.

GIERL, M. S., GRUHN, W. H., VON SEGGERN, A., MALTRY, N. & NIEHRS, C. 2012. GADD45G functions in male sex determination by promoting p38 signaling and Sry expression. *Dev Cell,* 23**,** 1032-42.

GINSBURG, M., SNOW, M. H. & MCLAREN, A. 1990. Primordial germ cells in the mouse embryo during gastrulation. *Development,* 110**,** 521-8.

GODMANN, M., KATZ, J. P., GUILLOU, F., SIMONI, M., KAESTNER, K. H. & BEHR, R. 2008. Kruppel-like factor 4 is involved in functional differentiation of testicular Sertoli cells. *Dev Biol,* 315**,** 552-66.

GONEN, N., FUTTNER, C. R., WOOD, S., GARCIA-MORENO, S. A., SALAMONE, I. M., SAMSON, S. C., SEKIDO, R., POULAT, F., MAATOUK, D. M. & LOVELL-BADGE, R. 2018a. Sex reversal following deletion of a single distal enhancer of Sox9. *Science*.

GONEN, N., FUTTNER, C. R., WOOD, S., GARCIA-MORENO, S. A., SALAMONE, I. M., SAMSON, S. C., SEKIDO, R., POULAT, F., MAATOUK, D. M. & LOVELL-BADGE, R. 2018b. Sex reversal following deletion of a single distal enhancer of Sox9. *Science,* 360**,** 1469-1473.

GONEN, N., QUINN, A., O'NEILL, H. C., KOOPMAN, P. & LOVELL-BADGE, R. 2017. Normal Levels of Sox9 Expression in the Developing Mouse Testis Depend on the TES/TESCO Enhancer, but This Does Not Act Alone. *PLoS Genet,* 13**,** e1006520.

GRAU, D. J., CHAPMAN, B. A., GARLICK, J. D., BOROWSKY, M., FRANCIS, N. J. & KINGSTON, R. E. 2011. Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes Dev,* 25**,** 2210-21.

GREENFIELD, A., CARREL, L., PENNISI, D., PHILIPPE, C., QUADERI, N., SIGGERS, P., STEINER, K., TAM, P. P., MONACO, A. P., WILLARD, H. F. & KOOPMAN, P. 1998. The UTX gene escapes X inactivation in mice and humans. *Hum Mol Genet,* 7**,** 737-42.

GRUENBAUM, Y., NAVEH-MANY, T., CEDAR, H. & RAZIN, A. 1981. Sequence specificity of methylation in higher plant DNA. *Nature,* 292**,** 860-2.

GUBBAY, J., COLLIGNON, J., KOOPMAN, P., CAPEL, B., ECONOMOU, A., MUNSTERBERG, A., VIVIAN, N., GOODFELLOW, P. & 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**,** 245-50.

GUO, Y., CHENG, H., HUANG, X., GAO, S., YU, H. & ZHOU, R. 2005. Gene structure, multiple alternative splicing, and expression in gonads of zebrafish Dmrt1. *Biochem Biophys Res Commun,* 330**,** 950-7.

HACKER, A., CAPEL, B., GOODFELLOW, P. & LOVELL-BADGE, R. 1995. Expression of Sry, the mouse sex determining gene. *Development,* 121**,** 1603-14.

HAMMES, A., GUO, J. K., LUTSCH, G., LEHESTE, J. R., LANDROCK, D., ZIEGLER, U., GUBLER, M. C. & SCHEDL, A. 2001. Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell,* 106**,** 319-29.

HARLEY, V. R., JACKSON, D. I., HEXTALL, P. J., HAWKINS, J. R., BERKOVITZ, G. D., SOCKANATHAN, S., LOVELL-BADGE, R. & GOODFELLOW, P. N. 1992. DNA binding activity of recombinant SRY from normal males and XY females. *Science,* 255**,** 453-6.

HASHIMOTO, N., BROCK, H. W., NOMURA, M., KYBA, M., HODGSON, J., FUJITA, Y., TAKIHARA, Y., SHIMADA, K. & HIGASHINAKAGAWA, T. 1998. RAE28, BMI1, and M33 are members of heterogeneous multimeric mammalian Polycomb group complexes. *Biochem Biophys Res Commun,* 245**,** 356-65.

HENDRICH, B. & BIRD, A. 1998. Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol,* 18**,** 6538-47.

HERZ, H. M., MOHAN, M., GARRETT, A. S., MILLER, C., CASTO, D., ZHANG, Y., SEIDEL, C., HAUG, J. S., FLORENS, L., WASHBURN, M. P., YAMAGUCHI, M., SHIEKHATTAR, R. & SHILATIFARD, A. 2012. Polycomb repressive complex 2-dependent and -independent functions of Jarid2 in transcriptional regulation in Drosophila. *Mol Cell Biol,* 32**,** 1683-93.

HIRAMATSU, R., MATOBA, S., KANAI-AZUMA, M., TSUNEKAWA, N., KATOH-FUKUI, Y., KUROHMARU, M., MOROHASHI, K., WILHELM, D., KOOPMAN, P. & KANAI, Y. 2009. A critical time window of Sry action in gonadal sex determination in mice. *Development,* 136**,** 129-38.

HONG, S., CHO, Y. W., YU, L. R., YU, H., VEENSTRA, T. D. & GE, K. 2007. Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc Natl Acad Sci U S A,* 104**,** 18439-44.

HU, Y. C., OKUMURA, L. M. & PAGE, D. C. 2013. Gata4 is required for formation of the genital ridge in mice. *PLoS Genet,* 9**,** e1003629.

IKAWA, T., MASUDA, K., ENDO, T. A., ENDO, M., ISONO, K., KOSEKI, Y., NAKAGAWA, R., KOMETANI, K., TAKANO, J., AGATA, Y., KATSURA, Y., KUROSAKI, T., VIDAL, M., KOSEKI, H. & KAWAMOTO, H. 2016. Conversion of T cells to B cells by inactivation of polycomb-mediated epigenetic suppression of the B-lineage program. *Genes Dev,* 30**,** 2475-2485.

IMAMURA, T., KERJEAN, A., HEAMS, T., KUPIEC, J. J., THENEVIN, C. & PALDI, A. 2005. Dynamic CpG and non-CpG methylation of the Peg1/Mest gene in the mouse oocyte and preimplantation embryo. *J Biol Chem,* 280**,** 20171-5.

INOUE, K., OGONUKI, N., MEKADA, K., YOSHIKI, A., SADO, T. & OGURA, A. 2009. Sex-reversed somatic cell cloning in the mouse. *J Reprod Dev,* 55**,** 566-9.

ITOH, Y., KAMPF, K. & ARNOLD, A. P. 2011. Possible differences in the two Z chromosomes in male chickens and evolution of MHM sequences in Galliformes. *Chromosoma,* 120**,** 587-98.

IWASE, S., LAN, F., BAYLISS, P., DE LA TORRE-UBIETA, L., HUARTE, M., QI, H. H., WHETSTINE, J. R., BONNI, A., ROBERTS, T. M. & SHI, Y. 2007. The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell,* 128**,** 1077-88.

JACOB, E., HOD-DVORAI, R., BEN-MORDECHAI, O. L., BOYKO, Y. & AVNI, O. 2011. Dual function of polycomb group proteins in differentiated murine T helper (CD4+) cells. *J Mol Signal,* 6**,** 5.

JAGER, R. J., HARLEY, V. R., PFEIFFER, R. A., GOODFELLOW, P. N. & SCHERER, G. 1992. A familial mutation in the testis-determining gene SRY shared by both sexes. *Hum Genet,* 90**,** 350-5.

JAKOB, S. & LOVELL-BADGE, R. 2011. Sex determination and the control of Sox9 expression in mammals. *FEBS J,* 278**,** 1002-9.

JAMESON, S. A., LIN, Y. T. & CAPEL, B. 2012a. Testis development requires the repression of Wnt4 by Fgf signaling. *Dev Biol,* 370**,** 24-32.

JAMESON, S. A., NATARAJAN, A., COOL, J., DEFALCO, T., MAATOUK, D. M., MORK, L., MUNGER, S. C. & CAPEL, B. 2012b. Temporal transcriptional profiling of somatic and germ cells reveals biased lineage priming of sexual fate in the fetal mouse gonad. *PLoS Genet,* 8**,** e1002575.

JEONG, Y. H., LU, H., PARK, C. H., LI, M., LUO, H., KIM, J. J., LIU, S., KO, K. H., HUANG, S., HWANG, I. S., KANG, M. N., GONG, D., PARK, K. B., CHOI, E. J., PARK, J. H., JEONG, Y. W., MOON, C., HYUN, S. H., KIM, N. H., JEUNG, E. B., YANG, H., HWANG, W. S. & GAO, F. 2016. Stochastic anomaly of methylome but persistent SRY hypermethylation in disorder of sex development in canine somatic cell nuclear transfer. *Sci Rep,* 6**,** 31088.

JESKE, Y. W., BOWLES, J., GREENFIELD, A. & KOOPMAN, P. 1995. Expression of a linear Sry transcript in the mouse genital ridge. *Nat Genet,* 10**,** 480-2.

JO, A., DENDULURI, S., ZHANG, B., WANG, Z., YIN, L., YAN, Z., KANG, R., SHI, L. L., MOK, J., LEE, M. J. & HAYDON, R. C. 2014. The versatile functions of Sox9 in development, stem cells, and human diseases. *Genes Dis,* 1**,** 149-161.

JOST, A. 1952. [Investigation of hormonal control of genesis of the sex organs in rabbit and notes on certain malformations of the genital apparatus in man]. *Ginecol Obstet Mex,* 7**,** 477-92.

KARL, J. & CAPEL, B. 1998. Sertoli cells of the mouse testis originate from the coelomic epithelium. *Dev Biol,* 203**,** 323-33.

KASHIMADA, K., PELOSI, E., CHEN, H., SCHLESSINGER, D., WILHELM, D. & KOOPMAN, P. 2011. FOXL2 and BMP2 act cooperatively to regulate follistatin gene expression during ovarian development. *Endocrinology,* 152**,** 272-80.

KATOH-FUKUI, Y., MIYABAYASHI, K., KOMATSU, T., OWAKI, A., BABA, T., SHIMA, Y., KIDOKORO, T., KANAI, Y., SCHEDL, A., WILHELM, D., KOOPMAN, P., OKUNO, Y. & MOROHASHI, K. 2012. Cbx2, a polycomb group gene, is required for Sry gene expression in mice. *Endocrinology,* 153**,** 913-24.

KATOH-FUKUI, Y., TSUCHIYA, R., SHIROISHI, T., NAKAHARA, Y., HASHIMOTO, N., NOGUCHI, K. & HIGASHINAKAGAWA, T. 1998. Male-to-female sex reversal in M33 mutant mice. *Nature,* 393**,** 688-92.

KERNOHAN, K. D., VERNIMMEN, D., GLOOR, G. B. & BERUBE, N. G. 2014. Analysis of neonatal brain lacking ATRX or MeCP2 reveals changes in nucleosome density, CTCF binding and chromatin looping. *Nucleic Acids Res,* 42**,** 8356-68.

KETTLEWELL, J. R., RAYMOND, C. S. & ZARKOWER, D. 2000. Temperature-dependent expression of turtle Dmrt1 prior to sexual differentiation. *Genesis,* 26**,** 174-8.

KIM, G. J., SOCK, E., BUCHBERGER, A., JUST, W., DENZER, F., HOEPFFNER, W., GERMAN, J., COLE, T., MANN, J., SEGUIN, J. H., ZIPF, W., COSTIGAN, C., SCHMIADY, H., ROSTASY, M., KRAMER, M., KALTENBACH, S., ROSLER, B., GEORG, I., TROPPMANN, E., TEICHMANN, A. C., SALFELDER, A., WIDHOLZ, S. A., WIEACKER, P., HIORT, O., CAMERINO, G., RADI, O., WEGNER, M., ARNOLD, H. H. & SCHERER, G. 2015. Copy number variation of two separate regulatory regions upstream of SOX9 causes isolated 46,XY or 46,XX disorder of sex development. *J Med Genet,* 52**,** 240-7.

KIM, J., DANIEL, J., ESPEJO, A., LAKE, A., KRISHNA, M., XIA, L., ZHANG, Y. & BEDFORD, M. T. 2006a. Tudor, MBT and chromo domains gauge the degree of lysine methylation. *EMBO Rep,* 7**,** 397-403.

KIM, T. H., ABDULLAEV, Z. K., SMITH, A. D., CHING, K. A., LOUKINOV, D. I., GREEN, R. D., ZHANG, M. Q., LOBANENKOV, V. V. & REN, B. 2007a. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell,* 128**,** 1231-45.

KIM, Y., BINGHAM, N., SEKIDO, R., PARKER, K. L., LOVELL-BADGE, R. & CAPEL, B. 2007b. Fibroblast growth factor receptor 2 regulates proliferation and Sertoli differentiation during male sex determination. *Proc Natl Acad Sci U S A,* 104**,** 16558-63.

KIM, Y., KOBAYASHI, A., SEKIDO, R., DINAPOLI, L., BRENNAN, J., CHABOISSIER, M. C., POULAT, F., BEHRINGER, R. R., LOVELL-BADGE, R. & CAPEL, B. 2006b. Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination. *PLoS Biol,* 4**,** e187.

KIRMIZIS, A., BARTLEY, S. M., KUZMICHEV, A., MARGUERON, R., REINBERG, D., GREEN, R. & FARNHAM, P. J. 2004. Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. *Genes Dev,* 18**,** 1592-605.

KOOPMAN, P., GUBBAY, J., VIVIAN, N., GOODFELLOW, P. & LOVELL-BADGE, R. 1991. Male development of chromosomally female mice transgenic for Sry. *Nature,* 351**,** 117-21.

KUROKI, S., MATOBA, S., AKIYOSHI, M., MATSUMURA, Y., MIYACHI, H., MISE, N., ABE, K., OGURA, A., WILHELM, D., KOOPMAN, P., NOZAKI, M., KANAI, Y., SHINKAI, Y. & TACHIBANA, M. 2013. Epigenetic regulation of mouse sex determination by the histone demethylase Jmjd1a. *Science,* 341**,** 1106-9.

KUROKI, S., OKASHITA, N., BABA, S., MAEDA, R., MIYAWAKI, S., YANO, M., YAMAGUCHI, M., KITANO, S., MIYACHI, H., ITOH, A., YOSHIDA, M. & TACHIBANA, M. 2017. Rescuing the aberrant sex development of H3K9 demethylase Jmjd1a-deficient mice by modulating H3K9 methylation balance. *PLoS Genet,* 13**,** e1007034.

KUSAKA, M., KATOH-FUKUI, Y., OGAWA, H., MIYABAYASHI, K., BABA, T., SHIMA, Y., SUGIYAMA, N., SUGIMOTO, Y., OKUNO, Y., KODAMA, R., IIZUKA-KOGO, A., SENDA, T., SASAOKA, T., KITAMURA, K., AIZAWA, S. & MOROHASHI, K. 2010. Abnormal epithelial cell polarity and ectopic epidermal growth factor receptor (EGFR) expression induced in Emx2 KO embryonic gonads. *Endocrinology,* 151**,** 5893-904.

KWOK, C., WELLER, P. A., GUIOLI, S., FOSTER, J. W., MANSOUR, S., ZUFFARDI, O., PUNNETT, H. H., DOMINGUEZ-STEGLICH, M. A., BROOK, J. D., YOUNG, I. D. & ET

AL. 1995. Mutations in SOX9, the gene responsible for Campomelic dysplasia and autosomal sex reversal. *Am J Hum Genet,* 57**,** 1028-36.

LAGHA, M., BOTHMA, J. P. & LEVINE, M. 2012. Mechanisms of transcriptional precision in animal development. *Trends Genet,* 28**,** 409-16.

LAHN, B. T., TANG, Z. L., ZHOU, J., BARNDT, R. J., PARVINEN, M., ALLIS, C. D. & PAGE, D. C. 2002. Previously uncharacterized histone acetyltransferases implicated in mammalian spermatogenesis. *Proc Natl Acad Sci U S A,* 99**,** 8707-12.

LAMBETH, L. S., RAYMOND, C. S., ROESZLER, K. N., KUROIWA, A., NAKATA, T., ZARKOWER, D. & SMITH, C. A. 2014. Over-expression of DMRT1 induces the male pathway in embryonic chicken gonads. *Dev Biol,* 389**,** 160-72.

LAN, F., BAYLISS, P. E., RINN, J. L., WHETSTINE, J. R., WANG, J. K., CHEN, S., IWASE, S., ALPATOV, R., ISSAEVA, I., CANAANI, E., ROBERTS, T. M., CHANG, H. Y. & SHI, Y. 2007. A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature,* 449**,** 689-94.

LAU, M. S., SCHWARTZ, M. G., KUNDU, S., SAVOL, A. J., WANG, P. I., MARR, S. K., GRAU, D. J., SCHORDERET, P., SADREYEV, R. I., TABIN, C. J. & KINGSTON, R. E. 2017. Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. *Science,* 355**,** 1081-1084.

LEE, C. K., SHIBATA, Y., RAO, B., STRAHL, B. D. & LIEB, J. D. 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet,* 36**,** 900-5.

LEE, M. G., NORMAN, J., SHILATIFARD, A. & SHIEKHATTAR, R. 2007. Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. *Cell,* 128**,** 877-87.

LEIPOLDT, M., ERDEL, M., BIEN-WILLNER, G. A., SMYK, M., THEURL, M., YATSENKO, S. A., LUPSKI, J. R., LANE, A. H., SHANSKE, A. L., STANKIEWICZ, P. & SCHERER, G. 2007. Two novel translocation breakpoints upstream of SOX9 define borders of the proximal and distal breakpoint cluster region in campomelic dysplasia. *Clin Genet,* 71**,** 67-75.

LESCH, B. J., SILBER, S. J., MCCARREY, J. R. & PAGE, D. C. 2016. Parallel evolution of male germline epigenetic poising and somatic development in animals. *Nat Genet,* 48**,** 888-94.

LI, E., BESTOR, T. H. & JAENISCH, R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell,* 69**,** 915-26.

LI, Y., ZHENG, M. & LAU, Y. F. 2014. The sex-determining factors SRY and SOX9 regulate similar target genes and promote testis cord formation during testicular differentiation. *Cell Rep,* 8**,** 723-33.

LIN, J. M., COLLINS, P. J., TRINKLEIN, N. D., FU, Y., XI, H., MYERS, R. M. & WENG, Z. 2007. Transcription factor binding and modified histones in human bidirectional promoters. *Genome Res,* 17**,** 818-27.

LIU, C. F., BINGHAM, N., PARKER, K. & YAO, H. H. 2009. Sex-specific roles of beta-catenin in mouse gonadal development. *Hum Mol Genet,* 18**,** 405-17.

LOPES, R., KORKMAZ, G. & AGAMI, R. 2016. Applying CRISPR-Cas9 tools to identify and characterize transcriptional enhancers. *Nat Rev Mol Cell Biol,* 17**,** 597-604.

LU, T., HEYNE, S., DROR, E., CASAS, E., LEONHARDT, L., BOENKE, T., YANG, C., SAGAR, ARRIGONI, L., DALGAARD, K., TEPERINO, R., ENDERS, L., SELVARAJ, M., RUF, M., RAJA, S. J., XIE, H., BOENISCH U., ORKIN, S. H., LYNN, F., HOFFMAN, B. G., GRUN, D., VAVOURI, T., LEMPRADL A., POSPISILIK, A. 2017. The Polycomb-

dependent epigenome controls B-cell dysfunction, dedifferentiation and diabetes. *bioRxiv* DOI: 205641.

LUO, X., IKEDA, Y. & PARKER, K. L. 1994. A Cell-Specific Nuclear Receptor Is Essential for Adrenal and Gonadal Development and Sexual Differentiation. *Cell,* 77**,** 481-490.

LUX, A., KROPF, S., KLEINEMEIER, E., JURGENSEN, M., THYEN, U. & GROUP, D. S. D. N. W. 2009. Clinical evaluation study of the German network of disorders of sex development (DSD)/intersexuality: study design, description of the study population, and data quality. *BMC Public Health,* 9**,** 110.

LYBAEK, H., DE BRUIJN, D., DEN ENGELSMAN-VAN DIJK, A. H., VANICHKINA, D., NEPAL, C., BRENDEHAUG, A. & HOUGE, G. 2014. RevSex duplication-induced and sex-related differences in the SOX9 regulatory region chromatin landscape in human fibroblasts. *Epigenetics,* 9**,** 416-27.

MAATOUK, D. M., DINAPOLI, L., ALVERS, A., PARKER, K. L., TAKETO, M. M. & CAPEL, B. 2008. Stabilization of beta-catenin in XY gonads causes male-to-female sex-reversal. *Hum Mol Genet,* 17**,** 2949-55.

MAATOUK, D. M., NATARAJAN, A., SHIBATA, Y., SONG, L., CRAWFORD, G. E., OHLER, U. & CAPEL, B. 2017. Genome-wide identification of regulatory elements in Sertoli cells. *Development,* 144**,** 720-730.

MALONE, C. S., MINER, M. D., DOERR, J. R., JACKSON, J. P., JACOBSEN, S. E., WALL, R. & TEITELL, M. 2001. CmC(A/T)GG DNA methylation in mature B cell lymphoma gene silencing. *Proc Natl Acad Sci U S A,* 98**,** 10404-9.

MANUYLOV, N. L., FUJIWARA, Y., ADAMEYKO, II, POULAT, F. & TEVOSIAN, S. G. 2007. The regulation of Sox9 gene expression by the GATA4/FOG2 transcriptional complex in dominant XX sex reversal mouse models. *Dev Biol,* 307**,** 356-67.

MARCHAND, O., GOVOROUN, M., D'COTTA, H., MCMEEL, O., LAREYRE, J. J., BERNOT, A., LAUDET, V. & GUIGUEN, Y. 2000. DMRT1 expression during gonadal differentiation and spermatogenesis in the rainbow trout, Oncorhynchus mykiss. *Biochim Biophys Acta,* 1493**,** 180-7.

MARGUERON, R. & REINBERG, D. 2011. The Polycomb complex PRC2 and its mark in life. *Nature,* 469**,** 343-9.

MASUYAMA, H., YAMADA, M., KAMEI, Y., FUJIWARA-ISHIKAWA, T., TODO, T., NAGAHAMA, Y. & MATSUDA, M. 2012. Dmrt1 mutation causes a male-to-female sex reversal after the sex determination by Dmy in the medaka. *Chromosome Res,* 20**,** 163-76.

MATSON, C. K., MURPHY, M. W., SARVER, A. L., GRISWOLD, M. D., BARDWELL, V. J. & ZARKOWER, D. 2011. DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature,* 476**,** 101-4.

MATSON, C. K. & ZARKOWER, D. 2012. Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. *Nat Rev Genet,* 13**,** 163-74.

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

MCQUEEN, H. A. & CLINTON, M. 2009. Avian sex chromosomes: dosage compensation matters. *Chromosome Res,* 17**,** 687-97.

MINOUX, M., HOLWERDA, S., VITOBELLO, A., KITAZAWA, T., KOHLER, H., STADLER, M. B. & RIJLI, F. M. 2017. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. *Science,* 355.

MITSUHASHI, T., WARITA, K., SUGAWARA, T., TABUCHI, Y., TAKASAKI, I., KONDO, T., HAYASHI, F., WANG, Z. Y., MATSUMOTO, Y., MIKI, T., TAKEUCHI, Y., EBINA, Y.,

YAMADA, H., SAKURAGI, N., YOKOYAMA, T., NANMORI, T., KITAGAWA, H., KANT, J. A. & HOSHI, N. 2010. Epigenetic abnormality of SRY gene in the adult XY female with pericentric inversion of the Y chromosome. *Congenit Anom (Kyoto),* 50**,** 85-94.

MONIOT, B., DECLOSMENIL, F., BARRIONUEVO, F., SCHERER, G., ARITAKE, K., MALKI, S., MARZI, L., COHEN-SOLAL, A., GEORG, I., KLATTIG, J., ENGLERT, C., KIM, Y., CAPEL, B., EGUCHI, N., URADE, Y., BOIZET-BONHOURE, B. & POULAT, F. 2009. The PGD2 pathway, independently of FGF9, amplifies SOX9 activity in Sertoli cells during male sexual differentiation. *Development,* 136**,** 1813-21.

MOUSAVI, K., ZARE, H., WANG, A. H. & SARTORELLI, V. 2012. Polycomb protein Ezh1 promotes RNA polymerase II elongation. *Mol Cell,* 45**,** 255-62.

MUNGER, S. C., NATARAJAN, A., LOOGER, L. L., OHLER, U. & CAPEL, B. 2013. Fine time course expression analysis identifies cascades of activation and repression and maps a putative regulator of mammalian sex determination. *PLoS Genet,* 9**,** e1003630.

MURPHY, E. C., ZHURKIN, V. B., LOUIS, J. M., CORNILESCU, G. & CLORE, G. M. 2001. Structural basis for SRY-dependent 46-X,Y sex reversal: modulation of DNA bending by a naturally occurring point mutation. *J Mol Biol,* 312**,** 481-99.

NAN, X., HOU, J., MACLEAN, A., NASIR, J., LAFUENTE, M. J., SHU, X., KRIAUCIONIS, S. & BIRD, A. 2007. Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proc Natl Acad Sci U S A,* 104**,** 2709-14.

NAN, X., NG, H. H., JOHNSON, C. A., LAHERTY, C. D., TURNER, B. M., EISENMAN, R. N. & BIRD, A. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature,* 393**,** 386-9.

NEF, S., SCHAAD, O., STALLINGS, N. R., CEDERROTH, C. R., PITETTI, J. L., SCHAER, G., MALKI, S., DUBOIS-DAUPHIN, M., BOIZET-BONHOURE, B., DESCOMBES, P., PARKER, K. L. & VASSALLI, J. D. 2005. Gene expression during sex determination reveals a robust female genetic program at the onset of ovarian development. *Dev Biol,* 287**,** 361-77.

NIEHRS, C. & SCHAFER, A. 2012. Active DNA demethylation by Gadd45 and DNA repair. *Trends Cell Biol,* 22**,** 220-7.

NISHINO, K., HATTORI, N., SATO, S., ARAI, Y., TANAKA, S., NAGY, A. & SHIOTA, K. 2011. Non-CpG methylation occurs in the regulatory region of the Sry gene. *J Reprod Dev,* 57**,** 586-93.

NISHINO, K., HATTORI, N., TANAKA, S. & SHIOTA, K. 2004. DNA methylation-mediated control of Sry gene expression in mouse gonadal development. *J Biol Chem,* 279**,** 22306-13.

NORA, E. P., LAJOIE, B. R., SCHULZ, E. G., GIORGETTI, L., OKAMOTO, I., SERVANT, N., PIOLOT, T., VAN BERKUM, N. L., MEISIG, J., SEDAT, J., GRIBNAU, J., BARILLOT, E., BLUTHGEN, N., DEKKER, J. & HEARD, E. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature,* 485**,** 381-5.

OH, H. J., LI, Y. & LAU, Y. F. 2005. Sry associates with the heterochromatin protein 1 complex by interacting with a KRAB domain protein. *Biol Reprod,* 72**,** 407-15.

OHNESORG, T., CROFT, B., TAN, J. & SINCLAIR, A. H. 2016. Using ROADMAP Data to Identify Enhancers Associated with Disorders of Sex Development. *Sex Dev,* 10**,** 59-65.

OKANO, M., BELL, D. W., HABER, D. A. & LI, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell,* 99**,** 247-57.

OSTERWALDER, M., BAROZZI, I., TISSIERES, V., FUKUDA-YUZAWA, Y., MANNION, B. J., AFZAL, S. Y., LEE, E. A., ZHU, Y., PLAJZER-FRICK, I., PICKLE, C. S., KATO, M., GARVIN, T. H., PHAM, Q. T., HARRINGTON, A. N., AKIYAMA, J. A., AFZAL, V., LOPEZ-RIOS, J., DICKEL, D. E., VISEL, A. & PENNACCHIO, L. A. 2018. Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature,* 554**,** 239-243.

OTTOLENGHI, C., PELOSI, E., TRAN, J., COLOMBINO, M., DOUGLASS, E., NEDOREZOV, T., CAO, A., FORABOSCO, A. & SCHLESSINGER, D. 2007. Loss of Wnt4 and Foxl2 leads to female-to-male sex reversal extending to germ cells. *Hum Mol Genet,* 16**,** 2795-804.

PAGE, Y.-C. H. L. M. O. D. C. 2013. *Gata4* Is Required for Formation of the Genital Ridge in Mice. *PLoS Genet,* 9.

PAN, G., TIAN, S., NIE, J., YANG, C., RUOTTI, V., WEI, H., JONSDOTTIR, G. A., STEWART, R. & THOMSON, J. A. 2007. Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell,* 1**,** 299-312.

PARMA, P., RADI, O., VIDAL, V., CHABOISSIER, M. C., DELLAMBRA, E., VALENTINI, S., GUERRA, L., SCHEDL, A. & CAMERINO, G. 2006. R-spondin1 is essential in sex determination, skin differentiation and malignancy. *Nat Genet,* 38**,** 1304-9.

PASINI, D., BRACKEN, A. P., HANSEN, J. B., CAPILLO, M. & HELIN, K. 2007. The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol Cell Biol,* 27**,** 3769-79.

PENG, H., IVANOV, A. V., OH, H. J., LAU, Y. F. & RAUSCHER, F. J., 3RD 2009. Epigenetic gene silencing by the SRY protein is mediated by a KRAB-O protein that recruits the KAP1 co-repressor machinery. *J Biol Chem,* 284**,** 35670-80.

PINARBASI, E., ELLIOTT, J. & HORNBY, D. P. 1996. Activation of a yeast pseudo DNA methyltransferase by deletion of a single amino acid. *J Mol Biol,* 257**,** 804-13.

PINTER, S. F., SADREYEV, R. I., YILDIRIM, E., JEON, Y., OHSUMI, T. K., BOROWSKY, M. & LEE, J. T. 2012. Spreading of X chromosome inactivation via a hierarchy of defined Polycomb stations. *Genome Res,* 22**,** 1864-76.

POLANCO, J. C., WILHELM, D., DAVIDSON, T. L., KNIGHT, D. & KOOPMAN, P. 2010. Sox10 gain-of-function causes XX sex reversal in mice: implications for human 22q-linked disorders of sex development. *Hum Mol Genet,* 19**,** 506-16.

POLANCO, J. C., WILHELM, D., MIZUSAKI, H., JACKSON, A., BROWNE, C., DAVIDSON, T., HARLEY, V., SINCLAIR, A. & KOOPMAN, P. 2009. Functional analysis of the SRY-KRAB interaction in mouse sex determination. *Biol Cell,* 101**,** 55-67.

PONTIGGIA, A., RIMINI, R., HARLEY, V. R., GOODFELLOW, P. N., LOVELL-BADGE, R. & BIANCHI, M. E. 1994. Sex-reversing mutations affect the architecture of SRY-DNA complexes. *EMBO J,* 13**,** 6115-24.

QIN, Y., KONG, L. K., POIRIER, C., TRUONG, C., OVERBEEK, P. A. & BISHOP, C. E. 2004. Long-range activation of Sox9 in Odd Sex (Ods) mice. *Hum Mol Genet,* 13**,** 1213-8.

RADA-IGLESIAS, A., BAJPAI, R., SWIGUT, T., BRUGMANN, S. A., FLYNN, R. A. & WYSOCKA, J. 2011. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature,* 470**,** 279-83.

RAZIN, A. & CEDAR, H. 1984. DNA methylation in eukaryotic cells. *Int Rev Cytol,* 92**,** 159-85.

READ, C. M., CARY, P. D., CRANE-ROBINSON, C., DRISCOLL, P. C. & NORMAN, D. G. 1993. Solution structure of a DNA-binding domain from HMG1. *Nucleic Acids Res,* 21**,** 3427-36.

REFAI, O., FRIEDMAN, A., TERRY, L., JEWETT, T., PEARLMAN, A., PERLE, M. A. & OSTRER, H. 2010. De novo 12;17 translocation upstream of SOX9 resulting in 46,XX testicular disorder of sex development. *Am J Med Genet A,* 152A**,** 422-6.

SANTOS-ROSA, H., SCHNEIDER, R., BANNISTER, A. J., SHERRIFF, J., BERNSTEIN, B. E., EMRE, N. C., SCHREIBER, S. L., MELLOR, J. & KOUZARIDES, T. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature,* 419**,** 407-11.

SCHMAHL, J. & CAPEL, B. 2003. Cell proliferation is necessary for the determination of male fate in the gonad. *Dev Biol,* 258**,** 264-76.

SEKIDO, R., BAR, I., NARVAEZ, V., PENNY, G. & LOVELL-BADGE, R. 2004. SOX9 is up-regulated by the transient expression of SRY specifically in Sertoli cell precursors. *Dev Biol,* 274**,** 271-9.

SEKIDO, R. & LOVELL-BADGE, R. 2008. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature,* 453**,** 930-4.

SHAO, C., LI, Q., CHEN, S., ZHANG, P., LIAN, J., HU, Q., SUN, B., JIN, L., LIU, S., WANG, Z., ZHAO, H., JIN, Z., LIANG, Z., LI, Y., ZHENG, Q., ZHANG, Y., WANG, J. & ZHANG, G. 2014. Epigenetic modification and inheritance in sexual reversal of fish. *Genome Res,* 24**,** 604-15.

SHARROCKS, A. D. 2001. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol,* 2**,** 827-37.

SHLYUEVA, D., STAMPFEL, G. & STARK, A. 2014. Transcriptional enhancers: from properties to genome-wide predictions. *Nat Rev Genet,* 15**,** 272-86.

SHPARGEL, K. B., SENGOKU, T., YOKOYAMA, S. & MAGNUSON, T. 2012. UTX and UTY demonstrate histone demethylase-independent function in mouse embryonic development. *PLoS Genet,* 8**,** e1002964.

SIMON, J. A. & KINGSTON, R. E. 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol,* 10**,** 697-708.

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. & 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**,** 240-4.

SINGH, N. P., MADABHUSHI, S. R., SRIVASTAVA, S., SENTHILKUMAR, R., NEERAJA, C., KHOSLA, S. & MISHRA, R. K. 2011. Epigenetic profile of the euchromatic region of human Y chromosome. *Nucleic Acids Res,* 39**,** 3594-606.

SINHA, P., SINGH, K. & SACHAN, M. 2017. Heterogeneous pattern of DNA methylation in developmentally important genes correlates with its chromatin conformation. *BMC Mol Biol,* 18**,** 1.

SLEUTELS, F., SOOCHIT, W., BARTKUHN, M., HEATH, H., DIENSTBACH, S., BERGMAIER, P., FRANKE, V., ROSA-GARRIDO, M., VAN DE NOBELEN, S., CAESAR, L., VAN DER REIJDEN, M., BRYNE, J. C., VAN IJCKEN, W., GROOTEGOED, J. A., DELGADO, M. D., LENHARD, B., RENKAWITZ, R., GROSVELD, F. & GALJART, N. 2012. The male germ cell gene regulator CTCFL is functionally different from CTCF and binds CTCF-like consensus sites in a nucleosome composition-dependent manner. *Epigenetics Chromatin,* 5**,** 8.

SMITH, C. A., ROESZLER, K. N., OHNESORG, T., CUMMINS, D. M., FARLIE, P. G., DORAN, T. J. & SINCLAIR, A. H. 2009. The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature,* 461**,** 267-71.

SMYK, M., SZAFRANSKI, P., STARTEK, M., GAMBIN, A. & STANKIEWICZ, P. 2013. Chromosome conformation capture-on-chip analysis of long-range cis-interactions of the SOX9 promoter. *Chromosome Res,* 21**,** 781-8.

SONG, L., ZHANG, Z., GRASFEDER, L. L., BOYLE, A. P., GIRESI, P. G., LEE, B. K., SHEFFIELD, N. C., GRAF, S., HUSS, M., KEEFE, D., LIU, Z., LONDON, D., MCDANIELL, R. M., SHIBATA, Y., SHOWERS, K. A., SIMON, J. M., VALES, T., WANG, T., WINTER, D., ZHANG, Z., CLARKE, N. D., BIRNEY, E., IYER, V. R., CRAWFORD, G. E., LIEB, J. D. & FUREY, T. S. 2011. Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res,* 21**,** 1757-67.

STOEVA, R., GROZDANOVA, L., SCHERER, G., KRASTEVA, M., BAUSCH, E., KRASTEV, T., LINEV, A. & STEFANOVA, M. 2011. A novel SOX9 nonsense mutation, q401x, in a case of campomelic dysplasia with XY sex reversal. *Genet Couns,* 22**,** 49-53.

SYMON, A. & HARLEY, V. 2017. SOX9: A genomic view of tissue specific expression and action. *Int J Biochem Cell Biol,* 87**,** 18-22.

TAM, P. P. & SNOW, M. H. 1981. Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *J Embryol Exp Morphol,* 64**,** 133-47.

TERANISHI, M., SHIMADA, Y., HORI, T., NAKABAYASHI, O., KIKUCHI, T., MACLEOD, T., PYM, R., SHELDON, B., SOLOVEI, I., MACGREGOR, H. & MIZUNO, S. 2001. Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. *Chromosome Res,* 9**,** 147-65.

TEVOSIAN, S. G. & MANUYLOV, N. L. 2008. To beta or not to beta: canonical beta-catenin signaling pathway and ovarian development. *Dev Dyn,* 237**,** 3672-80.

TOMASELLI, S., MEGIORNI, F., LIN, L., MAZZILLI, M. C., GERRELLI, D., MAJORE, S., GRAMMATICO, P. & ACHERMANN, J. C. 2011. Human RSPO1/R-spondin1 is expressed during early ovary development and augments beta-catenin signaling. *PLoS One,* 6**,** e16366.

TSUDA, M., TAKAHASHI, S., TAKAHASHI, Y. & ASAHARA, H. 2003. Transcriptional co-activators CREB-binding protein and p300 regulate chondrocyte-specific gene expression via association with Sox9. *J Biol Chem,* 278**,** 27224-9.

UHLENHAUT, N. H., JAKOB, S., ANLAG, K., EISENBERGER, T., SEKIDO, R., KRESS, J., TREIER, A. C., KLUGMANN, C., KLASEN, C., HOLTER, N. I., RIETHMACHER, D., SCHUTZ, G., COONEY, A. J., LOVELL-BADGE, R. & TREIER, M. 2009. Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. *Cell,* 139**,** 1130-42.

VAINIO, S., HEIKKILA, M., KISPERT, A., CHIN, N. & MCMAHON, A. P. 1999. Female development in mammals is regulated by Wnt-4 signalling. *Nature,* 397**,** 405-9.

VAN GALEN, P., VINY, A. D., RAM, O., RYAN, R. J., COTTON, M. J., DONOHUE, L., SIEVERS, C., DRIER, Y., LIAU, B. B., GILLESPIE, S. M., CARROLL, K. M., CROSS, M. B., LEVINE, R. L. & BERNSTEIN, B. E. 2016. A Multiplexed System for Quantitative Comparisons of Chromatin Landscapes. *Mol Cell,* 61**,** 170-80.

VIDAL, V. P., CHABOISSIER, M. C., DE ROOIJ, D. G. & SCHEDL, A. 2001. Sox9 induces testis development in XX transgenic mice. *Nat Genet,* 28**,** 216-7.

VISEL, A., RUBIN, E. M. & PENNACCHIO, L. A. 2009. Genomic views of distant-acting enhancers. *Nature,* 461**,** 199-205.

WADDINGTON, C. 1956. *Principles of development.,* London, Allen and Unwin.

WADDINGTON, C. H. 1940. *Organisers and genes*, Cambridge University Press.

WADDINGTON, C. H. 1942. Canalization of development and the inheritance of aquired characters. *Nature,* 150**,** 563-565.

WAGNER, T., WIRTH, J., MEYER, J., ZABEL, B., HELD, M., ZIMMER, J., PASANTES, J., BRICARELLI, F. D., KEUTEL, J., HUSTERT, E., WOLF, U., TOMMERUP, N., SCHEMPP, W. & SCHERER, G. 1994. Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell,* 79**,** 1111-20.

WALPORT, L. J., HOPKINSON, R. J., VOLLMAR, M., MADDEN, S. K., GILEADI, C., OPPERMANN, U., SCHOFIELD, C. J. & JOHANSSON, C. 2014. Human UTY(KDM6C) is a male-specific N-methyl lysyl demethylase. *J Biol Chem,* 289**,** 18302-13.

WANG, H., HU, Y. C., MARKOULAKI, S., WELSTEAD, G. G., CHENG, A. W., SHIVALILA, C. S., PYNTIKOVA, T., DADON, D. B., VOYTAS, D. F., BOGDANOVE, A. J., PAGE, D. C. & JAENISCH, R. 2013. TALEN-mediated editing of the mouse Y chromosome. *Nat Biotechnol,* 31**,** 530-2.

WARR, N., CARRE, G. A., SIGGERS, P., FALEATO, J. V., BRIXEY, R., POPE, M., BOGANI, D., CHILDERS, M., WELLS, S., SCUDAMORE, C. L., TEDESCO, M., DEL BARCO BARRANTES, I., NEBREDA, A. R., TRAINOR, P. A. & GREENFIELD, A. 2012. Gadd45gamma and Map3k4 interactions regulate mouse testis determination via p38 MAPK-mediated control of Sry expression. *Dev Cell,* 23**,** 1020-31.

WELSTEAD, G. G., CREYGHTON, M. P., BILODEAU, S., CHENG, A. W., MARKOULAKI, S., YOUNG, R. A. & JAENISCH, R. 2012. X-linked H3K27me3 demethylase Utx is required for embryonic development in a sex-specific manner. *Proc Natl Acad Sci U S A,* 109**,** 13004-9.

XI, H., SHULHA, H. P., LIN, J. M., VALES, T. R., FU, Y., BODINE, D. M., MCKAY, R. D., CHENOWETH, J. G., TESAR, P. J., FUREY, T. S., REN, B., WENG, Z. & CRAWFORD, G. E. 2007. Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. *PLoS Genet,* 3**,** e136.

XU, C. R., COLE, P. A., MEYERS, D. J., KORMISH, J., DENT, S. & ZARET, K. S. 2011. Chromatin "prepattern" and histone modifiers in a fate choice for liver and pancreas. *Science,* 332**,** 963-6.

XU, J., DENG, X. & DISTECHE, C. M. 2008a. Sex-specific expression of the X-linked histone demethylase gene Jarid1c in brain. *PLoS One,* 3**,** e2553.

XU, J., DENG, X., WATKINS, R. & DISTECHE, C. M. 2008b. Sex-specific differences in expression of histone demethylases Utx and Uty in mouse brain and neurons. *J Neurosci,* 28**,** 4521-7.

YANG, F., BABAK, T., SHENDURE, J. & DISTECHE, C. M. 2010a. Global survey of escape from X inactivation by RNA-sequencing in mouse. *Genome Res,* 20**,** 614-22.

YANG, X., ZHENG, J., XU, G., QU, L., CHEN, S., LI, J. & YANG, N. 2010b. Exogenous cMHM regulates the expression of DMRT1 and ER alpha in avian testes. *Mol Biol Rep,* 37**,** 1841-7.

YAO, B., WANG, Q., LIU, C. F., BHATTARAM, P., LI, W., MEAD, T. J., CRISH, J. F. & LEFEBVRE, V. 2015. The SOX9 upstream region prone to chromosomal aberrations causing campomelic dysplasia contains multiple cartilage enhancers. *Nucleic Acids Res,* 43**,** 5394-408.

YAO, H. H., MATZUK, M. M., JORGEZ, C. J., MENKE, D. B., PAGE, D. C., SWAIN, A. & CAPEL, B. 2004. Follistatin operates downstream of Wnt4 in mammalian ovary organogenesis. *Dev Dyn,* 230**,** 210-5.

YELLAJOSHYULA, D., LIM, J. W., THOMPSON, D. M., JR., WITT, J. S., PATTERSON, E. S. & KROLL, K. L. 2012. Geminin regulates the transcriptional and epigenetic status of

neuronal fate-promoting genes during mammalian neurogenesis. *Mol Cell Biol,* 32**,** 4549-60.

YODER, J. A., WALSH, C. P. & BESTOR, T. H. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet,* 13**,** 335-40.

YOSHIMOTO, S., OKADA, E., UMEMOTO, H., TAMURA, K., UNO, Y., NISHIDA-UMEHARA, C., MATSUDA, Y., TAKAMATSU, N., SHIBA, T. & ITO, M. 2008. A W-linked DM-domain gene, DM-W, participates in primary ovary development in Xenopus laevis. *Proc Natl Acad Sci U S A,* 105**,** 2469-74.

ZHAO, J., SUN, B. K., ERWIN, J. A., SONG, J. J. & LEE, J. T. 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science,* 322**,** 750-6.

ZHAO, L. & KOOPMAN, P. 2012. SRY protein function in sex determination: thinking outside the box. *Chromosome Res,* 20**,** 153-62.

Sara Alexandra García-Moreno
303 E Superior 7-250, Chicago IL 60611
sagm@u.northwestern.edu • (312) 503 0858

EDUCATION

Northwestern University   Graduate studies 2013-2018
Driskill Graduate Program in the Life Sciences

Universidad Nacional Autónoma de México   Undergraduate studies 2006-2011
B.S. with honors in Biology

Copenhagen University   Exchange program 2009-2010

RESEARCH EXPERIENCE

Duke University, Durham NC   2017-2018
Laboratory of Blanche Capel, Ph.D.
Finishing Ph.D. research focused on investigating the epigenetic regulation of mammalian sex determination.

Northwestern University, Chicago IL   2014-2017
Laboratory of Danielle Maatouk, Ph.D.
Ph.D. research focused on investigating the epigenetic mechanisms and cis-regulatory elements involved in the regulation of mammalian sex determination.

National Institute of Respiratory Diseases (INER), Mexico City   2011-2013
Laboratory of Immunobiology and Genetics, Joaquín Zúñiga, Ph.D.
Characterization of the microRNAs expressed by human pulmonary epithelial cells A549 infected with the 2009 pandemic virus A/H1N1 and the seasonal influenza virus A/PR/8/34, under the supervision of Gustavo Ramírez, Ph.D.

Laboratory of Microbiology, Martha Torres Rojas, Ph. D.
Social service (480 hours): standardizing a protocol for intracellular cytokine detection in T-lymphocytes stimulated with tuberculosis specific antigens CFP-10 and ESAT-6, under the supervision of Yolanda González, Ph.D.

Faculty of Medicine, UNAM, Mexico City   2010-2011
Laboratory of Immunotherapy and Tissue Engineering, Andrés Castell, Ph.D.

Undergraduate research focused on developing immunotherapy against melanoma in mice by re-introduction of differentiated dendritic cells cultured in melanoma-specific antigen MAGE-X and melatonin.

Laboratory of Immunotherapy and Tissue Engineering, Andrés Castell, Ph.D.   2008-2009
Undergraduate research focused on tissue engineering of guinea pig corneal cells and of mouse epithelial tissue.

PUBLICATIONS

García-Moreno SA, Futtner CR, Salamone IM, Gonen N, Lovell-Badge R, Maatouk DM. *Supporting gonadal cells acquire sex-specific chromatin landscapes during sex determination*. In review.

Gonen N, Futtner CR, Wood S, **García-Moreno SA,** Salamone IM, Samson SC, Sekido R, Poulat F, Maatouk DM, Lovell-Badge R, 2018. *Sex reversal following deletion of a single distal enhancer of Sox9*. Science 360(6396): 1469-1473.

García-Moreno SA, Plebanek MP, Capel B, 2018. *Epigenetic regulation of male fate commitment from an initially bipotential system*. Molecular and Cellular Endocrinology 468: 19-30.

Ramírez-Martínez G, Cruz-Lagunas A, Jiménez-Álvarez L, Ortíz-Quintero B, Santos-Mendoza T, Herrera MT, Canché-Pool E, Mendoza C, Bañales JL, **García-Moreno SA,** Morán J, Cabello C, Orozco L, Aguilar-Delfín I, Hidalgo-Miranda A, Romero S, Suratt BT, Selman M, Zúñiga J, 2013. *Seasonal and pandemic influenza H1N1 viruses induce differential expression of SOCS-1 and RIG-1 genes and cytokine/chemokine production in macrophages*. Cytokine 62(1): 151-9.

Bautista E, Arcos M, Jiménez-Álvarez, García-Sancho MC, Vázquez ME, Peña E, Higuera A, Ramírez G, Fernández-Plata R, Cruz-Lagunas A, **García-Moreno SA**, Urrea F, Ramírez R, Correa-Rotter R, Pérez-Padilla JR, Zúñiga J, 2013. *Angiogenic and inflammatory markers in acute respiratory distress syndrome and renal injury associated to A/H1N1 virus infection*. Experimental and Molecular Pathology, 94(3): 486-92.

FELLOWSHIPS

The Cellular and Molecular Basis of Disease Training Program   March 2015-March 2018
Northwestern University (T32 NIH T32 GM08061)

POSTERS, PRESENTATIONS and AWARDS

The 8th International Symposium on the Biology of Vertebrate   April 2018
Sex Determination
King Kamehameha's Kona Beach Hotel, Kona HI
Oral Presentation: Epigenetic Regulation of Sex Determination in the Mouse Fetal Gonad.

The 36th Center for Reproductive Science Minisympoisum   April 2017
Northwestern University, Chicago IL
Award Winning Oral Presentation: Epigenetic regulation of sex determination in the mouse fetal gonad.

The 75th Annual Meeting for the Society for Developmental Biology   August 2016
Marriot Copey Place, Boston, MA
Poster presentation: García-Moreno SA, Salamone IM, Maatouk DM, *Polycomb-mediated chromatin remodeling during sex determination.*

The 7th Illinois Symposium on Reproductive Sciences   October 2015
University of Urbana-Champaign, Champaign IL
Member of the organizing committee

Genetics and Genomics Cluster Annual Symposium   August 2015
Poster Presentation: García-Moreno SA, Salamone IM, Maatouk DM, *Polycomb-mediated chromatin remodeling during sex determination.*

The Mouse Genetics Group Lecture and Poster Competition   May 2015
Northwestern University, Chicago IL

<u>2nd Place Award Winning Poster Presentation:</u> García-Moreno SA, Salamone IM, Maatouk DM, *Polycomb-mediated chromatin remodeling during sex determination.*

The 7th International Symposium on the Biology of Vertebrate   April 2015
Sex Determination
King Kamehameha's Kona Beach Hotel, Kona HI
Poster Presentation: García-Moreno SA, Salamone IM, Futtner CR, Natarajan A, Capel B, Maatouk DM, Epigenetic Regulation of Sex Determination in the Mouse Fetal Gonad.

The 34th Center for Reproductive Science Minisympoisum   January 2015
Northwestern University, Chicago IL
Oral Presentation: *Polycomb-mediated chromatin remodeling during sex determination.*

The 6th Illinois Symposium on Reproductive Sciences   October 2014
University of Illinois at Chicago, Chicago IL
<u>3rd Place Award Winning Poster Presentation:</u> García-Moreno SA, Salamone IM, Maatouk DM, *Polycomb-mediated chromatin remodeling during sex determination.*

Cell Symposia: Transcriptional Regulation in Development   August 2014
Northwestern University, Chicago IL
Volunteer