Regulatory region upstream of SOX9 and 46,XX DSD. Hyon et al.

Title: Refining the regulatory region upstream of SOX9 associated with 46,XX testicular Disorders of Sex Development (DSD)

Running title: Regulatory region upstream of SOX9 and 46,XX DSD

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

\textsuperscript{1}AP-HP, Hôpitaux Universitaires Est Parisien, Hôpital Trousseau, Service de Génétique et d’Embryologie médicales, 75012, Paris, France
\textsuperscript{2}INSERM UMR_S933, 75012, Paris, France
\textsuperscript{3}UPMC Univ Paris 06, UFR de Médecine Pierre et Marie Curie, 75005, Paris, France
\textsuperscript{4}Service Génétique Médicale, CHU Poitiers, France
\textsuperscript{5}AP-HP, Hôpitaux Universitaires Est Parisien, Hôpital Tenon, Department of Radiology, 75020, Paris, France
\textsuperscript{6}AP-HP, Hôpitaux Universitaires Est Parisien, Hôpital Tenon, Department of Pathology, 75020, Paris, France
\textsuperscript{7}Institut Pasteur, Human Developmental Genetics, 75724, Paris, France
\textsuperscript{8}Service Endocrinologie, CHU Poitiers, France
\textsuperscript{9}Centre de Référence Anomalies du Développement Ouest, CHU Poitiers, France

Corresponding author
Capucine Hyon
Service de Génétique et d’Embryologie Médicales
26 rue du Dr Arnold Netter, 75012 Paris
Tel 00 33 1 44 73 52 80
Fax 00 33 1 44 73 52 82
capucine.hyon@trs.aphp.fr
Abstract

Disorders of Sex Development (DSD) are a heterogeneous group of disorders affecting gonad and/or genito-urinary tract development and usually the endocrine-reproductive system. A genetic diagnosis is made in only around 20% of these cases. The genetic causes of 46,XX-SRY negative testicular DSD as well as ovotesticular DSD are poorly defined. Duplications involving a region located ~600 kb upstream of SOX9, a key gene in testis development, were reported in several cases of 46,XX DSD. Recent studies have narrowed this region down to a 78 kb interval that is duplicated or deleted respectively in 46,XX or 46,XY DSD. We identified three phenotypically normal patients presenting with azoospermia and 46,XX testicular DSD. Two brothers carried a 83,8 kb duplication located ~600 kb upstream of SOX9 that overlapped with the previously reported rearrangements. This duplication refines the minimal region associated with 46,XX-SRY negative DSD to a 40.7-41.9 kb element located ~600 kb upstream of SOX9. Predicted enhancer elements and evolutionary-conserved binding sites for proteins known to be involved in testis determination are located within this region.

Key words
46,XX DSD, duplication, regulatory element, SOX9 gene, CNV
INTRODUCTION

Disorders of sex development (DSD) are congenital conditions in which the development of chromosomal, gonadal, or anatomical sex is atypical [Hughes et al., 2006]. DSD covers a wide spectrum of phenotypes. 46,XY DSD includes 46,XY complete or partial gonadal dysgenesis as well as undermasculinization of an XY male due to a defect in androgen synthesis or action. At the opposite, most 46,XX DSD correspond to the virilization of an XX female due to androgen excess. Most often, 46,XX testicular DSD, a rare pathology affecting 1 in 20,000-25,000 newborn males, [Kousta et al., 2010] are SRY positive and have Y-to-X translocations while most cases of 46,XX ovotesticular DSD are mosaics or chimeras [López et al., 1995]. However, some 46,XX males do not carry the SRY gene and the genetic cause of these DSD cases remains poorly defined.

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

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

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

**MATERIEL AND METHODS**

**Clinical reports**

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

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

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

Genomic DNA was extracted for each patient from peripheral blood samples using FlexiGene DNA Kit (Qiagen, Valencia, CA) according to manufacturer’s instructions. For patients 1, 2 and 3, SNP array analysis was performed using the HumanCytoSNP-12 BeadChip from Illumina (San Diego, CA). Data analysis was performed using Illumina’s Genome Studio Genotyping Module software allowing the identification of both copy number variations (CNVs) with the Log R ration (LRR) and regions of copy-neutral loss of heterozygosity (LOH) with the B allele frequency. Previous array CGH for patient 3 was performed using the Human Genome CGH Microarray Kit 105A from Agilent (Santa Clara, CA). DNA sequence information was made according to the UCSC Genome Browser (http://genome.ucsc.edu/; February 2009, Assembly, hg19).

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

Quantitative PCR

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

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

RESULTS

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

In patient 1 and 2, SNP array analysis identified a 77,078 to 85,175 bp duplication in chromosome 17 long arm, upstream from SOX9 gene, extending between 69,490,856 (normal)-69,493,863 (duplicated) to 69,570,941 (duplicated)-69,576,031 (normal) (Figure 1). For patient 3, array CGH identified an overlapping duplication of 194 kb in the same region which was reduced to a 140,572-152,536 bp between 69,435,809 (normal)-69,441,277 (duplicated) and 69,581,849 (duplicated)-69,588,345 (normal) using the SNP array (Figure 1). These duplications were confirmed by independent qPCR analysis for patients 1 and 2 (data not shown) and sequencing the boundaries of the duplication for patient 2 which allowed us to identify the exact size of the duplication extending between 69,491,366 and 69,575,195 (83,829 bp duplication) (Figure 2).

Xiao et al, described an XX male with a 74 kb duplication overlapping with this region which proximal breakpoint was located between 69,533,305 bp (normal) and 69,534,526 bp (duplicated) [Xiao et al., 2013]. Taking this data and our breakpoint into account we can redefine the minimal region duplicated to a minimum
size of 40,669 bp and a maximum size of 41,890 bp (Figure 3). Detailed analysis of this region indicates that it contains two non-overlapping putative enhancer elements. The Encyclopedia of DNA Elements (ENCODE) project datasets indicate a weak or poised enhancer at chr17:69,568,380-69,569,850. Chromatin at this position is enriched for experimentally determined CTCF-binding from human mammary epithelial cells (HMEC). The primary role of CTCF is thought to be in regulating the 3D structure of chromatin. CTCF creates boundaries between topologically-associating domains in chromosomes and anchors DNA to cellular structures like the nuclear lamina and thereby facilitates interactions between transcription regulatory sequences [Murrell, 2011].

This region is also associated with acetylation of histone H3 at lysine 27 (H3K27ac) and enriched in monomethylation of histone H3 at lysine 4 (H3K4me1) from a variety of cell lines (Figure 3). Both H3K4 methylation and H3K27 acetylation are epigenetic marks that are generally associated with gene activation [Eissenberg and Shilatifard, 2010; Smith and Shilatifard, 2014].

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

Multiz alignment results obtained from the UCSC genome browser show 12 short evolutionary-conserved regions that are enriched for transcription factors binding sites that are present in the human, mouse and rat (Figure 3). These include factors that are strongly expressed in the male supporting cell lineages of the mouse testis at the time of sex-determination (E11.5 to E13.5) and include NFIL3 (also known as E4BP4; chr17:69,544,862-69,544,873), PBX1 (chr17:69,563,419-69,563,427) and GATA1 (chr17:69,560,189-69,560,198 and chr17:69,560,189-69,560,198). In addition, to these factors TRANSFAC and TFSEARCH analysis of the human sequence predicted multiple DNA-binding sites for transcription factors known to be involved in the early stages of sex-determination such as SRY, WT1, SOX9, NR5A1 and LHX9 [Bashamboo and McElreavey, 2013]. The region also contains 4 predicted binding sites for the evolutionary conserved transcription factor DMRT1 at chr17:69,543,454-69,543,468, chr17:69,554,542-69,554,559, chr17:69,559,047-69,559,062 and chr17:69,569,634-69,569,648.

**DISCUSSION**

Multiple tissue specific enhancers located upstream and downstream from SOX9 have been described [Velagaleti et al., 2005; Bagheri-Fam et al., 2006; Gordon et al., 2009]. In the human, a large duplication
Regulatory region upstream of SOX9 and 46,XX DSD. Hyon et al.

(46,XX,dup(17)(q23.1q24.3))/46,XX) involving SOX9 was associated with female to male sex reversal in a 46,XX individual [Huang et al., 1999], however mutations or translocations involving the TESCO region have not yet been reported [Georg et al., 2010]. This suggests that in human gonad additional regulatory elements may be involved in the regulation of SOX9 expression. This hypothesis is supported by recent reports that identified rearrangements involving elements situated ~600 kb upstream of SOX9 in association with 46,XX testicular DSD and termed the RevSex element [Benko et al., 2011; Cox et al., 2011; Vetro et al., 2011; Xiao et al., 2013]. The phenotypes associated with these rearrangements are reported in Table 1. In this study, through the analyses of three XX males and comparison with the published data, we redefined the minimal critical region located ~600 kb upstream of SOX9 associated with XX sex-reversal. The two microduplications that we describe overlap those previously reported in 46,XX testicular DSD as well as deletions in 46,XY undermasculinized DSD patients. Therefore, we can define the critical duplicated region to 69,533,305 bp (normal) and 69,534,526 bp (duplicated) to 69,575,195 bp. This corresponds to a minimum size of 40,669 bp and a maximum size of 41,890 bp. Although this considerably reduces the extent of the RevSex regulatory element, it is unclear precisely how the rearrangements of this region result in changes in the expression of SOX9 in the early developing gonad.

The duplicated region itself contains two predicted enhancer motifs. Although interpretation of ENCODE datasets should be taken with caution since the data was generated using cell lines that are not necessarily relevant for DSD, there are potentially interesting observations. The proximal strong enhancer motif (chr17:69,544,206-69,546,005) is enriched for H3K4 methylation and H3K27 acetylation that are epigenetic marks characteristic of gene activation [Eissenberg and Shilatifard 2010; Smith and Shilatifard, 2014]. This enhancer element binds the histone acetyltransferase EP300 that regulates transcription via chromatin remodelling and it is important in the processes of cell proliferation and differentiation [Ogryzko et al., 1996]. EP300 is strongly expressed in the somatic and germ cell lineages of the XX and XY gonad during sex-determination and it can act as a co-activator of both NR5A1 and SOX9 [Ito et al., 1998; Furumatsu et al., 2009; Munger et al., 2013]. The CH3 domain of EP300 directly associates with the C-terminal PQ-rich transactivation domain of Sox9, and activates Sox9-dependent transcription in chondrogenesis by induction of histone acetylation [Furumatsu et al., 2009]. NR5A1 (also known as SF-1) was originally identified as a master-regulator of steroidogenic enzymes in the early 1990s and controls many key aspects of adrenal and reproductive functions [El-Khairi and Achermann, 2012]. Mutations involving NR5A1 are one of the most common causes of 46,XY DSD and they are associated with a range of phenotypes including 46,XY gonadal dysgenesis and adrenal
Regulatory region upstream of SOX9 and 46,XX DSD. Hyon et al.

failure, 46,XY DSD with apparently normal adrenal function, 46,XY infertile male and 46,XX female with ovarian insufficiency [El-Khairi and Achermann, 2012]. It is possible that SOX9 expression may be regulated through this element involving both NR5A1 and EP300 proteins. Furthermore, this enhancer motif is located between two predicted binding sites for DMRT1 (chr17:69,543,454-69,543,468 and chr17:69,554,542-69,554,559). DMRT1 controls many aspects of testicular development in the mouse and human, including the differentiation, proliferation, migration, and pluripotency of germ cells, and also proliferation and differentiation of Sertoli cells [Matson and Zarkower, 2012]. Haploinsufficiency of DMRT1 in the human is associated with a failure of testis-determination, whilst in the mouse DMRT1 is essential for maintaining mammalian sex-determination by antagonising the ovarian gene regulatory pathway [Ottolenghi and McElreavey, 2000; Matson et al., 2011]. In avians, where DMRT1 is the best candidate testis-determining gene, it is necessary for the expression of SOX9 [Smith et al., 2009; Lambeth et al., 2014]. ChIPseq experiments using mice using cells from P28 testis have demonstrated that Dmrt1 binds to both upstream and downstream of Sox9 and bind near to the genes Wnt4, Foxl2 and Rspo1 suggesting that DMRT1 may initiate or maintain male cell fate by directly activating Sox9, whilst repressing female-promoting genes [Murphy et al., 2010; Matson et al., 2011]. The localisation of DMRT1-binding sites within a putative enhancer element within the RevSex element opens the possibility that the testis formations associated with these duplications could be initiated through DMRT1 activation.

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

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

Acknowledgements

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

The authors declare no conflict of interest.
References

Bagheri-Fam, S, Barrionuevo, F, Dohrmann, U, Günther, T, Schüle, R, Kemler, R, Mallo, M, Kanzler, B,


Kilpatrick, N, Lees, MM, Picard, A, Temple, IK, Thomas, P, Vazquez, M-P, Vekemans, M, Crollius,

Benko, S, Gordon, CT, Mallet, D, Sreenivasan, R, Thauvin-Robinet, C, Brendehaug, A, Thomas, S, Bruland, O,
David, M, Nicolino, M, Labalme, A, Sanlaville, D, Callier, P, Malan, V, Huet, F, Molven, A, Dijoud, F,

Bhagavath, B, Layman, LC, Ullmann, R, Shen, Y, Ha, K, Rehman, K, Looney, S, McDonough, PG, Kim, H-G,


Regulatory region upstream of SOX9 and 46,XX DSD. Hyon et al.


Regulatory region upstream of SOX9 and 46,XX DSD. Hyon et al.


Table 1. Phenotypic features and characteristics of the rearrangement of individuals with rearrangement upstream of SOX9.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SRY</td>
<td>46,XX</td>
<td>46,XX</td>
<td>46,XX</td>
<td>46,XY</td>
<td>46,XY</td>
<td>46,XX</td>
</tr>
<tr>
<td>Phenotype</td>
<td>- Male</td>
<td>- Male</td>
<td>- Male</td>
<td>- Male</td>
<td>- Male</td>
<td>- Male</td>
</tr>
<tr>
<td></td>
<td>- Male</td>
<td>- Male</td>
<td>- Male</td>
<td>- Male</td>
<td>- Male</td>
<td>- Male</td>
</tr>
<tr>
<td></td>
<td>- Female</td>
<td>- Female</td>
<td>- Female</td>
<td>- Female</td>
<td>- Female</td>
<td>- Male</td>
</tr>
<tr>
<td>Internal</td>
<td>- NA</td>
<td>- NA</td>
<td>- NA</td>
<td>- NA</td>
<td>- NA</td>
<td>- NA</td>
</tr>
<tr>
<td>genitilia</td>
<td>- NA</td>
<td>- NA</td>
<td>- NA</td>
<td>- NA</td>
<td>- NA</td>
<td>- NA</td>
</tr>
<tr>
<td>Sex of rearing</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
</tr>
<tr>
<td>Gonadal histology</td>
<td>- Sertoli and Leydig cells</td>
<td>- Germinal cell aplasia</td>
<td>- NA</td>
<td>- NA</td>
<td>- Streak gonad</td>
<td>- Eosinophilic Sertoli cells</td>
</tr>
<tr>
<td></td>
<td>- Atrophied seminiferous tubules</td>
<td>- Atrophied seminiferous tubules</td>
<td>- NA</td>
<td>- NA</td>
<td>- Eosinophilic Leydig cells</td>
<td></td>
</tr>
<tr>
<td>Type of rearrangement</td>
<td>Duplication</td>
<td>Duplication</td>
<td>Duplication</td>
<td>Duplication</td>
<td>Duplication</td>
<td>Duplication</td>
</tr>
<tr>
<td>Size of the rearrangement</td>
<td>178 kb</td>
<td>96 kb</td>
<td>693 kb</td>
<td>148 kb</td>
<td>780 kb</td>
<td>240 kb</td>
</tr>
</tbody>
</table>

NA : Not available
Regulatory region upstream of SOX9 and 46,XX DSD. Hyon et al.

Legends to figures

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

**Figure 2.** Exact boundaries of the duplication for patient 2 (according to the UCSC Genome Browser (http://genome.ucsc.edu/; February 2009, Assembly, hg19)).
**Figure 3 (A).** UCSC genome browser panel displaying the minimal (red bar) and maximal (blue bar) extend of the revised RevSex duplication upstream of SOX9. Panels show H3K4me1, H3K4me3 and H3K27Ac sites within this region from a variety of cell lines. DNaseI hypersensitive sites and transcription factor ChIP-seq binding sites are indicated. Proximal and distal enhancer elements are shown as well as Multiz conservation alignments. **(B)** UCSC genome browser panel displaying evolutionary conserved DNA transcription factor binding sites (see text for details). Histone modifications correspond to the following cells and cell lines –

<table>
<thead>
<tr>
<th>GM12878</th>
<th>H1-hESC</th>
<th>HSMM</th>
<th>HUVEC</th>
<th>K562</th>
<th>NHEK</th>
<th>NHLF</th>
</tr>
</thead>
</table>
Regulatory region upstream of SOX9 and 46,XX DSD. Hyon et al.

GM12878 - B-lymphocyte, lymphoblastoid, International HapMap Project - CEPH/Utah - European Caucasian, Epstein-Barr Virus

H1-hESC – human embryonic stem cells

HSMM – human skeletal muscle myoblasts

HUVEC – human umbilical vein endothelial cells

K562 – human leukaemia cell line

NHEK – normal human epidermal keratinocytes

NHLF – normal lung fibroblasts

Color figure can be viewed in the online issue, which is available at

http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1552-4833.}