Human sex reversal is caused by duplication or deletion of core enhancers upstream of SOX9

Disorders of sex development (DSDs) are conditions affecting development of the gonads or genitalia. Variants in two key genes, SRY and its target SOX9, are an established cause of 46,XY DSD, but the genetic basis of many DSDs remains unknown. SRY-mediated SOX9 upregulation in the early gonad is crucial for testis development, yet the regulatory elements underlying this have not been identified in humans. Here, we identified four DSD patients with overlapping duplications or deletions upstream of SOX9. Bioinformatic analysis identified three putative enhancers for SOX9 that responded to different combinations of testis-specific regulators. All three enhancers showed synergistic activity and together drive SOX9 in the testis. This is the first study to identify SOX9 enhancers that, when duplicated or deleted, result in 46,XX or 46,XY sex reversal, respectively. These enhancers provide a hitherto missing link by which SRY activates SOX9 in humans, and establish SOX9 enhancer mutations as a significant cause of DSD.
gonadal sex differentiation begins in the embryo with the development of the bipotential gonads into either testes or ovaries. In most mammals including humans, testis determination is initiated by the Y-linked gene SRY. Sry is expressed transiently from 10.5 days post coitum (dpc) in mouse embryos while in humans expression begins at 6 weeks post ovulation and is maintained throughout gestation. SRY acts by upregulating the gene encoding the transcription factor SOX9, which is both necessary and sufficient for testis development and hence male differentiation in mice and humans. How SRY activates SOX9 transcription is a major unresolved issue. SOX9 tissue specific expression is influenced by long-range regulatory elements (including enhancers, which are orientation-independent controllers of expression) within the 2 Mb upstream of the transcriptional start site. In mice, it is known that SRY co-operates with NR5A1 (SF1) to bind to and activate a Sox9 enhancer known as Testis Specific Enhancer of Sox9 (TES) and its core sequence known as TESCO. Once Sox9 expression begins, Sox9 auto-regulates its own transcription via TESCO, creating a positive feedback loop. However, loss of function of TESCO in mice results in reduced Sox9 expression but this is insufficient to cause sex reversal. This indicates that TESCO is not the only enhancer required for Sox9 expression. In light of this, a recent study has identified a new mouse long-range enhancer for Sox9 gonadal expression (denoted Enh13). When this enhancer is deleted, Sox9 expression is markedly reduced resulting in complete XY sex reversal.

In recent years, several isolated 46,XX and 46,XY DSD patients with copy number variants (CNVs) within the 2 Mb putative upstream regulatory region have been identified, denoted XYSR and RevSex (reviewed by Croft et al.). In addition, despite extensive research, no DSD patients with pathogenic variants in TESCO have been identified. Thus, while studies have defined two enhancers necessary for Sox9 regulation in the mouse, to date, the testis-specific enhancers have yet to be identified in humans.

In this study we redefine the upstream regulatory landscape of human SOX9. Using new patient data, we refined the 32.5 kb XYSR and 24 kb RevSex intervals and analysed the genomic regions using bioinformatic and luciferase tiling approaches, to identify three putative enhancers 5’ of Sox9. In cell-based reporter assays these enhancers responded to different combinations of testis-specific regulators including SRY, SF1 and SOX9 itself. When combined, all three enhancers show synergistic activity, significantly increasing their individual enhancer activity. In vivo, deletion of these three enhancers in mice resulted in different outcomes ranging from: no apparent effect to reduced Sox9 transcription and complete sex reversal. Given that duplication or deletion of these sequences results in sex reversal in SRY-negative 46,XX males and 46,XY females, respectively, our results suggest a mechanism by which these enhancers have crucial roles in human sex development and DSD.

Results

A human SOX9 enhancer associated with 46,XX and 46XY DSD. We focused on genetic intervals in the 2 Mb upstream regulatory region that were previously associated with 46,XY and 46,XX DSD. First, we analysed a 32.5 kb region known as 46XY-SRY and 24 kb RevSex intervals and analysed genomic regions using bioinformatic and luciferase tiling approaches, to identify three putative enhancers 5’ of Sox9. In cell-based reporter assays these enhancers responded to different combinations of testis-specific regulators including SRY, SF1 and SOX9 itself. When combined, all three enhancers show synergistic activity, significantly increasing their individual enhancer activity. In vivo, deletion of these three enhancers in mice resulted in different outcomes ranging from: no apparent effect to reduced Sox9 transcription and complete sex reversal. Given that duplication or deletion of these sequences results in sex reversal in SRY-negative 46,XX males and 46,XY females, respectively, our results suggest a mechanism by which these enhancers have crucial roles in human sex development and DSD.

A human SOX9 enhancer contributes to SOX9 autoregulation. Next, we focused on a previously identified upstream regulatory region of SOX9, RevSex. When the RevSex region was first described its minimal length was 178 kb. The analysis of subsequent DSD patients carrying informative CNVs has allowed the minimal critical region to be narrowed to 41 kb. This region contains a strong enhancer signature previously confirmed by chromatin immunoprecipitation followed by sequencing (ChIP-seq). Recently, we further refined the minimal RevSex region to 24 kb by analysing a duplication in a 46,XX ovotesticular DSD patient (patient 3) overlapping those previously described. To locate the enhancer element within RevSex we employed an unbiased screen, generating 16 overlapping ~2 kb fragments covering this 24 kb region and testing each for enhancer activity using luciferase assays as described above. Two fragments (b8 and b9) showed a significant increase in enhancer activity in response to SF1 and SOX9. These two fragments have a 416 bp overlap that contains a highly conserved SOX9 binding motif. ENCODE data show a binding site for p300 (a general transcription factor often associated with enhancers)
Fig. 1 Duplication or deletion of the human SOX9 testicular enhancer eSR-A is associated with DSD. 

- **a**: The 600 kb genomic region upstream of human SOX9 showing the XYSR, RevSex and TESCO candidate regulatory regions. 
- **b**: The XYSR region was defined by deletions in two previously published 46,XY DSD patients (blue)⁷. Two novel duplications in 46,XX DSD patients (grey), allowed us to redefine the minimal overlap to 5.2 kb (green). 
- **c**: Sub-cloning strategy, dark blue lines indicate sub-clones analysed for enhancer activity in the pGL4.10 Beta-globin (βg) plasmid using luciferase assays. Predicted transcription factor binding motifs for SF1 (SF1-a and SF1-b) and SOX9 are shown with vertical light blue and magenta lines. Bioinformatic tracks from the UCSC genome browser are shown including the ENCODE track of enhancers present in human mammary epithelial cells (HMEC) (yellow denotes a weak enhancer)²⁸ and DNaseI hypersensitivity data from human foetal testis and ovary (ROADMap)²³. This shows a testis-specific peak over the a4 fragment. The 100-vertebrate conservation track shows a spike of conservation beneath the a4 fragment DNase peak. 
- **d**: Enhancer activities of sub-cloned fragments a1-a5 as measured by luciferase assays transfected with SF1 and SOX9 (n = 4). 
- **e**: Mutation of the SF1-a (eSR-A ΔSF1-a) and SOX9 (eSR-A ΔSOX9) binding motifs either separately or together results in a loss of enhancer activity compared to non-mutated eSR-A as assessed by luciferase assays with co-transfection of SF1 and SOX9 (n = 4). 
- **f**: Co-transfection of FOXL2 with SF1 + SOX9 shows a repression of eSR-A activity compared to SF1 + SOX9 (n = 3). All luciferase assays carried out in COS7 cells. Error bars are s.e.m. P-values (two-tailed t tests): *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. Source data are provided as a Source Data file.
and a DNaseI hypersensitive sites highlighting the regulatory potential of this genomic region\textsuperscript{27,28} (Fig. 2c). The 416 bp fragment also showed significant reporter activity when transfected constructs expressing SOX9 alone (Fig. 2e), and mutation of a SOX9 binding motif led to a 60% reduction in reporter expression levels ($P = 0.0053$, t-test) (Fig. 2f). We designated this fragment Sex Reversal Enhancer-B (eSR-B).

In addition, FOXL2 almost abolished eSR-B enhancer activity despite the stimulatory effect of SOX9 (Fig. 2g). Human eSR-B reporter transgenic mice showed equivalent expression in both testis and ovaries at E12.5 however, by E13.5 and E14.5 human eSR-B levels decreased in male gonads (Supplementary Figure 4a–c). Immunofluorescence on sections of E12.5 XY gonads showed that transgene expression was restricted to the
All three enhancers synergise to upregulate SOX9 expression. Previous reports of chondrocyte-specific SOX9 enhancers have shown that interaction of multiple enhancers can lead to synergistic gene activation\(^29\). Therefore, we assessed whether this was also the case with our three gonadal enhancers. We cloned different combinations of enhancers (eSR-A, eSR-B and eALDI) in tandem and tested these in our luciferase reporter assay. SRY is the initiator of SOX9 expression in the early gonad. In combination with SFI, SRY strongly activates the eALDI enhancer, but not eSR-A or eSR-B (Fig. 4a). eSR-A and eSR-B combined showed no response to SRY plus SFI (Fig. 4a) and when eALDI is present in tandem with eSR-A, eSR-B, or both, there was no synergistic increase in activity (beyond the accumulative total), suggesting that for initiation of SOX9 expression by SRY, eALDI is the primary enhancer.

Following SRY initiation, SOX9 in combination with SFI can regulate its own expression. Analysis of the tandem enhancers in the presence of SOX9 and SFI shows that in pairs, eSR-A and eSR-B, eSR-A and eALDI, and eSR-B and eALDI all synergise (1.6, 1.5 and 1.8 fold respectively) (Fig. 4b). Together, all three enhancers showed the greatest synergy (2 fold) (Fig. 4b), suggesting that all three human enhancers have a combined role in the positive autoregulation of SOX9 expression.

To confirm whether the three enhancers were active in a human chromatin context, we performed ChIP-PCR using H3K27ac, a marker for active enhancers\(^30\) on a human testis carcinoma cell line (NT2-D1) which, in the absence of human tissues, acts as a proxy for human embryonic testicular material. The ChIP revealed that eSR-A, eSR-B and eALDI are all present in active enhancer chromatin, with eSR-A and eSR-B showing strongest H3K27ac enrichment relative to the control (Supplementary Figure 8).

Together, our results suggest a model by which each enhancer has an essential yet distinct role in initiating and maintaining human SOX9 expression. In this model, eALDI is the primary enhancer by which SRY and SFI act to initiate SOX9 expression; SOX9 alone or in combination with SFI can then upregulate expression via all three enhancers, which may together act on the SOX9 promoter through chromatin looping events (Fig. 4c).

**Discussion**
In this study, we have identified human enhancers that contribute to SOX9 expression, and in which genetic aberrations lead to sex reversal. The first two enhancers, eSR-A and eSR-B, were located using new patient CNV data. Both of these enhancers show activation by SOX9 alone or in combination with SFI, but little

testis cords, in particular in the Sertoli cells (indicated by co-
staining with the Sertoli marker SOX9) (Supplementary Figure 4h & i). In the ovary at E12.5 and E13.5, the reporter gene was co-expressed with the granulosa cell marker FOXL2 (Supplementary Figure 4d–g). This suggests that the human eSR-B is responsive to factors in the gonads from E12.5 onwards, consistent with response to SF1 + SOX9 and SOX9 alone in vitro. eSR-B shares 75% sequence conservation in mouse. Like human eSR-B, mouse eSR-B demonstrated activation by mouse SOX9 in in vitro luciferase assays, albeit to a lesser extent (Supplementary Figure 5a, b). We therefore generated a CRISPR/Cas9 deletion of the mouse eSR-B region. These mice showed no obvious gonadal or sex reversal phenotype at either embryonic or adult stages (Supplementary Figure 6a–p) and no significant changes in Sox9, Wnt4, Foxl2 or Amh mRNA expression levels (Supplementary Figure 6q-t). This suggests that eSR-B is a human-specific SOX9 gonadal enhancer that when disrupted can cause sex reversal.

**Human SOX9 testis enhancer eALDI is activated by SRY and SF1.** As both eSR-A and eSR-B enhancers are duplicated in SRY-negative 46,XX DSD patients, they can be activated without SRY. Indeed, our in vitro data have shown that eSR-A and eSR-B have a greater response to SOX9 + SF1 than to SRY + SF1. Therefore, in an effort to locate an SRY-responsive enhancer of SOX9 in humans we performed a bioinformatics screen, looking for conserved enhancer marks upstream of SOX9 from publicly available databases, with particular interest the in region in which TESCO lies. 1.4 kb upstream of human TESCO, we discovered a locus containing a DNaseI hypersensitive site in embryonic testis and with strong active enhancer marks (Fig. 3a). Importantly, this region contains a putative SRY binding site and is highly conserved (Fig. 3a).

This 1259 bp fragment was significantly activated by co-
transfection with SRY + SF1 or SOX9 + SF1, but did not respond to SOX9 alone (Fig. 1b). It showed significantly stronger transcriptional activation in vitro than human TESCO, which showed very little activity in this assay (Fig. 3b). Mutation of the SRY binding site significantly reduced activity when co-
transfected with SRY + SF1 or SOX9 + SF1 (Fig. 3c), and addition of FOXL2 markedly repressed this activity (Fig. 3d). This 1259 bp human fragment shows 78% sequence conservation with mouse. CRISPR/Cas9 mediated deletion of the orthologous 1509 bp mouse sequence resulted in a significant reduction in Sox9 expression levels at E11.5 and E14.5 (Fig. 3e). However, this did not alter the expression levels of Amh, Wnt4 or Foxl2 (Supplementary Figure 7). Together, these data indicate that the 1259 bp fragment contains a third human enhancer important for SRY initiation of SOX9 expression in the developing testis.

We have designated this enhancer Alternate Long-Distance Initiator (eALDI) of SOX9.

**Fig. 2** Identification and analysis of the SOX9 enhancer eSR-B. a Schematic of a number of previously published 46,XX DSD duplication/triplication (grey)\(^{16,18,20,21}\) and 46,XY DSD deletion (blue)\(^{17,22}\) patients upstream of SOX9. The minimal overlap (RevSex) of the six patients (24 kb) in magma. b Unbiased luciferase tiling strategy of the minimal RevSex region (16 constructs—blue) to test for enhancer activity using the PGL4-fg vector. c Two fragments, b8 and b9, showed significant enhancer activity and their overlapping fragment, eSR-B was sub-cloned (blue lines). A conserved SOX9 binding site was identified (magenta line and box) and UCSC genome browser ENCODE data (HMEC) highlighted a strong active enhancer site (orange), conservation track indicated a conserved sequence (blue peaks). No DNaseI peaks were detected in human foetal testis or ovaries (green). d Enhancer activities of small fragments (b1-b16) of the minimal overlap in COS7 cells transfected with SFI and SOX9 in luciferase assays. Only the fragments b8 and b9 showed significant enhancer activity compared to the empty vector PGL4-fg (n = 3). e Enhancer activities of b8 and b9 overlapping region eSR-B as assessed by luciferase assays in COS7 cells co-transfected with SFI and SOX9 alone (grey bars) and in combination (blue bars) (n = 4). Significance is compared to the PGL4-fg empty vector control. f Enhancer activity of eSR-B fragment with mutation of the SOX9 (eSR-B ΔSOX9) binding site motif in response to SOX9 transfection (n = 7) significance compared to the eSR-B non-mutated fragment. g Luciferase reporter assays show FOXL2 transfected with SOX9 represses eSR-B enhancer activity compared to SOX9 alone (n = 4). All luciferase assays are carried out in COS7 cells. Error bars are s.e.m. P-values derived from two-tailed t tests: *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. Source data are provided as a Source Data file.
Fig. 3 Identification and analysis of the SOX9 enhancer eALDI. a Schematic of the upstream region of SOX9 highlighting human TESCO (hTESCO) and the proposed enhancer eALDI. Bioinformatic data from UCSC genome browser. ENCODE data (HMEC) indicated a strong active enhancer (orange) covering the proposed eALDI region. DNaseI peaks were observed in one foetal testis track (dark green), and strong conservation across the region was observed (blue peaks). A predicted SRY transcription factor binding motif is presented by magenta vertical line and box. By comparison ENCODE and DNaseI data for hTESCO show less enhancer potential and lower levels of conservation. b Enhancer activity of eALDI and hTESCO measured by luciferase assays in COS7 cells when co-transfected with SOX9 alone (grey bars), SOX9 + SF1 (blue bars) and SF1 + SRY (green bars). The eALDI region shows significant enhancer activity in response to both SF1 + SOX9 and SF1 + SRY transfection when compared to the empty pgl4-β (n = 4). c Mutation of the SRY (eALDIΔSRY) transcription factor motif in the eALDI enhancer reduced activity when co-transfected with SF1 + SOX9 or SF1 + SRY compared to unmuted eALDI. Error bars represent s.e.m. P-values derived from two-tailed t tests: **P ≤ 0.01, ****P ≤ 0.0001. d Luciferase reporter assays in COS7 show FOXL2 transfected with SF1 + SOX9 represses eALDI enhancer activity compared to co-transfection with SOX9 + SF1 alone (n = 3) Error bars are s.e.m. P-values derived from two-tailed t tests: ****P ≤ 0.0001. e CRISPR/Cas9 mediated knock-out of eALDI resulted in reduction of Sox9 expression to 40% of wild-type at E11.5 and 51% of wild-type levels in mouse testis at E14.5 as assessed by qRT-PCR. At 11.5 n = 6 XY WT, XY eALDI n = 6 and XX WT n = 3. At E14.5 XY WT n = 6, XY eALDI n = 6 and XX WT n = 4. Error bars are s.e.m. P-values derived from t-tests: *P ≤ 0.05 and ****P < 0.0001 compared to wild-type males. Source data are provided as a Source Data file.
Saw the greatest enhancer synergy with activation by SOX9 and eSR-B and eALDI to amplify testicular Sox9 expression. In vivo, activation by SRY, the initiator of testis differentiation. eSR-A and eSR-B must logically be active in the absence of SRY.

Following this SRY-dependent development, bind eALDI to initiate Sox9 expression. We found a 60% reduction of Sox9 levels, respectively, with no sex reversal.

We have shown that the mouse eSR-A enhancer is activated by mouse SRY and SF1 in luciferase assays, in contrast to the human enhancer eSR-A which does not respond to SRY and is more likely involved in the upregulation and maintenance of Sox9. Thus, while both the human and mouse enhancers have a key role in testis-specific expression of Sox9, the way in which they are activated differs between the species. Unlike eSR-A, loss of function analysis of mouse eSR-B was uninformative and ChIP-seq for H3K27ac and Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) analysis of E13.5 Sertoli cells in the mouse showed no potential enhancer elements in the mouse equivalent eSR-B.

These human enhancers also show variable conservation in mouse. eSR-B drove reporter gene (LacZ) expression in mouse testis, and also strikingly in embryonic ovaries at different stages of development. It is possible that eSR-B drove gonad specific expression but require as yet unidentified repressor elements to ensure silencing in the ovary. A recent study has shown that loss of the mouse orthologue of eSR-A results in complete sex reversal (Enh13), demonstrating that this is a highly conserved essential testis enhancer. We have shown that the mouse eSR-A enhancer is activated by mouse SRY and SF1 in luciferase assays, in contrast to the human enhancer eSR-A which does not respond to SRY, and is more likely involved in the upregulation and maintenance of Sox9. Thus, while both the human and mouse enhancers have a key role in testis-specific expression of Sox9, the way in which they are activated differs between the species. Unlike eSR-A, loss of function analysis of mouse eSR-B was uninformative and ChIP-seq for H3K27ac and Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) analysis of E13.5 Sertoli cells in the mouse showed no potential enhancer elements in the mouse equivalent eSR-B.

These human enhancers also show variable conservation in mouse. eSR-B drove reporter gene (LacZ) expression in mouse testis, and also strikingly in embryonic ovaries at different stages of development. It is possible that eSR-B drove gonad specific expression but require as yet unidentified repressor elements to ensure silencing in the ovary. A recent study has shown that loss of the mouse orthologue of eSR-A results in complete sex reversal (Enh13), demonstrating that this is a highly conserved essential testis enhancer. We have shown that the mouse eSR-A enhancer is activated by mouse SRY and SF1 in luciferase assays, in contrast to the human enhancer eSR-A which does not respond to SRY, and is more likely involved in the upregulation and maintenance of Sox9. Thus, while both the human and mouse enhancers have a key role in testis-specific expression of Sox9, the way in which they are activated differs between the species. Unlike eSR-A, loss of function analysis of mouse eSR-B was uninformative and ChIP-seq for H3K27ac and Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) analysis of E13.5 Sertoli cells in the mouse showed no potential enhancer elements in the mouse equivalent eSR-B.

As mouse eSR-B shows some enhancer activity in luciferase assays, mouse eSR-B may have a redundant or shadow enhancer role in mice at a time point yet to be examined by ChIP or ATAC-seq.

In mice, Testis Specific Enhancer of Sox9 (TES) with its 1.4 kb core, TESCO, lies 13 kb upstream of the Sox9 gene, and can enhance Sox9 expression in response to SF1, SRY and Sox9 in mice. By contrast, we have shown in vitro that sequences homologous to TES/TEESCO do not appear to play a role in human Sox9 regulation. Instead, the functional characteristics of human eALDI show a striking resemblance to those of TESCO in mice. Both are activated most strongly by Sox9 and SRY, and to a lesser extent, Sox9 and SF1 alone. Both eALDI and TESCO are important for Sox9 expression levels but are not critical for male sex determination in mice, as revealed by CRISPR/Cas9 targeting. We found a 60% reduction of Sox9 at 11.5 dpc in the eALDI null mice which is very similar to that observed in the TESCO/TEESCO mutant mice, which express Sox9 at 62% and 44% of wild-type Sox9 levels, respectively, with no sex reversal.

Threshold levels for Sox9 vary between human and mice as human patients heterozygous for Sox9 mutations often display 46,XY DSD sex reversal while mice lacking one copy of Sox9 do not. It is possible that the functional similarities between eALDI and TESCO may mean that eALDI is the TESCO equivalent in humans/primates. To date we have not identified CNVs in the
eALDI region in DSD patients, yet given the close proximity of eALDI to TESCO and the SOX9 gene, any CNV would likely affect both enhancers and/or the gene itself. However, we will continue to screen DSD patients for alterations in this region.

In summary, this is the first report of human testis-specific enhancers that, if duplicated or deleted lead to 46,XX or 46,XY DSD (sex reversal), respectively, with no effect on other developmental processes that require SOX9 (i.e absence of skeletal abnormalities or other features of campomelic dysplasia). For both eSR-A and eSR-B, loss of one copy in 46,XY DSD patients appears sufficient to prevent upregulation or maintenance of SOX9 expression to the levels required to ensure proper testis development. By contrast, a single additional copy of either of these enhancers promotes the expression of SOX9 to a level that is sufficient to override the ovarian programme causing 46,XX testicular or ovotesticular DSD. We envisage that this occurs through the activation of the duplicated enhancer(s) by SF1 and basal levels of SOX9 present in the XX foetal gonad to ultimately human male development.

Luciferase assay data analysis. For all luciferase assays we normalised the Firefly luciferase values to the Renilla luciferase values (control for transfection efficiency) and then to the average of the empty transcription factor vector (pcDNA3.1) control. All experiments were carried out in duplicate. We calculated standard deviations of the mean (s.e.m.) and two-tailed Students t-test from at least three biological replicates (independent transfections) using Graphpad Prism V7.

Enhancer-lacZ transgenics for eALDI and eSR-B. Transgenic constructs were prepared by cloning the putative enhancer (eALDI or eSR-B) in front of a human β-galactosidase promoter and the lacZ reporter. Transgene expression was then confirmed using standard procedures, and amplified expression was prepared by standard42 or piggybac-assisted43 means and gonads were collected and analysed between E11.5-E14.5. In the case of eSR-B and eALDI stable transgenic lines were also generated (a total of 11 and 3, respectively) and gonads were isolated from founder offspring embryos at various time points and analysed. One eSR-B-lacZ line robustly expressed the transgene and was analysed further; none of the three eALDI lines showed expression.

β-galactosidase staining. Transgenic gonads with associated mesonephric from control embryos, or from the eSR-B-lacZ line at various time points, were stained for β-galactosidase activity using standard methods and incubating overnight at 37 °C, as previously described44.

Sequence conservation of the three enhancers. Sequence conservation analysis (www.ncbi.nlm.nih.gov/BLAST/) of the 1259 bp human eALDI sequence, located 16 kb 5’ to SOX9 (Homo sapiens chromosome 17: 70,100,240-70,101,498: ref NC00017.11), identified a 1176 bp sequence of 75% similarity located 14.4 kb 5’ to mouse Sox9 (Mus musculus strain C57BL/6), chromosome 11: 112,669,990-112,768,165: ref NC00017.11. The remaining sequence was protected by standard42 or piggybac-assisted43 means and gonads were collected and analysed between E11.5-E14.5. In the case of eSR-B and eALDI stable transgenic lines were also generated (a total of 11 and 3, respectively) and gonads were isolated from founder offspring embryos at various time points and analysed. One eSR-B-lacZ line robustly expressed the transgene and was analysed further; none of the three eALDI lines showed expression.

Sequence conservation analysis of the enhancers. Sequence conservation analysis (www.ncbi.nlm.nih.gov/BLAST/) of the 1314 bp human eSR-A sequence to the mouse genome identified that eSR-A (hg19, chr17:69480137-69481650) aligned to mouse chromosome 17: 71,548,908: ref NC00007.6). Sequence conservation of the 402 bp human eSR-B sequence, located 572 kb 5’ to SOX9 (H. sapiens chromosome 17: 71,548,507-71,548,908: ref NC00017.11), identified a 392 bp sequence of 75% similarity located 511 kb 5’ to mouse Sox9 (M. musculus strain C57BL/6) chromosome 11: 112,270,933-112,271,324: ref NC00007.6).

Mutation of mouse eALDI and eSR-B. Using the homology analysis described above two 16-mer single guide RNAs (sgRNAs) were designed (GT-Scan, http://gt-scan.broadinstitute.org/), to flank and delete the 1167 bp mouse sequence by CRISPR/Cas9 editing45. A perfect excision between the two targeted Cas9 cleavage sites would remove 1368 bp of sequence encompassing the 1176 bp mouse ‘eALDI’ sequence.

Using the homology analysis described above two 16-mer single guide RNAs (sgRNAs) were designed (www.crispr.mit.edu/) to flank and delete the 392 bp mouse eSR-B sequence by CRISPR/Cas9 editing. A perfect excision between the two sgRNA target cleavage sites would remove 541 bp of sequence encompassing the 392 bp mouse eSR-B element.

sgRNAs and Cas9 mRNA were prepared as described elsewhere46. Briefly, for each sgRNA the complementary pair of oligos (Extended Data Table 2) was annealed and cloned into pX330 (Addgene #42230). A T7-sgRNA PCR product was amplified with a T7 promoter sequence introduced on the forward primer in conjunction with a universal reverse primer sgRNA-unr (Supplementary Table 2). This product was used as the template for in vitro transcription (IVT) using the MEGASHortscript T7 IVT kit (Life Technologies). The Cas9 coding region was released from pX330 and sub-cloned into pbLuscript II (pB-Cas9). XhoI linearised pBS-Cas9 was used as the template for IVT using the mMESSAGE mMACHINE T7 ULTRA kit (Life Technologies). Both sgRNAs and Cas9 mRNA were purified using the MEGAClear kit (Life Technologies) and eluted in RNase-free water. 30 ng/μl of Cas9 mRNA and 15 ng/μl of each sgRNA diluted in RNase-free water were injected into one pronucleus of C57Bl/6 x CBA F1 hybrid one-cell

Mouse studies. All animal work was conducted according to protocols approved by the University of Queensland Animal Ethics Committee. Embryos were collected with noon of the day a mating plug was observed designated 0.5 days post coitum (dpc). Chromosomal sex was determined by PCR genotyping as described41.
embryos. Injected embryos were cultured overnight to the two-cell stage, then surgically transferred into oviducts of day-of-plug pseudopregnant CD1 mice. Transgenic founders with a full deletion of the eSR-B and eALDI sequence were identified amongst live born mice by PCR genotyping (all genotyping primers are listed in Supplementary Table 2). All genotyping was performed using genomic DNA extracted from tail tip of embryonic or adult tissue of progeny animals. Male founders were bred to C57BL/6 females to verify that deleted alleles were carried by germline transmitted to progeny. eSR-B and eALDI deletions were verified by cloning and sequencing of amplified products from heterozygous offspring of founders. Stable lines were established from two male founders (eSR-B), one bearing a 502 bp deletion and the other bearing a 566 bp deletion, both of which encompassed the deletion. While only one male founder eALDI bearing a 1590 bp deletion. For all embryos and pups, chromosomal sex was determined by genotyping with primers Sex-F and Sex-R (Supplementary Table 2).

qRT-PCR to assess CRISPR knockouts. Gonadal tissue was isolated and subjected to RNA synthesis (Quagen RNeasy micro kit including DNase treatment) and cDNA synthesis (Applied Biosystems (ABI), High Capacity cDNA Archive kit). Relative cDNA levels were determined by the 2^(-ΔΔCt) method with reactions including cDNA master mix (ABI) and Taqman gene expression sets and carried out on an ABI Vii7 machine. The control gene used for normalisation was Tbp encoding the ubiquitously expressed TATA box binding protein, Mm00446973_m1. Primers are listed in Supplementary Table 2. Statistical significance was determined using Prism GraphPad software (t-test).

Haematoxylin and eosin (H&E) and immunofluorescence (IF). Whole embryos were recovered and IF analyses were carried out on fixed, paraffin-embedded 7 mm sections. Immunohistochemical staining was performed on 7μm-thick sagittal sections (E14.5 and 6 weeks) according to standard protocols. IF was conducted using high pH antigen retrieval and blocking in 10% heat inactivated horse serum in PBTX. Primary antibodies used were: goat anti-AMH (1:500, Santa Cruz Biotechnology), rabbit anti-FOXL2 (1:650, generated as described previously), mouse anti-MVH (1:1000, code ab27591, Abcam), goat anti-FOXL2 (1:200, code ab5096, Abcam), rabbit anti-SOX9 (1:200, code AB5535, Millipore), chicken anti-b-galactosidase (1:200, code ab9361, Abcam). Secondary antibodies (all from Molecular Probes, Invitrogen) were goat anti-mouse IgG Alexa Fluor 594, goat anti-rabbit IgG Alexa Fluor 488, and goat anti-chicken IgY Alexa Fluor 488, donkey anti-rabbit Alexa Fluor 647, donkey anti-goat Alexa Fluor 488. Slides were imaged on an Olympus BX-51 fluorescence microscope.

Chromatin immunoprecipitation. The human Sertoli-like cell line NT2/D1 was cultured in DMEM/F12 medium (Invitrogen, Carlsbad, CA, USA) with 10% foetal bovine serum and 1% penicillin-streptomycin at 37 °C with 5% CO2. Cells (5 × 10⁷) cultured in DMEM/F12 medium (Invitrogen, Carlsbad, CA, USA) with 10% foetal bovine serum and 1% penicillin-streptomycin at 37 °C with 5% CO2. Cells (5 × 10⁷)

ChIP-PCR analysis. The purified immunoprecipitated DNA was subjected to PCR analysis to identify the degree of histone H3K27ac-binding within eSR-B, eSR-B and eALDI DNA extracted from chromatin fractions prior to immunoprecipitation (i.e. input DNA) was included in the ChIP analysis. Primers were designed using Primer-BLAST to amplify ~150 bp products (Supplementary Table 3). Exon 1 of the GAPDH gene was used as a positive control (primers from Abcam plc, Cambridge, UK), as it has high levels of histone modifications associated with active gene transcription and low levels of histone modifications associated with gene silencing. A negative control for the PCR reactions with no template added was also included. PCR products were electrophoresed on a 2% agarose gel and imaged using a Gel Doc XR+ System (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

Data availability
CGH-array patient data have been submitted to NCBI ClinVar. Patient 1 accession SCV000845738, Patient 2 accession SCV000845739. All unique materials and data are readily available from the authors. The source data underlying Figs. 1d–f, 2b–f, 4a, b and Supplementary Figures 2, 3, 5, 6, 7 and 8 are provided as a Source Data file.

Received: 12 February 2018 Accepted: 22 November 2018
Published online: 14 December 2018

References
1. Sinclair, A. H. et al. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature 346, 240–244 (1990).
2. Koopman, P. M. et al. Male development of chromosomally female mice transgenic for Sry. Nature 351, 117–121 (1991).
3. Jeske, Y. W., Bowles, J., Greenfield, A. & Koopman, P. Expression of a linear Sry transcript in the mouse genital ridge. Nat. Genet. 10, 480–482 (1995).
4. Koopman, P. M. et al. Expression of a candidate sex-determining gene during mouse testis differentiation. Nature 348, 450–452 (1990).
5. Hacker, A., Capel, B., Goodfellow, P. & Lovell-Badge, R. Expression of Sry, the mouse sex determining gene. Development 121, 1603–1614 (1995).
6. Hanley, N. A. et al. SRY, SOX9, and DAX1 expression patterns during human sex determination and gonadal development. Mech. Dev. 91, 403–407 (2000).
7. Moraes da Silva, S. et al. Sox9 expression during gonadal development implies a conserved role for the gene in testis differentiation in mammals and birds. Nat. Genet. 14, 62–68 (1996).
8. Vidal, V., Chaboissier, M.-C., de Rooij, D. G. & Schedl, A. Sox9 induces testis development in XX transgenic mice. Nat. Genet. 28, 216–217 (2001).
9. Bishop, C. E. et al. A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. Nat. Genet. 26, 490–494 (2000).
10. Symon, A. & Harley, V. SOX9; A genomic view of tissue specific expression and action. Int. J. Biochem. Cell. Biol. 87, 18–22 (2017).
11. Sekido, R. & Lovell-Badge, R. Sex determination involves synergistic action of SRY and SEFI on a specific sox9 enhancer. Nature 453, 930–934 (2008).
12. Gonen, N. et al. Normal levels of sox9 expression in the developing mouse testis depend on the TES/TESCO enhancer, but this does not act alone. PLoS Genet. 13, e1006520 (2017).
13. Gonen, N. et al. Sex reversal following deletion of a single distal enhancer of Sox9. Science 360, 1469–1473 (2018).
14. Croft, B., Ohnsgren, T. & Sinclair, A. H. The role of copy number variants in disorders of sex development. Sex. Dev. 12, 19–29 (2018).
15. Georg, I. et al. Mutations of the SRY-responsive enhancer of SOX9 are uncommon in XY gonadal dysgenesis. Sex. Dev. 4, 321–325 (2010).
16. Ohnsgren, T. et al. A duplication in a patient with 46,XX ovo-testicular disorder of sex development refines the SOX9 testis-specific regulatory region to 24kb. Clin. Genet. 92, 347–349 (2017).
17. Kim, G. J. et al. Copy number variation of two separate regulatory regions upstream of SOX9 causes isolated 46,XY or 46,XX disorder of sex development. J. Med. Genet. 52, 240–247 (2015).
18. Hyon, C. et al. Refining the regulatory region upstream of SOX9 associated with 46,XX testicular disorders of sex development (DSD). Am. J. Med. Genet. 167, 1851–1858 (2015).
19. Vetro, A. et al. Tests development in the absence of SRY: chromosomal rearrangements at SOX9 and SOX3. Eur. J. Hum. Genet. 23, 1025–1032 (2014).
20. Xiao, B. et al. A rare case of 46, XX SRY-negative male with a similar to 74-kb duplication in a region upstream of SOX9. Eur. J. Med. Genet. 56, 695–698 (2013).
21. Vetro, A. et al. XX males SRY negative: a confirmed case of infertility. J. Med. Genet. 48, 710–712 (2011).
22. Benko, S. et al. Disruption of a long distance regulatory region upstream of SOX9 in isolated disorders of sex development. J. Med. Genet. 48, 825–830 (2011).
23. Roadmap Epigenomics Consortium et al. Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330 (2015).
24. Uda, M. et al. Foxl2 disruption causes mouse ovarian failure by pervasive blockage of follicle development. Hum. Mol. Genet. 13, 1171–1181 (2004).
25. Schinzel, D. et al. The murine winged-helix transcription factor Foxi2 is required for granulosa cell differentiation and ovary maintenance. Dev. Biol. 313, 933–942 (2004).
26. Cox, J. J., Willatt, L., Homfray, T. & Woods, C. G. A SOX9 duplication and familial 46,XX developmental testicular disorder. New Engl. J. Med. 364, 91–93 (2011).
27. Dunham, I. et al. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74 (2012).
28. Ernst, J. et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 473, 43–49 (2011).
29. Yao, B. et al. The SOX9 upstream region prone to chromosomal aberrations causing campomelic dysplasia contains multiple caricature enhancers. Nucleic Acids Res. 43, 5394–5408 (2015).
30. Creyghton, M. P. et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl Acad. Sci. USA 107, 21931–21936 (2010).
31. Cannavó, E. et al. Shadow enhancers are pervasive features of developmental regulatory networks. Curr. Biol. 26, 38–51 (2016).
32. Osterwalder, M. et al. Enhancer redundancy provides phenotypic robustness in mammalian development. Nature 554, 239–243 (2018).
33. Maatouk, D. M. et al. Genome-wide identification of regulatory elements in Sertoli cells. Development 144, 720–730 (2017).
34. López-Hernández, B. et al. Duplication of SOX9 associated with 46,XX ovotesticular disorder of sex development. Reprod. Biomed. Online 37, 107–112 (2018).
35. Houmard, B. et al. Global gene expression in the human fetal testis and ovary. Biol. Reprod. 81, 438–443 (2009).
36. Eggers, S. et al. Disorders of sex development: insights from targeted gene sequencing of a large international patient cohort. Genome Biol. 17, 243–264 (2016).
37. Kent, W. J. et al. The human genome browser at UCSC. Genome Res. 12, 996–1006 (2002).
38. Yee, S.-P. & Rigby, P. The regulation of myogenin gene expression during the embryonic development of the mouse. Genes Dev. 7, 1277–1289 (1993).
39. Mathelier, A. et al. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res. 42, D142–D147 (2013).
40. Ohnesorg, T., Croft, B., Tan, J. & Sinclair, A. H. Using ROADMAP data to identify enhancers associated with disorders of sex development. Sex. Dev. 10, 59–65 (2016).
41. McFarlane, L., Truong, V., Palmer, J. & Wilhelm, D. Novel PCR assay for determining the genetic sex of mice. Sex. Dev. 7, 207–211 (2013).
42. Bowles, J., Cooper, L., Berkman, J. & Koopman, P. Sry requires a CAG repeat domain for male sex determination in Mus musculus. Nat. Genet. 22, 405–408 (1999).
43. Zhao, L., Ng, E. T. & Koopman, P. A piggyBac transposon-and gateway-enhanced system for efficient RAC transgenesis. Dev. Dyn. 243, 1086–1094 (2014).
44. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R. Staining frozen mouse embryo sections for β-galactosidase (lacZ) activity. Cold Spring Harb. Protoc. 2007, pprot4726 (2007).
45. O’Brien, A. & Bailey, T. L. GT-scan: identifying unique genomic targets. Bioinformatics 30, 2673–2675 (2014).
46. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 818–823 (2013).
47. Bowles, J. et al. FGFI9 suppresses meiosis and promotes male germ cell fate in mice. Dev. Cell. 19, 440–449 (2010).
48. Polanco, J. C. et al. SOX10 gain-of-function causes XX sex reversal in mice: Implications for human 22q-linked disorders of sex development. Hum. Mol. Genet. 19, 509–516 (2009).
49. Ye, J. et al. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13, 134–145 (2012).

Acknowledgements
This work was supported by a National Health and Medical Research Council (NHMRC) Program grant 1074258 to A.S., P. K. and an NHMRC Project grant 1031214 to A.S., T.O., ARC grants DP150102120 and DP160104948 to P.K. B.C. was supported by an Australian Government Research Training Program Scholarship. R.S. was supported by NHMRC Early Career Fellowship No. 1126995. NHMRC Fellowships supported A.S. (1062854), P.K. (1059006) and V.H. (102034). Thanks to Dan Belluccio for assistance with design and analysis of the custom CGH-array and the Victorian Government’s Operational Infrastructure Support Program.

Author contributions
A.S., T.O. and B.C. conceived and designed the project to identify and analyse SOX9 enhancers. J.A.H., D.V.C. and J.A.H. recruited and clinically phenotyped patients. J.A.H. and V.C. mapped patient 2 duplication. K.A., J.v.d.B., G.R. and I.K., analysed patient data, designed and ran the CGH-array to map patient deletions and duplications. B.C. and T.O. carried out bioinformatics analysis and as well as cloning and luciferase assays with J.T. J.B., A.Q., E.P. and P.K. designed and analysed mouse experiments. V.H., R.S. and B.C. ran the CHIP analysis. A.S., B.C., T.O. and K.A., compiled and analysed data and wrote the manuscript with help from all authors. All authors contributed intellectual input.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07784-9.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.