

REVIEW

Switching on sex: transcriptional regulation of the testis-determining gene *Sry*

Christian Larney, Timothy L. Bailey and Peter Koopman*

ABSTRACT

Mammalian sex determination hinges on the development of ovaries or testes, with testis fate being triggered by the expression of the transcription factor sex-determining region Y (*Sry*). Reduced or delayed *Sry* expression impairs testis development, highlighting the importance of its accurate spatiotemporal regulation and implying a potential role for *SRY* dysregulation in human intersex disorders. Several epigenetic modifiers, transcription factors and kinases are implicated in regulating *Sry* transcription, but it remains unclear whether or how this farrago of factors acts co-ordinately. Here we review our current understanding of *Sry* regulation and provide a model that assembles all known regulators into three modules, each converging on a single transcription factor that binds to the *Sry* promoter. We also discuss potential future avenues for discovering the *cis*-elements and *trans*-factors required for *Sry* regulation.

KEY WORDS: Disorders of sex development, Sex determination, *Sry*, Testis development, Transcriptional regulation

Introduction

Sex in mammals is usually determined by the presence or absence of *Sry*, a small, single-exon gene located in the male-specific region of the Y chromosome (Gubbay et al., 1990; Sinclair et al., 1990). The expression of *Sry* in the genital ridges typically results in their development into testes, whereas the absence or dysfunction of *Sry* leads to the development of ovaries (Fig. 1A). *Sry* is the only gene from the Y chromosome required for testis determination, as a 14 kb genomic DNA fragment containing *Sry* and no other genes is able to direct male development in XX transgenic mice (Koopman et al., 1991).

Sry encodes a transcription factor that acts by binding to and activating the testis-specific enhancer (TES) of the related gene *Sox9* (Sekido and Lovell-Badge, 2008). A network of gene activity downstream of *Sox9* then promotes male development while simultaneously impeding the gene network that drives ovarian development (reviewed by Warr and Greenfield, 2012) (Fig. 1B). These two alternative gene regulatory networks give the bipotential gonad its unique ability to differentiate into two morphologically and functionally distinct organs.

Even though *Sry* triggers a cascade that influences sex-specific development throughout the organism and has lifelong consequences, its expression is limited to only a few cells and, at least in mice, to a narrow time window (Fig. 1C,D). *Sry* is expressed exclusively within the supporting cells of the genital ridges, directing them towards the fate of a Sertoli cell (Albrecht and Eicher, 2001), which are the primary constituents of testis cords and which sequester primordial germ cells and facilitate spermatogenesis. Other testicular cell types do not

express *Sry* and adopt a male fate in response to molecular signals arising from Sertoli cells. *Sry* expression is detectable in mice from 10.5 days post-coitum (dpc) (Koopman et al., 1990; Hacker et al., 1995; Jeske et al., 1996) and is initially restricted to the central region of the gonad (Bullejos and Koopman, 2001). This region of *Sry* expression then rapidly expands to encompass the entire gonad at ~11.5 dpc (Fig. 1D) before being extinguished, first in the central region and then outwards towards the poles. Expression is undetectable by 12.5 dpc. This expression profile appears to place testis determination on a knife edge, with slight decreases in *Sry* expression levels, or delays of peak expression by as little as a few hours, leading to failure of the testis-determining pathway and malapropos development of either ovotestes or ovaries (Hiramatsu et al., 2009; Wilhelm et al., 2009).

Accurate transcriptional upregulation of *Sry* is, therefore, essential for triggering male development, and malfunction of this mechanism is likely to be responsible for at least some of the many unexplained cases of human XY disorders of sexual development (DSD). Many attempts have been made to understand how *Sry* expression is so precisely upregulated, but owing to a unique combination of circumstances, a full explanation has not been forthcoming (see Box 1). Even less is known about how *Sry* expression is extinguished (see Box 2). Despite these obstacles, the clear requirement for timely, localised expression of *Sry* in the genital ridges during sex determination demands that a robust regulatory system must exist. In this Review, we summarise our current understanding of the molecular mechanisms contributing to the regulation of *Sry* transcription. Although epigenetic modifications within the *Sry* locus have been identified and implicated in regulating *Sry* expression (see Box 3), we focus here on transcription factor pathways and propose a model in which more than a dozen different factors fall into three modules responsible for *Sry* upregulation. We also suggest potential avenues for closing the remaining gaps.

Transcription factors and proximal *cis*-elements that regulate *Sry*: an integrative model of *Sry* regulation

Biochemical analyses and/or studies in transgenic mice have implicated upwards of a dozen factors in regulating *Sry*. However, a coherent model to explain the organisation of these factors has been lacking. Our model proposes that three key transcription factors bind to the *Sry* locus, and that each is the focus of a regulatory module comprising various and overlapping combinations of other factors, which function as partner proteins, post-translational modifiers and upstream regulators of the primary factors (Fig. 2). Evidence for this model is discussed in the following sections.

Module 1: pathways converging on GATA4

The expression of *Sry* is known to be dependent on the presence of the transcription factor GATA4 (Tevosian et al., 2002). *Gata4* knockout induced at 10.5 dpc, but not at later times, results in sex reversal,

Institute for Molecular Bioscience, The University of Queensland, Brisbane QLD 4072, Australia.

*Author for correspondence (p.koopman@imb.uq.edu.au)

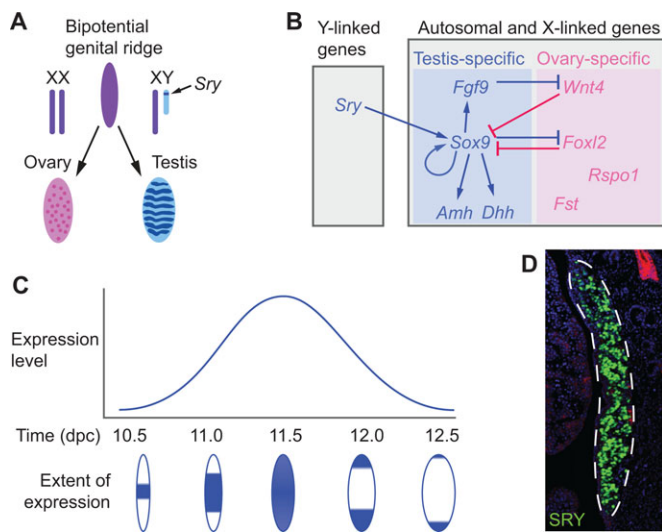


Fig. 1. Function and expression of *Sry*. (A) Testis fate is determined in most mammals by the presence of an *Sry*-containing Y chromosome. In individuals lacking a Y chromosome, genital ridges instead develop as ovaries. (B) *Sry* engages a sequence of regulatory steps that simultaneously promotes testis development and suppresses ovary development. (C) *Sry* is expressed in the supporting cells of the genital ridges between 10.5 and 12.5 dpc, with peak expression occurring at 11.5 dpc. Beneath is illustrated the centre-to-pole wave of *Sry* expression (blue) in the developing testis. (D) Immunostaining for SRY (green) showing expression in pre-Sertoli cells throughout the XY mouse genital ridge (outlined) at 11.5 dpc. Nuclei are counterstained with DAPI (blue); anti-E-cadherin (red) marks mesonephric tubules and (weakly) germ cells.

indicating that GATA4 is important for the earliest steps of male sex determination (Manuylov et al., 2011). Co-transfection experiments in HeLa cells indicate that GATA4 transactivates *Sry* promoter constructs from both mouse and pig, but not human (Miyamoto et al., 2008). In addition, PCR amplification of sequences obtained through chromatin immunoprecipitation (ChIP) has confirmed that GATA4 binds to DNA in two regions (Fig. 2) of the mouse *Sry* promoter (Gierl et al., 2012). Interaction between GATA4 and its co-factor FOG2 (also known as ZFPM2) is also required for *Sry* expression. Transgenic mice with a null *Fog2* allele, or those with a targeted mutation in *Gata4* that eliminates its interaction with FOG2, display reduced levels of *Sry* expression (Tevosian et al., 2002). *Fog2* expression, in turn, appears to be regulated by the homeobox transcription factors SIX1 and SIX4. Accordingly, *Six1/Six4* double-knockout mice express significantly lower levels of both FOG2 and SRY, and exhibit XY sex reversal (Fujimoto et al., 2013).

Recent studies have also shown that GATA4 binds to the *Sry* promoter only after the protein has been phosphorylated (Gierl et al., 2012). GATA4 phosphorylation is evidently carried out by p38 MAP kinases, as transgenic mice with a *p38β* (also known as *Mapk11*) null allele and a *p38α* (*Mapk14*) allele that is conditionally deleted in epiblast-derived tissues exhibit reduced *Sry* expression leading to sex reversal (Warr et al., 2012). Upstream components of the p38 MAPK cascade have also been implicated in regulating *Sry* expression; the gene encoding mitogen activated protein kinase kinase 4 (MAP3K4) was found in a forward genetic screen in mice to be mutated in XY gonadal sex reversal (Bogani et al., 2009). A premature stop codon in *Map3k4* in these mice resulted in a protein that lacked the kinase domain, rendering it non-functional. Quantitative RT-PCR showed a significant reduction in gonadwide *Sry* expression in these mutant mice, and immunofluorescence microscopy confirmed that this was partly due to reduced expression per cell.

Mutations in the related protein MAP3K1 also lead to human XY DSD (Pearlman et al., 2010), which evidently results from downregulation of testis-inducing genes, including *SRY*, and upregulation of ovary-inducing genes (Loke et al., 2013). However, homozygosity for a non-functional allele of *Map3k1* leads to only minor testis abnormalities in XY mice (Warr et al., 2011), leaving its role in regulating *Sry* unclear. It is possible that different members of the MAP3K family are responsible for regulating *Sry* in different species, a possibility that is given weight by the lack of *MAP3K4* mutations in human DSD. Alternatively, because human DSD phenotypes result from single nucleotide changes in *MAP3K1*, whereas complete loss of function has little

Box 1. Understanding *Sry* regulation: why has it been so difficult?

A number of idiosyncrasies of *Sry* have combined to render the question of its regulation impenetrable to date.

(1) ***Sry* is located on the Y chromosome.** This location excludes it from repair mechanisms available to autosomal or X-linked genes, so potential regulatory sequence motifs are easily degraded. Moreover, XX rather than XY samples are used in most genome sequencing projects (Hughes and Rozen, 2012), such that relatively few Y chromosome sequences are available, thus further impeding studies.

(2) **Mouse *Sry* is embedded in a large inverted repeat.** As such, any potential regulatory element will normally reside at a similar distance both 5' and 3' of the *Sry* ORF (Gubbay et al., 1992), complicating loss-of-function analyses.

(3) **Mouse *Sry* has more than one promoter.** In addition to its principal linear mRNA transcript, *Sry* is able to generate transcripts that arise from a second, more distal 5' promoter in the inverted repeat region. The RNA generated forms a stem-loop structure that can be spliced to form circular transcripts, which are not loaded onto polyribosomes and are mostly produced in germ cells of the adult testis, so a role in sex determination is unlikely (Capel et al., 1993a; Dolci et al., 1997). The question of how this second promoter is regulated has received little attention.

(4) ***Sry* is amplified in several species.** Rats have 11 non-identical copies of *Sry* (Prokop et al., 2013), at least six of which are expressed (Turner et al., 2007); rabbits (Gerald and Ferrand, 2006), pigs (Groenen et al., 2012) and dogs (Li et al., 2013) also have multiple *Sry* genes. It is not known which copy (or copies) carries the regulatory elements required for testis determination, confounding efforts to identify these sequences.

(5) ***Sry* expression in mice is limited in time and space.** Only a limited quantity of *Sry*-expressing tissue can be collected, making it difficult to study *Sry* regulation experimentally. Moreover, regulatory models must account not only for *Sry* activation and silencing, but also for the asynchronous, centre-to-pole wave of expression that occurs.

(6) ***Sry* regulation is intertwined with cell proliferation in the genital ridges.** A number of mutations in candidate *Sry*-regulating genes result in decreased *Sry* expression (Bradford et al., 2009; Fujimoto et al., 2013; Katoh-Fukui et al., 2012). In these situations, it is necessary to determine whether reduced expression is a result of lower *Sry* expression in each cell, and hence supportive of a defect in *Sry* regulation, or simply reflects a reduced population of *Sry*-expressing cells.

(7) ***Sry* expression is variable between species.** Compared with mice, *Sry* expression is more widespread in humans and wallabies and more durable in sheep and rabbits. This implies that the regulatory sequences are not necessarily conserved in evolution, further limiting the usefulness of comparative studies between species.

(8) **A suitable cell culture model is lacking.** Cell lines corresponding to endogenous *Sry*-expressing pre-Sertoli cells remain unavailable, and primary cell lines are quick to extinguish *Sry*. These limitations hamper biochemical analyses of transcription factor binding to putative regulatory sequences flanking *Sry*.

Box 2. Switching off *Sry* expression

In mice, *Sry* has a 'cameo role', with expression in the genital ridges disappearing only 2 days after the gene is activated. What switches *Sry* off and why? Because *Sry* is downregulated just as SOX9 expression reaches its peak, it is commonly assumed that SOX9 represses *Sry* transcription. Prolonged expression of *Sry* in mice conditionally null for *Sox9* supports this concept (Barrionuevo et al., 2006; Chaboissier et al., 2004), although the observation that *Sry* is also downregulated in the ovarian portion of ovotestes, where *Sox9* is not expressed, suggests that other factors are in play (Wilhelm et al., 2009). If SOX9 does contribute to *Sry* downregulation, this factor must be capable of acting as both a transcriptional activator and repressor in the same cell type. Nonetheless, two observations suggest that this overall phenomenon is not important for sex determination. First, the rapid downregulation of *Sry* is observed only in mice; in several other mammalian species, *Sry* expression has been shown to continue for much longer, even into adult life. Second, in transgenic mouse models in which *Sry* is upregulated correctly but fails to be downregulated on time or at all, male development can proceed normally (Kidokoro et al., 2005). For these reasons, the *cis*-sequences and *trans*-acting factors involved in *Sry* downregulation in mice have not been characterised.

effect in mice, it is possible that the human mutations cause a gain-of-function or even a dominant-negative effect that interferes with other steps in the regulatory pathway. Even more uncertainty surrounds the related MAP kinase kinase (MAP2K) family, the participating member(s) of which remain to be identified. Regardless, available evidence supports the MAPK cascade as a crucial regulator of GATA4 activity and, hence, *Sry* regulation (Fig. 2).

The activity of the MAPK cascade, in turn, is triggered by the stress-response protein GADD45G, which is able to bind to the N-terminus of MAP3K4, thereby disrupting an autoinhibitory state and instead activating the enzymatic activity of MAP3K4 (Miyake et al., 2007). *Gadd45g* is required for normal *Sry* expression (Gierl et al., 2012) and its loss leads to sex reversal (Johnen et al., 2013; Warr et al., 2012). Importantly, *Gadd45g* expression in mice proceeds in a centre-to-pole wave that foreshadows the expression of *Sry* by ~5 h (Warr et al., 2012). GADD45G also appears to link the insulin receptor family to the MAPK pathway upstream of *Sry*. The expression of both *Gadd45g* and *Sry* is perturbed in mice that are null for members of the insulin receptor tyrosine kinase family (Nef et al., 2003; Pitetti et al., 2013). For example, a peak in *Sry* expression in insulin receptor (*Insr*)/insulin-like growth factor 1 receptor (*Igflr*) double-knockout mice is not reached until 13.5 dpc, 2 days later than usual, and greatly reduced levels of *Gadd45g* are observed at 11.5 dpc, highlighting that GADD45G and insulin signalling may be linked. In support of this, insulin signalling has previously been shown to be capable of inducing another member of the Gadd45 family, *Gadd45b* (Bortoff et al., 2010). It is not known if expression of the insulin receptor family genes replicates the centre-to-pole wave exhibited by *Sry*.

In summary, the available evidence suggests that insulin receptors, GADD45G, the MAPK pathway, SIX1/4 and FOG2 are components of an integrated module that converges on GATA4 to regulate *Sry* transcription (Fig. 2). The wave-like expression of at least one of these components (*Gadd45g*) might explain the spatial profile of *Sry* expression during genital ridge development.

Module 2: pathways converging on NR5A1

A second group of factors implicated in *Sry* regulation has NR5A1 (nuclear receptor subfamily 5, group A, member 1; also known as

SF-1) at its centre (Fig. 2). NR5A1 is known to play a number of roles in sex development, including roles in establishing the genital ridges (Luo et al., 1994), regulating *Sox9* expression (Sekido and Lovell-Badge, 2008) and steroidogenesis (Morohashi et al., 1992). Accordingly, *NR5A1* mutations result in XY sex reversal in humans (Achermann et al., 1999), whereas homozygous null *Nr5a1* mutations in mice result in gonadal agenesis (Luo et al., 1994). Studies using an electrophoretic mobility shift assay (EMSA) indicate that NR5A1 binds to specific sites in the pig and human *Sry* promoters (Pilon et al., 2003; De Santa Barbara et al., 2001). Directed mutagenesis of these sites leads to reduced *Sry* expression in cell-based reporter assays, although verifying these results *in vivo* is difficult because the sites in pig and human are not found in the corresponding position in mouse. Instead, sequences resembling the short NR5A1 core binding motif are found at many positions in the mouse *Sry* 5' flanking sequence. It is assumed that NR5A1 binds somewhere in this region, but it is not clear where.

The expression of NR5A1, in turn, appears to be regulated by the Polycomb group protein CBX2. XY sex reversal occurs in humans carrying *CBX2* mutations (Biason-Lauber et al., 2009) and in mice lacking *Cbx2* function (Katoh-Fukui et al., 2012). In these mice, *Sry* is expressed at negligible levels and far fewer cells express *Nr5a1* compared with wild-type gonads, suggesting that, rather than directly regulating *Sry*, CBX2 might mediate its activities through the regulation of *Nr5a1* (Katoh-Fukui et al., 2012) (Fig. 2). Also, it is not clear whether it is the number of *Sry*-expressing cells or the level of *Sry* expression per cell (or indeed both) that is reduced in *Cbx2* knockout gonads (see Box 1). In addition, given that other genes with roles in gonad development, including *Gata4* and *Lhx9*, are also affected in *Cbx2* knockout mice, CBX2 might well have multiple roles in this system. It is also important to note that Polycomb repressive complex 1, of which CBX2 is a member, is known to have activating potential, for example of *Nr5a1* during adrenal development (Katoh-Fukui et al., 2005), despite more commonly being considered a transcriptional silencer.

A second factor likely to interact with *Nr5a1* to regulate *Sry* is CITED2 (Cbp/p300-interacting transactivator with Glu/Asp-rich carboxy-terminal domain 2). The expression of both *Nr5a1* and *Sry*

Box 3. Epigenetic regulation of *Sry*

Correct gene regulation generally depends on chromatin being made accessible through epigenetic modifications, both in the region of the promoter and at regulatory elements. For *Sry*, dimethylation of histone H3 at lysine nine (H3K9me2) (Barski et al., 2007), which is a repressive epigenetic mark, has been shown to affect its promoter in gonadal cells, and is partly removed by the histone demethylase JMJD1A (also known as KDM3A) (Kuroki et al., 2013). Accordingly, higher levels of H3K9me2 are present in the *Sry* locus of *Jmjd1a*-deficient mice at the crucial 11.5 dpc time point and these mice show both reduced levels of *Sry* expression and XY sex reversal. Permissive histone marks, including H3 lysine four trimethylation (H3K4me3) and H3 acetylation (H3ac) (Gierl et al., 2012), are also enriched in the *Sry* promoter at the same stage in gonad cells, but not in cells of the adjacent mesonephros, where *Sry* is not expressed. How JMJD1A and the other histone modifiers presumably required for these processes are recruited to the *Sry* promoter remains an open question. Activation of *Sry* in mice is also preceded by gonad-specific DNA hypomethylation of five CpG sites in the 400 bp immediately proximal to the TSS (Gierl et al., 2012; Nishino et al., 2004). These sites are largely hypermethylated again by 15.5 dpc (Nishino et al., 2004), suggesting an involvement in *Sry* temporal regulation. The mechanism underlying these events, as for histone modifications, remains unknown.

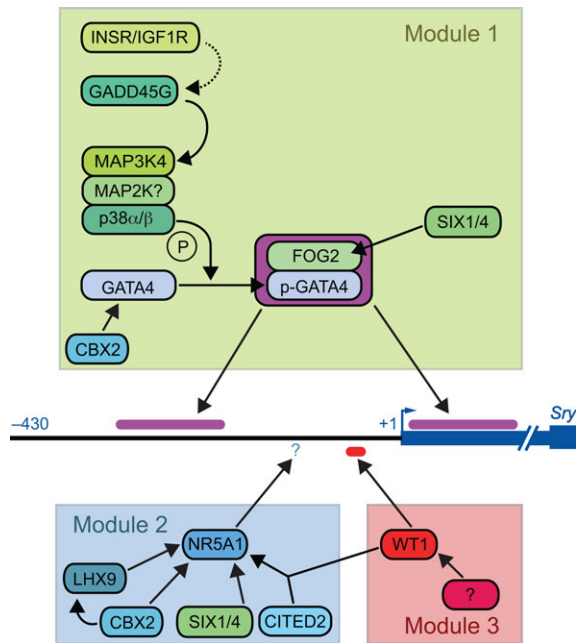


Fig. 2. A model of *Sry* regulation in mouse. The many factors implicated in regulating *Sry* can be partitioned into three modules. In Module 1, the insulin (INSR/IGF1R) pathway induces *Gadd45g* expression, which activates a MAP kinase cascade involving MAP3K4, an as yet unknown member of the MAP2K family, and p38 α/β , leading to phosphorylation (P) of GATA4. Simultaneously, SIX1 and SIX4 upregulate FOG2, which partners with GATA4. The resulting complex is able to bind to two specific regions of the *Sry* promoter. In Module 2, CITED2 cooperates with WT1 to upregulate NR5A1, a process that also involves transcription factors including CBX2, LHX9 and SIX1/4. NR5A1 then binds to an unknown position on the *Sry* locus where it contributes to *Sry* upregulation. Module 3 involves the transcription factor WT1, which is also able to bind directly to DNA, with *in vitro* evidence suggesting a specific binding site in mouse *Sry* between -56 and -47, near the TSS (+1, blue arrow). Possible regulators of WT1 remain unknown.

is reduced in *Cited2*^{-/-} mixed strain mice that contain a weakened *Sry* allele (*Sry*^{POS}), such that either ovary or ovotestis development can result (Buaas et al., 2009). Similar results are seen in mice that lack *Cited2* and harbour a heterozygous mutation in *Nr5a1*.

How CITED2 regulates *Nr5a1* is not clear, although it appears to involve interaction with Wilms tumor 1 protein (WT1) (Fig. 2). Both XX and XY mice lacking the -KTS isoform of WT1 [which lacks the additional lysine-threonine-serine (KTS) tripeptide between zinc fingers three and four that is present in the related +KTS isoform] exhibit streak gonads, probably owing to a failure to upregulate *Nr5a1* in conjunction with LHX9 (Birk et al., 2000; Hammes et al., 2001; Wilhelm and Englert, 2002). Moreover, testis development is compromised in mice doubly heterozygous for *Wt1* and *Nr5a1* (Correa et al., 2012). Gonadal defects are seen in XY *Cited2*^{-/-} mice (Combes et al., 2010) and are exacerbated by a heterozygous mutation of *Wt1*, leading to total and partial XY sex reversal, respectively (Buaas et al., 2009), suggesting that WT1 interacts with CITED2 to regulate *Nr5a1* expression. Hence, WT1 contributes to the function of Module 2 while also directly regulating *Sry* transcription, as discussed below.

Finally, SIX1 and SIX4, which, as discussed above, are likely regulators of *Fog2*, also play a role in activating *Nr5a1*. Both XX and XY *Six1/Six4* double-knockout mice have reduced *Nr5a1* expression and smaller gonads, in addition to reduced levels of FOG2 (Fujimoto et al., 2013). Again, it is not clear whether the number of *Nr5a1*⁺ cells or the level of *Nr5a1* and/or *Fog2*

expression per cell is affected. The molecular details of how SIX1/SIX4 might act in this system are currently unknown.

Module 3: gene regulation via WT1

The third major regulatory module is provided directly by WT1 (Fig. 2). *In vitro* cell co-transfection experiments in HeLa cells indicate that WT1 is able to transactivate *Sry* promoter constructs from mouse, pig and human (Miyamoto et al., 2008). Subsequent mutational analyses confirmed a requirement for direct DNA binding in this regulation in both mice and humans (Miyamoto et al., 2008; Hossain and Saunders, 2001). The importance of WT1 has also been shown *in vivo*, as XY mice lacking the +KTS isoform express *Sry* at only 25% of wild-type levels and fail to develop testes (Hammes et al., 2001). Reduced *Sry* expression is caused both by lower expression levels of *Sry* per cell and by a reduction in the number of *Sry*-expressing cells, which occurs as a result of reduced cell proliferation (Bradford et al., 2009). This reduced cell proliferation is not observed in similarly mutated XX embryos, implying that it occurs downstream of reduced *Sry* expression in addition to being part of the cause.

It has been speculated that CITED2 participates in the direct regulation of *Sry* by WT1, but compelling evidence remains lacking (Buaas et al., 2009). By contrast, co-transfection of *Sry* promoter constructs with both WT1 and GATA4 shows synergistic effects in mice, pigs and human (Miyamoto et al., 2008), suggesting that GATA4 and WT1 might physically interact to regulate *Sry*.

The search for distal cis-regulatory elements

Above, we have proposed that various regulators converge on three distinct factors – GATA4, NR5A1 and WT1 – to directly regulate the expression of *Sry*. Although binding sites that might allow both WT1 and GATA4 to regulate *Sry* have been identified (Gierl et al., 2012; Miyamoto et al., 2008), these sites are restricted to the proximal promoter, within just a few hundred base pairs of the transcription start site (TSS) (Fig. 2). The presumed binding site(s) for NR5A1 have not been identified, and might be proximal or more distal in position. In the following sections, we consider evidence from transgenic mice, chromosomal rearrangements, and comparative analyses of flanking sequences that suggests the presence of other cis-regulatory elements at more distal positions than those currently known.

Sry transgene constructs that induce XX sex reversal

Sry was initially confirmed as the testis-determining factor by making XX sex-reversed transgenic mice using a 14.6 kb genomic DNA fragment known as L741 (Fig. 3A) that contains the *Sry* gene, ~8 kb of 5' flanking sequence and ~3 kb of 3' flanking sequence (Koopman et al., 1991). As a result, the assumption that all regulatory sequences required for *Sry* function in sex determination reside within that fragment has become the cornerstone for attempts to understand *Sry* regulation. More specifically, it has been assumed that regulatory sequences will lie in the 5' region of the construct, with little consideration of possible regulatory sequences located 3'. This assumption is supported by the fact that the 3 kb of 3' flanking sequence in L741 duplicates sequences in the 5' flank (Gubbay et al., 1992). Subsequent studies showed that the same 8 kb of 5' regulatory sequence driving an enhanced green fluorescent protein (EGFP) reporter transgene (Fig. 3B) accurately recapitulates the endogenous *Sry* centre-to-pole expression profile described previously (Albrecht and Eicher, 2001). Although original reports suggested that the expression of this reporter transgene continued after 12.5 dpc, a more recent study found it to be sex-specifically extinguished in XY embryos at the correct time (Jameson et al., 2012). Together, these

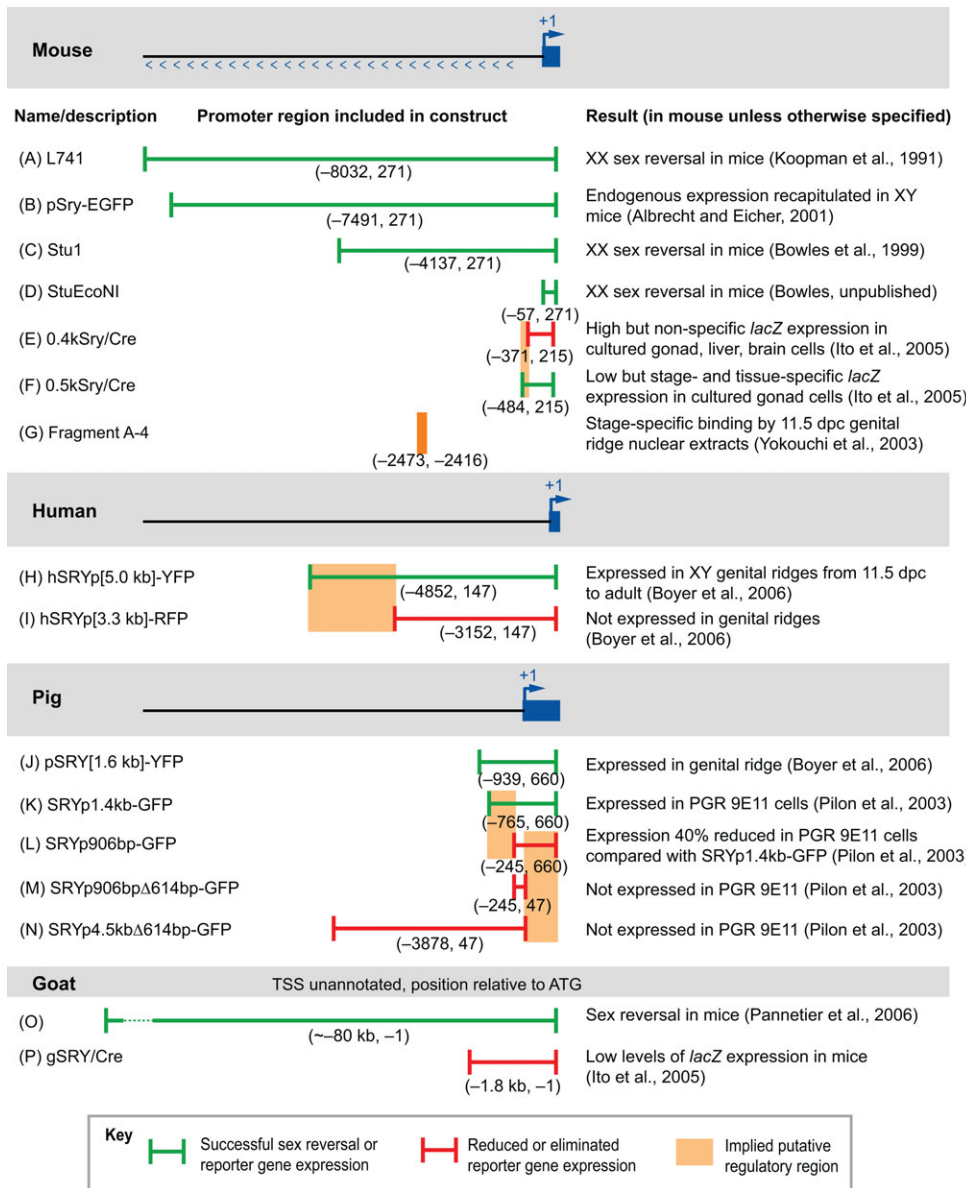


Fig. 3. Sry promoter constructs reveal putative Sry regulatory regions. Promoter constructs have been generated using the regulatory regions of mouse (A-F), human (H,I), pig (J-N) and goat (O,P) *Sry*. These constructs regulate the expression of either the endogenous *Sry* ORF or a fluorescent reporter protein, typically in transgenic mice or a transfected cell line. The TSS is illustrated for the first three of these species; the *Sry* TSS has not been characterised in goat. The region included in each construct is indicated in green or red and the locations of putative *cis*-regulatory regions suggested by these studies are marked (orange). Also included is the result of Yokouchi et al. (2003) (G), which does not involve a promoter construct but assays DNase hypersensitivity of the mouse *Sry* promoter region.

studies strongly suggest that the sequences required to correctly regulate mouse *Sry* expression are located in the 8 kb immediately 5' to the *Sry* ORF.

Shorter *Sry* transgenic constructs that use an endogenous ORF and nested deletions of 5' flanking sequence are also able to induce sex reversal in mice with reasonable efficiency (Fig. 3C,D) (Bowles et al., 1999; Koopman, 2002). Although this ability might be linked to the presence of multiple copies of the transgene in the genome of these animals, an alternative explanation for the efficacy of these constructs is that they contain the recognised WT1 binding site (Miyamoto et al., 2008) (Fig. 2). However, the pattern of expression that they generate has not been detailed.

Another putative regulatory region within the mouse *Sry* locus has been indicated by experiments using *Sry/Cre* fusion constructs to drive *lacZ* expression (Ito et al., 2005). Echoing the above results obtained using transgenes containing the endogenous *Sry* ORF, these fusion constructs were found to be expressed in the gonad regardless of the length of 5' sequence that they contain. However, whereas a construct with 400 bp of 5' sequence (Fig. 3E) was highly expressed in cultured gonadal cells at 11.5 and 13.5 dpc, as well as

in samples from liver and brain, a construct with 500 bp of 5' sequence (Fig. 3F) was exclusively expressed in 11.5 dpc gonadal cells, mirroring endogenous *Sry* expression. Although expression of the longer construct occurred at much lower levels than that of the shorter construct, the different stage- and tissue-specific patterns that they give rise to suggest that an important cell-specific regulatory element is present between positions -484 and -371 (Fig. 3, orange box between E and F). A function for this region might be restricted to the mouse, given that it is not conserved in any other species studied (Ito et al., 2005).

Also in mouse, a 58 bp region ~2.4 kb 5' to *Sry* was found to be bound by nuclear extracts from 11.5 dpc gonads when assayed by EMSA (Fig. 3G) (Yokouchi et al., 2003). However, extracts from both 12.5 dpc and 13.5 dpc gonads failed to bind to this region, suggesting stage-specific regulation during the sex-determining period.

The regulatory regions within human, pig and goat *Sry* loci have also been examined. For example, it was shown that human *SRY* regulatory regions remain functional in mice. When driven by 5 kb of human 5' sequence (Fig. 3H), a reporter transgene is expressed in

the murine gonad, but this expression is absent when using a short 3.3 kb construct (Fig. 3I) (Boyer et al., 2006), suggesting the presence of an important element in the -3.1 to -4.8 kb region (Fig. 3, orange box between H and I). Pig *Sry* regulatory regions are also functional in mice. Fluorescent reporters driven by either 1.6 kb (Fig. 3J) or 2.6 kb of pig 5' sequence were shown to be expressed in the genital ridge of mice (Boyer et al., 2006). Furthermore, pig 5' sequences of 4.5 kb and 1.4 kb (Fig. 3K) can drive robust reporter gene expression in porcine genital ridge (PGR) 9E11 cells (Pilon et al., 2003). Reporter expression was reduced by 40% when a shorter construct of 906 bp (leaving just 245 bp before the TSS) was used (Fig. 3L). In addition, removal of the 5' UTR from the reporter construct virtually eliminated expression, even if the remainder of the 4.5 kb construct was included (Fig. 3M,N). Elements that are essential for *Sry* expression are thus likely to be found in the 5' UTR between positions +47 and +660 (Fig. 3, orange box between L, M and N), while an additional element located somewhere between -245 and -765 (Fig. 3, orange box between K and L) appears to enhance the levels of transcription. Finally, it has been shown that goat *Sry* also has some function in mice, as XX transgenic mice with a 90 kb insert containing goat *Sry*, some 80 kb of 5' sequence and 6 kb of 3' sequence developed testes in two out of six cases (Fig. 3O) (Pannetier et al., 2006). However, 1.8 kb of a goat *Sry* promoter was unable to drive a *Cre/loxP/lacZ* reporter when transfected into cells from 11.5 dpc genital ridges (Fig. 3P) (Ito et al., 2005), suggesting the presence of regulatory sequences in the goat beyond this region.

In summary, these results argue for the presence of regulatory regions up to several kilobases from the TSS (Fig. 3, orange bands). However, as discussed below, large-scale rearrangements of the Y chromosome involving deletions and inversions suggest either that even longer range regulation is occurring or that *Sry* is susceptible to position effects.

Deletions and rearrangements that result in XY sex reversal

Deletions in the regions flanking human *SRY* (Fig. 4) are associated with sex reversal but how they mediate this reversal is poorly understood. A large deletion, from ~ 1.8 kb 5' of *SRY* to between 23 kb and 50 kb 5' of the gene (McElreavey et al., 1992), has been noted in a patient (NV) with streak gonads and external female appearance (Fig. 4). This deletion potentially removes almost the entire sequence between *SRY* and the gene encoding RPS4Y1 (ribosomal protein S4, Y-linked 1), which is the nearest neighbour to *SRY* in the 5' direction. Another patient (SC) with dysgenetic testes and external female appearance was reported to have a 2.5-7 kb deletion (Fig. 4) beginning between 2 and 3 kb 3' of *SRY* and

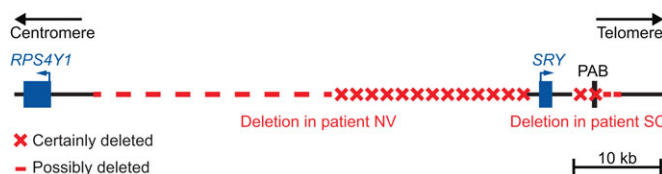


Fig. 4. Deletions in human *SRY* flanking regions associated with XY disorders of sexual development (DSD). A large deletion 5' of *SRY* has been reported in a patient (NV) and leads to streak gonads and external female genitalia (McElreavey et al., 1992). This deletion may remove almost the entire region between *SRY* and *RPS4Y1*. A second deletion, 3' of *SRY*, has also been reported (patient SC) and is associated with dysgenetic testes and female external genitalia (McElreavey et al., 1996). This deletion crosses the pseudoautosomal boundary (PAB) and extends some way into the pseudoautosomal region.

extending into the pseudoautosomal region (McElreavey et al., 1996). Although the lack of corroborating patients makes proving a causal link between these mutations and the associated phenotypes difficult, it is known that both unaffected fathers lack similar mutations, which are thus *de novo* in the affected individuals.

Large deletions of the Y chromosome, estimated to be 3-4 Mb in size (Mahadevaiah et al., 1998), also cause sex reversal in mice, but in this case the deletions are further away from the gene and are thought to act through a generalised position effect, bringing *Sry* closer to a heterochromatic region that prevents its expression (Capel et al., 1993b; Laval et al., 1995). They are, therefore, not instructive in locating potentially important regulatory intervals.

Two recent reports discuss XY sex reversal resulting from pericentric inversion of the Y chromosome (Fig. 5). In the first case, the breakpoint on the short arm was located ~ 93 kb 3' of *SRY*, while the long arm breakpoint was located in the heterochromatin of the q11.2 cytoband (Gimelli et al., 2006) (Fig. 5A). This inversion places heterochromatin relatively close to the 3' end of *SRY*, apparently silencing the gene completely. Accordingly, CpG sites located in the promoter of *SRY* in this patient were more highly methylated than the same sites in the patient's unaffected father, suggesting that, as in the mouse, DNA methylation might be involved in *SRY* regulation in humans. In the second case, the inversion breakpoints were located in p11.2 and q11.2 (Fig. 5B). The short arm breakpoint was located some 350 kb 5' of *SRY*, distinguishing this case from the previous one and suggesting some form of long-range 5' regulation. RT-PCR of the patient's streak gonads revealed *SRY* expression (Mitsuhashi et al., 2010), suggesting a more subtle misregulation than in the first case. Although both of these cases illustrate the sensitivity of *SRY* expression to its chromosomal position, it remains unclear how they might further inform the search for regulating factors and *cis*-regulatory regions.

Finally, it should be noted that a number of single nucleotide polymorphisms (SNPs) have been associated with XY sex reversal phenotypes (Kwok et al., 1996; Poulat et al., 1998; Veitia et al., 1997a). However, ambiguous results and small sample sizes make it difficult to assess whether these mutations are causative of sex reversal or simply a coincidence in otherwise affected individuals.

Insights from bioinformatic comparative analyses

A common strategy used to identify potential regulatory sites is comparative analysis of a gene's flanking sequences. Unfortunately, the peculiar difficulties associated with *Sry* (Box 1) mean that these studies have been limited by the availability of only short sequences from the 5' flank, with little consideration given to the possibility of regulatory sequences located in the 3' region, even in non-mouse species in which the two sequences are independent. In addition, the use of different sets of species in each study – only four species appear in more than two studies – makes comparisons difficult. Finally, the use of pairwise, rather than multiple, alignments in a number of studies makes it difficult to assess just how well conserved the sequences are across a broad spectrum of species. Nonetheless, a number of comparative studies have revealed conserved sequences adjacent to the *Sry* coding sequence (Fig. 6, Table 1)

The region around the TSS has been found in all comparative studies to be somewhat conserved (Hacker et al., 1995; Margarit et al., 1998; Pilon et al., 2003; Ross et al., 2008; Veitia et al., 1997a). This region includes putative binding sites for the transcription factor SP1 in all species studied, albeit with varying degrees of conservation. The function of two of these SP1 sites has been verified in the human NT2-D1 cell line (Desclozeaux et al., 1998). Although the role of SP1 as a ubiquitous transcription factor

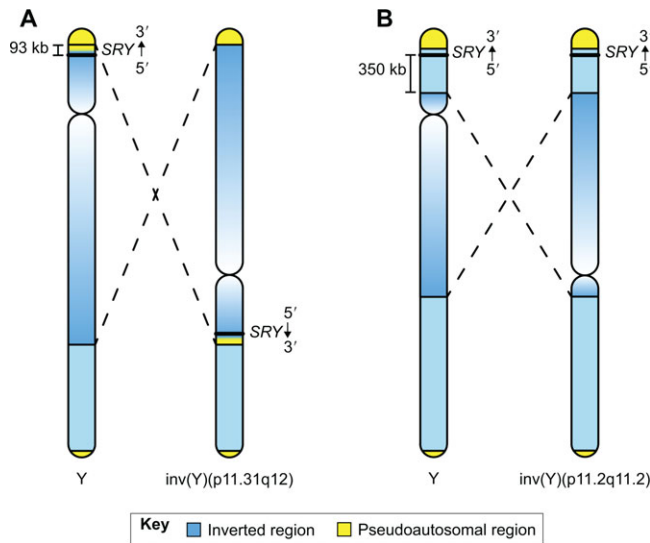


Fig. 5. Pericentric inversions of the Y chromosome associated with XY DSD. (A) In one reported case of XY DSD, *SRY* is translocated via a pericentric inversion (indicated by dashed lines) to the long arm of the Y chromosome. (B) In a second case, which is also associated with XY DSD, a different pericentric inversion leaves *SRY* in place but modifies the sequence some 350 kb 5' of *SRY* (Mitsuhashi et al., 2010). Chromosomes are not shown to scale.

rules it out as the primary driver of the stage- and tissue-specific regulation of *Sry*, the presence of these sites suggests that the factors that control these aspects need to be competent to interact with SP1. Motifs for additional transcription factor families, including GATA and Oct factors, are also found in this region (Margarit et al., 1998), although the function of these motifs has not been reported.

In this same region around the TSS, Hacker et al. (1995) identified a 146 bp region that is conserved between humans (−58 to +88) and mice (−132 to +14; Fig. 6). Independently, Veitia et al. (1997b) aligned sequences from human, four other primate species, cow and pig and found a short conserved interval, which they

labelled Motif C, that is located between −57 and −43 in human (Fig. 6). The fact that the regions identified in these two studies overlap suggests strong conservation of this region across a range of species. The only species found to lack Motif C to date has been sheep (Margarit et al., 1998), and a comparison of this motif with known transcription factor binding motifs suggests that it is competent to bind aryl hydrocarbon receptor nuclear translocator (ARNT), or possibly SOX9.

Veitia et al. (1997b) also identified two more distally located conserved motifs, which they termed Motif A and Motif B (Fig. 6). In human, these motifs are located at −373 to −362 and −349 to −340, respectively. Whereas Motif A is considered competent to bind HinfA or Oct factors, no transcription factor motifs could be associated with Motif B, and neither motif has been identified in other comparative studies.

Moving further away from *Sry*, the next significantly conserved region is a cluster of five transcription factor consensus binding motifs located within a region of just 90 bp (Ross et al., 2008) (Fig. 6). Located between positions −907 and −817 in human, this cluster is conserved between cow, human, goat and pig and is particularly remarkable because the five motifs are so densely clustered yet only nine putative transcription factor binding sites were found to be conserved between all four species elsewhere in the entire 5 kb of sequence that was analysed (Ross et al., 2008). The clustering of transcription factor motifs in this manner is often used to predict functional *cis*-regulatory elements (Spitz and Furlong, 2012). Of the five clustered sites, two could potentially bind caudal type homeobox (CDX) factors. Another is competent to bind a member of the zinc fingers, C₂H₂-type with BTB/POZ domain (ZBTB) family, and a further motif is predicted to bind NF-κB. The final conserved element is a serum-response element capable of binding serum response factor (SRF). However, there are no reports to date of any of these factors being involved in *Sry* regulation, and the functional significance of this cluster remains to be investigated.

Further 5' of *Sry*, Hacker et al. (1995) recognised a second, ~100 bp region of conservation between human and mouse sequences (Fig. 6). This region lies within the 5' arm of the

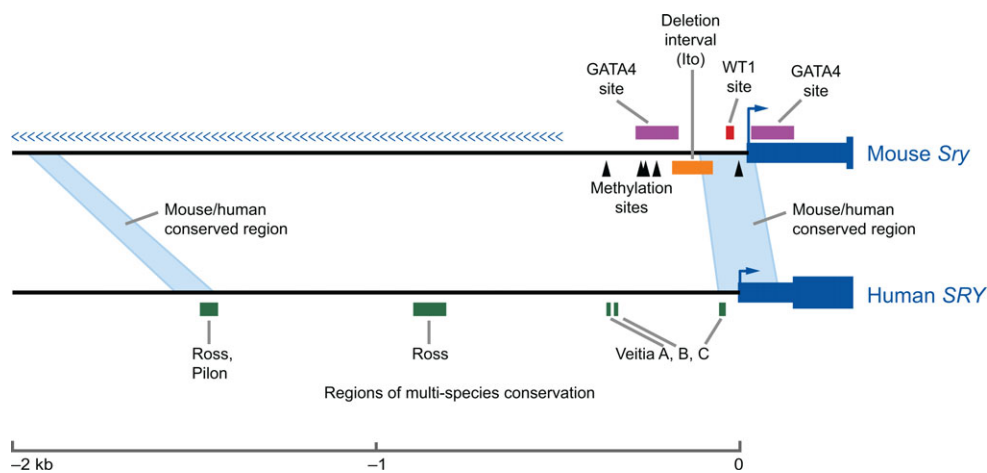


Fig. 6. Summary of the regulatory regions and sequence motifs that lie 5' of *Sry*. *Sry* transcripts in mouse (top) and human (bottom) are shown in dark blue. Regions implicated in the binding of GATA4 (purple) and WT1 (red) are illustrated, along with five CpG sites that exhibit gonad-specific demethylation (black arrowheads) and a regulatory region inferred by Ito et al. (2005) (orange; see Fig. 3E,F). Two regions of moderate homology with human are present in the proximal 2 kb of the mouse *Sry* 5' region (Hacker et al., 1995) (light blue shading). The inverted repeat sequence flanking mouse *Sry* is indicated by chevrons. Additionally, a number of short motifs within the *Sry* 5' region are conserved between human and various other species [green rectangles; the relevant studies are indicated (Ross et al., 2008; Pilon et al., 2003; Veitia et al., 1997b)]. The most proximal (Veitia C) and distal (Ross, Pilon) of these regions each overlap one of the two regions conserved between human and mouse.

Table 1. Species, sequence lengths and alignment methods used in comparative studies of *Sry* regulatory regions

Study	Hacker et al. (1995)	Veitia et al. (1997b)	Margarit et al. (1998)	Pilon et al. (2003)	Ross et al. (2008)
Length of sequences	Several (~8) kb	205 bp to 4.5 kb	346 bp to 4.5 kb	4.5 kb	4.6-17 kb
Alignment method	Pairwise	Multiple (Clustal W)	Multiple (Clustal)	Pairwise	Pairwise (AVID, LAGAN)
Human	•	•	•	•	•
Pig		•	•	•	•
Mouse	•		•	•	•
Cow		•	•		•
Chimpanzee		•	•		
Sheep			•	•	
Goat					•
Rhesus macaque		•			
Olive baboon		•			
Green monkey		•			
Gorilla			•		
Gazelle			•		
Rat			•		
Guinea pig			•		

mouse inverted repeat, between –1992 and –1897, and corresponds to the human sequence between –1558 and –1453. Independently, Ross et al. (2008) found that positions –1502 to –1452 of the human sequence are conserved with the cow sequence, and roughly the same portion of the human sequence has been found to share homology with the pig sequence (Pilon et al., 2003). Thus, a number of studies point to a region that is somewhat conserved between human, cow, pig and mouse (Fig. 6, leftmost green rectangle), although the functionality of this region, as for other potentially conserved regions, remains untested.

Conclusions and future directions

The various approaches described above, aimed at identifying and locating sequence motifs and factors important for *Sry* regulation, have yielded only modest progress to date. Genetic studies have identified three pivotal transcription factors – GATA4, WT1 and NR5A1 – that play important roles, each assisted by a suite of accessory factors arranged into proposed modules. For each module, it is clear that additional components might come to light through further studies, and suggested roles for the various components do not preclude the possibility of other roles for them in *Sry* regulation.

In addition, the specific binding sites occupied by these three factors, and the functionality of those sites, have yet to be confirmed with *in vivo* evidence. The evidence for DNA binding is, instead, restricted to regions of 100 bp or more (GATA4; Fig. 2), or else is based only on *in vitro* evidence (WT1; Fig. 2). Alignment-based comparative analysis has also only been able to provide hints of conserved regions that might indicate likely regulatory *cis*-elements. As a result, it is not known whether GATA4, WT1 and NR5A1 each bind single critical sites, nor whether the binding sites for these factors are clustered or dispersed. Furthermore, the putative regulatory sites for GATA4 and WT1 are located adjacent to the TSS, so all that is currently understood about *Sry* regulation is restricted to the proximal promoter region.

The existence of further distal regulatory elements seems likely from transgenesis and DNA rearrangement studies, even though the location of these elements remains unknown. Findings from alignment-based comparative analyses suggest that potential regulatory intervals might not be conserved between species, or at least might not occupy a similar position in the species being compared. Particularly problematic, given the role of the mouse as a

primary model organism, are the different expression profiles of *Sry* and the high level of sequence variability between mouse and human. Furthermore, the presence of large inverted repeats surrounding the mouse *Sry* locus suggests that mouse *Sry* might have been transposed from a different position on the chromosome at some point in the past. Taken together, these observations suggest that there are some differences between *Sry* regulation in humans and mice, even though the primary function of the gene has clearly been conserved. Such differences may limit the usefulness of the mouse model as a means of identifying regulatory motifs that can be examined in detail for mutations in human XY gonadal dysgenesis cases.

Future studies of *Sry* regulation must both identify and functionally confirm putative regulatory sites. The comparative analysis of genomic DNA sequences from different species is likely to yield insights when longer *Sry* flanking sequences from a greater range of species become available. Such analyses would also benefit from the availability of *Sry* flanking sequences in a broader variety of mouse strains than the B6 sample currently used as the mouse reference genome. The sequencing of Y chromosomes from a variety of mouse strains is therefore a clear priority.

In addition to identifying regulatory sequences solely through analysis of genomic sequences, it is now also possible to directly assay cell-specific regulatory activity. Perhaps the most widely utilised of these methods is ChIP followed by sequencing (ChIP-seq) for histone modifications associated with active regulatory regions, in particular histone H3 monomethylation at lysine 4 (H3K4me1) and acetylation at lysine 27 (H3K27ac) (Creyghton et al., 2010). Alternatively, the more recent finding that regulatory regions can be transcribed, producing so-called enhancer-derived RNAs (eRNAs) (Andersson et al., 2014; Hah et al., 2013; Melgar et al., 2011), shows that some gene expression assays, in particular global run-on sequencing (GRO-seq) (Core et al., 2008) and cap analysis of gene expression (CAGE) (Kodzius et al., 2006), can also predict the locations of *cis*-regulatory elements. Finally, promoter-enhancer interactions can be captured *in situ* through the use of chromosome conformation capture methods such as Hi-C (Lieberman-Aiden et al., 2009). Whereas ChIP-seq experiments continue to require in the order of 10⁴ cells (Adli and Bernstein, 2011), presenting a formidable target when working with embryonic gonads, protocols exist to perform CAGE (Islam et al., 2012) and Hi-C (Nagano et al., 2013) on single cells. Not only does this scale make these strategies more amenable to the small samples

available from nascent organs, but, if applied to groups of cells from different parts of the developing gonad, it also provides the potential for understanding intra-tissue transcriptional dynamics, such as the wave pattern of *Sry* expression.

Whatever methods are used to identify putative regulatory regions, verifying the function of these regions remains a challenge. The availability of genome editing techniques, such as transcription activator-like effector nuclease (TALEN) (Ding et al., 2013) or clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR associated (Cas) (Wang et al., 2013a) technology, is likely to speed this effort. The utility of these techniques for manipulating the Y chromosome has already been proven (Kato et al., 2013; Wang et al., 2013b). Ultimately, identifying which *cis*-elements and *trans*-factors contribute to the correct regulation of *Sry* might have both immediate medical utility for those living with variations or disorders of sex development and an enduring relevance for basic biology.

Acknowledgements

We thank Dagmar Wilhelm for the image used in Fig. 1D and the anonymous referees for valuable suggestions.

Competing interests

The authors declare no competing financial interests

Funding

This work was supported by research grants from the Australian Research Council (ARC), the National Health and Medical Research Council of Australia (NHMRC), and the US National Institutes of Health. C.L. is the recipient of an Australian Postgraduate Award and a University of Queensland Scholarship. P.K. is a Senior Principal Research Fellow of the NHMRC. Deposited in PMC for release after 12 months.

References

- Achermann, J. C., Ito, M., Ito, M., Hindmarsh, P. C. and Jameson, J. L. (1999). A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal and adrenal failure in humans. *Nat. Genet.* **22**, 125-126.
- Adli, M. and Bernstein, B. E. (2011). Whole-genome chromatin profiling from limited numbers of cells using nano-CHIP-seq. *Nat. Protoc.* **6**, 1656-1668.
- Albrecht, K. H. and Eicher, E. M. (2001). Evidence that *sry* is expressed in pre-sertoli cells and sertoli and granulosa cells have a common precursor. *Dev. Biol.* **240**, 92-107.
- Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T. et al. (2014). An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455-461.
- Barriounevo, F., Bagheri-Fam, S., Klattig, J., Kist, R., Taketo, M. M., Englert, C. and Scherer, G. (2006). Homozygous inactivation of *Sox9* causes complete XY sex reversal in mice. *Biol. Reprod.* **74**, 195-201.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837.
- Biason-Lauber, A., Konrad, D., Meyer, M., deBeaufort, C. and Schoenle, E. J. (2009). Ovaries and female phenotype in a girl with 46,XY karyotype and mutations in the *CBX2* gene. *Am. J. Hum. Genet.* **84**, 658-663.
- Birk, O. S., Casiano, D. E., Wassif, C. A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J. A., Parker, K. L. et al. (2000). The LIM homeobox gene *Lhx9* is essential for mouse gonad formation. *Nature* **403**, 909-913.
- Bogani, D., Siggers, P., Brixey, R., Warr, N., Beddow, S., Edwards, J., Williams, D., Wilhelm, D., Koopman, P., Flavell, R. A. et al. (2009). Loss of mitogen-activated protein kinase kinase 4 (*MAP3K4*) reveals a requirement for MAPK signalling in mouse sex determination. *PLoS Biol.* **7**, e1000196.
- Bortoff, K. D., Keeton, A. B., Franklin, J. L. and Messing, J. L. (2010). Anti-inflammatory action of insulin via induction of *Gadd45-β* transcription by the mTOR signaling pathway. *Hepatic Med. Evid. Res.* **2001**, 79-85.
- Bowles, J., Cooper, L., Berkman, J. and Koopman, P. (1999). *Sry* requires a CAG repeat domain for male sex determination in *Mus musculus*. *Nat. Genet.* **22**, 405-408.
- Boyer, A., Pilon, N., Raiwet, D. L., Lussier, J. G. and Silversides, D. W. (2006). Human and pig *SRY* 5' flanking sequences can direct reporter transgene expression to the genital ridge and to migrating neural crest cells. *Dev. Dyn.* **235**, 623-632.
- Bradford, S. T., Wilhelm, D., Bandiera, R., Vidal, V., Schedl, A. and Koopman, P. (2009). A cell-autonomous role for *WT1* in regulating *sry* in vivo. *Hum. Mol. Genet.* **18**, 3429-3438.
- Buaas, F. W., Val, P. and Swain, A. (2009). The transcription co-factor *CITED2* functions during sex determination and early gonad development. *Hum. Mol. Genet.* **18**, 2989-3001.
- Bullejos, M. and Koopman, P. (2001). Spatially dynamic expression of *Sry* in mouse genital ridges. *Dev. Dyn.* **221**, 201-205.
- Capel, B., Swain, A., Nicolis, S., Hacker, A., Walter, M., Koopman, P., Goodfellow, P. and Lovell-Badge, R. (1993a). Circular transcripts of the testis-determining gene *Sry* in adult mouse testis. *Cell* **73**, 1019-1030.
- Capel, B., Rasberry, C., Dyson, J., Bishop, C. E., Simpson, E., Vivian, N., Lovell-Badge, R., Rastan, S. and Cattanach, B. M. (1993b). Deletion of Y chromosome sequences located outside the testis determining region can cause XY female sex reversal. *Nat. Genet.* **5**, 301-307.
- Chaboissier, M.-C., Kobayashi, A., Vidal, V. I. P., Lützkendorf, S., van de Kant, H. J. G., Wegner, M., de Rooij, D. G., Behringer, R. R. and Schedl, A. (2004). Functional analysis of *Sox8* and *Sox9* during sex determination in the mouse. *Development* **131**, 1891-1901.
- Combes, A. N., Spiller, C. M., Harley, V. R., Sinclair, A. H., Dunwoodie, S. L., Wilhelm, D. and Koopman, P. (2010). Gonadal defects in *Cited2*-mutant mice indicate a role for *SF1* in both testis and ovary differentiation. *Int. J. Dev. Biol.* **54**, 683-689.
- Core, L. J., Waterfall, J. J. and Lis, J. T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845-1848.
- Correa, S. M., Washburn, L. L., Kahlon, R. S., Musson, M. C., Bouma, G. J., Eicher, E. M. and Albrecht, K. H. (2012). Sex reversal in *C57BL/6J* XY mice caused by increased expression of ovarian genes and insufficient activation of the testis determining pathway. *PLoS Genet.* **8**, e1002569.
- Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A. et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21931-21936.
- de Santa Barbara, P., Méjean, C., Moniot, B., Malclès, M.-H., Berta, P. and Boizet-Bonhoure, B. (2001). Steroidogenic factor-1 contributes to the cyclic-adenosine monophosphate down-regulation of human *SRY* gene expression. *Biol. Reprod.* **64**, 775-783.
- Desclozeaux, M., Poulat, F., de Santa Barbara, P., Soullier, S., Jay, P., Berta, P. and Boizet-Bonhoure, B. (1998). Characterization of two *Sp1* binding sites of the human sex determining *SRY* promoter. *Biochim. Biophys. Acta* **1397**, 247-252.
- Ding, Q., Lee, Y.-K., Schaefer, E. A. K., Peters, D. T., Veres, A., Kim, K., Kuperwasser, N., Motola, D. L., Meissner, T. B., Hendriks, W. T. et al. (2013). A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* **12**, 238-251.
- Dolci, S., Grimaldi, P., Geremia, R., Pesce, M. and Rossi, P. (1997). Identification of a promoter region generating *Sry* circular transcripts both in germ cells from male adult mice and in male mouse embryonal gonads. *Biol. Reprod.* **57**, 1128-1135.
- Fujimoto, Y., Tanaka, S. S., Yamaguchi, Y. L., Kobayashi, H., Kuroki, S., Tachibana, M., Shinomura, M., Kanai, Y., Morohashi, K.-i., Kawakami, K. et al. (2013). Homeoproteins *Six1* and *Six4* regulate male sex determination and mouse gonadal development. *Dev. Cell* **26**, 416-430.
- Geraldes, A. and Ferrand, N. (2006). A 7-bp insertion in the 3' untranslated region suggests the duplication and concerted evolution of the rabbit *SRY* gene. *Genet. Sel. Evol.* **38**, 313.
- Gierl, M. S., Gruhn, W. H., von Seggern, A., Maltry, N. and Niehrs, C. (2012). *GADD45G* functions in male sex determination by promoting p38 signaling and *sry* expression. *Dev. Cell* **23**, 1032-1042.
- Gimelli, G., Giorda, R., Beri, S., Gimelli, S. and Zuffardi, O. (2006). A 46,X,inv(Y) young woman with gonadal dysgenesis and gonadoblastoma: cytogenetics, molecular, and methylation studies. *Am. J. Med. Genet. A* **140A**, 40-45.
- Groenen, M. A. M., Archibald, A. L., Uenishi, H., Tuggle, C. K., Takeuchi, Y., Rothschild, M. F., Rogel-Gaillard, C., Park, C., Milan, D., Megens, H.-J. et al. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **491**, 393-398.
- Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* **346**, 245-250.
- Gubbay, J., Vivian, N., Economou, A., Jackson, D., Goodfellow, P. and Lovell-Badge, R. (1992). Inverted repeat structure of the *sry* locus in mice. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7953-7957.
- Hacker, A., Capel, B., Goodfellow, P. and Lovell-Badge, R. (1995). Expression of *Sry*, the mouse sex determining gene. *Development* **121**, 1603-1614.
- Hah, N., Murakami, S., Nagari, A., Danko, C. G. and Kraus, W. L. (2013). Enhancer transcripts mark active estrogen receptor binding sites. *Genome Res.* **23**, 1210-1223.
- Hammes, A., Guo, J.-K., Lutsch, G., Leheste, J.-R., Landrock, D., Ziegler, U., Gubler, M.-C. and Schedl, A. (2001). Two splice variants of the *wilms'* tumor 1

- gene have distinct functions during sex determination and nephron formation. *Cell* **106**, 319-329.
- Hiramatsu, R., Matoba, S., Kanai-Azuma, M., Tsunekawa, N., Katoh-Fukui, Y., Kurohmaru, M., Morohashi, K.-i., Wilhelm, D., Koopman, P. and Kanai, Y. (2009). A critical time window of Sry action in gonadal sex determination in mice. *Development* **136**, 129-138.
- Hossain, A. and Saunders, G. F. (2001). The human sex-determining gene SRY is a direct target of WT1. *J. Biol. Chem.* **276**, 16817-16823.
- Hughes, J. F. and Rozen, S. (2012). Genomics and genetics of human and primate Y chromosomes. *Annu. Rev. Genomics Hum. Genet.* **13**, 83-108.
- Islam, S., Kjällquist, U., Moliner, A., Zajac, P., Fan, J.-B., Lönnberg, P. and Linnarsson, S. (2012). Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat. Protoc.* **7**, 813-828.
- Ito, M., Yokouchi, K., Naito, K., Endo, H., Hakamata, Y., Miyazaki, J.-I. and Tojo, H. (2005). Detection of elements responsible for stage- and tissue-specific expression of mouse Sry using an in vitro Cre/loxP system. *Biochem. Biophys. Res. Commun.* **337**, 264-270.
- Jameson, S. A., Natarajan, A., Cool, J., DeFalco, T., Maatouk, D. M., Mork, L., Munger, S. C. and Capel, B. (2012). Temporal transcriptional profiling of somatic and germ cells reveals lineage priming of sexual fate in the fetal mouse gonad. *PLoS Genet.* **8**, e1002575.
- Jeske, Y. W. A., Mishina, Y., Cohen, D. R., Behringer, R. R. and Koopman, P. (1996). Analysis of the role of Amh and Fra1 in the Sry regulatory pathway. *Mol. Reprod. Dev.* **44**, 153-158.
- Johnen, H., González-Silva, L., Carramolino, L., Flores, J. M., Torres, M. and Salvador, J. M. (2013). Gadd45g is essential for primary sex determination, male fertility and testis development. *PLoS ONE* **8**, e58751.
- Kato, T., Miyata, K., Sonobe, M., Yamashita, S., Tamano, M., Miura, K., Kanai, Y., Miyamoto, S., Sakuma, T., Yamamoto, T. et al. (2013). Production of Sry knockout mouse using TALEN via oocyte injection. *Sci. Rep.* **3**, 3136.
- Katoh-Fukui, Y., Owaki, A., Toyama, Y., Kusaka, M., Shinohara, Y., Maekawa, M., Toshimori, K. and Morohashi, K.-i. (2005). Mouse Polycomb M33 is required for splenic vascular and adrenal gland formation through regulating Ad4BP/SF1 expression. *Blood* **106**, 1612-1620.
- Katoh-Fukui, Y., Miyabayashi, K., Komatsu, T., Owaki, A., Baba, T., Shima, Y., Kidokoro, T., Kanai, Y., Schedl, A., Wilhelm, D. et al. (2012). Cbx2, a polycomb group gene, is required for Sry gene expression in mice. *Endocrinology* **153**, 913-924.
- Kidokoro, T., Matoba, S., Hiramatsu, R., Fujisawa, M., Kanai-Azuma, M., Taya, C., Kurohmaru, M., Kawakami, H., Hayashi, Y., Kanai, Y. et al. (2005). Influence on spatiotemporal patterns of a male-specific Sox9 activation by ectopic Sry expression during early phases of testis differentiation in mice. *Dev. Biol.* **278**, 511-525.
- Kodzius, R., Kojima, M., Nishiyori, H., Nakamura, M., Fukuda, S., Tagami, M., Sasaki, D., Imamura, K., Kai, C., Harbers, M. et al. (2006). CAGE: cap analysis of gene expression. *Nat. Methods* **3**, 211-222.
- Koopman, P. (2002). General discussion. In *The Genetics and Biology of Sex Determination* (ed. D. Chadwick and J. Goode), pp. 240-249. Chichester, UK: John Wiley and Sons.
- Koopman, P., Münsterberg, A., Capel, B., Vivian, N. and Lovell-Badge, R. (1990). Expression of a candidate sex-determining gene during mouse testis differentiation. *Nature* **348**, 450-452.
- Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. and Lovell-Badge, R. (1991). Male development of chromosomally female mice transgenic for Sry. *Nature* **351**, 117-121.
- Kuroki, S., Matoba, S., Akiyoshi, M., Matsumura, Y., Miyachi, H., Mise, N., Abe, K., Ogura, A., Wilhelm, D., Koopman, P. et al. (2013). Epigenetic regulation of mouse sex determination by the histone demethylase Jmjd1a. *Science* **341**, 1106-1109.
- Kwok, C., Tyler-Smith, C., Mendonca, B. B., Hughes, I., Berkovitz, G. D., Goodfellow, P. N. and Hawkins, J. R. (1996). Mutation analysis of the 2 Kb 5' to SRY in XY females and XY intersex subjects. *J. Med. Genet.* **33**, 465-468.
- Laval, S. H., Glenister, P. H., Rasberry, C., Thornton, C. E., Mahadevaiah, S. K., Cooke, H. J., Burgoyne, P. S. and Cattanach, B. M. (1995). Y chromosome short arm-Sxr recombination in Xsry/Y males causes deletion of Rbm and XY female sex reversal. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 10403-10407.
- Li, G., Davis, B. W., Raudsepp, T., Wilkerson, A. J. P., Mason, V. C., Ferguson-Smith, M., O'Brien, P. C., Waters, P. D. and Murphy, W. J. (2013). Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. *Genome Res.* **23**, 1486-1495.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O. et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289-293.
- Loke, J., Pearlman, A., Radi, O., Zuffardi, O., Giussani, U., Pallotta, R., Camerino, G. and Ostrer, E. (2013). Mutations in MAP3K1 tilt the balance from SOX9/FGF9 to WNT/ β -catenin signaling. *Hum. Mol. Genet.* **23**, 1073-1083.
- Luo, X., Ikeda, Y. and Parker, K. L. (1994). A cell-specific nuclear receptor is essential for adrenal and gonadal development and sexual differentiation. *Cell* **77**, 481-490.
- Mahadevaiah, S. K., Odoriso, T., Elliott, D. J., Rattigan, Á., Szot, M., Laval, S. H., Washburn, L. L., McCarrey, J. R., Cattanach, B. M., Lovell-Badge, R. et al. (1998). Mouse homologues of the human AZF candidate gene RBM are expressed in spermatogonia and spermatids, and map to a Y chromosome deletion interval associated with a high incidence of sperm abnormalities. *Hum. Mol. Genet.* **7**, 715-727.
- Manuylov, N. L., Zhou, B., Ma, Q., Fox, S. C., Pu, W. T. and Tevosian, S. G. (2011). Conditional ablation of Gata4 and Fog2 genes in mice reveals their distinct roles in mammalian sexual differentiation. *Dev. Biol.* **353**, 229-241.
- Margarit, E., Guillén, A., Rebordosa, C., Vidal-Taboada, J., Sánchez, M., Ballesta, F. and Oliva, R. (1998). Identification of conserved potentially regulatory sequences of the SRY gene from 10 different species of mammals. *Biochem. Biophys. Res. Commun.* **245**, 370-377.
- McElreavey, K., Vilain, E., Abbas, N., Costa, J. M., Souleyreau, N., Kucheria, K., Boucekkine, C., Thibaud, E., Brauner, R. and Flamant, F. (1992). XY sex reversal associated with a deletion 5' to the SRY "HMG box" in the testis-determining region. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11016-11020.
- McElreavey, K., Vilain, E., Barbaux, S., Fuqua, J. S., Fehner, P. Y., Souleyreau, N., Doco-Fenzy, M., Gabriel, R., Quereux, C., Fellous, M. et al. (1996). Loss of sequences 3' to the testis-determining gene, SRY, including the Y pseudoautosomal boundary associated with partial testicular determination. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8590-8594.
- Melgar, M. F., Collins, F. S. and Sethupathy, P. (2011). Discovery of active enhancers through bidirectional expression of short transcripts. *Genome Biol.* **12**, R113.
- Mitsuhashi, T., Warita, K., Sugawara, T., Tabuchi, Y., Takasaki, I., Kondo, T., Hayashi, F., Wang, Z.-Y., Matsumoto, Y., Miki, T. et al. (2010). Epigenetic abnormality of SRY gene in the adult XY female with pericentric inversion of the Y chromosome. *Congenit. Anom.* **50**, 85-94.
- Miyake, Z., Takekawa, M., Ge, Q. and Saito, H. (2007). Activation of MTK1/MEKK4 by GADD45 through induced N-C dissociation and dimerization-mediated trans autophosphorylation of the MTK1 kinase domain. *Mol. Cell. Biol.* **27**, 2765-2776.
- Miyamoto, Y., Taniguchi, H., Hamel, F., Silversides, D. W. and Viger, R. S. (2008). A GATA4/WT1 cooperation regulates transcription of genes required for mammalian sex determination and differentiation. *BMC Mol. Biol.* **9**, 44.
- Morohashi, K., Honda, S., Inomata, Y., Handa, H. and Omura, T. (1992). A common trans-acting factor, Ad4-binding protein, to the promoters of steroidogenic P-450s. *J. Biol. Chem.* **267**, 17913-17919.
- Nagano, T., Lubling, Y., Stevens, T. J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E. D., Tanay, A. and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59-64.
- Nef, S., Verma-Kurvari, S., Merenmies, J., Vassalli, J.-D., Efstratiadis, A., Accili, D. and Parada, L. F. (2003). Testis determination requires insulin receptor family function in mice. *Nature* **426**, 291-295.
- Nishino, K., Hattori, N., Tanaka, S. and Shiota, K. (2004). DNA methylation-mediated control of Sry gene expression in mouse gonadal development. *J. Biol. Chem.* **279**, 22306-22313.
- Pannetier, M., Tilly, G., Kocer, A., Hudrisier, M., Renault, L., Chesnais, N., Costa, J., Le Provost, F., Vaiman, D., Vilotte, J.-L. et al. (2006). Goat SRY induces testis development in XX transgenic mice. *FEBS Lett.* **580**, 3715-3720.
- Pearlman, A., Loke, J., Le Caignec, C., White, S., Chin, L., Friedman, A., Warr, N., Willan, J., Brauer, D., Farmer, C. et al. (2010). Mutations in MAP3K1 cause 46,XY disorders of sex development and implicate a common signal transduction pathway in human testis determination. *Am. J. Hum. Genet.* **87**, 898-904.
- Pilon, N., Daneau, I., Paradis, V., Hamel, F., Lussier, J. G., Viger, R. S. and Silversides, D. W. (2003). Porcine SRY promoter is a target for steroidogenic factor 1. *Biol. Reprod.* **68**, 1098-1106.
- Pitetti, J.-L., Calvel, P., Romero, Y., Conne, B., Truong, V., Papaioannou, M. D., Schaad, O., Docquier, M., Herrera, P. L., Wilhelm, D. et al. (2013). Insulin and IGF1 receptors are essential for XX and XY gonadal differentiation and adrenal development in mice. *PLoS Genet.* **9**, e1003160.
- Poulat, F., Desclozeaux, M., Tuffery, S., Jay, P., Boizet, B. and Berta, P. (1998). Mutation in the 5' noncoding region of the SRY gene in an XY sex-reversed patient. *Hum. Mutat.* **11**, S192-S194.
- Prokop, J. W., Underwood, A. C., Turner, M. E., Miller, N., Pietrzak, D., Scott, S., Smith, C. and Milsted, A. (2013). Analysis of Sry duplications on the Rattus norvegicus Y-chromosome. *BMC Genomics* **14**, 792.
- Ross, D. G., Bowles, J., Koopman, P. and Lehnert, S. (2008). New insights into SRY regulation through identification of 5' conserved sequences. *BMC Mol. Biol.* **9**, 85.
- Sekido, R. and Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* **453**, 930-934.
- Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frischauf, A.-M., Lovell-Badge, R. and Goodfellow, P. N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346**, 240-244.
- Spitz, F. and Furlong, E. E. M. (2012). Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613-626.
- Tevosian, S. G., Albrecht, K. H., Crispino, J. D., Fujiwara, Y., Eicher, E. M. and Orkin, S. H. (2002). Gonadal Differentiation, sex determination and normal Sry

- expression in mice require direct interaction between transcription partners GATA4 and FOG2. *Development* **129**, 4627-4634.
- Turner, M. E., Martin, C., Martins, A. S., Dunmire, J., Farkas, J., Ely, D. L. and Milsted, A.** (2007). Genomic and expression analysis of multiple Sry loci from a single *Rattus norvegicus* Y chromosome. *BMC Genet.* **8**, 11.
- Veitia, R., Ion, A., Barbaux, S., Jobling, M. A., Souleyreau, N., Ennis, K., Ostrer, H., Tosi, M., Meo, T., Chibani, J. et al.** (1997a). Mutations and sequence variants in the testis-determining region of the Y chromosome in individuals with a 46,XY female phenotype. *Hum. Genet.* **99**, 648-652.
- Veitia, R. A., Fellous, M. and McElreavey, K.** (1997b). Conservation of Y chromosome-specific sequences immediately 5' to the testis determining gene in primates. *Gene* **199**, 63-70.
- Wang, H., Yang, H., Shivalila, C. S., Dawlaty, M. M., Cheng, A. W., Zhang, F. and Jaenisch, R.** (2013a). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-918.
- Wang, H., Hu, Y.-C., Markoulaki, S., Welstead, G. G., Cheng, A. W., Shivalila, C. S., Pyntikova, T., Dadon, D. B., Voytas, D. F., Bogdanove, A. J. et al.** (2013b). TALEN-mediated editing of the mouse Y chromosome. *Nat. Biotechnol.* **31**, 530-532.
- Warr, N. and Greenfield, A.** (2012). The molecular and cellular basis of gonadal sex reversal in mice and humans. *Wiley Interdiscip. Rev. Dev. Biol.* **1**, 559-577.
- Warr, N., Bogani, D., Siggers, P., Brixey, R., Tateossian, H., Dopplapudi, A., Wells, S., Cheeseman, M., Xia, Y., Ostrer, H. et al.** (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, J. V., Brixey, R., Pope, M., Bogani, D., Childers, M., Wells, S., Scudamore, C. L. et al.** (2012). Gadd45γ and Map3k4 interactions regulate mouse testis determination via p38 MAPK-mediated control of Sry expression. *Dev. Cell* **23**, 1020-1031.
- Wilhelm, D. and Englert, C.** (2002). The Wilms tumor suppressor WT1 regulates early gonad development by activation of Sf1. *Genes Dev.* **16**, 1839-1851.
- Wilhelm, D., Washburn, L. L., Truong, V., Fellous, M., Eicher, E. M. and Koopman, P.** (2009). Antagonism of the testis- and ovary-determining pathways during ovotestis development in mice. *Mech. Dev.* **126**, 324-336.
- Yokouchi, K., Ito, M., Nishino, K., Yamanouchi, K., Naito, K., Suzawa, M., Kato, S., Hakamata, Y., Endo, H. and Tojo, H.** (2003). Stage-specific regulatory element of mouse Sry gene. *Mol. Reprod. Dev.* **64**, 389-396.