Published online 2 May 2017

Nucleic Acids Research, 2017, Vol. 45, No. 12 ***7191–7211***
doi: 10.1093/nar/gkx328

In mammalian foetal testes, SOX9 regulates expression of its target genes by binding to genomic regions with conserved signatures

Massilva Rahmoun[1,†], Rowena Lavery[2,†], Sabine Laurent-Chaballier[3,‡], Nicolas Bellora[4,‡], Gayle K. Philip[5,‡], Moïra Rossitto[1], Aleisha Symon[2], Eric Pailhoux[6], Florence Cammas[3], Jessica Chung[5], Stefan Bagheri-Fam[2], Mark Murphy[7], Vivian Bardwell[7], David Zarkower[7], Brigitte Boizet-Bonhoure[1], Philippe Clair[8], Vincent R. Harley[2,*,§] and Francis Poulat[1,*,§]

[1]Institute of Human Genetics, CNRS-University of Montpellier UMR9002, 34396 Montpellier cedex 5, France, [2]The Hudson Institute of Medical Research and Department of Anatomy, Monash University, Melbourne, Australia, [3]Institut de Recherche en Cancérologie de Montpellier, IRCM, INSERM U1194, Université de Montpellier, Institut régional du Cancer de Montpellier, Montpellier F-34298, France, [4]Instituto Andino Patagónico de Tecnologías Biológicas y Geoambientales (IPATEC), Universidad Nacional del Comahue - CONICET, Bariloche, Argentina, [5]VLSCI, LAB-14, 700 Swanston Street, Carlton 3053, Victoria, Australia, [6]INRA Biologie du Développement et Reproduction, Domaine de Vilvert, 78352 Jouy-en-Josas Cedex, France, [7]Department of Genetics, Cell Biology and Development, University of Minnesota, 6-160 Jackson hall, 321 Church St, SE, Minneapolis, MN 55455, USA and [8]University of Montpellier, Montpellier GenomiX, bat 24, Place Eugène Bataillon, 34095 Montpellier cedex 5, France

Received August 24, 2016; Revised March 23, 2017; Editorial Decision April 07, 2017; Accepted April 17, 2017

ABSTRACT

In mammalian embryonic gonads, SOX9 is required for the determination of Sertoli cells that orchestrate testis morphogenesis. To identify genetic networks directly regulated by SOX9, we combined analysis of SOX9-bound chromatin regions from murine and bovine foetal testes with sequencing of RNA samples from mouse testes lacking *Sox9*. We found that SOX9 controls a conserved genetic programme that involves most of the sex-determining genes. In foetal testes, SOX9 modulates both transcription and directly or indirectly sex-specific differential splicing of its target genes through binding to genomic regions with sequence motifs that are conserved among mammals and that we called 'Sertoli Cell Signature' (SCS). The SCS is characterized by a precise organization of binding motifs for the Sertoli cell reprogramming factors SOX9, GATA4 and DMRT1. As SOX9 biological role in mammalian gonads is to determine Sertoli cells, we correlated this genomic signature with the presence of SOX9 on chromatin in foetal testes, therefore equating this signature to a genomic bar code of the fate of foetal Sertoli cells. Starting from the hypothesis that nuclear factors that bind to genomic regions with SCS could functionally interact with SOX9, we identified TRIM28 as a new SOX9 partner in foetal testes.

INTRODUCTION

Sex determination in mammals provides an ideal *in vivo* model to study cell fate regulation by transcription factors. Indeed, embryonic gonads originate from a bipotential primordium (the genital ridge) that can differentiate into testis or ovary via the action of distinct transcription factors. In the mouse embryo, sex determination occurs between embryonic day (E) 10.5 and E12.5 (1) by differentiation of progenitors into a specific cell population (the supporting cells) within the genital ridge. Supporting cells organize the embryonic gonads morphologically and induce sex determination in the germinal lineage. In males, differentiation of supporting cells, also known as Sertoli cells, is driven by the expression of the Y chromosome gene *Sry* that activates the autosomal gene *Sox9* (2), a SOX E group family member. In the absence of *Sry* in females, progenitors differentiate into granulosa cells, the female supporting cell lineage. Genetic experiments have demonstrated that *Sox9*

*To whom correspondence should be addressed. Tel: +33 4 34 35 99 40; Fax: +33 4 34 35 99 01; Email: francis.poulat@igh.cnrs.fr
Correspondence may also be addressed to Vincent R Harley. Email: vincent.harley@hudson.org.au
†These authors contributed equally to this work as first authors.
‡These authors contributed equally to this work as second authors.
§These authors contributed equally to this work as last authors.

is the essential direct target gene of SRY (3) and the central effector of the male pathway. Heterozygous missense or nonsense *SOX9* mutations in 46, XY human patients cause Disorders of Sex Development (DSD) and induce a highly penetrant (75%) male-to-female sex-reversal phenotype as well as campomelic dysplasia (4). Similarly, targeted disruption of *Sox9* in mouse embryonic testes leads to their development into ovaries (5,6). Conversely, *SOX9* duplication in XX patients (7) or ectopic expression of *Sox9* in mouse embryonic XX gonads induces testis formation (8,9). Upon *Sox9* activation in mouse embryonic gonads, *Sox8* and *Sox10*, the two other SOX E group members, are also induced and either gene can replace *Sox9* function and masculinize the gonad (10,11).

Four other transcription factors are required for Sertoli cell differentiation: DMRT1, GATA4, WT1 and SF1 (also known as NR5A1). Like for SOX9, alterations or point mutations in the genes encoding these factors have been detected in XY individuals with DSD ((12) and reviewed by (13)) and ablation of each of these genes affects testis development in the mouse (14–17). Together, DMRT1, GATA4, WT1, SF1 and SOX9 induce direct reprogramming of mouse embryonic fibroblasts into functional embryonic Sertoli-like cells, thereby defining the transcription factor module required for the Sertoli cell fate (18). However, *Sox9* is the only one with male-specific expression and induces female-to-male sex reversal when ectopically expressed in XX human and mouse gonads (7–9), while *Dmrt1*, *Gata4*, *Wt1* and *Sf1* are expressed in both male and female progenitors. By contrast, *Dmrt1* overexpression in female embryonic gonads induces incomplete Sertoli differentiation without testis cords formation (19), while *Dmrt1* expression at a physiological level cannot induce the male pathway in XX embryonic gonads (20). Taken together, genetic, transcriptomic and direct reprogramming experiments suggest that the fate choice towards Sertoli cells is controlled by SOX9 and its set of target genes. This implies that the differentiation of supporting cell progenitors into male-specific Sertoli cells or female-specific granulosa cells is controlled by SOX9 presence or absence.

In the present study, to better understand the mechanisms underlying SOX9 role in foetal Sertoli cell differentiation, we used chromatin immunoprecipitation followed by sequencing (ChIP-seq) of samples from foetal testes of two distant mammals (mouse and cattle). We found that SOX9 binds to 4293 genes in common between the mouse and bovine genomes. Most of these genes are already known to be involved in sex determination. Moreover, transcriptomic (RNA-seq) analysis of foetal testes from *Sox9* knockout mice showed that SOX9 not only regulates transcription of its target genes directly, but also influences their RNA splicing.

Finally, analysis of the ChIP-seq data from these two distant mammals allowed the identification of genomic DNA motifs that characterize genes bound by SOX9 in differentiated Sertoli cells from foetal testes. This 'Sertoli cell signature' (SCS) is conserved among mammals and is characterized by the organized clustering of SOX9, GATA4 and DMRT1 binding sites. At the chromatin level, we observed that DMRT1 and GATA4 can be co-localized with SOX9 on its target genes. *In silico* prediction of SCS identified a potential link between SOX9 and the nuclear factor TRIM28 (TIF1ß or KAP1). We confirmed this prediction experimentally by showing the physical interaction between SOX9 and TRIM28 in foetal testes, their co-localization on the foetal testis chromatin and their potential functional interaction in transcriptional regulation.

MATERIALS AND METHODS

Cells and animals

NT_2D_1 cells were obtained from the American Type Culture Collection (ATCC) and cultivated in DMEM/F12 medium with Glutamax (Life Technologies) supplemented with 10% foetal bovine serum (Life Technologies). Animal care and handling (mouse foetal gonads used for ChIP-seq) were according to the 'Réseau des Animaleries de Montpellier' (RAM). For the foetal gonads from wild type and $Sox9^{\Delta/\Delta}$ animals (RNAseq), all procedures involving mice were approved by the Animal Ethics Committee of Monash University, Australia. Experiments on bovine foetuses reported in this work were performed in agreement with the ethical guidelines of the French National Institute for Agricultural Research (INRA). Foetuses were produced by artificial insemination of Holstein females with semen of Holstein males (day 0) and collected at E90 at the INRA slaughterhouse (France). The protocol (no. 12/046) was approved by the local ethics committee (COMETHEA) and Eric Pailhoux is the recipient of an official authorization for animal experimentation (no. B91-649).

Tissue fixation and ChIP-seq analysis

Micro-dissected gonads were separated from the mesonephros of E13.5 (mouse) or E90 (bovine) embryos, snap-frozen and stored at −80°C. Frozen gonads were crushed in liquid nitrogen using a mortar. Powdered tissue samples were immediately fixed with PBS containing 2 mM of disuccinimidyl glutarate (DSG) (Pierce ref 20693) for 30 min as described previously (21). After three washes with PBS, samples were fixed in PBS/1% formaldehyde at room temperature for 30 min. Protein A coupled to Dynabeads magnetic beads (Life Technologies) was pre-incubated at 4°C with saturating buffer [1 mg/ml BSA, 0.2 mg/ml glycogen (Roche), 1 mg/ml yeast tRNA (Roche) in 1× PBS] for 1 h. Beads were then incubated at 4°C with 2 μg of purified home-made rabbit polyclonal anti-SOX9 IgG antibody (22) in 50 μl of PBS/0.02% Tween for 4 h. After washing, antibody-bound beads were used for immunoprecipitation (IP) of 20 μg of sonicated chromatin overnight. IP and washing buffers were as described (2). Unbound sonicated DNA was sequenced as input. Six independent IPs were pooled in two samples (three IPs/each) and used for construction of two independent ChIP-seq libraries that were sequenced at the BGI facility (BGI Shentzen, China). ChIP-seq reads were aligned to the mouse genome using SOAP2.21 (mm9, build 37) or to the bovine genome (Bostau6, UMD3.1). For peak calling, as the number of reads in the inputs was twice that in the ChIP-seq experiments, 20 million reads were randomly sampled from each input, using a homemade script that respects the proportion of reads per chromosome of the initial input.

With this method, 10 input samples were generated for each ChIP-seq analysis that were used for ten rounds of peak calling with the MACS software (1.4) (23) (bw = 250, mfold = 10 and *P* value = 1e−5). Peaks common to the 10 operations with FDR<0.05 were kept for further analysis. Peaks were assigned to neighbouring genes using the PAVIS annotation tool (http://manticore.niehs.nih.gov/pavis2/) with the UCSC annotation for mouse mm9 and the Ensembl annotation UMD3.1r82 for bovine Bostau6. For annotation of the bovine peaks using the mouse genome as reference, bovine coordinates were translated to mouse mm9 coordinates, as previously described (24), using the UCSC LiftOver tool. Coordinates translated to mm9 where then annotated using PAVIS. The correspondence of orthologous genes between mouse and bovine were analyzed using BIOMART (http://www.ensembl.org/biomart/martview/5761d5751336b65a1461f18d849facfe). Pathway analysis was performed using the MetaCore™ GeneGo software to map genes to a global database of known networks (https://portal.genego.com/). The top 50 enriched ($P < 0.05$) pathway networks were analyzed from genes associated to the mouse and bovine SOX9 ChIP-seq peaks.

ChIP-qPCR and sequential ChIP-qPCR

Experiments were carried out with the same protocol as for ChIP-seq. The anti-DMRT1 antibody was described previously (25); the anti-GATA4 antibody and control IgG were purchased from Santa-Cruz Biotech, Heidelberg, Germany (Sc-1237and sc-66830, respectively). For sequential ChIP, the rabbit-anti SOX9 antibody was cross-linked to protein A prior to immunoprecipitation in order to prevent its dissociation from protein A magnetic beads. After incubation of the antibody with protein A Dynabeads, beads were washed with 0.2 M sodium borate pH 9.0 and incubated in the same buffer in the presence of 20 mM dimethylpimelimidate (DMP, Pierce ref 21666) at room temperature for 45 min. Beads were then washed and incubated at RT with 0.2 M ethanolamine pH 8.0 for 1 h, washed three times with 0.58% (v/v) acetic acid/150 mM NaCl and then three times with PBS. In each experiment, the equivalent of 2 μg of coupled antibody was used. Chromatin was eluted from antibodies after immunoprecipitation with 10 mM Tris pH 8, 1 mM EDTA, 2% SDS, 15 mM DTT at 37°C for 1 h. Supernatants were diluted 1:20 in 50 mM Tris pH8, 150 mM NaCl, 2 mM EDTA, 1% Triton X100 and processed for the second immunoprecipitation as before.

After de-crosslinking, DNA was purified with the MinElute PCR Purification Kit (QIAGEN GmbH, D-40725 Germany, ref 28004). Quantitative PCR was performed with the SYBR Select Master Mix (Applied Biosystems ref 4472908) and a LightCycler 480 apparatus (Roche). Primer sequences are listed in Supplementary Table S9 of the Additional Methods.

Embryonic tissue collection and RNA preparation

Micro-dissected gonads were separated from the mesonephros of E13.5 embryos, snap-frozen and stored at −80°C. Amh-*Cre; Sox9*$^{flox/flox}$ mice were obtained as described previously (26). All animal work was approved by Monash Medical Centre Animal Ethics (MMCB/2012/23). RNA was extracted using the RNeasy Kit (QIAGEN GmbH, Germany). The developmental stage and phenotype of each embryo was visually analyzed and recorded. A tail piece was collected for genotyping and genetic sex was determined as previously described (27). Each RNA sample (Amh-*Cre; Sox9*$^{flox/flox}$ XY and wild type *Sox9*$^{flox/flox}$ XX) consisted of a minimum of six pooled gonads for each genotype. An equal number of age-matched mouse gonads for each genotype was collected and divided in two additional groups to provide three biological replicates for each RNA-seq analysis. RNA quality was assessed by measuring the optical density at 230, 260 and 280 nm using a ND-1000 Spectrophotometer (Nanodrop, Wilmington, DE, USA) and by evaluating their integrity using the Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA). RNA samples with a RIN value of ≥7.5 were used to generate cDNA libraries for sequencing. Typically, 800 ng of high quality total RNA was obtained from six pooled Amh-*Cre; Sox9*$^{flox/flox}$ gonads.

RNA-seq library preparation, sequencing and analysis

Libraries were generated using the NuGEN Mondrian Technology and the SPIA amplification methodology that specifically selects and linearly reverse transcribes non-ribosomal RNA. 6pM of bar-coded libraries per lane were sequenced using the Illumina HiSeq1500 and paired-end RNA sequencing chemistry (150nt read lengths at 50 million reads per sample). All datasets were processed with the RNA-seq pipeline (http://dx.doi.org/10.5281/zenodo.47479) as follows. Sequencing reads were quality-filtered and adapters were removed using *Trimmomatic* (version 0.30) (28). Reads shorter than 36 bp were discarded from subsequent analyses and the remaining trimmed reads were mapped to the *Mus musculus* genome (Ensembl version 38.77) using *Tophat2* (version 2.0.8) (29). *HTSeq-count* (version 0.6.1p1) (30) was used to count the reads overlapping with genes using the union overlap resolution mode. A statistical analysis of the count data was then performed using '*voom*'(31) from the '*limma*' (version 3.24.0) (32) *R* package to assess the differential expression between the different experimental conditions (KO versus XX, KO versus XY, and XX versus XY). Specifically, filtering was performed to include only genes the expression of which in counts-per-million (CPM) was greater than 1 in at least two samples. To eliminate composition biases between libraries, the Trimmed Mean of *M*-values (TMM) method was used to calculate normalization factors between samples (33). The *voomWithQualityWeights* function (34), which combines observational-level and sample-specific weights, was used in the subsequent linear modelling to down-weight more variable samples (35). Differentially expressed genes were examined with an adjusted p value cut-off of 0.05.

For alternative splicing analysis, RNA-seq datasets were analyzed independently as follows. Paired-end raw reads were cleaned of sequencing contaminants with Scythe (https://github.com/vsbuffalo/scythe.) and quality trimmed with Sickle (https://github.com/najoshi/sickle) to keep only

reads with a minimum length of 21nt. Mouse genomic sequences (mm10) and annotations (Ensembl 75) were retrieved from the UCSC genome browser. Alignments to the mouse genome were performed with STAR v2.4.1c (36) and quantification of Ensembl isoforms with Salmon v0.3.0 (http://biorxiv.org/content/early/2015/10/03/021592). Both programs ran with default read-matching parameters. Differential expression was calculated between pairs of conditions using one-way ANOVA and *t*-tests on square-root transformed TPM data of the three replicates (37). MISO (38) was used for alternative splicing analysis and results were gathered with the following parameters: –num-inc 0 –num-exc 0 –num-sum-inc-exc 5 –delta-psi 0.20 –bayes-factor 1. SOX9 ChIP-seq peaks were translated from the mm9 genome assembly to mm10 using LiftOver for correspondence of splicing events with ChIP-seq peaks.

Motifs analyses

Detailed in supplemental methods.

Cell culture, generation of tet-SOX9 cell lines and gene expression analysis by RT-qPCR

NT_2D_1 and HEK293T cells were cultured as previously described (39,40). For generation of lentiviral vectors, the human SOX9 open reading frame was cloned in a modified version of the pTRIPZ vector (Open Biosystems) to introduce a HA-tag in the SOX9 N-terminus. Viral particles were produced at the vectorology facility (PVM, Plateforme de Vectorologie de Montpellier). After infection with the pTRIPZ-SOX9 plasmid, NT_2D_1 and HEK293T cells were grown in the presence of 1 μg/ml puromycin (Invitrogen). SOX9 expression was inducted by addition of 2 μg/ml tetracycline (Sigma) for 16 h. Tet-SOX9-HEK293 cells were transfected with empty plasmid or pCX-Flag-mTRIM28 using Lipofectamine® 2000 (Invitrogen) following the manufacturer's protocol. Total RNA was isolated using the TRIzol® reagent (Life Technologies) and the DNA-free Kit (Invitrogen). Total RNA was reverse transcribed with SuperScript® IV (Invitrogen) and quantitative PCR performed as described in the ChIP-qPCR chapter. Primer sequences are listed in Supplementary Table S8 of Additional Methods.

Gel mobility shift assay

Proteins were produced using TNT™ Quick Coupled /Translation System (Promega) according to the manufacturer's instructions. Probe annealing and experiments were carried out as described before (41) for mouse DMRT1 and with some modifications for GATA4 and SOX9. For SOX9 binding assay, the non-specific competitor dI-dC was replaced by poly(dG-dC) (GE Healthcare). Mouse GATA4 was *in vitro* translated using as template the pCS2$^+$mGATA4 plasmid, a generous gift from Dr Sergei Tevosian.

For Figure 4, the top strand probes were:

Lats2: 5′-ggCCGAGCGGGACATTCGCTACATTG TTGGCATTCCACGGGCG-3′

Lmo4: 5′-ggCATCTCCCATTATTGTTCCAAATCTC ATTTCATTTTGAA-3′

Mrpl45: 5′-ggGTTTTTTTCACCGATTGTAAATAA GGTGTAACAATGTGTTAAGGAACCAGGA-3′

DMRT1 BS (12): 5′-ggGGGAGATTTGATACATTG TTACTTTATGG-3′

GATA4 BS (42): 5′-ggGGGGCTTTGGTCTCAGCT TATCAAACTGCCCTG-3′

SOX9 BS (43): 5′-ggGTTGACAGAACAATGGCTGC TAGA-3′

The 'g' nucleotides at C-terminus were added for probe labelling with Klenow DNA polymerase and α [^{32}P]-dCTP

Immunoprecipitation using NT_2D_1 cell and embryonic gonad tissue extracts

Co-immunoprecipitation using NT_2D_1 cell extracts was performed as previously described (39) with 2 μg of purified rabbit antibodies against the N- (44) or C-terminal (22) regions of SOX9. For immunoprecipitations using male and female E13.5 gonad tissues, nuclear extracts were prepared with the Nuclear Extract Kit (Active Motifs, La Hulpe, Belgium). Nuclear extracts corresponding to 30 pairs of gonads were diluted in lysis buffer in the same conditions as for NT_2D_1 cells. After extensive washes (wash buffer: 150 mM NaCl, 50 mM Tris pH 7.5, 2 mM EDTA, 1% NP40), immune complexes were analyzed by western blotting using a monoclonal antibody against mouse TRIM28 (45). As control, 5% of immunoprecipitated supernatant was used for western blotting with a control IgG antibody (Santa-Cruz Biotech, Heidelberg, Germany).

Immunostaining and antibodies

Gonads were dissected from staged embryos, fixed in 4% paraformaldehyde and paraffin-embedded. Expression of SOX9, TRIM28 and DMRT1 was assessed by incubating 5-μm sections with a rabbit polyclonal anti-SOX9 (1:300) (22), a mouse monoclonal anti-TRIM28 (1:2000) (46) and a goat polyclonal anti-DMRT1 (1:25) antibody (Santa Cruz Cat. No. sc-104885) at 4°C overnight, followed by the appropriate secondary antibodies (1/800) (Alexa-Ig, Molecular Probe). Nuclei were stained with DAPI (Sigma). Images were captured with a Zeiss Axioimager Apotome microscope or a DeltaVision OMX microscope.

In vitro pull-down assay

The bacterially expressed GST-mTRIM28 fusion protein was produced and purified as previously described (40). The open reading frame of human SOX9 was cloned in the pet28a vector (Novagen) and the histidine-tagged protein (6xHis-SOX9) was bacterially expressed and purified as previously described (47). For *in vitro* interaction, purified GST-TRIM28 or GST alone was incubated with 6xHis-SOX9 in 150 mM NaCl, 50 mM Tris pH 7.5, 0.5 mM DTT and 0.5% NP40 supplemented with 1 mg/ml of BSA (Sigma) at room temperature for 20 min. Protein complexes were captured using glutathione–agarose beads (Sigma). After extensive washes, proteins were analyzed by western blotting with the previously described anti-SOX9 and -TRIM28 antibodies.

Availability of data and computer programs

Datasets generated for this manuscript are accessible at GEO (http://www.ncbi.nlm.nih.gov/geo/) with the accession number GSE81490. Computer programs, source codes and scripts are freely available at GitHub (https://github.com/fpoulat/SCS-motifs-scanning)

RESULTS

Most genes interacting with SOX9 contribute to Sertoli cell determination in foetal testis

ChIP-seq experiments were performed using chromatin from E13.5 mouse and E90 bovine foetal testes. In both species, this represents the stage of testis development when Sertoli cells differentiate. The specificity of the homemade anti-rabbit SOX9 antibody used for ChIP was demonstrated by the finding that the antibody recognized only chromatin from male, but not female E13.5 gonads in a slot-blot experiment (Supplementary Figure S1).

ChIP-seq experiments gave 6419 (mouse) and 11 988 (bovine foetal testis) peaks with FDR<0.05 (Supplementary Table S1: Sheets 'ChIP-seq mouse mm9' and 'ChIP-seq bovine Bostau6'). To overcome the problem of limited gene annotations in the *Bos taurus* genome (BosTau6), bovine peaks were attributed either by direct annotation using the bovine genome assembly (BosTau6, Ensembl UMD3.1) (Supplementary Table S1: Sheet 'ChIP-seq bovine Bostau6'), or by translating the BosTau6 genomic coordinates of the bovine ChIP-seq peaks to the mouse genome assembly mm9 using the LiftOver tool from UCSC (Supplementary Table S1: Sheet 'ChIP-seq bovine liftover mm9'). The 6419 ChIP-seq peaks from the mouse samples were assigned to 5844 genes and the 11 988 peaks from bovine samples were assigned to 8229 genes by direct annotation using the bovine genome and to 8957 genes using LiftOver from BosTau6 to mm9. Among the mouse peaks, 70% were preferentially located within the gene body [5′ untranslated regions (UTRs): 19%. 3′ UTRs: 1%. exons: 15%; introns: 35%] and 30% were outside [24% upstream the Transcription Start Site (TSS) and 6% downstream the Transcription End Site (TES)] (Figure 1A, left panel). Conversely, direct annotation using the BosTau6 genome assembly indicated that 62% of bovine peaks were outside the gene body (48% upstream of the TSS and 14% downstream of the TES) (Figure 1A, middle panel). However, when using the mouse genome as reference after translation of the bovine coordinates, the bovine and mouse peak distributions were comparable (Figure 1A, right panel).

Cross-species comparison (24) of the regions bound by SOX9 in bovine and mouse foetal testes, identified 4293 peaks that were common to both ChIP-seq datasets, 75% of which (3236 peaks) overlapped by >500 bp (Figure 1B). From these 4293 peaks, 3849 neighboring genes were annotated that showed strict orthology using Biomart. This suggests that SOX9 binds to and regulate a similar core set of genes in two evolutionarily distant mammals. Some of these genes (for instance, *Col27a1*/*COL27A1* (48), *Fgf9*/*FGF9* (49) and *Sox10*/*SOX10* (10)) displayed striking similarity of peak location in the mouse and bovine ChIP-seq datasets. Conversely, the peak profile of *Lmo4*/*LMO4* (50) was relatively different between species.

Gene Ontology (GO) term enrichment analysis (Supplementary Table S1: Sheets GO) of genes bound by SOX9 showed that the male sex differentiation pathway was enriched in the mouse ($P = 3.454 \times 10^{-7}$), bovine ($P = 2.972 \times 10^{-7}$) and the combined mouse and bovine gene sets ($P = 3.811 \times 10^{-6}$). Moreover, other enriched pathways, such as FGF, NOTCH, Hedgehog and WNT signalling cascades, also contribute to male sex determination and gonad function.

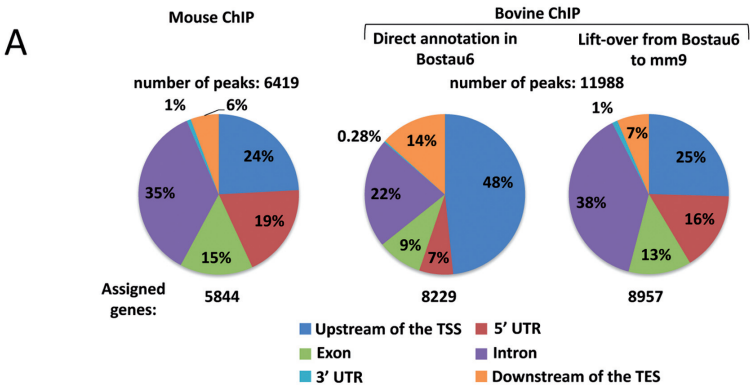
Most genes that have been involved in mammalian gonad determination and differentiation by genetic or functional studies were identified here as SOX9 targets (Table 1). This included genes with early roles, such as *Sry* activation (*Kdm3a* (51), *Gadd45g* (52), *Gata4* (53) and *Cbx2* (54)), and genes acting downstream of SOX9 (*Tcf21* (55), *Sox10* (10), *Fgf9* (56), *Dhh* (57), *Hhat* (58) and *Amh* (59)). Some genes were bound by SOX9 in the mouse, but not in cattle (*Ctnnb1, Cited2, Dmrt1, Hhat, Notch1, Pdgfa* and *Wnt4*), or *vice-versa* (*CYP26B1, PTGDS, SOX8*), suggesting a divergent regulation between species, or SOX9 redundancy with other SOX E group factors (SOX8 and 10) (10,60). SOX9 also bound to genes involved in the female programme of sex determination, such as *Wnt4* (61), *Fst, Bmp2* (62) and *FoxL2* (63), suggesting that it could negatively regulate their transcription in Sertoli cells.

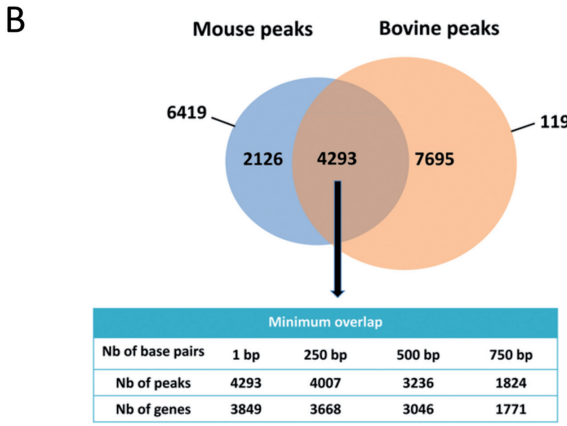
Transcriptional regulation by SOX9 in foetal testis at E13.5

Comparison of the mouse SOX9 ChIP-seq dataset with previously published transcriptomic datasets from male and female supporting cell lineages (Sertoli and granulosa cells, respectively) in mouse E13.5 gonads (64) enabled us to identify candidate genes that could be up- or down-regulated by SOX9 cell autonomously at this stage. Specifically, 38% (557 out of 1451) of genes that were upregulated and 44% (539 out of 1225) of genes that were repressed in Sertoli cells compared with granulosa cells at E13.5 were bound by SOX9 (Figure 2A; see also Supplementary Table S1: Sheet 'ChIP-seq mouse mm9').

Then, to better understand SOX9 transcriptional function in foetal mouse testes, we analysed by RNA-seq the transcriptome of E13.5 gonads in which SOX9 was present or not. *Sox9* conditional knock-out in differentiating Sertoli cells was performed soon after testis commitment by crossing mice carrying the *Sox9* allele ($Sox9^{flox}$) (65) with mice in which the Cre recombinase was controlled by the *Amh* gene promoter (66). In this context, a mild phenotype is observed and the effect on E13.5 gonad transcriptome can be essentially attributed to transcriptional mis-regulation in Sertoli cells due to *Sox9* loss (67). *Sox9* mRNA quantification in the RNA-seq datasets obtained using E13.5 wild type (WT) XY and XX gonads and XY mutant gonads (subsequently named XY $Sox9^{\Delta/\Delta}$ or KO in figures) showed that *Sox9* transcript levels were comparable in $Sox9^{\Delta/\Delta}$ XY and XX gonads (Supplementary Figure S2A), as previously reported (67).

Detailed analysis of the $Sox9^{\Delta/\Delta}$ testis transcriptome showed the downregulation of 240 mRNAs and the upregulation of 107 transcripts compared with WT XY testes





Minimum overlap				
Nb of base pairs	1 bp	250 bp	500 bp	750 bp
Nb of peaks	4293	4007	3236	1824
Nb of genes	3849	3668	3046	1771

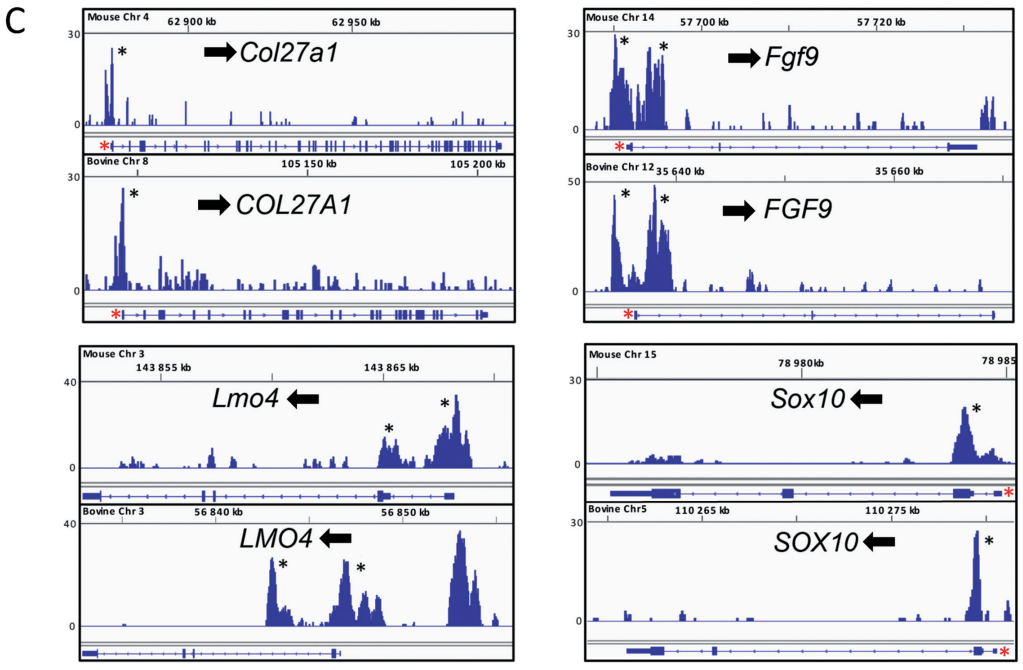


Figure 1. SOX9 peaks in mouse and bovine foetal testes. (**A**) Distribution of SOX9 peaks relative to neighbouring genes and gene components: upstream of the transcription start site (TSS), 5′ untranslated region (5′ UTR), exon, intron, 3′ untranslated region (3′ UTR) and downstream of the transcription end site (TES). For the bovine ChIP-seq dataset, annotations were made directly using the BosTau6 genome assembly (Ensembl UMD3.1) (middle pie-chart) or indirectly using LiftOver and the mouse genome (mm9) as the reference genome (right pie chart). (**B**) Venn diagram showing that SOX9 binds to 4293 orthologous mouse and bovine foetal testis regions. Bottom table: number of overlapping peaks (and of the respective genes) in the mouse and bovine ChIP-seq datasets (overlaps from 1 base pair (bp) to 750 bp). (**C**) Examples of peak position conservation in the mouse and bovine ChIP-seq datasets. For each gene, the upper panel represents the mouse gene and the lower panel the bovine gene. The input background was subtracted using IGV. Binding peaks called by MACS are indicated by asterisks. As in the BosTau6 genome assembly the TSS of bovine *FGF9* are missing, they were determined using LiftOver with mm9 and added by hand using IGV. Black arrows, direction of transcription; red stars, transcription start sites.

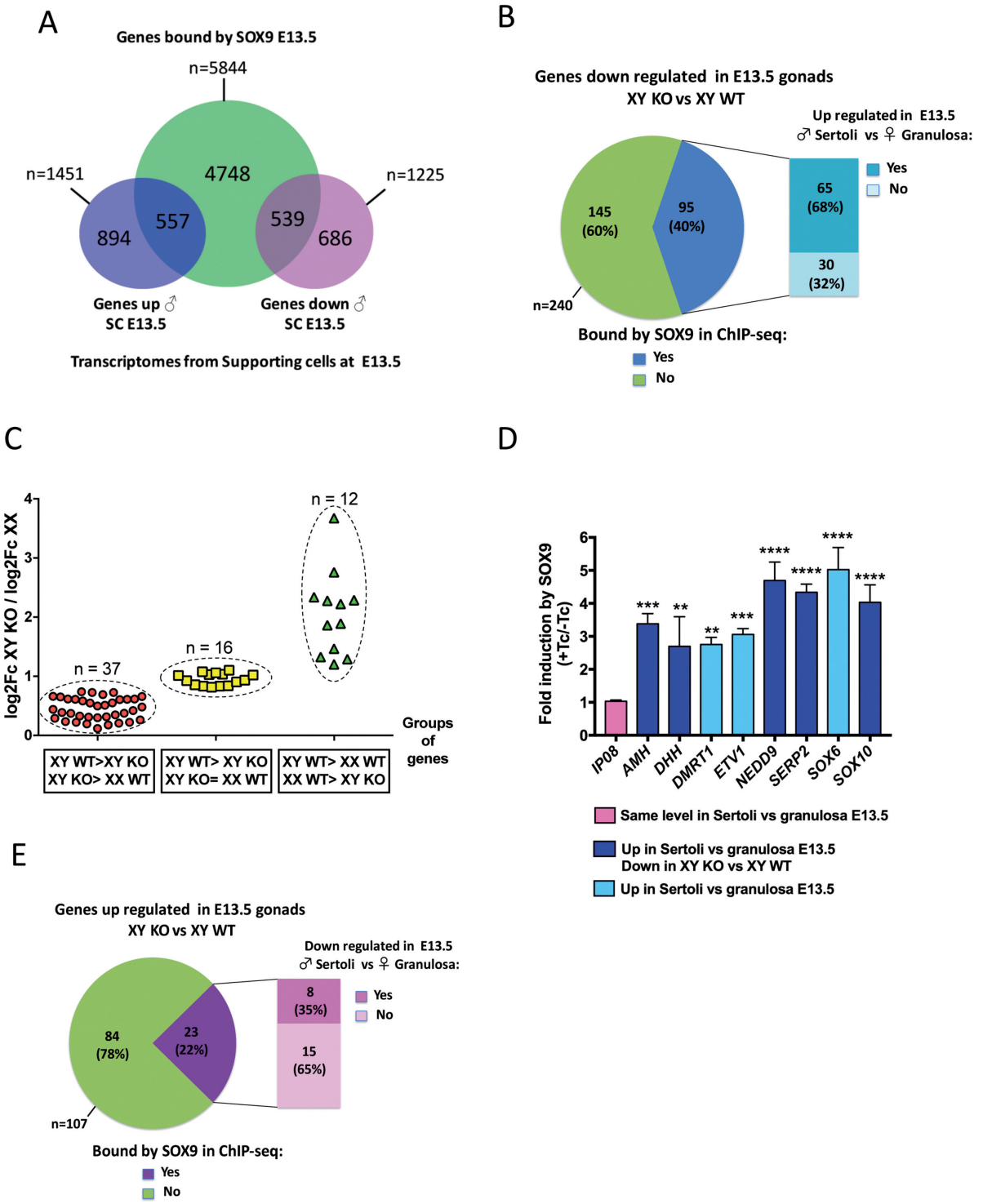


Figure 2. SOX9 role in the expression of its target genes in foetal testis (**A**) Venn diagram showing the association between mouse SOX9-bound genes (ChIP-seq dataset) and genes that are differentially expressed in E13.5 Sertoli and granulosa cells. Genes up or down: genes that are upregulated or downregulated in E13.5 male and female supporting cells (SC) (64). (**B**) Pie chart showing the classification of genes that are downregulated in $Sox9^{\Delta/\Delta}$ XY (XY KO) versus wild type XY (XY WT) gonads in genes bound or not by SOX9 (mouse ChIP-seq dataset). Among the 95 SOX9-bound genes, 68% were upregulated at E13.5 in wild type Sertoli versus granulosa cells (64). (**C**) The ratio between the log2 fold change (FC) of XY KO versus XY WT and the log2FC of XX WT versus XY WT allowed classifying the differentially regulated genes in three groups according to their levels of expression in the XY WT, XY KO and XX WT RNA-seq datasets. (**D**) RT-qPCR analysis of SOX9-mediated induction of gene expression in tet-SOX9-NT_2D_1 cells upon addition of tetracycline (Tc). Results are shown as the mean ± SD ($n = 3$); $^{****}P < 0.0001$; $^{***}P < 0.001$; $^{**}P < 0.01$ compared with cells without tetracycline (ANOVA with the Geisser–Greenhouse correction for multiple comparisons). (**E**) Pie chart showing the classification of genes that are upregulated in XY $Sox9^{\Delta/\Delta}$ testis versus XY WT in genes bound or not by SOX9 (mouse ChIP-seq dataset). Only 8% of these SOX9-bound genes are down-regulated at E13.5 in wild type Sertoli (35%) versus granulosa cells (64).

Table 1. Most genes that have been involved in mammalian gonad determination and differentiation are SOX9 targets in mouse and cattle

Genes (Mouse/Bovine)	Nb of peaks in mouse	Nb of peaks in bovine	Bovine peaks orthologous to mouse peaks
Amh /AMH	1	1	0/1
Atrx/ATRX	1	1	1/1
Bmp2/BMP2	1	2	1/2
Cbx2/CBX2	1	1	1/1
Cited2/CITED2	1	0	-
Ctnnb1/CTNNB1 (ß Catenin)	1	0	-
Cyp26b1/CYP26B1	0	3	-
Dhh/DHH	1	1	0/1
Dmrt1/DMRT1	1	0	-
Emx2/EMX2	1	3	1/3
Fgf9/FGF9	2	3	2/3
Fgfr2/FGFR2	1	1	0/1
FoxL2/FOXL2	2	2	0/1
Fst/FST	4	1	1/1
Gadd45g/GADD45G	1	1	1/1
Gata4/GATA4	3	3	1/3
Hhat/HHAT	1	1	0/1
Igf1r/IGF1R	1	3	0/3
Insr/INSR	1	2	1/2
Kdm3a/KDM3A (Jmjd1a/JMJD1A)	1	1	0/1
Lhx9/LHX9	2	5	3/5
Lmo4/LMO4	2	3	2/3
Map3K4/MAP3K4	0	0	-
Notch1/NOTCH1	1	0	-
Notch2/NOTCH2	1	1	0/1
Notch3/NOTCH3	0	2	-
Nr0b1/NR0B1(Dax1/DAX1)	1	1	1/1
Nr5a1/NR5A1 (Sf-1/SF-1)	2	1	0/1
Pdgfa/PDGFA	2	0	-
Ptgds/PTGDS	0	1	-
Rspo1/RSPO1	0	0	-
Six1/SIX1	1	2	1/1
Six4/SIX4	1	1	0/1
Sox8/SOX8	0	1	-
Sox9/SOX9	2	3	2/3
Sox10/SOX10	1	1	1/1
Sry/SRY	0	0	-
Tcf21/TCF21 (Pod1)	1	4	0/4
Wnt4/WNT4	1	0	-
Wt1/WT1	1	2	0/2
Wwox/WWOX	1	4	0/4
Zfpm2/ZFPM2(Fog2/FOG2)	1	5	2/5

The number of peaks per gene are shown in both species. For each gene is shown the number of bovine peaks having orthologous corresponding peaks in mouse.

(Supplementary Table S2: Sheet 'KO versus WT XY, adjusted $P < 0.05$'). Among the downregulated genes, 40% (95/240) were bound (based on our ChIP-seq experiment results) and presumably activated by SOX9 (Figure 2B and Supplementary Table S2: Sheet 'summary'). Therefore, the 240 genes downregulated in $Sox9^{\Delta/\Delta}$ testes were highly enriched in SOX9-bound genes (χ^2 test with Yates correction: $P < 0.0001$). Moreover, most of the SOX9-bound, downregulated genes (65 out of 95) were normally more strongly expressed at E13.5 in Sertoli cells than in granulosa cells (their female counterpart) (64) (see also http://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS3995) (Figure 2B and Supplementary Table S2: Sheet 'summary'). This makes of them strong candidates for positive and direct transcriptional regulation by SOX9. Comparison of the expression level of these 65 genes in the WT XY, $Sox9^{\Delta/\Delta}$ XY and WT XX gonad datasets allowed their classification into three groups. The first group represented genes whose expression in the XY mutant was intermediate between WT XY and XX gonads (Figure 2C, red circles), suggesting that in the absence of SOX9, these genes are still partially active. This group included genes with an important role in sexual differentiation, such as *Amh* (encoding anti-Müllerian hormone) (68), *Sox10* (10) or *Dhh* (69) (Supplementary Figure S2B, upper panels). The second group (Figure 2C, yellow squares) included genes that were similarly expressed in XY $Sox9^{\Delta/\Delta}$ and WT XX gonads (e.g. *Rnd2*, *Mybl1* or *Gjb2*) (Supplementary Figure S2B, middle panels). This indicates that SOX9 regulates only the male-specific expression of these genes. The third group of genes displayed lower expression in XY $Sox9^{\Delta/\Delta}$ than in WT XX gonads (Figure 2C, green triangles and Supplementary Figure S2B, lower panels), suggesting that in E13.5 WT XX gonads, these genes are regulated by factors that are absent in XY $Sox9^{\Delta/\Delta}$ testes.

The finding that some SOX9-bound genes that were upregulated in E13.5 Sertoli cells (Figure 2A) were not downregulated in XY $Sox9^{\Delta/\Delta}$ gonads could be explained by

functional redundancy between SOX9 and SOX8. To verify that SOX9 could activate these genes, we used NT_2D_1 cells, a human cell line that displays some characteristics of Sertoli cells (70), to generate a line in which HA-tagged SOX9 expression can be induced by tetracycline (tet-SOX9-NT_2D_1 cells) (Supplementary Figure S2D shows the strong tetracycline-dependent expression of HA-tagged SOX9). We then used RT-qPCR analysis of tet-SOX9-NT_2D_1 cells before and after SOX9 induction to quantify the transcripts of genes that were bound by SOX9 in our ChIP-seq experiments (Figure 2D). Genes bound by SOX9 that were upregulated in E13.5 Sertoli cells and downregulated in E13.5 XY $Sox9^{\Delta/\Delta}$ testes (strong criteria to consider them as activated by SOX9 in Sertoli cells) were also upregulated upon induction of SOX9 by tetracycline (*AMH, DHH, NEDD9, SERP2, SOX10*) (dark blue bars in Figure 2D). Moreover, the expression of genes identified as SOX9 targets that were only upregulated in E13.5 Sertoli cells (light blue bars: *DMRT1*, *ETV1, SOX6*) also was increased by SOX9 expression in tet-SOX9-NT_2D_1 cells. As control, *IP08,* a SOX9-bound gene but without sexually dimorphic expression in the supporting lineage, was not upregulated by SOX9 induction in tet-SOX9-NT_2D_1 cells. These findings indicate that in foetal testes, SOX9 can bind to and upregulate the expression of genes that are only expressed in male supporting cells (i.e. Sertoli cells).

Finally, the finding that 107 genes were upregulated in KO XY gonads compared with WT XY gonads suggested that SOX9 could repress them in normal foetal testis, directly or indirectly. However, only 22% (23/107) of these genes (χ^2 test with Yates correction, $P = 0.94$) were bound by SOX9 and only eight were also upregulated in E13.5 female supporting cells compared with male supporting cells (Figure 2E and Supplementary Figure S2C). This suggests that genes repressed by SOX9 are less likely to be bound by SOX9 than genes activated by SOX9.

Role of SOX9 in RNA processing in foetal testis

SOX proteins could have a role in splicing regulation (71) and a direct link between SOX9 and the RNA processing of *Col2a1* (one of the SOX9 targets identified in our bovine ChIP-seq experiment) was demonstrated in chondrocytes (72). Therefore, we used our RNA-seq dataset to investigate SOX9 role in splicing of its target genes in foetal testis. Splicing of 154 mRNAs, of which 70 (45%) were bound by SOX9, was quantitatively different in $Sox9^{\Delta/\Delta}$ XY and WT XY testes (Figure 3A). All five classes of known alternative splicing events (73) were affected by SOX9 absence: exon cassette (also known as exon skipping), intron retention, alternative 3′ acceptor site, alternative 5′ donor site and mutually exclusive exons. Quantitative analysis of the WT XY and $Sox9^{\Delta/\Delta}$ XY RNA-seq datasets allowed calculating the Percent-Spliced-In (PSI) index that indicates the splicing efficiency for transcripts and can range from +1 (totally include) to –1 (totally exclude). The PSI values for transcripts of genes bound by SOX9 indicated that splicing efficiency differed between the WT XY and $Sox9^{\Delta/\Delta}$ XY datasets (Figure 3A; for details see Supplementary Table S3: Sheet 'XY WT versus XY KO'). Specifically, SOX9 absence affected the splicing of genes with important roles in sex determination and differentiation, such as *Atrx* (74), *Fgfr2* (75) and *Gadd45g* (52). Importantly, in 20 of the 70 differentially spliced genes bound by SOX9 (28%), SOX9 peaks overlapped with the exon/intron boundary involved in the splicing events (gene names underlined in Figure 3A, and examples in Figure 3B). Analysis of somatic and germ cell-specific transcriptomic datasets from foetal male and female mouse gonads (64) showed that the expression of SOX9-bound genes with alternatively spliced transcripts was not particularly sex- or cell-specific (not shown). This suggests that SOX9 may regulate some discrete molecular mechanisms that require finer transcriptomic analyses to be revealed.

We next investigated SOX9 mediated sex-specific splicing by searching for common splicing events between WT XY and $Sox9^{\Delta/\Delta}$ XY and between WT XY and WT XX datasets. Compared with WT XY gonads, several transcripts showed differential splicing events that involved the same exons in $Sox9^{\Delta/\Delta}$ XY and WT XX gonads (Figure 3C). Their respective PSI values were always in a similar range when comparing WT XY with $Sox9^{\Delta/\Delta}$ XY or WT XX gonads (Supplementary Table S3: Sheets 'XY WT versus XY KO' and 'XY WT versus XX WT'). This indicates that splicing of these exons is similar in $Sox9^{\Delta/\Delta}$ XY and WT XX gonads. We conclude that SOX9 regulates male-specific splicing of several of its target genes.

In all cases, these splicing events will modify the resulting protein sequences. Moreover, intron retention can also modify the intrinsic level of the translated proteins. For instance, in XX WT and $Sox9^{\Delta/\Delta}$ XY, retention of the second intron of the *Ubfd1* gene leads to the disruption of the open reading frame and to a shorter truncated protein of 204 amino acids instead of full length 368 amino acids UBFD1 (Figure 3B, panel Ubfd1). Moreover, at the transcriptional level, published transcriptomic datasets (64) and our RNA-seq data showed no variation in *Ubfd1* gene expression between male and female gonads, suggesting that SOX9 could control the protein level of active UBFD1 via a discrete splicing mechanism that does not affect the transcript level. Similarly, the second intron of *Gadd45g*, which has a role in sex-determination (52,76), was retained in $Sox9^{\Delta/\Delta}$ XY testes, leading to a shorter truncated protein of 52 amino acids instead of full length GADD45g (159 amino acids in length). As for the *Ubfd1* transcript, *Gadd45g* mRNA level at E13.5 was not significantly different between male and female E13.5 gonads (64).

Supporting these findings on splicing, we observed that induction of SOX9 expression in tet-SOX9-NT_2D_1 cells modified the ratio of splicing events for *FGFR2* and *ENAH* transcripts as observed in $Sox9^{\Delta/\Delta}$ XY / WT XY mouse foetal gonads (Figure 3D).

SOX9-bound genomic regions in foetal testes are enriched in binding motifs for three Sertoli cell differentiation/reprogramming factors

Our ChIP-seq analysis using chromatin from mouse and bovine foetal testes identified several regions (4293 peaks) that were bound by SOX9 in both species (Figure 1B). As DMRT1, GATA4, WT1, SF1 and SOX9 can reprogram embryonic fibroblasts into embryonic Sertoli cells (18), we

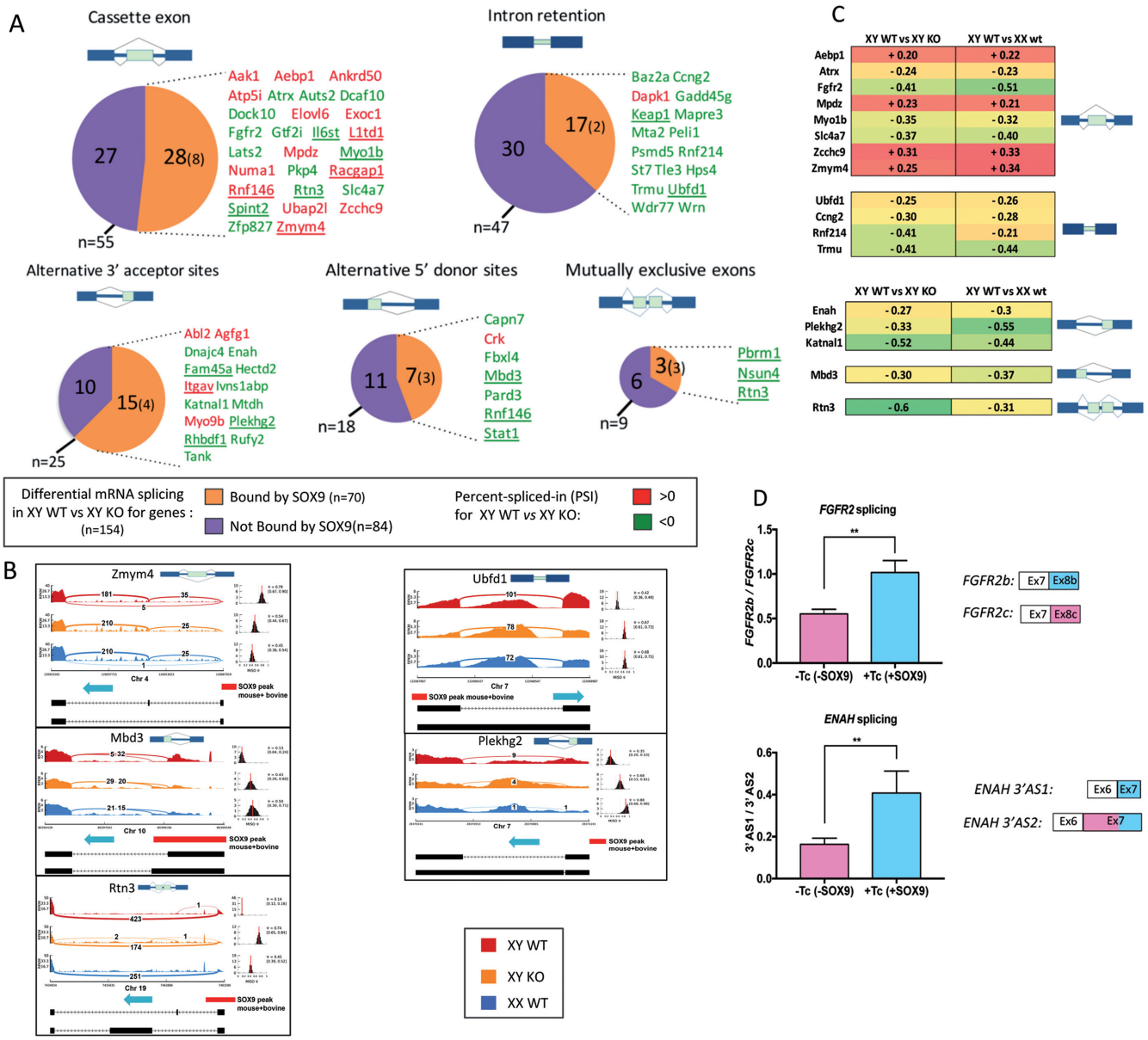


Figure 3. Analysis of differential splicing events in Sox9$^{\Delta/\Delta}$ XY gonads. (**A**) Pie charts for each of the splicing events deregulated in KO XY gonads compared with WT XY gonads. Transcripts of SOX9-bound genes (orange area) are listed in green when their Percent-Spliced-In (PSI) index is negative and in red when the PSI index is positive. For a splicing event in a given condition, PSI is between 0 (splicing event totally excluded) and 1 (splicing event totally included). The PSI values reported here correspond to the PSI value difference between genotypes. Underlined names represent genes where SOX9-bound peaks overlapped with the exons/introns involved in the splicing event; their total number is indicated between brackets in the relevant pie-chart. (**B**) Sashimi plots showing examples of the splicing events altered in *Sox9*$^{\Delta/\Delta}$ testes and described in Figure 3A, obtained from RNA-seq datasets. These transcripts correspond to genes that are bound by SOX9 in mouse *or* mouse and cattle, with ChIP-seq peaks in the vicinity of regions of splicing events (genes underlined in Figure 3A). Red bars, position of SOX9 peaks; blue arrows, direction of transcription. For each genotype (WT XY, KO XY and WT XX gonads), the quantification of the PSI value (Ψ) with 95% confidence intervals (histograms on the on right side of each panel) obtained using MISO are reported. Values reported on Sashimi plots correspond to the number of sequencing reads per splice junction. (**C**) Comparison of transcripts displaying differential splicing in WT XY versus KO XY gonads and in WT XY versus WT XX gonads. The table lists transcripts with the same differential splicing in the two situations. The color scale goes from red (inclusion of the splicing event; PSI>0) to green (exclusion of the splicing event; PSI<0). (**D**) RT-qPCR analysis of *FGFR2* and *ENAH* transcript splicing ratio using mRNA samples from tet-SOX9-NT$_2$D$_1$ cells in which SOX9 expression was induced (+Tc) or not (–Tc). Data are the mean ± SD ($n = 3$); $^{**}P < 0.01$ compared with the –Tc condition (Student's t test).

asked whether conserved binding motifs for these factors could be present in the vicinity of the identified SOX9 binding sites. Therefore, we developed a normalization bioinformatics method to detect sequences enriched in known consensus DNA motifs (41,43,77–79) recognized by the five Sertoli cell differentiation factors in the mouse and bovine SOX9 ChIP-seq datasets. The motif enrichment fold was the ratio of the number of binding motifs found in the ChIP-seq datasets versus the number found in the corresponding scrambled sequences (30 sets). The mouse and bovine ChIP-seq datasets were similarly enriched in DNA-binding motifs for SOX9, DMRT1 and GATA4 (i.e. the ratio ChIP/scrambled was >1) (Figure 4A, red and orange bars, left histogram) compared with unrelated genomic regions (blue bars). Conversely, SF1 motifs were not enriched and WT1 motifs were non-specifically enriched in the ChIP-seq datasets as well as in the unrelated genomic regions. As a control, binding motifs of transcription factors not related to Sertoli cell differentiation were not enriched in the two ChIP-seq datasets (Figure 4A, right histogram). SOX9/DMRT1/GATA4 binding motif enrichment was neither due to repetitive sequences (Supplementary Figure S3A) nor dependent on the position of the peaks relative to the gene body (Supplementary Figure S3B). Thus, SOX9 target sequences are enriched in binding motifs for SOX9, GATA4 and DMRT1. The same motif enrichment was also found in the orthologous regions of twelve other mammal genomes (Supplementary Figure S3C), suggesting that this genomic feature is conserved among mammals.

To determine whether this global motif enrichment reflected a situation existing at each genomic fragment bound by SOX9, we investigated the co-occurrence of SOX9, GATA4, DMRT1, SF1 and WT1 binding motifs in both SOX9 ChIP-seq datasets. The number of occurrences of DNA motifs for the five Sertoli cell differentiation/reprogramming factors was plotted against each other to assess whether a linear correlation existed between motifs (examples of scatterplots with Pearson correlation coefficients, PCC, are presented in Supplementary Figure S3D). The PCC obtained for the ChIP-seq datasets and control scrambled sequences showed a strong correlation (PCC between 0.7 and 0.9) among SOX9, GATA4 and DMRT1 motifs in both mouse and bovine ChIP-seq datasets (Figure 4B). By contrast, SF1 motifs were poorly represented (Figure 4B and Supplementary Figure S3D; PCC: 0.30–0.37) and WT1 motifs showed almost no correlation with the other four motifs. No correlation was observed between SOX9 motifs and TF motifs (negative control) and for any combination with the corresponding scrambled sequences. The co-occurrence of SOX9, GATA4 and DMRT1 DNA-binding motifs was also observed in the orthologous regions of SOX9 ChIP-seq datasets from the twelve mammalian genomes (Figure 4C). This indicates that the association of SOX9, GATA4 and DMRT1 DNA-binding motifs in these genomic regions is conserved among mammals. Compilation of the three motifs enriched in both ChIP-seq datasets showed that they were highly similar in rodent and bovine species (Supplementary Figure S3E), but displayed a milder selectivity than the consensus motifs obtained by *in vitro* selection for SOX9 (43), GATA4 (77) and DMRT1 (41).

Characterization of a Sertoli cell signature (SCS) in SOX9 genomic targets in foetal testis

To determine whether this SOX9, GATA4 and DMRT1 DNA-binding motif enrichment at SOX9-bound genomic regions was unique to foetal Sertoli cells, we analysed SOX9, GATA4 and DMRT1 binding motif enrichment in randomly selected ChIP-seq datasets for 13 transcription factors and eight SOX factors expressed in various tissue and cell types (http://www.ncbi.nlm.nih.gov/geo/). Compared with the mouse SOX9 ChIP-seq foetal testis dataset, we did not detect any motif enrichment in ChIP-seq datasets for the unrelated transcription factors (Figure 4D) and other SOX factors (Figure 4E).

Analysis of SOX9 ChIP-seq datasets from seven other mouse tissues highlighted a slight SOX9/DMRT1/GATA4 binding motif enrichment in rib chondrocytes (80), atrioventricular canal (AVC), limb (81) and hair follicle stem cells (HF-SCs) (82) (Figure 4F, blue bars). This enrichment was similar to the one observed for other transcription factors and SOX factors. On the other hand, in mouse basal cell carcinoma (BCC) (83) and human colorectal cancer cells (84), SOX9/DMRT1/GATA4 binding motif enrichment was comparable to the one observed in foetal testis, but for the DMRT1 motif (Figure 4F, green bars). After conversion of the coordinates of the relevant human regions to the mouse orthologous regions, we found an important peak overlap between our SOX9 ChIP-seq testis datasets and the colorectal cancer cell and BCC datasets, but not with the rib chondrocyte, limb, AVC and HF-SC SOX9 ChIP-seq datasets (Supplementary Figure S4). This shows that, at least in these two cancer types, SOX9 binds to regions with genomic features very similar to those observed in differentiated foetal Sertoli cells.

We named this genomic feature the 'Sertoli Cell Signature' (SCS) because we detected enrichment for SOX9/DMRT1/GATA4 binding motifs specifically in genomic regions bound by SOX9 in Sertoli cells from foetal testis.

Binding motifs in the SCS are clustered and organized

We then asked whether the SOX9/DMRT1/GATA4 binding motifs were organized within ChIP-seq peaks. In most of the cases, adjacent binding motifs were separated by less than 50 bp both in the mouse and bovine ChIP-seq datasets, indicating that they clustered together (Supplementary Figure S5A). By analyzing nucleotide spacing between motifs in both murine and bovine datasets (Supplementary Figure S5B) we observed that the distance between adjacent motifs was lower than ten nucleotides and remarkably conserved from mouse to cattle in term of distance and orientation. For instance, SOX9 motifs were preferentially on the same strand and separated by one or three nucleotides. Altogether, our results show that SOX9/DMRT1/GATA4 motifs are strongly clustered in both species. To confirm this, the distance of each motif relative to the nearby TSS was measured in peaks positioned

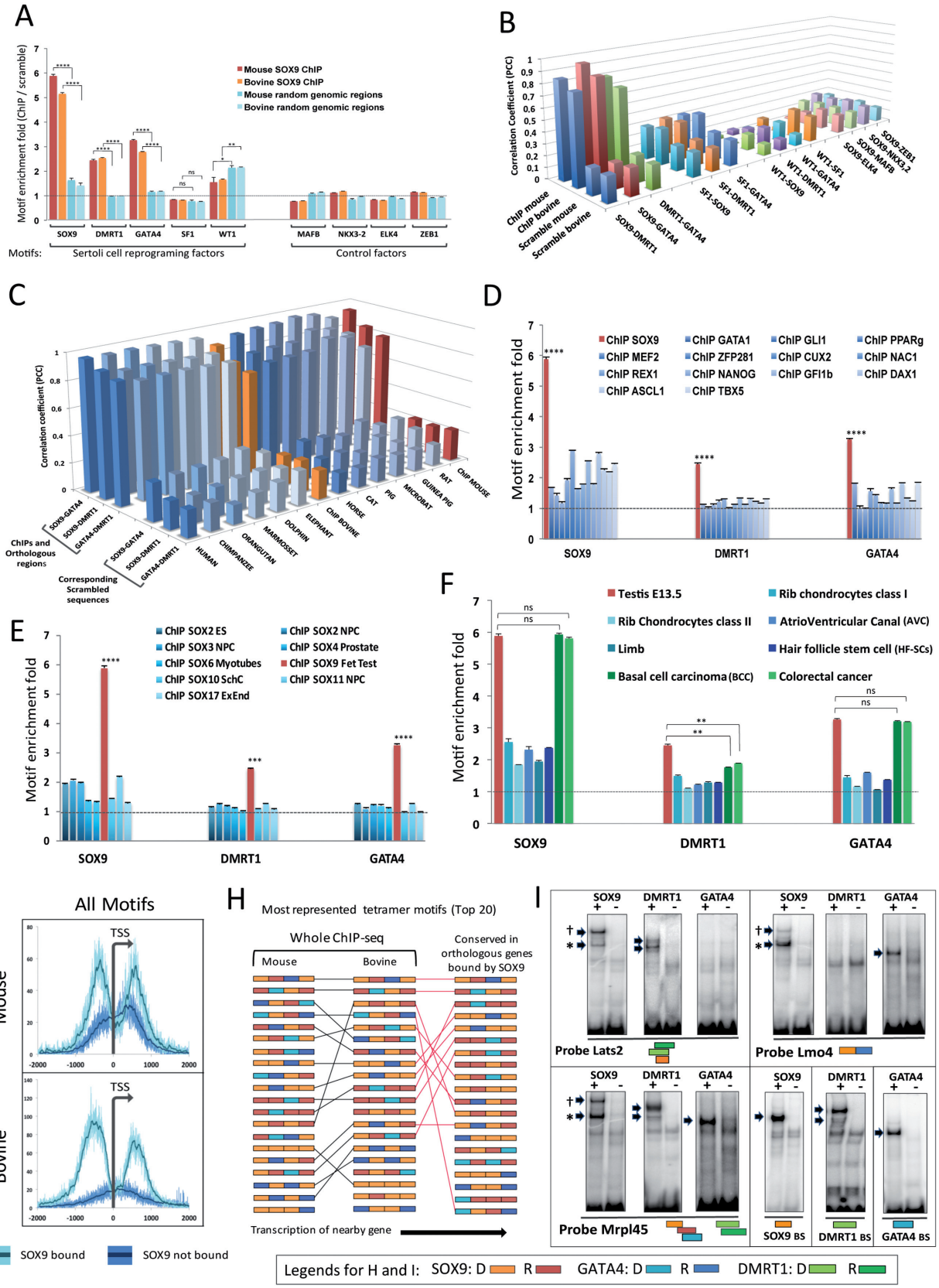


Figure 4. In foetal testis, SOX9 binds to genomic regions enriched in SOX9, DMRT1 and GATA4 binding motifs. (**A**) Global enrichment of binding motifs for the five Sertoli cell differentiation/reprograming factors (SOX9, GATA4, DMRT1, SF1 and WT1) and four unrelated transcription factors (control factors) in the mouse and bovine SOX9 ChIP-seq datasets (red and orange bars). Randomly selected mouse and bovine genomic regions (blue

near the TSS (i.e. from –1 kb to +1 kb around the TSS). The distribution of the three binding motifs analysed together (Figure 4G) or separately (Supplementary Figure S5C) followed a bimodal pattern around the TSS of both mouse and bovine SOX9 target genes, but not of unrelated genes (not bound by SOX9). This shows that in the mouse and cattle, SOX9/DMRT1/GATA4 binding motifs cluster and that for peaks around TSS, SCSs are located preferentially at –500/400 bp or at +600 bp from the TSS.

Next, we asked whether the three binding motifs showed a preferential order within SCSs and whether this order was conserved in mouse and bovine orthologous regions bound by SOX9. To this aim, we counted the occurrence of every possible multimer of each binding motif (from trimers to octamers) in both ChIP-seq datasets, taking in account their direct (D) or reverse orientation (R) relative to the direction of transcription of the neighboring gene. Analysis of the occurrence of motif tetramers (for detailed analysis see Supplementary Figure S6A and B) showed that among the 1296 possible tetramer combinations, 16 of the 20 most represented tetramers (top 20) were common between mouse and bovine SOX9 target regions (Figure 4H, black lines in left panels). This demonstrates the presence of a common, preferential organization of the motifs composing the SCS in both species. Of note, the two most represented tetramers were the same in both species, but in the reverse orientation. However, among the 3849 orthologous genes bound by SOX9 in both species (Figure 1B), only 670 (17.4%) displayed a strictly conserved organization of SCS tetramer motifs. Among these genes, we ranked the 20 most represented tetramers (Figure 4H, right panels and Supplementary Figure S6C) and found that 12 belonged to the previously identified top 20 tetramers (Figure 4H, red lines).

Surprisingly, in both mammals, the top 20 tetramers included only SOX9 and GATA4 motifs in both orientations, and the first tetramers that included a DMRT1 motif were ranked at positions below 50. This suggests that in regions bound by SOX9, the position of DMRT1 motifs, even if strongly correlated with either SOX9 or GATA4 motifs (see Figure 4B), is less constrained than that of the two other motifs. In summary, this analysis highlighted a strong preferential SOX9/DMRT1/GATA4 motif organization in mouse and bovine SCSs, although orthologous genes displayed only a relatively low conservation of the motif organization. Among the three motifs, SOX9 and GATA4 showed the strongest association within SCSs.

To test whether the predicted DNA motifs present in SCS were bound by SOX9, DMRT1 and GATA4, we performed *in-vitro* gel shift assays using *in-vitro* translated proteins and different probes with various combinations of motifs derived from SOX9 peaks harbouring SCS (Figure 4I) or with the SOX9, DMRT1 and GATA4 consensus binding motifs (controls; panels 'BS' in Figure 4I). SOX9 and DMRT1, but not GATA4, could bind to the *Lats2* probes (only DNA motifs for SOX9 and DMRT1). The *Lmo4* probe (SOX9 and GATA4 motifs) interacted only with SOX9 and GATA4, but not with DMRT1. Finally, the *Mrpl45* probe (all three motifs) bound to all three factors. In each case, we observed always two distinct complexes for SOX9, suggesting that SOX9 could binds either as monomeric (*) or dimeric (†) forms, in contrast with the BS probe that contains the *in vitro* defined SOX9 consensus motif (panel SOX9 BS on Figure 4I). However, we did not observe a clear preference for monomeric or dimeric SOX9, it depends of the probe used. Indeed, SOX9 bound preferentially as a homodimer to the *Lats2* probe, and as a monomer to the *Lmo4* and *Mrpl45* probes. Interestingly, SOX9 could bind as a homodimer also to a probe with a single predicted motif (probe *Lats2*). We also observed at least two protein-DNA complexes for DMRT1, as already published (41), and only one for GATA4. We conclude that the *in silico* predicted DNA motifs of the SCS are effectively bound, at least *in vitro*, by SOX9, DMRT1 and GATA4.

Genes bound by DMRT1 in foetal testis harbour SCSs only if they are also bound by SOX9

In foetal testis, *Dmrt1* is expressed in Sertoli and germ cells (85). As DMRT1 motifs were highly correlated with SOX9 motifs within SCSs, we hypothesized that these two transcription factors have common target genes in foetal Sertoli cells. We first confirmed that both SOX9 and DMRT1 were co-expressed in the nucleus of Sertoli cells at E13.5 (Figure 5A). Then, to determine whether SOX9 and DMRT1 had common target genes, we compared E13.5 mouse DMRT1

←

bars) were used as controls. (**B**) Pearson correlation coefficients (PCC; Y axis) for the indicated combinations of DNA binding motifs in the mouse and bovine ChIP-seq datasets and scrambled sequences. (**C**) PCC (Y axis) for the three combinations of DNA binding motifs (SOX9-GATA4; SOX9-DMRT1 and GATA4-DMRT1) in the mouse (red bars), bovine (orange bars) ChIP-seq datasets and in orthologous regions from twelve other mammals (blue bars). Genome builds are detailed in Supplementary Figure S3C. (**D**) SOX9, DMRT1 and GATA4 binding motif enrichment in ChIP-seq datasets for transcription factors (GATA1, GLI1, PPARg, MEF2, ZFP281, CUX2, NAC1, REX1, NANOG, GFI1b, ASCL1 and TBX5) expressed in various tissue and cell types. (**E**) SOX9, DMRT1 and GATA4 binding motif enrichment in ChIP-seq for other SOX factors (SOX2, SOX2, SOX3, SOX4, SOX6, SOX10, SOX11 and SOX17). ES, ES cells; NPC, neural progenitor cells; SchC, Schwann cells; ExEnd, extra-embryonic endoderm; Fet Test, foetal testis. (**F**) SOX9, DMRT1 and GATA4 binding motif enrichment in SOX9 ChIP-seq datasets from various tissues. For (D), (E) and (F), see description of tissue/cell types used to obtain the datasets in Supplementary Tables S4 to S6. For panels (A), (D), (E), (F): data are the mean ± SD; $****P < 0.0001$; $***P < 0.001$; $**P < 0.01$; ns: not significant (ANOVA with the Geisser–Greenhouse correction for multiple comparisons). (**G**) Graph showing for peaks close to the TSS in the mouse and bovine SOX9 ChIP-seq datasets, the averaged curves of the localization of the SOX9, GATA4 and DMRT1 binding motifs relative to the TSS. X-axis: nucleotide distance from the TSS. Y-axis: count of the three motifs. (**H**) The 20 most represented tetramer motifs among peaks bound by SOX9 in the mouse (left) and bovine (middle) datasets. On the right: the top 20 conserved tetramers among orthologous genes bound by SOX9 in the mouse and bovine ChIP-seq datasets. Motifs are in the direct (D) or reverse (R) orientation relative to the transcription of the nearby gene (black arrow). Black and red lines show identical tetramers in the two mammalian species. (I) Electrophoretic mobility shift assays showing binding of *in vitro* translated SOX9, DMRT1 and GATA4 proteins to DNA probes derived from mouse ChIP-seq peaks assigned to the *Lats2*, *Lm04* and *Mrpl45* genes and containing SCSs. +/ indicate programmed and un-programmed lysates. Arrows indicate protein/DNA complexes. †: SOX9 dimers; *: SOX9 monomers; BS, control probes with the consensus binding motifs for SOX9, DMRT1 and GATA4. Coloured bars indicate the direct (D) or reverse (R) orientation of motifs in each DNA probe.

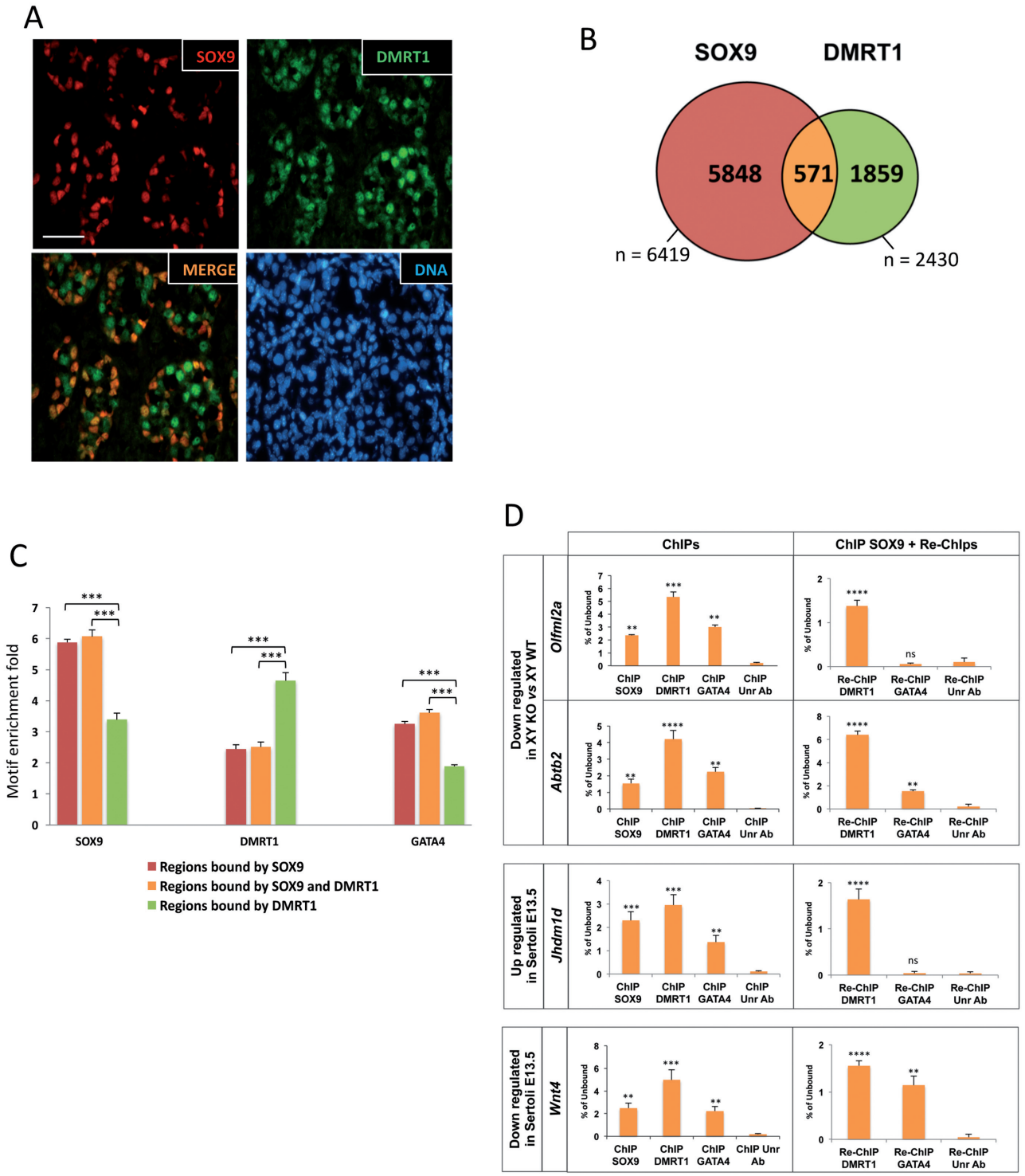


Figure 5. Genes bound by DMRT1 in foetal testis have SCSs only if they are bound also by SOX9. (**A**) Immunofluorescence experiments with antibodies against SOX9 (red) and DMRT1 (green) on sections of E13.5 wild type XY gonads. Merged view shows Sertoli cells that express both SOX9 and DMRT1 in the nucleus (orange). DNA was stained with DAPI (blue). Scale bar, 30 μm. (**B**) Venn diagram showing the genomic overlap between peaks bound both by SOX9 and DMRT1 from independent SOX9 and DMRT1 ChIP-seq datasets from E13.5 foetal testes. (**C**) SOX9, DMRT1 and GATA4 binding motif enrichment in regions bound by SOX9 only (red bars), SOX9 and DMRT1 (orange bars) and DMRT1 only (green bars). (**D**) ChIP-qPCR (left panels) and sequential ChIP-qPCR (right panels) assays of sheared chromatin from E13.5 foetal mouse testes. For ChIPs, chromatin samples were immunoprecipitated in parallel with anti-SOX9, -DMRT1 or -GATA4 antibodies. For sequential ChIPs, samples were first immunoprecipitated with an anti-SOX9 antibody and then with an anti-DMRT1 or anti-GATA4 antibody. The analysed genomic regions correspond to overlapping SOX9 and DMRT1 ChIP-seq peaks. *Olfml2a* and *Abtb2* are genes upregulated in E13.5 Sertoli cells versus granulosa cells and downregulated in XY KO gonads. *Jhdm1d* and *Wnt4* are genes up- and down-regulated, respectively, in E13.5 Sertoli cells versus granulosa cells. Data are the mean ± SD ($n = 3$); $^{****}P < 0.0001$; $^{***}P < 0.001$; $^{**}P < 0.01$ (ANOVA with the Geisser–Greenhouse correction for multiple comparisons).

(25) and SOX9 ChIP-seq datasets and found that among the 2430 and 6419 respective peaks, 671 overlapped (Figure 5B). As a control, 30 sets of 2430 regions of equivalent size and randomly selected in the mouse genome displayed very little overlap (<10 regions) with the 6419 peaks of the SOX9 ChIP-seq dataset (not shown). From these 571 common regions, we also identified 411 (71%) orthologous regions in the bovine SOX9 ChIP-seq dataset (not shown).

We then analyzed separately the SCS content of the two populations of peaks in the DMRT1 ChIP-seq dataset (i.e. regions bound or not by SOX9). Overall, regions bound by DMRT1 displayed the same global enrichment for SOX9, DMRT1 and GATA4 binding sites as observed in the SOX9 ChIP-seq dataset. However, in peaks bound only by DMRT1, SOX9 and GATA4 sites were not as enriched as DMRT1 binding sites (Figure 5C). Similar results were obtained by calculating the PCCs (Supplementary Figure S7). From these analyses, we conclude that regions bound by DMRT1 in foetal testis display SCSs only when they are also bound by SOX9.

On the basis of these results, we hypothesized that on Sertoli cells chromatin, DMRT1 and GATA4 could cohabit in proximity of SOX9 on it target genes. To verify this hypothesis, we performed ChIP and sequential ChIP-qPCR assays using E13.5 foetal testis chromatin extracts. For sequential ChIP (also known as ChIp re-ChIP), we took advantage of the fact that in foetal testis, SOX9 is expressed only in Sertoli cells, while DMRT1 and GATA4 are expressed in Sertoli cells and also in germinal (Figure 5A) and interstitial compartment cells, respectively (Supplementary Figure S7B). Therefore, we first immunoprecipitated chromatin with an anti-SOX9 antibody to retain only chromatin from Sertoli cells, and after immune complex dissociation, we performed the second immunoprecipitation with an anti-DMRT1 or GATA4 antibody. For qPCR analysis, we chose probes that amplify the following four regions (among the 571 overlapping peaks in SOX9 and DMRT1 ChIP-seqs; orange area and bars in Figure 5B and C): the *Olfml2a* and *Abtb2* genes that are downregulated in $Sox9^{\Delta/\Delta}$ XY testes, and the *Jhdm1d* and *Wnt4* genes that are up- and down-regulated, respectively, in E13.5 Sertoli cells compared with granulosa cells (64). Single ChIP assay showed that the four genes were bound by SOX9, DMRT1 and GATA4 in E13.5 whole testis extracts (Figure 5D, left panels). Sequential ChIPs (Figure 5D, right panels) showed that in Sertoli cells, *Olfml2a*, *Abtb2, Jhdm1d* and *Wnt4* were bound by SOX9 and DMRT1, while only *Abtb2* and *Wnt4* were bound by SOX9, DMRT1 and GATA4. The fact that the genes *Olfml2a* and *Jhdm1d* were immunoprecipitated in the GATA4-ChIP, but not by SOX9-ChIP with GATA4-re-ChIP is explained by the expression of GATA4 by non-Sertoli cells (i.e. SOX9-negative cells) (Supplementary Figure S7B). As control, ChIP-seq peaks from genes that are strongly expressed in germ cells and bound only by DMRT1 (green circle and bars in Figure 5B and C), like *Cdk12* and *Nek3,* were not bound by SOX9 (not shown).

Our results show that, in SOX9 target genomic regions, DMRT1 and GATA4 bind near SOX9 in foetal Sertoli cells.

SCS scanning identified TRIM28 as a new SOX9 partner in foetal testis

As regions bound by DMRT1 with a SCS were also bound by SOX9 in foetal testes, we hypothesized that by searching for SCSs in ChIP-seq datasets of various unrelated transcription factors we might identify new proteins that contribute to SOX9-mediated transcriptional regulation in Sertoli cells. By analyzing ChIP-seq datasets for 50 different transcription factors obtained from GEO, we identified a SCS in the dataset for the nuclear factor TRIM28 in mouse embryonic stem (ES) cells (TRIM28 ChIP-seq dataset) (86) (Figure 6A) where 1501 peaks (44.5% of 3370) overlapped with those of the mouse SOX9 ChIP-seq dataset (Figure 6B). As this finding indicates that TRIM28 can bind to the same genomic regions as SOX9, the two might physically interact. Co-immunofluorescence experiments showed that TRIM28 and SOX9 were both expressed in the nucleus of foetal Sertoli cells (Supplementary Figure S8A) where they co-localize, as indicated by super-resolution three-dimensional structured-illumination (3D-SIM) microscopy with z-sections of 0.12 μm (87) (Figure 6C, white arrows). Moreover, two different antibodies against the N- or C-terminal moieties of SOX9 could co-immunoprecipitate TRIM28 with SOX9 in NT_2D_1 cells (Supplementary Figure S8B). Pull-down assays with purified recombinant proteins showed that GST-TRIM28 interacted directly with SOX9 (Figure 6D). SOX9 and TRIM28 were also co-immunoprecipitated in nuclear extracts from E13.5 mouse foetal testes, but not from foetal ovaries that do not express SOX9 (Figure 6E). This confirms that this interaction also occurs in foetal Sertoli cells. To test whether SOX9 and TRIM28 co-localized on foetal Sertoli cell chromatin, we then performed sequential ChIP-qPCR experiments. Among the 1501 genes that are bound by SOX9 and TRIM28 (orange in Figure 6B), we chose *Sox13* and *Sox10,* two genes that are downregulated in $Sox9^{\Delta/\Delta}$ XY gonads, and *Fgf18* and *Wnt4* that are up- and down-regulated, respectively, in E13.5 Sertoli cells compared with granulosa cells (64). Single ChIP-qPCR (Figure 6F, left panels) showed that both SOX9 and TRIM28 bound to these genes in E13.5 testes, and sequential ChIP (ChIP with anti-SOX9 antibody followed by re-ChIP with an anti-TRIM28 antibody) gave similar results. This demonstrates that, in Sertoli cells, for these SOX9 target genes TRIM28 binds to chromatin near SOX9.

To investigate whether SOX9 and TRIM28 could collaborate on transcriptional regulation, we transduced HEK293 human embryonic kidney cells with a tet-SOX9 lentiviral vector (tet-SOX9-HEK293 cells), as previously done in NT_2D_1 cells. We chose HEK293 cells for this experiment, rather than NT_2D_1 cells, because of their low endogenous SOX9 and TRIM28 expression levels. Tetracycline-induced expression of the *SOX9* transgene strongly induced the expression of endogenous *SOX10* compared with cells without tetracycline. This effect was significantly increased in tet-SOX9-HEK293 cells that overexpressed also TRIM28 (Figure 6G, upper panel). In agreement, SOX9-mediated downregulation of *EMX2,* a gene that is strongly repressed in differentiating foetal Sertoli cells (Figure 6F), also was significantly increased by co-expression of TRIM28 (Figure

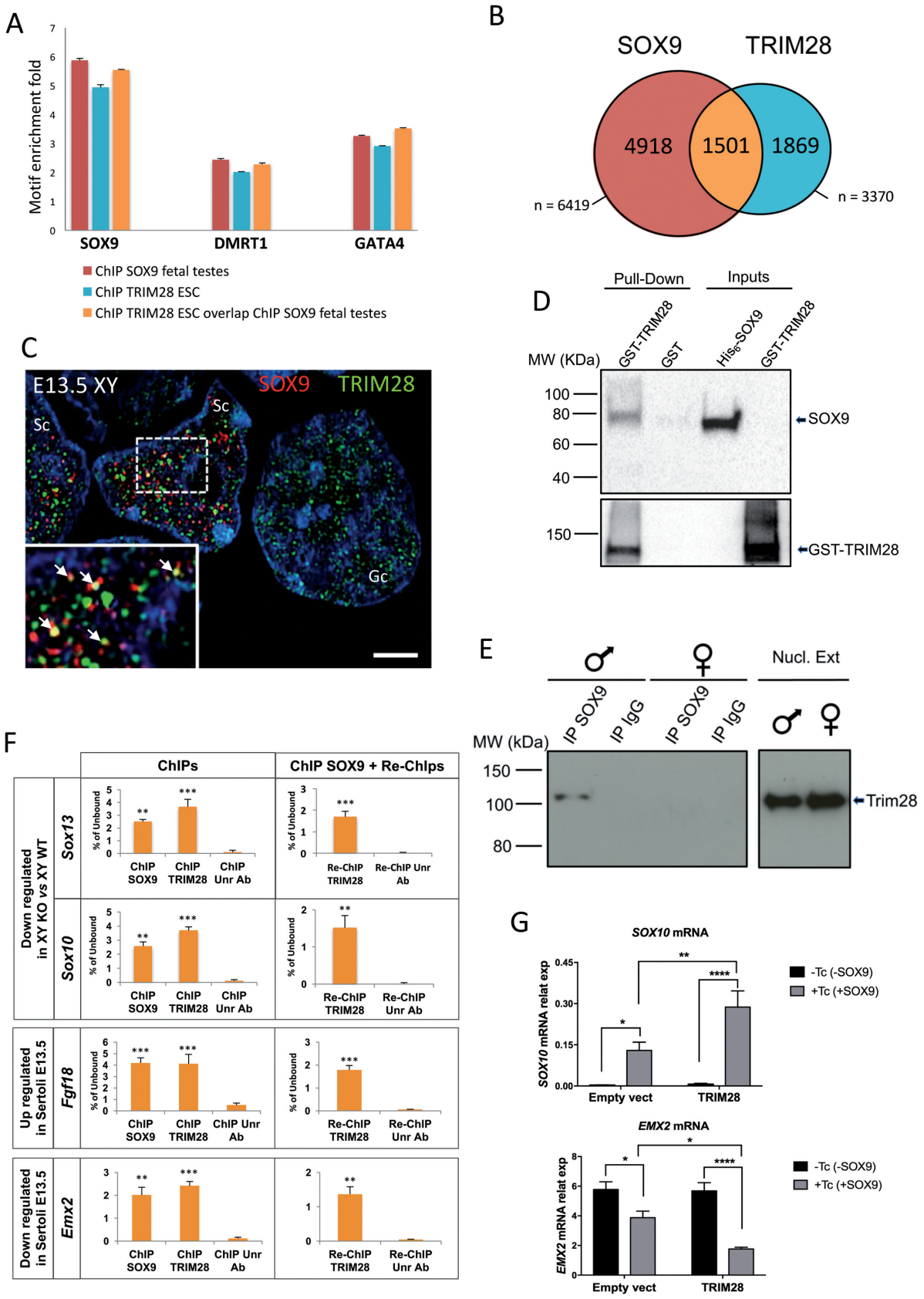


Figure 6. SCS screening allowed the identification of TRIM28 as a new SOX9 partner in foetal testis. (**A**) The three SOX9/DMRT1/GATA4 binding motifs are similarly enriched in the TRIM28 ChIP-seq dataset from mouse ES cells (blue bars) (Supplementary Table S7) as the SCSs observed in the

6G, lower panels). These results suggest that TRIM28 could contribute to SOX9-mediated activation or repression of its target genes.

These findings indicate that TRIM28 binds together with SOX9 to chromatin regions containing SCSs and may contribute to the transcriptional regulation SOX9 target genes in foetal Sertoli cells. This strongly supports a novel functional relationship between SOX9 and TRIM28 and suggests that TRIM28 might be a new factor of testis differentiation with a role that remains to be fully characterized.

DISCUSSION

In this study, we tried to elucidate how the transcription factor SOX9 controls Sertoli cell fate by regulating the expression/transcription of its target genes. First, we identified the set of genes bound by SOX9 in two distant mammals (mouse and cattle) at an equivalent stage of foetal testis development. The 4293 conserved SOX9-bound genes include many molecules with previously described roles in male and female somatic sex determination. Moreover, a considerable proportion of genes that are up- or downregulated specifically in Sertoli cells at E13.5 (64) are bound by SOX9 (38% and 44% respectively), suggesting that SOX9 could simultaneously activate the male and repress the female differentiation pathway. Supporting this, we observed that in both mouse and bovine foetal testes, SOX9 also binds to typical female genes, such as *Wnt4* (61), *Fst, Bmp2* (62) and *FoxL2* (63).

To investigate the functional role of SOX9 in Sertoli cells, we then knocked-out the *Sox9* gene after primary sex determination specifically in Sertoli cells. This allowed creating a hypomorphic mutant of the SOX E group. Moreover, in this genetic background, *Sox8* remains active and plays a partial compensatory role (67). Therefore, the effects on the E13.5 gonad transcriptome should result from transcriptional mis-regulation only in Sertoli cells.

In $Sox9^{\Delta/\Delta}$ XY gonads, 240 genes were down- and 107 genes up-regulated compared with wild type XY gonads. SOX9-bound genes were more frequently found among the downregulated than the upregulated genes (40% versus 22%), suggesting that SOX9 acts mainly as a transcriptional activator. Alternatively, *Sox8* could replace *Sox9* for this function because in the double *Sox9* and *Sox*8 knock-out, *FoxL2*, a female gene bound by SOX9 in the mouse and bovine ChIP-seq datasets, shows ectopic testicular expression (88). Among the 95 SOX9-activated genes (*i.e.,* downregulated in the absence of SOX9), 65 are also differentially expressed in male and female supporting cells and the expression of 37 of them in $Sox9^{\Delta/\Delta}$ XY gonads was intermediate between that observed in wild type XY and XX gonads. For example, *Sox10* displayed residual expression, probably the result of its activation by SOX8, as already reported (88). Also, our results identify *Sox10* as a direct target gene of SOX9 during sexual differentiation. The expression of other downregulated genes in $Sox9^{\Delta/\Delta}$ XY gonads was comparable or lower than in wild type XX gonads, showing that they are specifically regulated by SOX9 and not rescued by SOX8. Notably, the expression of 30 SOX9-bound genes that were downregulated in $Sox9^{\Delta/\Delta}$ XY gonads compared with wild type XY gonads was comparable in Sertoli cells of XY gonads and granulosa cells of XX gonads (64). This suggests that SOX9 contributes to the regulation of the expression of non-sex-specific genes involved in the fate of supporting cells. This role would depend on distinct transcription factors in female granulosa cells. Using an *in vitro* approach, we confirmed the activating role of SOX9 on genes that are downregulated in $Sox9^{\Delta/\Delta}$ XY gonads (such as *SOX10*). We also confirmed that genes (such as SOX6) upregulated in Sertoli cells and bound by SOX9 but not affected by in $Sox9^{\Delta/\Delta}$ XY gonads could be activated by SOX9. This observation is explained by SOX9 redundancy with SOX8 in mutant gonads.

We then took advantage of the RNA-seq technology to investigate whether SOX9 ablation could perturb RNA processing of its target genes. In $Sox9^{\Delta/\Delta}$ XY gonads, splicing of 154 transcripts (among which 70 of SOX9-bound genes) was significantly affected. In $Sox9^{\Delta/\Delta}$ XY gonads, splicing variations encompassed the five major splicing events. For example, incorporation of exon 2 in the *Atrx* transcript was more frequently observed in mutant testis and ovary than in wild type testis. *Fgfr2,* which is crucial for sex-determination (75), has two spliced isoforms that include either exon 8 (*Fgfr2b*) or exon 9 (*Fgfr2c* that encodes the high affinity FGF9 receptor) (89). Our results show that at E13.5, SOX9 inhibits the inclusion of exon 9 in the *Fgfr2* transcript and promotes the production of *Fgfr2b*, as already described in pancreas (90) and in agreement with our previous observation that the *Fgfr2c* isoform is more expressed in XY sex-reversed ovaries than in XY testes (91).

SOX9 ChIP-seq dataset from mouse foetal testes (red bars). Orange bars, TRIM28 ChIP-seq peaks that overlap with SOX9 ChIP-seq peaks as shown in the Venn diagram (**B**). (**C**) Super-resolution OMX fluorescence microscopy using 0.12 μm Z-sections showing the co-localization of SOX9 (red) and TRIM28 (green) in the nucleus of E13.5 Sertoli cells (orange/yellow spots indicated by white arrows in the insets). DNA was stained with DAPI (blue). Sc: Sertoli cells; Gc: germ cells. Bar: 2 μm. (**D**) In vitro pull-down experiments using purified mouse GST-TRIM28 or GST alone and 6xHis-tagged human recombinant SOX9. Proteins were analysed by western blotting using anti-SOX9 (upper panel) and anti-TRIM28 antibodies (lower panel). (**E**) Left panel: nuclear extracts from E13.5 male and female foetal gonads were immunoprecipitated using the anti-CT-SOX9 antibody or IgG (control) and analysed by western blotting with the anti-TRIM28 antibody. Right panel: detection of TRIM28 in nuclear extracts (Nucl. Ext) before immunoprecipitation (input). (**F**) ChIP-qPCR (left panels) and sequential ChIP-qPCR (right panels) of sheared chromatin from E13.5 foetal mouse testes with anti-SOX9 and –TRIM28 antibodies. The analysed genomic regions correspond to overlapping SOX9 and TRIM28 ChIP-seq peaks. *Sox13* and *Sox10* are genes upregulated in E13.5 Sertoli cells versus granulosa cells and downregulated in XY KO gonads. *Fgf18* and *Emx2* are genes up- and down-regulated, respectively, in E13.5 Sertoli cells versus granulosa cells. Data are the mean ± SD ($n = 3$); $^{***}P < 0.001$; $^{**}P < 0.01$ compared to control unrelated antibody (Unr Ab) (ANOVA with the Geisser–Greenhouse correction for multiple comparisons for ChIP data; Student's *t* test for sequential ChIP data). (**G**) RT-qPCR analysis showing transcriptional activation or repression of the *SOX10* and *EMX2* genes, respectively) in tet-SOX9-HEK293T cells incubated (+Tc) or not (–Tc) with tetracycline to induce SOX9 expression and transfected or not (empty vector) with a TRIM28 expression plasmid. *SOX10* and *EMX2* expression data are relative to the *IPO8* control. Data are the mean ± SD ($n = 3$); $^{****}P < 0.0001$; $^{**}P < 0.01$; $^{*}P < 0.1$ (ANOVA with the Geisser–Greenhouse correction for multiple comparisons).

More generally, our results suggest that SOX9 could influence the proteome of a specific cell type by controlling its splicing programme. Beside the production of specific isoforms of proteins involved in the epithelial cell fate (*Ubfd1* and *Gadd45g*), intron retention could be a powerful mechanism to finely regulate the stoichiometry of active proteins without modifying transcription.

Concerning the molecular mechanisms whereby SOX9 might control RNA processing, recent studies showed that epigenetic regulation plays a direct role in the control of alternative splicing (for review see:(92)). For instance, alternative splicing of *Fgfr2* transcripts is regulated at the chromatin level (93), suggesting a possible link between SOX9 binding to its target genes and the splicing machinery. The finding that the SOX9 ChIP-seq peaks of 28% of SOX9-bound genes with perturbed mRNA splicing in $Sox9^{\Delta/\Delta}$ XY gonads overlapped with the exons/introns involved in the splicing events (Figure 3B) is strongly in favour of this hypothesis. As transcription and splicing are coupled, we could hypothesize that DNA-bound SOX9 might regulate the cross-talk between chromatin and the splicing machinery during transcription of its target genes. Alternatively, and as previously suggested for SOX proteins (71,94), direct binding of SOX9 to pre-mRNA cannot be excluded. Sex-specific RNA splicing is a key mechanism in *Drosophila* (95), but was not known in mammals. Although we showed that *in vitro* SOX9 can regulate splicing, this mechanism requires additional studies at the molecular level.

We then investigated whether SOX9-bound genes in mammalian testis shared genomic determinants that characterize the Sertoli cell fate developmental program. For this, we took advantage of the ChIP-seq analysis in two distant mammals that highlighted an important conservation of the SOX9 peak distribution in the two species. Therefore, instead of comparing the nucleotide conservation of putative binding sites for transcription factors among orthologous regions, we chose to analyse the distribution of pre-defined binding sites for five Sertoli cell differentiation/reprogramming factors (SOX9, GATA4, DMRT1, SF1 and WT1) (18) in the whole ChIP-seq datasets from both mammals. This led to the definition of a Sertoli cell signature (SCS) in which SOX9, GATA4 and DMRT1 motif enrichment characterizes the genomic regions bound by SOX9 in Sertoli cells. This genomic signature is conserved, as indicated by its presence in orthologous regions from twelve other mammals. We also observed this signature in other vertebrate genomes, for instance, in the chicken genome (not shown) where DMRT1 plays a crucial role in sex determination (96). The fact that SOX, GATA and DMRT motifs co-segregate in a set of genes involved in male gonad differentiation could explain why either DMRT domains (birds, some fish species) or SOX domains (mammals) are involved in testis determination in vertebrates.

We found that SCSs are not a common feature of regions bound by other transcription factors or other SOX factors. Moreover, SCSs are not present in SOX9 target regions in the other tissue/cell types we analyzed (80–83), but for mouse basal cell carcinoma and human colorectal cancer cells. This suggests that in these pathological conditions SOX9 genomic features show strong similarity with those observed in foetal Sertoli cells. As SOX9 overexpression has an oncogenic role in basal cell carcinoma (83) and colorectal cancer (97), the presence of this genomic signature in other cancer types should now be investigated.

Analysis of SCS organization revealed a strong motif clustering, where motifs are separated by fewer than 50 nucleotides. Moreover, in peaks close to the gene body, motifs clustered at a precise distance upstream or downstream the TSS. This suggests that this genomic feature has functional significance in transcriptional regulation. Moreover, SOX9 and GATA4 binding motifs were preferentially associated, while the position of DMRT1 motifs was less constrained. Importantly using sequential ChIP-qPCR we showed that DMRT1 and GATA4 co-localize with SOX9 on chromatins of its target genes. Additional analyses are required to determine how these factors collaborate in gene regulation and how they bind to SCSs.

As DMRT1 is expressed in Sertoli and germ cells at E13.5, we searched for potential SCSs in DMRT1 ChIP-seq datasets (25). We found that peaks bound by both SOX9 and DMRT1 had SCSs, while those only bound by DMRT1 presented a different motif enrichment pattern. This suggests that a genomic region bound by DMRT1 will display a SCS only if it is also bound by SOX9.

As a proof of concept, we used our global motif-enrichment scanning method to identify new factors with SCSs in ChIP-seq datasets that could have functional relationship with SOX9. We found that in ES cells, TRIM28 interacts with many regions that are also bound by SOX9 in foetal testis and showed that the two proteins form a complex within the foetal testis. Moreover, as TRIM28 co-localizes on SOX9 target genes *in vivo* and cooperate with SOX9 *in vitro* to regulate gene expression, our results highlight an important role for TRIM28 in SOX9-dependent transcriptional regulation. Interestingly, our *in vitro* results suggest a bi-potential action for SOX9-TRIM28 complexes, either activating, either repressing the target genes. This observation fits well with the ChIP-seq results where SOX9 binds to male and also female genes.

Concerning the potential role of SCSs, they might facilitate the attraction of SOX9 or of other factors involved in male sexual differentiation, such as DMRT1 or GATA4, towards genes involved in the Sertoli cell programme, enhancing the probability of building multi-protein complexes on chromatin. This would make of SCSs important components of Sertoli cell fate. Additional studies with gain- or loss-of-function mouse models are required to validate this hypothesis. Moreover, it is not yet known whether in Sertoli cells SOX9 binds to all regions harbouring SCSs.

The motif-enrichment scanning method we developed for SCS identification uses matrices to characterize binding-motifs of a set of transcription factors defined by a reprogramming experiment, here the transcription factors required for the direct reprogramming of embryonic fibroblasts into Sertoli cells (18). This approach could be used also in other biological systems because many transcription factor combinations can directly reprogram the fate of different cell types (98).

In conclusion, our work shows that in mouse and bovine foetal testes, SOX9 binds to a core group of ∼4000 genes that includes all those with an already known role in somatic sex determination. Moreover, despite redundancy

with SOX8, SOX9 regulates specifically the transcription or splicing of a subset of target genes. However, we cannot rule out that SOX9 activity on splicing might be indirect. Finally, the presence of SOX9 on chromatin in foetal testes is correlated with a genomic signature (SCS) comprised of binding motifs for transcription factors involved in Sertoli cell programming, suggesting that this signature represents a Sertoli cell regulatory code. This genomic feature allowed us to discover TRIM28, a new SOX9 protein partner.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the staff of the Montpellier Imaging Facility (MRI) for their help with microscopy experiments. We thank Morgane Thomas-Chollier and Jacques van Helden for helpful discussion on motif scanning. We thank Sergei Tevossian for the mouse GATA4 expression plasmid. We are grateful to Dominique Giorgi for critical reading of the manuscript. We thank Elisabetta Andermarcher for manuscript editing.

FUNDING

Agence Nationale pour la Recherche (ANR blanc Testis-Dev to B.B. and F.P.); National Health and Medical Research Council Program [1074258 to V.R.H.]; National Health and Medical Research Council Fellowship [1020034 to V.R.H.]; Victorian Government's Operational Infrastructure Support Program (to V.R.H.); Australian Postgraduate Award (to A.S.). Funding for open access charge: Agence Nationale pour la Recherche.
Conflict of interest statement. None declared.

REFERENCES

1 Svingen,T. and Koopman,P. (2013) Building the mammalian testis: origins, differentiation, and assembly of the component cell populations. *Genes Dev.*, **27**, 2409–2426.
2 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.
3 Qin,Y. and Bishop,C.E. (2005) Sox9 is sufficient for functional testis development producing fertile male mice in the absence of Sry. *Hum. Mol. Genet.*, **14**, 1221–1229.
4 Wagner,T., Wirth,J., Meyer,J., Zabel,B., Held,M., Zimmer,J., Pasantes,J., Bricarelli,F.D., Keutel,J., Hustert,E. *et al.* (1994) Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell*, **79**, 1111–1120.
5 Chaboissier,M.C., Kobayashi,A., Vidal,V.I., Lutzkendorf,S., van de Kant,H.J., Wegner,M., de Rooij,D.G., Behringer,R.R. and Schedl,A. (2004) Functional analysis of Sox8 and Sox9 during sex determination in the mouse. *Development*, **131**, 1891–1901.
6 Barrionuevo,F., Bagheri-Fam,S., Klattig,J., Kist,R., Taketo,M.M., Englert,C. and Scherer,G. (2006) Homozygous inactivation of Sox9 causes complete XY sex reversal in mice. *Biol. Reprod.*, **74**, 195–201.
7 Huang,B., Wang,S., Ning,Y., Lamb,A.N. and Bartley,J. (1999) Autosomal XX sex reversal caused by duplication of SOX9. *Am. J. Med. Genet.*, **87**, 349–353.
8 Bishop,C.E., Whitworth,D.J., Qin,Y., Agoulnik,A.I., Agoulnik,I.U., Harrison,W.R., Behringer,R.R. and Overbeek,P.A. (2000) A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. *Nat. Genet.*, **26**, 490–494.
9 Vidal,V.P., Chaboissier,M.C., de Rooij,D.G. and Schedl,A. (2001) Sox9 induces testis development in XX transgenic mice. *Nat. Genet.*, **28**, 216–217.
10 Polanco,J.C., Wilhelm,D., Davidson,T.L., Knight,D. and 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.
11 Lavery,R., Chassot,A.A., Pauper,E., Gregoire,E.P., Klopfenstein,M., de Rooij,D.G., Mark,M., Schedl,A., Ghyselinck,N.B. and Chaboissier,M.C. (2012) Testicular differentiation occurs in absence of R-spondin1 and Sox9 in mouse sex reversals. *PLoS Genet.*, **8**, e1003170.
12 Murphy,M.W., Lee,J.K., Rojo,S., Gearhart,M.D., Kurahashi,K., Banerjee,S., Loeuille,G.A., Bashamboo,A., McElreavey,K., Zarkower,D. *et al.* (2015) An ancient protein-DNA interaction underlying metazoan sex determination. *Nat. Struct. Mol. Biol.*, **22**, 442–451.
13 Ono,M. and Harley,V.R. (2013) Disorders of sex development: new genes, new concepts. *Nat. Rev. Endocrinol.*, **9**, 79–91.
14 Raymond,C.S., Murphy,M.W., O'Sullivan,M.G., Bardwell,V.J. and Zarkower,D. (2000) Dmrt1, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. *Genes Dev.*, **14**, 2587–2595.
15 Padua,M.B., Jiang,T., Morse,D.A., Fox,S.C., Hatch,H.M. and Tevosian,S.G. (2015) Combined loss of the GATA4 and GATA6 transcription factors in male mice disrupts testicular development and confers adrenal-like function in the testes. *Endocrinology*, **156**, 1873–1886.
16 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.
17 Luo,X., Ikeda,Y. and Parker,K.L. (1995) The cell-specific nuclear receptor steroidogenic factor 1 plays multiple roles in reproductive function. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, **350**, 279–283.
18 Buganim,Y., Itskovich,E., Hu,Y.C., Cheng,A.W., Ganz,K., Sarkar,S., Fu,D., Welstead,G.G., Page,D.C. and Jaenisch,R. (2012) Direct reprogramming of fibroblasts into embryonic Sertoli-like cells by defined factors. *Cell Stem Cell*, **11**, 373–386.
19 Zhao,L., Svingen,T., Ng,E.T. and Koopman,P. (2015) Female-to-male sex reversal in mice caused by transgenic overexpression of Dmrt1. *Development*, **142**, 1083–1088.
20 Lindeman,R.E., Gearhart,M.D., Minkina,A., Krentz,A.D., Bardwell,V.J. and Zarkower,D. (2015) Sexual cell-fate reprogramming in the ovary by DMRT1. *Curr. Biol.*, **25**, 764–771.
21 Nowak,D.E., Tian,B. and Brasier,A.R. (2005) Two-step cross-linking method for identification of NF-kappaB gene network by chromatin immunoprecipitation. *Biotechniques*, **39**, 715–725.
22 Gasca,S., Canizares,J., De Santa Barbara,P., Mejean,C., Poulat,F., Berta,P. and Boizet-Bonhoure,B. (2002) A nuclear export signal within the high mobility group domain regulates the nucleocytoplasmic translocation of SOX9 during sexual determination. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 11199–11204.
23 Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
24 Bardet,A.F., He,Q., Zeitlinger,J. and Stark,A. (2012) A computational pipeline for comparative ChIP-seq analyses. *Nat. Protoc.*, **7**, 45–61.
25 Krentz,A.D., Murphy,M.W., Zhang,T., Sarver,A.L., Jain,S., Griswold,M.D., Bardwell,V.J. and Zarkower,D. (2013) Interaction between DMRT1 function and genetic background modulates signaling and pluripotency to control tumor susceptibility in the fetal germ line. *Dev. Biol.*, **377**, 67–78.
26 Barrionuevo,F., Georg,I., Scherthan,H., Lecureuil,C., Guillou,F., Wegner,M. and 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**, 301–312.
27 McFarlane,L., Truong,V., Palmer,J.S. and Wilhelm,D. (2013) Novel PCR assay for determining the genetic sex of mice. *Sex Dev.*, **7**, 207–211.
28 Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

29 Kim,D., Pertea,G., Trapnell,C., Pimentel,H., Kelley,R. and Salzberg,S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.
30 Anders,S., Pyl,P.T. and Huber,W. (2015) HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.
31 Law,C.W., Chen,Y., Shi,W. and Smyth,G.K. (2014) voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.*, **15**, R29.
32 Ritchie,M.E., Phipson,B., Wu,D., Hu,Y., Law,C.W., Shi,W. and Smyth,G.K. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*, **43**, e47.
33 Robinson,M.D. and Oshlack,A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.*, **11**, R25.
34 Liu,R., Holik,A.Z., Su,S., Jansz,N., Chen,K., Leong,H.S., Blewitt,M.E., Asselin-Labat,M.L., Smyth,G.K. and Ritchie,M.E. (2015) Why weight? Modelling sample and observational level variability improves power in RNA-seq analyses. *Nucleic Acids Res.*, **43**, e97.
35 Ritchie,M.E., Diyagama,D., Neilson,J., van Laar,R., Dobrovic,A., Holloway,A. and Smyth,G.K. (2006) Empirical array quality weights in the analysis of microarray data. *BMC Bioinformatics*, **7**, 261.
36 Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
37 Nnamani,M.C., Plaza,S., Romero,R. and Wagner,G.P. (2013) Evidence for independent evolution of functional progesterone withdrawal in primates and guinea pigs. *Evol. Med.Public Health*, **2013**, 273–288.
38 Katz,Y., Wang,E.T., Airoldi,E.M. and Burge,C.B. (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat. Methods*, **7**, 1009–1015.
39 Thevenet,L., Mejean,C., Moniot,B., Bonneaud,N., Galeotti,N., Aldrian-Herrada,G., Poulat,F., Berta,P., Benkirane,M. and Boizet-Bonhoure,B. (2004) Regulation of human SRY subcellular distribution by its acetylation/deacetylation. *EMBO J.*, **23**, 3336–3345.
40 Thevenet,L., Albrecht,K.H., Malki,S., Berta,P., Boizet-Bonhoure,B. and Poulat,F. (2005) NHERF2/SIP-1 interacts with mouse SRY via a different mechanism than human SRY. *J. Biol. Chem.*, **280**, 38625–38630.
41 Murphy,M.W., Zarkower,D. and Bardwell,V.J. (2007) Vertebrate DM domain proteins bind similar DNA sequences and can heterodimerize on DNA. *BMC Mol. Biol.*, **8**, 58.
42 Rojas,A., Schachterle,W., Xu,S.M. and Black,B.L. (2009) An endoderm-specific transcriptional enhancer from the mouse Gata4 gene requires GATA and homeodomain protein-binding sites for function in vivo. *Dev. Dyn.*, **238**, 2588–2598.
43 Mertin,S., McDowall,S.G. and Harley,V.R. (1999) The DNA-binding specificity of SOX9 and other SOX proteins. *Nucleic Acids Res.*, **27**, 1359–1364.
44 Notarnicola,C., Malki,S., Berta,P., Poulat,F. and Boizet-Bonhoure,B. (2006) Transient expression of SOX9 protein during follicular development in the adult mouse ovary. *Gene Expr. Patterns*, **6**, 695–702.
45 Nielsen,A.L., Ortiz,J.A., You,J., Oulad-Abdelghani,M., Khechumian,R., Gansmuller,A., Chambon,P. and Losson,R. (1999) Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *EMBO J.*, **18**, 6385–6395.
46 Cammas,F., Oulad-Abdelghani,M., Vonesch,J.L., Huss-Garcia,Y., Chambon,P. and Losson,R. (2002) Cell differentiation induces TIF1beta association with centromeric heterochromatin via an HP1 interaction. *J. Cell Sci.*, **115**, 3439–3448.
47 D'Astolfo,D.S., Pagliero,R.J., Pras,A., Karthaus,W.R., Clevers,H., Prasad,V., Lebbink,R.J., Rehmann,H. and Geijsen,N. (2015) Efficient intracellular delivery of native proteins. *Cell*, **161**, 674–690.
48 Munger,S.C. and Capel,B. (2012) Sex and the circuitry: progress toward a systems-level understanding of vertebrate sex determination. *Wiley Interdiscipl. Rev. Syst. Biol. Med.*, **4**, 401–412.
49 Colvin,J.S., Green,R.P., Schmahl,J., Capel,B. and Ornitz,D.M. (2001) Male-to-female sex reversal in mice lacking fibroblast growth factor 9. *Cell*, **104**, 875–889.
50 Munger,S.C., Natarajan,A., Looger,L.L., Ohler,U. and Capel,B. (2013) Fine time course expression analysis identifies cascades of activation and repression and maps a putative regulator of Mammalian sex determination. *PLoS Genet.*, **9**, e1003630.
51 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.
52 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) Gadd45gamma and Map3k4 interactions regulate mouse testis determination via p38 MAPK-mediated control of Sry expression. *Dev. Cell*, **23**, 1020–1031.
53 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.
54 Katoh-Fukui,Y., Miyabayashi,K., Komatsu,T., Owaki,A., Baba,T., Shima,Y., Kidokoro,T., Kanai,Y., Schedl,A., Wilhelm,D. *et al.* (2011) Cbx2, a polycomb group gene, is required for Sry gene expression in mice. *Endocrinology*, **153**, 913–924.
55 Cui,S., Ross,A., Stallings,N., Parker,K.L., Capel,B. and Quaggin,S.E. (2004) Disrupted gonadogenesis and male-to-female sex reversal in Pod1 knockout mice. *Development*, **131**, 4095–4105.
56 Kim,Y., Kobayashi,A., Sekido,R., DiNapoli,L., Brennan,J., Chaboissier,M.C., Poulat,F., Behringer,R.R., Lovell-Badge,R. and Capel,B. (2006) Fgf9 and Wnt4 act as antagonistic signals to regulate mammalian sex determination. *PLoS Biol.*, **4**, e187.
57 Yao,H.H. and Capel,B. (2002) Disruption of testis cords by cyclopamine or forskolin reveals independent cellular pathways in testis organogenesis. *Dev. Biol.*, **246**, 356–365.
58 Callier,P., Calvel,P., Matevossian,A., Makrythanasis,P., Bernard,P., Kurosaka,H., Vannier,A., Thauvin-Robinet,C., Borel,C., Mazaud-Guittot,S. *et al.* (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.
59 Morais da Silva,S., Hacker,H., Harley,V., Goodfellow,P.N., Swain,A. and Lovell-Badge,R. (1996) Sox9 expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. *Nat. Genet.*, **13**, 62–68.
60 Schepers,G.E., Bullejos,M., Hosking,B.M. and Koopman,P. (2000) Cloning and characterisation of the Sry-related transcription factor gene Sox8. *Nucleic Acids Res.*, **28**, 1473–1480.
61 Vainio,S., Heikkila,M., Kispert,A., Chin,N. and McMahon,A.P. (1999) Female development in mammals is regulated by Wnt-4 signalling. *Nature*, **397**, 405–409.
62 Yao,H.H., Matzuk,M.M., Jorgez,C.J., Menke,D.B., Page,D.C., Swain,A. and Capel,B. (2004) Follistatin operates downstream of Wnt4 in mammalian ovary organogenesis. *Dev. Dyn.*, **230**, 210–215.
63 Boulanger,L., Pannetier,M., Gall,L., Allais-Bonnet,A., Elzaiat,M., Le Bourhis,D., Daniel,N., Richard,C., Cotinot,C., Ghyselinck,N.B. *et al.* (2014) FOXL2 is a female sex-determining gene in the goat. *Curr. Biol.*, **24**, 404–408.
64 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 biased lineage priming of sexual fate in the fetal mouse gonad. *PLoS Genet.*, **8**, e1002575.
65 Kist,R., Schrewe,H., Balling,R. and Scherer,G. (2002) Conditional inactivation of Sox9: a mouse model for campomelic dysplasia. *Genesis*, **32**, 121–123.
66 Lecureuil,C., Fontaine,I., Crepieux,P. and Guillou,F. (2002) Sertoli and granulosa cell-specific Cre recombinase activity in transgenic mice. *Genesis*, **33**, 114–118.
67 Barrionuevo,F., Georg,I., Scherthan,H., Lecureuil,C., Guillou,F., Wegner,M. and Scherer,G. (2008) 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.

68 Behringer,R.R., Finegold,M.J. and Cate,R.L. (1994) Mullerian-inhibiting substance function during mammalian sexual development. *Cell*, **79**, 415–425.
69 Bitgood,M.J., Shen,L. and McMahon,A.P. (1996) Sertoli cell signaling by Desert hedgehog regulates the male germline. *Curr. Biol.*, **6**, 298–304.
70 Knower,K.C., Sim,H., McClive,P.J., Bowles,J., Koopman,P., Sinclair,A.H. and Harley,V.R. (2007) Characterisation of urogenital ridge gene expression in the human embryonal carcinoma cell line NT2/D1. *Sex Dev.*, **1**, 114–126.
71 Ohe,K., Lalli,E. and Sassone-Corsi,P. (2002) A direct role of SRY and SOX proteins in pre-mRNA splicing. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 1146–1151.
72 Hata,K., Nishimura,R., Muramatsu,S., Matsuda,A., Matsubara,T., Amano,K., Ikeda,F., Harley,V.R. and Yoneda,T. (2008) Paraspeckle protein p54nrb links Sox9-mediated transcription with RNA processing during chondrogenesis in mice. *J. Clin. Invest.*, **118**, 3098–3108.
73 Li,Q., Lee,J.A. and Black,D.L. (2007) Neuronal regulation of alternative pre-mRNA splicing. *Nat. Rev. Neurosci.*, **8**, 819–831.
74 Bagheri-Fam,S., Argentaro,A., Svingen,T., Combes,A.N., Sinclair,A.H., Koopman,P. and Harley,V.R. (2011) Defective survival of proliferating Sertoli cells and androgen receptor function in a mouse model of the ATR-X syndrome. *Hum. Mol. Genet.*, **20**, 2213–2224.
75 Kim,Y., Bingham,N., Sekido,R., Parker,K.L., Lovell-Badge,R. and Capel,B. (2007) Fibroblast growth factor receptor 2 regulates proliferation and Sertoli differentiation during male sex determination. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 16558–16563.
76 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.
77 Yamagata,T., Nishida,J., Sakai,R., Tanaka,T., Honda,H., Hirano,N., Mano,H., Yazaki,Y. and Hirai,H. (1995) Of the GATA-binding proteins, only GATA-4 selectively regulates the human interleukin-5 gene promoter in interleukin-5-producing cells which express multiple GATA-binding proteins. *Mol. Cell. Biol.*, **15**, 3830–3839.
78 Hartwig,S., Ho,J., Pandey,P., Macisaac,K., Taglienti,M., Xiang,M., Alterovitz,G., Ramoni,M., Fraenkel,E. and Kreidberg,J.A. (2010) Genomic characterization of Wilms' tumor suppressor 1 targets in nephron progenitor cells during kidney development. *Development*, **137**, 1189–1203.
79 Ferraz-de-Souza,B., Lin,L., Shah,S., Jina,N., Hubank,M., Dattani,M.T. and Achermann,J.C. (2011) ChIP-on-chip analysis reveals angiopoietin 2 (Ang2, ANGPT2) as a novel target of steroidogenic factor-1 (SF-1, NR5A1) in the human adrenal gland. *FASEB J.*, **25**, 1166–1175.
80 Ohba,S., He,X., Hojo,H. and McMahon,A.P. (2015) Distinct transcriptional programs underlie Sox9 regulation of the mammalian chondrocyte. *Cell Rep.*, **12**, 229–243.
81 Garside,V.C., Cullum,R., Alder,O., Lu,D.Y., Vander Werff,R., Zhao,Y., Jones,S.J., Marra,M.A., Underhill,T.M. *et al.* (2015) SOX9 modulates the expression of key transcription factors required for heart valve development. *Development*, **142**, 4340–4350.
82 Kadaja,M., Keyes,B.E., Lin,M., Pasolli,H.A., Genander,M., Polak,L., Stokes,N., Zheng,D. and Fuchs,E. (2014) SOX9: a stem cell transcriptional regulator of secreted niche signaling factors. *Genes Dev.*, **28**, 328–341.
83 Larsimont,J.C., Youssef,K.K., Sanchez-Danes,A., Sukumaran,V., Defrance,M., Delatte,B., Liagre,M., Baatsen,P., Marine,J.C., Lippens,S. *et al.* (2015) Sox9 controls self-renewal of oncogene targeted cells and links tumor initiation and invasion. *Cell Stem Cell*, **17**, 60–73.
84 Shi,Z., Chiang,C.I., Labhart,P., Zhao,Y., Yang,J., Mistretta,T.A., Henning,S.J., Maity,S.N. and Mori-Akiyama,Y. (2015) Context-specific role of SOX9 in NF-Y mediated gene regulation in colorectal cancer cells. *Nucleic Acids Res.*, **43**, 6257–6269.
85 Raymond,C.S., Kettlewell,J.R., Hirsch,B., Bardwell,V.J. and Zarkower,D. (1999) Expression of Dmrt1 in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. *Dev. Biol.*, **215**, 208–220.
86 Hu,G., Kim,J., Xu,Q., Leng,Y., Orkin,S.H. and Elledge,S.J. (2009) A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev.*, **23**, 837–848.
87 Gustafsson,M.G., Shao,L., Carlton,P.M., Wang,C.J., Golubovskaya,I.N., Cande,W.Z., Agard,D.A. and Sedat,J.W. (2008) Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophys. J.*, **94**, 4957–4970.
88 Georg,I., Barrionuevo,F., Wiech,T. and Scherer,G. (2012) Sox9 and Sox8 are required for basal lamina integrity of testis cords and for suppression of FOXL2 during embryonic testis development in mice. *Biol. Reprod.*, **87**, 1–11.
89 Santos-Ocampo,S., Colvin,J.S., Chellaiah,A. and Ornitz,D.M. (1996) Expression and biological activity of mouse fibroblast growth factor-9. *J. Biol. Chem.*, **271**, 1726–1731.
90 Seymour,P.A., Shih,H.P., Patel,N.A., Freude,K.K., Xie,R., Lim,C.J. and Sander,M. (2012) A Sox9/Fgf feed-forward loop maintains pancreatic organ identity. *Development*, **139**, 3363–3372.
91 Bagheri-Fam,S., Ono,M., Li,L., Zhao,L., Ryan,J., Lai,R., Katsura,Y., Rossello,F.J., Koopman,P., Scherer,G. *et al.* (2015) FGFR2 mutation in 46,XY sex reversal with craniosynostosis. *Hum. Mol. Genet.*, **24**, 6699–6710.
92 Luco,R.F., Allo,M., Schor,I.E., Kornblihtt,A.R. and Misteli,T. (2011) Epigenetics in alternative pre-mRNA splicing. *Cell*, **144**, 16–26.
93 Luco,R.F., Pan,Q., Tominaga,K., Blencowe,B.J., Pereira-Smith,O.M. and Misteli,T. (2010) Regulation of alternative splicing by histone modifications. *Science*, **327**, 996–1000.
94 Ohe,K., Tamai,K.T., Parvinen,M. and Sassone-Corsi,P. (2009) DAX-1 and SOX6 molecular interplay results in an antagonistic effect in pre-mRNA splicing. *Dev. Dyn.*, **238**, 1595–1604.
95 Salz,H.K. (2011) Sex determination in insects: a binary decision based on alternative splicing. *Curr. Opin. Genet. Dev.*, **21**, 395–400.
96 Smith,C.A., Roeszler,K.N., Ohnesorg,T., Cummins,D.M., Farlie,P.G., Doran,T.J. and Sinclair,A.H. (2009) The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature*, **461**, 267–271.
97 Matheu,A., Collado,M., Wise,C., Manterola,L., Cekaite,L., Tye,A.J., Canamero,M., Bujanda,L., Schedl,A., Cheah,K.S. *et al.* (2012) Oncogenicity of the developmental transcription factor Sox9. *Cancer Res.*, **72**, 1301–1315.
98 Xu,J., Du,Y. and Deng,H. (2015) Direct lineage reprogramming: strategies, mechanisms, and applications. *Cell Stem Cell*, **16**, 119–134.