Deciphering the molecular mechanisms of gonadal development

Sandra Elena Rojo Mendoza

Université Pierre et Marie Curie

Ecole doctorale 394. Physiologie, Physiopathologie et Thérapeutique

Unit of Human Developmental Genetics

Department of Developmental and Stem Cell Biology
Institut Pasteur,
25 Rue du Dr. Roux
75015 Paris

Deciphering the molecular mechanisms of gonadal development

Par Sandra Rojo

Pour obtenir le grade de

Docteur de l'Université Pierre et Marie Curie (UPMC-Paris 6)

Dirigée par Ken McElreavey & Anu Bashamboo

Présentée et soutenue publiquement le 25 Septembre 2015

Devant un jury composé de :

Pr. Siffroi Jean Pierre – President du Jury

Dr. Pailhoux Eric - Rapporteur

Dr. Poulat Francis - Rapporteur

Dr. Brauner Raja – Examinateur

Dr Hiort Olaf – Examinateur

Acknowledgements

Though only my name appears on the cover of this PhD dissertation, a lot of great people have contributed to its production. I owe my gratitude to all those people who have made this dissertation possible and because of whom my graduate experience has been one that I will cherish forever.

I thank all of my jury for taking the time of reading and work in my manuscript. Thank you Dr Siffroi for accepting to be the president of my jury. Thank you Dr. Francis Poulat and Dr. Eric Pailhoux for accepting to be my reporters and taking the time to read this manuscript. Many thanks also to Dr. Hiort Olaf and Dr. Raja Brauner, for being part of my jury as examinators of work I have done.

I want to sincerely thank Ken McElreavey and Anu Bashamboo for accepting me into their lab, for introducing me into the research of a wonderful and enormously passionating subject. My PhD has been a wonderful experience that has changed me and help me to grow not only as a scientist but as a more complete person.

I thank all the people that have been part of the lab while I was there. Joelle, you have been there from day 0 until the very end, thank you for cheering me up when I was sad, stressed or sick. Thank you for the candies, for sharing our coffee addiction, and sharing your advices and all ingenuous tricks. For the mutagenesis, sequencing and western blots! Thank you for being very positive and caring person. Thank you Jo! I cannot imagine my PhD without you being around. Pierre and Namya for the laughs, for sharing your promoters, I wish you a lot of success in your next step. MC thank you for sharing our lab room, for the replicas you have done, I hope everything goes very well and that you can finish your PhD as best as possible.. Remember... everything will be fine at the end!! You are almost there! Caro and Capucine, thank you a lot for your help while I was writing!! You are so sweet and I wish you have a lot of

papers, a lot of learning and great moments in the lab. I am sure it will be this way! To Laurence, thank you for all your help! You are the best secretary the lab could have, thank you for always been there when I needed something, even when you didn't have to help me, for the university and other very confusing French paper work I didn't understood. Merci beaucoup!
To my tutor Monica Sala, thank you for being there, for listening me and for helping me whenever it was possible and for your precious advice that made my PhD much easier.

Tanguy, muchas gracias por todo tu amor, y sobre todo tu infinita paciencia al final de mi tesis. No creo que nadie podría haberme dado tanta paciencia como tu lo has hecho. Gracias por tu ayuda con mis indexes, tablas, pdfs y francés.. al final, mi testis también tiene una parte tuya. No hay palabras que puedan agradecerte tu apoyo incondicional. Te amo mon amour, mon ami.

Gracias a mi familia francesa adoptiva: Marion, Ariane, Robin, Virginie, et Francis merci beaucoup pour tout votre support et les beaux moments. Vous etes les meilleurs. Je vous remercie infiniment.

A mis amigos, que aunque algunos están lejos, siempre están. Fede, Yoalli, Mari, Lu, Sofi, Cori, Mario, Merypi, Yare, los quiero, gracias por crecer conmigo, por las experiencias y por estar ahí.

A mi familia mexicana, gracias por creer en mi. Por todo su amor y apoyo incondicional, por acompañarme en esta aventura a la distancia, siempre están conmigo en mi corazón. A mi mamá, tampoco hay palabras que pueda decir para expresar la gratitud y amor que tengo. Todos y cada uno de mis logros tienen una parte de tí, no solo me diste la vida, cada día me sigues dando alas y fuerza para creer que puedo hacer las cosas bien y lograr cada uno de mis objetivos. Te amo mami, esta tesis es tuya también.

Finally, I appreciate the financial support from STEM DIM Pole and REVIVE that funded parts of the research discussed in this PhD dissertation.

Table of Contents

ABSTRACT

In mammals, sex determination and development of the embryonic gonads is a complex molecular process that involves a large number of genes and networks acting synergistically or antagonistically. Testis formation is initiated by *SRY* in synergy with NR5A1/SF1 by up-regulation of *SOX9* expression beyond a critical threshold. This results in the positive regulation of male-specific program and repression of the ovarian pathway. In females, the absence of *SRY* results in the activation of signaling pathways initiating ovarian development and repression of the testis-determining network. Errors in this process may result in either 46,XY gonadal dysgenesis, or testis development in a XX individual. Although, we are beginning to build up a picture of the genes involved in gonad development, mutations in these genes can explain only a minority of patients presenting with errors in sex development, and the mechanism leading to inappropriate sexual development remain largely unknown.

The DMRT1 protein, contains a DNA-binding domain, regulates sex-determination in Drosophila, the nematodes, as well as in some species of fish and birds, but its role in mammalian sex-determination in unclear. In mice, *Dmrt1* does not play a role in primary sex-determination but is required for the maintenance of adult testis. In this work we identified, for the first time, a point mutation in the human *DMRT1* gene associated with a lack of testis-determination. A series of functional analyses demonstrated the mechanism by which this mutation could lead to XY gonadal dysgenesis. The results show that, in contrast to the mouse, *DMRT1* is involved in

primary testis-determination in humans and also indicate that at the molecular level sex-determination may be a much more conserved process than previously thought.

Further studies on patients with 46,XY gonadal dysgenesis using a number of approaches, including exome sequencing, lead to the identification of potentially pathogenic mutations in two genes related to *SRY* and *SOX9* - *SOX7* and *SOX8.* This suggests that there may be a functional redundancy amongst SOX genes in human sex-determination. Furthermore, we identified mutations in the transcription factor *GATA4* associated with 46,XY gonadal dysgenesis – helping to establish an important role for this gene as a cause of gonadal dysgenesis.

The functional consequences of these mutations, on the biological activity of the protein, were assessed using a number of classical approaches – cellular localization, reporter assays, DNA-protein interactions, protein-protein interactions etc. These approaches have indicated a change in biological activity of the mutated proteins and in most cases failed to reveal the mechanisms involved in the development of the phenotype. Therefore, the development of novel *in-situ* cellular models may provide us with a tool to identify these, yet illusive, mechanisms. Using mouse embryonic stem cells (ESCs) as the model system we are developing novel cellular models to understand the biological context of these mutations in the appropriate environment. Once established, these models would be applicable to understand the development and progression of disease using cells derived from patients presenting with DSD.

Résumé

Chez les mammifères, la détermination sexuelle et le développement embryonnaire des gonades est un processus moléculaire complexe qui implique un grand nombre de gènes et de réseaux agissant en synergie ou antagonisme. La formation de testicules est initiée par SRY en synergie avec NR5A1/SF1 par sur-activation de l'expression de SOX9 au delà d'un seuil critique. Cela a pour résultat la régulation positive du programme masculin et la répression de la voie de différenciation de l'ovaire. Chez la femme, l'absence de SRY entraîne l'activation des voies initiant le développement ovarien et la répression de la signalisation de différenciation des testicules. Des erreurs dans ce processus peuvent entraîner des pathologies de type femme 46,XY avec dysgénésie gonadique, ou à l'inverse le développement de testicules chez des individus 46,XX. Bien que nous commençons à avoir une vision précise des gènes impliqués dans le développement des gonades, les mutations de ces gènes n'expliquent qu'une minorité de patients présentant des erreurs dans le développement du sexe, et les mécanismes conduisant à un développement sexuel inapproprié restent encore largement méconnus.

La protéine DMRT1, qui contient une zone de liaison à l'ADN, est un facteur crucial de la détermination du sexe chez la drosophile, les nématodes, ainsi que chez certaines espèces de poissons et chez les oiseaux, mais son rôle dans la détermination du sexe chez les mammifères n'est pas clair. Chez la souris, *Dmrt1* ne joue pas un rôle dans la détermination du sexe primaire, mais est nécessaire

pour le maintien de l'identité testiculaire à l'âge adulte. Dans ce travail, nous avons identifié, pour la première fois, une mutation ponctuelle dans le gène *DMRT1* humain associée à une absence de détermination des testicules. Une série d'analyses fonctionnelles démontrent le mécanisme par lequel cette mutation pourrait conduire à une pathologie de type 46,XY avec dysgénésie gonadique. Les résultats montrent que, contrairement à ce qui est décrit chez la souris, DMRT1 est impliqué dans la détermination primaire des testicules chez l'homme. Ces résultats indiquent également qu'au niveau moléculaire, la détermination du sexe peut être un processus beaucoup plus conservé que précédemment envisagé.

D'autres études concernant des patientes 46,XY avec dysgénésie gonadique, s'appuyant sur un certain nombre d'approches y compris le séquençage de l'exome, conduisent à l'identification de mutations potentiellement pathogènes dans deux gènes liés à *SRY* et *SOX9*: *SOX7* et *SOX8*. Cela amène à penser qu'il peut y avoir une redondance fonctionnelle entre les gènes SOX dans la détermination du sexe chez l'homme. En outre, nous avons identifié des mutations dans le facteur de transcription GATA4 associées également à un phénotype 46,XY avec dysgénésie gonadique, laissant supposer un rôle important de ce gène comme cause de dysgénésie gonadique.

Les conséquences fonctionnelles de ces mutations sur l'activité biologique des protéines mutantes, ont été évaluées en utilisant un certain nombre d'approches classiques: localisation cellulaire, dosages de gènes rapporteurs, interactions ADN-protéine, interactions protéine-protéine, etc... Ces approches ont parfois indiquées un changement dans l'activité biologique des protéines mutées mais dans la plupart des cas, n'ont pas réussi à révéler les mécanismes impliqués dans

le développement du phénotype morbide. Par conséquent, le développement de nouveaux modèles cellulaires *in vitro* pourrait fournir un outil plus performant pour identifier ces mécanismes. En utilisant des cellules souches embryonnaires de souris (ESCs en anglais) comme système modèle, nous avons développé de nouveaux modèles cellulaires pour comprendre le contexte biologique de ces mutations dans l'environnement approprié. Une fois établis, ces modèles seraient applicables pour comprendre le développement et la progression de la maladie en utilisant des cellules provenant directement de patients présentant des dysfonctionnements du déterminisme du sexe nommés DSD (Disorders of Sex Development).

AIMS and OBJECTIVES

The aim of this work was to understand the fundamental mechanisms of human sex determination through the detailed genetic and molecular analysis of human cases of sex reversal (46,XY gonadal dysgenesis, 46,XX testicular or ovotesticular DSD) as well as a large and unique ancestry matched cohort population. This thesis review current knowledge in sex determination, development and maintenance of mammalian gonads. It describes novel mutations found in patients with DSD, mostly by using exome sequencing and direct sequencing methods. For many if these mutations detailed *in silico* and functional analysis was performed in order to understand how this mutations caused the error of testis development, or testis formation in a XX individual. These genes include:

- *DMRT1*
- *SOX7*
- *SOX8*
- *GATA4*
- *RevSex* regulatory region of *SOX9*

Moreover it will describe and discuss our preliminary results of the derivation of a mESCs stable steroidogenic cell line that would be developed as a suitable cellular model for studying sex development and DSD.

INTRODUCTION

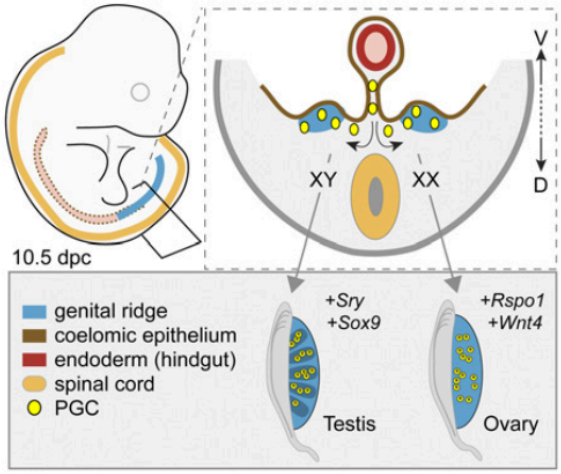
The propagation of the vertebrate species relies on the development of reproductive organs that support the differentiation of germ cell lineage into two types of functional gametes that carry the genetic material: sperm and oocytes. (DeFalco and Capel, 2009). In mammals, the testis and the ovary develop from the embryonic gonad, which is initially bipotential. Depending on the molecular signals during development, the bipotential anlage must decide between two mutually exclusive cell fates, referred as primary sex determination, to differentiate into supporting cells of either a testis or an ovary.

Bipotential Gonad.

Primordial cells in bipotential gonad in both sexes have the potential of give rise to lineages with similar functions. XX ridges can be stimulated to differentiate as testis by the expression of male sex-determining genes such as *Sry* or *Sox9* and XY ridges can differentiate into ovaries by the absence of testis-determining genes (Fig 1) (Barrionuevo, 2005; Vidal et al., 2001; Chaboissier et al., 2004). The molecular pathways operating to determine formation of testis and ovaries are mutually antagonistic. (Kim and Capel, 2006; Jameson et al., 2012; Ottolenghi et al., 2007).

Formation of the Bipotential gonads

In mouse, the bipotential gonad appears around embryonic day E10.5, whereas in humans the gonadal precursor can be detected at day 32 post conception. The indifferent gonad arises as parallel thickening of the epithelium at the ventromedial surfaces of the mesonephros. The coleomic epithelial proliferation covers each mesonephros and this results in the outgrowth of the genital ridge (Karl and Capel, 1998; Schmahl et al., 2000). The genital ridge mesenchyme grows by proliferation, recruitment of cells from the adjacent mesonephros and the ingression of cells from the coleomic epithelium (Fig 1). Several transcription factors are critical for the development of the undifferentiated bipotential gonad during embryogenesis. Mainly by mouse knock out studies, several genes including *Lhx9, Emx2, Wt1, Cbx2, Nr5a1, Gata4 and Six1/4* have been implicated in establishment of the bipotential gonad in both males and females (Fig 2) (Svingen and Koopman, 2013).



INTRO - Figure 1. <u>Bipotential gonadal ridges establishment and sex determination.</u>

Genital ridges (blue) appear as an outgrowth along the mesoneprhoi in the coleomic epithelium cavity. In mouse it can distinguish from E10, cells are recruited from the overlying coleomic epithelium (brown). Primordial germ cells (yellow) leave the hindgut (red) via the dorsal mesentery to colonize the genital ridges (arrows). At this stage, genital ridges are bipotential and depending on the genetic signals can differentiate in either testis or ovaries. At E10.5, *Sry* is expressed in the XY genital ridges and initiate Sox9 expression leading to testis differentiation. In absence of *Sry*, XX gonads initiate Rspo1 and Wnt4 signaling and develop towards ovaries. (D) Dorsal; (V) ventral. (Fig from T. Svingen and P. Koopman, Genes and Dev, 2013)

Homeobox proteins **Emx2** and Lim/homeobox protein **Lhx9** have essential roles in the early development of the bipotential gonad. In humans, ***EMX2*** is located on chromosome 10; it encodes a transcription factor containing a pentapeptide motif that may facilitate its interaction with TALE homeobox proteins. EMX2 has a critical role in early development of the urogenital and central nervous system. *Emx2* mutant mice lack kidneys, gonads and genital tracts (Miyamoto et al., 1997). In the *Emx2* $^{-/-}$ forming genital ridges, cells of the coleomic epithelium lose their polarity leading to aberrant tight junction assembly. Moreover, *Emx2*$^{-/-}$ mice coleomic epithelial cells that migrate in the mesenchymal compartment are reduced in numbers, this results in the impairment of the gonad formation (Kusaka et al., 2010).

LHX9 is located on human chromosome 1, it encodes a 397 amino acid protein with two amino terminal domains, a DNA binding homeobox and another involved in protein-protein interaction (Birk et al., 2000). Lhx9 regulates the expression of a key element required for early gonad development: *Nr5a1.* Lhx9 can bind directly to *Nr5a1* promoter *in vitro*, and has an additive effect on the Wt1 induced activation (Wilhelm and Englert, 2002). *Lhx9* deficient mice develop discrete gonads and by E13.5 show a regression of the gonads due to markedly reduced cell proliferation (Birk et al., 2000).

WT1 is located on chromosome 11 in humans and encodes for Wilms' tumor 1 zinc finger transcription factor. WT1 has two isoforms: with or without an additional three amino acids (lysine-threonine and serine) between the third and the fourth zinc finger (+KTS and -KTS); both have different roles during embryogenesis. $WT1^{-KTS}$ is essential for the formation and development of the bipotential gonads, and binds preferably DNA, whereas $WT1^{+KTS}$ has more affinity for RNA. Wt1 regulates the expression of different genes involved in urogenital development including *AmhR2, Sox9, Nr5a1, Wnt4* and *Sry* (Wilhelm and Englert, 2002). $Wt1^{-/-}$ mice of both sexes lack kidneys and gonads, due to apoptosis of the gonadal primordium (Kreidberg et al., 1993). In humans, heterozygous deletions of *WT1* are linked to a predisposition to Wilms tumor and mild genitourinary anomalies; whereas heterozygous missense *WT1* mutations are linked to Denys-Drash syndrome (Niaudet and Gubler, 2006; Clarkson et al., 1993; Ludbrook and Harley, 2004).

Human ***CBX2*** is located on chromosome 17 and encodes for Chromobox homolog 2, a component of the polycomb group complex of regulatory proteins. *Cbx2* is involved in the formation and development of the genital ridges. Mice lacking *Cbx2* present delayed gonad development and male to female sex reversal. In XX mice, *Cbx2* deficiency results in a reduced size of ovaries (Katoh-Fukui et al., 1998). In humans, heterozygous loss of function mutations in *CBX2* have been reported in a 46,XY individual with female internal and external genitalia and normal ovaries (Biason-Lauber et al., 2009). Evidence indicates that Cbx2 is involved in the regulation of several genes involved in testis development such as *Sry*, *Nr5a1, Wt1* and *Sox9* (Katoh-Fukui et al., 2012).

NR5A1 is present on chromosome 9 in humans and encodes the nuclear receptor Steroidogenic factor 1 (SF1) and is also known as NR5A1 or AD4BP**.** In mice *Nr5a1* is expressed in the developing urogenital ridge, pituitary and hypothalamus (Ikeda 1996, 2001). Deletion of *Nr5a1* in mouse results in complete agenesis of the gonad and adrenal gland (Luo et al., 1994; Sadovsky et al., 1995). In humans, *NR5A1* mutations result in primary adrenal failure and complete XY sex reversal (Achermann et al., 1999). In addition to its function in the bipotential gonad development, Nr5a1 plays multiple roles during testis development by regulating the expression of genes such as *Sox9* and *Amh* (See section on NR5A1 below).

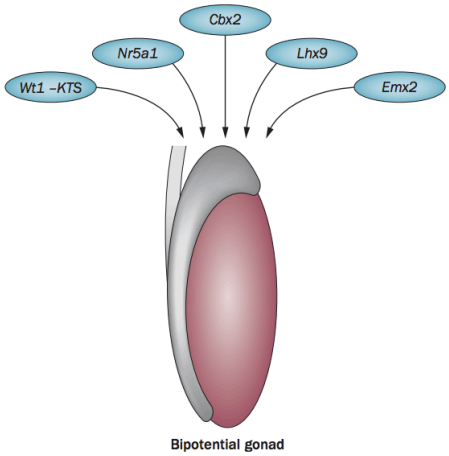
GATA4 is a member of the GATA zinc finger transcription factor family. It is located on human chromosome 8. In mouse embryos, *Gata4* expression is detected in the gonadal primordial prior and during sex determination (Ketola et

al., 2000) and it remains visible until E11.5 in the somatic cells of both XX and XY gonads. Around E13.5 *Gata4* expression becomes dimorphic, with upregulation in the male gonad. *Gata4* $^{-/-}$ mice die before the genital ridge forms due to cardiac defects (Molkentin et al., 1997). *Gata4* conditional knockout mice do not show signs of initiation of genital ridge formation at E8.75, because their coleomic epithelium remains as an undifferentiated monolayer (Hu et al., 2013), suggesting an early role for *Gata4* in gonadogenesis. In the XY *Gata4*^{ki} mouse, which abrogates the interaction of GATA4 with its cofactor FOG2, the genital ridge forms but testis differentiation is blocked due to a reduction in the levels of *Sry* expression (Tevosian et al., 2002). In humans, *GATA4* mutations have been reported in cases of abnormal cardiac development (Garg et al., 2003) and a familial case with CGD and 46,XY DSD (Lourenço et al., 2011a) (See section on GATA4 below)

Six1 and ***Six4*** genes belong to the mammalian homolog of the Drosophila sine oculis homeobox (Six) family, including six members (*Six1-Six6*) in the mouse. Six family genes encode transcriptional factors with both Six domains and homeodomains (Kawakami et al., 2000). In humans *SIX1* and *SIX4* are closely located on chromosome 14, and, mouse *Six1* and *Six4* are located in the same genomic region on chromosome 12. *Six1 and Six4* are expressed in overlapping tissues during mouse embryogenesis including the gonads (Kumar, 2009). Mice lacking either *Six1* or *Six4* have normal gonadal development. However, in the *Six1;Six4* double knockout, XY mice show reduced size of the gonads, decreased levels of *Sry* expression and features of ovarian development. Moreover, both, XY and XX mice offspring show reduced gonad size and a diminished number of

gonadal precursor cells before the onset of *Sry* expression. *Fog2* and *Nr5a1* are downstream targets of Six1/Six4 suggesting that the role of Six genes in the growth of the gonad precursor cells in the bipotential gonad is achieved by regulating *Nr5a1* and also *Sry* via *Fog2* (Fujimoto et al., 2013).

In the developing bipotential gonad, the supporting cell lineages are the first to acquire a sex specific fate. In males, the first cells to differentiate are Sertoli cells and in females their counterpart are granulosa cells.

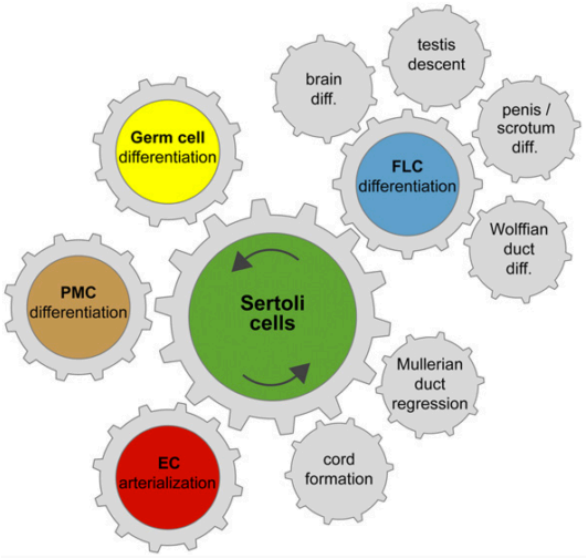


INTRO - Figure 2. Bipotential gonad development genetic network.

Genes essential for the correct development of the bipotential gonad identified by knockout mouse strains show complete absence of gonad development. WT1 (-KTS) in synergy with LHX9 bind and activate *Sf1* promoter. Cbx2 regulates in the same way Sf1 expression in the spleen and adrenal gland, suggesting a similar effect in the bipotential gonad. Emx2 is thought to have a role in proliferation, differentiation and cell survival pathways in bipotential gonad development (Fig from S. Eggers et al., 2014)

Testis differentiation

The differentiation of the testis begins with the expression of the Y chromosome specific gene *SRY* in a subset of somatic cells that will give rise to Sertoli cells (Koopman et al., 1990). The appearance of Sertoli cells orchestrates the differentiation of the other cell types of the testis: Leydig cells and peritubular myoid cells (PMCs) (Fig 3).



INTRO - Figure 3. Sertoli cells as a hub of testis differentiation and organization.

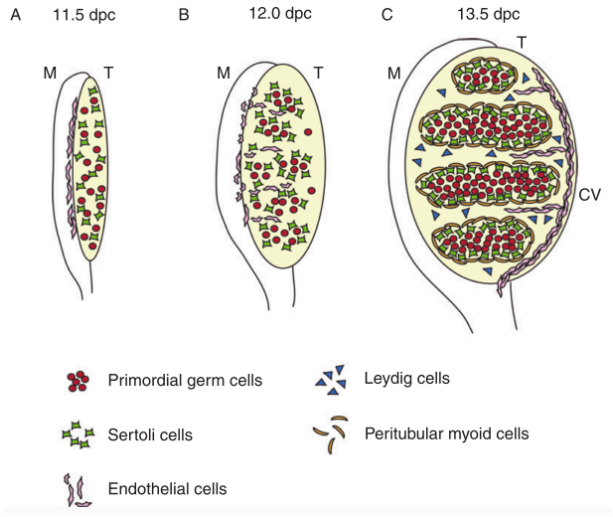
Sertoli cells are the first somatic cells to appear in testis development. They influence testis cord formation, Müllerian duct regression, and differentiation of FLCs, PMCs, germ cells and endothelial cells. Fetal Leydig cells are in charge of androgen production and thus masculinization of the organism. (Fig from T. Svingen and P. Koopman, Genes and Dev, 2013)

Testis Morphology

Sertoli cells.

Sertoli cells are the first cells to differentiate within the bipotential gonad, and thus the first indicator of the commitment of the undifferentiated gonad towards male development. Murine fetal Sertoli cells derive from the coleomic epithelium and can be detected around E11.2-11.4 (Lin and Capel, 2015).

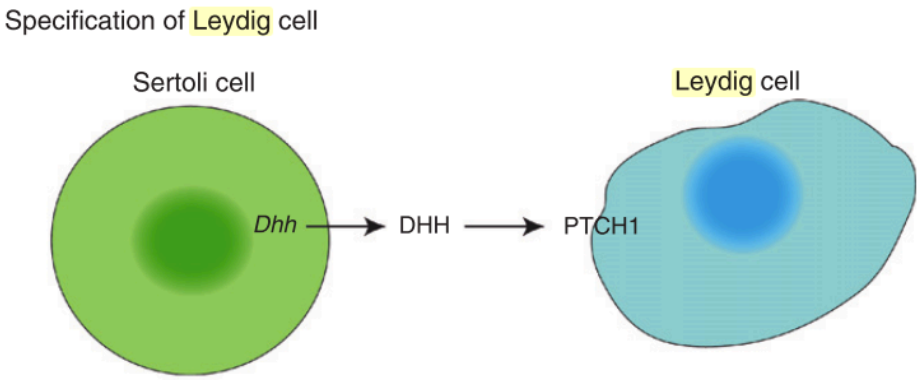
The specification of precursors of Sertoli cells towards mature Sertoli cells is marked by the expression of *Sox9* and their polarization to form epithelial aggregates around germ cells that will shape testis cords (Brennan and Capel, 2004) (Fig 4).



INTRO - Figure 4. <u>Model of testis development in mouse.</u>

(A) At E11.5 SRY is expressed in pre-Sertoli cells (green), equally distributed along with PGCs (red) in the genital ridge. Endothelial cells (pink) start to migrate into the gonad. (B) At E12 Sertoli cells become polarized and form epithelial aggregates around germ cells. Endothelial cells separate these clusters into the future testis cords. (C) At E13.5 testis cords are defined: clusters of GCs are surrounded by Sertoli cells and one layer of peritubular myoid cells (yellow). In the interstitium, Leydig cells (blue) produce testosterone. T: testis; M: mesonephros, CV: coelomic vessel. (Figure from E. Wainwright and D. Wilhelm, 2010)

Signals from developing Sertoli cells initiate the appearance of Leydig cells in the interstitial space (Habert et al., 2001). The differentiation of Leydig cells depends on signals expressed by Sertoli cells. *Dhh*, a secreted signaling factor by Sertoli cells acts via Patched receptor (PTCH1), expressed in interstitial cells at the time of FLC differentiation to specify FLC lineage. It has been suggested that Dhh/Ptch1 signaling pathway triggers the expression of *Nr5a1* and *Cyp11a1* in the interstitial precursor cells. (Barsoum and Yao, 2006). *DHH* mutations are associated with complete or partial 46,XY gonadal dysgenesis (Canto et al., 2004, 2005). At E12.5, PDGF factors and receptors are expressed in fetal XY gonads, PdgfA in Sertoli cells and Pdgf*α* in the interstitium. *Pdgfα* null mice show a reduced Sertoli cell and fetal Leydig cell proliferation and differentiation, and disruption in mesonephric cell migration (Brennan et al., 2003; Pierucci-Alves et al., 2001) (Fig 5).



INTRO - Figure 5. Specification of fetal Leydig cells.

The morphogene DHH expressed by Sertoli cells induces Leydig cell specification through its receptor PTCH1. (Figure from E. Wainwright & D. Wilhelm, 2010)

Once Sertoli cells mature they produce AMH, inducing the regression of the "female" Müllerian duct and establishing the development of the "male" Wolffian duct (Wilhelm and Koopman, 2006).

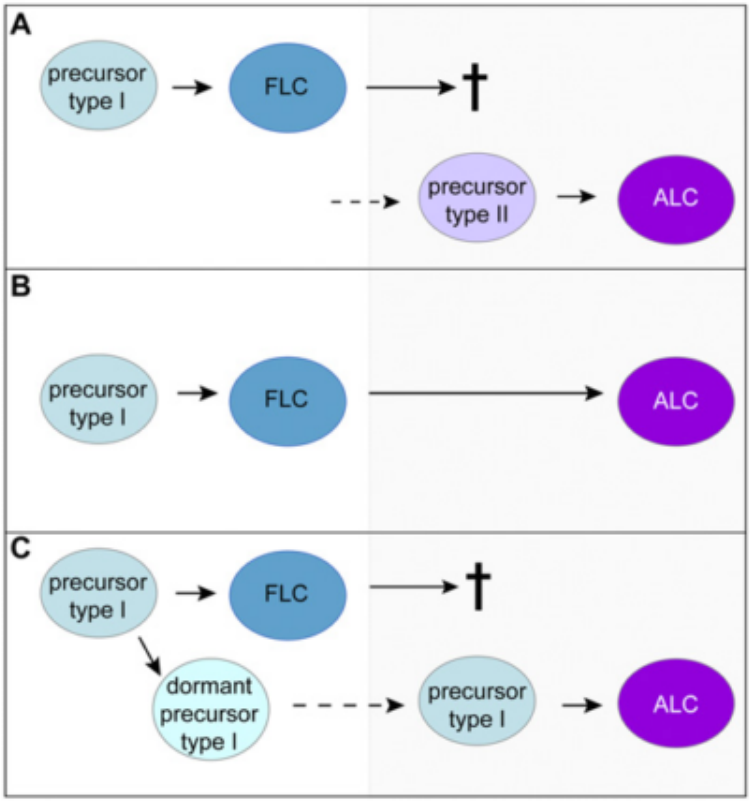
Adult Sertoli cells support germ cells and provide them the needed factors to become mature and functional sperm. Each Sertoli cell can cope with a determined number of germ cells. Therefore, the number of Sertoli cells and sperm production is directly correlated (Petersen and Söder, 2006).

Leydig cells.

Fetal Leydig cells (FLCs) appear in the interstitium at E12.5 in mouse, shortly after testis determination, and remain until they are replaced by the adult Leydig cells around birth (Habert et al., 2001). Their primary function is the production of testosterone, which virilizes the internal and external genitalia. They also produce insulin like-3 (INSL3) that participates in the descent of the testis. The origin and differentiation of FLCs is controversial, but we know that a Nr5a1 positive population is shared and divided between the adrenal cortex and the gonad. This cell population expands along the anterior of the mesonephros and into the gonad before E11.5 (Hatano et al., 1996). At this time point expression of *Nr5a1* is active in both pre-Sertoli and pre-FLCs. Afterwards, *Nr5a1* expression diminish in fetal Sertoli cells and increases in differentiating FLCs, when they regulate the elements involved in testosterone production (Sugawara et al., 1996; Morohashi et al., 1992).

Adult Leydig cells (ALCs) differentiate after birth; whether they have or not a different origin than their fetal counterparts is still a subject of debate. It seems that ALCs derive form peritubular like stem cells (Ge et al., 2006). In rodents, the Leydig stem cells proliferate during fetal development to become Leydig progenitor cells (around 10-14 post natal day); these ALCs intermediaries produce low levels of androgens, steroidogenic enzymes, and express luteinizing hormone (LH). Leydig progenitors cells divide and differentiate into immature Leydig cells and decay rapidly in numbers. Immature Leydig cells persist for a short duration and secrete high quantities of testosterone metabolites. Immature Leydig cells divide once and give rise to ALCs (Chen et al., 2009). Mature ALCs produce high levels of testosterone, needed for pre-pubertal development of external male genitalia and the onset of spermatogenesis, among other functions (Teerds and Huhtaniemi, 2015).

The origin of Leydig cell populations is not completely understood. There are many morphological and genetic studies suggesting they arise from different populations. However, this idea has recently been challenged and different models for the origin of FLCS and ALCs are now considered (Fig 6). LCs arise in two separate phases of development in eutherian animals. It is possible that, similar to Sertoli cells, LCs may have a different morphology and transcriptional profile during different developmental stages; but they share a common precursor. The differences between fetal, immature and mature LC populations may be linked to stage-dependent androgen levels in the testis (Svingen and Koopman, 2013).



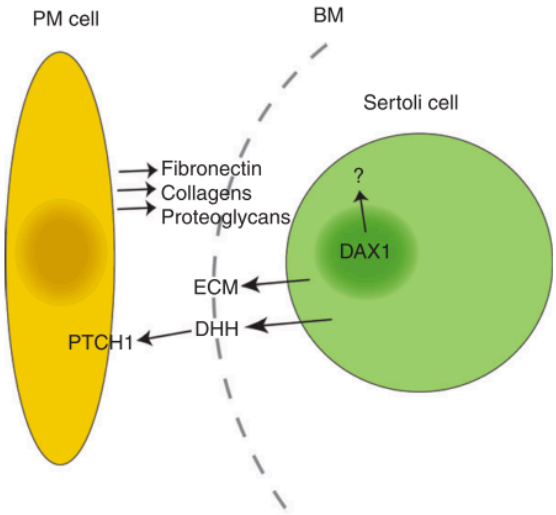
INTRO - Figure 6. Different models of FLCs and ALCs origin.

In most mammals, the FLC population seems to be lost after birth, and replaced by a new ALC population. However, different hypotheses exist to explain the origins of the two populations. (A) The most popular hypothesis: Both populations arise from two unique and different precursor cell populations unrelated to each other, with the FLCs disappearing after birth. (B) The least-regarded hypothesis: Both cell lines are the same cell lineage; the FLC population is almost completely lost after birth, but few remaining cells divide and differentiate to give rise to ALCs. (C) Alternative hypothesis: FLCs population decreases after birth, the ALC population is established from the same precursor cells that FLCs after lying dormant throughout prepubertal stage. (Figure from T. Svingen and P. Koopman, Genes & Development 2013)

Peritubular myoid cells.

Peritubular myoid cells (PMCs) origin is quite unclear. They were thought to be one of the three type cells that migrate from the mesonephros into the male gonads (Merchant-Larios et al., 1993). However, recent evidence suggests that the only cells migrating from the mesonephros are endothelial cells that contribute to testis vasculature (Cool et al., 2008; Combes et al., 2009). PMCs probably originate from an unknown cell population in the gonad. PMCs

differentiation is thought to be started by factors secreted by Sertoli cells, such as DHH, since *Dhh* null mice, develop abnormal PMCs and malformed testis cords (Pierucci-Alves et al., 2001). In addditon, Dax1 expression in Sertoli cells is required for PMC differentiation (Meeks et al., 2003). PMCs form a single layer of flattened cells adjacent to Sertoli cells, marking the perimeter of the testis cords (Fig 7). They have two main functions: (i) to contribute to the structural conformation of the testis cords in conjunction with Sertoli cells and (ii) to promote the movement of mature sperm through the adult seminiferous tubules to the seminal vesicles, a function mediated by their smooth muscle like character. There is no clear evidence for specific markers for this cell line, although they express α-smooth actin *(αSma*) and desmin. (Capel et al., 1999; Martineau et al., 1997).

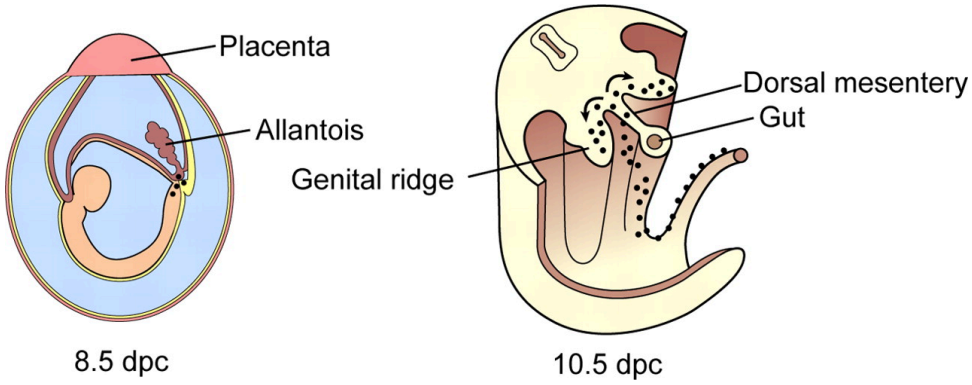


INTRO - Figure 7. <u>Interaction between PMC and Sertoli cells</u>.

Sertoli cell regulation of DHH signaling and interaction of PTCH1 receptor, results in differentiation of PMCs. Dax1 expression from Sertoli cells is also necessary for PMCs differentiation. (Figure from E.Wainwright & D. Wilhelm, 2010)

Primordial Germ Cells

Primordial germ cells (PGCs) do not arise form the same cell population as the somatic cells within the genital ridge. Because they are alkaline phosphatase positive, it is possible to track them from their place of origin, the base of the allantois at the posterior end of the primitive streak (Lawson and Hage, 1994). PGCs are first seen in the mouse at E7, in the developing hindgut. When the hindgut invaginates, PGCs are swept into the embryo, in a process that seems to be passive, since PGCs are still non motile. By E9.5, PGCs move out from the hindgut and pass into the forming urogenital ridges. As development proceeds, the hindgut descends into the coleomic cavity and PGCs migrate through the dorsal mesentery before entering the developing gonads (Fig 8). The survival and migration of PGCs is dependent on the interaction of the tyrosine kinase receptor c-Kit, expressed by germ cells, and its ligand Stem Cell Factor (SCF) expressed by the somatic cells in the gonads (Rossi, 2013). During their migration, PGCs undergo several rounds of cell division. Once they colonize the gonadal ridge, they lose their ability to move and aggregate. PGCs continue to proliferate within the undifferentiated gonad and maintain their bipotentiality until E13, when cells within the male gonad become enclosed within the forming testis cords and enter mitotic arrest as T1 prospermatogonia in male sex development.



INTRO - Figure 8. <u>Primordial germ cells migratory pathway.</u>

PGCs are in the base of the allantois around the hindgut at E8.5 in mouse embryo. B) Their migration along the hindgut, dorsal mesentery and into the gonadal ridge at E10.5 (Fig from Wilhelm et al., 2007)

Genetic regulation of sex determination

Male sex determination

Sox transcription factors

The *Sox* (Sry related HMG box) gene family encodes more than 20 transcription factors involved in development of a variety of cell lineages. Sox proteins are characterized by the presence of a 79 amino acid motif that encodes an HMG (High mobility group) domain. SOX proteins can bind and bend to the minor groove in DNA by their HMG domain, recognizing a consensus DNA sequence: ATTGTT (Kamachi and Kondoh, 2013). They are classified into groups A-H, depending on conservation of the amino acid sequence of the HMG domain, strong homology and whole organization of protein domains within a group (Bowles et al., 2000; Schepers et al., 2003)(Table 1). Sox factors play widespread roles during development including in neural crest development, chondrogenesis, CNS neurogenesis, and testis development among others (Bowles et al., 2000; Chew and Gallo, 2009).

The *Sox* gene family was originally identified based on molecular conservation of the 79 amino acid HMG DNA binding domain found in the gene for the mammalian testis-determining factor SRY. Almost all *Sox* genes have at least 50% of amino acid similarity with Sry HMG box (Chew and Gallo, 2009).

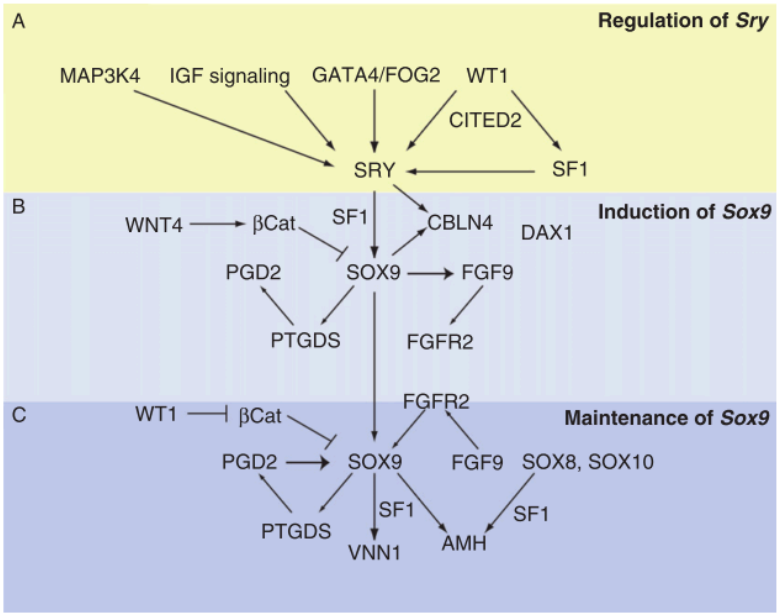
Table 1. Sox subgroups in *Mus musculus*. Different members of Sox transcription factors and their subgroups category in *Mus musculus*.

Sox subgropus (Mus musculus)	
SoxA	Sry
SoxB1	Sox1
	Sox2
	Sox3
SoxB2	Sox14
	Sox21
SoxC	Sox4
	Sox11
	Sox12
SoxD	Sox5
	Sox6
	Sox13
SoxE	Sox8
	Sox9
	Sox10
SoxF	Sox7
	Sox17
	Sox18
SoxG	Sox15
SoxH	Sox30

SRY

The first Sox family member identified was *Sry* (Sex determining region on Y chromosome). (Sinclair et al., 1990; Koopman et al., 1990). Located on the short arm of the Y chromosome immediately adjacent to the pseudoautosomal region, *SRY* is necessary and sufficient to induce testis development in mouse and humans. When expression levels of *Sry* reach a critical threshold, expression of its major downstream target *Sox9* is upregulated, which initiates a cascade of signaling events resulting in male sex determination. In mice, Sry binds to a testis specific enhancer sequence (Tesco) upstream of the *Sox9* gene and in synergy

with Nr5a1 regulates its expression (Sekido and Lovell-badge, 2008) (Fig 9). *Sry* expression is detected in mice at E10.5 (Koopman et al., 1990; Jeske et al., 1996), initially at the central region of the gonad, which then rapidly expands outwards among all the gonad at E11.5 before reducing in the same wave-like manner (from the center and towards the boundaries) (Bullejos and Koopman, 2001). *Sry* expression is completely diminished by E12.5 in mouse. In other species, such as the human the onset of *SRY* expression is similar to the mouse however the expression persists in Sertoli cells throughout testis development (Hanley et al., 2000; Pannetier et al., 2006). The spatiotemporal profile of *Sry* expression is critical, since a delay in the timing or a reduction in the levels of *Sry* expression can lead to a failure of testis determination and inappropriate development of ovotestis or ovaries (Wilhelm et al., 2009).



INTRO – Figure 9. Molecular pathway underlying Sertoli cell specification and differentiation.

There are a three different phases. A) the regulation of Sry expression. B) induction of Sox9 expression and C) the maintenance of Sox9 expression. (Figure from E. Wainwright & D. Wilhelm, 2010)

Apart from *Sox9*, several other genes have been identified as downstream targets of Sry, such as *Pod1, Cbln1* and *Ntf3* (Bhandari et al., 2012; Clement et al., 2011; Bradford et al., 2009a). How or if they are involved in testis determination under *Sry* regulation in mice and humans needs to be defined.

Several transcription factors are involved in the regulation of *Sry* expression in the future supporting cell precursors (See Fig 10) These include ***Wt1*** **(–KTS)** that binds to *Sry* and *Nr5a1* promoters *in vitro*, although, the cell population preceding testis differentiation is *Wt1* (-KTS) negative. Mice lacking the Wt1 (-KTS) isoform have small embryonic gonads and low but detectable expression of *Sox9* and *Amh*, implying that the male pathway is not completely blocked (Hammes et al., 2001). **Wt1 (+KTS)** isoform can not bind *in vitro* to *Sry* promoter (Hossain and Saunders, 2001), however, its abolition results in considerably reduced levels of *Sry* expression and sex reversal (Bor et al., 2006; Bradford et al., 2009b). Wt1 +KTS may also be involved in the post-transcriptional regulation of *Sry* (Hammes et al., 2001).

In ***Nr5a1***$^{-/-}$ mice, XX and XY gonads regress by E12.5, along with a complete absence of *Sry* expression suggesting that *Nr5a1* may be one of the upstream regulators of *Sry* (de Santa Barbara et al., 2001; Pilon et al., 2003). In humans, *NR5A1* haploinsufficiency results in sex reversal (Achermann et al., 2002).

Mouse XY embryos lacking ***Fog2*** (also known as *Zfpm2*) and the homozygous ***Gata4*** mutant: Gata4^{ki}, that abolishes the interaction between Gata4 and Fog2, exhibits reduced levels of *Sry* expression and sex reversal (Tevosian et al., 2002).

Sry expression in *Gata4*^{ki} gonads isolated at E11.5 was 25% of that detected in wild type gonads (Manuylov et al., 2008). This suggests that Gata4-Fog2 dimers are required for *Sry* regulation. Moreover, *Gata4* and *Fog2* seem to have separate and distinct roles during gonadal development, since the absence of *Fog2* leads to a partial block of testis differentiation, whereas *Gata4* seems to be required for testis cord formation and *Dmrt1* expression in early sex differentiation (Manuylov et al., 2011).

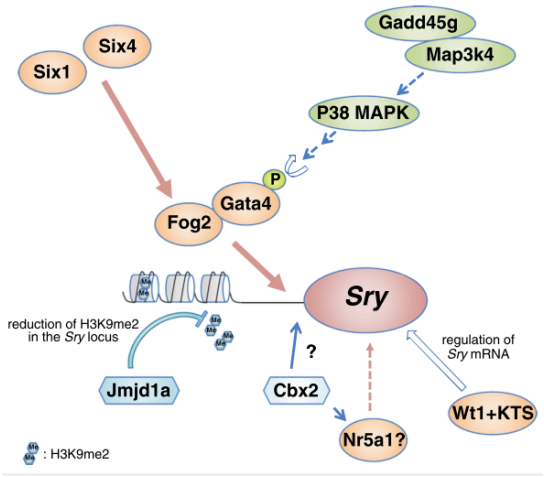
Gadd45g$^{-/-}$ mice present XY gonadal sex reversal due to loss of *Sry* expression (associated with reduction in phosphorylation of p38 MAPK and Gata4). Transient overexpression of *Map3k4* can rescue the gonadal defects suggesting that *Gadd45g* is involved in male sex determination (See section MAPK below).

Six1 *and* ***Six4,*** are proposed as positive regulators of *Sry* expression by upregulation of *Fog2* in the coleomic epithelium at the onset of *Sry* expression. *Six1*$^{-/-}$*; Six4*$^{-/-}$ XY mice gonads show reduced size due to disruption of *Nr5a1* expression and low levels of *Fog2* and thus *Sry* expression, accompanied by impaired testis differentiation. The forced expression of *Sry* transgene can rescue testis development, but not the precursor cell growth (Fujimoto et al., 2013).

Mice lacking the polycomb ***Cbx2*** gene show reduced levels of *Sry* expression, causing XY sex reversal in mice and humans (Katoh-Fukui et al., 2012), the impaired testis development can be rescued by transgenes *Sry* or *Sox9*. Cbx2 can bind specifically to *Nr5a1* promoter, regulating its expression (Katoh-Fukui et al.,

2005). However, the exact interaction between Cbx2 and *Sry* remains to be clearly understood.

Recent data suggests that *Sry* expression is also under epigenetic regulation. *Kdm3a* (*Jmjd1a*) deficient mice show varying degrees of XY sex reversal; at E13.5 these mice present ovotestes with both Sox9 and Foxl2 positive cells, indicating a failure of testis determining pathway. In wild type mice, *Kdm3a* mRNA levels are increased from E10.5 to E11.5 in the somatic cells of the gonads; *Kdm3a* deficient mice show increased levels of H3K9me2 in regulatory regions of *Sry* but not in *Sox9,* however, levels of both *Sry* and *Sox9* are reduced at E11.5 (Kuroki et al., 2013).



INTRO - Figure 10. *Sry* regulation.

Kdm3a/Jmjd1a upregulates the H3K9 demethylation of *Sry* locus, allowing the upregulation of *Sry* expression by different transcriptional factors. *Gata4* and *Fog2* are critical for *Sry* activation. *Fog2* is upregulated by Six1 and Six4 in the coleomic epithelium. *Gata4* is transiently activated by Gadd45g-Map3k4-p38 MAPK pathway. Phosphorylated GATA4 and FOG2 complex may bind to Sry promoter activating its expression. The WT1 (+KTS) may contribute to the post transcriptional regulation of *Sry* mRNA. CBX2 is required for *Sry* upregulation but still remains to be elucidated how this interaction is. Additionally CBX2 upregulates *Sf1* expression and SF1 is an upstream regulator of *Sry.* (Figure from S. Tanaka & R. Nishinakamura, Cell.Mol.Life.Sci 2014)

SOX9

Human *SOX9 is* located on chromosome 17 and encodes a protein of 509 amino acids. It is a transcription factor containing a HMG domain, which belongs to the SOX subgroup E (including SOX8 and SOX10). *Sox9* is considered a pro-testis gene in all mammals where it has been studied. In mice *Sox9* is expressed in the undifferentiated gonad of both sexes (E10.5), but becomes dramatically up regulated in pre-Sertoli cells immediately after the onset of *Sry* expression (E11.5) in a similar wave-like pattern to *Sry*. Unlike *Sry* in mice*, Sox9* expression is maintained in Sertoli cells after birth. SOX9 is not only required but sufficient for male sex development, because *Sox9* loss is associated with XY sex-reversal in mouse and human due to failure of Sertoli cell differentiation (Barrionuevo, et al., 2005; Chaboissier et al., 2004). Additionally, the forced expression of *Sox9* in XX mouse embryonic gonads results in testis development (Vidal et al., 2001; Bishop et al., 2000). In humans, the duplication of the *SOX9* gene has been reported to cause 46,XX testicular DSD (Huang et al., 1999; Cox et al., 2011).

In mice, a gonad-specific regulatory element of *Sox9* has been defined as a 3.2kb enhancer (TES) located 13kb upstream of *Sox9*. Further studies have delimited the TES element to a 1.4kb conserved core enhancer region (TESCO). Sry and Nr5a1 bind synergistically to TES and enhance *Sox9* expression. Since in the mouse *Sry* is only expressed for a short period, later on Sry binding is replaced by the Sox9 protein, thereby maintaining its own expression levels (See Fig 11)(Sekido and Lovell Badge, 2008). In humans, TES homologous sequence have been described 14kb upstream of *SOX9*, however no mutations or rearrangements in it have been reported in association with 46,XY DSD (Georg et al., 2010) suggesting the existence of additional testis specific enhancer regions

in humans. Recent reports have identified specific isolated regulatory regions for 46,XY and 46,XX DSD within a long distance regulatory region upstream of *SOX9* associated with DSD called *RevSex* (Benko et al., 2011; Cox et al., 2011; Vetro et al., 2011). 46,XY DSD (SRY positive) have been linked to a 32.5kb deleted region at 607.1-639.6kb upstream of *SOX9*, which was further narrowed to a 1.9kb Sertoli cell specific enhancer. 46,XX DSD (S*RY*-negative) duplications defined a 68kb region, 516-584kb upstream of *SOX9* (Kim et al., 2015). Furthermore, a recent collaboration in our lab has defined a regulatory element of ≈40kb region, that when duplicated is associated with 46,XX DSD (Hyon et al., 2015).

Sox9 is also under miRNAs regulation. In the developing mouse ovarian cells the expression of miR124 down-regulates *Sox9* expression (Real et al., 2013). Whereas in the developing testis, *Sox9* triggers the over-expression of miR202-5p and miR202-3p (Wainwright et al., 2013). The precise roles for miR202-5p/3p during testis determination remain to be understood.

Once *Sox9* expression reaches a critical threshold in fetal Sertoli cells, several positive regulatory loops are activated to maintain its expression, including its autoregulation via TESCO and other regulatory pathways including Fgf9 and PGD_2 (See Fig 11). Sox9 upregulates *Fgf9* expression and Fgf9 then, maintains *Sox9* expression, by establishing a feed forward loop via FGF receptor 2 (FGFR2) (Kim et al., 2007). Mutations in *Fgf9* or *Fgfr2* cause the reduction of *Sox9* expression levels and thus lead to XY sex reversal (Colvin et al., 2001; Kim et al., 2007). Moreover, a duplication of *FGF9* has been reported in a 46,XX male (Chiang et al., 2013) and a deletion including *FGFR2* in a 46,XY female with

ambiguous genitalia has been reported (Tannour-Louet et al., 2010), suggesting an involvement of both genes in human male gonad development.

Another amplification loop for *Sox9* expression is given by the direct Sox9 binding to the promoter region of *Ptgds*, which encodes for prostaglandin D synthase and mediates the production of **PGD_2** (Wilhelm et al., 2007). Autocrine and/or paracrine PGD_2 regulation stimulates Sox9 transcription and nuclear localization and thus reinforce the differentiation of Sertoli cells (Wilhelm et al., 2005; Malki et al., 2005). Targeted null mutations of *Ptgds* in mouse XY gonads, result in decreased and delayed expression, and aberrant cellular localization of *Sox9;* however no sex reversal was observed, suggesting that another signaling pathway can compensate the lack of *Ptgds* and rescue the levels of *Sox9* expression in Sertoli cells differentiation (Moniot et al., 2009).

Other direct targets of SOX9 include the anti-Müllerian hormone (*AMH*) gene, which encodes a TGF-β like molecule secreted by fetal Sertoli cells; AMH is responsible for the regression of the Müllerian ducts (see section on AMH). Cerebellin 4 precursor gene: *Cbln4* (Bradford et al., 2009a), which encodes a transmembranal protein; and Vanin-1 (*Vnn*) (Wilson et al., 2005) have been proposed as putative targets for Sox9, however their role during male gonadal development remains to be clarified.

SOX8

SOX8 belongs to the subgroup E of SOX family of transcription factors. *Sox8* is expressed in a similar temporal and spatial pattern to *Sox9* in the developing gonads. Its expression is up regulated in Sertoli cells around 12h

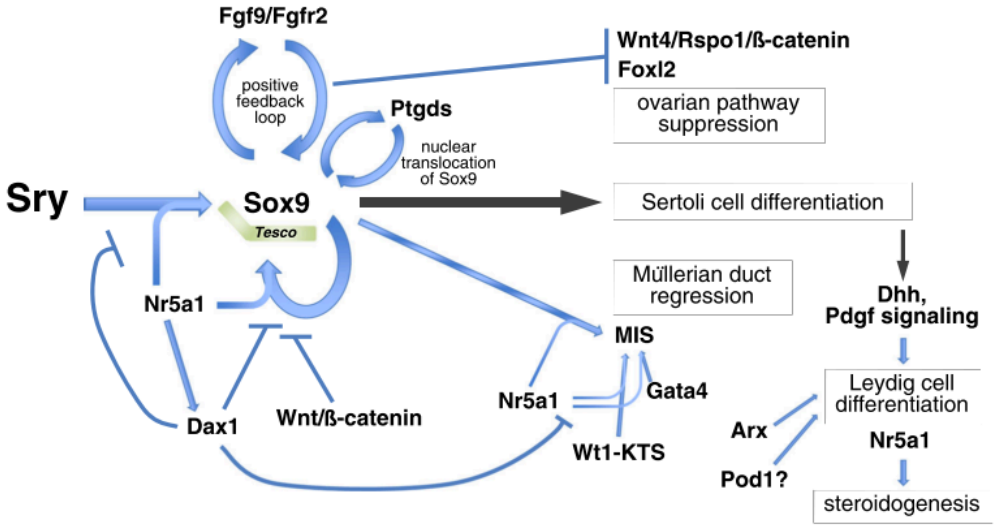
following *Sox9*, however it precedes *AMH* expression (Schepers et al., 2003). Like Sox9, Sox8 can bind to *AMH* promoter and interact in synergy with Nr5a1 in order to activate it *in vitro* (Schepers et al., 2003). *In vivo*, the redundant role of *Sox8* and *Sox9* has been confirmed by mouse double knock out studies, where Sox8 is able to reinforce Sox9 function during testis determination (Chaboissier et al., 2004).

Sox8 along with *Sox9* have an important role in maintenance of male fertility. *Sox8* mutant mice, which are initially fertile, develop progressive seminiferous tubule failure and infertility with age (Barrionuevo et al., 2009). Germ cells depend on Sertoli cells for their correct movement towards the seminiferous epithelium and $Sox8^{-/-}$ mice show a progressive degeneration of the seminiferous tubules through perturbed physical interactions between Sertoli cells and developing germ cells (O'Bryan et al., 2008).

SOX10

During mouse embryonic testis development, spatio-temporal expression of *Sox10* is similar to the other members of SOX-E subgroup (Cory et al., 2007). XX mice expressing *Sox10* under the promoter of *Wt1*, develop testis and a male physiology depending on the expression levels of *Sox10* (Polanco et al., 2010). *Sox10* is expressed in primordial gonads at low levels before sex is determined and expression becomes male specific during testis differentiation. Duplication of 22q including *SOX10* in human is associated with 46,XX DSD (Polanco et al., 2010).

The expression patterns of *Sox8* and *Sox10* during embryonic development are similar to the expression profile of *Sox9*. Both Sox8 and Sox10 have the *in vitro* ability to regulate known Sox9 targets (they both activate TESCO and *AMH* promoter in synergy with Nr5a1 (Sekido and Lovell-badge, 2008; Lin and Achermann, 2008)), suggesting a functional redundancy between the SOXE members during the development of the male gonad. Importantly, in XX *Sox9*^{cKO},*Rspo1*^{cKO} mice, testis development occurs in the absence of *Sry* and *Sox9*. These mice show upregulated expression of both *Sox8* and *Sox10*, suggesting that these *Sox* genes can induce male differentiation in the absence of a female pathway activator such as *Rspo1* (Lavery et al., 2012). The existence of a highly conserved HMG-box domain that mediates DNA binding among SOX proteins may explain the observed redundancy. SOX proteins acquire their cell specific regulatory functions by interacting with different protein partners. The functional redundancy between members of Sox family, regulating the same target genes, preserves the developmental processes against genetic variation between individuals (Kamachi and Kondoh, 2013).



INTRO - Figure 11. Gene regulatory network of embryonic male gonadal development.

In mice, the expression of Sry triggers the pre Sertoli cell differentiation. Sry directly activates Sox9 through the core element of Testis specific enhancer region of Sox9 (TESCO) in synergy with SF1. Sox9 afterwards maintain its own expression through TESCO together with SF1. Sox9 upregulates Fgf9 expression, and FGF9 in turn establishes the Sox9–FGF9 positive feedback loop through FGFR2, maintaining a high level of Sox9 expression. The Sox9–FGF9 positive feedback loop also suppress ovary-specific WNT4/Rsp1/b-catenin signaling activity. Sox9 expression also suppresses the expression of ovarian gene Foxl2. In addition, Sox9 upregulates Ptgds, which promotes nuclear translocation of Sox9 and facilitate Sertoli cell differentiation. In synergy SOX9 and SF1 promote expression of MIS/AMH and this activates the regression of the Müllerian ducts. AMH is regulated by SF1 in synergy with WT1 (-KTS) and also Gata4. Then, Sertoli cells express Dhh and Pdgf, which are involved in Leydig cells development Probably Arx and Pod1 are also involved) and SF1 regulates the steroidogenesis in Leydig cells. (Figure from S. Tanaka & R. Nishinakamura, Cell.Mol.Life.Sci 2014)

NR5A1

Steroidogenic factor 1 (*SF1/NR5A1/AD4BP)* is a member of the orphan nuclear receptor superfamily and it is expressed in the developing urogenital ridge, hypothalamus and the anterior pituitary gland (Luo et al., 1994; Morohashi et al., 1992). Human *NR5A1* encodes a 461 amino acid protein comprising of a DNA binding domain with two zinc fingers, an A box that mediates specific DNA binding, a flexible hinge region, a ligand-domain and two activation function domains (AF-1/2). Nr5a1 binds to DNA as a monomer, it is expressed in Sertoli and Leydig cells of the developing testis and it can be

detected in both, non steroidogenic lineage and steroidogenic cell lineages in the testis after sex differentiation, as well as in multiple cell lineages in the fetal, postnatal, prepubertal, and mature ovary (Hanley et al., 1999; Ikeda et al., 1994). In mammalian testis determination and differentiation, NR5A1 is a positive regulator of sex determining genes including *SRY* (de Santa Barbara et al., 2001), *SOX9* (via TESCO, Nr5a1 synergizes sequentially with Sry and Sox9) (Sekido and Lovell-Badge, 2008; Knower et al., 2011) and *AMH* (Also in synergy with Sox9)(Sekido and Lovell-badge, 2008; Lin and Achermann, 2008) (See Fig 11). NR5A1 also regulates the expression of several genes involved in endocrine development and function, as *STAR* and several CYPs (Honda et al., 1993; Morohashi et al., 1993; Val et al., 2003). Mice lacking *Nr5a1* show a complete failure of gonadal and adrenal development, and a female phenotype at birth but die at postnatal day 8 due adrenal failure (Luo et al., 1994). In human, mutations in *NR5A1* are known to cause 46,XY DSD with a wide range of phenotypes (from hypospadias to complete gonadal dysgenesis) with or without adrenal failure (Achermann et al., 2002; Allali et al., 2011; Ferraz-de-Souza et al., 2011). Mutations in *NR5A1* are also associated with 46,XX DSD, 46,XX primary ovarian insufficiency, and severe oligozoospermia and azoospermia in 46,XY individuals (Bashamboo et al., 2010a). Taken together the data indicates that *NR5A1* plays diverse roles in the formation of the undifferentiated gonads and in subsequent gonadal sex determination, steroidogenesis and germ cell maintenance.

GATA Family.

GATA proteins belong to a class of evolutionary conserved lineage-limited zinc finger transcription factors that participate in cell fate determination, proliferation and maturation (Molkentin, 2000; Morceau et al., 2004). GATA members have two conserved type IV zinc fingers domains C-X_2-C-X_{17}-C-X_2-C (Lowry and Atchley, 2000). These factors recognize a consensus sequence of DNA to which they specifically bind: (A/T)GATA(A/T) (Lowry and Atchley, 2000). Six members compose the GATA family: GATA 1-6. GATA 1/2/3 are expressed in hematopoietic lineages while GATA 4/5/6 are expressed mostly in mesoderm and endoderm derived lineages including heart, gastrointestinal tract and gonads. In the gonads there is expression of *Gata1* (Yomogida et al., 1994; Ito et al., 1993; Viger et al., 1998), *Gata2* (Siggers et al., 2002), *Gata4* (Viger et al., 1998; Ketola et al., 1999a; McCoard et al., 2001) and *Gata6* (Heikinheimo et al., 1997; Ketola et al., 1999b). With the exception of *Gata2*, they are all expressed in the somatic cells. GATA proteins regulate multiple downstream targets to achieve specificity and transcription control in the different tissues where they are expressed (Merika and Orkin, 1993; Gordon et al., 1997; Chlon and Crispino, 2012).

GATA4

Gata4 is expressed prior to and during sex determination (Ketola et al., 2000), and is seen until E11.5 in the somatic cells of both XX and XY gonads. At E13.5, *Gata4* expression becomes dimorphic; it is increased in the Sertoli cells in XY gonads, while in the interstitial cells of the XY gonads and in XX gonads the

expression decrease (but remains detectable). GATA4 is important for Müllerian duct regression through the regulation of *Amh* (Anti Müllerian hormone) gene. *In vitro* experiments show that GATA4 bind to consensus activation sites in *Amh* promoter resulting in its activation (Tremblay and Viger, 1999; Watanabe et al., 2000). GATA4 can cooperate with several factors, including NR5A1, to regulate expression of key genes involved in sex determination and differentiation: *SRY, SOX9, AMH, CYP19A1, STAR, INHA* (aytouni et al., 2011). *Gata4* expression is maintained through adulthood in the testis and increases in the granulosa cells of adult ovaries (Viger et al., 1998).

Gata4 null mice die between E7 and E9.5 due to severe abnormalities of heart tube and ventral morphogenesis prior to sex determination (Molkentin et al., 1997). Homozygous $Gata4^{ki}$ mice, carrying a mutation that disrupts interaction between Gata4 and Fog2, develop genital ridges; nonetheless further differentiation of the testis is obstructed. In the $Gata4^{ki-/-}$ mice, *Sry* expression at E11.5 was 25% of the expression in wild type gonads (Manuylov et al., 2008), confirming a role for Gata4 upstream of *Sry* expression during testis development. The downstream gene targets of Sry required for Sertoli (*Sox9, Amh, Dhh*) and Leydig cells (*Cyp11a1, Hsd3b1, Cyp17a1*) formation and function were not expressed in *Gata4* mutants (Tevosian et al., 2002). Gata4-Fog2 dimers are required for *Sry* regulation, and for ovary development (Manuylov et al., 2008), however both proteins seem to have independent and different roles through gonad development (Manuylov et al., 2011). Gata4 expression seems to be required for testis differentiation, testis cord formation and the transcriptional regulation of *Dmrt1*.

In humans, mutations in *GATA4* are associated with congenital heart defects. In addition, previously in our lab we have described a heterozygous missense mutation in *GATA4* associated to a familial case of 46,XY DSD. Mutant *GATA4* had compromised activation of *AMH* promoter, independently and synergistically with NR5A1. Moreover mutant GATA4 had lost its ability to interact with its cofactor FOG2, this interaction is essential for gonad formation, thus confirming a role for GATA4 in human testis development (Lourenço et al., 2011b).

GATA6

GATA6 is located on human chromosome 18. *Gata6* is expressed in the developing testis and ovary, in both somatic and germ cells (Ketola et al., 1999a; Lavoie et al., 2004). The timing at which *Gata6* starts to be expressed in the gonads is not precisely known due to lower levels of its expression compared to that of *Gata4*, but it can be detected as early as E13.5 in fetal mouse testis (Robert et al., 2002).

Promoter studies have identified several putative target genes for Gata6, including the steroid biosynthetic genes *Cyp11a1* (Jimenez et al., 2003), *Cyp17a1* (Jimenez et al., 2003; Nakamura et al., 2007), *Hsd3b2* (Martin et al., 2005), *CYB5* (Huang et al., 2006), and *Sult2a1* (Jimenez et al., 2003; Saner et al., 2005; Nakamura et al., 2009).

In human fetal gonads, *GATA6* expression is detected in a subset of fetal Sertoli and interstitial Leydig cells (Ketola et al., 2003). *GATA6* is also expressed in human fetal and adult ovary. During the ovarian development and folliculogenesis, the depletion of *GATA4* is partially compensated by *GATA6*

(Bennett et al., 2012). Moreover, it has been shown that GATA6 and GATA4 have a cooperative role in early ovarian development and function since ablation of both genes results in strong down-regulation of genes involved in ovarian developmental pathway (such as *FST)* and down-regulation of granulosa cell markers (such as FOXL2) resulting in impairment of granulosa cell proliferation and arrested follicular development (Padua et al., 2014).

FOG2

ZPFM2 or *FOG2* is a member of the Friend of GATA (FOG) family of zinc finger proteins. FOG members are known for their interaction with the N-terminal zinc fingers of different GATA proteins to regulate gene expression (Robert et al., 2002; Cantor and Orkin, 2005). Human *FOG2* is located on chromosome 8; it encodes a protein containing eight zinc fingers. *FOG2* is co-expressed with *GATA4* in the developing heart, brain and gonads (Ketola et al., 1999a; Robert et al., 2002; Anttonen et al., 2003; Laitinen et al., 2000) and acts as a cofactor, enhancing or repressing GATA transcriptional activity depending on the cellular context.

Fog2 null mice die from complications associated with heart defects similar to the *Gata4* null mice, however, these mice survive until E14.5, allowing the study of developing gonads (Tevosian et al., 2000, 2002). Gonadal loss of *Fog2* leads to an early arrest of testis differentiation and consequently male to female sex reversal (Manuylov et al., 2011). Human mutations of *FOG2* have been associated with congenital heart disease, and Tetralogy of Fallot. A translocation of 8q23.1-8q24.1 including *FOG2* has been reported in a case of a male gonadal

dysgenesis (Finelli et al., 2007). Moreover, using exome sequencing technologies we have recently described different missense point mutations in *FOG2* associated with 46,XY DSD; where *FOG2* mutants show an impaired protein interaction with GATA4, thus confirming a role for *FOG2* during human testis development (Bashamboo et al., 2014).

AMH

Anti Müllerian hormone (*Amh*) or Müllerian inhibition substance (*Mis*) is a member of the transforming growth factor-β (TGF-β) family (Josso et al., 2006; Teixeira et al., 2001; Münsterberg and Lovell-Badge, 1991), which functions by binding and assembling two related serine/threonine kinase receptors. Human *AMH* is located on chromosome 19 and encodes a 560 amino acid glycoprotein. AMH signaling is essential and sufficient for the regression of the Müllerian ducts during testis development (Behringer et al., 1994; Belville et al., 1999; Tsuji et al., 1992) (Fig 1 and Fig 11). In mice, *Amh* is expressed only in Sertoli cells; it can be detected in pre-Sertoli cells around E12.5 and continues throughout male fetal development until puberty when Sertoli cells mature and secrete lower levels of *Amh*. In females, *Amh* expression can be detected in granulosa cells around birth, stays low levels throughout reproductive life and becomes nonexistent after menopause (Rey et al., 2003).

Nr5a1 interacts with several genetic factors to positive regulate *AMH* expression in male embryos including Sox9, Sox8, Gata4 and Wt1 (Nachtigal et al., 1998; Arango et al., 1999; Watanabe et al., 2000; Schepers et al., 2003;

Chaboissier et al., 2004). Dax1 inhibits *Amh* transcription *in vitro* (Nachtigal et al., 1998; Tremblay and Viger, 2001). FOG2 inhibits *AMH* during female ovarian development and male puberty (Anttonen et al., 2003; Tremblay et al., 2001).

In XX mice, the ectopic expression of *Amh* causes the regression of the Müllerian ducts, which give rise to the female reproductive tract, including oviducts, uterus and upper vagina (Behringer et al., 1994). Mutations in mouse and human *AMH* or its receptor are related to the Persistent Müllerian Duct Syndrome (Josso et al., 2005) and is often associated with cryptorchidism and infertility.

DAX1

The gene *DAX1* (DSS-AHC critical region on the chromosome X), also known as *NR0B1* (Nuclear receptor subfamily 0, group B, gene 1) is located at chromosome Xp21.2 (anaria et al., 1994). Dax1 has a transcriptional silencing domain by which it can interact to co-repress its targets. *Dax1* is expressed in many cell types throughout the hypothalamic pituitary gonadal axis (Ikeda et al., 1996). In the mouse gonad, *Dax1* is expressed in the primordial ridge of both sexes after *Nr5a1* is expressed, it increases in Sertoli cells of the XY gonads at E12.5, decreases rapidly after, increases again in the interstitial cells around E13.5-E14.5, and decreases again (Ikeda et al., 2001). The function of DAX1 is controversial; it is considered both a pro-testis (Meeks et al., 2003) and an anti-testis gene (Swain et al., 1998). *In vitro* studies have shown that β-catenin in synergy with Nr5a1 can activate *Dax1* transcription in XX gonads (Mizusaki et al., 2003). Although several studies suggest Dax1 antagonizes Nr5a1 during gonadogenesis and steroidogenesis (Tremblay and Viger, 2001), others suggest

Dax1 and Nr5a1 act together to synergistically activate testis specific genes (Park et al., 2005). Transgenic mice overexpressing *Dax1* show male to female sex reversal suggest an anti-testis role (Jiménez and Burgos, 1998). However, there are reports showing its role as a pro-testis gene: hemizygous *Dax1* XY mice have abnormal testis, whereas *Dax1* null XX mice are normal and fertile (Yu et al., 1998). In humans, mutations in *DAX1* result in an X-linked syndrome known as Congenital Adrenal Hypoplasia (CAH) which causes adrenal insufficiency and hypogonadotropic hypogonadism in XY individuals (anaria et al., 1994). Duplications of X chromosome regions containing *DAX1* lead to dosage sensitive sex reversal (Bardoni et al., 1994). Very small or high levels of *Dax1* expression have the same effect on testis development indicating it must be regulated very strictly in quantities and time during gonadal development (Ludbrook and Harley, 2004). However, the exact mechanisms by which *Dax1* contribute to testis development remains to be clearly understood.

DHH

Desert Hedgehog (*DHH*) is located on human chromosome 12 and encodes a 396 amino acid protein of the Hedgehog family (Tate et al., 2000). *Dhh* is the only member of the hedgehog signaling family that is expressed in the developing mouse gonads. In fetal testis, *Dhh* expression can be detected in pre-Sertoli cells at E11.5, after the onset of *Sry*, being thereafter present in Sertoli cells. In contrast *Dhh* is not detected in fetal ovaries (Yao et al., 2002; Bitgood et al., 1996). Mice $Dhh^{-/-}$ testes from adult animals lacked adult-type Leydig cells and displayed numerous undifferentiated fibroblast cells in the interstitium. The

basal lamina, normally present between the myoid cells and Sertoli cells, was absent (Clark et al., 2000; Pierucci-Alves et al., 2001). Dhh/Ptch1 is required for Leydig cell and PMC differentiation pathways by activating *Nr5a1* and *Cyp11a1* expression in the interstitial precursor cells (Yao et al., 2002). Human *DHH* mutations have been reported in several cases of 46,XY gonadal dysgenesis, usually in association with formation of minifascicles within the endoneurium of the sural nerve (Umehara et al., 2000; Canto et al., 2004, 2005; Castro et al., 2013; Das et al.; Paliwal et al., 2011; Werner et al., 2015).

HHAT

The hedgehog acyl-transferase (HHAT) protein, a member of the MBOAT family of membrane-bound acyltransferases, catalyzes amino-terminal palmitoylation of Hh proteins. Palmitoylation is crucial for biological activity and plays a major role in guiding Hh proteins to specific membrane domains. Recently, a homozygous G287V missense mutation in the MBOAT domain was reported in a single case of 46,XY gonadal dysgenesis and chondrodysplasia (Callier et al., 2014). The mutation disrupted the ability of HHAT to palmitoylate Hh proteins including DHH and SHH suggesting that it is pathogenic. *HHAT* is expressed in the somatic cells of both XX and XY gonads at the time of sex determination. The absence of Hhat in the XY gonad did not affect testis-determination but from E12.5 to E15.5 there was a drastic reduction in testis size. Sertoli cell specification appeared to be normal but the development of fetal Leydig cells as well as the proper formation of testis cords was severely impaired, leading ultimately to testicular dysgenesis. These data suggest that HHAT is not involved

in primary testis-determination but is required for the initiation of Leydig cell formation. It is unlikely that mutations in *HHAT* are a common cause of DSD. The phenotype of the girl carrying homozygous mutation is a complex, consisting of a rare combination of gonadal dysgenesis and chondrodysplasia. Moreover, a *de novo* dominant mutation the MBOAT domain of HHAT was reported in association with intellectual disability and apparently normal testis development (Agha et al., 2014) suggesting that the *HHAT* genotype-phenotype relationship is likely to be complex.

Mitogen-Activated Protein Kinase Pathway

Two members of the mitogen activated protein kinase (MAPK) signaling pathway family, *MAP3K1* and *Map3k4* are part of the growing list of genes involved in testis development. *MAP3K1* mutations were first identified in independent cases of 46,XY DSD (Pearlman et al., 2010). However, in mice, *Map3k1* does not have a significant effect during testis development, *Map3k1* null mice do not show gonadal abnormalities (Warr et al., 2011). Further, an ENU screen identified a *Map3k4* mutation that caused male-to-female sex reversal on C57BL/6J background mice. XY *Map3k4* null gonads show reduced levels of *Sry* and *Sox9* expression (Bogani et al., 2009). It has been proposed that Gadd45g (growth arrest and DNA damage inducible 45 γ) activates Map3k4, which stimulates p38 MAP kinases (MAPK11/MAPK14) and subsequently leads to phosphorylation and activation of Gata4, which activates different testis genes, including *Sry* and *Sox9* (See Sections on Regulation of Sry and GATA4) (Warr et al., 2012). Given the current evidence, it seems that the role of MAPK signaling

pathway may have diverged between mice and human, nevertheless in both species MAPKs plays an important roles during testis development.

ATRX

The α-thalassemia/mental retardation syndrome X linked gene (*ATRX*) is a member of the SNF-2 like helicase super family subgroup, that contains genes involved in DNA recombination, repair and regulation of transcription (Bérubé, 2011). Human *ATRX* is located in chromosome X and encodes multiple transcripts where the longest encodes for a 2492 amino acid protein. This nuclear protein has two highly conserved domains, a N-terminal ADD (ATRX-DMNT3L-DMNTL3A) domain comprising a GATA-like zinc finger and a plant homeodomain generally present in chromatin-associated proteins and a C-terminal domain containing a switch/sucrose non-fermenting like ATPase domain that has nucleosome remodeling activity (Bérubé, 2011). ATRX provides a link between chromatin remodeling, DNA methylation and gene expression in developmental processes.

However, during gonadal development the exact role of ATRX remains to be clearly understood. *Atrx* null mice present reduced size of testis due to apoptosis of the proliferating fetal Sertoli cells (Bagheri-Fam et al., 2011). In humans, *ATRX* mutations cases ATR-X syndrome, a condition characterized by α-thalassaemia, severe psychomotor retardation, microcephaly, short stature, characteristic facial features, cardiac, skeletal and urogenital abnormalities (Ion et al., 1996). Almost all known mutations linked to ATR-X syndrome fall within the ADD and plant homeodomains. 80% of XY patients have urogenital anomalies with a wide

spectrum of phenotypes (from complete gonadal dysgenesis to relatively mild microspadias or micropenis). Moreover, within families carrying the same *ATRX* mutation there is phenotypic variation. ATRX regulates gene expression by binding to G-rich tandem repeat sequences. Variation in the quantities of tandem repeats are associated with differences in the level of gene expression (Law et al., 2010) which could be linked to the variation in the phenotypes observed in patients with same *ATRX* mutations.

DMRT1

DMRT1 (dsx- and mab3 related transcription factor 1) is a member of the DM domain transcription factor family, which was named after the homology of the DNA binding domain between *Drosophila dsx* (doublesex) and the *C.elegans mab-3* (male abnormal) gene. Human *DMRT1* is located on chromosome 9p and encodes a highly conserved 373 amino acid protein. DM transcription factors bind to consensus palindromic DNA sequence of 7bp via a conserved zinc finger like DNA binding domain, called DM domain (Raymond et al., 1998; Murphy et al., 2007, 2010). DM domain genes encode transcription factors involved in sex differentiation in different organisms; for example, Dm containing genes are the sex-determining switch in chicken (Smith et al., 2009) and medaka fish (Matsuda et al., 2002).

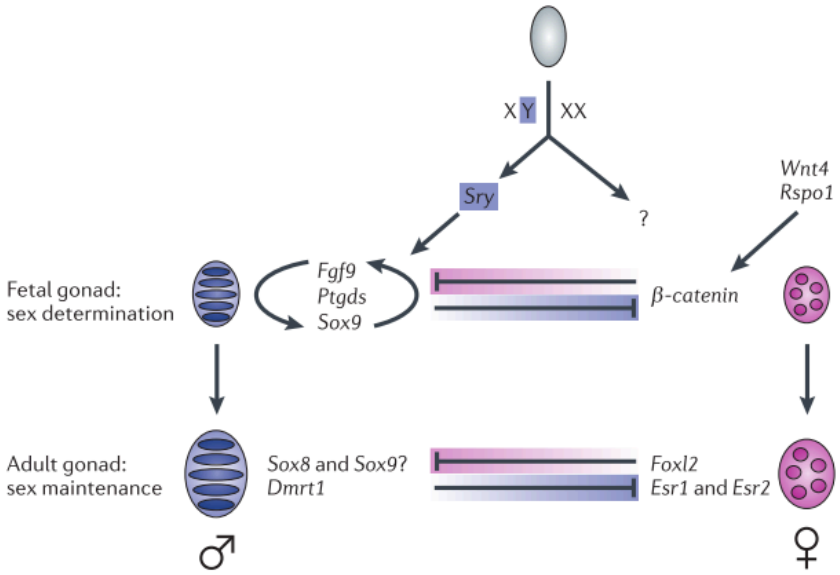
In mice, *Dmrt1* is expressed in the developing gonad in both XX and XY mice; around E12.5-E13.5 its levels decline in ovary and increase in the testis. The high levels of expression are thus male specific and maintained until adulthood in Sertoli cells and pre-meiotic germ cells (Raymond et al., 2000; De Grandi et al.,

2000). *Dmrt1* transcription in Sertoli cells is regulated by Gata4; distinct proximal promoters mediate the expression in germ and somatic cells (Lei and Heckert, 2004a; Lei et al., 2009; Manuylov et al., 2011). In humans, *DMRT1* expression has been detected around the sixth week of gestation in the undifferentiated XY gonadal primordium, at the seventh week, *DMRT1* expression is limited to the seminiferous tubules. *DMRT1* has not been detected at any stage of the development of female gonads (Moniot et al., 2000). *Dmrt1* null mice do not present anomalies during sex-determination and embryonic gonadal development, suggesting this gene may not have a determinant role in the early stages of gonadal development in mammals. However, they show dramatic defects in testicular differentiation after birth. Sertoli cells fail to complete their differentiation and proliferate excessively causing a high degree of disorganization of the testis (Raymond et al., 2000). In humans, deletions in chromosome 9p, comprising *DMRT1*, are associated with monosomy 9p syndrome, which is characterized by mental retardation and somatic anomalies. Some of these patients present with 46,XY DSD with a wide range of gonadal phenotypes. (Ottolenghi and McElreavey, 2000; Raymond et al., 1999), whereas 46,XX cases with 9p deletions do not show anomalies in the ovarian development and function (Vialard et al., 2002).

Dmrt1 is also expressed in germ cells during fetal gonad development and after birth, where it has a central role in the shift of mitosis-meiosis or cell cycle arrest by regulating the expression of the meiosis inducer gene *Stra8*. In female, *Dmrt1*$^{-/-}$ germ cells have a dramatic reduction of *Stra8* and germ cells enter abnormal meiotic prophase, however they still produce around half of the normal quantities of oocytes and are fertile (Krentz et al., 2011). After birth, Dmrt1 has

different functions in male germ cells. It reinitiates mitosis versus meiosis by suppressing retinoic acid signaling and by direct transcriptional repression of *Stra8*, so that spermatogonial differentiation and proliferation can be completed before meiosis (Mark et al., 2008; Anderson et al., 2008). Thus, Dmrt1 has opposite effects on regulation of *Stra8* expression in germ cells of both sexes.

The function of DMRT1 during mammalian testis development remains to be completely understood. There is evidence indicating an anti-ovarian role for Dmrt1 in adult mice. Loss of function of *Dmrt1* in mice, in adult testis, results in the transdifferentiation of Sertoli and Leydig cells towards granulosa and theca cells, including the activation of *Foxl2*, *Cyp19a1*/aromatase, the production of estrogen and the feminization of germ cells (Matson et al., 2011). Moreover, CHIP experiments of 4-week-old testes have shown that Dmrt1 is able to bind directly to regulatory regions of both testis (as *Sox8, Sox9, Ptgdr*) and ovary stimulating genes (*Foxl2, Esr, Wnt4* and *Rspo1*), suggesting the role of *Dmrt1* as a pro-testis and anti-ovary gene. Some of these genes are involved in fetal sex determination and also participate in the maintenance of the gonad **(**See Fig 12**)** (Matson and arkower, 2012) (See Section Battle of the Sexes).



INTRO - Figure 12. <u>Model for sex determination and maintenance</u>.

In mammals, sex is decided in the gonad during fetal development by the presence or absence of Sry. In its presence, it activates Sox9 and determines sex fate towards male development. Sex determination can be viewed as a power struggle in the fetal gonad between a male regulatory gene network centered on Sox9 and a female network involving Wnt4, Rspo1 and their downstream effector β -catenin. SOX9 expression in Sertoli cells is necessary and sufficient for testis determination and is reinforced by positive feedback via Fgf9 and Ptgds. If SRY is not present in the supporting cell lineage during a brief crucial period, sustained SOX9 expression is not established, and the female regulatory network predominates and drives ovarian development. Postnatally, a long time after sex is determined, a distinct regulatory network functions in each sex to maintain sex. In males, this network involves Dmrt1, and in females it requires Foxl2. Loss of either of these genes can cause even fully differentiated Sertoli or granulosa cells to transdifferentiate into the other cell type, triggering extensive reprogramming of the gonad towards that of the opposite sex. (Figure from Matson and Zarkower, 2012.)

Ovarian determination

We know much less details about the regulation of ovary development than testis development. Until recently, the molecular pathway involved in ovarian development was considered to be the "passive" pathway. However, in recent years, with emerging technologies and functional genetic analyses, numerous ovary-specific and somatic cell-derived genes have been identified. It has become clear that ovarian pathway is not passive; more over it has become evident that components of both male and female pathways antagonize each other to ensure proper development (Fig 12). Genes participating in ovarian development can be classified into two categories: intracellular factors such as FOXL2, and β-catenin and extracellular factors: R-spondin1 and WNT4.

Foxl2

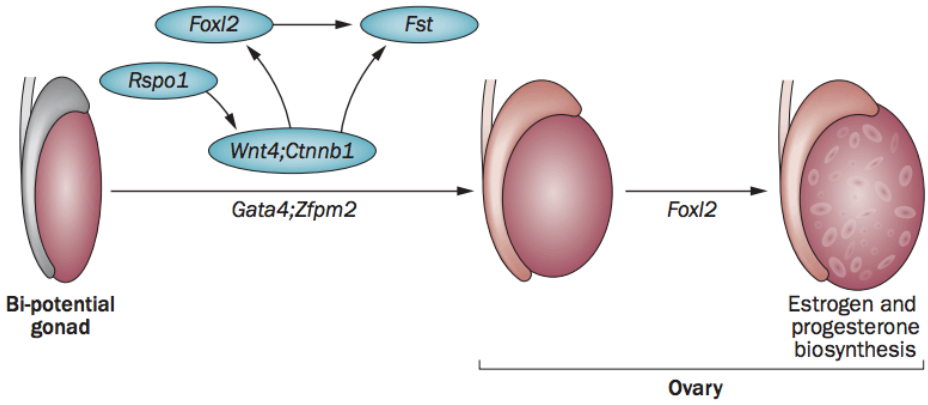
Forkhead box L2 (*Foxl2*) is a member of the forkhead box gene family, which encodes an evolutionarily conserved group of transcription factors. Human *FOXL2* is located on chromosome 3 and encodes for a 376 aa protein. Forkhead box head family is characterized by a 111 amino acid DNA binding domain known as forkhead (Carlsson and Mahlapuu, 2002). In the developing mouse female gonad, *Foxl2* is one of the first genes to be upregulated, it can be found from E12.5 in mesenchymal granulosa cells and later in granulosa cells. Ablation of *Foxl2* expression in XX mice and overexpression of *Foxl2* in XY mice results in gonad anomalies, but not sex reversal (Ottolenghi et al., 2005). However, in goats, mutations in *FOXL2* or its regulatory sequence results in goat

polled intersex syndrome (PIS) with XX male (Parma et al., 1999; Boulanger et al., 2014). Foxl2 suppresses genes involved in terminal granulosa cells differentiation, such as *Star* (Pollack 1997) and *Cyp19a1* (Baron 2004); it is thought to prevent the premature depletion of ovarian follicles by blocking the differentiation/proliferation of granulosa cells. Moreover, *Foxl2* has an essential role in postnatal maintenance and development in mice ovary. The targeted loss of *Foxl2* in adult ovary results in ovary-to-testis trans-differentiation, where male specific genes such as *Sox9, Dmrt1* and *Dhh* are upregulated, leading to a reprogramming of granulosa cells into Sertoli-like cells (Uhlenhaut et al., 2009). Foxl2 and estrogen receptors (Esr) α/β cooperate to maintain the granulosa cell lineage in adult mice (Uhlenhaut et al., 2009). XX mice lacking *Esr* α and *Esr* β or the aromatase enzyme that produce estrogen, show the sex reversal phenotype similar to that of *Foxl2*$^{-/-}$ (Couse et al., 1999; Britt et al., 2001; Uhlenhaut et al., 2009). Foxl2 and Esrα can synergistically repress *Sox9* expression *in vitro*, via Tesco and attenuate Nr5a1, Sry+Nr5a1 and Sox9+Nr5a1 activation. Moreover, Foxl2 negatively regulates *Nr5a1* expression by antagonizing Wt1 (-KTS isoform) during early ovarian development in mice (Takasawa et al., 2014).

In humans, *FOXL2* mutations in the human germline lead to the autosomal dominant blepharophimosis/ptosis/epicanthus inversus syndrome (BPES) which is characterized by eyelid abnormalities and often associated with premature ovarian failure in females (Crisponi et al., 2001), suggesting a role for FOXL2 in follicle maintenance during adult life (Harris et al., 2002). Recent work has shown that *FOXL2* mutations impairing the transcriptional activity (more specifically the ability to bind DNA) are responsible for POI in BPES (Dipietromaria et al., 2009; Todeschini et al., 2011). Indeed, a specific mutation

in *FOXL2* has ben identified in more than 95% of adult granulosa cell tumors, confirming the association of *FOXL2* and granulosa cell fate and function (Shah et al., 2009; Jamieson and Fuller, 2012).

There is accumulating evidence indicating the essential role of FOXL2 to prevent *Sox9* expression, thereby preventing Sertoli cell differentiation and allowing the maintenance of granulosa cell fate throughout life (See Fig 13).



INTRO - Figure 13. Molecular pathway required for ovary differentiation and development.

Genes as *Rspo1, Wnt4* and *Foxl2* are expressed during ovarian development. *Rspo1* mutant ovaries show reduced levels of *Wnt4* expression, suggesting it acts upstream of Wnt4, there is also evidence suggesting the synergistic action of bothWnt4 and Rspo1 to upregulate Ctnnb1 (β -catenin). Together Foxl2, Rspo1 and Wnt4 activate *Fst* expression. In addition, Gata4-Fog2 complex has been implicated in embryonic ovary development. In adult ovary, Foxl2 is essential for cell fate maintenance and follicle maturation. (Figure from Eggers et al., 2014)

Wnt4

Wingless-type MMTV Integration Site Family, Member 4 *(WNT4)* is located on human chromosome 1, and encodes a 351 amino acid protein. *Wnt4* belongs to a family of secreted lipid-modified glycoproteins that are evolutionarily conserved with essential roles in different developmental processes including proliferation, differentiation and cell migration among others (Chien et al., 2009; Miller, 2002). *Wnt4* encodes for a protein that act as a suppressor of male differentiation, it is expressed in the bipotential gonads from E9.5 and onwards. Around E11.5, after the onset of *Sry* expression, it becomes dimorphic, its expression is downregulated in XY gonads and maintained in XX gonads, becoming female specific (See Fig 14)(Vainio et al., 1999). Loss of *Wnt4* in XX mice, results in partial sex reversal, suggesting the positive regulatory role of *Wnt4* in ovary differentiation. *Wnt4*$^{-/-}$ female mice show masculinization of the gonads, including the formation of the coleomic blood vessel and the inward migration of adrenal steroidogenic cells that produce androgens. Moreover, these *Wnt4* null mice are unable to maintain germ cells. On the other hand, overexpression of *Wnt4* in XY mice results in disruption of testicular vasculature and arrest of Leydig cell development with inhibition of testosterone production, however there is no sex reversal (Jordan et al., 2003). In humans, homozygous mutations in *WNT4* are associated with 46,XX testicular DSD (Mandel et al., 2008).

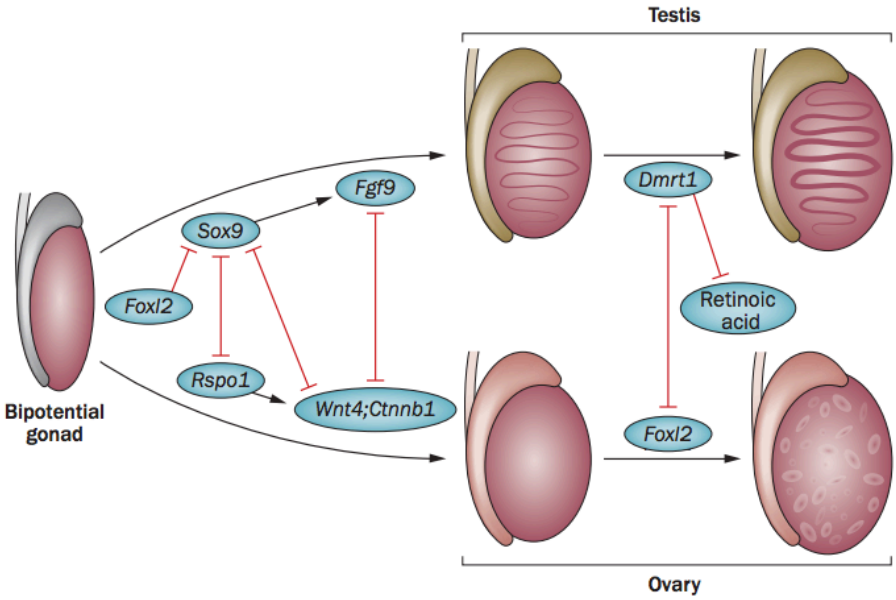
Rspo1

R-spondin family 1 (*RSPO1*) is located on human chromosome 1. RSPO1 belongs to the family of secreted furin-like domain containing proteins that activates *Wnt4/β-catenin* signaling and has pleiotropic functions in development. Rspo1 stimulates the expression of *Wnt4*, and synergize with it to activate β-catenin, the intracellular activator of WNT pathway, this regulates the expression of different genes, among them many ovarian components, for example *Wnt4* and *Fst* and thus enabling the female pathway to proceed (See Fig 13) (Yao et al., 2004; Chassot et al., 2008b; Liu et al., 2010; Maatouk et al., 2008). LOF mutations of *Rspo1* in mice result in disruption of ovarian development but not complete sex reversal. XX mice with gonads lacking *Rspo1* and *Wnt4* initially develop as ovaries but later after birth time, they acquire testis features as Sertoli cells and testis cords (Chassot et al., 2008a; Tomizuka et al., 2008). β-catenin expression in mouse XY gonads causes male to female sex reversal (Maatouk et al., 2008). In human, *RSPO1* null homozygous mutations cause 46,XX testicular DSD (Parma et al., 2006). Interestingly, a duplication of chromosome 1p, including *WNT4* and *RSPO1* had been reported in a 46,XY DSD patient with female phenotype (Jordan et al., 2001).

Rspo1 and *Wnt4* pathway function independently of *Foxl2* pathway. *Foxl2* deletions, do not affect *Rspo1/Wnt4* expression and vice-versa (Ottolenghi et al., 2007; Chassot et al., 2008a), however both pathways converge to suppress *Sox9* expression.

All the available information point to a model where FOXL2, WNT4 and RSPO1 act to antagonize *SOX9* expression, allowing ovarian development to occur and

be preserved, and at the same time repress the male testis differentiation pathway (see **Fig 14**).



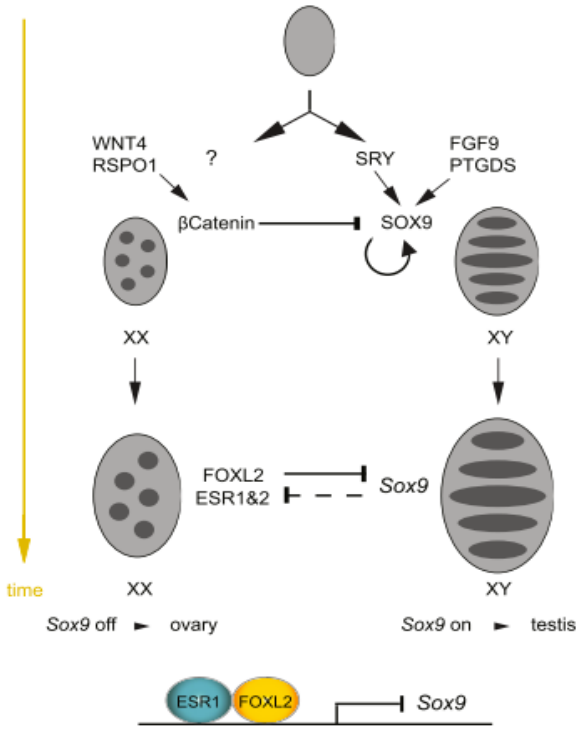
INTRO - Figure 14. Molecular interaction pathway between testis and ovary during embryonic development and adulthood.

Genes and proteins interact between the two pathways in a mutually antagonistic manner to ensure the maintenance of the cell fate and repression of the other pathway. During embryonic development Sox9 and Fgf9 inhibit the ovarian pathway by suppressing Rspo1 and Wnt4-Ctnnb1 expression and in the ovary Foxl2, Rspo1, Wnt4 and Ctnnb1 suppress the male pathway by inhibition of Sox9 and Fgf9. During adult life, testis and ovaries need to be actively maintained by suppression of the pathway of the opposite sex. In adult testis, Dmrt1 has central role in the maintenance of Sertoli and Leydig cells by suppressing ovarian Foxl2 and retinoic acid pathway; in adult ovaries, Foxl2 is required to suppress Dmrt1 expression and prevent transdifferentiation of ovarian cells into testicular like cells. (Figure from Eggers, et al., 2014)

Maintenance of testis and ovarian identity

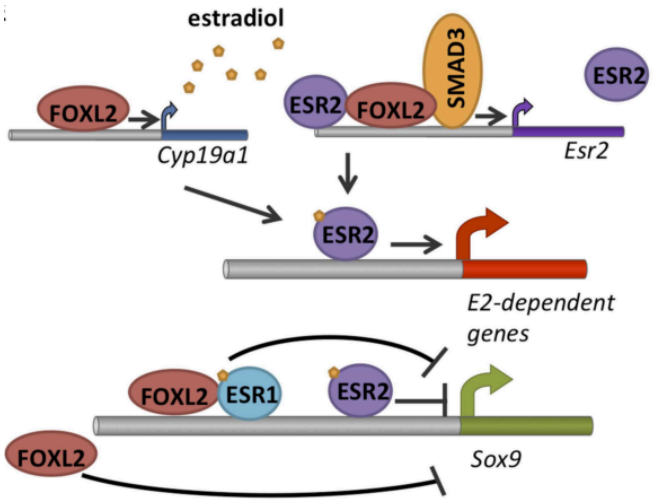
Components of both male and female sex differentiation pathways act antagonistically (See **Fig 14**). During embryonic development SOX9 and β-catenin are involved. In males, *SOX9* overexpression by SRY limits the expression of the female specific genes. The positive feedback loops of SOX9 and FGF9 actively supresses the activation of *WNT4*, moving the balance towards testis developmental pathway (Kim and Capel, 2006). On the other hand, WNT4 and RSPO1 activate and stabilize β-catenin, preventing the expression of *SOX9* by inhibiting the SOX9/FGF9 feedback loop, thus stabilizing ovarian development in females (Maatouk and Capel, 2008).

After birth, when sex is already determined, a different antagonistic maintenance regulatory network functions in each sex. In males, this network requires *Dmrt1*, and in females it involves *Foxl2* (Matson and arkower, 2012). Foxl2 is required to maintain the ovarian phenotype at the postnatal stage. Targeted ablation of *Foxl2* in the adult mice ovary leads to transdifferentiation of the supporting cells of the ovary into the supporting cells of the testes associated with the expression of pro-testis network including *Sox9*, *Dmrt1* and *Dhh* among others (See Fig 15) (Uhlenhaut et al., 2009). A recent study has shown that Foxl2 directly modulates *Esr2* expression through an intronic element in *Esr2* and that by doing so it represses *Sox9* expression (See Fig 16)(Georges et al., 2014).



INTRO Figure 15 - <u>Model of sex cell fate maintenance.</u>

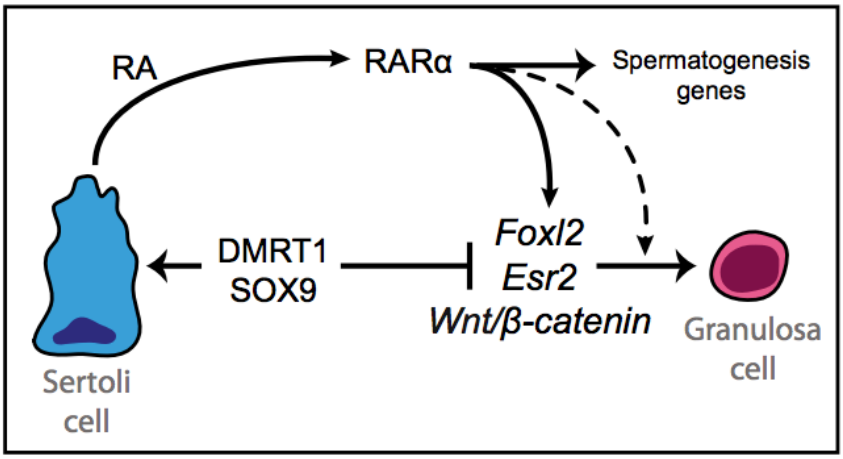
During sex determination, SRY upregulates *Sox9* expression and subsequent positive regulatory loops involving SOX9, FGF9 and PGD2 signaling, active and maintain *Sox9* expression in male gonads. In female gonads, b-catenin stabilized by WNT4 and RSPO1 actively repress *Sox9*. After birth, b-catenin declines in female gonads and FOXL2 and ESR1/2 actively repress Sox9 expression, ensuring ovarian somatic cell fate. This active transcriptional repression is necessary throughout life to prevent transdifferentiation of the somatic ovarian cells into testis like cells. (Figure from Ulenhaut, et al., 2009)



INTRO - Figure 16 <u>Model for FOXL2 regulation of estradiol signaling</u>.

FOXL2 may regulate estradiol signaling by A) regulating the production of E2 and b) regulation of estrogen receptivity by controlling the expression of Esr2. FOXL2 cooperates with Activin/SMAD3 and possibly ESR2 itself in the latter process. Sox9 repression seems to rely on multiple pathways, as FOXL2 and ESR1 have been described to co-bind a SOX9 regulatory element in the ovary. (Figure from Georges et al., 2014)

On the other hand, chromosome 9p deletions in human that remove the *DMRT1* gene are associated with a failure of testis determination (Matson et al., 2011). In mammals, DMRT1 functions by repressing female development. Knockout of *Dmrt1* in XY mouse Sertoli cells, in adult testis results in the activation of *Foxl2* and Sertoli cells reprogramming into granulosa and theca like cells, with production of oestrogen and germ cells which appear feminized (Matson et al., 2011). *Sox9* and *Sox8* are candidates for the maintenance of male fate since the loss of both genes after sex determination is associated to similar testicular anomalies as the ones seen in *Dmrt1* mutants (Barrionuevo et al., 2009). Recently, *Dmrt1* and *Sox9* have been shown to be required for the maintenance of postnatal testis via suppression of female specific determining genes, such as *Foxl2* (Minkina et al., 2014). More precisely, Dmrt1 blocks testicular retinoic acid signalling in the testis, controlling the expression of female sex determining genes (See Fig 17).



INTRO - Figure 17. Model of DMRT1 genetic regulation of cell fate in Sertoli cells.

DMRT1 direct repression of feminizing genes including Foxl2, Esr2, Wnt4 and Rspo1 in Sertoli cells (Matson et al., 2011). Sertoli cells express DMRT1 which allows the production of RA needed for spermatogenesis without causing RARa to activate these feminizing genes. The model also indicates that it is possible, based on data from other systems, that RARa synergizes with products of some of the feminizing genes to drive transdifferentiation. In addition to the genes shown, DMRT1 also represses Cyp19a1/aromatase, which makes estradiol that stimulates ER activity. (Figure from Minkina et al., 2014)

Taken together all information to date, indicate that both testis and ovaries are vulnerable to change, even after birth. Both require the constitutive expression of appropriate genes to actively supress the opposing pathway and prevent transdifferentiation. More genetic studies are needed to identify the genes/networks that are specific to post-natal sex maintenance and those that participate in foetal development and sex maintenance.

Disorders of Sex Development

Disorders of sex development (DSD) are defined as congenital conditions in which the development of chromosomal, gonadal or anatomical sex is atypical (Hughes et al., 2006). Truly ambiguous genitalia is rare, reported to occur in about 1:4500 births. However, a concern about the development of the external genitalia may exist in 1 in 300 newborn infants and in over 80% of cases the affected infants are raised as boys and have a presumed or actual XY karyotype (Ohnesorg et al., 2014).

46,XY DSD includes 46,XY complete or partial gonadal dysgenesis (errors of testis determination), or undervirilization or undermasculinization of a XY male due to defects in androgen synthesis or action. 46,XX DSD includes gonadal dysgenesis or, more commonly, overvirilization or masculinization of a XX individual due to androgen excess.

46, XY DSD

46,XY DSD includes 46,XY complete or partial gonadal dysgenesis (errors of testis determination), or undervirilization or undermasculinization of an XY male due to defects in androgen synthesis or action. 46,XY gonadal dysgenesis is characterized by a failure of testis determination. These cases can be subdivided into 46,XY complete gonadal dysgenesis (CGD) and 46,XY partial gonadal dysgenesis (PGD). The former is characterized by completely female external genitalia, well-developed Müllerian structures and a gonad composed of a streak

of fibrous tissue. 46,XY PGD is characterized by partial testis formation, usually a mixture of Wolffian and Müllerian ducts and varying degrees of masculinization of the external genitalia.

Of the mutations we know causing 46,XY CGD, about one third are point mutations or deletions involving *SRY* (Scherer et al., 1998). Mutations in other genes are associated with non-syndromic forms of 46,XY gonadal dysgenesis include *NR5A1, CBX2, DHH* (in a proportion of cases) and *MAP3K1*. Chromosomal deletions that include *DMRT1 (9p)* or *EMX2* (10q) as well as duplication of Xp (*DAX1)* are also associated with 46,XY DSD. Details about genes implicated in 46, XY DSD and their associated phenotypes can be found in each gene specific section of sex determination and in Table 2.

46, XX Ovotesticular DSD

46,XX testicular DSD refers to a male with testis and a normal male habitus, whereas 46,XX ovotesticular DSD refers to individuals that have both ovarian and testicular tissue in the gonads, usually as ovotestes but less commonly as separate gonads and azoospermia. A confirmed genetic diagnosis remains elusive in 60-70% of all of the cases carry the *SRY* gene usually as a translocation on the X chromosome (Délot and Vilain, 2015).

SRY and the X-linked gene *SOX3* are thought to share a common ancestor and during early mammalian evolution *SRY* arose from a gain-of-function mutation in the proto-Y allele of *SOX3* that resulted in testis specific expression. Although

Sox3 is normally not expressed in the urogenital ridge at the moment of sex-determination, it can substitute for *Sry* in testis determination. This is demonstrated by the ectopic expression of Sox3 in XX transgenic gonads that induces *Sox9* upregulation and consequent testicular development (Sutton et al., 2011). In the human *SOX3* loss-of-function mutations are not associated with either 46,XY gonadal dysgenesis nor 46,XX testicular/ovotesticular DSD but are associated with mental retardation and growth hormone deficiency (Laumonnier et al., 2002). However, rearrangements at the *SOX3* locus can lead to testis development in an XX background (Sutton et al., 2011). These studies raise the possibility that ectopic expression of other HMG-box containing proteins in the urogenital ridge at the moment of sex-determination may result in testicular development in a chromosomal female individual. One such example may be SOX10. Complete or partial duplications of chromosome 22 in 46,XX-*SRY* negative individuals are associated with various degrees of masculinization (Nicholl et al., 1994; Seeherunvong et al., 2004). Mutations involving *RSPO1* and *WNT4* are associated with exceptionally rare syndromic forms of 46,XX testicular or ovotesticular DSD. Human homozygous *RSPO1* mutations are associated with a rare recessive syndrome characterized by XX testicular DSD, palmoplantar hyperkeratosis and predisposition to squamous cell carcinoma of the skin (Parma et al., 2006). Mutations involving *RSPO1* have not been reported in non-syndromic cases of testicular and ovotesticular DSD. In the human, four dominant heterozygous missense mutations in *WNT4* have been reported in association with various degrees of virilization (with no overt evidence of testicular development) including androgen excess, abnormal development of Mullerian ducts but normal female genitalia (Biason-Lauber et al., 2004, 2006;

Philibert et al., 2008, 2011). A single homozygous *WNT4* mutation was reported in a consanguineous family with an embryonic lethal syndrome of 46,XX testicular DSD and dysgenesis of kidneys, adrenals, and lungs (SERKAL syndrome; SEx Reversion, Kidneys, Adrenal and Lung dysgenesis; (Mandel et al., 2008). The normal heterozygous carriers showed increased activity of 5 a-reductase suggesting that WNT4 may also play a role in prostate development (Mandel et al., 2008). Additional details about genes implicated in 46,XX DSD and their associated phenotypes can be found in Table 2.

Genomic studies of human patients presenting with DSD have revealed an increasing number of genes important for sex determination (Table 2). In addition murine studies have extended our understanding of the mechanism of action of these genes (Table 2). Understanding the molecular processes of testis development may also shed light on male infertility. However, despite these recent advances, with the exception of patients presenting with errors of testosterone biosynthesis, the genetic cause of 46,XY DSD remains unknown in the majority of cases.

Technological advances including SNPs arrays, NGS and epigenomic technologies are likely to lead to new and important findings allowing us to identify new genetic elements and the mechanisms of human sex determination and differentiation (Bashamboo et al., 2010b). Next generation sequencing (NGS) methods represent a paradigm shift in the use of sequencing in genetic diagnosis (Mardis, 2008). NGS have relative low costs and can sequence an entire human genome within a few hours and also allows the target sequencing of smaller

selections by capture of either a specific set of genes or the entire subset of the genome that is expressed, i.e., whole exome sequencing (WES) (Baxter and Vilain, 2013). Recent analysis of DSD patients using SNP arrays to identify genome wide duplications and deletions have identified many rearrangements in intergenic regions, some of them linked to already known sex genes (White et al., 2011). Whole exome sequencing of DSD patients to detect genomic rearrangements and mutations (up to single nucleotide level) is nowadays feasible and becoming the tool for excellence. WES has already proved by different research teams with positive and promising results (Bashamboo et al., 2014; Arboleda et al., 2013; Callier et al., 2014). These approaches offer affordable tools to detect already known and new mutations involved in disease, leading to a more efficient diagnosis and moreover the discovery of new genes involved in disease. Once novel mutations are identified, *in vitro* functional analyses of candidate genes are useful tools in order to better understand and validate their biological implication in DSD.

Gene/ Protein	Human genomic position	Mouse genomic position	Protein classification	Human phenotype		Mouse phenotype		Reference
				XY	XX	XY	XX	
Bipotential gonads								
EMX2	10q26	19D3	Transcription factor	Normal	Normal	No gonads present, No urogenital tract	No gonads present, no genital tract	Miyamoto et al. 19
CBX2	17q25	11E2	Polycomb protein	Normal ovaries, Müllerian structures present, Ovaries with oocytes	NR	*Cbx2*-/- : sex reversal		Katoh-Fukui et al.
LHX9	1q31-q32	1G-H2	Transcription factor	NR	NR	No gonads present, Müllerian ducts present	No gonads present	Birk et al. 2000
NR5A1/ SF1	9q33	2B	Transcription factor /Nuclear receptor	XY *NR5A1*-: sex reversal. Dysgenic testis, Müllerian ducts absent or present. Female, male or ambiguous external genitalia. Spermatogenic failure.	Dysgenic gonads (Primary ovarian failure). Müllerian structures present. Female like external genitalia	*Nr5a1*-/-: absense of gonads and adrenal glands. Nr5a1+/-: impaired adrenal stress response, small gonads	*Nr5a1*-/-: No gonads nor adrenal glands. *Nr5a*+/-: Small gonads, impaired adrenal glands response.	Luo et al. 1994; Ac et al. 1999; Louren 2009; Bashamboo 2010
Wt1	11p13	2E3	Transcription factor	Ambiguous or female external genitalia (Frasier), Dysgenetic gonads		No gonads present, No urogenital tract	No gonads present	Kreidberg et al. 19
Testis development								
DHH	12q13.1	15F1	Signaling molecule	Dysgenic testis. Müllerian ducts present. Female external genitalia	No gonadal phenotype	Dhh-/-: abnormal peritubular tissue, severly restricted spermatogenesis	No gonadal phenotype	Bitgood et al.
DMRT1	9p24.3	19C2-C3	Transcription factor	Dysgenic testis or ovotestis, Müllerian ducts absent or present. Female, male or ambiguous external genitalia	No gonadal phenotype	*Dmrt1*-/-: Severly impaired testis development (postnatal day 2). Loss of Sertoli cells and germ cells. Postnatal knockout of Dmrt1 in Sertoli cells leads to transdifferentiation of testis into ovaries	*Dmrt1* ectopic overexpression leads to transdifferentiation into testicular cells	Matson et al.
FGF9	13q22	14C3	Signaling molecule	NR	NR	*Fgf9* -/-: sex reversal, impaired development of Sertoli cells	No gonadal phenotype	Kim et al. 20
GATA4	8p23.1-p22	14D1	Transcription factor	Dysgenic testis. Müllerian ducts absent. Female or ambiguous external genitalia		*Gata4* -/-: embryonically lethal. *Gata4* ki: testis development severly affected	No gonadal phenotype	Tevosian et al.

KDM3A	2p11.2	6C1	Histone demethylase	NR	NR	XY male to female sex reversal with variable gonadal phenotypes	No gonadal phenotype	Kuorki et al. 20
MAP3K1	5q11.2	13D2	Kinase	Dysgenic testis, Müllerian ducts may be present. Female, male or ambiguous external genitalia	No gonadal phenotype	No gonadal phenotype	No gonadal phenotype	Warr et al. 201
HHAT	1q32	1H6	Membrane bound acyltransferase	Partial and complete Testicular dysgenesis	No gonadal phenotype	*Hhat*Creface/Creface:Gonadal dysgenesis, apolar Sertoli cells, decreased number of fetal Leydig cells, hypogonadism and feminized external genitalia	No gonadal phenotype	Callier et al. 20
MAP3K4	6q26	17A1	Kinase	No gonadal phenotype	No gonadal phenotype	XY male to female sex reversal, deficient growth and support of mesonephic cell migration	No gonadal phenotype	Bogani et al. 20
NR0B1 / DAX1	Xp21.3	XC1	Nuclear Receptor	*DAX1* duplication in XY: male to female sex reversal with gonadal dysgenesis		*Dax1* duplication in XY: male to female sex reversal, gonadal dysgenesis. *Dax1* mutant mice show testicular dysgenesis depending of genetic background		Meeks et al. 2003, S al. 1998
SIX1-SIX4	14q23	12C3	Transcription factor	NR	No gonadal phenotype	*Six1-/-, Six4-/-* XY male to female sex reversal, ovarian development in the majority of XY animals. *Six1-/- or Six4-/-* mice do not have gonadal phenotype	No gonadal phenotype	Fujimoto et al. 2
SOX3	Xq27.1	XA7	Transcription factor	NR	XX with rearrangements (duplication) within *SOX3*: Male or ambigous external genitalia. Atrophic testis with loss of spermatogenesis	*Sox3* ectopic expression cause complete male to female sex reversal	*Sox3* knockout have deficient gonadal function	Weiss et al. 2003; Su al. 2011
SOX8	16p13.3	17A3.3	Transcription factor	NR	NR	*Sox8* -/-XY mice: reduced fertility. *Sox8;Sox9* double knockout XY mice present variable degree of male to female sex reversal	No gonadal phenotype	Chaboissier et al. Barrionuevo et al.
SOX9	17q24-q25	11E2	Transcription factor	*SOX9* +/- show male to female sex reversal	*SOX9* Duplications:testicular DSD. Müllerian ducts absent. Male or ambiguous external genitalia	*Sox9*-/-XY mice: complete male to female sex reversal	Ods (deletion upstream *Sox9*) in XX mice: complete female-to-male sex reversal. *Sox9* ectopic expression: female-to male sex reversal	Barrionuevo et al. Bishop et al. 2000; V al.2001; Wagner et a Huang et al. 19
SOX10	22q13.1	15E1	Transcription factor	No gonadal phenotype	*SOX10* duplication in XX patient: gonadal dysgenesis	No gonadal phenotype	*Sox10* overexpression: Female to male sex reversal	Polanco et al. 20

SRY	Yp11.3	YA1	Transcription factor	*SRY*-: male to female sex reversal. Dysgenic testis. Müllerian ducts absent or present	*SRY*+: female to male sex reversal. Male or ambiguous external genitalia.	*Sry*-/-: testicular dysgenesis. Müllerian ducts present	*Sry* translotaction: Complete sex reversal	Berta et al. 1990; al. 1990
PTGDS	9q34.3	2A3	Signaling molecule	NR	NR	*Ptgds*-/- XY mice: delayed testis cord formation		
ZFPM2 /FOG2	8q23	15C	Cofactor of GATA4	*FOG2*-/-: gonadal dysgenesis	No gonadal phenotype	*Fog2*-/- XY mice: low Sr expression, male to female sex reversal	No gonadal phenotype	Bouma et al. Bashamboo et
varian Development								
CTNNB1 (b-catenin)	3p21	9F4	Transcription factor	NR	NR	XY mice with stabilized b-cat expression present male to female sex reversal	*Ctbnnb1* -/- show down regulation of ovary promoting genes	Maatouk et a
FOXL2	3q23	9E4	Transcription factor	BPES	Mutated *FOXL2* patients show premature ovarian failure	No gonadal phenotype	*Foxl2*-/-: infertility. *Foxl2* ablation in adult ovaries: transdifferentiation ovary-to-testis	Uhlenhaut et a Boulanger et a
RSPO1	1p34.3	4D2	Signaling molecule	No gonadal phenotype	Ovotestis or tesis. Müllerian structures absent. Male or ambigous external genitalia		*Rspo1* -/-: partial sex reversal, ovaries with male features like coleomic vessel formation, depletion of fetal oocytes, ectopic testosterone production	Tomizuka et a
WNT4	1p35	4D3	Signaling molecule	*WNT4* duplication: male to female sex reversal. Dysgenic testis. Müllerian ducts present. Ambiguous external genitalia	Testis or ovotestis or ovaries. Müllerian structures absent. Male/ambigous/female external genitalia	*Wnt4* transgenic model: abnormal testicular vasculature and function, no sex reversal	*Wnt4*-/-: partial sex reversal, ovaries presenting male features like coelomic vessel formation, Müllerian ducts absence, presence of Wolffian ducts and ectopic testosterone production	Vainio et al. 2004; al. 2001
R= No Reported								

able 2 . Genes involved in regulation of bipotential gonads, tesis and ovaries development. Chromosomal position, protein function, gonadal phenotype in human and mice. NR: No ported

Animal models

Our current understanding of the genetic regulation of sex determination and differentiation has arisen mostly from the mutational and functional analysis from patient presenting with DSDs. Among mammals, the mechanisms that underlie sex development have been most extensively studied and are best understood in mice. The pertinent use of mouse models of DSD and *in vitro* cellular models, in particular mouse models employing targeted mutagenesis and transgenesis have greatly contribute to our understanding of gene function and to decipher the molecular networks in sex development and maintenance (Table 2). However, it is important to consider that although mouse may be an excellent model system in many cases, it does not mimic exactly the mechanisms of sex determination and maintenance *in vivo* in humans. Differences between both organisms during sex determination and development are evident, for example, spatio-temporal expression of *Sry* expression strictly limited in mouse, while in humans *SRY* expression is maintained throughout lifetime (Kashimada and Koopman, 2010). Moreover, there are genes that do not show similar biological function on gonad development in the two species this is the case of *MAP3K1*, which is associated with 46,XY DSD gonadal dysgenesis in human, but in mouse does not have a gonadal phenotype (Pearlman et al., 2010). However, another member of MAPK family: *Map3k4* has been identified to have a gonadal phenotype in mouse when ablated (Bogani et al., 2009). Similarly, deletions of chromosome 9p including *DMRT1* have been linked to 46,XY DSD in human

(Herpin et al., 2010), whereas mutations *Dmrt1* knockouts in mice do not show primary sex reversal since mouse born with testis, however they later undergo abnormal differentiation (Raymond et al., 2000).

The detailed genetic analysis of human cases with anomalies in sex-determination has been a powerful tool in the identification of genes involved in sex-determination. *SRY* was mapped and eventually cloned through the analyses of XY females with gonadal dysgenesis and XX males carrying a portion of the Y chromosome on one of their X chromosomes (Guellaen et al., 1984). *WT1* and *SOX9* were identified through the analyses of human chromosomal rearrangements (Gessler et al., 1990; Call et al., 1990; Wagner et al., 1994; Foster et al., 1994). Furthermore, our lab and others have demonstrated a contribution of NR5A1, GATA4 and FOG2 in human DSD (Bashamboo et al., 2010a; Lourenço et al., 2011b; Bashamboo et al., 2014). The use of omics technologies such as exome sequencing has substantially increased our knowledge of the genetic mutations that can result in an error of sex-determination and discovery of mutations in novel genes can be expected as more and more patients are sequenced. Within the limitations of the current mouse models and *in-vitro* assays, we can at best predict the effect(s) the mutation may have on the biological activity of the protein in the context of known interactions. However, if we want to fully understand the mechanisms involved in decisions of cell fate choice in the developing gonad a series of complementary studies are required using different model systems- modeling cells derived from the patient carrying the mutation of interest and human fetal testis xenograft models.

Cellular models

Recently, ectopic xenografting has become a valid model to induce spermatogenesis in immature testicular tissue. The subcutaneous transplantation of pieces of testis tissue in immunodeficient hosts allows complete spermatogenesis in fresh and cryopreserved neonatal and prepubertal testicular tissue from different species including rhesus macaque and human into mouse host. (Rodriguez-Sosa and Dobrinski, 2009). Xenografts are useful models to study spermatogenesis, toxicological effects, influence of irradiation and cryopreservation studies on testicular tissues. (Mitchell et al., 2010). This technique may be modified to study genes involved in sex determination and testis development *in vitro.* In humans, studies showing that fetal human testicular tissue xenografts in nude mice have expression of testis cell markers however, they have low survival and post-natal tissue xenografts survival is also limited, although not completely mature sperm has been reported until now. Moreover, the study of human testis xenografts has been limited by the lack of donor tissue.

Different gonad cell lines have been produced, characterized and used with limited success. Murine cell lines such as TM4 (Sertoli like) and TM3 (Leydig like) (Mather, 1980), Leydig-like clonal cell lines (Ascoli, 1981), primary cultured cells from E 11.5 gonad (Nishino et al., 2004) and immortalized gonadal cell lines from transgenic mice or embryo that overexpress the SV40 T-antigen (Hofmann et al., 1992) exist. Moreover, many cell from rodent immature or adult testis, have been immortalized as cell lines, however, due to the species specific

temporal expression pattern of *Sry* after sex determination, these immortalized cell lines do not express *Sry* or other stage specific factors involved in the differentiation of the testis. TM3 and TM4 cell lines do not express *Sry* in long term culture (Beverdam et al., 2003). There are also immortalized primary cell lines established from isolated cells of the genital ridge of rodent gonads during sex determination time window (Haqq et al., 1994; Capel et al., 1996; Nishino et al., 2004). In these cell lines *Sry* expression decreases and is finally lost after long period of culture. Transient overexpression of *SRY* in these cell lines was used to study downstream targets of *SRY*, however the results obtained are not consistent. BCG, TM3 and TM4 cells transiently overexpressing SRY show an absence of *Sox9* expression (Beverdam et al., 2003). The use of a human cell line would be more desirable given the major differences in human and mice SRY proteins. Indeed, NT2/D1 a human multipotent testicular embryonic carcinoma cell line has been shown to express 35 of 39 testis-specific genes including markers for Sertoli cells, germ cells and steroidogenic enzymes present in Leydig cells, however, genes such as *WT1* and *DAX1* were not expressed and some ovarian genes such as *Fst* were expressed, suggesting the possibility of NT2/D1 cells may mimic an early stage of mammalian sex determination; however further studies are necessary to clearly understand the profile of gene expression in NT2/D1 cells (Knower et al., 2007).

Until now, there is no established cell line that exhibits all the properties of Sertoli cells. Primary immature, mature Sertoli cells and established cell lines lose their characteristics during long-term culture. They loss *Sry* expression, they do not bind to spermatids *in vitro*, making them unsuitable for studies on

spermatogenesis, or the interaction between Sertoli cells–germ cells, they are genomically unstable, as most of the reported cell lines are aneuploid and present chromosome rearrangements (Gassei and Schlatt, 2007). Recently, Buganim et al., have generated induced embryonic Sertoli like cells (ieSCs) by direct reprogramming of mouse embryonic fibroblasts (MEFs) (Buganim et al., 2012) by the dox induced expression of the five transcription factors *Nr5a1, Wt1, Gata4, Sox9* and *Dmrt1*. Their studies showed that expression of *Nr5a1*, *Wt1*, and *Dmrt1* initiated cell proliferation and mesenchymal to epithelial transition (MET), whereas *Nr5a1, Wt1*, and *Sox9* induced cell aggregation. ieSCs have a unique but similar expression pattern to fetal Sertoli cells, a normal karyotype, are able to maintain germ cells survival in culture and contribute to *in vivo* Sertoli cell population. Moreover, ieSCs are migratory epithelial cells that have the capability to aggregate, form tubule-like structures, and recruit endothelial cells (Buganim et al., 2012). However, ieSCs, are not able to grow for long periods in the absence of dox without losing the Sertoli cell characteristics. Moreover, the Sertoli cell marker, *Fgf9* was not highly expressed and *Sox9* expression was variable in ieSCs. The development of cell lines like these which mimic accurately *in vivo* gonad developmental process, represents not only the possibility of having cellular models to perform biochemical/genetic studies of sex determination and differentiation but also DSD.

Although the early genetic, cellular and morphological events during gonadal development have been characterized, the molecular mechanisms involved in human sex determination are poorly understood. Induced pluripotent stem cells (iPSCs) provide a practical system for directed differentiation of individual

specific cells into the cell-lineages of choice. Furthermore, transdifferentiation approaches may offer an alternative to iPSCs as they permit the conversion of committed cells into the lineage of choice, without the added manipulations required to revert them to the ground state. The development of novel cellular models based on biomaterial from patients with disorders of sex development (DSD) and their unaffected family members are one way forward to better understand the choice of cell fate during sex-determination and hence the mechanism of disease development in errors of sex determination.

RESULTS

An ancient protein-DNA interaction underlying metazoan sex determination

Introduction to Publication

The Human *DMRT1* gene is homologous to the *Drosophila* double sex (*dsx*) and *C. elegans mab-3* genes, which encode for transcription factors that bind DNA via a common DNA-binding domain, the DM domain (Cheng et al., 2006). Both *Dsx* and *mab-3* control sex specific differentiation. Mouse gonads express *Dmrt1* in the genital ridges of both sexes and then becomes male specific at the end of sex determination phase (Raymond et al., 2000). In mammals, DMRT1 binds and regulates genes with essential roles in mammalian sex development; it can activate male *Sox9* and repress female *Wnt4* and *Rspo1* (Matson et al., 2011). Deletions of chromosome 9p including *DMRT1* have been linked to 46,XY gonadal dysgenesis and occasional male to female sex reversal (Tannour-Louet et al., 2010). However, the role of DMRT1 during human sex determination has not been fully understood. Here, we describe how DMRT proteins employ a unique binding interaction with DNA, inserting two adjacent antiparallel recognition helices into a widened DNA major groove to make base-specific contacts. Human DMRT1 can use multiple DNA binding modes - tetramer, trimer, dimer – to interact with DNA. ChIP-Exo experiments indicated that multiple DNA binding modes also are used by DMRT1 *in vivo*. Finally in this publication, we show that mutations affecting residues crucial for DNA recognition that result in

an intersex phenotype in flies are also associated with 46,XY complete gonadal dysgenesis in humans. This was an international collaborative study. My contribution consisted of the identification and functional characterization of the mutant human DMRT1 protein. This includes results that were not part of the original publication but which are described here.

nature structural & molecular biology

An ancient protein-DNA interaction underlying metazoan sex determination

Mark W Murphy[1,8], John K Lee[2,8], Sandra Rojo[3,8], Micah D Gearhart[1,8], Kayo Kurahashi[2], Surajit Banerjee[4], Guy-André Loeuille[5], Anu Bashamboo[3], Kenneth McElreavey[3], David Zarkower[1,6,7], Hideki Aihara[2,6] & Vivian J Bardwell[1,6,7]

DMRT transcription factors are deeply conserved regulators of metazoan sexual development. They share the DM DNA-binding domain, a unique intertwined double zinc-binding module followed by a C-terminal recognition helix, which binds a pseudopalindromic target DNA. Here we show that DMRT proteins use a unique binding interaction, inserting two adjacent antiparallel recognition helices into a widened DNA major groove to make base-specific contacts. Versatility in how specific base contacts are made allows human DMRT1 to use multiple DNA binding modes (tetramer, trimer and dimer). Chromatin immunoprecipitation with exonuclease treatment (ChIP-exo) indicates that multiple DNA binding modes also are used *in vivo*. We show that mutations affecting residues crucial for DNA recognition are associated with an intersex phenotype in flies and with male-to-female sex reversal in humans. Our results illuminate an ancient molecular interaction underlying much of metazoan sexual development.

The sex of animals can be determined by varied cues in different species, including chromosomes, temperature, social status and photoperiod[1]. A common feature of sexual regulation across much of the animal kingdom is the involvement of DMRT proteins[2,3]. These are transcription factors related to Doublesex (Dsx) and Male abnormal-3 (MAB-3), key sexual regulators of insects and nematodes, respectively, and they share the highly conserved DM DNA-binding domain[4,5].

Genetic studies have found that DMRT-encoding genes can control the primary sex-determination decision, can act subsequently in sexual differentiation or, in some species, can do both[2]. DMRT-encoding genes are required for sexual development in planaria[6], insects[7], nematodes[8] and vertebrates[9], thus suggesting that their involvement in this process spans hundreds of millions of years. Vertebrates have six to seven DMRT-encoding genes, and at least one of these appears to regulate testis development in most or possibly all species, with *DMRT1* playing a leading part in the process. In some vertebrate groups, including birds[10,11], some fish[12] and amphibians[13], a *DMRT1* ortholog is located on a sex chromosome and has a sex-determining role[2]. In mammals, *DMRT1* is crucial for many aspects of testicular development[2]. Deletions of human chromosome 9p that cause *DMRT1* hemizygosity result in 46,XY gonadal dysgenesis, which can include sex reversal[14,15]. In mice, DMRT1 has been shown to regulate gonadal differentiation, and continuous DMRT1 expression is required to maintain the male cell fate of testicular Sertoli cells, by preventing their transdifferentiation to female granulosa cells[16]. Moreover, DMRT1 overexpression in mouse ovaries can cause male sex determination or female-to-male cell-fate transdifferentiation[17,18].

DMRT1 appears to be a bifunctional transcription factor, activating or repressing transcription of target genes. We previously found in mice that DMRT1 binds and regulates genes known to have key roles in mammalian sexual development, by activating the central male sex–determining gene *Sox9*, repressing the female sex–determining genes *Wnt4* and *Rspo1*, and regulating many other genes involved in subsequent sexual differentiation[16]. Here we have combined genomic, molecular, biochemical, structural and human genetics approaches to address how DMRT1 recognizes target-site DNA. We find that DMRT1 uses a unique type of protein-DNA interaction and can also use multiple distinct stoichiometries to discriminate among target sites with distinct DNA sequences. We also show that disrupting conserved residues in the DM domain that make base-specific contacts with DNA can severely reduce binding affinity and cause sex reversal in flies and in humans.

RESULTS

In vivo DMRT1-binding-site determination in mice and humans

DMRT1 binds *in vitro* to a pseudopalindromic 13-bp DNA sequence[19], but how the DM domain recognizes target DNAs is poorly understood because no composite protein–DNA structure has been described. For a genome-wide view of DMRT1-binding sites on DNA, we

[1]Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, Minnesota, USA. [2]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, Minnesota, USA. [3]Unit of Human Developmental Genetics, Institut Pasteur, Paris, France. [4]Northeastern Collaborative Access Team, Cornell University, Argonne, Illinois, USA. [5]Service de Pédiatrie, Centre Hospitalier de Dunkerque, Dunkerque, France. [6]Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota, USA. [7]Developmental Biology Center, University of Minnesota, Minneapolis, Minnesota, USA. [8]These authors contributed equally to this work. Correspondence should be addressed to V.J.B. (bardw001@umn.edu) or H.A. (aihar001@umn.edu).

Received 5 January; accepted 21 April; published online 25 May 2015; doi:10.1038/nsmb.3032

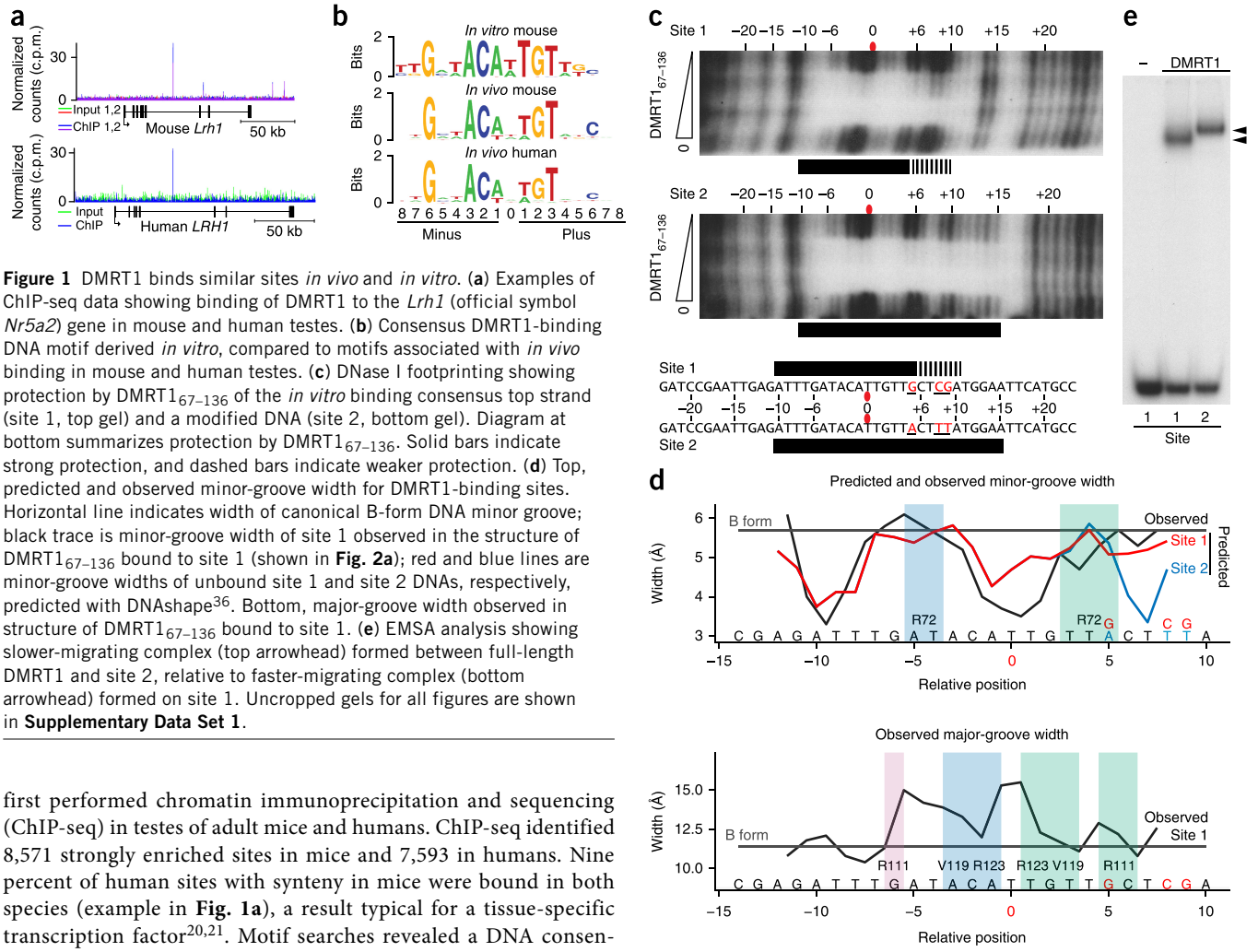


Figure 1 DMRT1 binds similar sites *in vivo* and *in vitro*. (**a**) Examples of ChIP-seq data showing binding of DMRT1 to the *Lrh1* (official symbol *Nr5a2*) gene in mouse and human testes. (**b**) Consensus DMRT1-binding DNA motif derived *in vitro*, compared to motifs associated with *in vivo* binding in mouse and human testes. (**c**) DNase I footprinting showing protection by $DMRT1_{67–136}$ of the *in vitro* binding consensus top strand (site 1, top gel) and a modified DNA (site 2, bottom gel). Diagram at bottom summarizes protection by $DMRT1_{67–136}$. Solid bars indicate strong protection, and dashed bars indicate weaker protection. (**d**) Top, predicted and observed minor-groove width for DMRT1-binding sites. Horizontal line indicates width of canonical B-form DNA minor groove; black trace is minor-groove width of site 1 observed in the structure of $DMRT1_{67–136}$ bound to site 1 (shown in **Fig. 2a**); red and blue lines are minor-groove widths of unbound site 1 and site 2 DNAs, respectively, predicted with DNAshape[36]. Bottom, major-groove width observed in structure of $DMRT1_{67–136}$ bound to site 1. (**e**) EMSA analysis showing slower-migrating complex (top arrowhead) formed between full-length DMRT1 and site 2, relative to faster-migrating complex (bottom arrowhead) formed on site 1. Uncropped gels for all figures are shown in **Supplementary Data Set 1**.

first performed chromatin immunoprecipitation and sequencing (ChIP-seq) in testes of adult mice and humans. ChIP-seq identified 8,571 strongly enriched sites in mice and 7,593 in humans. Nine percent of human sites with synteny in mice were bound in both species (example in **Fig. 1a**), a result typical for a tissue-specific transcription factor[20,21]. Motif searches revealed a DNA consensus element associated with *in vivo* binding in both species, which includes several nearly invariant nucleotides and resembles the *in vitro* consensus (**Fig. 1b**).

To examine binding to the *in vitro* DNA consensus (site 1), we next mapped DMRT1-DNA interactions by DNase I protection. A truncated DMRT1 protein containing the highly conserved human DM domain, $DMRT1_{67–136}$ (**Supplementary Fig. 1**), protected the top DNA strand beyond the central 13 bp (**Fig. 1c**). Protection was stronger on the left side, which is predicted[22] to have a narrower minor groove (**Fig. 1d**). We made base changes to site 1 that are expected to compress the right-side minor groove (site 2; **Fig. 1d**), and these resulted in extended protection and reduced electrophoretic mobility shift of full-length DMRT1 protein (**Fig. 1c**,**e**). The altered DNase I protection and electrophoretic mobility of site 2 relative to site 1 together suggest that DMRT1 can bind DNA with multiple stoichiometries or conformations.

DMRT1 inserts paired α-helices into the DNA major groove

For a detailed view of the DM-domain structure and insight into how DMRT1 interacts with DNA, we used X-ray crystallography, examining the interaction between the DMRT1 DM domain and site 1. Crystals of $DMRT1_{67–136}$ (**Supplementary Fig. 1**) and a 25-bp DNA corresponding to site 1 yielded a 3.8-Å-resolution structure (**Table 1**) containing three DM-domain protomers bound to a single DNA molecule (protomers A–C; **Fig. 2a**,**b**). The overall resolution of the $DMRT1_{67–136}$–DNA structure is not high; this probably reflects a combination of inherent flexibility of the complex, loose lattice contacts and the crystals' high solvent content and radiation sensitivity. We were able to mitigate these issues by validating the registers of DNA and protein residues, using crystals containing brominated DNA and selenomethionine-substituted protein (**Supplementary Fig. 2** and **Table 1**).

This first view of a DM domain bound to DNA revealed a unique type of DNA interaction. The zinc-binding module of each protomer spans the DNA minor groove, primarily through phosphate-backbone contacts, while a recognition helix inserts into the major groove, making base-specific contacts (**Fig. 2a**,**b**). Unexpectedly, recognition helices of protomers A and B lie antiparallel together in the major groove on one side of the consensus element, while a third (protomer C) lies in the major groove on the other side. We are unaware of any other protein that binds DNA by insertion of two adjacent α-helices into the same region of the major groove.

In the structure, protomers A and B bind DNA differently, reflecting different angles between their zinc-binding modules and recognition helices (**Fig. 2c**). Major-groove contacts on the left side of the binding site involve three amino acids (R111, V119 and R123) provided by protomers A and B (**Fig. 2d**,**f**). By contrast, although major-groove contacts on the right side also involve these same three amino acids, all are provided by protomer C (**Fig. 2d**,**g**,**i**). The left-side major groove is unusually wide (**Fig. 1d**), and it accommodates protomers A and B, which sit more perpendicular to the helical axis

Table 1 Data collection and refinement statistics

	Native (Zn-SAD)	BrDNA[b]	SeMet[b]
Data collection			
Space group	*I*222	*I*222	*I*222
Cell dimensions			
a, *b*, *c* (Å)	82.56, 138.49, 138.49	82.99, 138.31, 138.98	83.09, 139.48, 142.18
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Wavelength (Å)	1.28	0.92	0.9795
Resolution (Å)	3.81 (3.95–3.81)[a]	3.76 (3.96–3.76)	4.93 (5.15–4.93)
R_{merge}	0.070 (1.02)	0.052 (0.67)	0.043 (0.395)
$I / \sigma I$	12.9 (1.9)	16.7 (2.4)	10.5 (2.4)
Completeness (%)	99.9 (100.0)	99.9 (99.8)	97.0 (98.5)
Redundancy	7.1 (7.3)	6.6 (6.3)	3.0 (3.1)
Refinement			
Resolution (Å)	3.81		
No. reflections	8,170		
R_{work} / R_{free}	22.9 / 26.2		
No. atoms	2,534		
Protein/DNA	2,528		
Ligand/Ion	6		
B factors	133.0		
Protein/DNA	133.0		
Ligand/Ion	120.0		
r.m.s. deviations			
Bond lengths (Å)	0.004		
Bond angles (°)	0.90		

[a]Values in parentheses are for highest-resolution shell. [b]These data sets were not used for phasing but were used to guide the building of the model. SAD, single-wavelength anomalous dispersion; SeMet, selenomethionine.

than protomer C (**Fig. 2d**,**e**). In protomers B and C, R72, N terminal to the zinc-binding module, inserts into the minor groove to hydrogen-bond with base pairs (**Fig. 2d**,**f**–**h**); these interactions are consistent with use of arginine by other proteins to mediate minor-groove contacts[23]. The hydrogen-bond donor-acceptor pattern is almost indistinguishable in the minor groove for A-T versus T-A, and G-C versus C-G base pairs[24,25]. Thus, base readout by R72 probably involves base-pair recognition rather than base recognition and may also involve shape readout[25]. All three DMRT1 protomers extensively contact the DNA backbone (**Fig. 3a**). They also interact with each other: the recognition helices of protomers A and B are held in close apposition by an interdigitating hydrophobic zipper, and protomers B and C are hydrogen-bonded (**Fig. 3b**,**c**). The overall folding pattern of the zinc-binding module is very similar to that of Dsx (**Fig. 3d**), previously determined by NMR[26]. Critical DM-domain amino acids and protein-DNA and protein-protein contacts are summarized in **Figure 3e**,**f**. The contacts shown can explain virtually all of the conserved DM-domain amino acids and DNA nucleotides in the binding site (**Fig. 3f**). Most conserved amino acids without functions indicated have structural roles in maintaining the overall structure of the DM domain, for example by terminating helical domains, allowing bending or mediating folding of the zinc-binding module[26].

Sequence-specific binding is primarily via the DNA major groove

A prior study[26] found that DM-domain DNA binding tolerates extensive chemical modification of the DNA major groove but not the minor groove, and it proposed on this basis that binding is mainly mediated by sequence-specific minor-groove contacts. However, minor-groove contacts can distinguish only A-T from G-C base pairs but cannot distinguish specific sequences. Indeed, although our structure revealed potential hydrogen-bond interactions of R72 with the minor groove, the positions contacted by R72 do not show strong sequence conservation. By contrast, the structure revealed extensive sequence-specific major-groove interactions. These interactions involved highly conserved DNA base pairs (−6 and +6; −2 and +2) that were not specifically tested in the previous study. To verify the importance of these base pairs, we first changed the −6 and +6 positions from dG-dC to dA-dT (**Fig. 4a**), and this strongly reduced DMRT1 binding (**Fig. 4b**). To query the minor groove at these positions, we substituted dI-dC base pairs, removing minor-groove exocyclic amines without altering the major groove (**Fig. 4a**); these substitutions did not reduce binding (**Fig. 4b**). To query the major groove, we substituted 2-aminopurine (2AP)-dU base pairs, inverting the carbonyl oxygen and removing the exocyclic amine from the major groove without altering the minor-groove structure (**Fig. 4a**). 2AP-dU substitution virtually eliminated DMRT1 binding, thus demonstrating the importance of the major-groove sequence identity at these positions (**Fig. 4b**). The same major-groove modifications at the −2 and +2 positions also reduced binding, but minor-groove modifications at these positions did not (**Fig. 4c**). In summary, the −6 and +6 and the −2 and +2 positions are crucial for DNA binding, and these positions are recognized primarily via the major groove.

We also used protein-sequence substitutions to assess the importance of amino acid side chains that make major- or minor-groove contacts. Replacing R72, R111, V119 or R123 with alanine reduced or eliminated DMRT1 binding, results indicating that these residues are crucial for DNA recognition (**Fig. 4d** and **Supplementary Fig. 3**). DNA-backbone contacts also are important for DMRT1's DNA binding affinity, because K92A reduced binding (**Fig. 4e** and **Supplementary Fig. 3a**). These functional analyses were particularly important given the limited resolution of the X-ray structure.

DMRT1 can bind DNA as a tetramer, trimer or dimer

Next, we further examined DNA binding stoichiometry. The structure showed that site 1 can bind a DMRT1 trimer, and DNase I protection showed that site 2 is more extensively protected by DMRT1 than site 1. These data suggest that the slower-migrating electrophoretic mobility shift analysis (EMSA) complex on site 2 (**Figs. 1e** and **5a**) is a symmetric ABB′A′ tetramer. Site 2 differs from site 1 at +5, a position that is uniquely recognized by protomer C (**Figs. 2d**), and it also has changes at +8 and +9 that are predicted to narrow the minor groove between +6 and +8 (**Fig. 1d**). These differences suggest that DNA sequence and shape may dictate protein binding mode. We hypothesized that making other structure-guided sequence changes to site 1 might instead cause AB dimers to form. Indeed, modifying site 1 at +2 and +6 to alter bases recognized specifically by protomer C (site 3) generated a faster-migrating EMSA complex consistent with an AB dimer (**Fig. 5a** and **Supplementary Fig. 3b**). To confirm that DMRT1 can bind DNA with multiple stoichiometries, we performed additional EMSAs. Instead of full-length DMRT1 (as in **Fig. 5a**), we used $DMRT1_{67–136}$, which removes a multimerization domain (not shown) and reduces cooperative binding. $DMRT1_{67–136}$

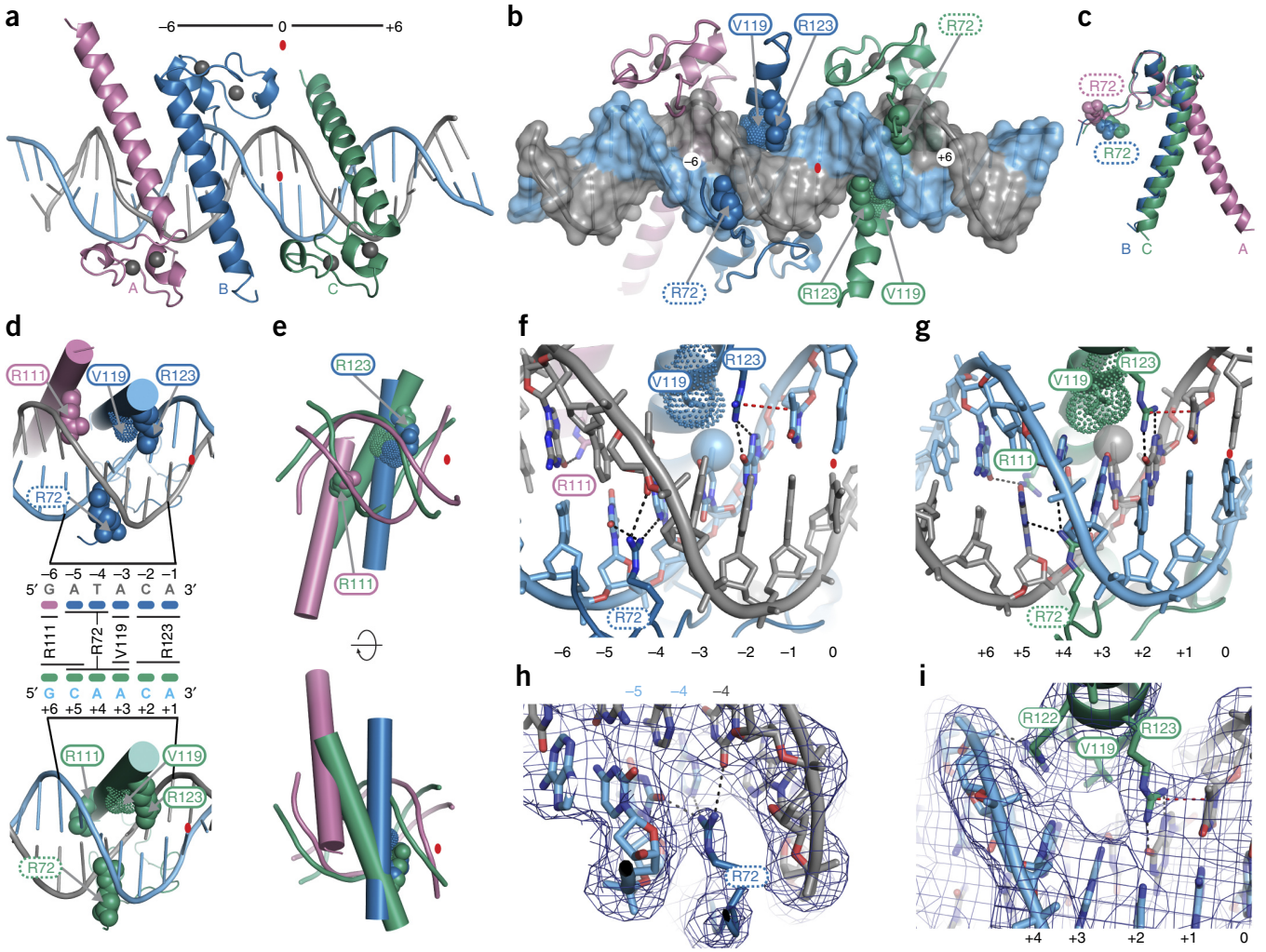


Figure 2 Major-groove interactions and use of multiple DNA binding modes by DMRT proteins. (**a**) Overview of structure, showing two DMRT1$_{67–136}$ protomers (A and B in pink and blue, respectively) inserted in DNA major groove on one side of site 1 and one protomer (C, green) in the major groove on the other side. Binding-site symmetry is indicated (–6 and +6). Red oval, central base pair; gray spheres, zinc ions. (**b**) Interaction of protomers A, B and C from the back side, highlighting insertion of R72 side chains of protomer B and C (dashed ovals) into the minor groove. Amino acids labeled with solid ovals make major-groove DNA contacts. (**c**) Overlay showing different orientation of R72 and a different angle between the zinc-binding module and recognition helix of protomer A relative to those of B and C. (**d**) DNA base contacts. Middle diagram summarizes contacts made by each protomer. (**e**) Overlaid views of DMRT1 recognition helices bound to DNA, aligned at R111 and R123 and viewed from front and back. Right- and left-side DNAs are color-coded: pink is bound by protomers A and B, and green by protomer C. (**f**) Major-groove and minor-groove interactions on the left side of site 1. (**g**) Major-groove and minor-groove hydrogen-bond interactions on the right side of site 1. (**h**) Interactions of protomer B R72 with the minor groove. (**i**) Interactions of protomer C with the major groove. In **h** and **I**, blue mesh shows the σ_A-weighted $2F_o - F_c$ electron density map contoured at 1.5σ. Black dashed lines in **f**–**i**, hydrogen bonds; red dashed line, arginine-thymine stacking interaction; stippled and solid spheres, van der Waals radii.

formed three distinct complexes with site 1 (**Fig. 5b**), which we interpret as monomers, dimers and trimers. Because binding to site 2 was highly cooperative even with DMRT1$_{67–136}$, we also assayed a site with reduced affinity and cooperativity (site 4). On site 4, DMRT1$_{67–136}$ formed four complexes (**Fig. 5b**), which we interpret as monomer through tetramer. To further confirm these stoichiometries, we performed protein cross-linking, using full-length DMRT1 bound to sites 1–3 (**Fig. 5c**). As predicted, DMRT1 formed DNA-dependent complexes of different maximum stoichiometries. Dimers formed on site 3; dimers and trimers formed on site 1, with traces of tetramer; and dimers, trimers and tetramers formed on site 2. Together, the structure, DNase I protection, EMSA analyses and protein cross-linking indicate that DMRT1 can bind DNA *in vitro* as a tetramer, a trimer or a dimer and, when the protein is truncated, as a monomer.

We next asked whether DMRT1 also binds DNA by using multiple stoichiometries *in vivo*. For a higher-resolution view of DMRT1-DNA interaction *in vivo* we used ChIP-exo[27], which uses strand-specific exonuclease digestion before sequencing to localize protein-DNA cross-links with higher precision than ChIP-seq. ChIP-exo did not reveal exact binding details at individual sites, so we used structural and *in vitro* DNA binding properties to group sites (**Fig. 5d**) and reveal their patterns of binding, a strategy that has also been used with other proteins[28]. We searched the genome for matches to the 7-bp core DMRT1 binding motif and selected those found under DMRT1 ChIP peaks. We then used minor-groove width predictions[29] to group peaks into those predicted to have bilateral narrowing of the minor groove (tetramers) and those with unilateral narrowing (dimers and trimers). Guided by the structure and EMSA analysis, we further selected sites on the basis of the sequence at positions –6, +5 and +6 (**Fig. 5d**). Finally, we

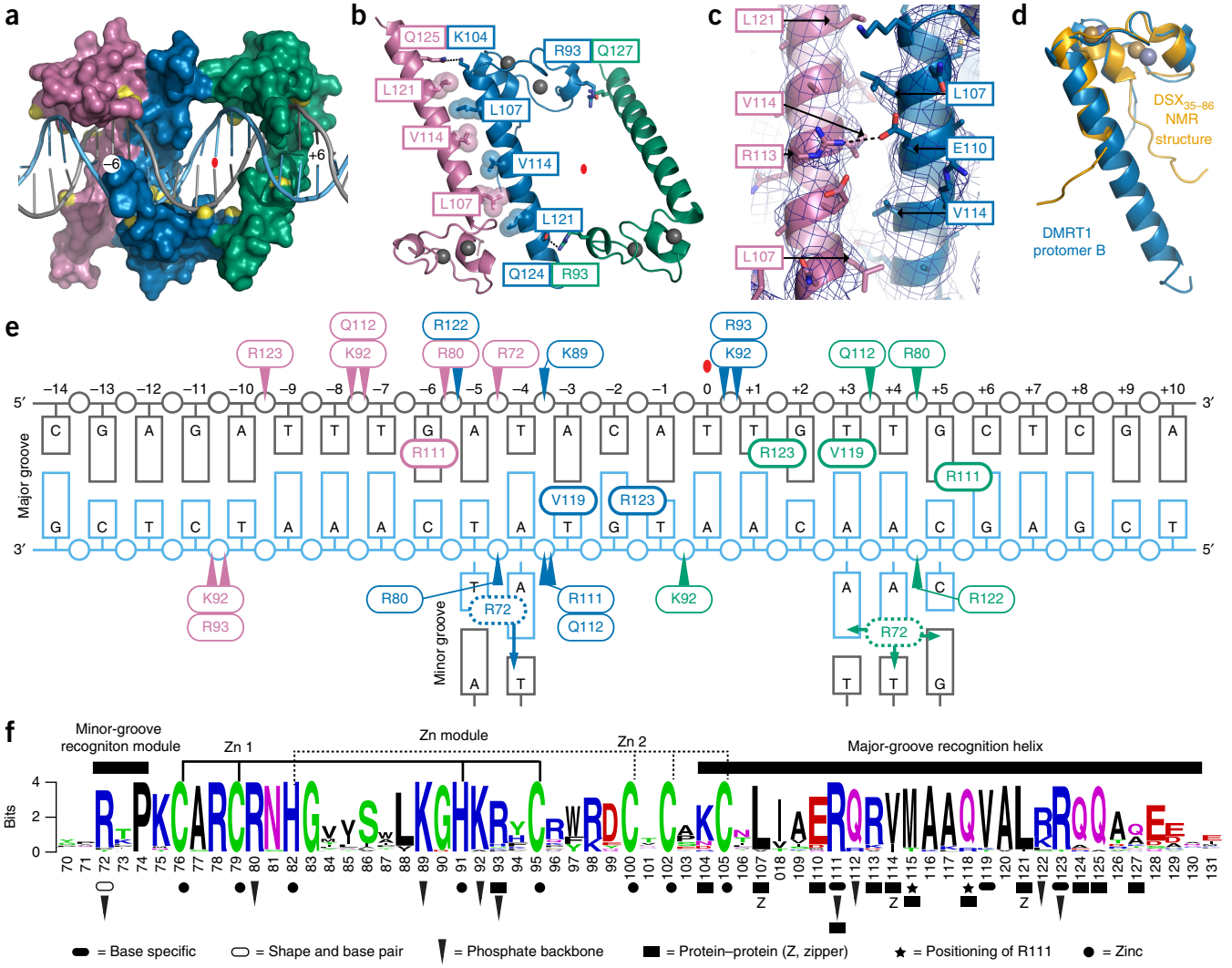


Figure 3 DNA backbone contacts, protein-protein interactions and binding summary. (**a**) Molecular surface of DMRT1$_{67–136}$ bound to DNA with charged groups contacting DNA phosphate backbone indicated in yellow. (**b**) Amino acids mediating protomer-protomer contacts. Interdigitating hydrophobic zipper and a Q-to-K hydrogen bond link protomers A and B. Two Q-to-R hydrogen bonds link B and C. (**c**) Close-up view of interaction between protomers A and B. Leucines and valines of interdigitating hydrophobic zipper are shown. In addition, R113 of protomer A and E110 of protomer B appear to form a salt bridge (dashed line). Blue mesh shows the σ_A-weighted $2F_o - F_c$ electron density map contoured at 1.0σ. (**d**) Overlay of DMRT1 protomer B structure with Dsx NMR structure[26], showing similar fold of zinc-binding domain. (**e**) Summary of DMRT1$_{67–136}$-DNA interactions. Colors indicate which protomer makes each contact. Thin-lined ovals with arrowheads identify amino acids that make DNA-backbone contacts, and thick-lined and dashed ovals identify amino acids that contact DNA bases. (**f**) Conservation of metazoan DM domains. Structural motifs and functional amino acids revealed by DMRT1 structural analysis are indicated for the region resolved by crystallography. Additional interactions could exist, particularly those bridged by water molecules. Amino acids are colored according to their chemical properties: green, polar (G,S,T,Y,C); blue, basic (K,R,H); red, acidic (D,E); black, hydrophobic (A,V,L,I,P,W,F,M); purple, neutral (Q,N).

plotted the ChIP-exo data in aggregate for each set of DMRT1-binding sites and compared the binding patterns, asking whether they differed and whether their differences were consistent with binding by each of the stoichiometries identified *in vitro* (**Fig. 5e**,**f**). Comparison of the compiled ChIP-exo data revealed a shared pattern on the left side in all three classes, as expected, but distinct patterns on the right. The predicted cross-linking patterns for each binding mode, according to the structure, conform well to the observed cross-linking patterns for the different groups of sites (**Fig. 5f** and **Supplementary Fig. 4a**). Predicted tetramers had symmetrical ChIP-exo patterns, whereas those of trimers and dimers were asymmetric, as expected. In trimers, the protomer C recognition helix sits at an angle in the DNA major groove that allows juxtaposition with more bases than in protomers A and B (**Fig. 2d**,**e**) and therefore has a higher density of potential cross-links (**Fig. 5f** and **Supplementary Fig. 4b**); consistently with this prediction, we observed stronger cross-links on the right side of the binding site, where protomer C would bind. Compilation of the selected DMRT1 consensus sequences did not reveal additional sequence or shape preferences, thus suggesting that the primary determinant of stoichiometry is the sequence and shape at the DMRT1-binding site rather than the presence of additional motifs (**Supplementary Fig. 4c**,**d**). Distinct patterns also were apparent in standard ChIP-seq, at lower resolution (**Supplementary Fig. 4a**). In summary, ChIP-exo suggests that DMRT1 binds as a tetramer, trimer or dimer *in vivo*, as *in vitro*, with the mode at each site determined by a combination of DNA sequence and shape.

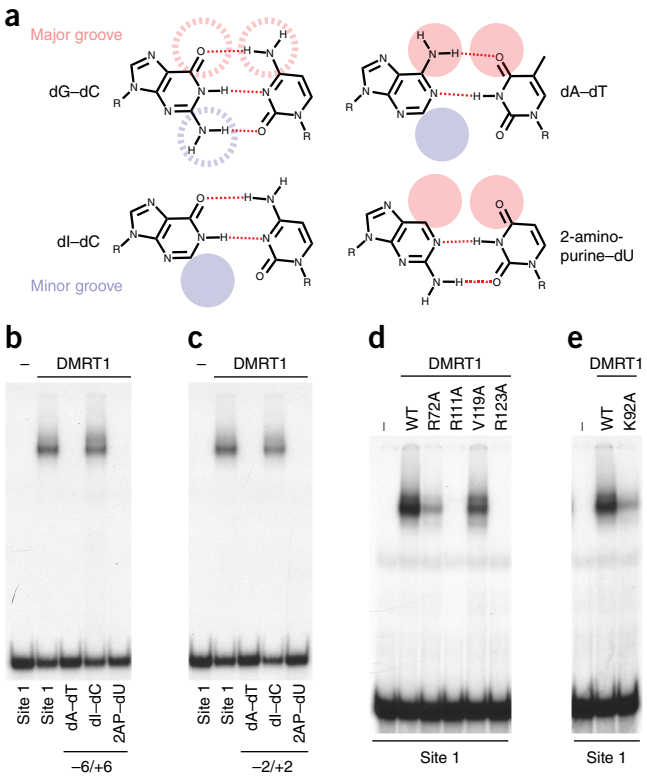


Figure 4 Confirmation of crucial protein-DNA contacts. (**a**–**c**) Confirmation of major-groove DNA contacts by chemical substitution *in vitro*. (**a**) Chemical structures of base pair analogs, with shaded circles indicating atoms altered in modified bases. (**b**) EMSA of DNAs modified at –6 and +6 positions, showing that major-groove, but not minor-groove, changes reduce DMRT1 binding. (**c**) EMSA of DNAs modified at –2 and +2 positions, showing that major-groove, but not minor-groove, modifications reduce binding. (**d**,**e**) Confirmation of critical protein contacts by amino acid substitution. (**d**) EMSAs showing effects of alanine substitution of DMRT1 amino acids making base contacts. (**e**) EMSA showing that substituting K92, which interacts with the DNA phosphate backbone, reduces DMRT1 DNA binding. WT, wild type.

Modeling suggests related binding modes for Dsx and MAB-3

Related but different DNA binding modes are probably used by the invertebrate sexual regulators Dsx and MAB-3. *In vitro*, Dsx and DMRT1 bind similar motifs, but Dsx has no sequence preference at –6 and +6 (**Fig. 6a**). EMSA confirmed that the –2 and +2 positions are important for binding of both Dsx and DMRT1, but –6 and +6 are important only for DMRT1 binding (**Fig. 6b**). This requirement for only the inner core of the binding motif suggests that Dsx binds as a symmetrical BB′-like dimer (modeled in **Fig. 6c**). *Caenorhabditis elegans* MAB-3 has tandem DM domains (**Supplementary Fig. 5**) and binds a site reminiscent of a DMRT1 half-site[30] (**Fig. 6a**). Molecular modeling suggests that the MAB-3 tandem DM domains might be equivalent to a DMRT1 AB dimer, with the truncated first recognition helix allowing looping so that both helices can bind adjacent to each other in the major groove (**Fig. 6d**).

DM-domain point mutations affect DNA binding in fly and human sex reversal

dsx determines sex in insects[7], and a number of *dsx* point mutations have been isolated that cause an intersex phenotype in *Drosophila*[5]. Most of these mutations alter residues required for zinc chelation, but one, R91Q, affects a recognition-helix residue equivalent to R123 in DMRT1 (**Supplementary Fig. 5**) and reduces DSX DNA binding[5]. We tested DMRT1^{R123Q} by EMSA and found that, like DMRT1^{R123A} (**Fig. 4a**), it eliminated DNA binding (**Fig. 7a** and **Supplementary Fig. 3c**). This result suggests that the Dsx^{R91Q} mutation disrupts a highly conserved sex-determining contact.

As discussed earlier, *DMRT1* determines gonadal sex in some vertebrates[2], but its role in human testis development has been less clear. In humans, primary XY male-to-female sex reversal results in female external genitalia and Mullerian structures (uterus and fallopian tubes) and undeveloped ('streak') gonads. This condition is also called 46,XY complete gonadal dysgenesis, or 46,XY CGD[31]. Human genetics has implicated DMRT1 as a key regulator of testis development: chromosome 9p deletions that remove one copy of *DMRT1* are associated with 46,XY feminization and gonadal dysgenesis, sometimes including 46,XY CGD[15,32]. Although they suggest that DMRT1 is haploinsufficient for testicular development, these deletions usually remove other genes, including the neighboring *DMRT2* and *DMRT3*. Also, most 9p deletions cause incomplete gonadal dysgenesis, so it has been unclear whether hemizygosity of *DMRT1* alone can cause full sex reversal. Although a *DMRT1* deletion removing exons 3 and 4, downstream of the DM domain, was found in a strongly feminized 46,XY individual[32], this deletion could have removed regulatory elements that affect other genes. Point mutations would help determine whether loss of *DMRT1* alone can cause sex reversal, but these have not been reported.

We therefore used exome resequencing to seek a DMRT1 point mutation. We were able to identify a 46,XY individual born fully feminized with complete gonadal dysgenesis (46,XY CGD) and carrying a heterozygous *de novo* point mutation (R111G) in the *DMRT1* recognition helix (**Fig. 7b**, **Supplementary Fig. 5a** and Online Methods). Genetic analysis found normal ploidy, and fluorescence *in situ* hybridization confirmed two copies of the regions containing *DMRT1* as well as the sex-determining genes *NR5A1*, *SOX9*, *WT1* and *DAX1* (official symbol *NR0B1*). No other potentially pathogenic mutations were apparent in the exome sequence, and the *DMRT1* mutation was not present in 240 ancestry-matched control individuals. Full details of the clinical and genetic characterization of this subject are provided in Online Methods. We conclude that the *de novo DMRT1*^{R111G} mutation is the most likely cause of the complete gonadal dysgenesis and 46,XY sex reversal in this subject. To our knowledge, this is the first human *DMRT1* point mutation associated with 46,XY sex reversal. The phenotype is very similar to that caused by mutations in the testis-determining gene *SRY*[33], and it strongly suggests that *DMRT1* is required for human sex determination.

We next examined the DNA binding properties of DMRT1^{R111G} and found that the mutant protein had strongly reduced DNA affinity, similar to DMRT1^{R111A} (**Figs. 7a** and **4d** and **Supplementary Fig. 3a**,**c**). In the structure, R111 of protomer C interacts with the +5 and +6 positions of site 1 (**Fig. 7g**). We found that DMRT1^{R111G} had altered sequence specificity: it bound a site with –6 and +6 dG-dC to dA-dT substitutions weakly but better than wild-type DMRT1 (**Fig. 7c**). Moreover, in an EMSA assay, when mixed with wild-type DMRT1, the mutant protein could promote tetramer binding on site 1, which normally is bound by trimers of wild-type DMRT1 (**Fig. 7d**). We also tested binding of DMRT1^{R111G} to *in vivo* DMRT1-binding sites from the *Sox9* gene (activated by DMRT1) and the *Foxl2* gene (repressed by DMRT1)[16,17]. The *Sox9* site is bound as a trimer by wild-type protein (**Fig. 7e**). DMRT1^{R111G} bound this site very weakly, but, when mixed with wild-type protein, it shifted the complex to a tetramer with much higher affinity. The *Foxl2* site was bound as a tetramer by wild-type

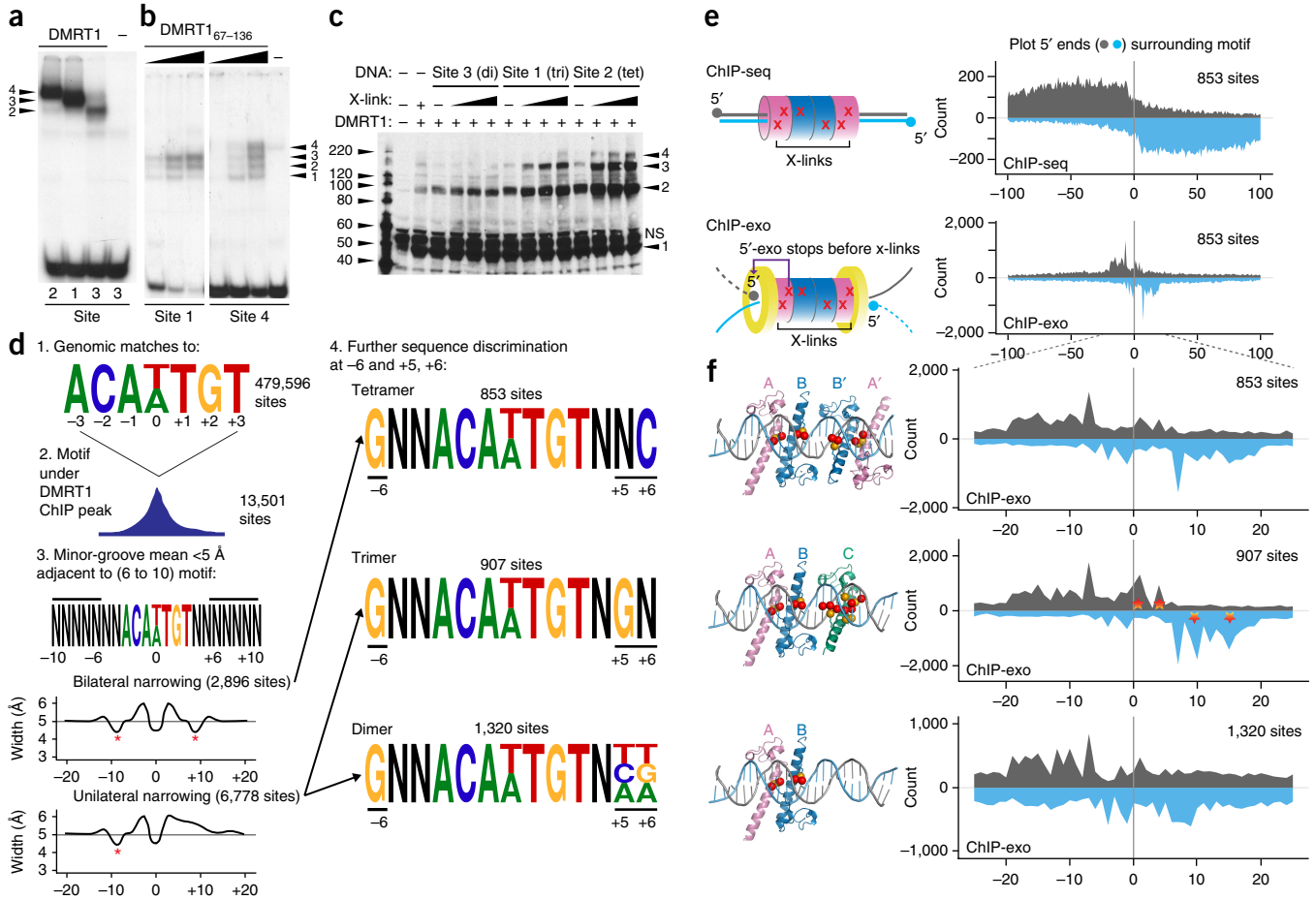


Figure 5 DMRT1 binds DNA with multiple stoichiometries *in vitro* and *in vivo*. (**a**) EMSA showing binding of DMRT1 tetramer, trimer and dimer to sites 2, 1 and 3, respectively. (**b**) EMSA of *in vitro*–translated SUMO-DMRT1$_{67–136}$ binding to sites 1 and 4, showing monomer-through-tetramer binding. (**c**) Protein cross-linking (x-link) showing interaction of DMRT1 to form higher-order complexes. Di, dimer; tri, trimer; tet, tetramer; NS, nonspecific band. (**d**) Workflow for evaluating DMRT1 binding stoichiometry by ChIP-exo. Sites under ChIP peaks were grouped on the basis of bilateral or unilateral minor-groove narrowing and sequence at positions –6, +5 and +6 (Online Methods). (**e**) Left, diagrams comparing ChIP-seq and ChIP-exo. Cross-linked protein blocks exonuclease digestion in ChIP-exo within several bases of the cross-link. Colors indicate binding by different protomers; 'X' illustrates potential cross-links. Right, compilation of 5′ ends of ChIP-seq (top) and ChIP-exo (bottom) reads aligned to DMRT1 tetramer-binding consensus, showing higher resolution of ChIP-exo. (**f**) ChIP-exo analysis of DMRT1 binding in mouse testes at sites sorted as indicated in **d**, showing three distinct patterns. Structural diagrams interpret DMRT1 binding modes on the basis of ChIP-exo patterns and indicate potential cross-links (red balls, cross-linkable DNA residues; yellow balls, cross-linkable protein residues). Stars highlight prominent differences in ChIP-exo pattern in putative trimer-binding sites relative to tetramer-binding sites.

DMRT1, and addition of DMRT1^{R111G} had little or no effect on binding (**Supplementary Fig. 6**). From the ability of DMRT1^{R111G} to alter binding stoichiometry of wild-type DMRT1 on a biologically relevant site *in vitro*, we suggest that the DMRT1^{R111G} mutation may combine severe loss of function and/or haploinsufficiency with a dominant disruption of normal binding stoichiometry at some DMRT1-binding sites. This combination of haploinsufficiency and dominant disruption may explain the severe phenotype caused by *DMRT1*^{R111G} heterozygosity. In the structure, R111 of protomer A is positioned to contact –6 by its own M115 and by M115 and Q118 of protomer B (**Fig. 7f**,**h**). We found that mutating these residues also reduced DNA binding (**Fig. 7a**). In summary, the severe effects of the DMRT1^{R111G} point mutation on DNA binding and its association with 46,XY male-to-female sex reversal strongly suggest that *DMRT1* has a role in human primary sex determination and identify another deeply conserved molecular interaction crucial for metazoan sexual development.

DISCUSSION

We have undertaken a structural analysis of DMRT protein–DNA interaction. We used ChIP-seq to define the DNA binding preference of DMRT1 in mice and humans and then used X-ray crystallography to determine a DMRT1–DNA structure. The structure revealed that binding of the human DMRT1 DM domain to DNA involves the recognition of specific bases primarily in the DNA major groove. We confirmed this finding by using chemical substitutions that selectively altered the major or minor groove of the DNA at key base pairs. A previous report concluded, on the basis of DNA substitutions, that Dsx binds DNA primarily via the minor groove[26]. From our structural analysis, it is apparent that the minor-groove modifications that reduced binding probably limited the ability of the major groove to expand and accommodate the DM-domain recognition helix rather than affecting sequence-specific base contacts; thus, the previous data are in accord with our structure.

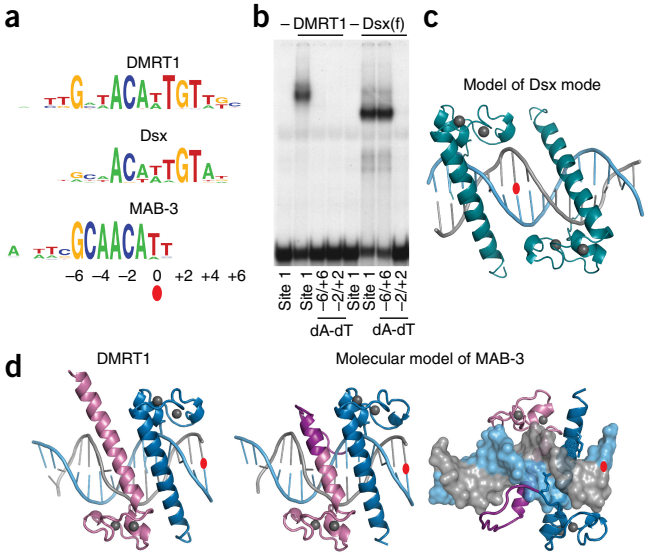


Figure 6 Modeling DNA interaction by Dsx and MAB-3 suggests related but different binding modes. (**a**) *In vitro* DNA binding motifs for DMRT1, Dsx and MAB-3, showing that the Dsx site[30] is symmetrical but lacks selection at –6 and +6 positions, whereas the MAB-3 motif[30] resembles the left side of the DMRT1 motif. (**b**) EMSA showing that binding by the female Dsx isoform Dsx(f) requires specific DNA base pairs at the –2 and +2 but not the –6 and +6 DNA positions, in agreement with the *in vitro* consensus. (**c**) Docking model of DMRT1 binding as a dimer to a previously determined Dsx-binding-site DNA structure[37], illustrating likely Dsx-binding mode. (**d**) A model of proposed interaction of MAB-3 DM domains with DNA, illustrating binding of MAB-3 as a covalently joined 'internal dimer'. MAB-3 (center and right) is proposed to form a structure on its consensus element similar to DMRT1 protomers A and B bound to the left side of the DMRT1 consensus element (left). The first DM domain of MAB-3 (DMa) is predicted to have a truncated recognition helix, with the remainder forming a linker joining DMa to DMb (**Supplementary Fig. 5**).

Binding of DMRT1 to DNA has two particularly noteworthy features. First, binding involves the insertion of paired recognition helices together into a widened DNA major groove. To our knowledge, this is the only example of two closely neighboring α-helices inserting into the same section of a major groove. Second, DMRT1 can bind DNA by using different stoichiometries. The basis of this versatility is that binding involves a small number of amino acid side chains that can make distinct sets of DNA interactions. As a result, different DNA sites can bind distinct configurations of protomers, ranging from dimers to tetramers. ChIP-exo analysis suggests that DMRT1 also binds *in vivo* with differing stoichiometries. Our ability to predict stoichiometry on the basis of DNA sequence preference and conformation (**Fig. 5**) suggests that the stoichiometry at a specific

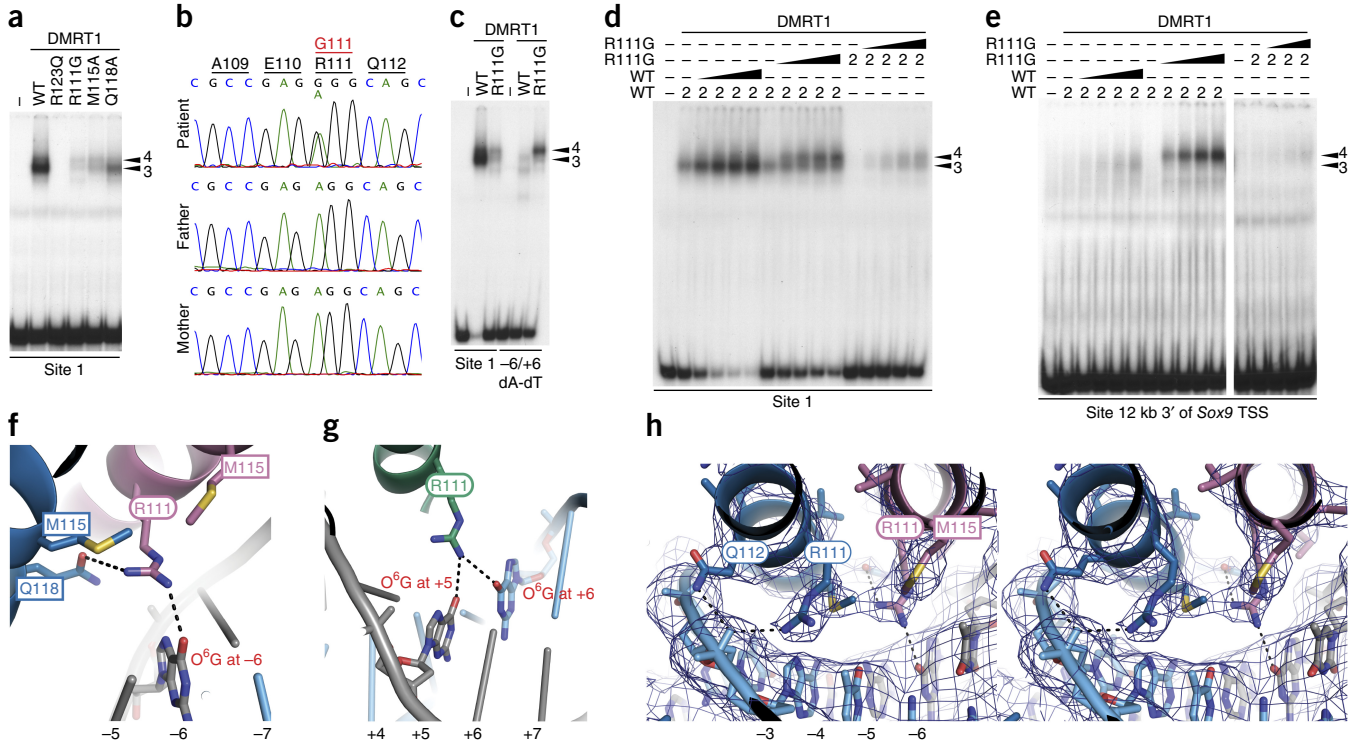


Figure 7 Disruption of crucial DMRT1 DNA contacts by a sex-reversing human mutation. (**a**) EMSA showing importance of R123, R111 and three amino acids that position R111 of protomer A for sequence-specific DNA interaction. (**b**) Sanger-sequencing chromatograms showing *de novo* A-to-G heterozygous sequence change causing an R111G coding change in a 46,XY female. (**c**) EMSA showing reduced binding affinity and altered binding specificity of the R111G mutation. (**d**) EMSA comparing DMRT1 and DMRT1^{R111G}. Left lanes, wild-type DMRT1 alone; middle, DMRT1 plus DMRT1^{R111G}; right, DMRT1^{R111G} alone. EMSAs contained 2 μl *in vitro*–translated protein; wedges indicate added increments up to a total of 6 μl. DMRT1^{R111G} can convert wild-type trimers on site 1 into slower-migrating tetramers (arrowheads), probably by occupying the right side of the binding site. (**e**) As in **d**, but with DMRT1-binding site from the *Sox9* gene. (**f**) Structure showing positioning of protomer A R111 (pink) to allow hydrogen-bonding with –6 guanine by M115 and Q118 of protomer B (blue) and M115 of protomer A (pink). Methionine methyl groups make van der Waals contacts with each other and R111. (**g**) Structure showing recognition of +5 or +6 by protomer C R111. (**h**) Walleyed stereo view of protomer A R111 interacting with –6G and protomer B R111 interacting with DNA backbone. Blue mesh, σ_A-weighted $2F_o - F_c$ electron density map contoured at 1.5σ.

binding site is determined largely by the sequence and shape of that site. A key remaining question is what biological significance the DMRT1 binding stoichiometries may have. Possibilities include association with transcriptional activation or repression, binding to different classes of regulatory elements (e.g., promoters or enhancers) or interaction with other regulatory proteins. We previously found that, in agreement with the third possibility, a subset of DMRT1-binding sites contain overlapping GATA1 and SOX9 consensus elements[34]. Distinguishing among these possibilities is an important goal, but this will require cell type–specific approaches, because DMRT1 has cell type–specific functions in germ cells and Sertoli cells.

Although a number of deletions removing part or all of DMRT1 have been found in people with 46,XY sex reversal, the DMRT1^{R111G} mutation that we report here is, to our knowledge, the only DMRT1 mutation shown to affect an essential functional domain. The severely reduced DNA binding affinity of DMRT1^{R111G} combined with the complete sex reversal and gonadal dysgenesis of the subject suggest that *DMRT1* plays a part in human sex determination. Our finding that the mutant protein can interfere with the binding stoichiometry of wild-type DMRT1 further suggests that the mutant protein may behave at least partially as a dominant negative. A dominant effect of DMRT1^{R111G} may help explain why the phenotype of this point mutation is more severe than those of most 9p deletions that completely remove *DMRT1*. The highly specific nature of point mutations such as *DMRT1*^{R111G} that can alter function of the remaining wild-type allele also may explain why *DMRT1* point mutations able to cause sex reversal are so rare. Because we observed reduced DNA binding specificity, we cannot exclude the possibility that DMRT1^{R111G} also binds and misregulates genes that are not normally controlled by DMRT1. An animal model of the DMRT1^{R111G} may help elucidate the *in vivo* effects of this mutation.

In summary, we have obtained a detailed view of how DMRT proteins recognize and associate with target DNA. We have defined crucial conserved atomic interactions that mediate DNA binding and found that these are required for sex determination in flies and humans. DMRT proteins have directed metazoan sexual differentiation for hundreds of millions of years[2,3]. Reproduction is the crucible of natural selection[35], and the long-term involvement of *DMRT* genes in sexual development suggests that they have substantially shaped metazoan evolution.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession code 4YJ0. Sequencing data have been deposited in the Gene Expression Omnibus repository under accession code GSE64892.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank K. Shi and J. Nix for help with crystallization trials and data collection; the University of Minnesota Supercomputing Institute for computational resources; L. Amble, the University of Minnesota Tissue Procurement Facility and anonymous donors for human testis tissue; D. Greenstein, C. Kim, M. Slattery, B.F. Pugh, J. Simon, H. Towle and members of our laboratories for helpful comments on the manuscript; and T. Gamble for help with phylogenetic analysis. X-ray data were collected at the Advanced Photon Source (APS) NE-CAT beamlines, which are supported by the US National Institute of General Medical Science (P41 GM103403). APS is a US Department of Energy Office of Science User Facility operated by Argonne National Laboratory under contract DE-AC02-06CH11357. This work was funded by the US National Institutes of Health (GM59152 and GM50399 to D.Z.; AI087098 and GM095558 to H.A.), European Cooperation in Science and Technology (COST) (Action DSDnet BM1303 to K.M.) and Program Blanc Assistance-Publique-Institut Pasteur (to K.M.).

AUTHOR CONTRIBUTIONS

M.W.M. performed and, with V.J.B., D.Z. and M.D.G. analyzed *in vitro* and *in vivo* DNA binding studies. M.D.G. performed bioinformatic analysis of ChIP data. J.K.L., K.K. and H.A. performed protein purification and crystallization. J.K.L., S.B. and H.A. collected X-ray diffraction data. J.K.L. processed the X-ray data and built and refined the atomic model. M.W.M., J.K.L., M.D.G., D.Z., H.A. and V.J.B. analyzed the structure and prepared the figures. G.-A.L. coordinated the clinical studies. A.B. and K.M. designed the human genetic studies and with S.R. analyzed the exome data sets. A.B. and S.R. performed Sanger sequencing. D.Z. and V.J.B. wrote the manuscript. M.W.M., M.D.G., H.A., A.B., K.M. and S.R. edited the manuscript. The first four authors made equivalent contributions.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Gamble, T. & Zarkower, D. Sex determination. *Curr. Biol.* **22**, R257–R262 (2012).
2. Matson, C.K. & Zarkower, D. Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. *Nat. Rev. Genet.* **13**, 163–174 (2012).
3. Kopp, A. Dmrt genes in the development and evolution of sexual dimorphism. *Trends Genet.* **28**, 175–184 (2012).
4. Raymond, C.S. *et al.* Evidence for evolutionary conservation of sex-determining genes. *Nature* **391**, 691–695 (1998).
5. Erdman, S.E. & Burtis, K.C. The *Drosophila* doublesex proteins share a novel zinc finger related DNA binding domain. *EMBO J.* **12**, 527–535 (1993).
6. Chong, T., Collins, J.J. III, Brubacher, J.L., Zarkower, D. & Newmark, P.A. A sex-specific transcription factor controls male identity in a simultaneous hermaphrodite. *Nat. Commun.* **4**, 1814 (2013).
7. Baker, B.S. & Ridge, K.A. Sex and the single cell. I. On the action of major loci affecting sex determination in *Drosophila melanogaster*. *Genetics* **94**, 383–423 (1980).
8. Shen, M.M. & Hodgkin, J. mab-3, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*. *Cell* **54**, 1019–1031 (1988).
9. Raymond, C.S., Murphy, M.W., O'Sullivan, M.G., Bardwell, V.J. & Zarkower, D. Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. *Genes Dev.* **14**, 2587–2595 (2000).
10. Smith, C.A. *et al.* The avian Z-linked gene *DMRT1* is required for male sex determination in the chicken. *Nature* **461**, 267–271 (2009).
11. Lambeth, L.S. *et al.* Over-expression of DMRT1 induces the male pathway in embryonic chicken gonads. *Dev. Biol.* **389**, 160–172 (2014).
12. Matsuda, M. *et al. DMY* is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature* **417**, 559–563 (2002).
13. Yoshimoto, S. *et al.* A W-linked DM-domain gene, DM-W, participates in primary ovary development in *Xenopus laevis*. *Proc. Natl. Acad. Sci. USA* **105**, 2469–2474 (2008).
14. Veitia, R. *et al.* Deletions of distal 9p associated with 46,XY male to female sex reversal: definition of the breakpoints at 9p23.3-p24.1. *Genomics* **41**, 271–274 (1997).
15. Tannour-Louet, M. *et al.* Identification of *de novo* copy number variants associated with human disorders of sexual development. *PLoS ONE* **5**, e15392 (2010).
16. Matson, C.K. *et al.* DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature* **476**, 101–104 (2011).
17. Lindeman, R.E. *et al.* Sexual cell-fate reprogramming in the ovary by DMRT1. *Curr. Biol.* **25**, 764–771 (2015).
18. Zhao, L., Svingen, T., Ng, E.T. & Koopman, P. Female-to-male sex reversal in mice caused by transgenic overexpression of Dmrt1. *Development* **142**, 1083–1088 (2015).
19. Murphy, M.W., Zarkower, D. & Bardwell, V.J. Vertebrate DM domain proteins bind similar DNA sequences and can heterodimerize on DNA. *BMC Mol. Biol.* **8**, 58 (2007).
20. Schmidt, D. *et al.* Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* **328**, 1036–1040 (2010).
21. Cheng, Y. *et al.* Principles of regulatory information conservation between mouse and human. *Nature* **515**, 371–375 (2014).
22. Rohs, R. *et al.* The role of DNA shape in protein-DNA recognition. *Nature* **461**, 1248–1253 (2009).
23. Rohs, R. *et al.* Origins of specificity in protein-DNA recognition. *Annu. Rev. Biochem.* **79**, 233–269 (2010).
24. Seeman, N.C., Rosenberg, J.M. & Rich, A. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. USA* **73**, 804–808 (1976).

25. Slattery, M. *et al.* Absence of a simple code: how transcription factors read the genome. *Trends Biochem. Sci.* **39**, 381–399 (2014).
26. Zhu, L. *et al.* Sexual dimorphism in diverse metazoans is regulated by a novel class of intertwined zinc fingers. *Genes Dev.* **14**, 1750–1764 (2000).
27. Rhee, H.S. & Pugh, B.F. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* **147**, 1408–1419 (2011).
28. Starick, S.R. *et al.* ChIP-exo signal associated with DNA-binding motifs provide insights into the genomic binding of the glucocorticoid receptor and cooperating transcription factors. *Genome Res.* doi:10.1101/gr.185157.114 (26 February 2015).
29. Chiu, T.P. *et al.* GBshape: a genome browser database for DNA shape annotations. *Nucleic Acids Res.* **43**, D103–D109 (2015).
30. Yi, W. & Zarkower, D. Similarity of DNA binding and transcriptional regulation by *Caenorhabditis elegans* MAB-3 and *Drosophila melanogaster* DSX suggests conservation of sex determining mechanisms. *Development* **126**, 873–881 (1999).
31. Ostrer, H. in *GeneReviews* (eds. Pagon, R.A. *et al.*) (2009).
32. Ledig, S., Hiort, O., Wunsch, L. & Wieacker, P. Partial deletion of DMRT1 causes 46,XY ovotesticular disorder of sexual development. *Eur. J. Endocrinol.* **167**, 119–124 (2012).
33. Jäger, R.J., Anvret, M., Hall, K. & Scherer, G. A human XY female with a frame shift mutation in the candidate testis-determining gene *SRY*. *Nature* **348**, 452–454 (1990).
34. Murphy, M.W. *et al.* Genome-wide analysis of DNA binding and transcriptional regulation by the mammalian Doublesex homolog DMRT1 in the juvenile testis. *Proc. Natl. Acad. Sci. USA* **107**, 13360–13365 (2010).
35. Darwin, C.R. *The Descent of Man, and Selection in Relation to Sex* (John Murray, 1871).
36. Zhou, T. *et al.* DNAshape: a method for the high-throughput prediction of DNA structural features on a genomic scale. *Nucleic Acids Res.* **41**, W56–W62 (2013).
37. Narayana, N. & Weiss, M.A. Crystallographic analysis of a sex-specific enhancer element: sequence-dependent DNA structure, hydration, and dynamics. *J. Mol. Biol.* **385**, 469–490 (2009).

ONLINE METHODS

Vertebrate animals. Experimental protocols were approved by the University of Minnesota Institutional Animal Care and Use Committee. Mice were adult males of mixed C57BL/6J and 129S1 genetic background. No statistical method was used to predetermine sample size. The experiments were not randomized and were not performed with blinding to the conditions of the experiments.

Figure preparation. Figures were prepared with Adobe Photoshop, Adobe Illustrator, and PyMOL (http://www.pymol.org/).

ChIP and ChIP-exo. Chromatin immunoprecipitation (ChIP) was performed as previously described[34] except that tissue was disaggregated with a Virtis Virtishear homogenizer (225318) in phosphate-buffered saline (PBS) containing 1% paraformaldehyde. Sonication times were extended to allow for smaller average-size products suitable for Illumina sequencing. Cross-links were reversed overnight at 55 °C. Illumina sequencing libraries were prepared according to the manufacturer's protocol except that end polishing of the ChIP fragments was by DNA terminator (Lucigen), and adapters were diluted 1:50 before ligation. For ChIP-exo, chromatin precipitation was performed as above and before elution of complexes from protein A–Sepharose beads. ChIP-exo libraries were prepared as previously described[27] except that primer sequences were modified to be compatible with the Illumina sequencing platform.

Primers for ChIP-exo library preparation are as follows: P2 adaptor, 5′-P-acactctttccctacacgacgctcttccgatct-3′ annealed to 5′-agatcggaagagcgtcgtgtag-3′. Primer extension oligonucleotide, acactctttccctacacgac. P1 adaptor, gtgactggagttcagacgtgtgctcttccgatct annealed to agatcggaagagcacacgtctg. PCR amplification was performed with primers P1 and P2 P1, caagcagaagacggcatacgagatc**gtgat**gtgactggagttcagacgtgtgc. P2, aatgatacggcgaccaccgagatctacactctttccctacacgacgc. Bold text in P1 indicates the sequence that was varied for multiplexing.

Human tissue for ChIP. Fresh testicular tissue from an orchiectomy was provided by the University of Minnesota Tissue Procurement Facility under IRB supervision and with informed consent. Normal histology of Bouin's fixed subsamples was confirmed by hematoxylin/eosin staining of paraffin-embedded sections.

DNA binding substrates. EMSAs to assay stoichiometry of DMRT1 binding used shorter (27–base pair) DNA duplexes for better-resolution of complexes. Sequences are as follows: site 1, top strand, 5′-gagatttgatacattgttgctcgatgg-3′. Site 2, top strand, 5′-gagatttgatacattgttactttatgg-3′. Site 3, top strand, 5′-gagatttgatacattattaatttatgg-3′. Site 4, top strand, 5′-ttgctatgatacattgtatcttgctgg-3′. *Sox9* site, top strand, 5′-gtggctgggcaccctgcagagacaatgtttccagctgcaggtcaggtct-3′. *Foxl2* site, top strand, 5′-gtggctgggcacaactctgtaacattgtttccaaggggaggtcaggtct-3′.

EMSA to evaluate effects of mutant DNA and protein on binding used longer (49–base pair) DNA duplexes based on the site 1 DNA duplex: 5′-GTGGCTGGGCAgagatttgatacattgttgctcgatggAGGTCAGGTCT-3′. Mutations were incorporated into hDMRT1 by overlap-extension PCR[38] with a T7-hDMRT1 (pDZ142) plasmid clone as template. The mutated products were subcloned back into pDZ142 and translated *in vitro* with the TNT Quick Coupled transcription/translation system (Promega).

***In vitro* DNA binding.** EMSA was performed as previously described[19] except that substrates were end labeled with T4 polynucleotide kinase (NEB). DNase I footprint analysis was performed with highly purified bacterially expressed $hDMRT1_{67–136}$ protein as previously described[39] except that after the DNase I–digestion step, the sample was phenol/chloroform extracted to remove protein before precipitation.

Protein cross-linking. Proteins were *in vitro* translated as for EMSA. Complexes were formed under the same conditions as for EMSA except with five times as much DNA at room temperature for 10 min before addition of glutaraldehyde to 0.0075% final concentration. Cross-linking was stopped at the indicated times by the addition of glycine to 0.125 M final concentration. Complexes were resolved on 4–12% NuPAGE Novex Bis-Tris mini gels (Invitrogen), and DMRT1 was detected by immunoblotting with a previously validated antibody[34].

X-ray crystallography. $hDMRT1_{67–136}$ was expressed as a SUMO fusion in *E. coli* Rosetta2 (DE3) and purified by metal-affinity chromatography. The His_6-tagged SUMO was removed by cleavage with the SUMO protease Ulp1. To form the protein–DNA complex, purified $hDMRT1_{67–136}$ (~1 mM) in 20 mM Tris-HCl, pH 7.4, 0.2 M NaCl, 10 μM $ZnCl_2$, and 2 mM β-mercaptoethanol, was mixed with a blunt-ended 25-bp target DNA (site 1, 5′-CGAGATTTGATACATTGTTGCTCGA-3′, and its complement) at a protein/DNA molar ratio of 2:1. The complex (~0.6 mM protein) was crystallized at 20 °C by the hanging-drop vapor-diffusion method with a reservoir solution (100 mM Bis-Tris, pH 6.5, 4–12% polyethylene glycol 3350, 4–10% 2-methyl-2,4-pentanediol (MPD), and 2–10 mM dithiothreitol). Crystals containing SeMet-labeled protein or the 5-BrdU–labeled oligonucleotide were grown under conditions similar to that for the native complex. The crystals were transferred in a stepwise fashion to the reservoir solution with increasing concentrations of glycerol (final concentration of 15%) and were flash cooled in liquid nitrogen for X-ray data collection.

X-ray diffraction data were collected at the Advanced Photon Source Northeastern Collaborative Access Team beamlines (24-ID-C/E) and the Advanced Light Source Molecular Biology Consortium (4.2.2) beamlines and processed with RAPD (https://rapd.nec.aps.anl.gov/rapd/), HKL2000 (ref. 40) or XDS[41]. X-ray wavelengths corresponding to the K-absorption edge of Zn, Se, and Br were used, respectively, for the native, SeMet-, and 5-Br-dU–labeled crystals. The structure was determined by SAD phasing with a 3.81-Å-resolution data set from a native crystal (**Table 1**) with PHENIX[42]. Six zinc sites were found, from which the structure-factor phases were calculated with a mean figure of merit of 0.66. The atomic model was built in COOT[43] and refined with REFMAC[44] and PHENIX, with Ramachandran and DNA restraints to maintain geometries for protein, DNA base pairs, and base stacking. The model building was facilitated by the Se and Br anomalous-difference Fourier peaks, which showed the correct register of amino acids and nucleotides, respectively. Protein residues with poor side chain electron density were modeled as alanines. The Ramachandran plot for the final model was generated by MolProbity[45] and showed 89.0%, 8.2%, and 2.8% of residues in favored, allowed, and outlier regions, respectively. The DNA structure was analyzed with 3DNA[46], the minor-groove width of the unbound DNA was estimated with DNAshape[36], and the molecular-graphics images were produced with PyMOL (http://www.pymol.org/).

Identification of a subject with $DMRT1^{R111G}$ mutation. *Overview of study.* The study was approved by the ethical board of Institut Pasteur (RBM 2003/8), and informed consent was obtained. Ancestry was determined by self-reporting, from responses to a personal questionnaire, which asked questions pertaining to the birthplace, languages and ethnicity of the participants, their parents and their grandparents. The control panel consisted of 240 unrelated 46,XY males of French ancestry who are either normospermic or have fathered at least two children and have no history of testicular anomalies (determined by self-reporting). All samples used for this study were collected with proper informed consent. Sequencing of the coding region of the *DMRT1* gene was performed as described previously[47].

Whole-exome sequencing. Exon enrichment was performed with Agilent SureSelect Human All Exon V4. Paired-end sequencing was performed on the Illumina HiSeq2000 platform with TruSeq v3 chemistry. Read files (fastq) were generated from the sequencing platform via the manufacturer's proprietary software. Reads were mapped with the Burrows-Wheeler Aligner[48], and local realignment of the mapped reads around potential insertion/deletion (indel) sites was carried out with GATK version 1.6 (ref. 49). Duplicate reads were marked with Picard version 1.62 (http://broadinstitute.github.io/picard/). Additional BAM file manipulations were performed with Samtools (0.1.18)[50]. SNP and indel variants were called with the GATK Unified Genotyper for each sample. SNP novelty was determined against dbSNP138. Novel variants were analyzed by a range of web-based bioinformatics tools with the EnsEMBL SNP Effect Predictor (http://www.ensembl.org/homo_sapiens/tools/VEP/). All variants were screened manually against the Human Gene Mutation Database Professional Biobase (http://www.biobase-international.com/product/hgmd/). *In silico* analysis was performed to determine the potential pathogenicity of the variants. Potentially pathogenic mutations were verified with classic Sanger sequencing.

Characterization of subject. The subject has two healthy brothers and a sister. A routine fetal karyotype was performed as part of a protocol for pregnancy with advanced maternal age. The karyotype was 46,XY, whereas the ultrasound showed a completely female fetus. The baby, born by cesarean section, was completely feminine. At day 1, serum testosterone levels were 57 ng/dl; dihydrotestosterone,

12 ng/dl; adrenocorticotropic hormone (ACTH), 70.8 ng/ml; and anti-Müllerian hormone (AMH), 0.02 ng/ml. The hormonal profile was consistent with gonadal dysgenesis. At 3 months of age, serum LH levels were 0.08 UI/l; FSH, 15.1 UI/l; inhibin B, <15 ng/ml; AMH, <0.15 ng/ml; testosterone, <0.05 ng/dl; androstenedione, 11 ng/dl; and ACTH, 11 ng/dl. At this time, ultrasound revealed the presence of a uterus and an apparent absence of gonads. At 18 months, gonadectomy revealed bilateral streak gonads with a gonadoblastoma on the right side. Histology of the gonads revealed ovarian-like stroma with no evidence of any testicular material. The diagnosis was 46,XY complete gonadal dysgenesis.

Genetic analysis. At three months of age, the karyotype of the subject on peripheral blood lymphocytes was 46,XY (50 cells). FISH analysis on lymphocyte spreads indicated two copies of the regions 9p24 (*DMRT1*), 9q22 (*NR5A1*), 11p13 (*SOX9*), 17q24 (*WT1*) and Xp21 (*DAX1*). Direct sequencing of the *SRY* and *NR5A1* genes revealed wild-type sequences. Array comparative genomic hybridization with the Agilent 44k platform confirmed a normal ploidy in the subject.

Whole-exome sequencing was performed on the parents and subject. The numbers of paired-end reads were 24,328,671 (father), 26,580,579 (mother) and 20,566,862 (child), with a mean coverage of 61.09, 68.01 and 53.41, respectively. The percentage of target bases with >10× coverage was 96.26%, 96.97% and 95.59%. The number of variants with predicted serious consequences (involving an essential splice site, a stop codon gained or lost, a complex indel, a frameshift mutation in the coding sequence or a nonsynonymous change with predicted deleterious effect on protein function) for the father, mother and child was 11,930, 11,944 and 11,758, respectively.

Analyses of the data sets revealed several *de novo* mutations that were predicted by PolyPhen2 (ref. 51) and/or SIFT[52] to be deleterious substitutions for protein function. These were the new heterozygous mutations c.644A>G (p.Glu215Gly) in *C2CD4C* (ENST00000332235), c.1309G>A (p.Glu437Lys) in *CEP104* (ENST00000378230), c.761G>C (p.Ala761Pro) in *DLGAP3* (ENST00000235180), c.331A>G (p.Arg111Gly) in *DMRT1* (ENST00000382276), c.3779G>C (p.Gly1260Ala) in *HSPG2* (ENST00000374695), c.58C<G (p.Leu20Val) in *MECR* (ENST00000263702) and c.560G>A (p.Gly187Asp) in *MPST* (ENST00000397225).

Assuming a recessive or X-linked model of inheritance and after filtering to remove variants with a minor allelic frequency of 0.05, there was only a single remaining gene with a serious mutation. This was a hemizygous c.262G>A (p.Arg88Trp) mutation in the X-chromosome gene *MID2* (ENST00000262843). This variant has previously been reported (rs375584547) with an allelic frequency of 1:6,727 in individuals of European-American ancestry. With the exception of *DMRT1*, there is an absence of a clear functional relationship between the variants in these genes and the absence of testis formation seen in the subject.

The *de novo* missense mutation in *DMRT1* was confirmed by direct sequencing of the *DMRT1* gene. No other potentially pathogenic mutations were identified in the exome-sequencing data set in other genes known to cause 46,XY gonadal dysgenesis (for example, *WT1*, *NR5A1* and *SRY*). The *DMRT1* mutation was not observed in direct sequence analysis of 240 unrelated ancestry-matched control individuals.

Bioinformatics workflow. ChIP libraries were sequenced on the Illumina HiSeq2000 platform with 50 cycles single-end (ChIP-seq) or 50 cycles paired-end (ChIP-exo) reads. Illumina fastq files were adaptor trimmed (Trimmomatic 0.32)[53] and mapped (Burrows-Wheeler Aligner – Mem, 0.7.10–r789)[48] to the mm9 or hg19 genome builds. Samtools (0.1.18)[50] was used to sort and convert aligned reads and to extract the R2 mate pairs in the paired-end ChIP-exo data set. Peaks with a *P* value less than 0.05 were identified in the ChIP-seq data sets (MACS 2.1.0.20140616)[54]. Peaks present in both replicates available for the mouse ChIP-seq were used for further analysis. The peaks that scored in the top sextile of the mouse and human peak lists contained 8,571 and 7,593 peaks respectively. 100 bp of DNA sequence surrounding the summits of these top peaks was analyzed by MEME (4.10.0)[55] to identify enriched sequence motifs. Genomic regions with DNA sequences compatible with dimer-, trimer- and tetramer-binding modes found underneath ChIP-seq peaks were used to aggregate read counts from the 5′ positions of the ChIP-exo reads with RSamtools (1.18.2; http://bioconductor.org/packages/release/bioc/html/Rsamtools.html/) and GenomicRanges (3.0)[56] with custom R (3.1.1; http://www.r-project.org/) scripts available at https://github.com/micahgearhart/exotools/.

Analysis of ChIP-exo stoichiometry. Nonoverlapping 51–base pair mouse genomic locations centered around the sequence ACA(A/T)TGT were identified with the BioStrings package in Bioconductor[57]. Sites coinciding with DMRT1 occupancy in the ChIP-seq data set from 8-week-old mice were selected for further analysis. The predicted minor-groove width at these locations was retrieved from the GBShape database[29]. The mean of the minor groove width at positions 6–10 on either side of the motif was used to classify sites containing either a unilateral or bilateral narrowing of the minor groove. A second classification based on the nucleotides at positions −6, +5 and +6 was used to predict which sites were likely to be bound by DMRT1 dimers, trimers, or tetramers, as indicated in **Figure 5d**.

38. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. & Pease, L.R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51–59 (1989).
39. Connaghan-Jones, K.D., Moody, A.D. & Bain, D.L. Quantitative DNase footprint titration: a tool for analyzing the energetics of protein-DNA interactions. *Nat. Protoc.* **3**, 900–914 (2008).
40. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
41. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
42. Adams, P.D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
43. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
44. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
45. Chen, V.B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
46. Lu, X.J. & Olson, W.K. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.* **3**, 1213–1227 (2008).
47. Vinci, G. *et al.* Association of deletion 9p, 46,XY gonadal dysgenesis and autistic spectrum disorder. *Mol. Hum. Reprod.* **13**, 685–689 (2007).
48. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
49. Zhu, P. *et al.* OTG-snpcaller: an optimized pipeline based on TMAP and GATK for SNP calling from ion torrent data. *PLoS ONE* **9**, e97507 (2014).
50. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
51. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
52. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
53. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
54. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
55. Bailey, T.L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).
56. Lawrence, M. *et al.* Software for computing and annotating genomic ranges. *PLOS Comput. Biol.* **9**, e1003118 (2013).
57. Huber, W. *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).

doi:10.1038/nsmb.3032

Additional Results

DMRT1 recognizes a consensus 13 bp DNA sequence *in vitro* (Murphy et al., 2007). Chip-seq analysis revealed that the consensus sequence recognized by DMRT1 *in vivo* is very similar to the one recognized *in vitro*. A complex of DMRT1 DM domain (DMRT1 $_{67\text{-}136}$) binding to DM recognition site (the one recognized *in vitro*) in a piece of DNA of 25bp was resolved using X-ray crystallography, and analysis of DMRT1 crystals revealed a unique kind of DNA interaction. EMSA and ChIP-exo analysis revealed that DMRT1 can bind DNA *in vitro* and *in vivo* as a dimer, trimer or tetramer, depending on the combination of DNA sequence and shape of the site.

Using exome sequencing we have found a heterozygous *de novo* point mutation (c.331A>G that caused the amino acid change Arg111Gly) in a completely feminized 46,XY individual with complete gonadal dysgenesis. The mutation was situated in the DNA recognition helix of DMRT1. Analysis of the exome sequencing data did not show any other pathogenic mutations relevant to the phenotype. Furthermore this mutation was not observed in 240 ancestry-matched controls. DMRT1p.R111G shows a reduced affinity to bind DNA. Moreover, mutant DMRT1 altered *in vitro* the binding stoichiometry of wild-type DMRT1 as shown by EMSA assay and structure analysis, probably effectuating a dominant disruption of normal binding activity of DMRT1. Analysis of the *in vivo* DMRT1 binding sites from *Sox9* show that WT DMRT1 bind as trimer, however DMRT1p.R111G bound weakly to the sites and when mixed with WT protein it altered the complex from a trimer to a tetramer with higher binding affinity.

In addition to the work published, using immunofluorescence we have analyzed the effect of DMRT1p.R111G on subcellular localization of the protein. Interestingly, when co-transfected with *FOXL2*, DMRT1p.R111G remains in the cytoplasm whereas the wild type DMRT1 co-localizes to the nucleus (FIG Results1A) Suggesting a probable interaction between DMRT1 and FOXL2, which is abolished by the DMRT1p.R111G mutation that results in the inability of mutant DMRT1 to localize to the nucleus. We also performed transient transactivation assays to asses the ability of the mutant protein to bind and activate the DM binding domain using an artificial reporter construct containing 4 copies of DM binding sites (Murphy et al., 2007). The mutant DMRT1p.R111G showed a statistically significant reduction in its ability to transactivate DM-binding sites (Fig Results1B).

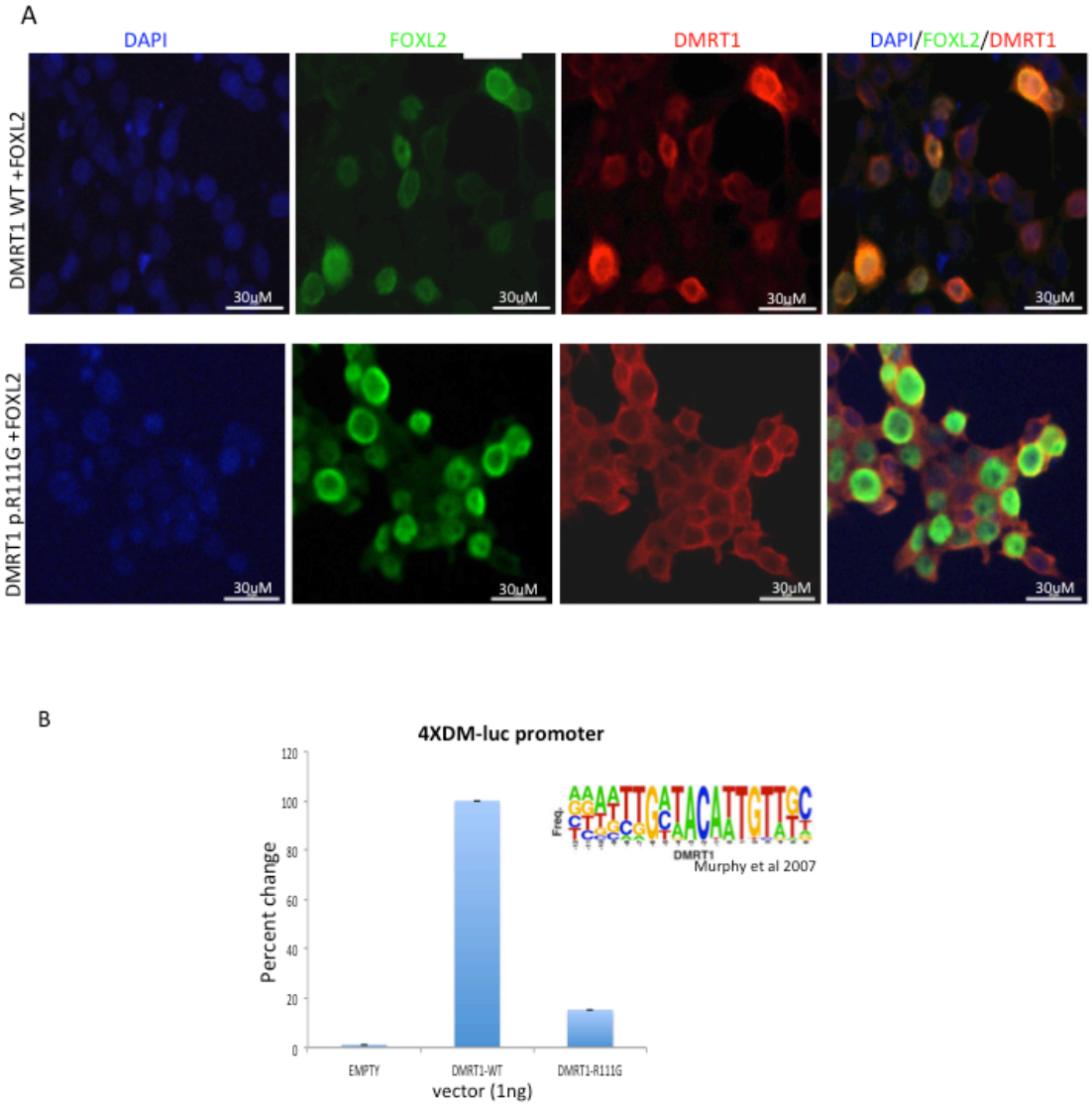


Figure Results 1. <u>DMRT1 supplementary results</u>.

(A) Cellular localization of DMRT1 with FOXL2. Human embryonic kidney HEK 293T cells were co-transfected with myc-tagged DMRT1 WT or p.R111G and FOXL2, 48hrs after transfection cells were fixed and stained with a DMRT1 antibody (RED) and FOXL2 (GREEN). DMRT1 shows a strong nuclear localization and co-localize with FOXL2 to the nucleus (DAPI/blue stained). While DMRT1p.R111G remains cytoplasmic when co transfected with FOXL2. (B) 4XDM transient gene expression analysis of DMRT1 WT and p.R111G using a vector containing 4X DM binding sites

Additional Materials and methods.

Plasmid constructions.

Plasmid *pCDNA6A-DMRT1-myc* was obtained by subcloning *DMRT1* from a T7-*DMRT1* vector (gift from Pr. David Zarkower, University of Minnesota) into pCDNA6A-myc (Life technologies V220-20) using the restriction sites NotI and XbaI. 4XDM vector containing 4 synthetic DM recognition site sequence, described previously (Murphy et al., 2007) was also a gift from Pr Zarkower. pCMV-XL5-FOXL2 vector was procured from Origene (SC126215)

Cell lines and transfections.

HEK293-T cells were a gift from Mlle Aurélie Claes (Institut Pasteur). For all trnsfections HEK293-T cells were transfected using Fugene 6 transfection reagent (Roche #1 814 443) using manufacturer's instructions.

Site-Directed Mutagenesis

Plasmid containing mutant DMRT1p.R111G was generated by site directed mutagenesis, using pCDNA6A-*DMRT1*-myc WT vector as template. Briefly, PCR reaction using oligonucleotides containing the mutation of interest in the middle of the sequence is done to amplify the complete plasmid sequence. PCR reaction is digested using *Dpn*I, a methylation sensitive enzyme that digests only template DNA. The reaction is transformed using XLBlue ultra-competent cells and seeded on selective NZY agar plates, colonies obtained are re-isolated for mini-prep and mutant plasmid are verified by direct sequencing before any functional analysis.

Luciferase/Renilla assays

Transient gene expression assays for DMRT1 function were performed in 96 well plates, using HEK 293T cells. The cells were at a concentration of 1*10^4 per well and allowed to grow for 24 hours. After 24 hours pCDNA6A-*DMRT1*-myc WT or p.R111G (1ng) were transfected with a reporter vector containing 4X DM binding sites (1ng/well) and pCMV-RL *Renilla* luciferase (Promega) as a marker of transfection efficiency, using Fugene 6 transfection reagent (Roche 1814443). Cells were lysed 48 hours later and luciferase assays were performed (Dual Luciferase Reporter Assay system, Promega) using a Centrox3 LB960 (Berthold Technologies). All data were standardized for Renilla activity. Results are shown as the mean ± SEM of five independent experiments, each performed atleast in triplicate.

Immunofluorescence

HEK 293T cells were seeded in an 8 well chamber slide using a concentration of 2 x 10^4 per well and transfected after 24 hours with pCDNA6A-*DMRT1*-myc WT or mutant with or without pcmvXL5-*FOXL2*. 48hrs after transfection, cells were fixed using 4%PFA (MSDS: 15735-605), blocked for 15 min, and incubated 1hr with primary antibody (DMRT1: ab126741, FOXL2: ab5096), washed 3X with PBS, incubated with secondary antibody (488 goat anti rabbit: a11034, anti goat: ab15132), washed 3X, covered with an antifading reagent with DAPI (Life Tech: P36931) and sealed. The cells were visualized using a microscope Leica DMI 4000B with an objective 40X and pictures analyzed using the software LAS V4.3

Discussion of the Publication

DMRT1 binds and bends DNA using a conserved mechanism for protein-DNA interaction. Using X-ray crystallography we have analyzed and described this interaction, which involves the recognition of a pseudo-palindromic target DNA mainly in DNA major groove. We also observed that DMRT1 can bind to DNA with multiple stoichiometries, from dimers to tetramers, depending on DNA sequence and shape of sites, that could explain the complex regulation of its targets. Using exome sequencing, we identified a heterozygous *de novo* mutation (R111G) in the DM domain of DMRT1, in a patient presenting with 46,XY CGD. This mutation is the only point mutation known, in human, in the functional domain of mammalian DMRT1 that abrogates the function of DMRT1 protein to bind DNA as shown by EMSA and transient transactivation assays using oligonucleotides and reporter constructs containing core DM binding elements.

The biological evidence and phenotype of the patient suggest a role for *DMRT1* in human primary sex-determination. A partial dominant negative effect of DMRT1p.R111G, on the WT DMRT1 protein may explain the severity of the phenotype that was observed. DMRT1p.R111G impairs the ability of DMRT1 to recognize DNA and as a consequence this could lead to the misregulation of DMRT1 specific gene targets. Moreover, we observed that WT DMRT1 and FOXL2 co-localised to the nucleus when co-transfected, but this was not seen with the DMRT1 mutant protein. It would be interesting to determine if there is a direct protein-protein interaction between DMRT1 and FOXL2. Since expression of *Dmrt1* in mice gonadal primordia of both sexes is present at E10.5, and at

E12.5 it decreases until it disappears in the ovaries and maintain high levels in the testis thereafter (De Grandi et al., 2000), and the *Foxl2* is expressed only in the ovary, beginning at E12.5 and continually expressed thereafter (Schmidt et al., 2004). It would also be of interest to develop a mouse model for DMRT1p.R111G to better understand what gene targets are being affected *in vivo,* even if *DMRT1* function seems to vary between two species, it constitutes a close model system to understand how *DMRT1* function. One point remains unresolved. Why have no other point mutations in the DMRT1 gene been reported in cases with a failure of testicular determination? Mutations in other sex-determining genes such as *SRY* or *NR5A1* are present in around 10-15% of cases of 46,XY gonadal dysgenesis, yet this is the only point mutation described to date in the *DMRT1* gene. One possibility is that the severe phenotype observed in this particular case is due to specific properties of the p.R111G amino acid change. The p.R111G mutation appears to have dominant negative activity on the WT-*DMRT1* allele essentially mimicking a homozygous mutation. However, this cannot be the only explanation. 75% of XY individuals who are hemizygous for *DMRT1* due to a deletion of 9p24 show various degrees of testicular anomalies with a proportion of cases presenting with XY complete gonadal dysgenesis. Therefore, one would expect to see heterozygous nonsense mutations in the *DMRT1* gene, which would be equivalent to the hemizygosity seen in 9p24 deletion cases.

In preparation: Mutations in SOX7 gene associated with 46, XY Gonadal dysgenesis.

Introduction to publication

In the laboratory we have been performing whole exome sequencing in patients who have unexplained and mainly non-syndromic 46,XY gonadal dysgenesis. For the most part these patients have been screened for mutations in the well-characterised testis-determining genes such as *SRY* and *NR5A1*. As part of this study we identified the *DMRT1* mutation described in the previous section. Here, we identified two heterozygous missense mutations in *SOX7* in two independent patients presenting with 46,XY gonadal dysgenesis. We have addressed the question of causality of the mutations by –

(i) Determining the presence or absence of these mutations in appropriate ancestry-matched control samples.

(ii) Examining where and when the *SOX7/Sox7* gene is expressed during gonad development.

(iii) Determining the cellular localisation of the wild type and mutated proteins

(iv) Performing functional analysis to characterise the changes in biological activity of the mutated protein compared to the wild type protein.

Our data are consistent with the hypothesis that *SOX7* mutations are a novel cause of 46,XY gonadal dysgenesis.

Publication in preparation

Mutations in the *SOX7* gene associated with 46,XY gonadal dysgenesis

Sandra Rojo[1], Caroline Eozenou[1], Marie-Charlotte Dumargne[1], Hassan Rouba[2], Joelle Bignon-Topalovic[1], Raja Brauner[3], Ken McElreavey[1], Anu Bashamboo[1]

[1]Human Developmental Genetics, Institut Pasteur, 75724 Paris Cedex 15 France
[2]Human Genetics Unit, Institut Pasteur of Morocco, Casablanca, 20100, Morocco

Keywords: Disorders of Sex Development (DSD), gonadal dysgenesis, SOX7, testis, Sertoli cells.

Abstract

Disorders of Sex Development (DSD) represent a wide spectrum of gonadal phenotypes. Identifying genetic causes of DSD is a major challenge since familial cases are rare and the genetic pathways involved are poorly conserved in evolution. Here, using an exome sequencing approach, we identified two individuals with 46,XY gonadal dysgenesis that carry independent novel missense mutations in the TAD domain of the *SOX7* gene. These mutations are absent from the public databases and in-house control samples. Functional analyses of the mutant proteins indicates that they lack the ability to synergise with NR5A1 to upregulate *SOX9* expression through the *Tesco* enhancer element. Although expression data in the mouse shows that Sox7 expression is limited to the endothelial cells of developing testis, in the human SOX7 is expressed in adult Leydig and Sertoli cells suggesting that it may be important in testis development.

Introduction

Mammalian testis-determination is a highly complex process where genetic pathways decide between two mutually opposing cell choice fates of the bipotential gonad anlage to develop into either a testis or an ovary (Svingen and Koopman, 2013). In the human, at 7 weeks in the XY gonad, the Y-located testis-determining *SRY* is expressed in pre-Sertoli cells and this, in synergy with the nuclear receptor NR5A1, results in the upregulation of *SOX9* expression beyond a critical threshold leading to the initiation of definitive Sertoli cell differentiation. Several positive regulatory loops are initiated for maintaining *SOX9* expression, including autoregulation of its own expression and formation of feed-forward loops via FGF9 or PGD2 signaling. In turn, SOX9 regulates the production of AMH from Sertoli cells, whilst at the same time represses the genetic pathways (WNT4/FOXL2) involved in ovarian development (Svingen and Koopman, 2013).

Errors in human testis-determination can give rise to Disorders of Sex Development that are defined as 'congenital conditions in which the development of chromosomal, gonadal, or anatomical sex is atypical' (Hughes et al., 2006). These include 46,XY complete or partial gonadal dysgenesis or 46,XX testicular or ovotesticular DSD.

The mammalian Sox transcription factor family comprises around 20 genes divided into 10 subgroups based on their sequence similarity and genomic organization (Bowles et al., 2000) and in recent years a role for other SOX family members in DSD phenotypes has been suggested. Mutations involving the *SOX3* and *SOX10* loci are associated with virilisation of 46,XX individuals (Sutton et al.,

2011; Polanco et al., 2010). Many other Sox family genes are also expressed in the developing testis (Schepers et al., 2003; Daigle et al., 2015).

SOX7 forms part of the subgroup F of the SOX family of transcription factors together with SOX17 and SOX18 (Francois et al., 2010). The human *SOX7* gene, located at 8p23.1, encodes a 388 AA protein that is characterized by a highly conserved N-terminal HMG-domain (aa 44-122), which contains two putative nuclear localization signals at both the N-terminal and C-terminal extremes. The protein also contains a C-terminal transactivation domain (TAD) (aa 177-387) conserved in SOX subgroup F (Takash et al., 2001). *SOX7* plays an important role as a developmental regulator in the differentiation of parietal endoderm by regulating expression of *Gata4* and *Gata6* (Murakami et al., 2004; Futaki et al., 2004; Costa et al., 2012), hematopoietic differentiation (Gandillet et al., 2009), cardiogenesis (Wat et al., 2012), arterial differentiation and maintenance and vasculargenesis (Herpers et al., 2008; Chung et al., 2011). As different members of SOX family, SOX7 can directly bind β-catenin catenin through its c-TAD domain and thereby negatively regulates Wnt signaling pathway (Guo et al., 2008)

Although *SOX7* function has largely been explored in vascular development and SOXF proteins appear to play a critical role in cardiovascular development (Wat et al., 2012; Wat and Wat, 2014), here, for the first time we show a role for SOX7 in human testis development. We identified novel missense mutations in the *SOX7* gene associated with 46,XY gonadal dysgenesis within the c-TAD domain.

Materials and Methods

The study was approved by the ethical board of Institut Pasteur (RBM 2003/8). Patient ancestry was determined by self reporting, based on responses to a personal questionnaire, which asked questions pertaining to the birthplace, languages and ethnicity of the participants, their parents and grandparents. The control panel consisted of 240 unrelated 46,XY males of French ancestry who are either normospermic or have fathered at least two children and have no history of testicular anomalies (determined by self reporting). All samples used for this study were collected with proper informed consent.

Whole Exome sequencing

Exon enrichment was performed using Agilent SureSelect Human All Exon V4. Paired-end sequencing was performed on the Illumina HiSeq2000 platform using TruSeq v3 chemistry. Read files (Fastq) were generated from the sequencing platform via the manufacturer's proprietary software. Reads were mapped using the Burrows-Wheeler Aligner14 and local realignment of the mapped reads around potential insertion/deletion (indel) sites was carried out with the GATK version 1.615. Duplicate reads were marked using Picard version 1.62 (http://picard.sourceforge.net). Additional BAM file manipulations were performed with Samtools (0.1.18)16. SNP and indel variants were called using the GATK Unified Genotyper for each sample. SNP novelty was determined against dbSNP138. Novel variants were analyzed by a range of web-based bioinformatics tools using the EnsEMBL SNP Effect Predictor (http://www.ensembl.org/homosapiens/userdata/uploadvariations). All

variants were screened manually against the Human Gene Mutation Database Professional Biobase (http://www.biobase-international.com/product/hgmd). *In silico* analysis was performed to determine the potential pathogenicity of the variants. Potentially pathogenic mutations were verified using classic Sanger sequencing.

Patient Descriptions.

Patient 1 was born to healthy non-consanguinous couple from North Africa. She was seen at the clinic at 17 years for primary amenorrhea. The external genitalia were completely female. There were no somatic anomalies. Breasts tanner stage 0, external genitalia Prader stage 0. Serum testosterone levels were 0.19 ng/ml (normal range 2.7-9ng/ml), FSH 74 U/Litre (ND-13.5 U/Litre) and LH 33 U/litre (2.4-13 U/Litre). Ultrasound showed has a normal uterus and 2 fallopian tubes. Laparotomy revealed two streak gonads. The histology of the testis revealed bilateral gonadal dysgenesis with homogenous ovarian stromal-like tissue and the presence of a gonadoblastoma. The proband has two 46,XX sisters with children. DNA of these individuals and the parents were unavailable for study. Karyotype was 46,XY. The diagnosis was 46,XY complete gonadal dysgenesis.

Patient 2 presented at 17 yrs with primary amenorrhea. Parents were healthy with no evidence of consanguinity and no family history of DSD or infertility. Breasts were not developed, Tanner stage 1 and external genitalia was female Prader stage 0. FSH levels were 50mUI/ml, LH levels were 20mUI/ml and E2 <13 pg/ml. Testosterone undetectable. Presence of a uterus. Histolopathology of the gonads. Right: Absence of Sertoli, Leydig and germ cells. Ovarian type stroma. No

tubule structures were observed. Left. Fallopian tube present. Ovarian stroma-like tissue. No evidence of Leydig cells or germ cells. Rare tubule-like structures present. Karyotype was 46,XY. The diagnosis was 46,XY complete gonadal dysgenesis.

For both patients direct sequencing of the *SRY, GATA4, WT1, FOG2/ZFPM2* and *NR5A1* genes revealed wild-type sequences. Array comparative genomic hybridization using the Nimblegene 2.1M aCGH platform confirmed a normal ploidy in both patients.

In silico analysis of 3D structure

The 3D structure models of wild type and mutants SOX7 c-terminal TAD were modeled using Phyre2 (Kelley and Sternberg, 2009). The predicted structure of AA177 to AA388 was analyzed using FirstGlance website interface based on Jmol (Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org/)

Plasmid constructions.

Plasmid pCMV6-*SOX7-myc* was obtained from Addgene (631604). All other vectors, pCMX-h*NR5A1* and reporters pGL3-*Tesco*, pGL3-*AMH*, pGL3-*NR5A1,* have been previously described in Bashamboo et al 2014. Reporter vector with *DMRT1 promoter was a gift from* from Dr. David Zarkower (Univ. of Minnesota). pCDNA-*SOX9*-Flag vector was a gift from Dr. Francis Poulat, Institut de Génétique Moléculaire de Montpellier.

Site-Directed Mutagenesis

Plasmids containing mutant *SOX7*p.D211A and p.Y381H were generated by site directed mutagenesis (QuikChange, Stratagene), using *pCMV6-SOX7-WT-myc* vector as a template as described elsewhere (Braman et al., 1996). Mutant plasmids were verified by direct sequencing before any functional analysis.

Transient expression analysis

Transient gene expression assays for SOX7 function were performed in 96 well plates, using HEK 293T cells at a concentration of $1x10^4$ per well, Fugene 6 transfection reagent (Roche 1814443) and a Dual-Glo luciferase reporter system (Dual Go Luciferase Substrate E297B) and a reporter vector for renilla activity (pRLSV40 Renilla luciferase). *pCMV6-SOX7-myc* WT, p.D211A or p.Y381H (10ng/well) were transfected with reporter vectors containing different sex determining genes promoters (10ng/well) and pRLSV40 Renilla (5ng) (Promega). 48hrs later, cells were lysed and luciferase assays were performed in a FLUOstar Optima fluorescence microplate reader (BMG Labtech). Data was normalized for Renilla activity. Results are shown as the mean of 3 independent experiments, performed in triplicate.

Immunochemistry

HEK 293T cells were seeded in an 8 well chamber slide using a concentration of 2×10^4 per well. Transfected using *pCMV6-SOX7-myc* WT, p.D211A or p.Y381H without and with *pCMX-NR5A1*-cherry. 48hrs after transfection, chamber slides were fixed using 4%PFA (15735-605), blocked for 15 min with 1X, and incubated 1hr with first antibody (anti-MYC: ab9186), washed 3X with PBS,

incubated with secondary fluorescent antibody (goat anti rabbit: Life-A11034), washed 3X, covered with mounting media with an anti-fading reagent (Life: P36931) and sealed.

Protein extraction

HEK 293T cells were seeded in 6 well plate at a concentration of $6x10^5$ per well. 24hrs later cells were transfected using the vectors *pCMV6-SOX7-myc* WT, p.D211A or p.Y381H along with *pacGFP-NR5A1*. 48hrs after transfection, cells were incubated with IP lysis buffer (Thermo 87788) and protease inhibitor (Life Tech 78429) and whole cell protein was transferred to a clean eppendorf, and measured by a Bradford assay.

Co-Immunoprecipitation assay

After Bradford assay to measure proteins, 250µg of protein extracts were incubated for 1 hr at 4°C with 1µg of rabbit IgG (Santa Cruz sc-2027) and 20µl of resuspended protein A/G PLUS agarose (Santa Cruz SC-2003). Samples were centrifuged (1000G, 5min at 4°C), the supernatant collected and incubated with 5µl of SF1 antibody (ab65815) for 1hr at 4°C. Subsequent, 20µl of A/GPLUS agarose (SC-2003) were added and samples were incubated on a rocking platform over night at 4°C. Samples were centrifuged (1000G, 5 min at 4°C), and the pellet was washed 4X with 1ml of cold PBS. After final centrifugation, pellet was resuspended in 40µl of SDS laemmli buffer + 4µl of DTT 1M. Samples were denatured for 5 min at 95°C and 20 µl were migrated on a fractionated on a 10% SDS-PAGE gel and transfer to a PVDF membrane. Western Blot was performed using SOX7 antibody for detection (ab49163), followed by a secondary antibody

HRP conjugated (ab6721) and visualized by chemioluminescence (Thermo-32106).

Results

SOX7 mutations associated with 46, XY gonadal dysgenesis.

We performed high throughput sequencing of DNA samples from the patient by exome capture. Bioinformatics analysis of WES data identified a novel heterozygous missense mutation, c.632A>C substitution that is predicted to result in a p.D211A amino acid substitution in the SOX7 transactivation domain for the first patient. The second patient carried a c.1141T>C mutation that is predicted to result in a p.Tyr381His (ENST00000304501). The presence of both mutations was confirmed by Sanger sequencing. The parents and other family members were unavailable for study, however, these missense mutations are absent from dbSNP138, Exome Variant Server (EVS), our internal database, and 1000 Genomes Project database. Furthermore, Sanger sequencing indicated that these mutations were absent from 400 healthy ancestry-matched control individuals. There was no other gene mutation found in the patients WES data that could be linked to testis development or 46,XY DSD.

In silico analysis of the SOX7p.D211A reveals an alteration of the secondary structure of SOX7 protein

Analysis of the secondary structure modeled with Phyre2 revealed that Wild-type SOX7 c-TAD was predicted to have 14.7% of α-helices, whereas SOX7p.D211A mutant has α-helices reduced to 6.2% and SOX7p.Y381H did have

the same percentage of predicted α-helices than WT: 14.7% (Fig 1). Suggesting that p.D211A could alters the normal structure of SOX7 protein. However, the accuracy of the structure models must be considered carefully, as these are only predictions since there is no actual model for TAD of SOX transcription factors.

Expression of SOX7

SOX7 is expressed in the nucleus of human adult Leydig cells, Sertoli cells and spermatogonia (Figure 2A; www.proteinatlas.org). During mouse development *Sox7* expression has been reported in the supporting endothelial cells of the testis as early as E13 (Figure 2B; EMAGE MGI: 98369) suggesting *Sox7* may play a role in testis development.

Missense mutations in SOX7 alter the biological activity.

To analyse the functional consequence of these mutation on the biological activity of the protein, we performed a series of transient gene expression assays using multiple gonadal promoters. Since SOX7 can directly inhibit Wnt4/β-catenin pathway (Guo et al 2008), we used the WNT sensitive TOP-Flash/FOP-Flash system to test the effect of *SOX7* mutants on WNT pathway. We did not observe a significant difference in the inhibition of WNT/β-catenin pathway between the wild-type and mutant SOX7p.D211A proteins **(Fig 3).**

NR5A1 is known to be a critical factor in testis development both by itself and in synergy with SOX9. We did not observe a significant difference in the activation of *NR5A1* reporter by the wild-type and SOX7p.D211A protein **(Fig 3)**. We then analyzed the *DMRT1* promoter, as it is also a key element for testis development

and the maintenance of testis cell fate. We did not detect a significant difference in activity between the WT SOX7 and SOX7pD211A **(Fig 3)**.

Both WT and SOX7p.D211A are able to activate the AMH reporter alone as well as synergistically with NR5A1 **(Fig 3).** We have seen that *SOX7* is able to recognize and activate/inhibit same targets as *SOX9*. Since SOX9 synergizes with NR5A1 to increase its own expression via Tesco, we performed co-transfections using *NR5A1* along with WT *SOX7* and *SOX7*p.D211A with reporter containing Tesco enhancer element of *Sox9*. Independent experiments show that both wild-type SOX7 and p.D211A have the ability to transactivate the *Sox9* Tesco enhancer element, however, SOX7p.D211A specifically failed to synergise with NR5A1 to transactivate the Tesco enhancer **(Fig 4**). However we did not observe a dominant negative effect of SOX7p.D211A over WT SOX7, SOX9 and NR5A1 on the Tesco enhancer **(Fig 4).**

SOX7p.D211A and SOX7p.Y381H interact with NR5A1

This failure of SOX7p.D211A mutant to synergise with NR5A1 for activation of Tesco enhancer element is not due to the loss of the ability of the mutant protein to physically interact with NR5A1. Independent co-immunoprecipitation experiments show that WT SOX7, SOX7p.D211A and SOX7p.Y381H proteins are able to physically interact with the NR5A1 protein **(Fig 5).**

SOX7p.D211A and SOX7p.Y381H mutations do not affect the nuclear localization of the protein or co-localization with NR5A1.

The human SOX7 protein is expressed in the nucleus (Séguin et al., 2008). We have found that WT SOX7, SOX7p.D211A, and SOX7p.Y381H show nuclear localization both WT and the two mutants SOX7 co-localize with NR5A1 (**Fig 6**).

Discussion

Large chromosomal rearrangements of 8p23.1 that contain both the *SOX7* and *GATA4* genes are associated with a complex phenotype that includes congenital diaphragmatic hernia and congenital heart disease. Some of the 46,XY affected individuals also presented with ambiguous external genitalia or hypospadias. (Barber et al., 2013 patient 6 (dup); Yu et al., 2011 (dup); Longoni et al., 2012, patients 11 (del), 13 (del), 17(del)). However, since these rearrangements also include the *GATA4* gene that is involved in Sertoli cell specification, it is difficult to specifically attribute the gonadal phenotypes seen in these individuals to the *SOX7* gene. Sequencing of the *SOX7* gene in a large series of 77 patients with congenital heart disease failed to identify pathogenic mutations (Wat et al., 2012). Here, using whole exome sequencing approach, we identified two novel heterozygous missense mutations involving *SOX7* in patients with 46,XY gonadal dysgenesis. No other variants were detected in these patients that could be linked to the phenotype of gonadal dysgenesis. In both cases other family members were unavailable for study and the mode of inheritance of the mutations is unknown, however, these mutations are absent from publically available SNP databases as well as our in-house ancestry-matched control samples suggesting a link between these mutations and the phenotype. Both

patients carrying the *SOX7* mutations had no history of cardiovascular anomalies.

SOX factors are expressed in multiple cell types and they achieve their specific regulatory functions by targeted binding to DNA in combination with their partner protein complexes. Non-HMG domains are directly involved in Sox-partner protein complexes. We have modeled *in silico* the secondary structure of SOX7 WT, SOX7p.D211A and SOX7p.Y381H and have found a difference in the quantity of α-helices in SOX7p.D211A model (WT: 14.7% and p.D211A: 6.2%); however, given the lack of an accurate structure of SOX7 TAD, this model prediction should be considered carefully. Importantly the ability to inhibit ovarian WNT/b-cat pathway of SOX7 was not disturbed by the mutations found in the c-TAD, as shown in transient gene expression using TOP/FOP system.

Analysis of the biological activity of mutant SOX7 proteins on male gonadal promoters did not show a difference between WT SOX7 and SOX7p.D211A to activate reporter gene activity through regulatory elements of the *NR5A1*, *DMRT1* and *AMH* genes. Although WT and p.D211A SOX7 were able to synergize with NR5A1 to over activate *AMH*, this was not observed with *Sox9 Tesco* enhancer element, where SOX7p.D211A is unable to synergize with NR5A1. The expression and physical ability of SOX7 WT and both mutants to interact at protein level with NR5A1 was confirmed by Co-IP analysis suggesting that the reduced ability of SOX7p.D211A to synergize with NR5A1 to activate *Tesco* enhancer is not due to a problem of partner complex formation with NR5A1. It could be possible that SOX7p.D211A is obstructing the correct interaction with

other factors required for the specific activation of the *Tesco* enhancer element at decisive stages for sex determination and thus may result in the 46,XY complete DSD phenotype observed in the patients. This would require expression of SOX7 in pre-Sertoli cells at the moment of sex-determination. Although the expression of murine *Sox7* appears to be specific for endothelial cells in the XY developing gonad at sex-determination (Figure 2A), corresponding data on the expression profile of SOX7 in the developing human gonad is unavailable. Interestingly, *SOX7* is expressed in human adult Sertoli and Leydig cells (Figure 2B). If *SOX7* is expressed in human developing Sertoli cells, then the functional data in this study may indicate a mechanism whereby these mutations could lead to gonadal dysgenesis. Conversely, if expression of the human gene is limited to endothelial cells during testis-determination, then the phenotype may be due to disrupted testicular vascularization leading to testicular dysgenesis.

In conclusion, our data show an association between *SOX7* mutation and 46,XY gonadal dysgenesis, indicating that S*OX7* expression is required at specific stages in human testicular development and/or maintenance and that *SOX7* mutation should be considered as a cause of 46,XY DSD.

References

- Bowles, J., Schepers, G., & Koopman, P. (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Developmental Biology*, *227*(2), 239–55. doi:10.1006/dbio.2000.9883
- Braman J, Papworth C, Greener A. Site-directed mutagenesis using double-stranded plasmid DNA templates. Methods Mol Biol. 1996;57:31-44. Review. PubMed PMID: 8849992.
- Chung, M. I. S., Ma, A. C. H., Fung, T.-K., & Leung, A. Y. H. (2011). Characterization of Sry-related HMG box group F genes in zebrafish hematopoiesis. *Experimental Hematology*, *39*(10), 986–998.e5. doi:10.1016/j.exphem.2011.06.010
- Costa, G., Mazan, A., Gandillet, A., Pearson, S., Lacaud, G., & Kouskoff, V. (2012). SOX7 regulates the expression of VE-cadherin in the haemogenic endothelium at the onset of haematopoietic development. *Development (Cambridge, England)*, *139*(9), 1587–98. doi:10.1242/dev.071282
- Daigle M, Roumaud P, Martin LJ. Expressions of Sox9, Sox5, and Sox13 transcription factors in mice testis during postnatal development. Mol Cell Biochem. 2015 Jun 5. [Epub ahead of print] PubMed PMID: 26045173.
- Francois, M., Koopman, P., & Beltrame, M. (2010). SoxF genes: Key players in the development of the cardio-vascular system. *The International Journal of Biochemistry & Cell Biology*, *42*(3), 445–8. doi:10.1016/j.biocel.2009.08.017
- Futaki, S., Hayashi, Y., Emoto, T., Weber, N., Sekiguchi, K., & Weber, C. N. (2004). Sox7 Plays Crucial Roles in Parietal Endoderm Differentiation in F9 Embryonal Carcinoma Cells through Regulating Gata-4 and Gata-6 Expression Sox7 Plays Crucial Roles in Parietal Endoderm Differentiation in F9 Embryonal Carcinoma Cells through Regulating . doi:10.1128/MCB.24.23.10492
- Gandillet, A., Serrano, A. G., Pearson, S., Lie-A-Ling, M., Lacaud, G., & Kouskoff, V. (2009). Sox7-sustained expression alters the balance between proliferation and differentiation of hematopoietic progenitors at the onset of blood specification. *Blood*, *114*(23), 4813–22. doi:10.1182/blood-2009-06-226290
- Guo, L., Zhong, D., Lau, S., Liu, X., Dong, X.-Y., Sun, X., … Zhou, W. (2008). Sox7 Is an independent checkpoint for beta-catenin function in prostate and colon epithelial cells. *Molecular Cancer Research : MCR*, *6*(9), 1421–30. doi:10.1158/1541-7786.MCR-07-2175
- Herpers, R., van de Kamp, E., Duckers, H. J., & Schulte-Merker, S. (2008). Redundant roles for sox7 and sox18 in arteriovenous specification in zebrafish. *Circulation Research*, *102*(1), 12–5. doi:10.1161/CIRCRESAHA.107.166066
- Hughes, I. A., Houk, C., Ahmed, S. F., & Lee, P. A. (2006). Consensus statement on management of intersex disorders. *Journal of Pediatric Urology*, *2*(3), 148–62. doi:10.1016/j.jpurol.2006.03.004
- Kelley, L. A., & Sternberg, M. J. E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols*, *4*(3), 363–71. doi:10.1038/nprot.2009.2
- Longoni M, Lage K, Russell MK, Loscertales M, Abdul-Rahman OA, Baynam G, Bleyl SB, Brady PD, Breckpot J, Chen CP, Devriendt K, Gillessen-Kaesbach G, Grix AW, Rope AF, Shimokawa O, Strauss B, Wieczorek D, Zackai EH, Coletti CM, Maalouf FI, Noonan KM, Park JH, Tracy AA, Lee C, Donahoe PK, Pober BR. Congenital diaphragmatic hernia interval on chromosome 8p23.1 characterized by genetics and protein interaction networks. Am J Med Genet A. 2012 Dec;158A(12):3148-58. doi: 10.1002/ajmg.a.35665. Epub 2012 Nov 19. PubMed PMID: 23165946; PubMed Central PMCID: PMC3761361.
- Murakami, A., Shen, H., & Dickson, C. (2004). Genes : Structure and Regulation : SOX7 and GATA-4 Are Competitive Activators of Fgf-3 Transcription SOX7 and GATA-4 Are

Competitive Activators of Fgf-3 Transcription *. doi:10.1074/jbc.M313814200

- Polanco JC, Wilhelm D, Davidson TL, Knight D, Koopman P. Sox10 gain-of-function causes XX sex reversal in mice: implications for human 22q-linked disorders of sex development. Hum Mol Genet. 2010 Feb 1;19(3):506-16. doi: 10.1093/hmg/ddp520. Epub 2009 Nov 20. PubMed PMID: 19933217.
- Séguin, C. A., Draper, J. S., Nagy, A., & Rossant, J. (2008). Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell*, *3*(2), 182–95. doi:10.1016/j.stem.2008.06.018
- Stovall, D. B., Cao, P., & Sui, G. (2014). SOX7: from a developmental regulator to an emerging tumor suppressor. *Histology and Histopathology*, *29*(4), 439–45. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4107680&tool=pmcentrez&rendertype=abstract
- Sutton E, Hughes J, White S, Sekido R, Tan J, Arboleda V, Rogers N, Knower K, Rowley L, Eyre H, Rizzoti K, McAninch D, Goncalves J, Slee J, Turbitt E, Bruno D, Bengtsson H, Harley V, Vilain E, Sinclair A, Lovell-Badge R, Thomas P. Identification of SOX3 as an XX male sex reversal gene in mice and humans. J Clin Invest. 2011 Jan;121(1):328-41. doi: 10.1172/JCI42580. Epub 2010 Dec 22. PubMed PMID: 21183788; PubMed Central PMCID: PMC3007141.
- Svingen, T., & Koopman, P. (2013). Building the mammalian testis: origins, differentiation, and assembly of the component cell populations. *Genes & Development*, *27*(22), 2409–26. doi:10.1101/gad.228080.113
- Takash, W., Cañizares, J., Bonneaud, N., Poulat, F., Mattéi, M. G., Jay, P., & Berta, P. (2001). SOX7 transcription factor: sequence, chromosomal localisation, expression, transactivation and interference with Wnt signalling. *Nucleic Acids Research*, *29*, 4274–4283. doi:10.1093/nar/29.21.4274
- Wat, J. J., & Wat, M. J. (2014). Sox7 in vascular development: review, insights and potential mechanisms. *The International Journal of Developmental Biology*, *58*(1), 1–8. doi:10.1387/ijdb.130323mw
- Wat, M. J., Beck, T. F., Hernández-García, A., Yu, Z., Veenma, D., Garcia, M., … Scott, D. A. (2012). Mouse model reveals the role of SOX7 in the development of congenital diaphragmatic hernia associated with recurrent deletions of 8p23.1. *Human Molecular Genetics*, *21*(18), 4115–25. doi:10.1093/hmg/dds241
- Weber, A., Köhler, A., Hahn, A., & Müller, U. (2014). 8p23.1 duplication syndrome: narrowing of critical interval to 1.80 Mbp. *Molecular Cytogenetics*, *7*(1), 94. doi:10.1186/s13039-014-0094-3
- Yu S, Zhou XG, Fiedler SD, Brawner SJ, Joyce JM, Liu HY. Cardiac defects are infrequent findings in individuals with 8p23.1 genomic duplications containing GATA4. Circ Cardiovasc Genet. 2011 Dec;4(6):620-5. doi: 10.1161/CIRCGENETICS.111.960302. Epub 2011 Sep 20. PubMed PMID: 21933911.

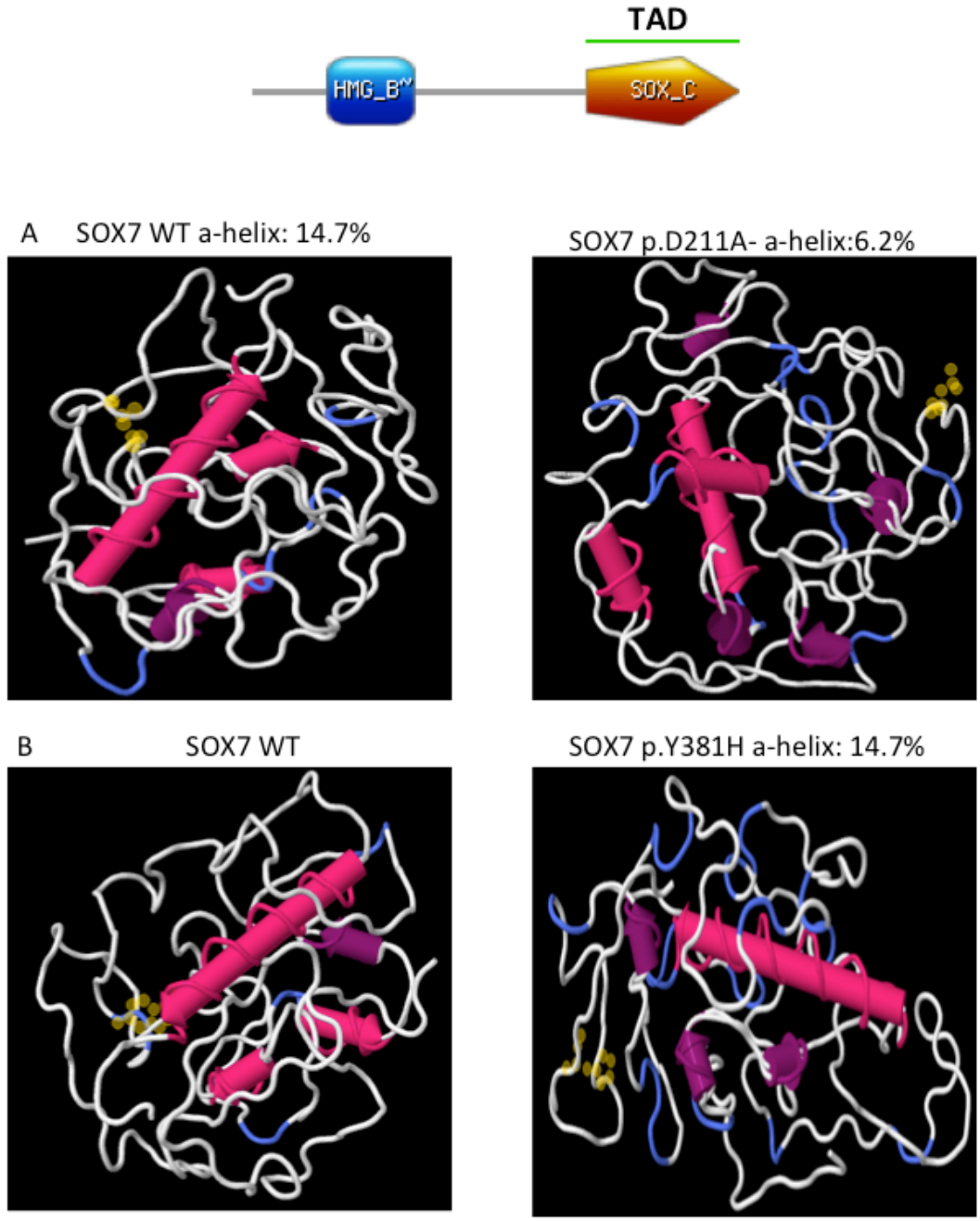
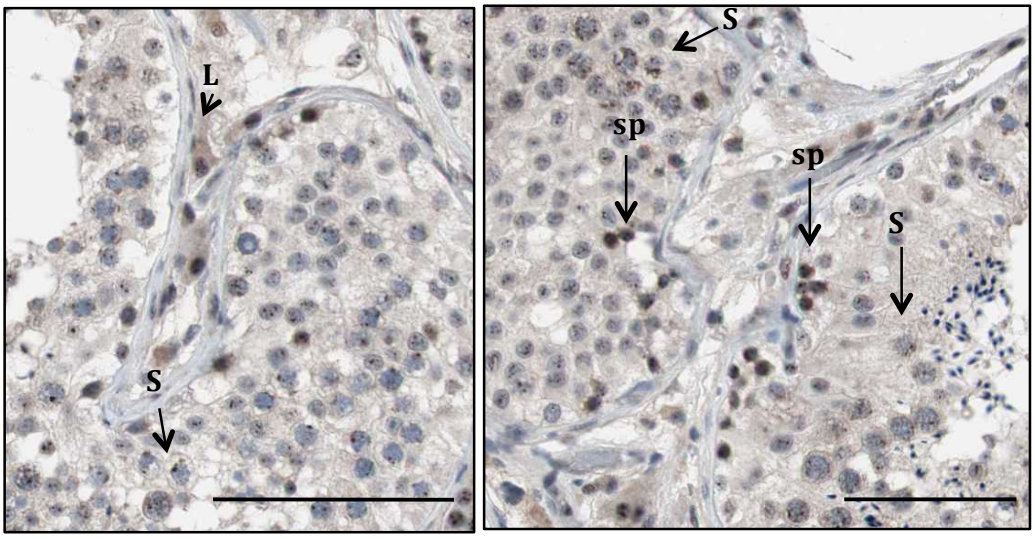


Figure 1. SOX7 TAD 3D MODELS. (A) Models of TAD domains of SOX7 WT and mutants constructed with Phyre2 algorithm and visualized with FirstGlance (Jmol). AA residues 211 and 381 labeled in yellow. Alpha helices are in pink and turns are in blue.

A



B

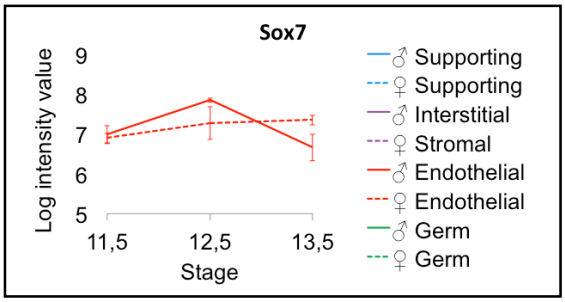


Figure 2. **Expression of SOX7 in (A) human adult testis, where expression is confined to Leydig cells (arrows). 26 yr old male testis, Leydig cells (L), Sertoli cells (S) and spermatogonia (sp) are indicated. bar 100 µm (www.proteinatlas.org) and (B) expression profile of *Sox7* from the developing mouse gonad showing expression in both male and female endothelial cells (from Jameson et al., 2012).**

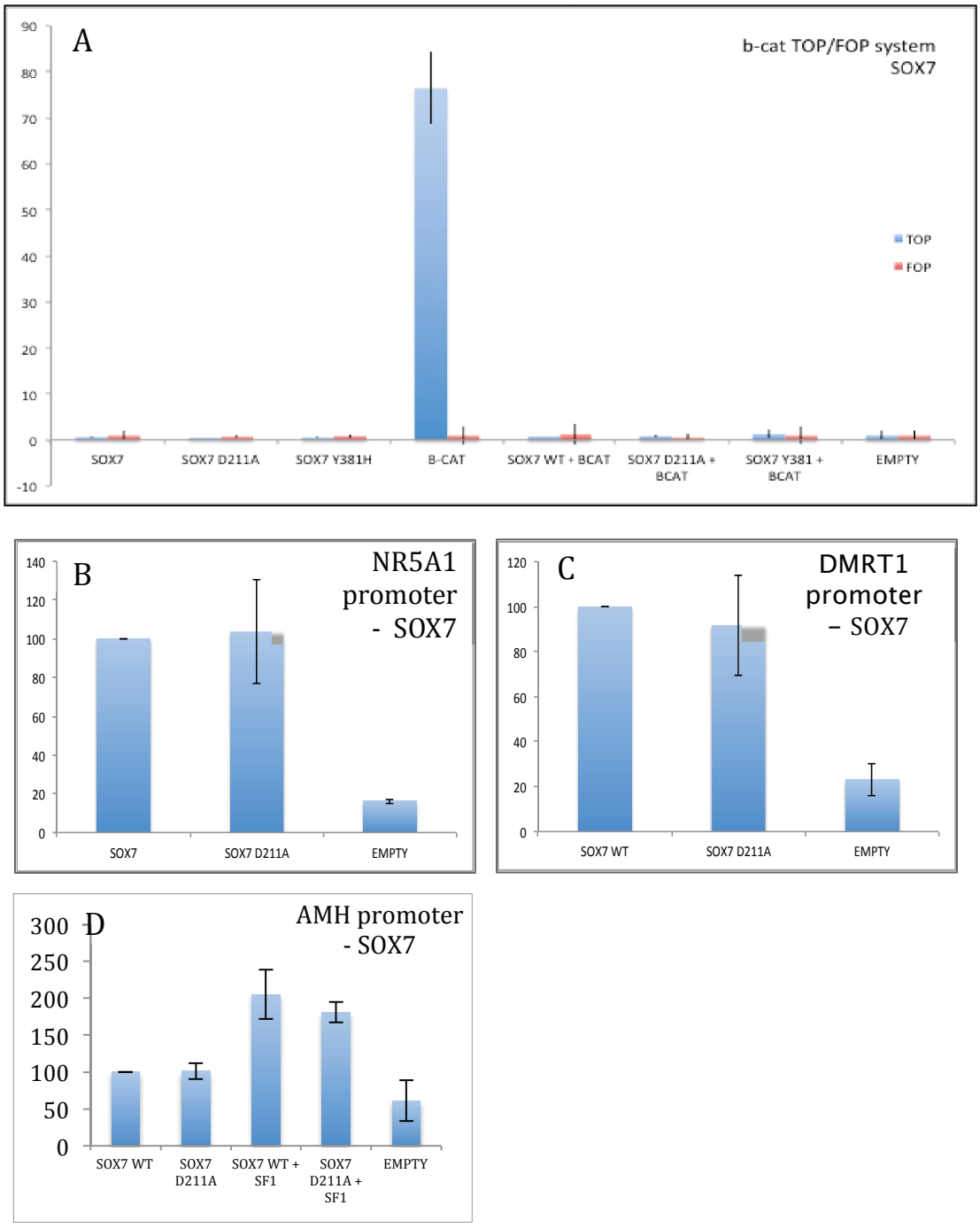


Figure 3 Transcriptional regulation of gonadal promoters by SOX7 (A) using the WNT sensitive TOP-Flash/FOP-Flash system both the wild-type and mutants SOX7 do not show a significant difference in the inhibition of WNT/β-catenin pathway B) both the wild-type and mutants SOX7 do not show a significant difference in the activation of *NR5A1* promoter (C) both the wild-type and mutants SOX7 do not show a significant difference in the activation of *NR5A1* promoter (D) Both WT and SOX7p.D221A are able to activate the AMH reporter alone as well as synergistically with *NR5A1*.

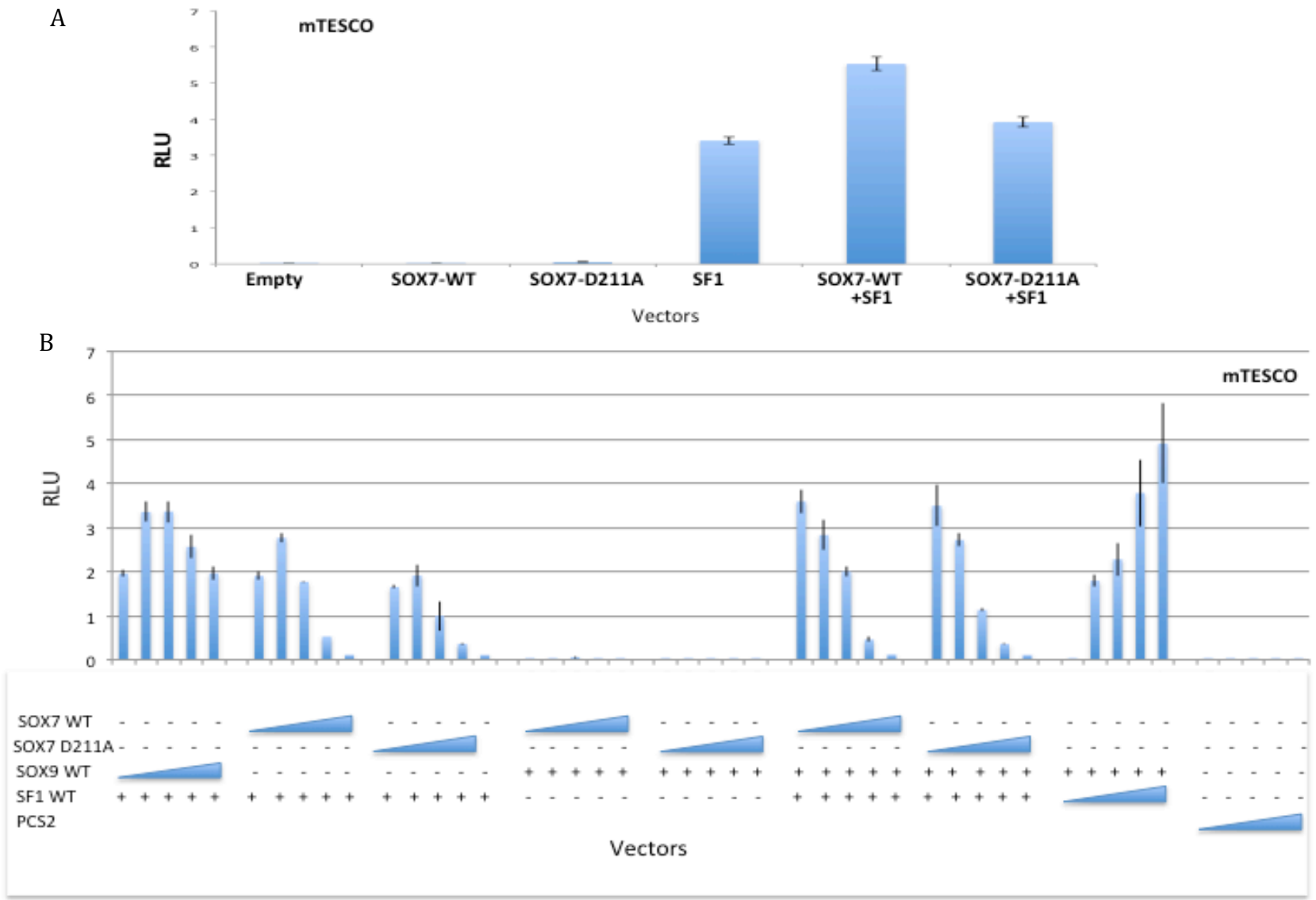


Figure 4 Transcriptional regulation of Tesco enhancer of *Sox9* by SOX7. A) Both wild-type and SOX7 and p.D211A have the ability to transactivate the *Sox9* Tesco enhancer element, however, the SOX7 p.D211A fails to synergise with NR5A1 to transactivate the Tesco enhancer B) SOX7p.D211A does not impart a dominant negative effect over WT SOX7, SOX9 and NR5A1 on the Tesco enhancer.

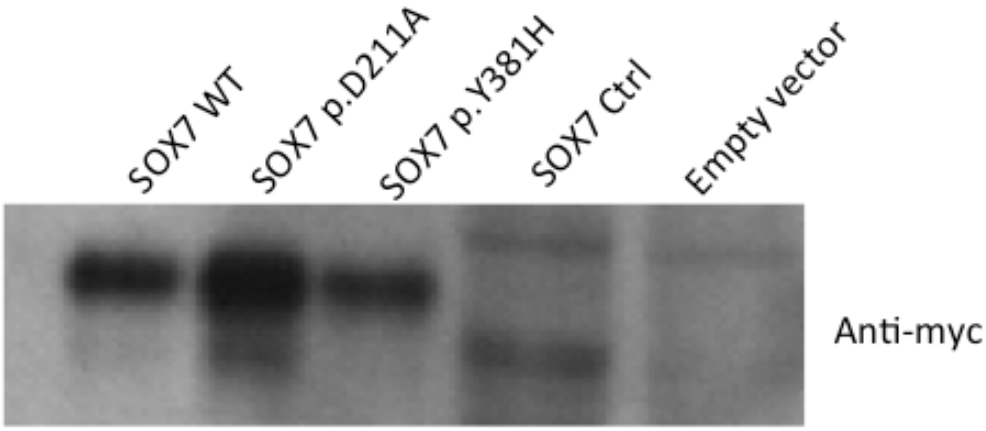


Figure5. SOX7 co-immunoprecipitates with NR5A1. Plasmids encoding SOX7 WT p.D211A and p.Y381H were transiently expressed in HEK293-T cells with WT NR5A1.. SOX7-NR5A1 complexes were immunoprecipitated from total protein extracts using A/G agarose beads (SC-2003) and an antibody specific for NR5A1 (ab65815) over night; the bound SOX7 protein was detected by western blot using an anti myc antibody (ab9106) followed by secondary antibody conjugated with HRP (ab6721), afterwards they were visualized by chemiluminescence (Thermo-80196). Blot shows that SOX7 WT (lane 1) and SOX7 p.D211A and p.Y381H (lane 2 and 3 respectively) can interact with NR5A1 protein.

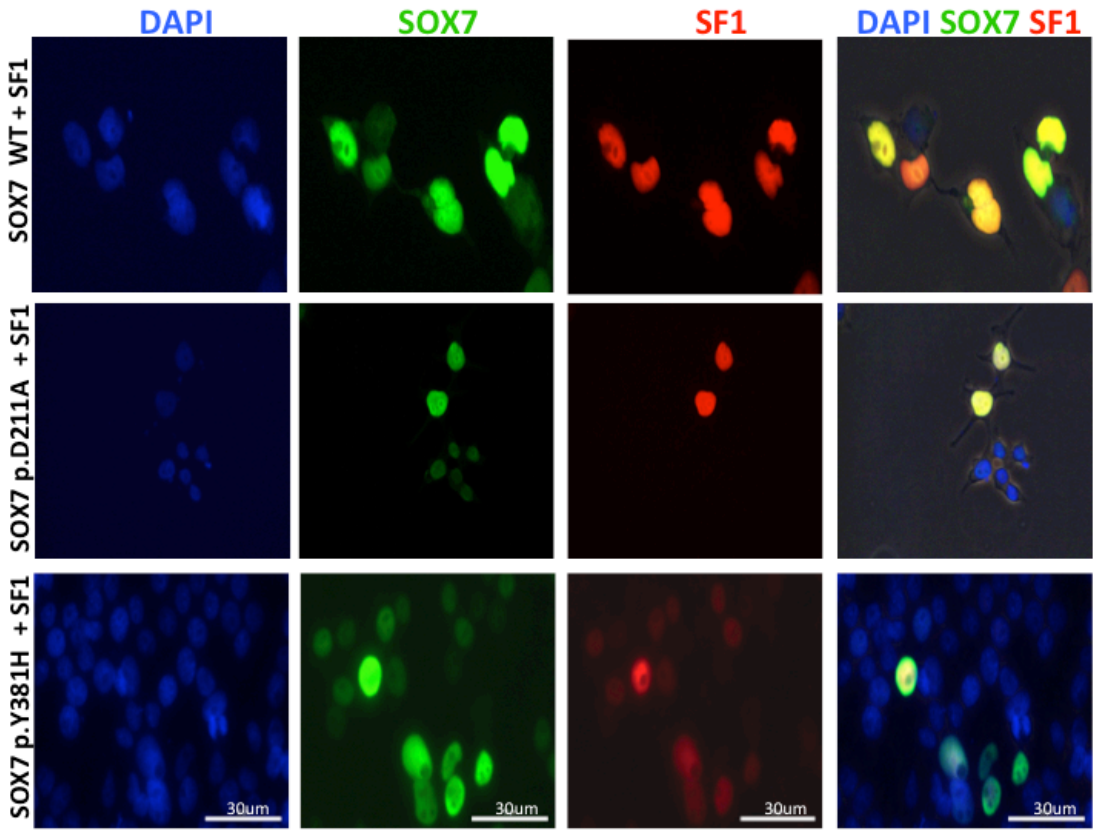


Figure 6 Cellular localization of SOX7. Human embryonic kidney HEK 293T cells were transfected with myc- SOX7 WT, myc- SOX7p.D211A, SOX7p.Y381H, and SF1-cherry, 48hrs after transfection cells were fixed and stained with a myc antibody. SOX7 WT, p.D211A and p.Y381H show a strong nuclear localization (green) and co-localize with SF1 (red) to the nucleus (DAPI/blue stained).

Discussion of Publication

We identified two novel missense mutations in TAD of SOX7 in patients presenting with 46,XY DSD. Mutations in *SOX7* gene have been previously associated with motor and speech development delay and/or learning problems and cancers (Stovall et al., 2014; Weber et al., 2014). This is the first time that point mutations in *SOX7* have been identified associated with non-syndromic 46,XY DSD.

SOX factors are expressed in multiple cell types. They achieve cell or lineage specific regulation by binding to target DNA in combination with their protein-partners in a cell specific manner. Using *in-vitro* assays to assess the effect of the two mutations, SOX7p.D211A and p.Y381H on the biological activity of protein we found that SOX7.pD211A activate the reporters for genes involved in early testicular development (*SF1, DMRT1* and *AMH*) similar to WT SOX7. Both mutants and the WT SOX7 can inhibit the ovarian specific pathway WNT. Moreover, both wild type and SOX7p.D211A were able to synergize with NR5A1 to activate *AMH* promoter. Experiments to further analyse the effect of SOX7p.Y381H mutation on transient activation of the downstream promoters are currently underway. However SOX7p.D211A cannot synergize with NR5A1 to upregulate the *Tesco* enhancer reporter. This loss of activation is not due to the disruption of partner complex formation with NR5A1 as confirmed by Co-IP analysis.

There is a possibility that the mutant proteins may occupy the SOX binding sites on the *Tesco* enhancer thus inhibiting the interaction of other SOX members at

decisive stages for sex determination. Alternatively, the mutation in *SOX7* may cause errors in testis determination by misregulation of other factors such as *GATA4* and *GATA6* similar to its role in parietal endoderm differentiation. We are currently testing this hypothesis.

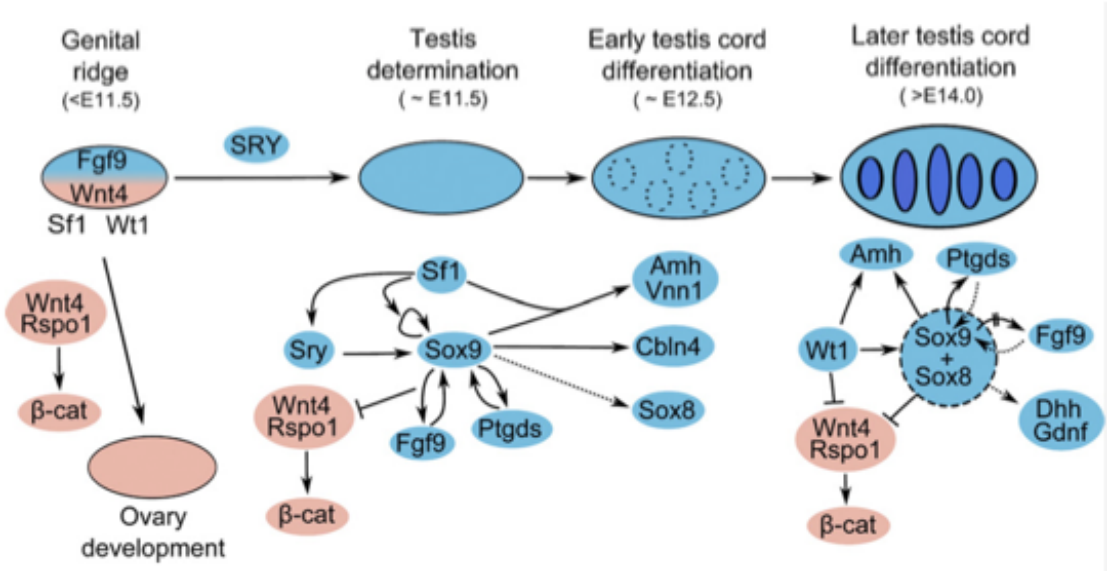
A key aspect and limitation of these findings is the expression profile of SOX7. Precisely, where and when is *SOX7* expressed in the developing human gonad. At the moment, we do not have access to suitable human foetal material to establish the expression profile although the gene is expressed in human germ cells, Sertoli Cells and Leydig cells in adults. Determining the expression profile of *SOX7* in the developing testis would enable us to understand the mechanism by which these mutations may contribute to the development of 46,XY gonadal dysgenesis.

Although the molecular mechanism still needs to be elucidated, for the first time we show that point mutations in *SOX7* may be associated with 46,XY DSD.

In preparation: Mutations involving the SRY-related gene SOX8 are associated with a spectrum of reproductive anomalies.

Introduction to publication

Sox8 is expressed during murine testis determination, first detected at E12.5, after *Sox9* expression and may possibly be initiated by it. Thereafter *Sox8* is expressed during embryonic and postnatal gonadal development (Fig Results2). The phenotype of mice lacking *Sox8* is relatively mild. *Sox8*$^{-/-}$ males are initially fertile but develop progressive spermatogenic failure followed by testicular dysgenesis, but there is no evidence for male-to-female sex-reversal (O'Bryan et al., 2008). The inactivation of both *Sox9* and *Sox8* results in a more severe phenotype. The deletion of one copy of *Sox9* at E11.5 in a *Sox8*$^{-/-}$ background causes a complete XY sex reversal with a few or absent sex cords at E15.5. When both genes are ablated at E14, it leads to a progressive testis cord degeneration with only residual levels of AMH and primary infertility (Barrionuevo et al., 2009) therefore indicating a functional redundancy between *Sox8* and *Sox9* in normal murine Sertoli cell differentiation and male fertility. The role of human SOX8 in gonad development or fertility is unknown. In this study, which shows that mutations in the *SOX8* gene are associated with a range of reproductive anomalies, my contribution was focused on the novel missense mutation in the HMG-box of SOX8 that we identified in association with 46,XY gonadal dysgenesis.



Results Figure 2. Roles of *Sox8* and *Sox9* in testis development.

WT1 and SF1 are expressed in the bipotential gonads. In the male gonad, SRY synergize with SF1 to upregulates *Sox9* expression. SOX9 start regulatory loops involving SOX9+SF1 and FGF9 and PGD2 signaling to keep its expression levels, and at the same time suppress the female pathway by repression of Wnt4 signaling. SOX9 activates *Amh*, and probably *Sox8*, thus establishing the testis developmental pathway. During testis cord differentiation, SOX9 and SOX8 co-regulate Amh, Dhh and Gdnf expression; and together SOX9 and SOX8 repress the ovarian genes Wnt4 and Rspo1. Solid and dotted arrows indicate proven and possible interactions, respectively. (Figure from Barrionuevo et al., 2010).

Publication in preparation

Mutations involving the *SRY*-related gene *SOX8* are associated with a spectrum of reproductive anomalies

*Portnoi, M-F[2], *Dumargne MC[1], Rojo S[1], Chantot-Bastaraud S[2], Hyon C[2], Louis-Sylvestre C[5], Validire P[6], de Malleray Pichard C[4], Christin-Maitre S[7], Brauner R[8], Persani L[9], Achermann J[3], Duncan A[3], Srivastava Y[10], Jauch R[10], Bignon-Topalovic J[1], Dain L1[1], Mazen I[12], Rouba H[13], Siffroi JP[2], McElreavey K[1], Bashamboo A[1]

1-Human Developmental Genetics, Institut Pasteur, Paris, France.
2-APHP-ER9 UPMC Service de Génétique et d'Embryologie médicales, Hôpital Armand Trousseau, Paris, 75012, France.
3-Developmental Endocrinology Research Group, Clinical and Molecular Genetics Unit, UCL Institute of Child Health, London, WC1N 1EH, U.K.
4-Service d'Endocrinologue, Hôpital Cochin, Paris, France.
5- Département Mère-Enfant, Institut Mutualiste Montsouris, Paris, France.
6- Département d'Anatomie Pathologique, Institut Mutualiste Montsouris, Paris, France.
7-Service d'endocrinologie, Hôpital Saint-Antoine, Paris, 75012, France.
8-Université Paris Descartes and Pediatric Endocrinology Unit, Fondation Ophtalmologique Adolphe de Rothschild, Paris, France.
9-Department of Clinical Sciences & Community Health, University of Milan, 20100 Milan, Italy Laboratory of Endocrine & Metabolic Research and Division of Endocrine and Metabolic Diseases, IRCCS Istituto Auxologico Italiano, Piazzale Brescia 20, 20149 Milan, Italy.
10-Genome Regulation Laboratory, Guangzhou Institutes of Biomedicine and Health, 190 Kai Yuan Avenue,Science Park, 510530 Guangzhou, China.
11-Centro Nacional de Genética Médica, Administración Nacional de Laboratorios e Institutos de Salud (ANLIS), Dr. Carlos G. Malbrán, Buenos Aires, Argentina; Instituto de Biología y Medicina Experimental, Consejo Nacional de Investigaciones Científicas y Técnicas (IBYME-CONICET), Buenos Aires, Argentina.
12-National research Centre, Cairo, Egypt
13-Human Genetics Unit, Institut Pasteur of Morocco, Casablanca, 20100, Morocco.

* Both the authors made an equal contribution to the study.

Human sex determination is a tightly controlled and highly complex process where the bipotential gonad anlage must choose between two mutually antagonistic fates - an ovary or a testis. In the XY gonad, *SRY* acts during embryonal development, in synergy with *NR5A1/SF1*, to activate the downstream effector *SOX9*[1], which in turn represses the ovarian pathways in the male gonad through multiple mechanisms[2]. *Sox8* shows an overlapping expression pattern with *Sox9* in foetal and adult mouse gonads[3] and functional redundancy between *Sox8* and *Sox9* may occur during testis development[4-6]. However, *Sox8* XY null mutants show normal testis development but develop postnatal progressive spermatogenic failure[7]. Here, we show that mutations involving human *SOX8* are associated with a failure of testis determination as well as male and female infertility.

Disorders of Sex Development (DSD) are defined as 'congenital conditions in which the development of chromosomal, gonadal, or anatomical sex is atypical[8]. This covers a broad range of phenotypes that include 46,XY DSD with complete (CGD) or partial (PGD) gonadal dysgenesis. In the newborn, it has been estimated that a specific molecular diagnosis is obtained in about 20% of cases of DSD and that only 50% of 46,XY children with DSD will receive a definitive clinical diagnosis[8]. Mutations in *SRY*, *NR5A1* and *MAP3K1* may explain 40% of the cases of 46, XY non-syndromic gonadal dysgenesis[9].

A role for *SOX8* in human primary testis-determination was suggested by the investigation of a 27 years old female with age appropriate external genitalia and secondary sexual characteristics, who presented with infertility and primary amenorrhea (table 1). Chromosome analysis showed a 46,XY *SRY*-positive karyotype with a derivative 16p chromosome in all metaphases that consisted of a paracentric inversion (**Fig 1a**). The karyotype was determined to be 46,XY,inv(16)(p13.3p13.1) and was confirmed by FISH analysis (**Fig. 1b**) Her parents and unaffected sister were not available for analysis and it is not known if this rearrangement is *de novo*. ArrayCGH (cytoChip Y3 BlueGnome) was normal as well as sequencing of the *SRY* and *NR5A1* genes. Using FISH with BAC

clones the centromeric breakpoint was mapped within the BAC clone RP11-609N14 between chr16;10428838-10600163 (UCSC Genome Browser; GRCh37/hg19), spanning the *ATF7IP2* gene (**Fig. 1c**). The telomeric breakpoint was mapped within the BAC clone RP11-728H8 between chr16; 814,190-962,809 (**Fig. 1d**). FISH analyses indicated that the BAC clone RP11 252I11 (chr16; 864894-1036488) was located within the inversion, thus localizing the breakpoint between chr16;814190-864894. The transcription initiation site of the *SOX8* gene is at chr16; 1031808. Therefore the breakpoint is located between 217,618-166,914 upstream of SOX8 **(Fig. 1e and supplementary Fig 1**). A bilateral gonadectomy revealed two small gonads, which on histological examination revealed bilateral undifferentiated gonadal tissue **(Fig. 1f**). These findings are similar to a 560 kb duplication located approximately 18 kb upstream of *SOX8* which was identified in a patient with 46,XY gonadal dysgenesis, skeletal and cardiac anomalies and developmental delay[10] and suggest that dysregulation of *SOX8* expression could negatively impact on sex-determination. Consistent with a role in sex-determination *Sox8* is expressed in supporting cell lineages during murine testis-determination (**Fig. 1g**) In the human, SOX8 is co-expressed with NR5A1 and SOX9 in Sertoli cells during human testis-determination as well as in adult Sertoli and Leydig cells (**Fig 1h and supplementary Fig. 2**).

To further determine if SOX8 play a role in primary sex-determination we sequenced the *SOX8* open-reading frame in a large series of individuals with 46,XY DSD (Table 2). The control cohorts consisted of the general unselected population with unknown fertility status and a highly selected ancestry-matched control sub-group of fertile and/or normospermic 46,XY individuals (Table 2). Three mutations were found in the panel of 204 cases of otherwise unexplained cases of 46,XY DSD. Two of these mutations were observed in control cohorts and their contribution to the phenotype is unlikely. The third mutation, a heterozygous c.468G>C is predicted to result in a p.Glu156Asp amino acid substitution located within the highly conserved DNA-binding HMG-box (**Fig. 2a**). These data show that not only large rearrangements but point mutations in *SOX8* also contribute to the phenotype. The Glu156 residue is evolutionary conserved not only within the SOX8 protein family members but also in the other

members of group E SOX proteins including SOX9 and SOX10, as well as the three members of the group F SOX family (**Fig. 2b**). The mode of inheritance of the mutation is unknown as the parents were unavailable for study, although the mutation was not observed in a large cohort of ancestry-matched controls. The SOX8 p.Glu156Asp mutation is not expected to disrupt DNA binding per se because this residue does not map to the DNA interaction interface (**Fig. 2c**). Furthermore, the equivalent residues in Sox2 and Sox17 (from Glu to Lys and vice versa) did not show any effect on DNA binding by the SOX HMG box in the absence of partner proteins[11]. However, this mutation could affect how SOX8 dimerises with its partner factors on DNA since changing the same residue in SOX2 and SOX17 affected the ability of these proteins to interact with OCT4 (**Fig 2c**)[11]. The *in-vitro* analyses to assess the effect of this mutant on the biological activity of the SOX8 protein support this hypothesis. Both the wild-type and SOX8p.Glu156Asp proteins co-localize to the nucleus along with NR5A1 with a small percentage being observed in the cytoplasm (**Fig 3a**). Both proteins can activate a series of gonadal promoters and can synergise with NR5A1 to activate the AMH promoter (**Fig. 3b**). Both wild-type and SOX8 p.Glu156Asp have the ability to transactivate the *Sox9* Tesco enhancer element, however, the SOX8p.Glu156Asp specifically failed to synergise with NR5A1 to transactivate the Tesco enhancer (**Fig 3b**). SOX8 and NR5A1 proteins can physically bind to each other as seen by protein interaction analysis (**Fig. 3c**). Recently it has been reported that SOX8 and SOX9 proteins heterodimerise and combinatorially regulate their target gene expression[12,13]. However, SOX8p.E156D exerts a dominant negative effect by preventing synergistic activation of Tesco enhancer by NR5A1 and SOX9 (**Fig. 3d**) as well as WT SOX8. Such a dominant negative effect may explain the failure of appropriate testis-determination in this patient and supports recent data that suggest Sox8 is a driver of testis formation, even in the absence of Sox9[4].

It is estimated that one in seven couples worldwide have problems conceiving but in the majority of cases the underlying causes of either male or female infertility is unknown[14]). Only a few single-gene defects that cause infertility have been identified in humans[15,16]. We and others, have previously reported

that mutations in another sex-determining gene, *NR5A1*, are associated with 3-5% cases of spermatogenic failure with normal male genital development[17] and female primary ovarian insufficiency (POI)[18]. POI is characterised by primary or secondary amenorrhea, hypergonadotropinism and estrogen deficiency in women under the age of 40 years[19]. This broad phenotypic spectrum includes ovarian dysgenesis and premature ovarian failure. To investigate whether mutations involving *SOX8* could be associated with other human reproductive disorders we directly sequenced the SOX8 gene in a large series of 427 patients with infertility. We identified rare or novel mutations in 1.47% of azoospermic men, 3.5% of oligozoospermic men and 6.5% of cases of POI. We identified mutations in 1.28% of the ancestry-matched general control population. In contrast, the incidence of mutations in the control sub-group of fertile and/or normospermic was 0.64% (Tables 2 and 3). The general control population contains samples from individuals with unknown fertility status. Considering the high incidence of human infertility in the general population, it is possible that at least some of the mutations within this group may be associated with infertility. All mutations associated with infertility are located either N- or C-terminal to the HMG-box of the protein. Transient transfection assays using the *AMH*, *Dmrt1* and *NR5A1* promoters as reporters in HEK-293 cells indicated biological differences between mutated and wild-type proteins that, in some cases, were promoter specific, e.g. the p.D382N mutation associated with POI shows a specific loss-of-function with the *AMH* promoter (**Supplementary Fig. 3**). The phenotype of male infertility is consistent with that observed in *Sox8*$^{-/-}$ mice[7]. In female mice, *Sox8* expression has been reported in preantral follicles, preovulatory follicles, cumulus granulosa cells and at high levels in mural granulosa cells, which line the wall of the follicle and are critical for steroidogenesis and ovulation[20]. The observation that Sox8 expression is higher in the mural cells that line the follicle wall than in the cumulus cells that surround the oocyte suggests that SOX8 may play a role in granulosa cell differentiation[20]. In the human adult ovary, SOX8 is highly expressed in granulosa cells (**Supplementary Fig. 2**). In the female, in the absence of S0X9 expression, SOX8 may be an important regulator of *AMH* expression in the adult ovary and thus explain the association of mutations in SOX8 in association with POI.

The results presented in this study provide further insight into the genetic mechanisms of human gonadal development and function and reinforce the suggestion that same genetic factors can contribute to the errors of gonadal development as well as function.

Acknowledgements

A.B is funded in part by the program Actions Concertees Interpasteuriennes (ACIP). A. B and K. McE are funded by a research grant from the EuroDSD in the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement n° 201444 as well as grant N°295097 entitled GM_NCD_in_Co—Reinforcing IPT capacities in Genomic Medicine, Non Communicable Diseases Investigation and international cooperation as part of the EU call FP7-INCO-2011-6. The work is also funded by an Franco-Egyptian AIRD-STDF grant. Y.S. thanks the Chinese Government Scholarship and UCAS for financial and infrastructure support. R.J. is supported by a 2013 MOST China-EU Science and Technology Cooperation Program, Grant No. 2013DFE33080, by the National Natural Science Foundation of China (Grant No. 31471238) and a 100 talent award of the Chinese Academy of Science**s.**

References

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

2. Lin, Y.T. & Capel, B. Cell fate commitment during mammalian sex determination. *Curr Opin Genet Dev*. **32**, 144-152 (2015)

3. Chaboissier, M. C. *et al.* Functional analysis of Sox8 and Sox9 during sex determination in the mouse. *Development*. **131**, 1891-901 (2004).

4. Lavery, R. *et al*. Testicular differentiation occurs in absence of R-spondin1 and Sox9 in mouse sex reversals. *PLoS Genet*. **8**, :e1003170 (2012)

5. Barrionuevo, F. *et al*. Testis cord differentiation after the sex determination stage is independent of Sox9 but fails in the combined absence of Sox9 and Sox8. *Dev Biol.* **327**, 301–312 (2009).

6. Nicol, B. & Yao, H. H. Gonadal Identity in the Absence of Pro-Testis Factor SOX9 and Pro-Ovary Factor Beta-Catenin in Mice. *Biol Reprod.* 2015 Jun 24. pii: biolreprod.115.131276. [Epub ahead of print]

7. O'Bryan, M. K. *et al*. Sox8 is a critical regulator of adult Sertoli cell function and male fertility. *Dev Biol.* **316**, 359-370 (2008).

8. Lee, P. A. Houk, C. P. Ahmed, S. F. & Hughes, I. A. International Consensus Conference on Intersex organized by the Lawson Wilkins Pediatric Endocrine Society and the European Society for Paediatric Endocrinology. Consensus statement on management of intersex disorders. International Consensus Conference on Intersex. *Pediatrics*. **118**, e488-500 (2006).

9. Bashamboo, A. & McElreavey, K. Gene mutations associated with anomalies of human gonad formation. *Sex Dev*. **7**, 126-146 (2013)

10. Erickson, R.P. Yatsenko, S.A. Larson, K. & Cheung, S. W. A Case of Agonadism, Skeletal Malformations, Bicuspid Aortic Valve, and Delayed Development with a 16p13.3 Duplication Including GNG13 and SOX8 Upstream Enhancers: Are Either, Both or Neither Involved in the Phenotype? *Mol Syndromol*. **1**, 185-191 (2011).

11. Ng, C.K., Li, N.X., Chee, S., Prabhakar, S., Kolatkar, P.R. & Jauch, R. Deciphering the Sox-Oct partner code by quantitative cooperativity measurements. *Nucleic Acids Res.* 40, 4933-41 (2012)

12. Huang, Y.H. Jankowski, A. Cheah, K.S. Prabhakar, S. & Jauch, R. SOXE transcription factors form selective dimers on non-compact DNA motifs through multifaceted interactions between dimerization and high-mobility group domains. *Sci Rep.* **5,** 10398 (2015)

13. Stolt, C.C. Lommes, P. Friedrich, R.P. &, Wegner, M. Transcription factors Sox8 and Sox10 perform non-equivalent roles during oligodendrocyte development despite functional redundancy. *Development*. 131, 2349-58 (2004)

14. Skakkebaek NE, *et al.* Is human fecundity declining? *Int J Androl*. **29**, 2-11 (2006)

15. Tahmasbpour, E. Balasubramanian, D. & Agarwal, A. A multi-faceted approach to understanding male infertility: gene mutations, molecular defects and assisted reproductive techniques (ART). *J Assist Reprod Genet*. **31**, 1115-1137 (2014)

16. Gordon, C. M, Kanaoka, T. & Nelson, L. M. Update on primary ovarian insufficiency in adolescents. *Curr Opin Pediatr*. 2015 Jun 17. [Epub ahead of print] PubMed PMID: 26087426.

17. Bashamboo, A. *et al*. Human male infertility associated with mutations in NR5A1 encoding steroidogenic factor 1. *Am J Hum Genet.* **87**, 505-512 (2010)

18. Lourenço, D. *et al.* Mutations in NR5A1 associated with ovarian insufficiency. *N Engl J Med.* **360**, 1200-1210 (2009).

19. Nelson, L. M. Clinical practice. Primary ovarian insufficiency. *N Engl J Med.* **360**, 606-614 (2009).

20. Salmon, N.A. Handyside, A.H. & Joyce, I. M. Expression of Sox8, Sf1, Gata4, Wt1, Dax1, and Fog2 in the mouse ovarian follicle: implications for the regulation of Amh expression. *Mol Reprod Dev.* **70**, 271-277 (2005).

21. Jameson, S.A.,. *et al.* Temporal transcriptional profiling of somatic and germ cells reveals biased lineage priming of sexual fate in the fetal mouse gonad. *PLoS Genet* 8:e1002575 (2012).

Patient	Phenotype	Clinical presentation (age)	Genetics	Hormonal profile (normal range)	Internal genitalia and gonads	Diagnosis
1	Female, height 183 cm, weight 96 kg	Primary amenorrhea with female external genitalia and small immature uterus, 43x22x30 mm (27 yrs).	46,XY,inv(16) (p13.3p13.1)	FSH: 59,5 Ui/l (2.0–15.0) LH: 23,6 Ui/l (2.1–14.9) T: 0.7 ng/ml (2.70–9.00)	Bilateral streak gonads with fallopian tubes. Stromal-like tissue with no evidence for testicular tissue.	46,XY complete gonadal dysgenesis
2	Female.	16 yr - primary amenorrhea	46,XY,	Elevated FSH, LH. Reduced T levels	Uterus absent. Streak gonads. Epididymus-like structure present.	46,XY partial gonadal dysgenesis

Table 1. Phenotypes, genetics and clinical investigations in two cases of 46,XY gonadal dysgenesis with mutations involving the *SOX8* locus (patient 1) or gene (patient 2).

Patient	Mutation (variant identifier dbSNP138; global minor allelic frequency; PolyPhen2 Score)	Ancestry	Phenotype and comments
1	p.Lys241Thr (rs553836905; 1/5007 alleles 0.762)	European	Azoospermia
2	p.Gly359Arg* (0.961)	Eastern European	Oligozoospermia $3.4x10^6$/ml
3	p.Gly359Arg* (0.961)	French	Oligozoospermia. $2x10^6$/ml
4	p.Ala373Ser (rs200828189; 14/4106 alleles; 0.000)	French	Oligozoospermia $9x10^6$/ml
5	p.Gly378Ser (rs576173763; 2/5006 alleles; 0.004)	Senegal	Severe Oligozoospermia $0.005x10^6$/ml
6	p.Asp382Asn (rs143203270; 6/8572 alleles; 0.986)	Tunisia	Oligozoospermia. $6x10^6$/ml
7	p.Asp382Asn (rs143203270; 6/8572 alleles; 0.986)	Algeria	Azoospermia
8	p.Arg8_Ser9del* -	Italian	Menarche at 13 yrs ; 1 gestation, secondary amenorrhea and POF at 37 yrs; Mother with menopause at 35 yrs
9	p.Ala32Val* (0.919)	Italian	Menarche at 14 yrs ; secondary amenorrhea and POF at 21 yrs
10	p.Pro242Leu (rs567451602; 2/5006 alleles; 0.001)	Paraguay	Primary Amenorrhea
11	p.Pro242Leu (rs567451602; 2/5006 alleles; 0.001)	Paraguay	Primary Amenorrhea (sister of case 10)
12	p.Ser267Leu (rs368972027; 1/8593 alleles; 0.942)	Italian	Amenorrhea at 30. Hypoparathyroidism during childhood; carries an *AIRE* gene mutation
13	p.Val282Ile* (0.990)	Argentinian-Italian	Secondary Amenorrhea
14	p.Asp382Asn (rs143203270; 6/8572 alleles; 0.986)	North African	First menses at the age of 12, Regular menses during 2 years. Secondary amenorrhea at 14 years; carries a heterozygous missense mutation in *NR5A1*
15	p.Asp382Asn (rs143203270; 6/8572 alleles; 0.986)	European	Primary amenorrhea at 16 yr. No breast development, impuberism, ovaries not seen by ultrasound
16	p.Asp382Asn (rs143203270; 6/8572 alleles; 0.986)	French	First menses at 11 yrs, regular menses up to 14 yrs, then secondary amenorrhea
17	p.Ala416Thr (rs201602067; 2/5006 alleles; 0.005)	Italian	Menarche at 9 yrs ; secondary amenorrhea and POF at 38 yrs

Table 2. Mutations in the *SOX8* gene associated with human male or female infertility. In cases 12 and 14 mutations were previously identified in the *AIRE* and *NR5A1* genes. In such cases the *SOX8* mutation may contribute to the severity of the pathology *Novel variant

46,XY DSD (n=204)	**Azoospermia (n=131)**	**Oligozoospermia (n=143)**	**POI (n=153)**	**Control I Ancestry-matched general population (n=946)**	**Control II Fertile normospermic controls (n=470; European 280; North African 190)**
p.Glu156Asp	p.Lys241Thr	p.Gly359Arg (x2)	p.Arg8_Ser9del deletion of 2 AA	p.Ala240Asp	p.Pro336Leu
p.Pro336Leu	p.Asp382Asn	p.Ala373Ser	p.Ala32Val	p.Asn263Ile	p.Ala428Thr
p.Thr340Ser		p.Gly378Ser	p.Pro242Leu (x2, sisters)	p.Pro290Leu	p.Thr340Ser
		p.Asp382Asn	p.Ser267Leu	p.Ala302Thr	
			p.Val282Ile	p.Pro336Leu	
			p.Asp382Asn (x3)	p.Thr340Ser	
			p.Ala416Thr	p.Asp382Asn (x4)	
				p.Ala428Thr (x2)	
1.47%	**1.53%**	**3.5%**	**6.5%**	**1.27%**	**0.64%**

Table 3. Frequency of *SOX8* mutations in 46,XY DSD, male infertility, primary ovarian insufficiency and in two control cohorts. Control cohort I includes both the ancestry-matched general population, where the fertility status is unknown and the control cohort II, which only includes ancestry-matched fertile/normospermic individuals.

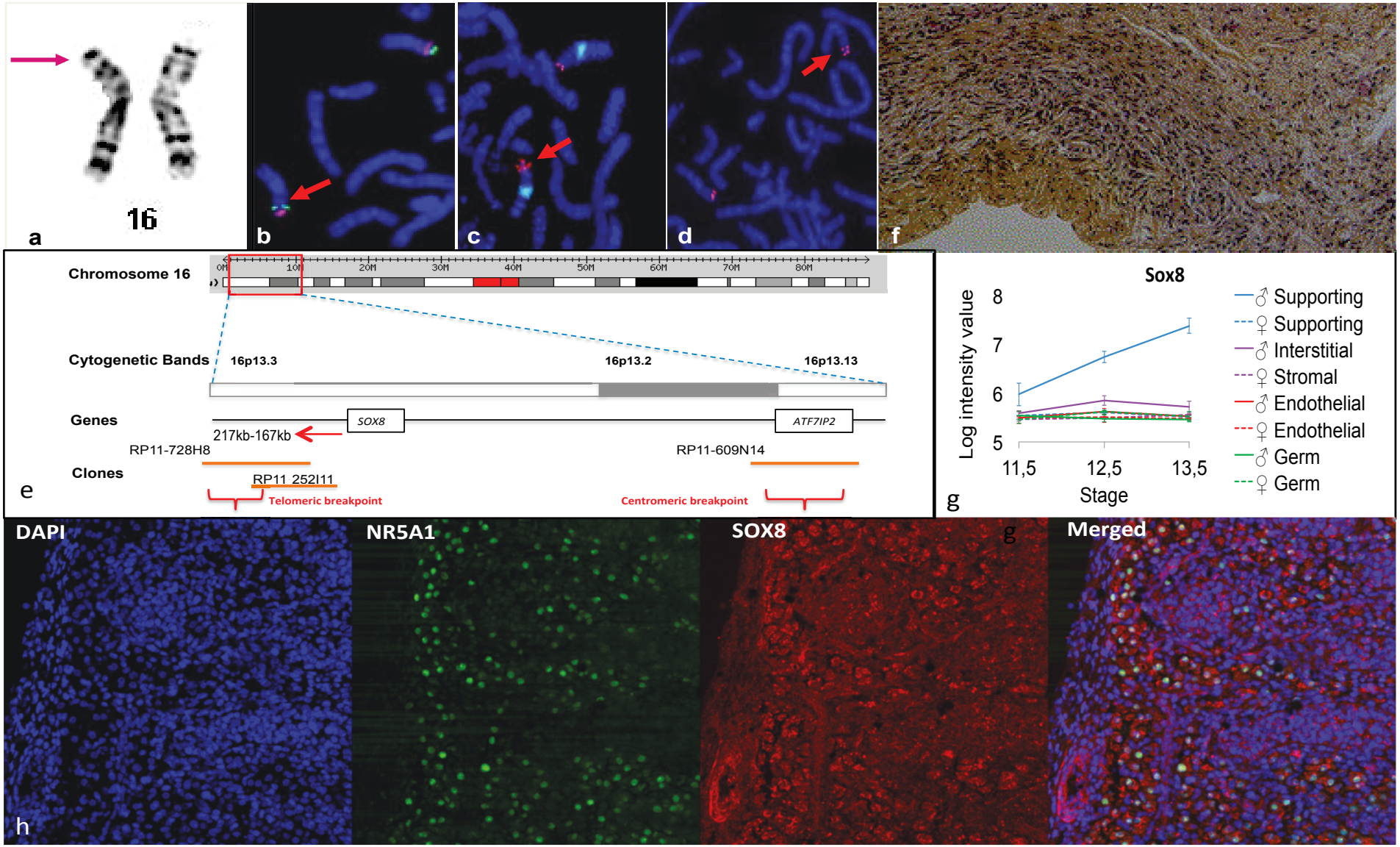
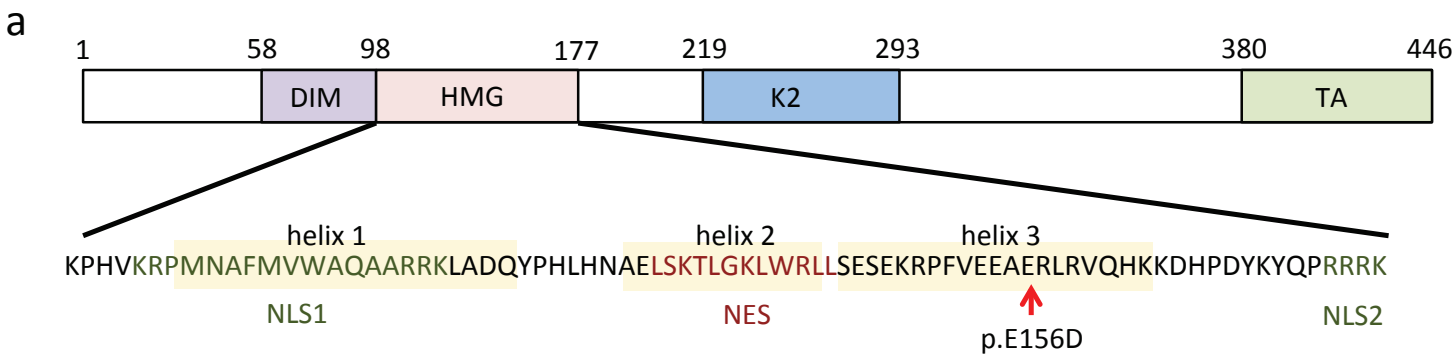

a
16
b
c
d
f
Chromosome 16
0M
10M
20M
30M
40M
50M
60M
70M
80M
Cytogenetic Bands
16p13.3
16p13.2
16p13.13
Genes
SOX8
ATF7IP2
217kb-167kb
RP11-728H8
RP11-609N14
Clones
RP11 252I11
Telomeric breakpoint
Centromeric breakpoint
e
Sox8
Log intensity value
8
7
6
5
11,5
12,5
13,5
Stage
♂ Supporting
♀ Supporting
♂ Interstitial
♀ Stromal
♂ Endothelial
♀ Endothelial
♂ Germ
♀ Germ
g
DAPI
NR5A1
SOX8
g
Merged
h

Figure 1. A chromosomal rearrangement of SOX8 associated with 46,XY complete gonadal dysgenesis (**a**) The patient chromosomes 16 (GTG banding) showing the rearranged chromosome 16 on the left (arrow). (**b**). FISH on metaphases from the patient with BAC RP11-161M6 (green) and RP11-297M9 (red) confirming a paracentric inversion on the der(16) (arrow). (**c**) The centromeric breakpoint is located within the clone RP11-728H8, on 16p13.3)(**d**). The telomeric inversion breakpoint mapped within RP11-609N14, on 16p13.13 (arrows (**e**). Cytogenetic band 16q13.3 with genes and clones of the region, indicating the distal breakpoint located 5' to the *SOX8* gene. **(f)** Histology of the gonad of the patient showing ovarian-like stromal cells with no evidence of testicular material (**g**) *Sox8* expression is highly up-regulated in the male supporting cell lineage of the mouse gonad during testis determination[21] (**h**) SOX8 expression is up-regulated in the human male supporting cell lineage together with NR5A1 during early testis formation at 9 weeks of gestation.



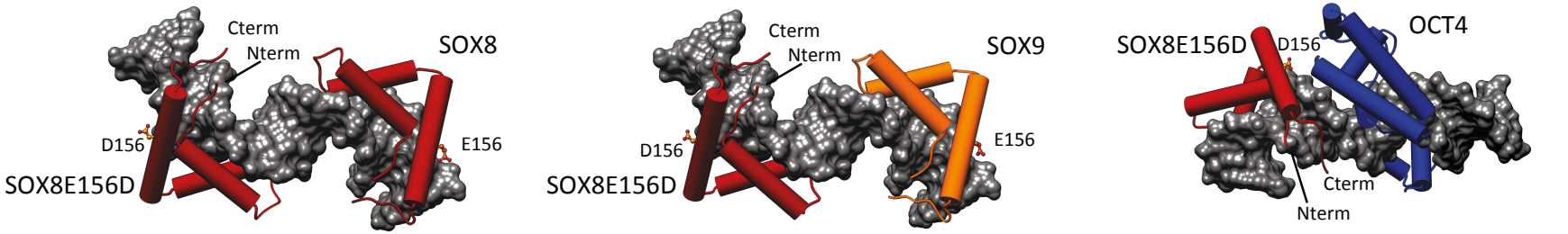
b

```
Homo sapiens           131-ELSKTLGKLWRLLSESEKRPFVEEAERLRVQHKKDHPDYK-170
Mus musculus           128-ELSKTLGKLWRLLSESEKRPFVEEAERLRVQHKKDHPDYK-167
Gallus gallus          134-ELSKTLGKLWRLLNESEKRPFVEEAERLRVQHKKDHPDYK-173
Oreochromis niloticus  129-ELSKTLGKLWRLLSESEKRPFVDEAERLRVQHKKDHPDYK-168
                           *************_********_*****************
```

```
SOX E  Human SOX8    132-LSKTLGKLWRLLSESEKRPFVEEAERLRVQHKKDHPDYK-170
       Human SOX9    135-LSKTLGKLWRLLNESEKRPFVEEAERLRVQHKKDHPDYK-173
       Human SOX10   134-LSKTLGKLWRLLNESDKRPFIEEAERLRMQHKKDHPDYK-172

SOX F  Human SOX17    98-LSKMLGKSWKALTLAEKRPFVEEAERLRVQHMQDHPNYK-136
       Human SOX18   115-LSKMLGKAWKELNAAEKRPFVEEAERLRVQHLRDHPNYK-153
       Human SOX7     75-LSKMLGKSWKALTLSQKRPYVDEAERLRLQHMQDYPNYK-113
```

c



CCCGGACAATGAAACATTGTCCGGG

CCCGGACAATGAAACATTGTCCGGG

GGCATTGTCATGCAAATCGGCGG

Figure 2 Mutation in the HMG-box of SOX8 associated with 46,XY gonadal dysgenesis. (**a**) Schematic representation of the SOX8 protein showing the functional domains of the SOX8 protein. The amino acid sequence of the HMG domain including the three alpha-helices, the 2 nuclear localization signals (NLS1, NLS2), and the nuclear export sequence (NES) together with the position of SOX8 p.Glu156Asp mutation are indicated. The DNA-dependent dimerization domain (DIM), the DNA-binding HMG domain, the K2 domain and the transactivation domain (TA) are shown. (**b**) Left, alignment of the distal portion of the SOX8 HMG-box domain containing the helix 3 compared with selected vertebrates. Right, alignment of the distal portion of the SOX8 HMG-box domain with other human SOXE and SOXF family members. The position of the p.Glu156Asp missense mutation is highlighted. (**c**) Structural models for ternary complexes of conjectured SOX8 p.Glu156Asp -SOX8, SOX8 p.Glu156Asp -SOX9 and SOX8 p.Glu156Asp -OCT4 dimers on composite DNA elements. SOX8 is shown in dark red, SOX9 in orange and OCT4 in blue. The HMG domains of SOX E proteins and the POU domain of OCT4 are shown as cartoon with cylindrical alpha-helices and the DNA as gray van-der-Waals surface. The SOX8p.Glu156Asp site is drawn as ball-and-sticks. The DNA sequences used to generate the models with color-coded binding elements are depicted underneath the models. The SOXE HMG box is N-terminally extended by a 40 amino acid dimerization (DIM) domain of unknown structure that mediates highly cooperative DNA dependent dimerization presumably by interactions with the HMG box.

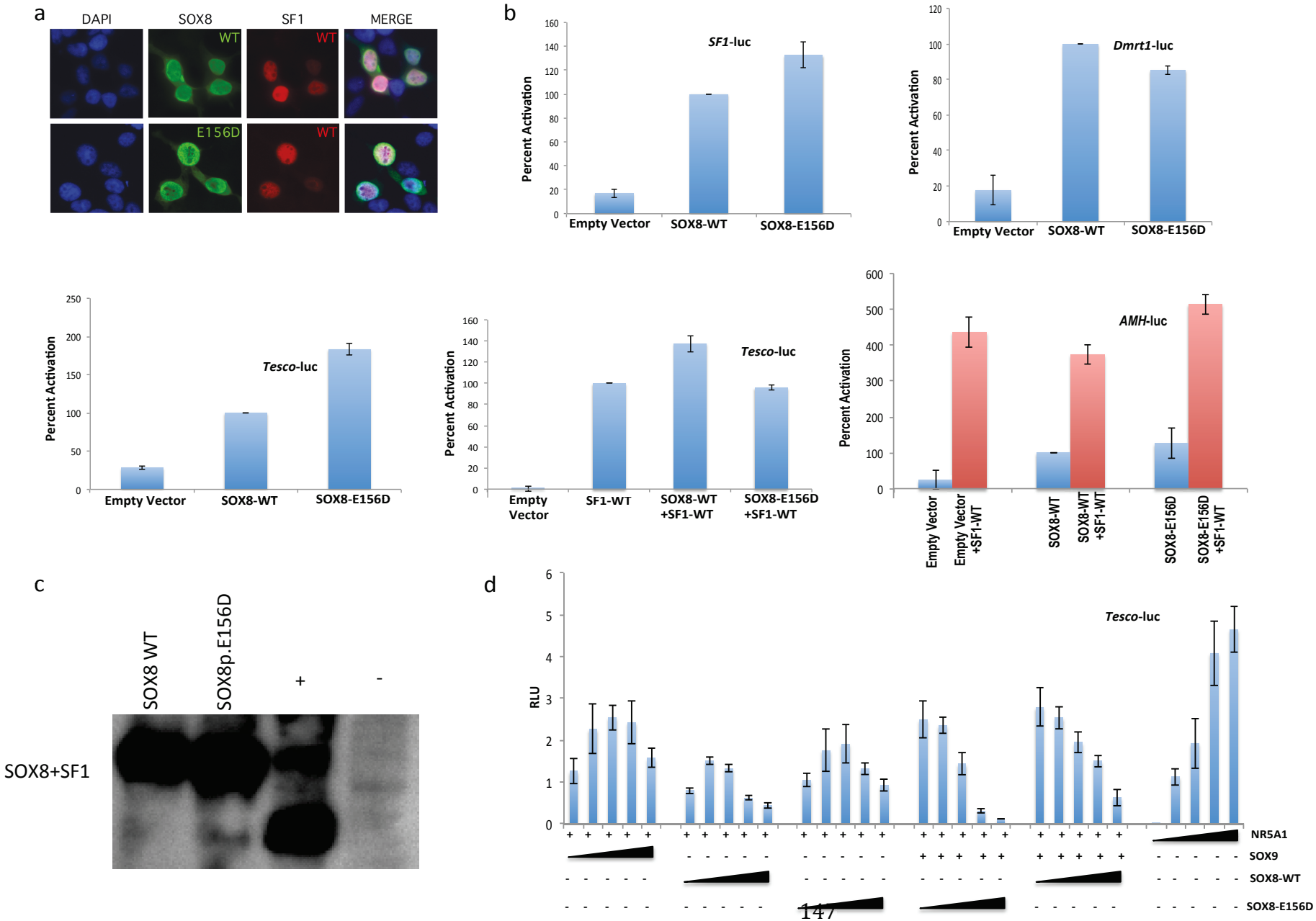

a
DAPI
SOX8
SF1
MERGE
WT
WT
E156D
WT
b
SF1-luc
Percent Activation
Empty Vector
SOX8-WT
SOX8-E156D
Dmrt1-luc
Tesco-luc
AMH-luc
Empty Vector +SF1-WT
SF1-WT
SOX8-WT +SF1-WT
SOX8-E156D +SF1-WT
c
SOX8 WT
SOX8p.E156D
+
-
SOX8+SF1
d
RLU
Tesco-luc
NR5A1
SOX9
SOX8-WT
SOX8-E156D

Figure 3 SOX8 p.Glu156Asp shows aberrant biological activity. **(a)** Human embryonic kidney (HEK 293-T) cells were transfected with myc-tagged SOX8-WT or SOX8 p.Glu156Asp with untagged NR5A1 expression vectors. 48 hours post transfection the cells were fixed and the co-expression of WT-SOX8 and SOX8 p.Glu156Asp along with NR5A1 was detected using anti-Myc (green) and anti-SF1 (red) antibody respectively. The nucleus is stained by DAPI (blue). When co-transfected both SOX8-WT or SOX8 p.Glu156Asp co-localize to the nucleus. However, when transfected alone SOX8 p.Glu156Asp shows and equal distribution between the nucleus and the cytoplasm (50% and 50%, respectively), as compared to WT protein, which is mostly nuclear (70%). Analyses included at least five biological replicates. Scale bars are 20 μm. Original magnification x40. **(b)** The transcriptional activities of SOX8-WT and SOX8 p.Glu156Asp were studied using the human *AMH* and *NR5A1* and mouse *Dmrt1* promoters and Tesco enhancer as reporter following transfection in HEK293-T cells. The data shown here represent the mean ± SEM of three independent experiments, each of which was performed in quadriplicate. The reporter constructs were transfected into HEK293-T cells with either the SOX8-WT or SOX8 p.Glu156Asp expression vector. The results are expressed as a percentage of WT SOX8 activity. The SOX8 p.Glu156Asp can activate the gonadal promoters similar to SOX8-WT. However, though SOX8 p.Glu156A can synergistically activate *AMH* promoter with NR5A1, it fails to synergise with *NR5A1* to activate Tesco enhancer. **(c)** Plasmids encoding WT-SOX8 and SOX8 p.Glu156Asp were transiently expressed with WT NR5A1 for 48hrs in HEK293T cells. SOX8-NR5A1 complexes were immunoprecipitated from total protein extracts using A/G agarose beads and an antibody specific for NR5A1; the bound SOX8 protein was detected by western blot using an anti SOX8 antibody followed by secondary antibody conjugated with HRP. Both SOX8-WT or SOX8 p.Glu156Asp can interact with NR5A1. **(d)** HEK 293-T cells were co-transfected with Tesco reporter construct (10ng), WT-NR5A1 (1ng) and WT-SOX9 (1ng) and increasing amount of SOX8-WT or SOX8 p.Glu156Asp (0,1,2,5,10 ng). Results are expressed as Relative luminescence units (RLU). The SOX8 p.Glu156Asp vector exhibits a dominant negative effect on synergistic of the Tesco reporter by NR5A1 and SOX9 even at three times higher concentrations.

Supplementary Material and Methods

Patient and control samples

All 46,XY DSD patients met the revised criteria of the Pediatric Endocrine Society (LWPES)/European Society for Paediatric Endocrinology (ESPE) Consensus Group. The study was approved by the Institut Pasteur (RBM 2003/8). Patient ancestry was determined by self reporting, based on responses to a personal questionnaire, which asked questions pertaining to the birthplace, languages and self-reported ethnicity of the participants, their parents and grandparents.

The control panel consisted of 280 unrelated normospermic 46,XY males of French ethnic origin with known fertility and no history of testicular anomalies (determined by self reporting). An additional 180 normospermic individuals from North Africa were included in the study. Further control samples consisted of ECACC Human Random Control (HRC) DNA from Public Health England (Panels 1 to 5), representing a control population of 480 randomly selected, non-related UK Caucasian blood donors of unknown fertility status. Ancestry-matched in-house controls of unknown fertility status were also included.

Whole Exome sequencing

40 patients presenting with 46,XY DSD were sequenced using the whole exome approach. Exon enrichment was performed with Agilent SureSelect Human All Exon V4. Paired-end sequencing was performed on the Illumina HiSeq2000 platform with TruSeq v3 chemistry. Read files (fastq) were generated from the sequencing platform via the manufacturer's proprietary soft- ware. Reads were mapped with the Burrows-Wheeler Aligner, and local realignment of the mapped

reads around potential insertion/deletion (indel) sites was carried out with GATK version 1.6. Duplicate reads were marked with Picard version 1.62 (http://broadinstitute.github.io/picard/). Additional BAM file manipulations were performed with Samtools (0.1.18). SNP and indel variants were called with the GATK Unified Genotyper for each sample. SNP novelty was determined against dbSNP138. Novel variants were analyzed by a range of web-based bioinformatics tools with the EnsEMBL SNP Effect Predictor (http://www. ensembl.org/homosapiens/userdata/uploadvariations). All variants were screened manually against the Human Gene Mutation Database Professional Biobase (http://www.biobase-international.com/product/hgmd/). *In silico* analysis was performed to determine the potential pathogenicity of the variants. Potentially pathogenic mutations were verified with classic Sanger sequencing

Sequencing of the entire coding sequence of *SOX8*:

The majority of patients and control samples were screened for mutations in the *SOX8* gene by direct sequencing. The coding exons of the *SOX8* gene were amplified from DNA extracted using conventional techniques from peripheral blood lymphocytes of each individual. We designed 5 pairs of primers to amplify by PCR the *SOX8* coding exons. Exon 1 amplicon 1, F1: 5'- CGCGGAGCTTTCTTTATGG-3', R1: 5'- GACTCCAGTCGTAGCCCTTG-3'. Exon 1 amplicon 2, F2: 5- CCCGATGCTGGACATGAG-3', R2: 5'- CTCGCAGTCCGAGACCAG-3' Exon 2, F3: 5'- AGAGGAGTGTACTGCCTGGTG-3', R3, 5'- CACAGCTGCCTCCCATACAG-3'. Exon 3 amplicon 1, F4: 5'- GCGCTTCATGGAATTTTCTC-3', R4 5'- GCTGCTCCGTCTTGATGTG-3'. Exon 3 amplicon 2, F5 5'- CCACGAGTTCGACCAGTACC-3', R5: 5'-

ACTTCAGCAGGCACTTGGAG-3'. The PCR amplification conditions were, F1/R1, F2/R2 and F3/R3 - 95°C for 5 min followed by 37 cycles of 98°C 30s, 60°C 30s, 72°C 30s. For the amplicon F4/R4 and F5/R5 the annealing temperature was 65°C. 5 µl of each PCR fragments were then electrophoresed in a 2% agarose gel stained in ethidium bromide (1µg/mL) to verify the expected length of the amplified fragments. DNA sequence analysis was performed using at least 200 ng of purified DNA, 20 ng of primer and fluorescently labelled Taq DyeDeoxy terminator reaction mix (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. DNA sequence was determined using an ABI 3700 automated DNA sequencer.

Plasmid construction:

Vector containing Full length human *SOX8* in pCMV6-AN-Myc vector was purchased from Origene (http://www.origene.com/). The pCMX-*NR5A1* vector with human *NR5A1,* and *AMH* and Tesco reporters has been previously described (Bashamboo et al. 2014). The pCDNA-*SOX9*-Flag vector was a gift from Dr. Francis Poulat, Institut de Génétique Moléculaire de Montpellier. Reporter construct with mouse *Dmrt1* promoter was a gift from Dr. David Zarkower (Univ. of Minnesota).

Site-Directed Mutagenesis:

SOX8 expression vectors containing the variants were generated by site-directed mutagenesis (QuikChange, Stratagene) with the use of wild-type (WT) human *SOX8* cDNA in pCMV6-AN-Myc expression vector. The entire coding sequence of all mutant plasmids was confirmed by direct sequencing prior to further studies.

Cell lines, transfections and protein preparation.

HEK293-T cells were a gift from Mlle Aurélie Claes (Institut Pasteur). For all trnsfections HEK293-T cells were transfected using Fugene 6 transfection reagent (Roche #1 814 443). Whole cellular protein was extracted using IP lysis buffer (PIERCE 87788). Briefly, the culture medium was removed from the wells and washed once with ice-cold phosphate buffered saline. 500µl ice cold lysis buffer was added to the cells and incubated on ice for 5 minutes with periodic mixing. Lysate was transferred to a microcentrifuge tube and centrifuged at ~13,000g for 10 minutes to pellet the cell debris at 4°C. supernatant was assayed for total protein concentration using Coomassie Plus (Bradford) Protein Assay (PIERCE 23236). The efficiency of transfection and presence of specific proteins was detected by western blot analysis using anti-SOX8 (rabbit, ab76196, Abcam®), anti-SOX9 (rabbit, #ab26414, Abcam®) or anti-SF1 (rabbit, #ab65815, Abcam®) antibodies.

Structural Modeling:

Binary SOXE/DNA complexes were generated with I-TASSER (Yang, et al., 2015) using human SOX8 (Uniprot = P57073) and SOX9 (Uniprot = P48436) sequences and the mouse Sox17/DNA complex (PDBID = 3F27) (Palasingam, et al., 2009) as a template. For the ternary SOX8–OCT4-DNA complex a previously generated SOX2–OCT4 model on the canonical *sox-oct* DNA element (GGCATTGTCATGCAAATCGGCGG) was used as a template (Merino, et al., 2014). The ternary complexes of Sox8E156D–Sox9, Sox8E156D–Sox8 were generated using chimera 1.10.1 (http://www.cgl.ucsf.edu/chimera/) (Pettersen, et al.,

2004) by superimposing DNA fragments to modelled ideal B-DNA corresponding to a palindromic *soxe* consensus sequence with 3 base-pair spacer between the *sox* elements (CCCG*GACA*<u>*ATG*</u>AAA<u>*CAT*</u>*TGTC*CGGG, base-pairs used for superposition are underlined and the core *sox* motifs are shown in italic) (Huang, et al., 2015). Next, sequences of the original *Lama1* DNA in 3F27 were converted into corresponding nucleotides in the ideal sequence and overlapping nucleotides were deleted. The phosphodiester bonds between juxtaposed SOXE/DNA complexes were created using the chimera 'adjust bond' option. The energy of the ternary complex models was minimized in chimera by using amber force fields ff14SB for the protein and Bsc0 for DNA. The final models were validated using the w3DNA server (Zheng, et al., 2009) and the integrity of protein-DNA contacts was checked using nucplot (Luscombe, et al., 1997). All figures were generated using chimera.

<u>Cellular Localization</u>

Cellular localization of both WT-SOX8 and mutants were assayed by transfecting the different plasmids into HEK 293-T cells with the use of FUGENE 6 (Roche) in chambers slides. 48 hours post transfection, cells were fixed in 3-4% paraformaldehyde in PBS (pH 7.4) for 15min at room temperature, washed twice with ice cold 1xPBS, permeabilized in PBS containing 0.25% Triton X-100, and then incubated with 1% BSA in PBS-T for 30 min to block non-specific binding of the antibodies. Cells were incubated with the primary antibody (diluted in 1% BSA in PBS-T) in a humid chamber overnight at +4°C. The following dilutions of primary antibodies were used: anti-Myc (rabbit, #ab9106, Abcam®), 1:100 and anti-SF1 (rabbit, #ab65815, Abcam®), 1:100. After 16 hour incubation, cells

were washed three times with 1xPBS for 5 min each. After the incubation with the secondary antibody in 1% BSA for 1h at room temperature in the dark (Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor® 488 conjugate, #A11034, and Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor® 594 conjugate, both diluted to 1:1000, Life Technologies), cells were washed three times with 1xPBS for 5 min each in the dark. The coverslips were mounted with a drop of mounting medium containing DAPI (DNA stain) for 1 min and sealed with nail varnish to prevent drying and movement under. Images were obtained with a Leica Microsystems DMI4000B microscope at 40x or 63x magnifications.

Transient Gene Expression Assays:

Transient gene expression assays to assess *SOX8* function were performed in 96-well plates (TPP) using either HEK293-T human embryonic kidney cells, Fugene6 and a Dual-Luciferase reporter assay system (Promega) with pCMV-RL *Renilla* luciferase (Promega) expression as a marker of transfection efficiency as described previously (Bashamboo et al; 2014). pCMV6-*SOX8*-Myc WT or mutant expression vectors (10 ng/well) were co-transfected with or without pCMX-*NR5A1* or pCDNA-*SOX9*-Flag into HEK293-T and mES cells with reporter vectors and vector containing renilla (10 ng/well). Cells were lysed 48 hours later and luciferase assays were performed (Dual Luciferase Reporter Assay system, Promega) using a Centrox3 LB960 (Berthold Technologies). All data were standardized for Renilla activity. Results are shown as the mean ± SEM of five independent experiments, each performed at least in triplicate.

Co-Immunoprecipitation assay

Whole cell lysate was prepared from the HEK293-T human embryonic kidney cells in 6-well plates (TPP) transfected using Fugene6. The cells were transfected with pCMV6-*SOX8*-Myc WT or mutant expression vectors (1 ug/well) with or without pCMX-*NR5A1* or pCDNA-*SOX9*-Flag. Transfection and protein extraction are described above.

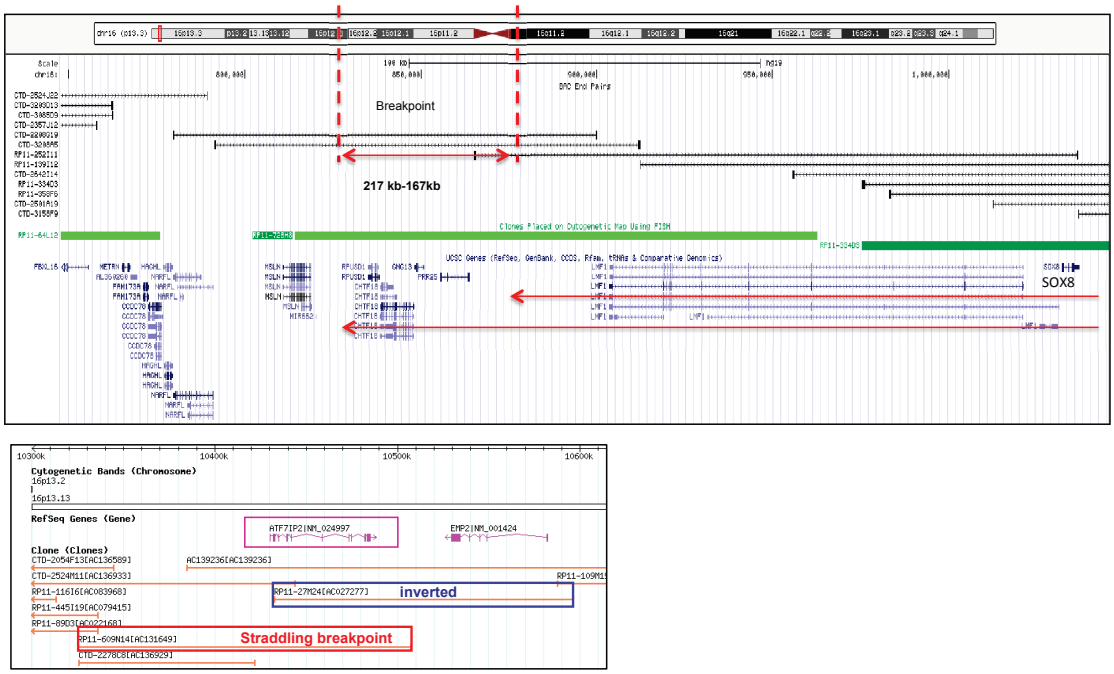
250 µg of total cellular protein extract was pre-cleared by adding 20 µl of resuspended Protein A/G PLUS-Agarose (SantaCruz sc2003) and 5ul normal mouse IgG (SantaCruz sc2025) followed by incubation at 4° C on a rocker platform for 1 hour. The mix was centrifuged at 1,000g for 5 minutes at 4° C and supernatant containing pre-cleared protein was collected in a new tube. The pre-cleared lysate was incubated overnight on a rocker platform at 4° C with 50 ng SOX8 antibody (rabbit, ab76196, Abcam®) and 25 µl of resuspended Protein A/G PLUS-Agarose. The antibody-protein-bead complex was centrifuged at 1,000g for 5 minutes at 4° C and the supernatant discarded. The pellet was washed 4 times with 1ml ice cold PBS and finally resuspended in 20 µl of SDS-lammeli buffer. The resuspended protein was boiled for 2 min and 25 µl was migrated on 10% SDS-PAGE gel and transferred to PVDF membrane. Proteins were detected with anti-SOX9 (rabbit, #ab26414, Abcam®) or anti-SF1 (rabbit, #ab65815, Abcam®) antibodies, followed by secondary antibody conjugated with horseradish peroxidase and visualized by chemiluminiscence using ECL reagent (Amersham).

References

1. Bashamboo, A, Brauner, R, Bignon-Topalovic, J, Lortat-Jacob, S, Karageorgou, V, Lourenco, D, Guffanti, A, McElreavey, K. 2014. Mutations in the FOG2/ZFPM2 gene are associated with anomalies of human testis determination. Hum. Mol. Genet. 23: 3657–65.
2. Huang, Y.H., *et al*. (2015) SOXE transcription factors form selective dimers on non-compact DNA motifs through multifaceted interactions between dimerization and high-mobility group domains, *Scientific reports*, **5**, 10398.
3. Luscombe, N.M., Laskowski, R.A. and Thornton, J.M. (1997) NUCPLOT: a program to generate schematic diagrams of protein-nucleic acid interactions, *Nucleic acids research*, **25**, 4940-4945.
4. Merino, F., *et al*. (2014) Structural basis for the SOX-dependent genomic redistribution of OCT4 in stem cell differentiation, *Structure*, **22**, 1274-1286.
5. Palasingam, P., *et al*. (2009) The structure of Sox17 bound to DNA reveals a conserved bending topology but selective protein interaction platforms, *J Mol Biol*, **388**, 619-630.
6. Pettersen, E.F., *et al*. (2004) UCSF Chimera--a visualization system for exploratory research and analysis, *Journal of computational chemistry*, **25**, 1605-1612.
7. Yang, J., *et al*. (2015) The I-TASSER Suite: protein structure and function prediction, *Nature methods*, **12**, 7-8.
8. Zheng, G., Lu, X.J. and Olson, W.K. (2009) Web 3DNA--a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures, *Nucleic acids research*, **37**, W240-246.

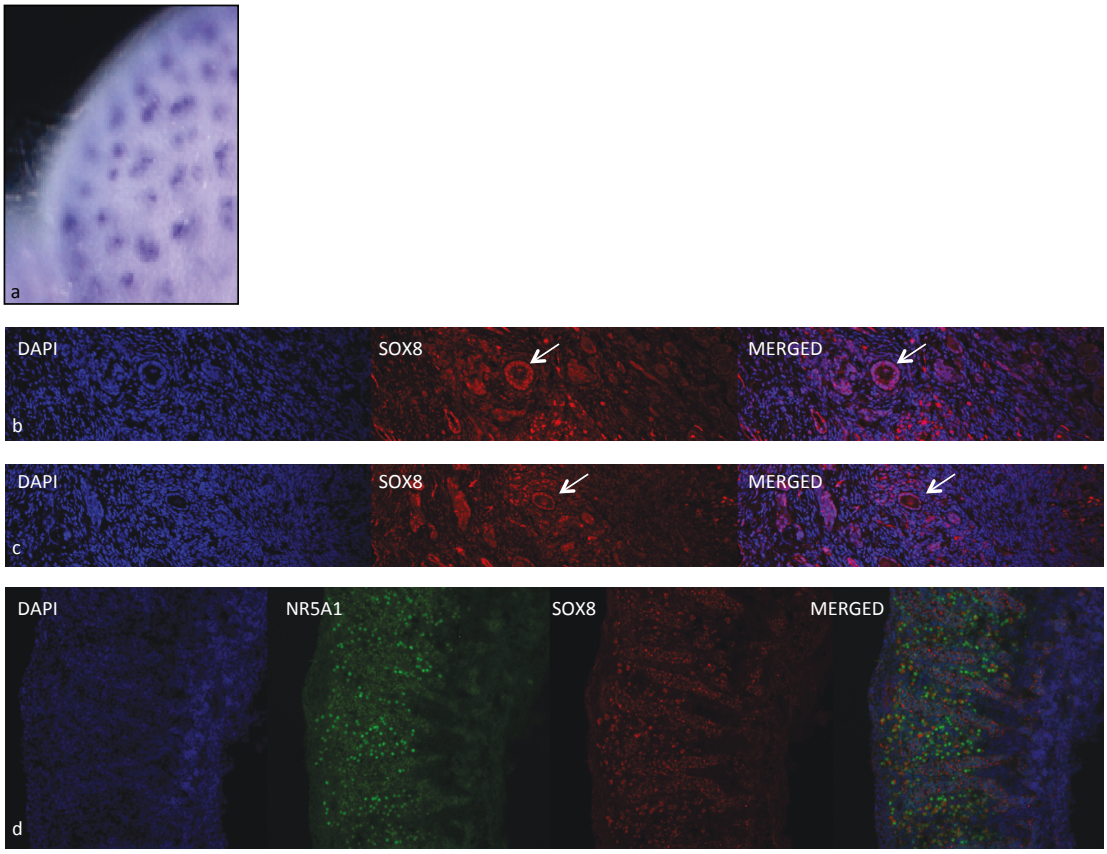
Supplementary Results

Supplementary Figure 1.



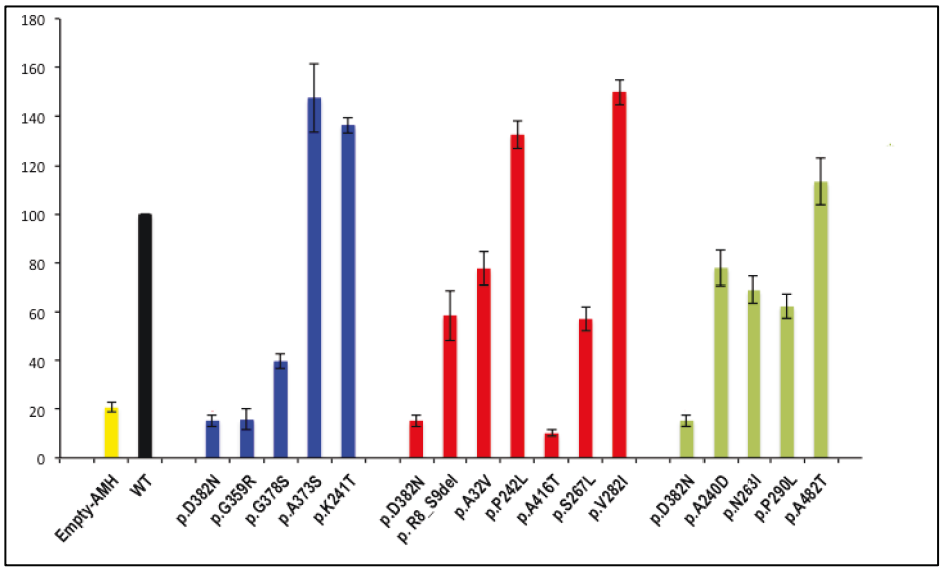
Supplementary Figure 1 Legend. Position of the telomeric breakpoint of the inversion Top, schematic representation of the position of the telomeric breakpoint showing all transcripts in the region and bottom, schematic representation showing the position of the BAC clones used in FISH experiments to map the breakpoint.

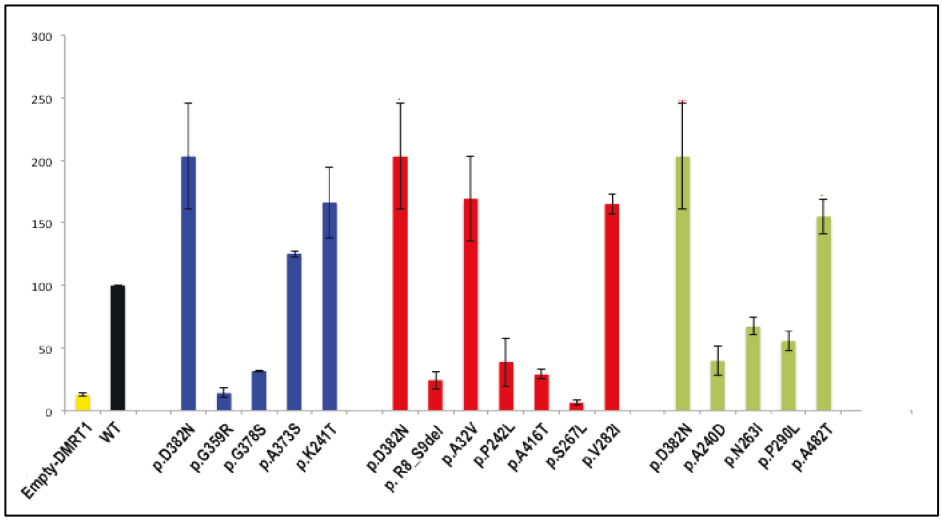
Supplementary Figure 2.

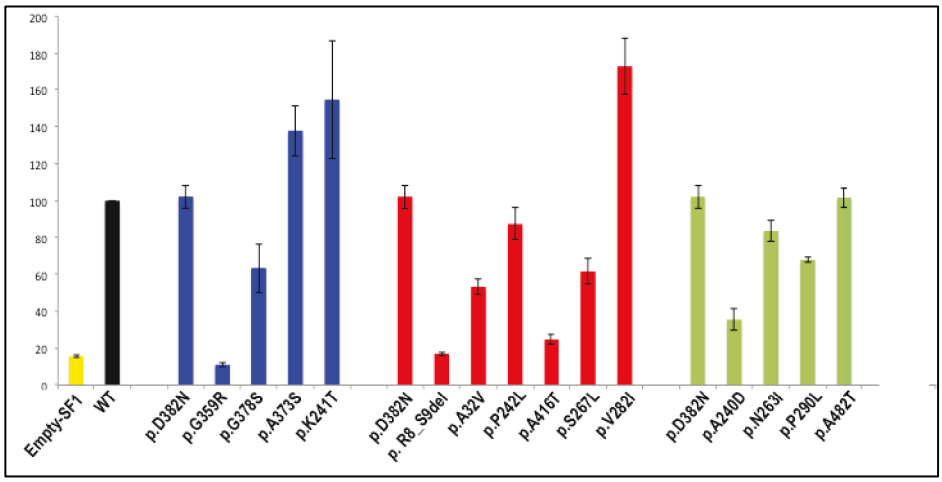


Supplementary Figure 2. Expression of SOX8 in murine and human tissues. (a) Murine *Sox8* expression by in situ hybridization in the developing testis cords at 15.5 d.p.c. (GUDMAP:6637, www.gudmap.org). (b, c) Immunohistochemistry showing SOX8 expression in granulosa cells (arrow) lining the follicles of the adult ovary (19 yrs old). (d) Expression of NR5A1 and SOX8 in human male fœtus at 11 weeks gestation. Expression is mainly nuclear with both Sertoli and Leydig cells expressing both proteins.

Supplementary Figure 3.







Supplementary Figure 3. The transcriptional activities of SOX8-WT and SOX8 mutants were studied using the human *AMH* and *NR5A1* and mouse *Dmrt1* promoters as reporters following transfection in HEK293-T cells. The data shown here represent the mean ± SEM of three independent experiments, each of which was performed in quadriplicate. The reporter constructs were transfected into HEK293-T cells with either the SOX8-WT or SOX8 mutants expression vectors. The results are expressed as a percentage of WT SOX8 activity. Data are shown as male infertility (blue), female infertility (red) or control (green). SOX8 p.D382N was observed in both male and female infertility and in 4 control samples of unknown fertility status. The p.A428T mutation was found exclusively in the normospermic control cohort. Although, the majority of mutants, including those in the general population cohort, showed absent or significantly reduced biological activity across all promoters studied (p.G359R, p.G378S, p.R8S9del, p.A416T, p.S267L, p.A240D, p.N263I, p.P290I), other mutants showed a promoter-dependent loss-of-function (p.D382N, p.A32V, p.P242L), whilst others showed a consistent and increased biological activity using different promoters (p.A373S, p.K241T, p.V282I). The p.A428T mutation showed WT activity and this mutation was found only general controls and a normospermic individual.

Discussion of publication

Sequencing the coding region of *SOX8* gene in 204 patients with idiopathic 46,XY DSD revealed a novel heterozygous mutation in a highly conserved Glu residue in the functional HMG box of the protein. The mutation imparted a dominant negative activity over both WT-SOX8 and WT-SOX9. Transient gene activation assays using Tesco enhancer element for *Sox9* show that SOX8p.Glu156Asp specifically failed to synergise with NR5A1 to transactivate the Tesco enhancer. Moreover SOX8p.E156D exerts a dominant negative effect by preventing synergistic activation of Tesco enhancer by NR5A1 and SOX9 as well as NR5A1 and SOX8. Its is possible, that this dominant negative effect imparted by mutant SOX8 alters the normal regulation of target genes involved in sex determination and development, thus leading to anomalous testis development. Mutations in *SOX8* may be rare in cases of DSD and the functional redundancy between *SOX8* and *SOX9* during testis development could be an evolutionary and biological strategy to reinforce male sex determination during development since at least in mouse, *Sox8* expression levels are dependent on *Sox9* levels (Chaboissier et al., 2004). The infertility data are consistent with mutations involving *SOX8* substantially contributing to human infertility.

In preparation: Mutations in GATA4 associated to 46,XY DSD

Introduction to publication

GATA binding protein 4 (GATA4) is a transcription factor that belongs to the GATA family of zinc finger proteins. GATA proteins play important role in cell fate choice, maturation and proliferation during embryonic development (Molkentin, 2000; Zaytouni et al., 2011).

Previously, our lab had reported a familial case of 46,XY DSD and CHD with affected individuals carrying a heterozygous mutation with variable penetrance, showing for the first time that point mutations in *GATA4* are associated with 46,XY DSD (Lourenço et al., 2011). Recently using whole exome sequencing we have identified a heterozygous GATA4p.D425N mutation in a woman with 46,XY DSD. This mutation has been previously described in association with CHD without a reported gonadal phenotype (Tomita-Mitchell et al., 2007; Zhang et al., 2009; Granados-Riveron et al., 2012). A second heterozygous mutation, GATA4p.S224P, has also been identified in the zinc finger domain of the protein in a patient presenting 46,XY DSD.

GATA4 mutation in association with 46,XY complete gonadal dysgenesis

Sandra Rojo[1], Rima Dada[2], Joelle Bignon-Topalovic[1], Ken McElreavey[1], Anu Bashamboo[1]

[1]Human Developmental Genetics, Institut Pasteur, 75724 Paris Cedex 15 France
[2]All India Institut of Medical Sciences, New Delhi, India

GATA-binding protein 4 (GATA4) is a transcription factor belonging to the GATA family of zinc finger proteins that play essential roles in cell fate, maturation and proliferation (Molkentin, 2000; Zaytouni et al., 2011). In humans and mice GATA4 is expressed in the somatic cells of the developing gonads before and during sex determination (Viger et al., 1998) and physically interacts with different proteins involved in sex-determination including FOG2 and NR5A1 to positively regulate the expression of crucial testis-determining genes *SRY, SOX9*, and *DMRT1* and as well as genes involved in steroidogenesis such as *STAR, CYP17A1, CYP19A1, HSD3B1, CYP11A1* and *INHA* (Miyamoto et al., 2008; Viger et al., 2008; Bergeron et al., 2015; Zaytouni et al., 2011; Lei and Heckert, 2004). Mice deficient in *Gata4* die *in utero* due to abnormalities in ventral morphogenesis and heart tube formation (Kuo et al., 1997; Molkentin et al., 1997). In humans, *GATA4* mutations have been described in patients with a range of congenital heart defects (CHD) (Garg et al., 2003). Additionally, deletions of 8p23 region comprising *GATA4* among other genes have been

associated with abnormalities such as hypospadias and cryptorchidism in a proportion of XY individuals with diaphragmatic hernia and CHD (Barber et al., 2013; Longoni et al., 2012; Wat et al., 2009). Recently, we reported a familial case of 46,XY DSD and CHD, where affected individuals carried a heterozygous mutation with incomplete penetrance, revealing the role for GATA4 in testis determination in human (Lourenço et al., 2011). Here, we describe a case of 46,XY complete gonadal dysgenesis with no evidence for CHD associated with a heterozygous GATA4p.D425N mutation. This mutation was previously reported in association with CHD without an apparent gonadal phenotype (Tomita-Mitchell et al., 2007; Zhang et al., 2009; Granados-Riveron et al., 2012). Our observations suggest that GATA4p.D425N affect the stability of GATA4 protein. This may create a haploinsufficiency leading to 46,XY DSD.

The study was approved by the Institutional Review Board of Institut Pasteur (RBM 2003/8) and Comité de Protection des Personnes Ile-de-France (N°IRB00003835). All participants in the study had approved written consent. The patient, of Indian origin, presented at 17 years of age with primary amenorrhea and elevated levels of FSH and LH and undetectable testosterone. Histology of the gonads revealed bilateral streak gonads. There were no other somatic anomalies. The karyotype was 46,XY. Direct sequencing of the *SRY* and *NR5A1* genes revealed wild-type sequences. Array comparative genomic hybridization using the Nimblegene 2.1M aCGH platform confirmed a normal ploidy. Exon enrichment was performed using Agilent SureSelect Human All Exon V4. Paired-end sequencing was performed on the Illumina HiSeq2000 platform with an average sequencing coverage of x50 as described elsewhere

(Bashamboo et al., 2014). Briefly, read files were generated from the sequencing platform via the manufacturer's proprietary software. Reads were mapped using the Burrows–Wheeler Aligner and local realignment of the mapped reads around potential insertion/deletion (indel) sites was carried out with the GATK version 1.6. SNP and indel variants were called using the GATK Unified Genotyper for each sample. SNP novelty was determined against dbSNP138. Datasets were filtered for novel or rare (MAF<0.01) variants. Candidate pathogenic mutations were confirmed by Sanger sequencing. Analysis of novel and rare variants (Minor allelic frequency <0.001) indicated one variant in the *GATA4* gene c.1273G>A (rs56208331) that is predicted to result in an p.D425N amino acid change. The parents were not available for study and consequently the mode of inheritance of the mutation is unknown. This mutation results in the replacement of an evolutionary conserved negatively charged aspartic acid residue by an uncharged asparagine residue.

We analyzed the effect of GATA4p.D425N mutation on the biological function of GATA4. TNT-*Gata4*WT and TNT-*Gata4*p.D425N vectors were transfected either independently or with the addition of an *NR5A1*-cherry vector in HEK 293-T cells. After 48hrs of transfection and using a GATA4 specific antibody, both wild-type and mutant GATA4 showed strong nuclear localization, alone and co-localize with NR5A1 (**Fig Sup1A**). We next analyzed the effect of p.D425N mutation on the ability of the protein to regulate the expression of different sex determining genes: *DMRT1, Sox9, AMH* and *NR5A1*. We do not observe a significant difference between the WT GATA4 and p.D425N in their ability to activate these promoters (**Fig Sup.2**). We next analyzed stability of

GATA4pD425N protein. Three independent experiments demonstrated a significant reduction in mutated versus WT GATA4 protein following the inhibition of protein synthesis (**Fig 1A**). The ability of GATA4p.D425N to physically interact with its protein partner SF1 analyzed by Co-immunoprecipitation (Co-IP) experiments showed that the GATA4p.D425N protein interacts with SF1, though reduced compared to the WT GATA4. This may be due to the instability of GATA4p.D425N (**Fig 1B**).

This mutation has been previously reported previously in four patients with ASD (2 patients), VSD, and tetralogy of Fallot although no functional analysis of the mutated protein has been performed (Tomita-Mitchell et al., 2007; Zhang et al., 2009; Granados-Riveron et al., 2012). The sex of the two patients carrying p.D425N in the report by Tomita-Mitchel et al., was not indicated but one of the patients inherited the mutation from a healthy mother. Similarly, the sex of the patient carrying the mutation was not indicated by Zhang et al., 2009. However, in the family reported by Granados-Riveron et al., 2012, the fertile mother carrying the p.D425N mutation, who presented with ASD, transmitted this mutation to her two healthy sons suggesting incomplete penetrance. Our functional analyses of biological activity of the mutant protein shows that GATA4p.D425N localizes to the nucleus, and is able to activate the expression of different genes involved in male sexual development including *DMRT1, NR5A1* and *AMH*, however the stability of the GATA4p.D425N protein was reduced compared to the WT protein. Such a reduction in stability of the mutated GATA4 protein may explain either the gonadal or heart defects that are seen in association with the mutation. Heart or male gonadal development may be

negatively impacted by the GATA4 mutation only when a critical biological threshold has been crossed either directly through reduced GATA4 protein levels or in combination with as yet unidentified genetic modifiers.

References

Barber, JCK, Rosenfeld, JA, Foulds, N, Laird, S, Bateman, MS, Thomas, NS, Baker, S, Maloney, VK, Anilkumar, A, Smith, WE, Banks, V, Ellingwood, S, Kharbutli, Y, Mehta, L, Eddleman, KA, Marble, M, Zambrano, R, Crolla, JA, Lamb, AN. 2013. 8p23.1 duplication syndrome; common, confirmed, and novel features in six further patients. Am. J. Med. Genet. A 161A: 487–500.

Bashamboo A, Brauner R, Bignon-Topalovic J, Lortat-Jacob S, Karageorgou V, Lourenco D, Guffanti A, McElreavey K. Mutations in the FOG2/ZFPM2 gene are associated with anomalies of human testis determination. Hum Mol Genet. 2014 Jul 15;23(14):3657-65. doi: 10.1093/hmg/ddu074. Epub 2014 Feb 18. PubMed PMID: 24549039.

Bergeron, F, Nadeau, G, Viger, RS. 2015. GATA4 knockdown in MA-10 Leydig cells identifies multiple target genes in the steroidogenic pathway. Reproduction 149: 245–57.

Garg, V, Kathiriya, IS, Barnes, R, Schluterman, MK, King, IN, Butler, CA, Rothrock, CR, Eapen, RS, Hirayama-Yamada, K, Joo, K, Matsuoka, R, Cohen, JC, Srivastava, D. 2003. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature 424: 443–7.

Granados-Riveron, JT, Pope, M, Bu'Lock, FA, Thornborough, C, Eason, J, Setchfield, K, Ketley, A, Kirk, EP, Fatkin, D, Feneley, MP, Harvey, RP, Brook, JD. 2012. Combined mutation screening of NKX2-5, GATA4, and TBX5 in congenital heart disease: Multiple heterozygosity and novel mutations. Congenit. Heart Dis. 7: 151–159.

Kuo, CT, Morrisey, EE, Anandappa, R, Sigrist, K, Lu, MM, Parmacek, MS, Soudais, C, Leiden, JM. 1997. GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. Genes Dev. 11: 1048–60.

Lei, N, Heckert, LL. 2004. Gata4 regulates testis expression of Dmrt1. Mol. Cell. Biol. 24: 377–88.

Longoni, M, Lage, K, Russell, MK, Loscertales, M, Abdul-Rahman, OA, Baynam, G, Bleyl, SB, Brady, PD, Breckpot, J, Chen, CP, Devriendt, K, Gillessen-Kaesbach, G, Grix, AW, Rope, AF, Shimokawa, O, Strauss, B, Wieczorek, D, Zackai, EH, Coletti, CM, Maalouf, FI, Noonan, KM, Park, JH, Tracy, AA, Lee, C, Donahoe, PK, Pober, BR. 2012. Congenital diaphragmatic hernia interval on chromosome 8p23.1 characterized by genetics and protein interaction networks. Am. J. Med. Genet. A 158A: 3148–58.

Lourenço, D, Brauner, R, Rybczynska, M, Nihoul-Fékété, C, McElreavey, K, Bashamboo, A. 2011. Loss-of-function mutation in GATA4 causes anomalies of human testicular development.

Miyamoto, Y, Taniguchi, H, Hamel, F, Silversides, DW, Viger, RS. 2008. A GATA4/WT1 cooperation regulates transcription of genes required for mammalian sex determination and differentiation. BMC Mol. Biol. 9: 44.

Molkentin, JD, Lin, Q, Duncan, SA, Olson, EN. 1997. Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. Genes Dev. 11: 1061–72.

Molkentin, JD. 2000. The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. J. Biol. Chem. 275: 38949–52.

Tomita-Mitchell, A, Maslen, CL, Morris, CD, Garg, V, Goldmuntz, E. 2007. GATA4 sequence variants in patients with congenital heart disease. J. Med. Genet. 44: 779–783.

Viger, RS, Guittot, SM, Anttonen, M, Wilson, DB, Heikinheimo, M. 2008. Role of the GATA family of transcription factors in endocrine development, function, and disease. Mol. Endocrinol. 22: 781–98.

Viger, RS, Mertineit, C, Trasler, JM, Nemer, M. 1998. Transcription factor GATA-4 is expressed in a sexually dimorphic pattern during mouse gonadal development and is a potent activator of the Müllerian inhibiting substance promoter. Development 125: 2665–75.

Wat, MJ, Shchelochkov, OA, Holder, AM, Breman, AM, Dagli, A, Bacino, C, Scaglia, F, Zori, RT, Cheung, SW, Scott, DA, Kang, S-HL. 2009. Chromosome 8p23.1 deletions as a cause of complex congenital heart defects and diaphragmatic hernia. Am. J. Med. Genet. A 149A: 1661–77.

Zaytouni, T, Efimenko, EE, Tevosian, SG. 2011. GATA transcription factors in the developing reproductive system. Adv. Genet. 76: 93–134.

Zhang, W, Li, X, Ma, Z, Zhang, J, Zhou, S, Li, T, Shi, L, Li, Z. 2009. GATA4 and NKX2.5 gene analysis in Chinese Uygur patients with congenital heart disease. Chin. Med. J. (Engl). 122: 416–9.

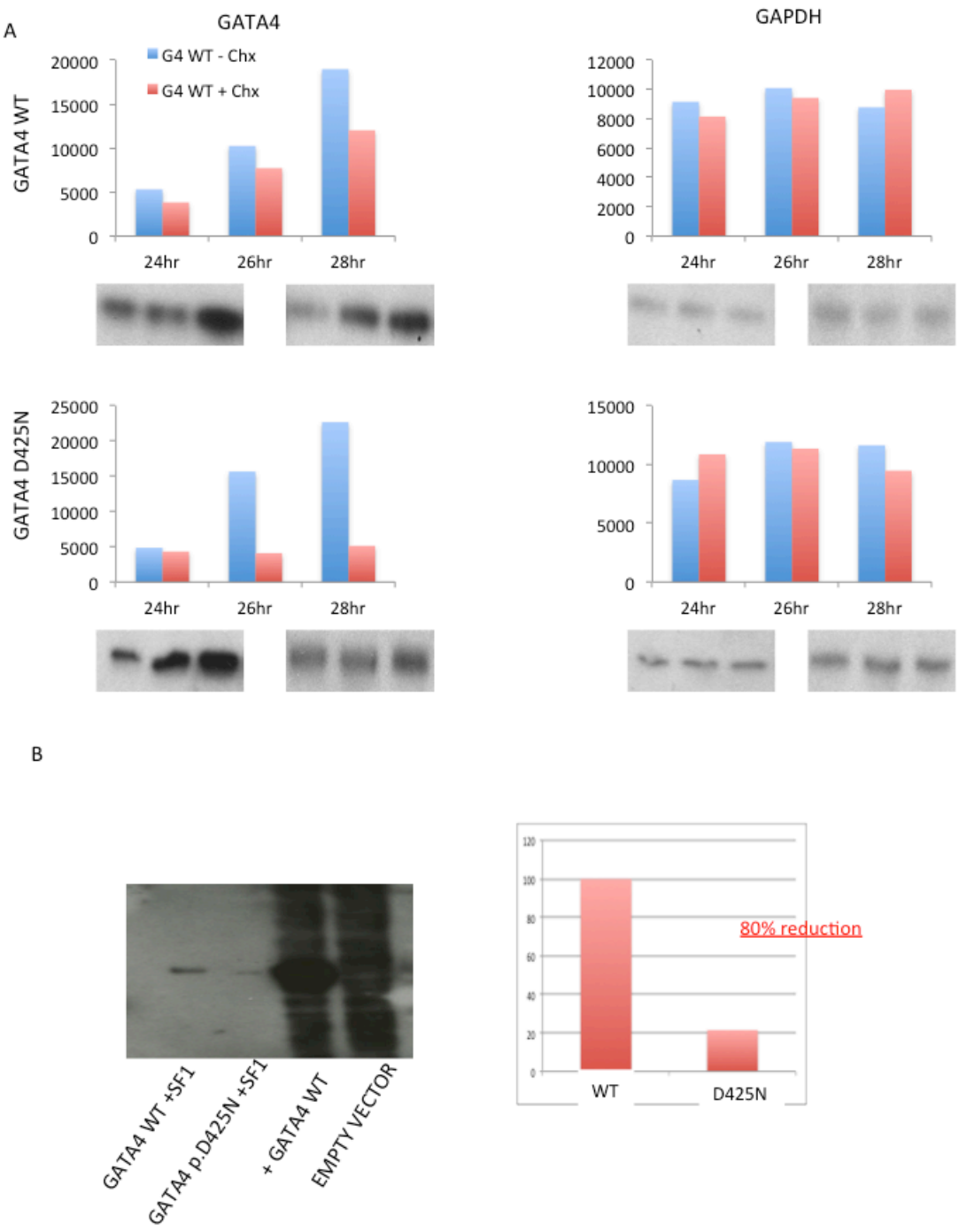
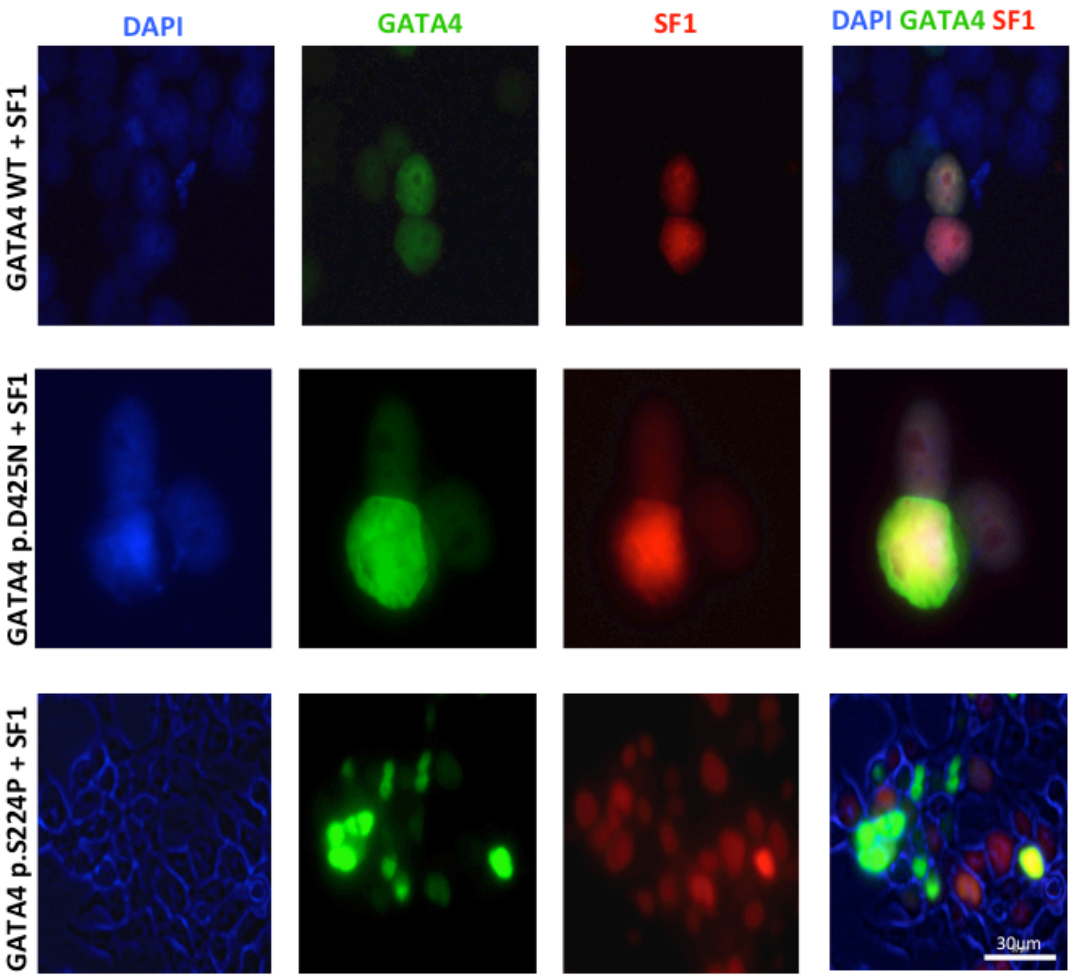
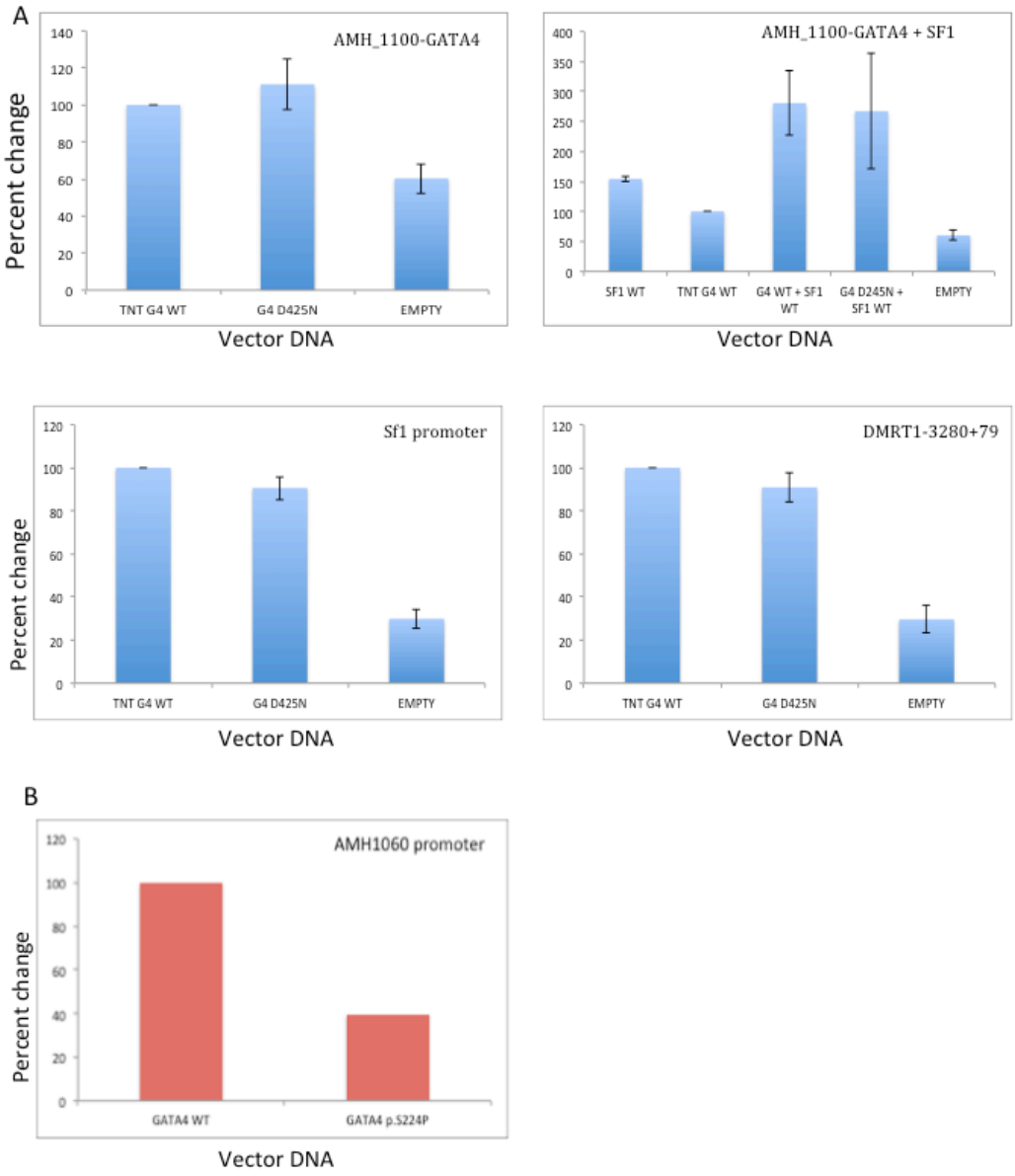


Figure 1 Protein analysis GATA4 mutants. (A) Time-course analysis of GATA4 protein production and stability by cyclohexamide inhibition of GATA4 WT and GATA4p.D425N. GATA4p.D425N is less stable than GATA4 WT. B) Protein-protein interaction of GATA4 and SF1 measured by Co-IP using SF1 antibody to co-immunoprecipitate SF1-GATA4 complexes, and analyzed by Western Blot using a GATA4 antibody. GATA4p.D425N show a reduced interaction with SF1, probably due to the instability of GATA4p.D425N



Supplementary Figure 1 Cellular localization of GATA4. Human embryonic kidney HEK 293T cells were transfected with *Gata4* WT or *Gata4* mutants and *NR5A1*-cherry, 48hrs after transfection cells were fixed and stained with a GATA4 antibody. Similar to GATA4 WT, p.D425N and p.S224P show a strong nuclear localization (green) and co-localize with NRA51/SF1 (red) to the nucleus (DAPI/blue stained).



Supplementary Figure 2 Transcriptional regulation reporters comprising of promoters of gonad specific genes (A) Both GATA4 WT and GATA4p.D425N show no difference in the activation of *AMH*, *NR5A1* or *DMRT1* promoters. (B) Preliminary results with GATAp.S224P show a reduction in the ability of the mutant to activate *AMH* promoter.

Supplementary Materials and Methods

Plasmid constructions.

Plasmid pCMV-TNT-*Gata4,* pCMX-*NR5A1*, pcs2-*FOG2*, Tesco and AMH reporter vectors are previously described in Lourenço et al, 2011. Reporter vector with *DMRT1 promoter was a gift from* from Dr. David Zarkower (Univ. of Minnesota).

Site-Directed Mutagenesis

Plasmids containing mutant Gata4p.D425N and p.S224P were generated by site directed mutagenesis (QuikChange, Stratagene), using *TNT-Gata4* vector as a template as described elsewhere (Braman et al., 1996). Mutant plasmids were verified by direct sequencing before any functional analysis.

Transient expression analysis

Transient gene expression assays for *Gata4* function were performed in 96 well plates, using HEK 293T cells at a concentration of $1x10^4$ per well, Fugene 6 transfection reagent (Roche 1814443) and a Dual-Glo luciferase reporter system (Dual Go Luciferase Substrate E297B) and a reporter vector for *Renilla* activity (pRLSV40 Renilla luciferase). pCMV-TNT-*Gata4* WT, p.D425N or p.S224P (10ng/well) were transfected with reporter vectors containing different sex determining genes promoters (10ng/well) and pRLSV40-*Renilla* (5ng) (Promega). 48hrs later, cells were lysed and luciferase assays were performed in a FLUOstar Optima fluorescence microplate reader (BMG Labtech). Data was

normalized for *Renilla* activity. Results are shown as the mean of 3 independent experiments, performed in triplicate.

Immunochemistry

HEK 293T cells were seeded in an 8 well chamber slide using a concentration of 2 x 10^4 per well. Transfected using pCMV-TNT-*Gata4* WT, p.D425N or p.S224P without and with pCMX-*NR5A1*-cherry. 48hrs after transfection, chamber slides were fixed using 4%PFA (15735-605), blocked for 15 min with 1X, and incubated 1hr with first antibody (anti-Gata4: sc25310), washed 3X with PBS, incubated with secondary fluorescent antibody (goat anti mouse: Life-R37120), washed 3X, covered with mounting media with an anti-fading reagent (Life: P36931) and sealed.

Protein extraction

HEK 293T cells were seeded in 6 well plate at a concentration of 6x10^5 per well. 24hrs later cells were transfected using the vectors pCMV-TNT-*Gata4* WT, p.D425N or p.S224P along with *pacGFP-NR5A1*. 48hrs after transfection, cells were incubated with IP lysis buffer (Thermo 87788) and protease inhibitor (Life Tech 78429) and whole cell protein was transferred to a clean eppendorf, and measured by a Bradford assay.

Co-Immunoprecipitation assay

After Bradford assay to measure proteins, 250µg of protein extracts were incubated for 1 hr at 4°C with 1µg of rabbit IgG (Santa Cruz sc-2027) and 20µl of resuspended protein A/G PLUS agarose (Santa Cruz SC-2003). Samples were

centrifuged (1000G, 5min at 4°C), the supernatant collected and incubated with 5µl of NR5A1 antibody (ab65815) for 1hr at 4°C. Subsequent, 20µl of A/GPLUS agarose (SC-2003) were added and samples were incubated on a rocking platform over night at 4°C. Samples were centrifuged (1000G, 5 min at 4°C), and the pellet was washed 4X with 1ml of cold PBS. After final centrifugation, pellet was resuspended in 40µl of SDS laemmli buffer + 4µl of DTT 1M. Samples were denatured for 5 min at 95°C and 20 µl were migrated on a fractionated on a 10% SDS-PAGE gel and transfer to a PVDF membrane. Western Blot was performed using GATA4 antibody for detection (sc25310), followed by a secondary antibody HRP conjugated (ab6728) and visualized by chemioluminescence (Thermo-32106).

Time-course analysis of GATA4 protein production and stability by cyclohexamide inhibition

We transfected HEK 293-T cells with pCMV-TNT-*Gata4* wild type and p.D425N vectors and 24hrs later, protein biosynthesis was inhibited by the addition of cyclohexamide (Sigma C4859) every 2hrs. Cellular protein (IPLysis buffer) was isolated 24hrs later at different time points (24h, 26h, 28h). Proteins were measured and equivalent amount of both WT and mutated proteins were fractioned in a 10% SDS/PAGE gel, transferred to a PVDF membrane (Amersham 10600023) and incubated with specific GATA4 antibody (sc25310) and GAPDH antibody (ab9484) and visualized by chemioluminescence (Thermo-32106).

Discussion of Publication

The mutant GATA4p.D425N localizes to the nucleus and is able to activate the reporter constructs for many sex determination genes including *AMH, SOX9*, and *NR5A1* in *in-vitro* assays. However, GATA4p.D425N severely impairs the stability of the protein resulting in less amount of protein for regulation of the downstream targets thus causing haploinsufficiency and possibly 46,XY DSD. It is known that human sex-determination is highly sensitive to small changes in gene dosages (Bouma et al., 2007). Moreover, in mice, reduction of *Gata4* expression in the testis of *siGata4* transgenic mice results in the significant reduction (almost complete abolition) of Gata4 downstream target genes in male gonads as *Amh* and *StAR*; in female gonads however the expression of downstream genes was minimally affected. Interestingly, although silencing of *Gata4* was efficient, Gata4 expression was detected in heart after 38 days of silencing and it didn't affect heart function at low levels of Gata4 (20-30% of normal expression) suggesting that male gonad development is more sensitive to small changes in gene dosage as compared to other organs (Thurisch et al., 2009).

WES analysis of a second case of 46,XY DSD has identified a novel heterozygous mutation in *GATA4*: GATA4p.S224P, this mutation is maternally acquired and affects the zinc finger domain of the protein. *In vitro* functional assays show that this mutation does not change the nuclear localization of GATA4 (**FigSup1**). Preliminary transient gene expression analysis of *AMH* reporter show that GATA4p.S224P show a reduction in its ability to activate *AMH* as compared to

WT-GATA4 (**Fig Sup2B**). Unlike the GATA4p.D425N mutation, GATA4p.S224P may contribute to the phenotype by dysregulation of the downstream targets.

By personal communication from different collaborators, we are aware of the existence of several other mutations in *GATA4* associated with 46,XY DSD. This suggests that mutations in *GATA4* may have a bigger contribution to 46,XY DSD than previously thought. A systematic analysis will establish the incidence of *GATA4* mutations in 46,XY DSD and possibly lead to a useful marker for routine molecular diagnosis of the phenotype.

Refining The Regulatory Region Upstream Of SOX9 Associated With 46, XX Testicular Disorders Of Sex Development (DSD)

Introduction to publication

In XY embryos, the correct time and dosage of *SOX9* expression is pivotal for testicular differentiation. If *SOX9* expression reaches a critical threshold then the bipotential gonad will differentiate into Sertoli cells and conversely if not then it will become ovarian granulosa cells. Genomic variations resulting in specific alterations of *SOX9* expression in the developing gonads are linked to disorders of sex development (DSD). Human DSD are congenital conditions in which the development of chromosomal, gonadal or anatomical sex is atypical. 46,XX DSD occur very rarely with a rate of 1 in 20,000-25,000 newborn males (Kousta et al.). Recent studies have reported cases of 46,XX DSD (*SRY* negative) with testis development carrying duplications in a region 600kb upstream of *SOX9*. In addition, 46,XY females with DSD carrying deletions in this region have also been reported, identifying a human gonad specific regulatory element involved in sex determination called *RevSex* (Benko et al., 2011; Cox et al., 2011; Xiao et al., 2013; Kim et al., 2015). In this study we redefined a minimal region of the *SOX9* putative enhancer element that is located approximately 600 kb upstream of *SOX9* that is associated with 46,XX DSD.

CLINICAL REPORT

AMERICAN JOURNAL OF medical genetics PART A

Refining the Regulatory Region Upstream of *SOX9* Associated With 46,XX Testicular Disorders of Sex Development (DSD)

Capucine Hyon,[1,2,3]* Sandra Chantot-Bastaraud,[1] Radu Harbuz,[4] Rakia Bhouri,[1] Nicolas Perrot,[5] Matthieu Peycelon,[2] Mathilde Sibony,[6] Sandra Rojo,[7] Xavier Piguel,[8] Frederic Bilan,[4] Brigitte Gilbert-Dussardier,[4,9] Alain Kitzis,[4] Ken McElreavey,[7] Jean-Pierre Siffroi,[1,2,3] and Anu Bashamboo[7]

[1]AP-HP, Hôpitaux Universitaires Est Parisien, Hôpital Trousseau, Service de Génétique et d'Embryologie médicales, Paris, France

[2]INSERM UMR_S933, Paris, France

[3]UPMC Univ Paris 06, UFR de Médecine Pierre et Marie Curie, Paris, France

[4]Service Génétique Médicale, CHU Poitiers, France

[5]Department of Radiology, AP-HP, Hôpitaux Universitaires Est Parisien, Hôpital Tenon, Paris, France

[6]Department of Pathology, AP-HP, Hôpitaux Universitaires Est Parisien, Hôpital Tenon, Paris, France

[7]Institut Pasteur, Human Developmental Genetics, Paris, France

[8]Service Endocrinologie, CHU Poitiers, France

[9]Centre de Référence Anomalies du Développement Ouest, CHU Poitiers, France

Manuscript Received: 29 December 2014; Manuscript Accepted: 5 March 2015

Disorders of Sex Development (DSD) are a heterogeneous group of disorders affecting gonad and/or genito-urinary tract development and usually the endocrine-reproductive system. A genetic diagnosis is made in only around 20% of these cases. The genetic causes of 46,XX-*SRY* negative testicular DSD as well as ovotesticular DSD are poorly defined. Duplications involving a region located ~600 kb upstream of *SOX9*, a key gene in testis development, were reported in several cases of 46,XX DSD. Recent studies have narrowed this region down to a 78 kb interval that is duplicated or deleted respectively in 46,XX or 46,XY DSD. We identified three phenotypically normal patients presenting with azoospermia and 46,XX testicular DSD. Two brothers carried a 83.8 kb duplication located ~600 kb upstream of *SOX9* that overlapped with the previously reported rearrangements. This duplication refines the minimal region associated with 46,XX-*SRY* negative DSD to a 40.7–41.9 kb element located ~600 kb upstream of *SOX9*. Predicted enhancer elements and evolutionary-conserved binding sites for proteins known to be involved in testis determination are located within this region. © 2015 Wiley Periodicals, Inc.

Key words: 46,XX DSD; duplication; regulatory element; *SOX9* gene; CNV

How to Cite this Article:
Hyon C,Chantot-Bastaraud S, Harbuz R, Bhouri R, Perrot N, Peycelon M, Sibony M, Rojo S, Piguel X, Bilan F, Gilbert-Dussardier B, Kitzis A, McElreavey K, Siffroi J-P, Bashamboo A. 2015. Refining the Regulatory Region Upstream of *SOX9* Associated With 46,XX Testicular Disorders of Sex Development (DSD).
Am J Med Genet Part A 9999A:1–8.

INTRODUCTION

Disorders of sex development (DSD) are congenital conditions in which the development of chromosomal, gonadal, or anatomical sex is atypical [Hughes et al., 2006]. DSD covers a wide spectrum of

Conflict of interest: None.
Grant sponsor: Agence Nationale de la Recherche-GIS Institut des Maladies Rares; Grant sponsor: March of Dimes Foundation Research; Grant number: 1-FY07-490; Grant sponsor: EuroDSD in the European Community's Seventh Framework Programme FP7/2007–2013; Grant number: 201444.
*Correspondence to:
Capucine Hyon, Service de Génétique et d'Embryologie Médicales, 26 rue du Dr Arnold Netter, 75012 Paris, France.
E-mail: capucine.hyon@trs.aphp.fr
Article first published online in Wiley Online Library (wileyonlinelibrary.com): 00 Month 2015
DOI 10.1002/ajmg.a.37101

phenotypes. 46,XY DSD includes 46,XY complete or partial gonadal dysgenesis as well as undermasculinization of an XY male due to a defect in androgen synthesis or action. At the opposite, most 46,XX DSD correspond to the virilization of an XX female due to androgen excess. Most often, 46,XX testicular DSD, a rare pathology affecting 1 in 20,000–25,000 newborn males, [Kousta et al., 2010] are *SRY* positive and have Y-to-X translocations while most cases of 46,XX ovotesticular DSD are mosaics or chimeras [López et al., 1995]. However, some 46,XX males do not carry the *SRY* gene and the genetic cause of these DSD cases remains poorly defined.

Human sex determination is a tightly controlled and highly complex process where the bipotential gonad anlage develops to form either a testis or an ovary. In 46,XY males the expression of the testis determining gene *SRY* leads to the upregulation of the *SOX9* gene and the development of Sertoli cells that produce AMH and initiate the series of events that lead to testis formation. Prior to the expression of *Sry* in mice, *Sox9* expression is at basal levels in both male and female gonadal primordial cells and at 11.5 days dpc, following the expression of *Sry*, it is up-regulated in males and down-regulated in females. Upregulation in males is due, at least in part, to a synergistic action of Sry with NR5A1/SF1, through binding to a testis-specific *Sox9* enhancer named TESCO (Testis-specific Enhancer of *SOX9* Core), located approximately 13 kb upstream of *Sox9* [Sekido and Lovell-Badge, 2008]. In XX-*SRY* negative males duplication of the entire *SOX9* gene was described in a boy with severe penile hypospadias and XX karyotype [Huang et al., 1999]. Rearrangements involving other closely related members of the *SOX* gene family have also been associated with testis development in 46,XX individuals including the *SOX3* and *SOX10* genes [Polanco et al., 2010; Sutton et al., 2011; Moalem et al., 2012]. Mutations or deletions in *RSPO1* and *WNT4*, genes involved in the *WNT4*-β catenin pathway leading to *SOX9* downregulation, have been described in syndromic forms of 46,XX *SRY*-negative testicular DSD [Parma et al., 2006; Mandel et al., 2008].

As well as being essential for testis formation, *SOX9* plays a key role in chondrogenesis [Bell et al., 1997; Ng et al., 1997]. Consequently, mutations in the coding sequences of *SOX9* are associated with campomelic dysplasia and, in about 70% of affected 46,XY individuals, with a range of anomalies of testis development [Foster et al., 1994; Wagner et al., 1994]. Translocations and copy number variation both 5′ and 3′ to *SOX9* gene, are associated with a milder phenotype as compared with the intragenic mutations. Large duplications (>1 Mb) 5′ to *SOX9* are associated with brachydactyly-anonychia (Cooks syndrome) and deletions located >1.3 Mb 5′ and 3′ to *SOX9* are associated with micrognathia, cleft palate and glossoptosis (Pierre–Robin sequence) [Benko et al., 2009; Kurth et al., 2009]. In both examples there is an apparently normal testis development in 46,XY men and the precise phenotype associated with *SOX9* mutations depends on the position and size of the rearrangement. Indeed, four recent studies have reported 46,XX-*SRY* negative individuals who presented with testis development. Each carried duplications involving a chromosomal region located approximately 600 kb upstream of *SOX9* [Benko et al., 2011; Cox et al., 2011; Vetro et al., 2011; Xiao et al., 2013]. In contrast, 46,XY females with gonadal dysgenesis have been reported to carry overlapping deletions of this region [Benko et al., 2011; Bhagavath et al., 2014]. These individuals presented with no other associated somatic anomalies, suggesting that a regulatory element involved specifically in human sex-determination (termed *RevSex)* is located in this region [Benko et al., 2011].

Here, we describe three novel 46,XX-*SRY* negative patients, two brothers and an unrelated man with testicular DSD and azoospermia, who carry a microduplication upstream of *SOX9*. The detailed analyses of these patients and the incorporation of the published datasets have enabled the minimal critical region located ∼600 kb upstream of *SOX9* to be resolved to an approximately 40kb element that contains putative enhancers elements and DNA-binding sites for known factors to be involved in early testis formation.

MATERIALS AND METHODS

Clinical Reports

Patients 1 and 2 were two brothers referred for infertility at the age of 30 and 31 respectively. Both had normal male genitalia without any signs of undervirilization. Clinical examination in Patient 1 revealed a reduced testicular volume (10 ml; normal range 18–30 ml). Semen analysis revealed an azoospermia. Ultrasound examination of the scrotum in Patient 1 was normal and pelvic ultrasound examination showed a normal male genital tract without Mullerian remnants. Testicular biopsy was performed to rule out a possible 46,XX/46,XY chimera. It revealed atrophic seminiferous tubules containing only eosinophilic Sertoli cells suggestive of testicular dysgenesis. There was no evidence of spermatogenesis and the interstitium showed hyperplasia of Leydig cells.

Patient 3 was a 45-year-old male who was referred for clinical investigation after five years of infertility. He presented with a normal male phenotype and had a normal libido and no erectile dysfunction. Ultrasound investigation showed bilaterally hypotrophic testes (right 15 mm × 7 mm and left 18 mm × 12 mm) with calcifications, seminal vesicle hypoplasia and normal prostate. Semen analysis showed azoospermia.

Genetic Analysis

Genetic analyses were performed after obtaining patient's informed consent. Cytogenetic analysis was performed using GTG-banding techniques on metaphase chromosomes obtained by standard procedures from peripheral blood lymphocytes. FISH analysis was carried out using chromosome X specific probe DXZ1 and probes for *SRY* and *SOX9* locus.

Genomic DNA was extracted for each patient from peripheral blood samples using FlexiGene DNA Kit (Qiagen, Valencia, CA) according to manufacturer's instructions.

For patients 1, 2 and 3, SNP array analysis was performed using the HumanCytoSNP-12 BeadChip from Illumina (San Diego, CA). Data analysis was performed using Illumina's Genome Studio Genotyping Module software allowing the identification of both copy number variations (CNVs) with the Log R ration (LRR) and regions of copy-neutral loss of heterozygosity (LOH) with the B allele frequency. Previous array CGH for patient 3 was performed using the Human Genome CGH Microarray Kit 105A from Agilent (Santa Clara,

CA). DNA sequence information was made according to the UCSC Genome Browser (http://genome.ucsc.edu/; February 2009, Assembly, hg19).

Boundaries of the duplication for Patient 2 were confirmed by long-range PCR using primers in the 5' and 3' parts of the putative duplicated region (Supplemental data). The long-range PCR amplicon was generated using Roche's Expand Long Template PCR System (Roche Diagnostics, Meylan, France) according to the following conditions:5 μl of 10× Expand Long Template Buffer 3 (including 27.5 mM MgCl2 and detergents), 8.75 μl of a 2 μmol/L of dNTPs, 0.75 μl of Expand Long Template Enzyme Mix, 3 μl of a 5 μmol/L of each primer, 150 ng of genomic DNA, 26.5 μl of water in a total reaction volume to 50 μl. Reaction was cycled under following conditions: 94 °C for 2 min; 94 °C for 10 sec; annealing temperature of 60 °C for 30 sec; 68 °C for 10 min; cycle 10× from step 2; 94 °C for 15 sec; annealing temperature of 60 °C for 30 sec; 68 °C for 10 min plus 20 sec per cycle; cycle 30× from step 5; 68 °C for 7 min; and hold at 4 °C. Internal primers were used to sequence the amplicon. Data were analyzed using SeqScape® Software v2.7 (Applied Biosystems, Foster City, CA) and Sequencing Analysis Software (Applied Biosystems).

Quantitative PCR

For the qPCR analysis, genomic DNA from the Patients 1 and 2 and two other individuals with two copies of the region (as indicated by the array analysis) were evaluated for quantity and quality via an ND-8000 Spectrophotometer (NanoDrop®) and agarose gel electrophoresis. Only intact genomic DNA was used for qPCR analysis. Six sets of Taqman® CNV assays were custom designed, each at approximately 9 kb interval, to span the entire duplication (Supplemental data, Table I S1).

Each DNA sample was analyzed in triplicate and reactions were conducted in a MicroAmp fast 96-well optical reaction plate (P/N 4346906; Applied Biosystems) sealed with an optical adhesive cover (P/N 4311971; Applied Biosystems) on 7900HT System (96-Well Block) real-time PCR System. Each individual reaction contained CNV assay (containing two primers and a FAM™ dye labeled MGB probe), RNaseP assay (containing two primers and aVIC® dye-labeled TAMRA™ probe), TaqMan® Genotyping Master Mix (containing AmpliTaq Gold® DNAPolymerase, and dNTPs), 15 ng genomic DNA and water to a final reaction volume to 20 μl. Reactions were held at 95 °C for 10 min and then cycled 40 times through 95 °C for 15 sec and 60 °C for 1 min. The data was collected using Sequence Detection Software v2.3 (Applied Biosystems). qPCR data was analyzed using a delta delta Ct algorithm with RnaseP as an endogenous control. Calibration with a control sample with two copies of the region was used to calculate delta delta Ct values and assign copy numbers by CopyCaller® software.

RESULTS

For all the patients initial cytogenetic standard G-banding analysis (550 bands) showed an homogenous 46,XX karyotype. FISH studies with both the *DXZ1* and *SRY* probes confirmed the existence of two X chromosomes and the absence of the *SRY* gene in all cells analyzed (Supplemental data, Figure S1). Hybridization with the *SOX9* probe was normal for Patients 1 and 2 (not performed for Patient 3) (Supplemental data, Figure S1).

In Patient 1 and 2, SNP array analysis identified a 77,078 to 85,175 bp duplication in chromosome 17 long arm, upstream from *SOX9* gene, extending between 69,490,856 (normal)–69,493,863 (duplicated) and 69,570,941 (duplicated)–69,576,031 (normal) (Fig. 1). For Patient 3, array CGH identified an overlapping

FIG. 1. SNP array results for Patients 1, 2 and 3 of the duplicated region. Note the increase in the log R Ratio within the duplicated region (grey box) (Scales: for Patients 1 and 2:1pixel=1 kb, Patient 3:1pixel= 2kb)

duplication of 194 kb in the same region which was reduced to 140,572–152,536 bp between 69,435,809 (normal)–69,441,277 (duplicated) and 69,581,849 (duplicated)–69,588,345 (normal) using the SNP array (Fig. 1). These duplications were confirmed by independent qPCR analysis for Patients 1 and 2 (data not shown) and sequencing the boundaries of the duplication for Patient 2 which allowed us to identify the exact size of the duplication extending between 69,491,366 and 69,575,195 (83,829 bp duplication) (Fig. 2).

Xiao et al, described an XX male with a 74 kb duplication overlapping with this region which proximal breakpoint was located between 69,533,305 bp (normal) and 69,534,526 bp (duplicated) [Xiao et al., 2013]. Taking this data and our breakpoint into account we can redefine the minimal region duplicated to a minimum size of 40,669 bp and a maximum size of 41,890 bp (Supplementary Figure S2). Detailed analysis of this region indicates that it contains two non-overlapping putative enhancer elements. The Encyclopedia of DNA Elements (ENCODE) project datasets indicate a weak or poised enhancer at chr17:69,568,380–69,569,850. Chromatin at this position is enriched for experimentally determined CTCF-binding from human mammary epithelial cells (HMEC). The primary role of CTCF is thought to be in regulating the 3D structure of chromatin. CTCF creates boundaries between topologically-associating domains in chromosomes and anchors DNA to cellular structures like the nuclear lamina and thereby facilitates interactions between transcription regulatory sequences [Murrell, 2011]. This region is also associated with acetylation of histone H3 at lysine 27 (H3K27ac) and enriched in monomethylation of histone H3 at lysine 4 (H3K4me1) from a variety of cell lines (Supplementary Figure S2). Both H3K4 methylation and H3K27 acetylation are epigenetic marks that are generally associated with gene activation [Eissenberg and Shilatifard, 2010; Smith and Shilatifard, 2014].

A strong putative enhancer element is located at chr17:69,544,206–69,546,005 and it is associated with H3K27Ac and H3K4me1 enrichment, as well as the repressive mark H3K9me3 which was identified experimentally by ChIP-seq. The histone acetyltransferase EP300, which is a known co-activator of *SOX9*-dependent gene expression, binds to this element (chr17:69,544,741–69,545,056) [Furumatsu et al., 2009]. This was determined experimentally using neuroblastoma cell line treated with retinoic acid (SK-N-SH_RA).

Multiz alignment results obtained from the UCSC genome browser show 12 short evolutionary-conserved regions that are enriched for transcription factors binding sites that are present in the human, mouse and rat (Supplementary Figure S2). These include factors that are strongly expressed in the male supporting cell lineages of the mouse testis at the time of sex-determination (E11.5–E13.5) and include NFIL3 (also known as E4BP4; chr17:69,544,862–69,544,873), PBX1 (chr17:69,563,419–69,563,427) and GATA1 (chr17:69,560,189–69,560,198 and chr17:69,560,189–69,560,198). In addition, to these factors TRANSFAC and TFSEARCH analysis of the human sequence predicted multiple DNA-binding sites for transcription factors known to be involved in the early stages of sex-determination such as SRY, WT1, SOX9, NR5A1, and LHX9 [Bashamboo and McElreavey, 2013]. The region also contains four predicted binding sites for the evolutionary conserved transcription factor DMRT1 at chr17:69,543,454–69,543,468, chr17:69,554,542–69,554,559, chr17:69,559,047–69,559,062, and chr17:69,569,634–69,569,648.

DISCUSSION

Multiple tissue specific enhancers located upstream and downstream from *SOX9* have been described [Velagaleti et al., 2005; Bagheri-Fam et al., 2006; Gordon et al., 2009]. In the human, a large duplication (46,XX,dup(17)(q23.1q24.3)/46,XX) involving *SOX9* was associated with female to male sex reversal in a 46,XX individual [Huang et al., 1999], however mutations or translocations involving the TESCO region have not yet been reported [Georg et al., 2010]. This suggests that in human gonad additional regulatory elements may be involved in the regulation of SOX9 expression. This hypothesis is supported by recent reports that identified rearrangements involving elements situated ~600 kb upstream of *SOX9* in association with 46,XX testicular DSD and termed the RevSex element [Benko et al., 2011; Cox et al., 2011; Vetro et al.,

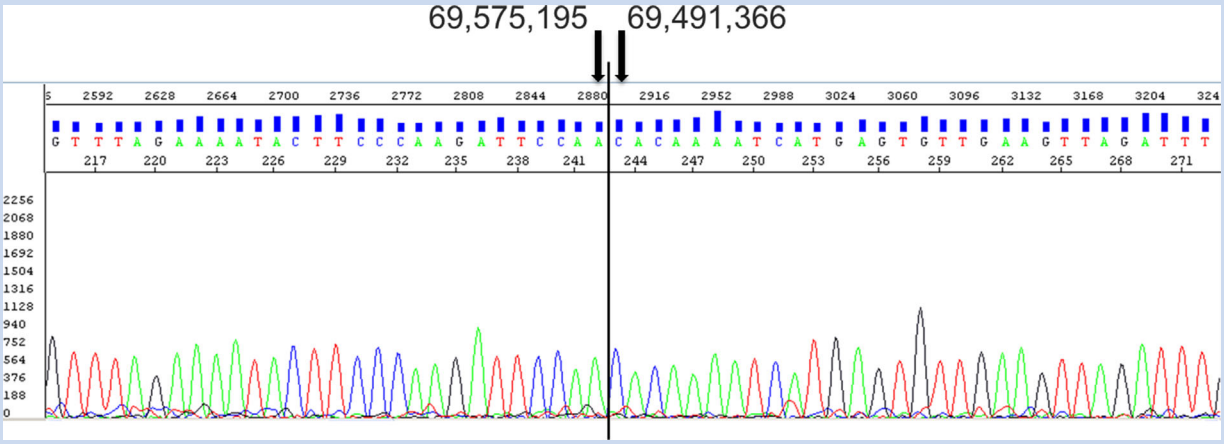


FIG. 2. Exact boundaries of the duplication for Patient 2 (according to the UCSC Genome Browser (http://genome.ucsc.edu/; February 2009, Assembly, hg19)).

TABLE I. Phenotypic Features and Characteristics of the Rearrangement of Individuals With Rearrangement Upstream of *SOX9*

			Benko et al. (2011)						Bhagavath et al. (2014)	Our patients	
	Cox et al. (2011)	Vetro et al. (2011)	DSD1	DSD2	DSD3	DSD4 Case 1	DSD4 Case 2	Xiao et al. (2013)	Case III3	Patients 1 & 2	Patient 3
Karyotype	46,XX	46,XX	46,XX	46,XX	46,XX	46,XY	46,XY	46,XX	46,XY	46,XX	46,XX
SRY	–	–	–	–	–	NA	NA	–	+	–	–
Phenotype	Male	Male	–	Perineal hypospadias	Perineal hypospadias	Urogenital sinus	Female	Male	Female	Male	Male
		–	Hypospadias	–	Incurved penis	Phallus		Hypotrophic testes	Left adnexal mass	Hypotrphic testes	Hypotrophic testis
		Hypotrophic testes	Incurved short penis	Asymmetric scrotum	Asymmetric scrotum			–	Primary amenorrhea		
			Bifid scrotum					Hypospadias			
Internal genitalia	NA	NA	Epididymal structures	Right : testis	Vagina, uterus	Right : normal testis	Right : streak gonad with gonadoblastoma	NA	Small uterus	No mullerian remnants	NA
			Bilateral fallopian tubes	Left : ovarian remnant with fallopian tube structures	Right : streak gonad partially differentiated toward ovary	Left : streak gonad, fallopian tube, hemi-uterus	Left : ovary		Left adnexal mass		
			Ovotestes		Left : ovotestis, epididymal structure and fallopian tube				Right streak gonad		
Sex of rearing	Male	Male	Male	Male	Female	Female	Female	Male	Female	Male	Male
Gonadal histology	Sertoli and Leydig cells	Germinal cell aplasia	NA	NA	Right : dispersed primodial ovocytes and rare follicles	Right : small testis and normal architecture with spermatogonia in seminiferous tubules	NA	NA	Streak gonad	Eosinophilic Sertoli cells	NA
	Atrophied seminiferous tubules				Left : Testicular portion : numerous seminiferous tubules and Sertoli cells;	Left : streak gonad				Atrophy of seminiferous tubules	
	No spermatogenesis				Ovarian portion : abundant					Leydig cells hyperplasia	

(*Continued*)

TABLE I. (*Continued*)

	Cox et al. (2011)	Vetro et al. (2011)	Benko et al. (2011)					Xiao et al. (2013)	Bhagavath et al. (2014)	Our patients		
			DSD1	DSD2	DSD3 primordial follicles	DSD4 Case 1	DSD4 Case 2		Case III3	Patients 1 & 2		Patient 3
										No spermatogenesis		
Type of rearrangement	Duplication	Duplication	Duplication	Duplication	Duplication	Deletion	Deletion	Duplication	Deletion	Duplication		Duplication
Size of the rearrangement	178 kb	96 kb	693 kb	148 kb	780 kb	240 kb	240 kb	74 kb	349 kb	83,8 kb		140 kb

NA, Not available.

2011; Xiao et al., 2013]. The phenotypes associated with these rearrangements are reported in Table I. In this study, through the analyses of three XX males and comparison with the published data, we redefined the minimal critical region located ~600 kb upstream of *SOX9* associated with XX sex-reversal. The two microduplications that we describe overlap those previously reported in 46,XX testicular DSD as well as deletions in 46,XY undermasculinized DSD patients. Therefore, we can define the critical duplicated region to 69,533,305 bp (normal) and 69,534,526 bp (duplicated) to 69,575,195 bp. This corresponds to a minimum size of 40,669 bp and a maximum size of 41,890 bp. Although this considerably reduces the extent of the RevSex regulatory element, it is unclear precisely how the rearrangements of this region result in changes in the expression of *SOX9* in the early developing gonad.

The duplicated region itself contains two predicted enhancer motifs. Although interpretation of ENCODE datasets should be taken with caution since the data was generated using cell lines that are not necessarily relevant for DSD, there are potentially interesting observations. The proximal strong enhancer motif (chr17:69,544,206–69,546,005) is enriched for H3K4 methylation and H3K27 acetylation that are epigenetic marks characteristic of gene activation [Eissenberg and Shilatifard 2010; Smith and Shilatifard, 2014]. This enhancer element binds the histone acetyltransferase EP300 that regulates transcription via chromatin remodeling and it is important in the processes of cell proliferation and differentiation [Ogryzko et al., 1996]. EP300 is strongly expressed in the somatic and germ cell lineages of the XX and XY gonad during sex-determination and it can act as a co-activator of both NR5A1 and SOX9 [Ito et al., 1998; Furumatsu et al., 2009; Munger et al., 2013]. The CH3 domain of EP300 directly associates with the C-terminal PQ-rich transactivation domain of Sox9, and activates Sox9-dependent transcription in chondrogenesis by induction of histone acetylation [Furumatsu et al., 2009]. NR5A1 (also known as SF-1) was originally identified as a master-regulator of steroidogenic enzymes in the early 1990s and controls many key aspects of adrenal and reproductive functions [El-Khairi and Achermann, 2012]. Mutations involving *NR5A1* are one of the most common causes of 46,XY DSD and they are associated with a range of phenotypes including 46,XY gonadal dysgenesis and adrenal failure, 46,XY DSD with apparently normal adrenal function, 46,XY infertile male and 46,XX female with ovarian insufficiency [El-Khairi and Achermann, 2012]. It is possible that *SOX9* expression may be regulated through this element involving both NR5A1 and EP300 proteins. Furthermore, this enhancer motif is located between two predicted binding sites for DMRT1 (chr17:69,543,454–69,543,468 and chr17:69,554,542–69,554,559). DMRT1 controls many aspects of testicular development in the mouse and human, including the differentiation, proliferation, migration, and pluripotency of germ cells, and also proliferation and differentiation of Sertoli cells [Matson and Zarkower, 2012]. Haploinsufficency of *DMRT1* in the human is associated with a failure of testis-determination, while in the mouse DMRT1 is essential for maintaining mammalian sex-determination by antagonizing the ovarian gene regulatory pathway [Ottolenghi and McElreavey, 2000; Matson et al., 2011]. In avians, where *DMRT1* is the best candidate testis-determining gene, it is necessary for the expression of SOX9 [Smith et al., 2009; Lambeth et al., 2014]. ChIPseq experiments using mice using cells from P28 testis have demonstrated

that Dmrt1 binds to both upstream and downstream of *Sox9* and bind near to the genes *Wnt4*, *Foxl2*, and *Rspo1* suggesting that DMRT1 may initiate or maintain male cell fate by directly activating Sox9, whilst repressing female-promoting genes [Murphy et al., 2010; Matson et al., 2011]. The localization of DMRT1-binding sites within a putative enhancer element within the RevSex element opens the possibility that the testis formations associated with these duplications could be initiated through DMRT1 activation.

Evolutionary conserved predicted DNA-binding sites for other proteins involved in sex-determination are located in the minimal region including PBX1. PBX1, encodes a TALE (three amino acid loop extension) class homeodomain protein that participates in multimeric transcriptional complexes to modulate gene expression [Longobardi et al., 2014]. A role for PBX1 in gonad development is highlighted by *Pbx1*-deficient mice which exhibit embryonic lethality at E15 and have severe adrenal hypoplasia together with pancreatic dysfunction, skeletal abnormalities and impaired gonadal development [Schnabel et al., 2003].

Narrowing the region encompassing gonad specific regulatory elements of *SOX9* gene to an approximately 40 kb interval containing evolutionarily conserved elements is of potential interest for further studies to delineate precisely the sequences and mechanism(s) by which this region may regulate the expression of *SOX9* during human testicular development. As hypothesized by Lybaek and colleagues, the RevSex element may interact with other known regulatory elements of *SOX9* such as the TESCO motif or the minimal promoter region of *SOX9*. In a study of a familial case of XX sex-reversal carrying a 148 kb RevSex duplication, the chromatin landscape at the TESCO enhancer element was found to be modified in cells from XX masculinized individuals [Lybæk et al., 2014]. The authors speculate that the RevSex duplication was acting to induce long-range epigenetic changes including a more open chromatin landscape at the TESCO element resulting in the upregulation of SOX9 expression. The analyses of further DSD patients with duplications or deletions of this region should delimit further the minimal region involved and provide insights into the mechanism of sex-reversal. Moreover, the description of two novel duplications of this region, in two different families, suggests that it is likely to be a recurrent genetic cause of 46, XX-*SRY* negative males in humans although the exact mechanism of duplications/deletions remains to be elucidated.

ACKNOWLEDGMENTS

We thank Joelle Bignon-Topalovic for technical assistance, Support for this work was provided by the Agence Nationale de la Recherche-GIS Institut des Maladies Rares (K.M.), March of Dimes Foundation Research Grant 1-FY07-490 (K.M.), and EuroDSD in the European Community's Seventh Framework Programme FP7/2007–2013 under Grant 201444 (K.M. and A.B.).

REFERENCES

Bagheri-Fam S, Barrionuevo F, Dohrmann U, Günther T, Schüle 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 291:382–397.

Bashamboo A, McElreavey K. 2013. Gene mutations associated with anomalies of human gonad formation. Sex Dev 7:126–146.

Bell DM, Leung KK, Wheatley SC, Ng LJ, Zhou S, Ling KW, Sham MH, Koopman P, Tam PP, Cheah KS. 1997. *SOX9* directly regulates the type-II collagen gene. Nat Genet 16:174–178.

Benko S, Fantes JA, Amiel J, Kleinjan D-J, Thomas S, Ramsay J, Jamshidi N, Essafi A, Heaney S, Gordon CT, McBride D, Golzio C, Fisher M, Perry P, Abadie V, Ayuso C, Holder-Espinasse M, Kilpatrick N, Lees MM, Picard A, Temple IK, Thomas P, Vazquez M-P, Vekemans M, Crollius HR, Hastie ND, Munnich A, Etchevers HC, Pelet A, Farlie PG, FitzPatrick DR, Lyonnet S. 2009. Highly conserved non-coding elements on either side of *SOX9* associated with Pierre Robin sequence. Nat Genet 41:359–364.

Benko S, Gordon CT, 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: 825–830.

Bhagavath B, Layman LC, Ullmann R, Shen Y, Ha K, Rehman K, Looney S, McDonough PG, Kim H-G, Carr BR. 2014. Familial 46,XY sex reversal without campomelic dysplasia caused by a deletion upstream of the *SOX9* gene. Mol Cell Endocrinol 393:1–7.

Cox JJ, Willatt L, Homfray T, Woods CG. 2011. A *SOX9* duplication and familial 46, XX developmental testicular disorder. N Engl J Med 364: 91–93.

Eissenberg JC, Shilatifard A. 2010. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. Dev Biol 339:240–249.

El-Khairi R, Achermann JC. 2012. Steroidogenic factor-1 and human disease. Semin Reprod Med 30:374–381.

Foster JW, Dominguez-Steglich MA, Guioli S, Kwok C, Weller PA, Stevanović M, Weissenbach J, Mansour S, Young ID, Goodfellow PN. 1994. Campomelic dysplasia and autosomal sex reversal caused by mutations in an *SRY*-related gene. Nature 372:525–530.

Furumatsu T, Ozaki T, Asahara H. 2009. Smad3 activates the *Sox9*-dependent transcription on chromatin. Int J Biochem Cell Biol 41:1198–1204.

Georg I, Bagheri-Fam S, Knower KC, Wieacker P, Scherer G, Harley VR. 2010. Mutations of the *SRY*-responsive enhancer of *SOX9* are uncommon in XY gonadal dysgenesis. Sex Dev 4:321–325.

Gordon CT, Tan TY, Benko S, Fitzpatrick D, Lyonnet S, Farlie PG. 2009. Long-range regulation at the *SOX9* locus in development and disease. J Med Genet 46:649–656.

Huang B, Wang S, Ning Y, Lamb AN, Bartley J. 1999. Autosomal XX sex reversal caused by duplication of *SOX9*. Am J Med Genet 87:349–353.

Hughes IA, Houk C, Ahmed SF, Lee PA. 2006. Consensus statement on management of intersex disorders. J Pediatr Urol 2:148–162.

Ito M, Yu RN, Jameson JL. 1998. Steroidogenic factor-1 contains a carboxy-terminal transcriptional activation domain that interacts with steroid receptor coactivator-1. Mol Endocrinol Baltim Md 12:290–301.

Kousta E, Papathanasiou A, Skordis N. 2010. Sex determination and disorders of sex development according to the revised nomenclature and classification in 46, XX individuals. Horm Athens Greece 9: 218–131.

Kurth I, Klopocki E, Stricker S, van Oosterwijk J, Vanek S, Altmann J, Santos HG, van Harssel JJT, de Ravel T, Wilkie AOM, Gal A, Mundlos S. 2009. Duplications of noncoding elements 5' of *SOX9* are associated with brachydactyly-anonychia. Nat Genet 41:862–863.

Lambeth LS, Raymond CS, Roeszler KN, Kuroiwa A, Nakata T, Zarkower D, Smith CA. 2014. Over-expression of DMRT1 induces the male pathway in embryonic chicken gonads. Dev Biol 389:160–172.

Longobardi E, Penkov D, Mateos D, De Florian G, Torres M, Blasi F. 2014. Biochemistry of the tale transcription factors PREP, MEIS, and PBX in vertebrates. Dev Dyn Off Publ Am Assoc Anat 243:59–75.

López M, Torres L, Méndez JP, Cervantes A, Pérez-Palacios G, Erickson RP, Alfaro G, Kofman-Alfaro S. 1995. Clinical traits and molecular findings in 46, XX males. Clin Genet 48:29–34.

Lybæk H, de Bruijn D, den Engelsman-van Dijk AHA, 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 Off. J DNA Methylation Soc 9:416–427.

Mandel H, Shemer R, Borochowitz ZU, Okopnik M, Knopf C, Indelman M, Drugan A, Tiosano D, Gershoni-Baruch R, Choder M, Sprecher E. 2008. SERKAL syndrome: An autosomal-recessive disorder caused by a loss-of-function mutation in WNT4. Am J Hum Genet 82:39–47.

Matson CK, Murphy MW, Sarver AL, Griswold MD, Bardwell VJ, Zarkower D. 2011. DMRT1 prevents female reprogramming in the postnatal mammalian testis. Nature 476:101–104.

Matson CK, Zarkower D. 2012. Sex and the singular DM domain: Insights into sexual regulation, evolution and plasticity. Nat Rev Genet 13: 163–174.

Moalem S, Babul-Hirji R, Stavropolous DJ, Wherrett D, Bägli DJ, Thomas P, Chitayat D. 2012. XX male sex reversal with genital abnormalities associated with a de novo SOX3 gene duplication. Am J Med Genet A 158A:1759–1764.

Munger SC, Natarajan A, Looger LL, 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 MW, Sarver AL, Rice D, Hatzi K, Ye K, Melnick A, Heckert LL, Zarkower D, Bardwell VJ. 2010. Genome-wide analysis of DNA binding and transcriptional regulation by the mammalian Doublesex homolog DMRT1 in the juvenile testis. Proc Natl Acad Sci USA 107:13360–13365.

Murrell A. 2011. Setting up and maintaining differential insulators and boundaries for genomic imprinting. Biochem Cell Biol Biochim Biol Cell 89:469–478.

Ng LJ, Wheatley S, Muscat GE, Conway-Campbell J, Bowles J, Wright E, Bell DM, Tam PP, Cheah KS, Koopman P. 1997. *SOX9* binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. Dev Biol 183:108–121.

Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y. 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell 87:953–959.

Ottolenghi C, McElreavey K. 2000. Deletions of 9p and the quest for a conserved mechanism of sex determination. Mol Genet Metab 71: 397–404.

Parma P, Radi O, Vidal V, Chaboissier MC, 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: 1304–1309.

Polanco JC, 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 19:506–516.

Schnabel CA, Selleri L, Cleary ML. 2003. Pbx1 is essential for adrenal development and urogenital differentiation. Genes NYN 37:123–130.

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

Smith CA, Roeszler KN, Ohnesorg T, Cummins DM, Farlie PG, Doran TJ, Sinclair AH. 2009. The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature 461:267–271.

Smith E, Shilatifard A. 2014. Enhancer biology and enhanceropathies. Nat Struct Mol Biol 21:210–219.

Sutton E, Hughes J, White S, Sekido R, Tan J, Arboleda V, Rogers N, Knower K, Rowley L, Eyre H, Rizzoti K, McAninch D, Goncalves J, Slee J, Turbitt E, Bruno D, Bengtsson H, Harley V, Vilain E, Sinclair A, Lovell-Badge R, Thomas P. 2011. Identification of SOX3 as an XX male sex reversal gene in mice and humans. J Clin Invest 121: 328–341.

Velagaleti GVN, Bien-Willner GA, Northup JK, Lockhart LH, Hawkins JC, Jalal SM, Withers M, Lupski JR, Stankiewicz P. 2005. Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of *SOX9* in two patients with campomelic dysplasia. Am J Hum Genet 76:652–662.

Vetro A, Ciccone R, Giorda R, Patricelli MG, Della Mina E, Forlino A, Zuffardi O. 2011. XX males *SRY* negative: A confirmed cause of infertility. J Med Genet 48:710–712.

Wagner T, Wirth J, Meyer J, Zabel B, Held M, Zimmer J, Pasantes J, Bricarelli FD, 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 S OX9. Cell 79:1111–1120.

Xiao B, Ji X, Xing Y, Chen Y-W, Tao J. 2013. A rare case of 46, XX SRY-negative male with a ~74 kb duplication in a region upstream of S OX9. Eur J Med Genet 56:695–698.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Discussion of publication

In mice the decision of the bipotential gonad to develop towards the male determination pathway is directed by the synergistic action of Sry with Nr5a1 to upregulate *Sox9* expression via *Tesco* (the testis specific Sox9 enhancer). *Tesco* lies 13kb upstream of *Sox9* (Sekido and Lovell-Badge, 2008). However, in human cases of DSD mutations in *Tesco* have not been reported until now. A possible explanation is the existence of additional regulatory elements involved in gonad specific regulation of *SOX9*. Data obtained until now indicate that *RevSex*, a region around 600kb upstream of *SOX9*, is critical for its expression during human gonad development. We identified duplications in this region associated with 3 cases of 46,XX DSD. These duplications and previously published reports helped to redefine a minimal critical region of around 40kb associated with XX sex-reversal. How the rearrangements of this region impact the expression of *SOX9* remains to be studied. Importantly, *in silico* analysis of this region revealed the presence of two binding sites for important factors involved in early testis formation: SRY, WT1, SOX9, NR5A1, LHX9 and DMRT1. DMRT1 binds to *Sox9* regulatory sequences *in vitro* and *in vivo*, and to regulatory sequences of feminizing genes including *Foxl2* (Matson et al., 2011; Murphy et al., 2010), suggesting its involvement in the initiation and maintenance of male cell fate by activation of *Sox9* and possible suppressing *Foxl2*. The presence of DMRT1 binding sites within the duplication region in *RevSex*, raises the possibility that regulation by DMRT1 could be linked to the sex reversal phenotype through its activation.

Derivation of a steroidogenic like cell lineage from mESCs.

Introduction

In vivo model systems have been useful to study sex determination and testis development in the embryo. Moreover, the existence of transgenic models and knock out mice has considerably increased our knowledge in the subject (Vaiman, 2003). However *in vivo* models are complex to analyze and demanding to maintain. In the other hand, *in vitro* model systems have been used to study biochemical and genomic effects of hormones, however most of them are unresponsive to androgens under culture conditions or showed low levels of response (Nurmio et al., 2009; Gassei and Schlatt, 2007). Although Sertoli cell lines have been immortalized, they loss their unique characteristics during long periods of culture, they do not bind to spermatids *in vitro*, cannot maintain SSCs, they loose expression of Sertoli markers and their genomes are unstable (Wolski et al., 2007; Hofmann et al., 1992; Cory et al., 2007; Knower et al., 2007). Recently, a model system was derived by the reprograming of mouse embryonic fibroblasts to induced embryonic Sertoli like cells (ieSCs) by the induction of five transcription factors: *Sox9, Gata4, Dmrt1, Nr5a1* and *Wt1* (Buganim et al., 2012). Altough ieSCs have a similar expression pattern to fetal Sertoli cells, a normal karyotype, can aggregate, form tubule like structures and are able to maintain germ cells in culture. ieSCs lose their Sertoli cell characteristics when induced expression was removed during long periods of culture.

Given the lack of an appropriate model system to study cell fate choice during gonad determination and therefore DSD our objective is to develop a cell line that may mimic the characteristics of the somatic cells of the testis during the early development. Using murine embryonic stem cells (mESCS) we have derived a mESC stable cell line overexpressing DMRT1, which we are studying further for its ability to differentiate into the fetal Sertoli like cells.

Materials and methods

Cell culture.

E14 mESCs were seeded in gelatin coated flasks, and cultured in DMEM +10%FBS + LIF (ESG1107). Cells were split every two days to $1.2x10^6$ cells in $25cm^2$ flasks. E14 mESCs were kept in culture for at least 3 passages before electroporation.

Plasmid construction

Plasmid pCDNA6A-*DMRT1*-myc with blasticidin resistance was obtained by subcloning *DMRT1* from a T7-*DMRT1* vector (gift from Pr. David Zarkower, University of Minnesota) into pCDNA6A-myc (Life technologies V220-20) using the restriction sites *NotI* and *XbaI.*

Electroporation.

pCDNA6A-*DMRT1*-myc was linearized using *StuI* enzyme. The digestion was verified by agarose gel electrophoresis. The linearized plasmid was excised out of the gel and purified using Gel extraction kit (Thermo: K0509). E14 WT mESCs were passaged twice before electroporation; $1x10^7$ cells in electroporation buffer (BioRad 165-2677) were electroporated with 100µg of linear pCDNA6A-*DMRT1*-myc. Electroporation was performed using a Biorad Gene Pulser Xcell machine under conditions: 240V, 500µF, TC∞. Recovered cells were seeded in previously gelatinized plates and incubated 24hrs before changing to selection medium with blasticidin (Fisher Sc. BP2647-100). Cells were selected for 12 days until colonies were formed. Individual colonies were picked and transferred to 96 well plates and expanded until confluent, after which they were seeded in duplicates in 6 well plate and frozen with 10%DMSO.

Thawing cell lines.

Cells were thawed in water bath by gently shaking. Resuspended in 5mls of pre warmed DMEM and centrifuged (850rpm- 2min). Pellets were resuspended in DMEM-complete + LIF and seeded onto $25cm^2$ flasks pre-coated with 0.1% gelatin. After 24hrs medium was changed to DMEM with blasticidin.

Genetic screening of clones

Genomic DNA was extracted from the recovered ES clones using ESCs lysis buffer, as described previously (Ramírez-Solis et al., 1993).

The oligos for the screening PCR were designed from exon 2 (GGGACGTATGGTCATCCAGG) to exon 3 (TCTCCATGTTCTTCATCTGCCA) of

DMRT1 gene and PCR reactions were performed using TAQ Polymerase BIOLINE (B1021040).

PCR conditions:
95C – 1min
(95C-30 secs, 60.6C-30secs, 72C-30secs) x28
72C - 3 min
10C ∞

Protein analysis

Whole protein lysate were extracted by incubating cells with IP lysis buffer (Thermo: 87788) and protease inhibitor (Life Tech: 78429) for 15 min at 4°C, centrifuged at 11000 rpm, and the supernatant transferred to a clean eppendorf tube. Proteins were quantified with a Bradford assay and equal amounts of protein for each clone were denatured and migrated on a 10% acrylamide gel. Proteins were transferred to PVDF membrane, blocked with 5% milk in TBST, incubated over night with primary antibody against DMRT1 (ab126741) and one hour with secondary antibody (ab6721). Western blots were revealed using ECL reagent.

Immunocytochemistry.

E14-DMRT1.4 and E14DMRT1.6 cell lines were seeded in gelatin coated 8 well chamber slides at a concentration of 10^5 cells /ml. 48hrs later, cells were fixed using 4%PFA (MSDS: 15735-605), blocked for 15 min, and incubated 1hr with primary antibody (rabbit pab DMRT1: ab126741), washed 3X with PBS, incubated with secondary antibody (488 goat anti rabbit: a11034) washed 3X,

covered with an antifading reagent with DAPI (Life Tech: P36931) and sealed. The cells were visualized using a microscope Leica DM1 4000B with an objective 40X, and using the software LAS V4.3, pictures were analyzed with the same software.

Results

We have recovered more than 10 clones that were positive for transfected *DMRT1* in PCR screening (ResultsFig-3A). From these, 4 were positive for DMRT1 protein expression as confirmed by WB (ResultsFig-3B). The two cell lines with higher DMRT1 protein expression (Cell lines 4 and 6) were used to continue and confirm the stability of the transgene, we froze and thawed these clones twice and tested for the expression of the transgene after every freeze-thaw cycle. Both cells lines were also analyzed by immunocytochemistry to study the correct expression and cellular localization of exogenous *DMRT1* (ResultsFig-3C).

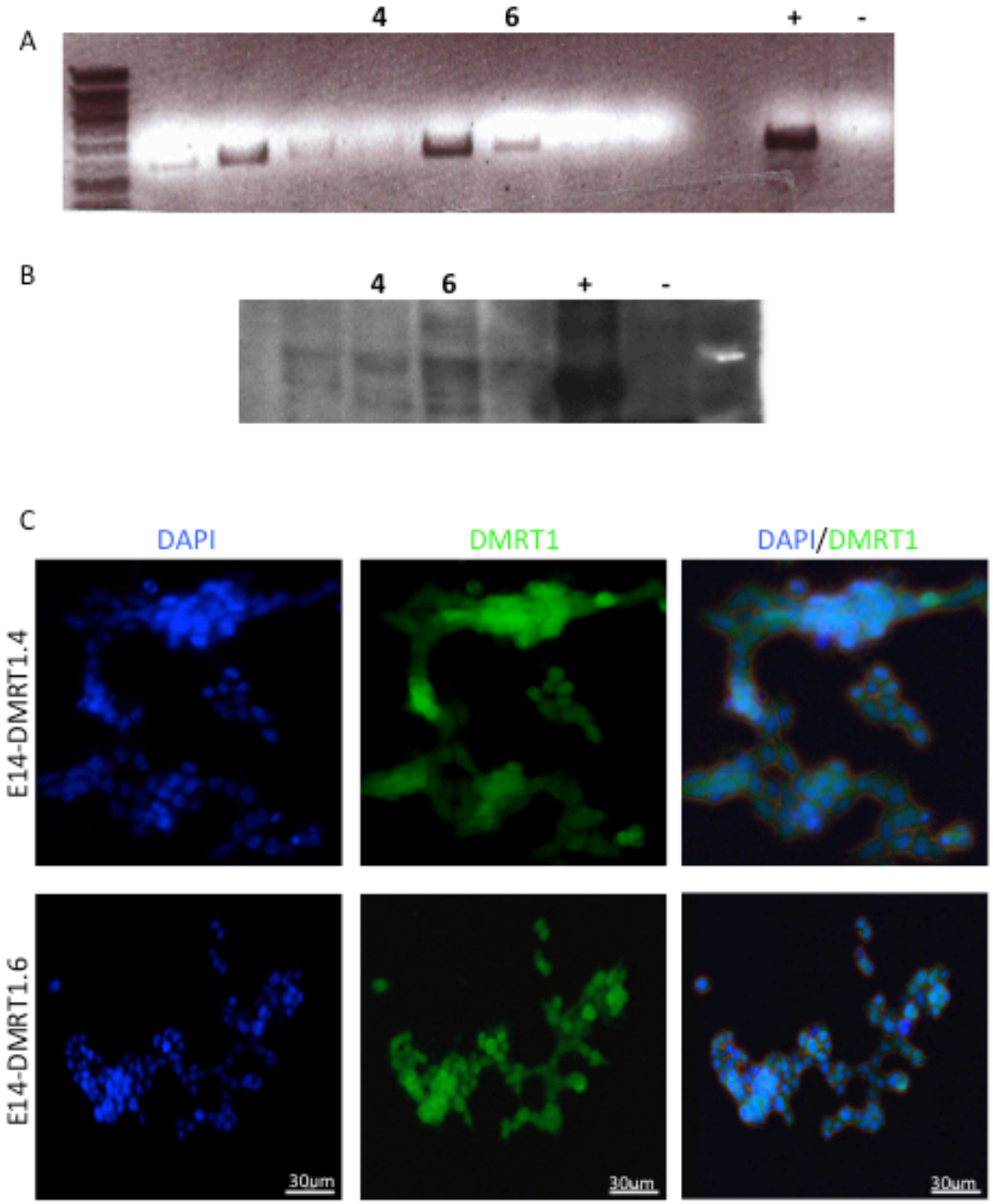


Figure Results3. <u>E14 DMRT1 screening.</u> (A) DMRT1 PCR screening reaction from genomic DNA of cell lines E14-DMRT1.4 and E14-DMRT1.6 (B) Western blot of proteins from cell lines E14-DMRT1.4 and E14-DMRT1.6 with a DMRT1 antibody (C) Cellular localization of E14-DMRT1 cell lines. E14 DMRT1.4 and E14 DMRT1.6 cell lines were fixed and stained with a DMRT1 antibody. Both cell lines overexpress DMRT1 (green) and can localize to the nucleus (DAPI/blue stained).

Conclusion and Perspectives

The E14-DMRT1 cell lines are being tested for their suitability to give rise to fetal Sertoli like cells in collaboration with Prof. Robin Lovell-Badge (Institute Crick), whose lab is developing protocols to derived SCs from murine pluripotent cells using conditioned medium with added growth factors. These protocols are being tested. Once the protocol and medium are optimized, the E14-DMRT1 cell lines will be directed to differentiate towards the fetal Sertoli like cell lineages. The resultant population will be compared for their expression pattern of fetal Sertoli cell markers by qPCR as compared to cells from the embryonic gonad of the *Nr5a1*-GFP transgenic mouse. The cell lines will also be assessed for their ability to produce AMH. Once confirmed we will immortalize the appropriate cell lines to give us a model system to use for studying the effect of the mutations causing errors of testis determination.

General discussion and Conclusions

Human sex determination is a complex biological process involving a number of genetic factors and regulatory mechanisms. It is a classical paradigm for cell fate choice where depending on the molecular signals, two mutually antagonistic pathways suppress the choice of the alternate fate from the bipotential gonadal primordium. In spite of recent advances in our knowledge of genes involved in the process, surprisingly little is known about the mechanism of vertebrate sex-determination.

In evolution most of the developmental processes are conserved both at genetic and molecular level, however contrary to this, sex determination is not. Some of the genes may show an evolutionary conservation but the molecular mechanism by which they regulate or participate in the process of sex determination is variable amongst different phyla. One such gene family is DM domain DNA binding motif transcriptional regulators. DM domain genes arose early in metazoan evolution probably to regulate sex differentiation and have been conserved during hundreds of millions of years. *Dmrt1* regulates sex development in some vertebrates and was identified by comparison of *D. melanogaster dsx* and *mab-3* in *C. elegans*. DM genes function in the gonads and promote male somatic gonad development in at least some vertebrates (Matson 2012). In vertebrates and some invertebrates, DM domain genes act in the somatic gonadal primordium to control the male specific differentiation. However, in others such as *Drosophila, dsx* controls the survival of mesodermal

cells that are later recruited to the male gonadal primordium. This suggests that even when sex developmental mechanisms differ, the key regulators may be conserved.

In mouse *Dmrt1* is not essential for primary sex determination, *Dmrt1* null XY mice are born with testis, however they undergo testicular dysgenesis 2 weeks after birth (Raymond et al., 2000). The ablation of *Dmrt1* in mice adult gonads results in the transdifferentiation of Sertoli cells into Granulosa cells with *Foxl2* expression (Matson et al., 2011). Concordantly, forced expression of *Dmrt1* in XX mice fetal gonads results in repression of *Wnt4* and *Foxl2* expression, the activation of *Sox9* expression and consequent postnatal development of male internal and external genitalia (Zhao et al., 2015). This suggests that even when *Dmrt1* is dispensable for primary sex determination in mice, being functionally replaced by *Sry*, it can still act as a primary testis-determining gene when expressed at right moment.

In different non-mammalian vertebrates DM genes are located on the sex chromosome and act as the regulator for sex determination, such as medaka fish *Oryzias latipes* where a duplication of the autosomal *Dmrt1a* lead to creation of a new Y chromosome on which *Dmy* act as the male sex determining gene (Matsuda et al., 2002; Shinomiya et al., 2010), in the amphibian *Xenopus laevis* where females have a duplicated copy of *Dmrt1* (dm-w) on their W chromosome that block autosomal *Dmrt1* to specify testis (Yoshimoto et al., 2010). In chicken, where females are (ZW) and males are homogametic (ZZ) *Dmrt1* gene is Z-chromosome linked, thus males have two copies and present a higher expression

of *Dmrt1* in the genital ridge, which is required for testis fate; where as females have just one copy (Lambeth et al., 2014; Smith et al., 2009). All evidence suggest that *Dmrt1* had an ancient role in primary sex determination and development in vertebrates and that in placental mammals the function of *Drmt1* in primary sex determination was recently replaced by *Sry* and *Sox9* (Matson and Zarkower, 2012). One possibility is that since *Dmrt1* participates in different biological functions (sex determination and sex maintenance), its pleiotropic nature could have given rise to gene duplication events, which may have lead to (i) divergence of DM genes and functions and (ii) other genes such as *Sry* to replacing its function.

We identified and characterized a *de novo* point mutation in *DMRT1*, which, to our knowledge, is the only human mutation known that abolishes the specificity and changes the mechanism of the DNA-protein interaction by the DMRT1 protein. This mutation interferes with DNA-binding by altering the stoichiometry of DMRT1, resulting in the misregulation of downstream targets of DMRT1. For example, the DM binding sites in *Sox9* sequence are recognized and bound by WT-DMRT1 as a trimer, whereas mutant DMRT1 weekly recognizes these sites. Moreover, when mixed both WT and mutant DMRT1 proteins, the recognition has higher affinity but the stoichiometry is shifted from trimer to a tetramer complex. This mutation indicates, in contrast to the mouse, a role for *DMRT1* in human primary sex determination. The apparent lack of a requirement for *Dmrt1* in mouse sex-determination may be the exception rather then the rule in mammalian sex-determination although this needs to be experimentally evaluated. For example, in the goat DMRT1 may control *SOX9* expression. In XX

goats lacking *FOXL2*, the *DMRT1* gene is expressed before *SOX9*. (Elzaiat et al., 2014). This suggests, albeit indirectly, that *DMRT1* could be sex-determining in the goat. Although a formal proof of this is required for this.

As well as SRY, other members of the SOX transcription factor family play a major role in multiple developmental processes including early embryogenesis, gastrulation, organogenesis and during specification and differentiation of various cell types. Sox genes arose in evolution before multicellularity since *Monosiga brevicollisa* an unicellular choanoflagellate, has two Sox-like genes with HMG domains exhibiting 49-50% amino acid identity to the HMG domain of mouse Sry and human SRY (Guth and Wegner, 2008). As *Sox* genes are randomly distributed and there is no a cluster of *Sox* genes, it seems that divergence of *Sox* seems to have arise from tandem duplication events (Bowles et al., 2000). The generation of new copies of *Sox* genes has caused the partition of conserved function by several Sox proteins. Genes in the same subgroup, besides having a high similarity at DNA sequence and protein level, often have overlapping if not identical expression patterns. This is observed in SoxE subgroup where *Sox8, Sox9* and *Sox10* are expressed in somatic cells of the developing gonads of the vertebrates in a male-specific sexually dimorphic pattern. *Sox9* initiates Sertoli cell differentiation, testis cords formation, and overall testis development. However, *Sox8* and *Sox10* are also expressed in the developing male gonad (Barrionuevo and Scherer, 2010). *Sox8* reinforces *Sox9* expression during testis differentiation and has also a role during adult Sertoli cell maintenance and *Sox10* overexpression is implicated in XX sex reversal not only in mice but in human 46,XX DSD with duplications of 22q13.1 that include *SOX10* (Polanco et

al., 2010). Moreover, in XX mice lacking *Sox9* and ovary determining genes such as *Rspo1* or b-catenin, *Sox8* expression seems to be enough to induce testis differentiation (Lavery et al., 2012; Nicol and Yao, 2015; Barrionuevo et al., 2009). This suggests that the functional redundancy between SoxE members during testis development is a biological strategy to reinforce male sex developmental pathway in the developing gonad.

During the course of this work we identified a number of mutations in *SOX8* in individuals with unexplained cases of DSD and both male and female infertility. Of special interest was a point mutation in *SOX8* in a patient presenting with 46,XY CGD. The heterozygous mutation c.468G>C changed a highly conserved residue p.E156D in the HMG domain of SOX8. Functional characterization of this mutation, showed that both wild-type and SOX8 p.Glu156Asp have the ability to transactivate the *Sox9* Tesco enhancer element, however, the SOX8p.Glu156Asp specifically failed to synergize with NR5A1 to transactivate the *Tesco* enhancer even though SOX8 and NR5A1 proteins can physically bind to each other. Moreover, SOX8p.E156D exerts a dominant negative effect by preventing synergistic activation of *Tesco* enhancer by NR5A1 and wild-type SOX8 as well as by NR5A1 and SOX9. Our results suggest that *SOX8* has not only a similar expression pattern to *SOX9* but also an overlapping function during human gonadal development, revealing the molecular and genetic mechanisms played by SOXE members during this process, and moreover associates pathogenic *SOX8* mutations to human DSD.

Using WES we have identified two new point mutations in *SOX7* associated with two individuals presenting with 46,XY DSD. SOX7 has an important role during different developmental processes. During vascular development, *Sox7* has overlapping expression with vascular markers such as *Flk-1* (the major receptor of vascular endothelial growth factor VEGF) in the endothelial precursors. In this cell population the expression of cardiac transcription factors such as *Gata4* emerge. Importantly, differences in *Sox7* regulation affect the cell fate decision between cardiac and vascular pathways by controlling Wnt/ β -catenin signaling pathway (Nelson et al., 2009), where SOX7 has an inhibitory effect of β -catenin through its c-TAD domain and thereby negatively regulates Wnt signaling pathway (Guo et al., 2008). In male gonads, testis genes act to repress their feminizing counterparts; β -catenin, *Gata4* and *Gata6* play important sexually dimorphic roles during gonad development; it is possible that SOX7 may play a regulating role of these genes in male gonads and that hemizygocity of *SOX7* caused by p.D211A or p.Y381H might result in the misregulation of its gene targets and thus DSD. Our functional assays demonstrated that the heterozygous mutation SOX7p.D211A abolishes the synergistic activation of *Sox9* expression via TESCO by SOX7-NR5A1 complex. This suggests that mutant SOX7 could abolish the correct binding of other SOX proteins which are normally activating *Tesco* at important stages of sex determination or development, for example when *SOX9* is being highly induced or to maintain its level of expression in Sertoli cells.

In mouse expression data available shows that *Sox7* is expressed only in the endothelial cells of the developing gonad and it expression can be detected on RNA seq data from E13 in the testis. However, in humans there is no current evidence about *SOX7* expression in the developing gonads, but we know that it is expressed in adult Sertoli cells and Leydig cells. it could be possible that *SOX7* is expressed in human embryonic Sertoli cells or Leydig cells, if so, our functional assays showing absence of interaction with NR5A1 at Tesco, would suggest a mechanism by which the mutations in *SOX7* could be linked to DSD. However, if *SOX7* is expressed only in the endothelial cells in the male developing gonads, it could be possible that mutant SOX7 misregulates the downstream targets for vascularization in the early testicular development and may perturb the testicular architecture and this could result in dysgenic testis.

As SOX7 is known to directly regulate *GATA* factors expression during differentiation of parietal endoderm (Murakami et al., 2004; Costa et al., 2012), and GATA factors are key genes in testis development, it is also possible that SOX7 mutations impair GATA factors activation in the testis. Experiments to verify this using transient gene activation assays would be useful to answer this question. Moreover, further analysis of the effect of SOX7p.Y381H on the biological activity of SOX7, such as the activation of different genes and regulatory elements important for sex determination such as TESCO, *AMH*, and *NR5A1* are needed and currently undergoing. Current evidence suggests that *SOX7* could be involved in maintaining the integrity of the male gonads. To have a better understanding of mechanism and role of SOX7 in developing human

gonads we need to understand the correct spatial and temporal expression pattern of SOX7.

GATA4 is well known to be involved in murine sex-determination (Viger et al., 2008). We identified two heterozygous mutations in *GATA4* associated to 46,XY DSD. The first mutation, GATA4p.D425N, has already been reported in cases of TOF and CHD without any reported gonadal phenotype (Tomita-Mitchell et al., 2007; Zhang et al., 2009; Granados-Riveron et al., 2012). Functional analyses of this mutation show a reduction in the stability of the mutant protein. Reduced stability of mutated GATA4 protein may explain the gonadal or heart defects seen in association with the mutation due to the gene dosage effects. It has been shown that the testis determination, at least in mouse, is very sensitive to correct doses of GATA4 as shown by *siGata4* transgenic mice where efficient silencing of *Gata4*, does not show an effect in heart function or development even at very low levels of Gata4 (20-30% of wt), but severely effects the male gonad development. More importantly in male gonads this reduction results in a significant reduction of Gata4 downstream target such as *Amh* and *StAR*. In contrast to the male, expression of *Gata4* downstream target in the female gonads was just slightly altered (Thurisch et al., 2009), suggesting that Gata4 acts in a dose dependent manner during development and testis development is very sensitive to small changes in gene doses. Similarly to this it is possible that the reduced stability of the GATA4p.D425N creates haploinsufficiency leading to reduction in the expression of targets of GATA4, which are needed for testis determination, and thus results in DSD. In the literature, several mutations have been reported in patients with cardiac defects and no gonadal phenotype (Tomita-Mitchell et al.,

2007; Zhang et al., 2009; Granados-Riveron et al., 2012). This could be due to two reasons (i) the gonadal development or function in the patients was not evaluated (ii) these mutations may specifically impair the interaction with genes/proteins required for cardiac development or function but not that of the gonad. Previously our lab has also identified a familial case with 46,XY DSD and 46,XX with normal ovarian function carrying the mutation in *GATA4*, where not all the 46,XY DSD individuals had heart anomalies (Lourenço et al., 2011). In this family the ovaries in 46,XX females did not appear to be affected by the mutation in *GATA4*. The mutation identified by me, GATA4p.D425N, has been previously reported in cases of CHD, where the sex of the individuals carrying the mutation was not reported (Tomita-Mitchell et al., 2007; Zhang et al., 2009). In Granados-Riveron et al., 2012, the fertile mother (presenting Atrial Septal Defect ASD) passed the mutation to her two sons who were completely healthy suggesting an incomplete penetrance of this mutation.

Preliminary analysis of a second mutation GATA4p.S224P showed that the mutant protein showed a reduction in the activation of the *AMH* reporter. More work is needed to determine the exact mechanism and downstream targets (such as *DMRT1* and *NR5A1*) that GATA4p.S224P may affect during sex determination. Another possibility is that GATA4 protein interaction with its partners NR5A1 or FOG2 could be altered by p.S224P mutation and may lead to DSDS. Experiments are currently undergoing to verify this.

We are also aware of the existence of multiple *GATA4* mutations associated with 46,XY DSD in different laboratories of collaborators. It is possible that *GATA4* mutations have a higher contribution to 46,XY DSD than what we currently think. The incidence of *GATA4* mutations linked to 46,XY DSD could be established by

systematic analysis and this could lead to better molecular tools for the identification of causes of DSD.

SOX9 is key during male sex determination in mammals. In mice, *Sox9* expression in male gonads is regulated via a testis specific enhancer of *SOX9*: *Tesco*. Although there is a region in human genome, homologous to mouse *Tesco*, we still do not understand the precise mechanism how *SOX9* is regulated in the human testis. To date there are no published reports of DSD associated with mutations or rearrangements in the *Tesco* element. Furthermore, *RevSex* a regulatory element of *SOX9* expression in humans ~600kb upstream of *SOX9* has been associated with different cases of DSD (Benko et al., 2011; Vetro et al., 2011; Xiao et al., 2013; Kim et al., 2015). Using aCGH technology, we identified three cases of 46,XX DSD with duplications that overlap *RevSex* regulatory element of *SOX9*, which in addition to previous reports allowed to delimit the region associated with 46,XX DSD to approximately 40kb lying ~600kb upstream of *SOX9* initiation site. Interestingly, the duplicated region contains enhancer elements and multiple conserved binding sites for key factors involved in sex determination and development including DMRT1. It would be interesting to determine if DMRT1 has a regulatory role for *SOX9* expression during sex development or maintenance through this regulatory region *in vitro* by mutation of the minimal enhancer region and co-transfection analysis with *DMRT1.*

Mammalian gonad development and maintenance is a very complex process regulated by different genetic and molecular events, where a fine balance between mutually antagonistic pathways is needed to determine, develop and

maintain testis or ovaries. When errors in the molecular network alter the balance this results in DSD. The detailed genetic analyses of human cases of errors in sex-determination have been a powerful tool in the identification of genes involved in sex-determination. *SRY* was mapped and eventually cloned through the analyses of XY females with gonadal dysgenesis and XX males carrying a portion of the Y chromosome on one of their X chromosomes. *WT1* and *SOX9* were identified through the analyses of human chromosomal rearrangements. Furthermore, others and we have demonstrated a contribution of *NR5A1, GATA4* and *FOG2* in human DSD. The use of -omics technologies such as exome sequencing has substantially increased our knowledge of the genetic mutations that can result in an error of sex-determination. The present work has identified and analyzed mutations in the *DMRT1, SOX7, SOX8 and GATA4* genes, and the regulatory *RevSex* region of *SOX9* and how they impact on DSD.

Although our knowledge has considerably increased last years to understand early genetic, cellular and morphological events during gonadal development, the molecular mechanisms involved in human sex determination remain poorly understood. Different reasons explain this: The existent cellular models we use for studying sex determination are not accurate and informative enough. The established cell lines until now do not have precise characteristics of Sertoli cells. Moreover, the established cell lines and the primary immature and adult Sertoli cells lose their "Sertoliness" during long periods of culture (Gassei and Schlatt, 2007). In addition, the identification of new genes involved in sex determination by classical genetic studies of familial cases presenting errors in sex determination is affected because these cases are infrequent (Ohnesorg et al.,

2014). Moreover, sex determination is not a conserved process during evolution (Bachtrog et al., 2014).

We are developing protocols for derivation of the steroidogenic lineages of the testis using mouse embryonic stem (mESs) cells in collaboration with Pr. Robin Lovell-Badge from Institute Crick.

Once established these protocols would be exploited for development of novel cellular models based on biomaterial from patients with DSD and their unaffected family members. Induced pluripotent stem cells (iPSCs) provide a practical system for directed differentiation of individual specific cells into the cell-lineages of choice. Furthermore, transdifferentiation approaches may offer an alternative to iPSCs as they permit the conversion of committed cells into the lineage of choice, without the added manipulations required to revert them to the ground state. From a clinical perspective this will lead, in the short term, to an improved diagnostic approach to DSD and, in the longer term, it could result in more appropriate knowledge-based therapies. From a developmental biology perspective the data generated by these approaches should improve our understanding of the mechanism of gonad formation and the mechanism underlying the choice of somatic cell sex in humans.

Bibliography

Achermann, JC, Ito, M, Hindmarsh, PC, Jameson, JL. 1999. A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal and adrenal failure in humans. Nat. Genet. 22: 125–6.

Achermann, JC, Ozisik, G, Ito, M, Orun, UA, Harmanci, K, Gurakan, B, Jameson, JL. 2002. Gonadal determination and adrenal development are regulated by the orphan nuclear receptor steroidogenic factor-1, in a dose-dependent manner. J. Clin. Endocrinol. Metab. 87: 1829–33.

Agha, Z, Iqbal, Z, Azam, M, Ayub, H, Vissers, LELM, Gilissen, C, Ali, SHB, Riaz, M, Veltman, JA, Pfundt, R, van Bokhoven, H, Qamar, R. 2014. Exome sequencing identifies three novel candidate genes implicated in intellectual disability. PLoS One 9: e112687.

Allali, S, Muller, J-B, Brauner, R, Lourenço, D, Boudjenah, R, Karageorgou, V, Trivin, C, Lottmann, H, Lortat-Jacob, S, Nihoul-Fékété, C, De Dreuzy, O, McElreavey, K, Bashamboo, A. 2011. Mutation analysis of NR5A1 encoding steroidogenic factor 1 in 77 patients with 46, XY disorders of sex development (DSD) including hypospadias. PLoS One 6: e24117.

Anderson, EL, Baltus, AE, Roepers-Gajadien, HL, Hassold, TJ, de Rooij, DG, van Pelt, AMM, Page, DC. 2008. Stra8 and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. Proc. Natl. Acad. Sci. U. S. A. 105: 14976–80.

Anttonen, M, Ketola, I, Parviainen, H, Pusa, A-K, Heikinheimo, M. 2003. FOG-2 and GATA-4 Are coexpressed in the mouse ovary and can modulate mullerian-inhibiting substance expression. Biol. Reprod. 68: 1333–40.

Arango, NA, Lovell-Badge, R, Behringer, RR. 1999. Targeted mutagenesis of the endogenous mouse Mis gene promoter: in vivo definition of genetic pathways of vertebrate sexual development. Cell 99: 409–19.

Arboleda, VA, Lee, H, Sánchez, FJ, Délot, EC, Sandberg, DE, Grody, WW, Nelson, SF, Vilain, E. 2013. Targeted massively parallel sequencing provides comprehensive genetic diagnosis for patients with disorders of sex development. Clin. Genet. 83: 35–43.

Ascoli, M. 1981. Characterization of several clonal lines of cultured Leydig tumor cells: gonadotropin receptors and steroidogenic responses. Endocrinology 108: 88–95.

Bachtrog, D, Mank, JE, Peichel, CL, Kirkpatrick, M, Otto, SP, Ashman, T, Hahn, MW, Kitano, J, Mayrose, I, Ming, R, Perrin, N. 2014. Sex Determination : Why So Many Ways of Doing It ? 12.

Bagheri-Fam, S, Argentaro, A, Svingen, T, Combes, AN, Sinclair, AH, Koopman, P, Harley, VR. 2011. Defective survival of proliferating Sertoli cells and androgen receptor function in a mouse model of the ATR-X syndrome. Hum. Mol. Genet. 20: 2213–24.

Barber, JCK, Rosenfeld, JA, Foulds, N, Laird, S, Bateman, MS, Thomas, NS, Baker, S, Maloney, VK, Anilkumar, A, Smith, WE, Banks, V, Ellingwood, S, Kharbutli, Y, Mehta, L, Eddleman, KA, Marble, M, Zambrano, R, Crolla, JA, Lamb, AN. 2013. 8p23.1 duplication syndrome; common, confirmed, and novel features in six further patients. Am. J. Med. Genet. A 161A: 487–500.

Bardoni, B, Zanaria, E, Guioli, S, Floridia, G, Worley, KC, Tonini, G, Ferrante, E, Chiumello, G, McCabe, ER, Fraccaro, M. 1994. A dosage sensitive locus at chromosome Xp21 is involved in male to female sex reversal. Nat. Genet. 7: 497–501.

Barrionuevo, F, Georg, I, Scherthan, H, Lécureuil, C, Guillou, F, Wegner, M, Scherer, G. 2009. Testis cord differentiation after the sex determination stage is independent of Sox9 but fails in the combined absence of Sox9 and Sox8. Dev. Biol. 327: 301–12.

Barrionuevo, F, Scherer, G. 2010. SOX E genes: SOX9 and SOX8 in mammalian testis development. Int. J. Biochem. Cell Biol. 42: 433–6.

Barrionuevo, F. 2005. Homozygous Inactivation of Sox9 Causes Complete XY Sex

Reversal in Mice. Biol. Reprod. 74: 195–201.

Barsoum, IB, Yao, HH-C. 2006. Fetal Leydig cells: progenitor cell maintenance and differentiation. J. Androl. 31: 11–15.

Bashamboo, A, Brauner, R, Bignon-Topalovic, J, Lortat-Jacob, S, Karageorgou, V, Lourenco, D, Guffanti, A, McElreavey, K. 2014. Mutations in the FOG2/ZFPM2 gene are associated with anomalies of human testis determination. Hum. Mol. Genet. 23: 3657–65.

Bashamboo, A, Ferraz-de-Souza, B, Lourenço, D, Lin, L, Sebire, NJ, Montjean, D, Bignon-Topalovic, J, Mandelbaum, J, Siffroi, J-P, Christin-Maitre, S, Radhakrishna, U, Rouba, H, Ravel, C, Seeler, J, Achermann, JC, McElreavey, K. 2010a. Human male infertility associated with mutations in NR5A1 encoding steroidogenic factor 1. Am. J. Hum. Genet. 87: 505–12.

Bashamboo, A, Ledig, S, Wieacker, P, Achermann, JC, Achermann, J, McElreavey, K. 2010b. New technologies for the identification of novel genetic markers of disorders of sex development (DSD). Sex Dev. 4: 213–24.

Baxter, RM, Vilain, E. 2013. Translational Genetics for Diagnosis of Human Disorders of Sex Development.

Behringer, RR, Finegold, MJ, Cate, RL. 1994. Müllerian-inhibiting substance function during mammalian sexual development. Cell 79: 415–25.

Belville, C, Josso, N, Picard, JY. 1999. Persistence of Müllerian derivatives in males. Am. J. Med. Genet. 89: 218–23.

Benko, S, Gordon, CT, 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.

Bennett, J, Wu, Y-G, Gossen, J, Zhou, P, Stocco, C. 2012. Loss of GATA-6 and GATA-4 in granulosa cells blocks folliculogenesis, ovulation, and follicle stimulating hormone receptor expression leading to female infertility. Endocrinology 153: 2474–85.

Bergeron, F, Nadeau, G, Viger, RS. 2015. GATA4 knockdown in MA-10 Leydig cells identifies multiple target genes in the steroidogenic pathway. Reproduction 149: 245–57.

Bérubé, NG. 2011. ATRX in chromatin assembly and genome architecture during development and disease. Biochem. Cell Biol. 89: 435–44.

Beverdam, A, Wilhelm, D, Koopman, P. 2003. Molecular characterization of three gonad cell lines. Cytogenet. Genome Res. 101: 242–9.

Bhandari, RK, Haque, M, Skinner, MK. 2012. Global Genome Analysis of the Downstream Binding Targets of Testis Determining Factor SRY and SOX9. 7: 1–17.

Biason-Lauber, A, De Filippo, G, Konrad, D, Scarano, G, Nazzaro, A, Schoenle, EJ. 2006. WNT4 deficiency--a clinical phenotype distinct from the classic Mayer-Rokitansky-Kuster-Hauser syndrome: A Case Report. Hum. Reprod. 22: 224–229.

Biason-Lauber, A, Konrad, D, Meyer, M, DeBeaufort, C, Schoenle, EJ. 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.

Biason-Lauber, A, Konrad, D, Navratil, F, Schoenle, EJ. 2004. A WNT4 mutation associated with Müllerian-duct regression and virilization in a 46,XX woman. N. Engl. J. Med. 351: 792–8.

Birk, OS, Casiano, DE, Wassif, CA, Cogliati, T, Zhao, L, Zhao, Y, Grinberg, A, Huang, S, Kreidberg, JA, Parker, KL, Porter, FD, Westphal, H. 2000. The LIM homeobox gene Lhx9 is essential for mouse gonad formation. Nature 403: 909–13.

Bishop, CE, Whitworth, DJ, Qin, Y, Agoulnik, AI, Agoulnik, IU, Harrison, WR, Behringer, RR, Overbeek, PA. 2000. A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. Nat. Genet. 26: 490–4.

Bitgood, MJ, Shen, L, McMahon, AP. 1996. Sertoli cell signaling by Desert hedgehog

regulates the male germline. Curr. Biol. 6: 298–304.

Bogani, D, Siggers, P, Brixey, R, Warr, N, Beddow, S, Edwards, J, Williams, D, Wilhelm, D, Koopman, P, Flavell, RA, 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.

Bor, Y, Swartz, J, Morrison, A, Rekosh, D, Ladomery, M, Hammarskjöld, M-L. 2006. The Wilms' tumor 1 (WT1) gene (+KTS isoform) functions with a CTE to enhance translation from an unspliced RNA with a retained intron. Genes Dev. 20: 1597–608.

Boulanger, L, Pannetier, M, Gall, L, Allais-Bonnet, A, Elzaiat, M, Le Bourhis, D, Daniel, N, Richard, C, Cotinot, C, Ghyselinck, NB, Pailhoux, E. 2014. FOXL2 is a female sex-determining gene in the goat. Curr. Biol. 24: 404–8.

Bouma, GJ, Washburn, LL, Albrecht, KH, Eicher, EM. 2007. Correct dosage of Fog2 and Gata4 transcription factors is critical for fetal testis development in mice. Proc. Natl. Acad. Sci. U. S. A. 104: 14994–14999.

Bowles, J, Schepers, G, Koopman, P. 2000. Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. Dev. Biol. 227: 239–55.

Bradford, ST, Hiramatsu, R, Maddugoda, MP, Bernard, P, Chaboissier, M-C, Sinclair, A, Schedl, A, Harley, V, Kanai, Y, Koopman, P, Wilhelm, D. 2009a. The cerebellin 4 precursor gene is a direct target of SRY and SOX9 in mice. Biol. Reprod. 80: 1178–88.

Bradford, ST, Wilhelm, D, Bandiera, R, Vidal, V, Schedl, A, Koopman, P. 2009b. A cell-autonomous role for WT1 in regulating Sry in vivo. Hum. Mol. Genet. 18: 3429–38.

Braman, J, Papworth, C, Greener, A. 1996. Site-directed mutagenesis using double-stranded plasmid DNA templates. Methods Mol. Biol. 57: 31–44.

Brennan, J, Capel, B. 2004. One tissue, two fates: molecular genetic events that underlie testis versus ovary development. Nat. Rev. Genet. 5: 509–21.

Brennan, J, Tilmann, C, Capel, B. 2003. Pdgfr-alpha mediates testis cord organization and fetal Leydig cell development in the XY gonad. Genes Dev. 17: 800–10.

Britt, KL, Drummond, AE, Dyson, M, Wreford, NG, Jones, ME, Simpson, ER, Findlay, JK. 2001. The ovarian phenotype of the aromatase knockout (ArKO) mouse. J. Steroid Biochem. Mol. Biol. 79: 181–5.

Buganim, Y, Itskovich, E, Hu, YC, Cheng, AW, Ganz, K, Sarkar, S, Fu, D, Welstead, GG, Page, DC, Jaenisch, R. 2012. Direct reprogramming of fibroblasts into embryonic sertoli-like cells by defined factors. Cell Stem Cell 11: 373–386.

Bullejos, M, Koopman, P. 2001. Spatially dynamic expression of Sry in mouse genital ridges. Dev. Dyn. 221: 201–5.

Call, KM, Glaser, T, Ito, CY, Buckler, AJ, Pelletier, J, Haber, DA, Rose, EA, Kral, A, Yeger, H, Lewis, WH. 1990. Isolation and characterization of a zinc finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. Cell 60: 509–20.

Callier, P, Calvel, P, Matevossian, A, Makrythanasis, P, Bernard, P, Kurosaka, H, Vannier, A, Thauvin-Robinet, C, Borel, C, Mazaud-Guittot, S, Rolland, A, Desdoits-Lethimonier, C, Guipponi, M, Zimmermann, C, Stévant, I, Kuhne, F, Conne, B, Santoni, F, Lambert, S, Huet, F, Mugneret, F, Jaruzelska, J, Faivre, L, Wilhelm, D, Jégou, B, Trainor, PA, Resh, MD, Antonarakis, SE, Nef, S. 2014. Loss of function mutation in the palmitoyl-transferase HHAT leads to syndromic 46,XY disorder of sex development by impeding Hedgehog protein palmitoylation and signaling. PLoS Genet. 10: e1004340.

Canto, P, Söderlund, D, Reyes, E, Méndez, JP. 2004. Mutations in the desert hedgehog (DHH) gene in patients with 46,XY complete pure gonadal dysgenesis. J. Clin. Endocrinol. Metab. 89: 4480–3.

Canto, P, Vilchis, F, Söderlund, D, Reyes, E, Méndez, JP. 2005. A heterozygous mutation in the desert hedgehog gene in patients with mixed gonadal dysgenesis. Mol. Hum. Reprod.

11: 833–6.

Cantor, AB, Orkin, SH. 2005. Coregulation of GATA factors by the Friend of GATA (FOG) family of multitype zinc finger proteins. Semin. Cell Dev. Biol. 16: 117–28.

Capel, B, Albrecht, KH, Washburn, LL, Eicher, EM. 1999. Migration of mesonephric cells into the mammalian gonad depends on Sry. Mech. Dev. 84: 127–31.

Capel, B, Hawkins, JR, Hirst, E, Kioussis, D, Lovell-Badge, R. 1996. Establishment and characterization of conditionally immortalized cells from the mouse urogenital ridge. J. Cell Sci. 109 (Pt 5: 899–909.

Carlsson, P, Mahlapuu, M. 2002. Forkhead transcription factors: key players in development and metabolism. Dev. Biol. 250: 1–23.

Castro, JJ, Méndez, JP, Coral-Vázquez, RM, Soriano-Ursúa, MA, Damian-Matsumura, P, Benítez-Granados, J, Rosas-Vargas, H, Canto, P. 2013. In vitro and molecular modeling analysis of two mutant desert hedgehog proteins associated with 46,XY gonadal dysgenesis. DNA Cell Biol. 32: 524–30.

Chaboissier, M-C, Kobayashi, A, Vidal, VIP, Lützkendorf, S, van de Kant, HJG, Wegner, M, de Rooij, DG, Behringer, RR, Schedl, A. 2004. Functional analysis of Sox8 and Sox9 during sex determination in the mouse. Development 131: 1891–1901.

Chassot, A-A, Ranc, F, Gregoire, EP, Roepers-Gajadien, HL, Taketo, MM, Camerino, G, de Rooij, DG, Schedl, A, Chaboissier, M-C. 2008a. Activation of beta-catenin signaling by Rspo1 controls differentiation of the mammalian ovary. Hum. Mol. Genet. 17: 1264–77.

Chassot, AA, Gregoire, EP, Magliano, M, Lavery, R, Chaboissier, MC. 2008b. Genetics of ovarian differentiation: Rspo1, a major player. Sex Dev. 2: 219–27.

Chen, H, Ge, R-S, Zirkin, BR. 2009. Leydig cells: From stem cells to aging. Mol. Cell. Endocrinol. 306: 9–16.

Cheng, HH, Ying, M, Tian, YH, Guo, Y, McElreavey, K, Zhou, RJ. 2006. Transcriptional diversity of DMRT1 (dsx- and mab3-related transcription factor 1) in human testis. Cell Res. 16: 389–393.

Chew, L-J, Gallo, V. 2009. The Yin and Yang of Sox proteins: Activation and repression in development and disease. J. Neurosci. Res. 87: 3277–87.

Chiang, H-S, Wu, Y-N, Wu, C-C, Hwang, J-L. 2013. Cytogenic and molecular analyses of 46,XX male syndrome with clinical comparison to other groups with testicular azoospermia of genetic origin. J. Formos. Med. Assoc. 112: 72–8.

Chien, AJ, Conrad, WH, Moon, RT. 2009. A Wnt survival guide: from flies to human disease. J. Invest. Dermatol. 129: 1614–27.

Chlon, TM, Crispino, JD. 2012. Combinatorial regulation of tissue specification by GATA and FOG factors. Development 139: 3905–16.

Chung, MIS, Ma, ACH, Fung, T-K, Leung, AYH. 2011. Characterization of Sry-related HMG box group F genes in zebrafish hematopoiesis. Exp. Hematol. 39: 986–998.e5.

Clark, AM, Garland, KK, Russell, LD. 2000. Desert hedgehog (Dhh) gene is required in the mouse testis for formation of adult-type Leydig cells and normal development of peritubular cells and seminiferous tubules. Biol. Reprod. 63: 1825–38.

Clarkson, PA, Davies, HR, Williams, DM, Chaudhary, R, Hughes, IA, Patterson, MN. 1993. Mutational screening of the Wilms's tumour gene, WT1, in males with genital abnormalities. J. Med. Genet. 30: 767–72.

Clement, TM, Bhandari, RK, Sadler-Riggleman, I, Skinner, MK. 2011. SRY directly regulates the neurotrophin 3 promoter during male sex determination and testis development in rats. Biol. Reprod. 85: 277–84.

Colvin, JS, Green, RP, Schmahl, J, Capel, B, Ornitz, DM. 2001. Male-to-female sex reversal in mice lacking fibroblast growth factor 9. Cell 104: 875–89.

Combes, AN, Wilhelm, D, Davidson, T, Dejana, E, Harley, V, Sinclair, A, Koopman, P. 2009. Endothelial cell migration directs testis cord formation. Dev. Biol. 326: 112–20.

Cool, J, Carmona, FD, Szucsik, JC, Capel, B. 2008. Peritubular myoid cells are not the migrating population required for testis cord formation in the XY gonad. Sex Dev. 2: 128–33.

Cory, AT, Boyer, A, Pilon, N, Lussier, JG, Silversides, DW. 2007. Presumptive pre-Sertoli cells express genes involved in cell proliferation and cell signalling during a critical window in early testis differentiation. Mol. Reprod. Dev. 74: 1491–504.

Costa, G, Mazan, A, Gandillet, A, Pearson, S, Lacaud, G, Kouskoff, V. 2012. SOX7 regulates the expression of VE-cadherin in the haemogenic endothelium at the onset of haematopoietic development. Development 139: 1587–98.

Couse, JF, Hewitt, SC, Bunch, DO, Sar, M, Walker, VR, Davis, BJ, Korach, KS. 1999. Postnatal sex reversal of the ovaries in mice lacking estrogen receptors alpha and beta. Science 286: 2328–31.

Cox, JJ, Willatt, L, Homfray, T, Woods, CG. 2011. A SOX9 duplication and familial 46,XX developmental testicular disorder. N. Engl. J. Med. 364: 91–3.

Crisponi, L, Deiana, M, Loi, A, Chiappe, F, Uda, M, Amati, P, Bisceglia, L, Zelante, L, Nagaraja, R, Porcu, S, Ristaldi, MS, Marzella, R, Rocchi, M, Nicolino, M, Lienhardt-Roussie, A, Nivelon, A, Verloes, A, Schlessinger, D, Gasparini, P, Bonneau, D, Cao, A, Pilia, G. 2001. The putative forkhead transcription factor FOXL2 is mutated in blepharophimosis/ptosis/epicanthus inversus syndrome. Nat. Genet. 27: 159–66.

Das, DK, Sanghavi, D, Gawde, H, Idicula-Thomas, S, Vasudevan, L. Novel homozygous mutations in Desert hedgehog gene in patients with 46,XY complete gonadal dysgenesis and prediction of its structural and functional implications by computational methods. Eur. J. Med. Genet. 54: e529–34.

De Grandi, A, Calvari, V, Bertini, V, Bulfone, A, Peverali, G, Camerino, G, Borsani, G, Guioli, S. 2000. The expression pattern of a mouse doublesex-related gene is consistent with a role in gonadal differentiation. Mech. Dev. 90: 323–6.

De Santa Barbara, P, Méjean, C, Moniot, B, Malclès, MH, Berta, P, Boizet-Bonhoure, B. 2001. Steroidogenic factor-1 contributes to the cyclic-adenosine monophosphate down-regulation of human SRY gene expression. Biol. Reprod. 64: 775–83.

DeFalco, T, Capel, B. 2009. Gonad morphogenesis in vertebrates: divergent means to a convergent end. Annu. Rev. Cell Dev. Biol. 25: 457–82.

Délot, EC, Vilain, EJ. 2015. Nonsyndromic 46,XX Testicular Disorders of Sex Development.

Dipietromaria, A, Benayoun, BA, Todeschini, A-L, Rivals, I, Bazin, C, Veitia, RA. 2009. Towards a functional classification of pathogenic FOXL2 mutations using transactivation reporter systems. Hum. Mol. Genet. 18: 3324–33.

Elzaiat, M, Jouneau, L, Thepot, D, Klopp, C, Allais-Bonnet, A, Cabau, C, Andre, M, Chaffaux, S, Cribiu, E-P, Pailhoux, E, Pannetier, M. 2014. High-Throughput Sequencing Analyses of XX Genital Ridges Lacking FOXL2 Reveal DMRT1 Up-Regulation Before SOX9 Expression During the Sex-Reversal Process in Goats. Biol. Reprod. 91: 153–153.

Ferraz-de-Souza, B, Lin, L, Achermann, JC. 2011. Steroidogenic factor-1 (SF-1, NR5A1) and human disease. Mol. Cell. Endocrinol. 336: 198–205.

Finelli, P, Pincelli, AI, Russo, S, Bonati, MT, Recalcati, MP, Masciadri, M, Giardino, D, Cavagnini, F, Larizza, L. 2007. Disruption of friend of GATA 2 gene (FOG-2) by a de novo t(8;10) chromosomal translocation is associated with heart defects and gonadal dysgenesis. Clin. Genet. 71: 195–204.

Foster, JW, Dominguez-Steglich, MA, Guioli, S, Kwok, C, Weller, PA, Stevanović, M, Weissenbach, J, Mansour, S, Young, ID, Goodfellow, PN. 1994. Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. Nature 372: 525–30.

Francois, M, Koopman, P, Beltrame, M. 2010. SoxF genes: Key players in the development of the cardio-vascular system. Int. J. Biochem. Cell Biol. 42: 445–8.

Fujimoto, Y, Tanaka, SS, Yamaguchi, YL, Kobayashi, H, Kuroki, S, Tachibana, M,

Shinomura, M, Kanai, Y, Morohashi, K-I, Kawakami, K, Nishinakamura, R. 2013. Homeoproteins Six1 and Six4 regulate male sex determination and mouse gonadal development. Dev. Cell 26: 416–30.

Futaki, S, Hayashi, Y, Emoto, T, Weber, N, Sekiguchi, K, Weber, CN. 2004. Sox7 Plays Crucial Roles in Parietal Endoderm Differentiation in F9 Embryonal Carcinoma Cells through Regulating Gata-4 and Gata-6 Expression Sox7 Plays Crucial Roles in Parietal Endoderm Differentiation in F9 Embryonal Carcinoma Cells through Regulating .

Gandillet, A, Serrano, AG, Pearson, S, Lie-A-Ling, M, Lacaud, G, Kouskoff, V. 2009. Sox7-sustained expression alters the balance between proliferation and differentiation of hematopoietic progenitors at the onset of blood specification. Blood 114: 4813–22.

Garg, V, Kathiriya, IS, Barnes, R, Schluterman, MK, King, IN, Butler, CA, Rothrock, CR, Eapen, RS, Hirayama-Yamada, K, Joo, K, Matsuoka, R, Cohen, JC, Srivastava, D. 2003. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature 424: 443–7.

Gassei, K, Schlatt, S. 2007. Testicular morphogenesis: comparison of in vivo and in vitro models to study male gonadal development. Ann. N. Y. Acad. Sci. 1120: 152–67.

Ge, R-S, Dong, Q, Sottas, CM, Papadopoulos, V, Zirkin, BR, Hardy, MP. 2006. In search of rat stem Leydig cells: identification, isolation, and lineage-specific development. Proc. Natl. Acad. Sci. U. S. A. 103: 2719–24.

Georg, I, Bagheri-Fam, S, Knower, KC, Wieacker, P, Scherer, G, Harley, VR. 2010. Mutations of the SRY-responsive enhancer of SOX9 are uncommon in XY gonadal dysgenesis. Sex Dev. 4: 321–5.

Georges, A, L'Hôte, D, Todeschini, AL, Auguste, A, Legois, B, Zider, A, Veitia, RA. 2014. The transcription factor FOXL2 mobilizes estrogen signaling to maintain the identity of ovarian granulosa cells. Elife 3.

Gessler, M, Poustka, A, Cavenee, W, Neve, RL, Orkin, SH, Bruns, GA. 1990. Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosome jumping. Nature 343: 774–8.

Gordon, DF, Lewis, SR, Haugen, BR, James, RA, McDermott, MT, Wood, WM, Ridgway, EC. 1997. Pit-1 and GATA-2 interact and functionally cooperate to activate the thyrotropin beta-subunit promoter. J. Biol. Chem. 272: 24339–47.

Granados-Riveron, JT, Pope, M, Bu'Lock, FA, Thornborough, C, Eason, J, Setchfield, K, Ketley, A, Kirk, EP, Fatkin, D, Feneley, MP, Harvey, RP, Brook, JD. 2012. Combined mutation screening of NKX2-5, GATA4, and TBX5 in congenital heart disease: Multiple heterozygosity and novel mutations. Congenit. Heart Dis. 7: 151–159.

Guellaen, G, Casanova, M, Bishop, C, Geldwerth, D, Andre, G, Fellous, M, Weissenbach, J. 1984. Human XX males with Y single-copy DNA fragments. Nature 307: 172–3.

Guo, L, Zhong, D, Lau, S, Liu, X, Dong, X-Y, Sun, X, Yang, VW, Vertino, PM, Moreno, CS, Varma, V, Dong, J-T, Zhou, W. 2008. Sox7 Is an independent checkpoint for beta-catenin function in prostate and colon epithelial cells. Mol. Cancer Res. 6: 1421–30.

Guth, SIE, Wegner, M. 2008. Having it both ways: Sox protein function between conservation and innovation. Cell. Mol. Life Sci. 65: 3000–18.

Habert, R, Lejeune, H, Saez, JM. 2001. Origin, differentiation and regulation of fetal and adult Leydig cells. Mol. Cell. Endocrinol. 179: 47–74.

Hammes, A, Guo, JK, Lutsch, G, Leheste, JR, Landrock, D, Ziegler, U, Gubler, MC, 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.

Hanley, NA, Ball, SG, Clement-Jones, M, Hagan, DM, Strachan, T, Lindsay, S, Robson, S, Ostrer, H, Parker, KL, Wilson, DI. 1999. Expression of steroidogenic factor 1 and Wilms' tumour 1 during early human gonadal development and sex determination. Mech. Dev. 87: 175–80.

Hanley, NA, Hagan, DM, Clement-Jones, M, Ball, SG, Strachan, T, Salas-Cortés, L,

McElreavey, K, Lindsay, S, Robson, S, Bullen, P, Ostrer, H, Wilson, DI. 2000. SRY, SOX9, and DAX1 expression patterns during human sex determination and gonadal development. Mech. Dev. 91: 403–7.

Haqq, CM, King, CY, Ukiyama, E, Falsafi, S, Haqq, TN, Donahoe, PK, Weiss, MA. 1994. Molecular basis of mammalian sexual determination: activation of Müllerian inhibiting substance gene expression by SRY. Science 266: 1494–500.

Harris, SE, Chand, AL, Winship, IM, Gersak, K, Aittomäki, K, Shelling, AN. 2002. Identification of novel mutations in FOXL2 associated with premature ovarian failure. Mol. Hum. Reprod. 8: 729–33.

Hatano, O, Takakusu, A, Nomura, M, Morohashi, K. 1996. Identical origin of adrenal cortex and gonad revealed by expression profiles of Ad4BP/SF-1. Genes Cells 1: 663–71.

Heikinheimo, M, Ermolaeva, M, Bielinska, M, Rahman, NA, Narita, N, Huhtaniemi, IT, Tapanainen, JS, Wilson, DB. 1997. Expression and hormonal regulation of transcription factors GATA-4 and GATA-6 in the mouse ovary. Endocrinology 138: 3505–14.

Herpers, R, van de Kamp, E, Duckers, HJ, Schulte-Merker, S. 2008. Redundant roles for sox7 and sox18 in arteriovenous specification in zebrafish. Circ. Res. 102: 12–5.

Herpin, A, Braasch, I, Kraeussling, M, Schmidt, C, Thoma, EC, Nakamura, S, Tanaka, M, Schartl, M. 2010. Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. PLoS Genet. 6: e1000844.

Hofmann, MC, Narisawa, S, Hess, RA, Millán, JL. 1992. Immortalization of germ cells and somatic testicular cells using the SV40 large T antigen. Exp. Cell Res. 201: 417–35.

Honda, S, Morohashi, K, Nomura, M, Takeya, H, Kitajima, M, Omura, T. 1993. Ad4BP regulating steroidogenic P-450 gene is a member of steroid hormone receptor superfamily. J. Biol. Chem. 268: 7494–502.

Hossain, A, Saunders, GF. 2001. The human sex-determining gene SRY is a direct target of WT1. J. Biol. Chem. 276: 16817–23.

Hu, Y-C, Okumura, LM, Page, DC. 2013. Gata4 is required for formation of the genital ridge in mice. PLoS Genet. 9: e1003629.

Huang, B, Wang, S, Ning, Y, Lamb, AN, Bartley, J. 1999. Autosomal XX sex reversal caused by duplication of SOX9. Am. J. Med. Genet. 87: 349–53.

Huang, Y-H, Lee, C-Y, Tai, P-J, Yen, C-C, Liao, C-Y, Chen, W-J, Liao, C-J, Cheng, W-L, Chen, R-N, Wu, S-M, Wang, C-S, Lin, K-H. 2006. Indirect regulation of human dehydroepiandrosterone sulfotransferase family 1A member 2 by thyroid hormones. Endocrinology 147: 2481–9.

Hughes, IA, Houk, C, Ahmed, SF, Lee, PA. 2006. Consensus statement on management of intersex disorders. J. Pediatr. Urol. 2: 148–62.

Hyon, C, Chantot-Bastaraud, S, Harbuz, R, Bhouri, R, Perrot, N, Peycelon, M, Sibony, M, Rojo, S, Piguel, X, Bilan, F, Gilbert-Dussardier, B, Kitzis, A, McElreavey, K, Siffroi, J-P, Bashamboo, A. 2015. Refining the regulatory region upstream of SOX9 associated with 46,XX testicular disorders of sex development (DSD). Am. J. Med. Genet. A.

Ikeda, Y, Shen, WH, Ingraham, HA, Parker, KL. 1994. Developmental expression of mouse steroidogenic factor-1, an essential regulator of the steroid hydroxylases. Mol. Endocrinol. 8: 654–62.

Ikeda, Y, Swain, A, Weber, TJ, Hentges, KE, Zanaria, E, Lalli, E, Tamai, KT, Sassone-Corsi, P, Lovell-Badge, R, Camerino, G, Parker, KL. 1996. Steroidogenic factor 1 and Dax-1 colocalize in multiple cell lineages: potential links in endocrine development. Mol. Endocrinol. 10: 1261–72.

Ikeda, Y, Takeda, Y, Shikayama, T, Mukai, T, Hisano, S, Morohashi, KI. 2001. Comparative localization of Dax-1 and Ad4BP/SF-1 during development of the hypothalamic-pituitary-gonadal axis suggests their closely related and distinct functions. Dev. Dyn. 220: 363–76.

Ion, A, Telvi, L, Chaussain, JL, Galacteros, F, Valayer, J, Fellous, M, McElreavey, K. 1996. A novel mutation in the putative DNA helicase XH2 is responsible for male-to-female sex reversal associated with an atypical form of the ATR-X syndrome. Am. J. Hum. Genet. 58: 1185–91.

Ito, E, Toki, T, Ishihara, H, Ohtani, H, Gu, L, Yokoyama, M, Engel, JD, Yamamoto, M. 1993. Erythroid transcription factor GATA-1 is abundantly transcribed in mouse testis. Nature 362: 466–8.

Jameson, S a., Lin, YT, Capel, B. 2012. Testis development requires the repression of Wnt4 by Fgf signaling. Dev. Biol. 370: 24–32.

Jamieson, S, Fuller, PJ. 2012. Molecular pathogenesis of granulosa cell tumors of the ovary. Endocr. Rev. 33: 109–44.

Jeske, YW, Mishina, Y, Cohen, DR, Behringer, RR, Koopman, P. 1996. Analysis of the role of Amh and Fra1 in the Sry regulatory pathway. Mol. Reprod. Dev. 44: 153–8.

Jimenez, P, Saner, K, Mayhew, B, Rainey, WE. 2003. GATA-6 is expressed in the human adrenal and regulates transcription of genes required for adrenal androgen biosynthesis. Endocrinology 144: 4285–8.

Jiménez, R, Burgos, M. 1998. Mammalian sex determination: joining pieces of the genetic puzzle. Bioessays 20: 696–9.

Jordan, BK, Mohammed, M, Ching, ST, Délot, E, Chen, XN, Dewing, P, Swain, A, Rao, PN, Elejalde, BR, Vilain, E. 2001. Up-regulation of WNT-4 signaling and dosage-sensitive sex reversal in humans. Am. J. Hum. Genet. 68: 1102–9.

Jordan, BK, Shen, JH-C, Olaso, R, Ingraham, HA, Vilain, E. 2003. Wnt4 overexpression disrupts normal testicular vasculature and inhibits testosterone synthesis by repressing steroidogenic factor 1/beta-catenin synergy. Proc. Natl. Acad. Sci. U. S. A. 100: 10866–71.

Josso, N, Belville, C, di Clemente, N, Picard, J-Y. 2005. AMH and AMH receptor defects in persistent Müllerian duct syndrome. Hum. Reprod. Update 11: 351–6.

Josso, N, Picard, JY, Rey, R, di Clemente, N. 2006. Testicular anti-Müllerian hormone: history, genetics, regulation and clinical applications. Pediatr. Endocrinol. Rev. 3: 347–58.

Kamachi, Y, Kondoh, H. 2013. Sox proteins: regulators of cell fate specification and differentiation. Development 140: 4129–44.

Karl, J, Capel, B. 1998. Sertoli cells of the mouse testis originate from the coelomic epithelium. Dev. Biol. 203: 323–33.

Kashimada, K, Koopman, P. 2010. Sry: the master switch in mammalian sex determination. Development 137: 3921–30.

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, Owaki, A, Toyama, Y, Kusaka, M, Shinohara, Y, Maekawa, M, Toshimori, K, Morohashi, K-I. 2005. Mouse Polycomb M33 is required for splenic vascular and adrenal gland formation through regulating Ad4BP/SF1 expression. Blood 106: 1612–20.

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.

Kawakami, K, Sato, S, Ozaki, H, Ikeda, K. 2000. Six family genes--structure and function as transcription factors and their roles in development. Bioessays 22: 616–26.

Kelley, LA, Sternberg, MJE. 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. 4: 363–71.

Ketola, I, Pentikäinen, V, Vaskivuo, T, Ilvesmäki, V, Herva, R, Dunkel, L, Tapanainen, JS, Toppari, J, Heikinheimo, M. 2000. Expression of transcription factor GATA-4 during human testicular development and disease. J. Clin. Endocrinol. Metab. 85: 3925–31.

Ketola, I, Rahman, N, Toppari, J, Bielinska, M, Porter-Tinge, SB, Tapanainen, JS,

Huhtaniemi, IT, Wilson, DB, Heikinheimo, M. 1999a. Expression and regulation of transcription factors GATA-4 and GATA-6 in developing mouse testis. Endocrinology 140: 1470–80.

Ketola, I, Rahman, N, Toppari, J, Bielinska, M, Porter-Tinge, SB, Tapanainen, JS, Huhtaniemi, IT, Wilson, DB, Heikinheimo, M. 1999b. Expression and regulation of transcription factors GATA-4 and GATA-6 in developing mouse testis. Endocrinology 140: 1470–1480.

Ketola, I, Toppari, J, Vaskivuo, T, Herva, R, Tapanainen, JS, Heikinheimo, M. 2003. Transcription factor GATA-6, cell proliferation, apoptosis, and apoptosis-related proteins Bcl-2 and Bax in human fetal testis. J. Clin. Endocrinol. Metab. 88: 1858–1865.

Kim, G-J, Sock, E, Buchberger, A, Just, W, Denzer, F, Hoepffner, W, German, J, Cole, T, Mann, J, Seguin, JH, Zipf, W, Costigan, C, Schmiady, H, Rostásy, M, Kramer, M, Kaltenbach, S, Rösler, B, Georg, I, Troppmann, E, Teichmann, A-C, Salfelder, A, Widholz, SA, 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, Y, Bingham, N, Sekido, R, Parker, KL, Lovell-Badge, R, Capel, B. 2007. Fibroblast growth factor receptor 2 regulates proliferation and Sertoli differentiation during male sex determination. Proc. Natl. Acad. Sci. U. S. A. 104: 16558–63.

Kim, Y, Capel, B. 2006. Balancing the bipotential gonad between alternative organ fates: a new perspective on an old problem. Dev. Dyn. 235: 2292–300.

Knower, KC, Kelly, S, Ludbrook, LM, Bagheri-Fam, S, Sim, H, Bernard, P, Sekido, R, Lovell-Badge, R, Harley, VR. 2011. Failure of SOX9 regulation in 46XY disorders of sex development with SRY, SOX9 and SF1 mutations. PLoS One 6: e17751.

Knower, KC, Sim, H, McClive, PJ, Bowles, J, Koopman, P, Sinclair, AH, Harley, VR. 2007. Characterisation of urogenital ridge gene expression in the human embryonal carcinoma cell line NT2/D1. Sex Dev. 1: 114–26.

Koopman, P, Münsterberg, A, Capel, B, Vivian, N, Lovell-Badge, R. 1990. Expression of a candidate sex-determining gene during mouse testis differentiation. Nature 348: 450–2.

Kousta, E, Papathanasiou, A, Skordis, N. Sex determination and disorders of sex development according to the revised nomenclature and classification in 46,XX individuals. Hormones (Athens). 9: 218–131.

Kreidberg, JA, Sariola, H, Loring, JM, Maeda, M, Pelletier, J, Housman, D, Jaenisch, R. 1993. WT-1 is required for early kidney development. Cell 74: 679–91.

Krentz, AD, Murphy, MW, Sarver, AL, Griswold, MD, Bardwell, VJ, Zarkower, D. 2011. DMRT1 promotes oogenesis by transcriptional activation of Stra8 in the mammalian fetal ovary. Dev. Biol. 356: 63–70.

Kumar, JP. 2009. The sine oculis homeobox (SIX) family of transcription factors as regulators of development and disease. Cell. Mol. Life Sci. 66: 565–83.

Kuo, CT, Morrisey, EE, Anandappa, R, Sigrist, K, Lu, MM, Parmacek, MS, Soudais, C, Leiden, JM. 1997. GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. Genes Dev. 11: 1048–60.

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.

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.

Laitinen, MP, Anttonen, M, Ketola, I, Wilson, DB, Ritvos, O, Butzow, R, Heikinheimo, M. 2000. Transcription factors GATA-4 and GATA-6 and a GATA family cofactor, FOG-2, are

expressed in human ovary and sex cord-derived ovarian tumors. J. Clin. Endocrinol. Metab. 85: 3476–83.

Lambeth, LS, Raymond, CS, Roeszler, KN, Kuroiwa, A, Nakata, T, Zarkower, D, Smith, CA. 2014. Over-expression of DMRT1 induces the male pathway in embryonic chicken gonads. Dev. Biol. 389: 160–72.

Laumonnier, F, Ronce, N, Hamel, BCJ, Thomas, P, Lespinasse, J, Raynaud, M, Paringaux, C, Van Bokhoven, H, Kalscheuer, V, Fryns, J-P, Chelly, J, Moraine, C, Briault, S. 2002. Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. Am. J. Hum. Genet. 71: 1450–5.

Lavery, R, Chassot, A-A, Pauper, E, Gregoire, EP, Klopfenstein, M, de Rooij, DG, Mark, M, Schedl, A, Ghyselinck, NB, Chaboissier, M-C. 2012. Testicular differentiation occurs in absence of R-spondin1 and Sox9 in mouse sex reversals. PLoS Genet. 8: e1003170.

Lavoie, HA, McCoy, GL, Blake, CA. 2004. Expression of the GATA-4 and GATA-6 transcription factors in the fetal rat gonad and in the ovary during postnatal development and pregnancy. Mol. Cell. Endocrinol. 227: 31–40.

Law, MJ, Lower, KM, Voon, HPJ, Hughes, JR, Garrick, D, Viprakasit, V, Mitson, M, De Gobbi, M, Marra, M, Morris, A, Abbott, A, Wilder, SP, Taylor, S, Santos, GM, Cross, J, Ayyub, H, Jones, S, Ragoussis, J, Rhodes, D, Dunham, I, Higgs, DR, Gibbons, RJ. 2010. ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. Cell 143: 367–78.

Lawson, KA, Hage, WJ. 1994. Clonal analysis of the origin of primordial germ cells in the mouse. Ciba Found. Symp. 182: 68–84; discussion 84–91.

Ledig, S, Hiort, O, Wünsch, L, Wieacker, P. 2012. Partial deletion of DMRT1 causes 46,XY ovotesticular disorder of sexual development. Eur. J. Endocrinol. 167: 119–24.

Lei, N, Heckert, LL. 2004. Gata4 Regulates Testis Expression of Dmrt1. 24: 377–388.

Lei, N, Heckert, LL. 2004. Gata4 regulates testis expression of Dmrt1. Mol. Cell. Biol. 24: 377–88.

Lei, N, Heckert, LL. 2004a. Gata4 Regulates Testis Expression of Dmrt1. 24: 377–388.

Lei, N, Heckert, LL. 2004b. Gata4 regulates testis expression of Dmrt1. Mol. Cell. Biol. 24: 377–88.

Lei, N, Karpova, T, Hornbaker, KI, Rice, DA, Heckert, LL. 2009. Distinct Transcriptional Mechanisms Direct Expression of the Rat Dmrt1 Promoter in Sertoli Cells and Germ Cells of Transgenic Mice 1. 125: 118–125.

Lin, L, Achermann, JC. 2008. Steroidogenic factor-1 (SF-1, Ad4BP, NR5A1) and disorders of testis development. Sex Dev. 2: 200–9.

Lin, Y-T, Capel, B. 2015. Cell fate commitment during mammalian sex determination. Curr. Opin. Genet. Dev. 32: 144–152.

Liu, C-F, Liu, C, Yao, HH-C. 2010. Building pathways for ovary organogenesis in the mouse embryo. Curr. Top. Dev. Biol. 90: 263–90.

Longoni, M, Lage, K, Russell, MK, Loscertales, M, Abdul-Rahman, OA, Baynam, G, Bleyl, SB, Brady, PD, Breckpot, J, Chen, CP, Devriendt, K, Gillessen-Kaesbach, G, Grix, AW, Rope, AF, Shimokawa, O, Strauss, B, Wieczorek, D, Zackai, EH, Coletti, CM, Maalouf, FI, Noonan, KM, Park, JH, Tracy, AA, Lee, C, Donahoe, PK, Pober, BR. 2012. Congenital diaphragmatic hernia interval on chromosome 8p23.1 characterized by genetics and protein interaction networks. Am. J. Med. Genet. A 158A: 3148–58.

Lourenço, D, Brauner, R, Rybczynska, M, Nihoul-Fékété, C, McElreavey, K, Bashamboo, A. 2011. Loss-of-function mutation in GATA4 causes anomalies of human testicular development.

Lourenço, D, Brauner, R, Rybczynska, M, Nihoul-Fékété, C, McElreavey, K, Bashamboo, A. 2011a. Loss-of-function mutation in GATA4 causes anomalies of human testicular development. Proc. Natl. Acad. Sci. U. S. A. 108: 1597–602.

Lourenço, D, Brauner, R, Rybczynska, M, Nihoul-Fékété, C, McElreavey, K, Bashamboo, A. 2011b. Loss-of-function mutation in GATA4 causes anomalies of human testicular development.

Lowry, JA, Atchley, WR. 2000. Molecular evolution of the GATA family of transcription factors: conservation within the DNA-binding domain. J. Mol. Evol. 50: 103–15.

Ludbrook, LM, Harley, VR. 2004. Sex determination: a “window” of DAX1 activity. Trends Endocrinol. Metab. 15: 116–21.

Luo, X, Ikeda, Y, Parker, KL. 1994. A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. Cell 77: 481–90.

Maatouk, DM, Capel, B. 2008. Sexual development of the soma in the mouse. Curr. Top. Dev. Biol. 83: 151–83.

Maatouk, DM, DiNapoli, L, Alvers, A, Parker, KL, Taketo, MM, Capel, B. 2008. Stabilization of beta-catenin in XY gonads causes male-to-female sex-reversal. Hum. Mol. Genet. 17: 2949–55.

Malki, S, Nef, S, Notarnicola, C, Thevenet, L, Gasca, S, Méjean, C, Berta, P, Poulat, F, Boizet-Bonhoure, B. 2005. Prostaglandin D2 induces nuclear import of the sex-determining factor SOX9 via its cAMP-PKA phosphorylation. EMBO J. 24: 1798–809.

Mandel, H, Shemer, R, Borochowitz, ZU, Okopnik, M, Knopf, C, Indelman, M, Drugan, A, Tiosano, D, Gershoni-Baruch, R, Choder, M, Sprecher, E. 2008. SERKAL syndrome: an autosomal-recessive disorder caused by a loss-of-function mutation in WNT4. Am. J. Hum. Genet. 82: 39–47.

Manuylov, NL, Smagulova, FO, Leach, L, Tevosian, SG. 2008. Ovarian development in mice requires the GATA4-FOG2 transcription complex. Development 135: 3731–43.

Manuylov, NL, Zhou, B, Ma, Q, Fox, SC, Pu, WT, Tevosian, SG. 2011. Conditional ablation of Gata4 and Fog2 genes in mice reveals their distinct roles in mammalian sexual differentiation. Dev. Biol. 353: 229–41.

Mardis, ER. 2008. The impact of next-generation sequencing technology on genetics. Trends Genet. 24: 133–41.

Mark, M, Jacobs, H, Oulad-Abdelghani, M, Dennefeld, C, Féret, B, Vernet, N, Codreanu, C-A, Chambon, P, Ghyselinck, NB. 2008. STRA8-deficient spermatocytes initiate, but fail to complete, meiosis and undergo premature chromosome condensation. J. Cell Sci. 121: 3233–42.

Martin, LJ, Taniguchi, H, Robert, NM, Simard, J, Tremblay, JJ, Viger, RS. 2005. GATA factors and the nuclear receptors, steroidogenic factor 1/liver receptor homolog 1, are key mutual partners in the regulation of the human 3beta-hydroxysteroid dehydrogenase type 2 promoter. Mol. Endocrinol. 19: 2358–70.

Martineau, J, Nordqvist, K, Tilmann, C, Lovell-Badge, R, Capel, B. 1997. Male-specific cell migration into the developing gonad. Curr. Biol. 7: 958–68.

Mather, JP. 1980. Establishment and characterization of two distinct mouse testicular epithelial cell lines. Biol. Reprod. 23: 243–52.

Matson, CK, Murphy, MW, Sarver, AL, Griswold, MD, Bardwell, VJ, Zarkower, D. 2011. DMRT1 prevents female reprogramming in the postnatal mammalian testis. Nature 476: 101–4.

Matson, CK, Zarkower, D. 2012. Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. Nat. Rev. Genet.

Matsuda, M, Nagahama, Y, Shinomiya, A, Sato, T, Matsuda, C, Kobayashi, T, Morrey, CE, Shibata, N, Asakawa, S, Shimizu, N, Hori, H, Hamaguchi, S, Sakaizumi, M. 2002. DMY is a Y-specific DM-domain gene required for male development in the medaka fish. Nature 417: 559–63.

McCoard, SA, Lunstra, DD, Wise, TH, Ford, JJ. 2001. Specific staining of Sertoli cell nuclei and evaluation of Sertoli cell number and proliferative activity in Meishan and White

Composite boars during the neonatal period. Biol. Reprod. 64: 689–95.

Meeks, JJ, Weiss, J, Jameson, JL. 2003. Dax1 is required for testis determination. Nat. Genet. 34: 32–3.

Merchant-Larios, H, Moreno-Mendoza, N, Buehr, M. 1993. The role of the mesonephros in cell differentiation and morphogenesis of the mouse fetal testis. Int. J. Dev. Biol. 37: 407–15.

Merika, M, Orkin, SH. 1993. DNA-binding specificity of GATA family transcription factors. Mol. Cell. Biol. 13: 3999–4010.

Miller, JR. 2002. The Wnts. Genome Biol. 3: REVIEWS3001.

Minkina, A, Matson, CK, Lindeman, RE, Ghyselinck, NB, Bardwell, VJ, Zarkower, D. 2014. DMRT1 protects male gonadal cells from retinoid-dependent sexual transdifferentiation. Dev. Cell 29: 511–20.

Mitchell, RT, Saunders, PTK, Childs, AJ, Cassidy-Kojima, C, Anderson, RA, Wallace, WHB, Kelnar, CJH, Sharpe, RM. 2010. Xenografting of human fetal testis tissue: a new approach to study fetal testis development and germ cell differentiation. Hum. Reprod. 25: 2405–14.

Miyamoto, N, Yoshida, M, Kuratani, S, Matsuo, I, Aizawa, S. 1997. Defects of urogenital development in mice lacking Emx2. Development 124: 1653–64.

Miyamoto, Y, Taniguchi, H, Hamel, F, Silversides, DW, Viger, RS. 2008. A GATA4/WT1 cooperation regulates transcription of genes required for mammalian sex determination and differentiation. BMC Mol. Biol. 9: 44.

Mizusaki, H, Kawabe, K, Mukai, T, Ariyoshi, E, Kasahara, M, Yoshioka, H, Swain, A, Morohashi, K-I. 2003. Dax-1 (dosage-sensitive sex reversal-adrenal hypoplasia congenita critical region on the X chromosome, gene 1) gene transcription is regulated by wnt4 in the female developing gonad. Mol. Endocrinol. 17: 507–19.

Molkentin, JD, Lin, Q, Duncan, SA, Olson, EN. 1997. Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. Genes Dev. 11: 1061–72.

Molkentin, JD. 2000. The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. J. Biol. Chem. 275: 38949–52.

Moniot, B, Berta, P, Scherer, G, Südbeck, P, Poulat, F. 2000. Male specific expression suggests role of DMRT1 in human sex determination. Mech. Dev. 91: 323–5.

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.

Morceau, F, Schnekenburger, M, Dicato, M, Diederich, M. 2004. GATA-1: friends, brothers, and coworkers. Ann. N. Y. Acad. Sci. 1030: 537–54.

Morohashi, K, Honda, S, Inomata, Y, Handa, H, Omura, T. 1992. A common trans-acting factor, Ad4-binding protein, to the promoters of steroidogenic P-450s. J. Biol. Chem. 267: 17913–9.

Morohashi, K, Zanger, UM, Honda, S, Hara, M, Waterman, MR, Omura, T. 1993. Activation of CYP11A and CYP11B gene promoters by the steroidogenic cell-specific transcription factor, Ad4BP. Mol. Endocrinol. 7: 1196–204.

Münsterberg, A, Lovell-Badge, R. 1991. Expression of the mouse anti-müllerian hormone gene suggests a role in both male and female sexual differentiation. Development 113: 613–24.

Murakami, A, Shen, H, Dickson, C. 2004. Genes : Structure and Regulation : SOX7 and GATA-4 Are Competitive Activators of Fgf-3 Transcription SOX7 and GATA-4 Are Competitive Activators of Fgf-3 Transcription *.

Murphy, MW, Sarver, AL, Rice, D, Hatzi, K, Ye, K, Melnick, A, Heckert, LL, Zarkower, D, Bardwell, VJ. 2010. Genome-wide analysis of DNA binding and transcriptional regulation by

the mammalian Doublesex homolog DMRT1 in the juvenile testis. Proc. Natl. Acad. Sci. U. S. A. 107: 13360–5.

Murphy, MW, Zarkower, D, Bardwell, VJ. 2007. Vertebrate DM domain proteins bind similar DNA sequences and can heterodimerize on DNA. BMC Mol. Biol. 8: 58.

Nachtigal, MW, Hirokawa, Y, Enyeart-VanHouten, DL, Flanagan, JN, Hammer, GD, Ingraham, HA. 1998. Wilms' tumor 1 and Dax-1 modulate the orphan nuclear receptor SF-1 in sex-specific gene expression. Cell 93: 445–54.

Nakamura, Y, Suzuki, T, Sasano, H. 2007. Transcription factor GATA-6 in the human adrenocortex: association with adrenal development and aging. Endocr. J. 54: 783–9.

Nakamura, Y, Xing, Y, Sasano, H, Rainey, WE. 2009. The mediator complex subunit 1 enhances transcription of genes needed for adrenal androgen production. Endocrinology 150: 4145–53.

Niaudet, P, Gubler, M-C. 2006. WT1 and glomerular diseases. Pediatr. Nephrol. 21: 1653–60.

Nicholl, RM, Grimsley, L, Butler, L, Palmer, RW, Rees, HC, Savage, MO, Costeloe, K. 1994. Trisomy 22 and intersex. Arch. Dis. Child. Fetal Neonatal Ed. 71: F57–8.

Nicol, B, Yao, HH-C. 2015. Gonadal Identity in the Absence of Pro-Testis Factor SOX9 and Pro-Ovary Factor Beta-Catenin in Mice. Biol. Reprod.

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.

Nurmio, M, Toppari, J, Kallio, J, Hou, M, Söder, O, Jahnukainen, K. 2009. Functional in vitro model to examine cancer therapy cytotoxicity in maturing rat testis. Reprod. Toxicol. 27: 28–34.

O'Bryan, MK, Takada, S, Kennedy, CL, Scott, G, Harada, S, Ray, MK, Dai, Q, Wilhelm, D, de Kretser, DM, Eddy, EM, Koopman, P, Mishina, Y. 2008. Sox8 is a critical regulator of adult Sertoli cell function and male fertility. Dev. Biol. 316: 359–70.

Ohnesorg, T, Vilain, E, Sinclair, AH. 2014. The genetics of disorders of sex development in humans. Sex Dev. 8: 262–72.

Ottolenghi, C, McElreavey, K. 2000. Deletions of 9p and the quest for a conserved mechanism of sex determination. Mol. Genet. Metab. 71: 397–404.

Ottolenghi, C, Omari, S, Garcia-Ortiz, JE, Uda, M, Crisponi, L, Forabosco, A, Pilia, G, Schlessinger, D. 2005. Foxl2 is required for commitment to ovary differentiation. Hum. Mol. Genet. 14: 2053–62.

Ottolenghi, C, Uda, M, Crisponi, L, Omari, S, Cao, A, Forabosco, A, Schlessinger, D. 2007. Determination and stability of sex. BioEssays 29: 15–25.

Padua, MB, Fox, SC, Jiang, T, Morse, DA, Tevosian, SG. 2014. Simultaneous gene deletion of gata4 and gata6 leads to early disruption of follicular development and germ cell loss in the murine ovary. Biol. Reprod. 91: 24.

Paliwal, P, Sharma, A, Birla, S, Kriplani, A, Khadgawat, R, Sharma, A. 2011. Identification of novel SRY mutations and SF1 (NR5A1) changes in patients with pure gonadal dysgenesis and 46,XY karyotype. Mol. Hum. Reprod. 17: 372–8.

Pannetier, M, Tilly, G, Kocer, A, Hudrisier, M, Renault, L, Chesnais, N, Costa, J, Le Provost, F, Vaiman, D, Vilotte, J-L, Pailhoux, E. 2006. Goat SRY induces testis development in XX transgenic mice. FEBS Lett. 580: 3715–20.

Park, SY, Meeks, JJ, Raverot, G, Pfaff, LE, Weiss, J, Hammer, GD, Jameson, JL. 2005. Nuclear receptors Sf1 and Dax1 function cooperatively to mediate somatic cell differentiation during testis development. Development 132: 2415–23.

Parma, P, Pailhoux, E, Cotinot, C. 1999. Reverse transcription-polymerase chain reaction analysis of genes involved in gonadal differentiation in pigs. Biol. Reprod. 61: 741–8.

Parma, P, Radi, O, Vidal, V, Chaboissier, MC, 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.

Pearlman, A, Loke, J, Le Caignec, C, White, S, Chin, L, Friedman, A, Warr, N, Willan, J, Brauer, D, Farmer, C, Brooks, E, Oddoux, C, Riley, B, Shajahan, S, Camerino, G, Homfray, T, Crosby, AH, Couper, J, David, A, Greenfield, A, Sinclair, A, Ostrer, H. 2010. Mutations in MAP3K1 cause 46,XY disorders of sex development and implicate a common signal transduction pathway in human testis determination. Am. J. Hum. Genet. 87: 898–904.

Petersen, C, Söder, O. 2006. The Sertoli Cell – A Hormonal Target and "Super" Nurse for Germ Cells That Determines Testicular Size. Horm. Res. 66: 153–161.

Philibert, P, Biason-Lauber, A, Gueorguieva, I, Stuckens, C, Pienkowski, C, Lebon-Labich, B, Paris, F, Sultan, C. 2011. Molecular analysis of WNT4 gene in four adolescent girls with mullerian duct abnormality and hyperandrogenism (atypical Mayer-Rokitansky-Küster-Hauser syndrome). Fertil. Steril. 95: 2683–2686.

Philibert, P, Biason-Lauber, A, Rouzier, R, Pienkowski, C, Paris, F, Konrad, D, Schoenle, E, Sultan, C. 2008. Identification and Functional Analysis of a New WNT4 Gene Mutation among 28 Adolescent Girls with Primary Amenorrhea and Müllerian Duct Abnormalities: A French Collaborative Study. J. Clin. Endocrinol. Metab. 93: 895–900.

Pierucci-Alves, F, Clark, AM, Russell, LD. 2001. A developmental study of the Desert hedgehog-null mouse testis. Biol. Reprod. 65: 1392–402.

Pilon, N, Daneau, I, Paradis, V, Hamel, F, Lussier, JG, Viger, RS, Silversides, DW. 2003. Porcine SRY promoter is a target for steroidogenic factor 1. Biol. Reprod. 68: 1098–106.

Polanco, JC, 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.

Ramírez-Solis, R, Davis, AC, Bradley, A. 1993. Gene targeting in embryonic stem cells. Methods Enzymol. 225: 855–78.

Raymond, CS, Murphy, MW, O'Sullivan, MG, Bardwell, VJ, Zarkower, D. 2000. Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. Genes Dev. 14: 2587–95.

Raymond, CS, Parker, ED, Kettlewell, JR, Brown, LG, Page, DC, Kusz, K, Jaruzelska, J, Reinberg, Y, Flejter, WL, Bardwell, VJ, Hirsch, B, Zarkower, D. 1999. A region of human chromosome 9p required for testis development contains two genes related to known sexual regulators. Hum. Mol. Genet. 8: 989–96.

Raymond, CS, Shamu, CE, Shen, MM, Seifert, KJ, Hirsch, B, Hodgkin, J, Zarkower, D. 1998. Evidence for evolutionary conservation of sex-determining genes. Nature 391: 691–5.

Real, FM, Sekido, R, Lupiáñez, DG, Lovell-Badge, R, Jiménez, R, Burgos, M. 2013. A microRNA (mmu-miR-124) prevents Sox9 expression in developing mouse ovarian cells. Biol. Reprod. 89: 78.

Rey, R, Lukas-Croisier, C, Lasala, C, Bedecarrás, P. 2003. AMH/MIS: what we know already about the gene, the protein and its regulation. Mol. Cell. Endocrinol. 211: 21–31.

Robert, NM, Tremblay, JJ, Viger, RS. 2002. Friend of GATA (FOG)-1 and FOG-2 differentially repress the GATA-dependent activity of multiple gonadal promoters. Endocrinology 143: 3963–73.

Rodriguez-Sosa, JR, Dobrinski, I. 2009. Recent developments in testis tissue xenografting. Reproduction 138: 187–94.

Rossi, P. 2013. Transcriptional control of KIT gene expression during germ cell development. 184: 179–184.

Sadovsky, Y, Crawford, PA, Woodson, KG, Polish, JA, Clements, MA, Tourtellotte, LM, Simburger, K, Milbrandt, J. 1995. Mice deficient in the orphan receptor steroidogenic factor 1 lack adrenal glands and gonads but express P450 side-chain-cleavage enzyme in the placenta and have normal embryonic serum levels of corticosteroids. Proc. Natl. Acad. Sci. U. S. A. 92: 10939–43.

Saner, KJ, Suzuki, T, Sasano, H, Pizzey, J, Ho, C, Strauss, JF, Carr, BR, Rainey, WE. 2005. Steroid sulfotransferase 2A1 gene transcription is regulated by steroidogenic factor 1 and GATA-6 in the human adrenal. Mol. Endocrinol. 19: 184–97.

Schepers, G, Wilson, M, Wilhelm, D, Koopman, P. 2003. SOX8 is expressed during testis differentiation in mice and synergizes with SF1 to activate the Amh promoter in vitro. J. Biol. Chem. 278: 28101–8.

Scherer, G, Held, M, Erdel, M, Meschede, D, Horst, J, Lesniewicz, R, Midro, AT. 1998. Three novel SRY mutations in XY gonadal dysgenesis and the enigma of XY gonadal dysgenesis cases without SRY mutations. Cytogenet. Cell Genet. 80: 188–92.

Schmahl, J, Eicher, EM, Washburn, LL, Capel, B. 2000. Sry induces cell proliferation in the mouse gonad. Development 127: 65–73.

Schmidt, D, Ovitt, CE, Anlag, K, Fehsenfeld, S, Gredsted, L, Treier, A-C, Treier, M. 2004. The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. Development 131: 933–42.

Seeherunvong, T, Perera, EM, Bao, Y, Benke, PJ, Benigno, A, Donahue, RP, Berkovitz, GD. 2004. 46,XX sex reversal with partial duplication of chromosome arm 22q. Am. J. Med. Genet. 127A: 149–151.

Séguin, CA, Draper, JS, Nagy, A, Rossant, J. 2008. Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. Cell Stem Cell 3: 182–95.

Sekido, R, Lovell-badge, R. 2008. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. 453.

Sekido, R, Lovell-Badge, R. 2008. Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature 453: 930–4.

Shah, SP, Köbel, M, Senz, J, Morin, RD, Clarke, BA, Wiegand, KC, Leung, G, Zayed, A, Mehl, E, Kalloger, SE, Sun, M, Giuliany, R, Yorida, E, Jones, S, Varhol, R, Swenerton, KD, Miller, D, Clement, PB, Crane, C, Madore, J, Provencher, D, Leung, P, DeFazio, A, Khattra, J, Turashvili, G, Zhao, Y, Zeng, T, Glover, JNM, Vanderhyden, B, Zhao, C, Parkinson, CA, Jimenez-Linan, M, Bowtell, DDL, Mes-Masson, A-M, Brenton, JD, Aparicio, SA, Boyd, N, Hirst, M, Gilks, CB, Marra, M, Huntsman, DG. 2009. Mutation of FOXL2 in granulosa-cell tumors of the ovary. N. Engl. J. Med. 360: 2719–29.

Shinomiya, A, Otake, H, Hamaguchi, S, Sakaizumi, M. 2010. Inherited XX sex reversal originating from wild medaka populations. Heredity (Edinb). 105: 443–8.

Siggers, P, Smith, L, Greenfield, A. 2002. Sexually dimorphic expression of Gata-2 during mouse gonad development. Mech. Dev. 111: 159–62.

Sinclair, AH, Berta, P, Palmer, MS, Hawkins, JR, Griffiths, BL, Smith, MJ, Foster, JW, Frischauf, AM, Lovell-Badge, R, Goodfellow, PN. 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature 346: 240–4.

Smith, CA, Roeszler, KN, Ohnesorg, T, Cummins, DM, Farlie, PG, Doran, TJ, Sinclair, AH. 2009. The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature 461: 267–71.

Stovall, DB, Cao, P, Sui, G. 2014. SOX7: from a developmental regulator to an emerging tumor suppressor. Histol. Histopathol. 29: 439–45.

Sugawara, T, Holt, JA, Kiriakidou, M, Strauss, JF. 1996. Steroidogenic factor 1-dependent promoter activity of the human steroidogenic acute regulatory protein (StAR) gene. Biochemistry 35: 9052–9.

Sutton, E, Hughes, J, White, S, Sekido, R, Tan, J, Arboleda, V, Rogers, N, Knower, K, Rowley, L, Eyre, H, Rizzoti, K, McAninch, D, Goncalves, J, Slee, J, Turbitt, E, Bruno, D, Bengtsson, H, Harley, V, Vilain, E, Sinclair, A, Lovell-Badge, R, Thomas, P. 2011. Identification of SOX3 as an XX male sex reversal gene in mice and humans. J. Clin. Invest. 121: 328–341.

Svingen, T, Koopman, P. 2013. Building the mammalian testis: origins, differentiation, and assembly of the component cell populations. Genes Dev. 27: 2409–26.

Swain, A, Narvaez, V, Burgoyne, P, Camerino, G, Lovell-Badge, R. 1998. Dax1 antagonizes Sry action in mammalian sex determination. Nature 391: 761–7.

Takasawa, K, Kashimada, K, Pelosi, E, Takagi, M, Morio, T, Asahara, H, Schlessinger, D, Mizutani, S, Koopman, P. 2014. FOXL2 transcriptionally represses Sf1 expression by antagonizing WT1 during ovarian development in mice. FASEB J. 28: 2020–8.

Takash, W, Cañizares, J, Bonneaud, N, Poulat, F, Mattéi, MG, Jay, P, Berta, P. 2001. SOX7 transcription factor: sequence, chromosomal localisation, expression, transactivation and interference with Wnt signalling. Nucleic Acids Res. 29: 4274–4283.

Tannour-Louet, M, Han, S, Corbett, ST, Louet, J-F, Yatsenko, S, Meyers, L, Shaw, CA, Kang, S-HL, Cheung, SW, Lamb, DJ. 2010. Identification of de novo copy number variants associated with human disorders of sexual development. PLoS One 5: e15392.

Tate, G, Satoh, H, Endo, Y, Mitsuya, T. 2000. Assignment of desert hedgehog (DHH) to human chromosome bands 12q12-->q13.1 by in situ hybridization. Cytogenet. Cell Genet. 88: 93–4.

Teerds, KJ, Huhtaniemi, IT. 2015. Morphological and functional maturation of Leydig cells: from rodent models to primates. Hum. Reprod. Update 21: 310–28.

Teixeira, J, Maheswaran, S, Donahoe, PK. 2001. Müllerian inhibiting substance: an instructive developmental hormone with diagnostic and possible therapeutic applications. Endocr. Rev. 22: 657–74.

Tevosian, SG, Albrecht, KH, Crispino, JD, Fujiwara, Y, Eicher, EM, Orkin, SH. 2002. Gonadal differentiation, sex determination and normal Sry expression in mice require direct interaction between transcription partners GATA4 and FOG2. Development 129: 4627–34.

Tevosian, SG, Deconinck, AE, Tanaka, M, Schinke, M, Litovsky, SH, Izumo, S, Fujiwara, Y, Orkin, SH. 2000. FOG-2, a cofactor for GATA transcription factors, is essential for heart morphogenesis and development of coronary vessels from epicardium. Cell 101: 729–39.

Thurisch, B, Liang, SY, Sarioglu, N, Schomburg, L, Bungert, J, Dame, C. 2009. Transgenic mice expressing small interfering RNA against Gata4 point to a crucial role of Gata4 in the heart and gonads. J. Mol. Endocrinol. 43: 157–69.

Todeschini, A-L, Dipietromaria, A, L'hôte, D, Boucham, FZ, Georges, AB, Pandaranayaka, PJE, Krishnaswamy, S, Rivals, I, Bazin, C, Veitia, RA. 2011. Mutational probing of the forkhead domain of the transcription factor FOXL2 provides insights into the pathogenicity of naturally occurring mutations. Hum. Mol. Genet. 20: 3376–85.

Tomita-Mitchell, A, Maslen, CL, Morris, CD, Garg, V, Goldmuntz, E. 2007. GATA4 sequence variants in patients with congenital heart disease. J. Med. Genet. 44: 779–783.

Tomizuka, K, Horikoshi, K, Kitada, R, Sugawara, Y, Iba, Y, Kojima, A, Yoshitome, A, Yamawaki, K, Amagai, M, Inoue, A, Oshima, T, Kakitani, M. 2008. R-spondin1 plays an essential role in ovarian development through positively regulating Wnt-4 signaling. Hum. Mol. Genet. 17: 1278–91.

Tremblay, JJ, Robert, NM, Viger, RS. 2001. Modulation of endogenous GATA-4 activity reveals its dual contribution to Müllerian inhibiting substance gene transcription in Sertoli cells. Mol. Endocrinol. 15: 1636–50.

Tremblay, JJ, Viger, RS. 1999. Transcription factor GATA-4 enhances Müllerian inhibiting substance gene transcription through a direct interaction with the nuclear receptor SF-1. Mol. Endocrinol. 13: 1388–1401.

Tremblay, JJ, Viger, RS. 2001. Nuclear receptor Dax-1 represses the transcriptional cooperation between GATA-4 and SF-1 in Sertoli cells. Biol. Reprod. 64: 1191–9.

Tsuji, M, Shima, H, Yonemura, CY, Brody, J, Donahoe, PK, Cunha, GR. 1992. Effect of human recombinant mullerian inhibiting substance on isolated epithelial and mesenchymal cells during mullerian duct regression in the rat. Endocrinology 131: 1481–8.

Uhlenhaut, NH, Jakob, S, Anlag, K, Eisenberger, T, Sekido, R, Kress, J, Treier, A-C, Klugmann, C, Klasen, C, Holter, NI, Riethmacher, D, Schütz, G, Cooney, AJ, Lovell-Badge, R, Treier, M. 2009. Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. Cell 139: 1130–42.

Umehara, F, Tate, G, Itoh, K, Yamaguchi, N, Douchi, T, Mitsuya, T, Osame, M. 2000. A novel mutation of desert hedgehog in a patient with 46,XY partial gonadal dysgenesis accompanied by minifascicular neuropathy. Am. J. Hum. Genet. 67: 1302–5.

Vaiman, D. 2003. Sexy transgenes: the impact of gene transfer and gene inactivation technologies on the understanding of mammalian sex determination. Transgenic Res. 12: 255–69.

Vainio, S, Heikkilä, M, Kispert, A, Chin, N, McMahon, AP. 1999. Female development in mammals is regulated by Wnt-4 signalling. Nature 397: 405–9.

Val, P, Lefrançois-Martinez, A-M, Veyssière, G, Martinez, A. 2003. SF-1 a key player in the development and differentiation of steroidogenic tissues. Nucl. Recept. 1: 8.

Veitia, R, Nunes, M, Brauner, R, Doco-Fenzy, M, Joanny-Flinois, O, Jaubert, F, Lortat-Jacob, S, Fellous, M, McElreavey, K. 1997. Deletions of distal 9p associated with 46,XY male to female sex reversal: definition of the breakpoints at 9p23.3-p24.1. Genomics 41: 271–4.

Vetro, A, Ciccone, R, Giorda, R, Patricelli, MG, Della Mina, E, Forlino, A, Zuffardi, O. 2011. XX males SRY negative: a confirmed cause of infertility. J. Med. Genet. 48: 710–2.

Vialard, F, Ottolenghi, C, Gonzales, M, Choiset, A, Girard, S, Siffroi, JP, McElreavey, K, Vibert-Guigue, C, Sebaoun, M, Joyé, N, Portnoï, MF, Jaubert, F, Fellous, M. 2002. Deletion of 9p associated with gonadal dysfunction in 46,XY but not in 46,XX human fetuses. J. Med. Genet. 39: 514–8.

Vidal, VP, Chaboissier, MC, de Rooij, DG, Schedl, A. 2001. Sox9 induces testis development in XX transgenic mice. Nat. Genet. 28: 216–7.

Viger, RS, Guittot, SM, Anttonen, M, Wilson, DB, Heikinheimo, M. 2008. Role of the GATA family of transcription factors in endocrine development, function, and disease. Mol. Endocrinol. 22: 781–98.

Viger, RS, Mertineit, C, Trasler, JM, Nemer, M. 1998. Transcription factor GATA-4 is expressed in a sexually dimorphic pattern during mouse gonadal development and is a potent activator of the Müllerian inhibiting substance promoter. Development 125: 2665–75.

Wagner, T, Wirth, J, Meyer, J, Zabel, B, Held, M, Zimmer, J, Pasantes, J, Bricarelli, FD, 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.

Wainwright, EN, Jorgensen, JS, Kim, Y, Truong, V, Bagheri-Fam, S, Davidson, T, Svingen, T, Fernandez-Valverde, SL, McClelland, KS, Taft, RJ, Harley, VR, Koopman, P, Wilhelm, D. 2013. SOX9 regulates microRNA miR-202-5p/3p expression during mouse testis differentiation. Biol. Reprod. 89: 34.

Warr, N, Bogani, D, Siggers, P, Brixey, R, Tateossian, H, Dopplapudi, A, Wells, S, Cheeseman, M, Xia, Y, Ostrer, H, Greenfield, A. 2011. Minor abnormalities of testis development in mice lacking the gene encoding the MAPK signalling component, MAP3K1. PLoS One 6: e19572.

Warr, N, Carre, G-A, Siggers, P, Faleato, JV, Brixey, R, Pope, M, Bogani, D, Childers, M, Wells, S, Scudamore, CL, Tedesco, M, del Barco Barrantes, I, Nebreda, AR, Trainor, PA, Greenfield, A. 2012. Gadd45γ and Map3k4 interactions regulate mouse testis determination via p38 MAPK-mediated control of Sry expression. Dev. Cell 23: 1020–31.

Wat, JJ, Wat, MJ. 2014. Sox7 in vascular development: review, insights and potential mechanisms. Int. J. Dev. Biol. 58: 1–8.

Wat, MJ, Beck, TF, Hernández-García, A, Yu, Z, Veenma, D, Garcia, M, Holder, AM, Wat, JJ, Chen, Y, Mohila, CA, Lally, KP, Dickinson, M, Tibboel, D, de Klein, A, Lee, B, Scott, DA. 2012.

Mouse model reveals the role of SOX7 in the development of congenital diaphragmatic hernia associated with recurrent deletions of 8p23.1. Hum. Mol. Genet. 21: 4115–25.

Wat, MJ, Shchelochkov, OA, Holder, AM, Breman, AM, Dagli, A, Bacino, C, Scaglia, F, Zori, RT, Cheung, SW, Scott, DA, Kang, S-HL. 2009. Chromosome 8p23.1 deletions as a cause of complex congenital heart defects and diaphragmatic hernia. Am. J. Med. Genet. A 149A: 1661–77.

Watanabe, K, Clarke, TR, Lane, AH, Wang, X, Donahoe, PK. 2000. Endogenous expression of Müllerian inhibiting substance in early postnatal rat sertoli cells requires multiple steroidogenic factor-1 and GATA-4-binding sites. Proc. Natl. Acad. Sci. U. S. A. 97: 1624–9.

Weber, A, Köhler, A, Hahn, A, Müller, U. 2014. 8p23.1 duplication syndrome: narrowing of critical interval to 1.80 Mbp. Mol. Cytogenet. 7: 94.

Werner, R, Merz, H, Birnbaum, W, Marshall, L, Schröder, T, Reiz, B, Kavran, JM, Bäumer, T, Capetian, P, Hiort, O. 2015. 46,XY Gonadal Dysgenesis due to a Homozygous Mutation in Desert Hedgehog (DHH) Identified by Exome Sequencing. J. Clin. Endocrinol. Metab. 100: E1022–9.

White, S, Ohnesorg, T, Notini, A, Roeszler, K, Hewitt, J, Daggag, H, Smith, C, Turbitt, E, Gustin, S, van den Bergen, J, Miles, D, Western, P, Arboleda, V, Schumacher, V, Gordon, L, Bell, K, Bengtsson, H, Speed, T, Hutson, J, Warne, G, Harley, V, Koopman, P, Vilain, E, Sinclair, A. 2011. Copy number variation in patients with disorders of sex development due to 46,XY gonadal dysgenesis. PLoS One 6: e17793.

Wilhelm, D, Englert, C. 2002. The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1. Genes Dev. 16: 1839–51.

Wilhelm, D, Hiramatsu, R, Mizusaki, H, Widjaja, L, Combes, AN, Kanai, Y, Koopman, P. 2007. SOX9 Regulates Prostaglandin D Synthase Gene Transcription in Vivo to Ensure Testis Development. J. Biol. Chem. 282: 10553–10560.

Wilhelm, D, Koopman, P. 2006. The makings of maleness: towards an integrated view of male sexual development. Nat. Rev. Genet. 7: 620–31.

Wilhelm, D, Martinson, F, Bradford, S, Wilson, MJ, Combes, AN, Beverdam, A, Bowles, J, Mizusaki, H, Koopman, P. 2005. Sertoli cell differentiation is induced both cell-autonomously and through prostaglandin signaling during mammalian sex determination. Dev. Biol. 287: 111–24.

Wilhelm, D, Washburn, LL, Truong, V, Fellous, M, Eicher, EM, Koopman, P. 2009. Antagonism of the testis- and ovary-determining pathways during ovotestis development in mice. Mech. Dev. 126: 324–36.

Wilson, MJ, Jeyasuria, P, Parker, KL, Koopman, P. 2005. The transcription factors steroidogenic factor-1 and SOX9 regulate expression of Vanin-1 during mouse testis development. J. Biol. Chem. 280: 5917–23.

Wolski, KM, Feig, C, Kirchhoff, C, Cameron, DF. 2007. Immortalized Sertoli cell lines sk11 and sk9 and binding of spermatids in vitro. Asian J. Androl. 9: 312–20.

Wu, Y, Li, Q, Chen, X-Z. 2007. Detecting protein-protein interactions by Far western blotting. Nat. Protoc. 2: 3278–84.

Xiao, B, Ji, X, Xing, Y, Chen, Y-W, Tao, J. 2013. A rare case of 46, XX SRY-negative male with approximately 74-kb duplication in a region upstream of SOX9. Eur. J. Med. Genet. 56: 695–8.

Yao, HH-C, Whoriskey, W, Capel, B. 2002. Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. Genes Dev. 16: 1433–40.

Yao, HHC, Matzuk, MM, Jorgez, CJ, Menke, DB, Page, DC, Swain, A, Capel, B. 2004. Follistatin operates downstream of Wnt4 in mammalian ovary organogenesis. Dev. Dyn. 230: 210–5.

Yomogida, K, Ohtani, H, Harigae, H, Ito, E, Nishimune, Y, Engel, JD, Yamamoto, M.

1994. Developmental stage- and spermatogenic cycle-specific expression of transcription factor GATA-1 in mouse Sertoli cells. Development 120: 1759–66.

Yoshimoto, S, Ikeda, N, Izutsu, Y, Shiba, T, Takamatsu, N, Ito, M. 2010. Opposite roles of DMRT1 and its W-linked paralogue, DM-W, in sexual dimorphism of Xenopus laevis: implications of a ZZ/ZW-type sex-determining system. Development 137: 2519–26.

Yu, RN, Achermann, JC, Ito, M, Jameson, JL. 1998. The Role of DAX-1 in Reproduction. Trends Endocrinol. Metab. 9: 169–75.

Index of Figures

Abbreviations

ALCs	Adult Leydig cells
AMH	Anti-Müllerian hormone
ATRX	Alpha thalassemia/mental retardation syndrome X-linked
BPES	Blepharophimosis/ptosis/epicanthus inversus syndrome
CAH	Congenital Adrenal Hypoplasia
Cbln4	Cerebellin 4
Cbx2	Chromobox homolog 2
CGD	Complete Gonadal Dysgenesis
CHD	Congenital Heart Disease
Co-IP	Co-Immunoprecipitation
DAX1	DSS-AHC critical region on the chromosome X
DHH	Desert Hedgehog
DMRT1	dsx-and mab3 related transcription factor 1
DSD	Disorders of Sex development
dsx	doublesex
E	Embryonic day
EMSA	Electrophoretic Mobility Shift Assay
Emx2	Empty spiracles homeobox 2
ESCs	Mouse embryonic stem cells
Esr α/β	Estrogen receptors α/β
FGFR2	FGF receptor 2
FLCs	Fetal Leydig cells
FOG2	Friend of GATA - 2
FOXL2	Forkhead box L2
Gadd45g	Growth arrest and DNA damage inducible 45 γ
GATA4	Gata binding protein 4
GATA6	Gata binding protein 6
GD	Gonadal dysgenesis
HHAT	Hedgehog acyl-transferase
HMG	High mobility group
ieSCs	Induced embryonic Sertoli like cells
INSL3	Insulin like-3
iPSCs	Induced pluripotent stem cells
Kdm3a	Lysine specific demethylase 3A
Lhx9	LIM homeobox 9
LOF	Loss of function
mab-3	Male abnormal-3
MAPK	Mitogen activated protein kinase
MEFs	Mouse embryonic fibroblasts
Mis	Müllerian inhibition substance
NGS	Next Generation Sequencing
NR0B1	Nuclear receptor subfamily 0, group B, gene 1
NR5A1/SF1	Nuclear receptor subfamily 5 group A, gene 1/ Steroidogenic factor 1
PCR	Polymerase chain reaction

PGCs	Primordial germ cells
PGD	Partial gonadal dysgenesis
PIS	Polled Intersex Syndrome
PMCs	Peritubular myoid cells
POI	Premature ovarian insufficiency
PTCH1	Patched receptor 1
REVSEX	Reversal Sex enhancer region
RSPO1	R-spondin family 1
SCF	Stem Cell Factor
Six1/4	SIX homeobox 1/4
SNPs	Single nucleotide polymorphism
SOX	Sry related HMG box
SOX10	Sry related HMG box – 10
SOX7	Sry related HMG box -7
SOX8	Sry related HMG box -8
SOX9	Sry related HMG box - 9
SRY	Sex determining region on Y chromosome
StAR	Steroidogenic Acute Regulatory protein
TAD	Transactivation domain
TESCO	Testis specific enhancer core region
TGF-β	Transforming growth factor-β
TOF	Tetralogy of Fallot
Vnn1	Vanin-1
WB	Western blot
WES	Whole exome sequencing
WNT4	Wingless-type MMTV Integration Site Family, Member 4
WT	Wild-Type
Wt1	Wilms' tumor 1
αSma	α-smooth actin