Copy Number Variation in Patients with Disorders of Sex Development Due to 46,XY Gonadal Dysgenesis

Stefan White1,2, Thomas Ohnesorg1, Amanda Notini1, Kelly Roeszler1, Jacqueline Hewitt1,2, Hinda Daggag1, Craig Smith1, Erin Turbitt1, Sonja Gustin1, Jocelyn van den Bergen1, Denise Miles1,2, Patrick Western1, Valerie Arboleda3, Valerie Schumacher4, Lavinia Gordon1, Katrina Bell1, Henrik Bengtsson5, Terry Speed5, John Hutson1,2, Garry Warne1,2, Vincent Harley6, Peter Koopman7, Eric Vilain8*, Andrew Sinclair1,2*

1 Murdoch Childrens Research Institute, Royal Children’s Hospital, Melbourne, Victoria, Australia, 2 Department of Paediatrics, The University of Melbourne, Melbourne, Victoria, Australia, 3 Department of Medical Genetics, University of California Los Angeles, Los Angeles, California, United States of America, 4 Pediatrics Department, Children’s Hospital, Boston, Massachusetts, United States of America, 5 Walter and Eliza Hall Institute, Melbourne, Victoria, Australia, 6 Prince Henry’s Institute of Medical Research, Melbourne, Victoria, Australia, 7 Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia

Abstract

Disorders of sex development (DSD), ranging in severity from mild genital abnormalities to complete sex reversal, represent a major concern for patients and their families. DSD are often due to disruption of the genetic programs that regulate gonad development. Although some genes have been identified in these developmental pathways, the causative mutations have not been identified in more than 50% 46,XY DSD cases. We used the Affymetrix Genome-Wide Human SNP Array 6.0 to analyse copy number variation in 23 individuals with unexplained 46,XY DSD due to gonadal dysgenesis (GD). Here we describe three discrete changes in copy number that are the likely cause of the GD. Firstly, we identified a large duplication on the X chromosome that included DAX1 (NR0B1). Secondly, we identified a rearrangement that appears to affect a novel gonad-specific regulatory region in a known testis gene, SOX9. Surprisingly this patient lacked any signs of campomelic dysplasia, suggesting that the deletion affected expression of SOX9 only in the gonad. Functional analysis of potential SRY binding sites within this deleted region identified five putative enhancers, suggesting that sequences additional to the known SRY-binding TES enhancer influence human testis-specific SOX9 expression. Thirdly, we identified a small deletion immediately downstream of GATA4, supporting a role for GATA4 in gonad development in humans. These CNV analyses give new insights into the pathways involved in human gonad development and dysfunction, and suggest that rearrangements of non-coding sequences disturbing gene regulation may account for significant proportion of DSD cases.

Introduction

A defining point during embryogenesis is the commitment to develop as male or female. In males this is initiated by the Y-linked SRY gene, which leads to testis development. Ovarian development occurs in the absence of the Y-linked SRY gene, and ultimately results in a female phenotype. These developmental pathways involve complex networks of genes, the precise regulation of which is vital for the correct development of the gonads and associated anatomical structures [1].

Disruption of these networks can lead to disorders of sex development (DSD), which are congenital conditions with atypical development of the chromosomal, gonadal or anatomical sex [2]. DSD can be divided into three etiological subclasses, namely syndromic, disorders of androgen action, and gonadal dysgenesis [3]. The focus here is on individuals with 46,XY DSD due to gonadal dysgenesis (hereafter referred to as 46,XY GD).

Individuals with 46,XY complete gonadal dysgenesis (CGD) are phenotypically female, have completely undeveloped streak gonads, and are often not diagnosed until puberty when secondary sexual characteristics fail to develop. Mutations in SRY and SOX9 account for approximately 20% of 46,XY CGD patients [4]. Causative mutations affecting several other genes have been identified [5], including DAX1 (NR0B1) [6], SF1 (NR5A1) [7], WNT4 [8], DHH [9] and MAP3K1 [10]. Mutations in these genes are responsible for fewer than 30% of cases [11] and little is known about the underlying genetic basis in the remaining 50% of patients.

Individuals with 46,XY partial gonadal dysgenesis (PGD) and ovotesticular DSD have a phenotype that can range from a mildly undervirilised male through ambiguous genitalia to a slightly virilised female phenotype, depending on the amount of residual testicular function present. The diagnosis is therefore usually made at birth of the child with anomalous genitalia. In 46,XY ovarian
DSD the phenotype is entirely female, including presence of female gonads. A diagnosis may be made due to presentation with early ovarian failure.

There have been intensive efforts in the last decade to identify novel genes involved in gonad differentiation, using a range of animal models [12,13]. Although a large number of candidate genes have been identified, few mutations have been identified in these genes in human DSD patients. Even less is known about the gonad-specific regulation of these genes, although recent work has identified one gonad-specific enhancer of SOX9 in the mouse[14]. It is currently unclear whether unexplained cases of DSD are due to mutations in novel gonadal genes or mutations in the regulatory regions of known gonadal genes.

The genetic basis of several diseases has been elucidated by rare cases involving large, cytogenetically visible deletions or translocations that identified specific chromosomal regions for further analysis. For example, the region of the Y chromosome carrying the SRY gene was pinpointed by examining 46,XX testicular DSD patients with translocations of Y chromosome material to the X chromosome and 46,XY CGD patients who had deletions of the Y chromosome [15].

Copy number variation (CNV) is a term used to describe rearrangements of the genome such as deletions and duplications that result in an increase or decrease in the effective copy DNA number. The development of microarrays has allowed copy number analysis of the genome at a much finer resolution than was possible by cytogenetic analysis. This approach has been used to identify new disease genes in a range of disorders [16,17]. This methodology has been applied to a few isolated DSD cases [18,19], and has recently been used for studying diverse cohorts of DSD patients[20,21]. Here we have used the Affymetrix Genome-Wide Human SNP Array 6.0 to identify copy number variants in 23 patients with DSD due to 46,XY gonadal dysgenesis.

### Materials and Methods

#### Patient Information

Twenty-three unrelated patients were diagnosed as having 46,XY GD. All patients with a diagnosis of XY complete gonadal dysgenesis met the clinical criteria for this diagnosis, including female external genitalia associated with a 46, XY karyotype, the presence of Mullerian structures suggesting the lack of functional testicular tissue, and high levels of gonadotropins suggesting a primary gonadal failure caused by gonadal dysgenesis. Additional clinical features, when present, (such as adrenal hypoplasia), are indicated in Table 1.

No mutation was detected in the SRY coding sequence in any of the 23 patients. For cases 14 and 21 (diagnosed with congenital adrenal hypoplasia) the DAX1 and SF1 genes had also been checked, and no mutation was identified. Genomic DNA from cases 1–8 was isolated from lymphoblastoid cell lines using standard methods. Genomic DNA from cases 9-23 was isolated

| Case # | Diagnosis Other clinical features | Best candidate CNV (size/type) | Candidate gene(s) |
|--------|----------------------------------|--------------------------------|-------------------|
| 1      | 46,XY ovotesticular DSD          |                                |                   |
| 2      | 46,XY partial gonadal dysgenesis |                                |                   |
| 3      | 46,XY complete gonadal dysgenesis|                                |                   |
| 4      | 46,XY complete gonadal dysgenesis|                                |                   |
| 5      | 46,XY partial gonadal dysgenesis |                                |                   |
| 6      | 46,XY gonadal dysgenesis         |                                |                   |
| 7      | 46,XY partial gonadal dysgenesis |                                |                   |
| 8      | 46,XY ovarian DSD                |                                |                   |
| 9      | 46,XY complete gonadal dysgenesis| Cleft palate, short stature    | SOX9              |
| 10     | 46,XY complete gonadal dysgenesis| Chr17:66200578-67393626        |                   |
| 11     | 46,XY complete gonadal dysgenesis| Chr10:12382107-12770026        | CAMK1D            |
| 12     | 46,XY complete gonadal dysgenesis| ChrX:30131772-30902339         | DAX1 (NR0B1)      |
| 13     | 46,XY complete gonadal dysgenesis| Chr8:11659702-11694481         | GATA4 (upstream of deletion) |
| 14     | 46,XY complete gonadal dysgenesis| Chr13:42568370-42610053        | DNAJC15           |
| 15     | 46,XY complete gonadal dysgenesis| Chr17:66200578-67393626        |                   |
| 16     | 46,XY complete gonadal dysgenesis| Galactosemia                    |                   |
| 17     | 46,XY complete gonadal dysgenesis| Short stature                   |                   |
| 18     | 46,XY complete gonadal dysgenesis| Amelia                          |                   |
| 19     | 46,XY complete gonadal dysgenesis| Adrenal Hypoplasia Congenita    |                   |
| 20     | 46,XY complete gonadal dysgenesis| IMAGE syndrome                  |                   |
| 21     | 46,XY complete gonadal dysgenesis| Adrenal Hypoplasia Congenita    |                   |
| 22     | 46,XY complete gonadal dysgenesis| IMAGE syndrome                  |                   |
| 23     | 46,XY complete gonadal dysgenesis| IMAGE syndrome                  |                   |

doi:10.1371/journal.pone.0017793.t001
from lymphocytes using standard methods. This study was approved by the Human Research Ethics Committee of the Royal Children’s Hospital Melbourne, Australia (# 22073D).

Affymetrix Microarray Analysis

All 23 DNA samples were hybridised onto the Affymetrix Genome-Wide Human SNP Array 6.0 microarrays. These arrays are composed of ~1.8 million markers, evenly targeted at single nucleotide polymorphisms (SNPs) and other genomic regions allowing the identification of copy number variants. Microarray hybridisations were performed at the Australian Genome Research Facility (Melbourne, Australia) following manufacturer’s instructions.

Data were analysed using an early-access version of the CRMAv2 total copy number (CN) method [22]. More precisely, crosstalk between alleles in (SNPs) was controlled for and global offset was removed from both SNPs and CN probes. The non-polymorphic SNP signals were estimated as the median probe-pair sum across replicates allele probe pairs (ignoring strand information). PCR fragment-length effects in the non-polymorphic SNP and CN probe signals were normalized. Total CNs were calculated as the ratio of the non-polynomial signals relative to the robust average of all hybridizations. Chromosomal aberrations were identified from log2 CN ratios using the Gain and Loss Analysis of DNA (GLAD) [22]. Duplications and deletions were called using log2-thresholds of +0.3 and −0.3, respectively, containing a minimum of 10 consecutive probes. The above analysis was conducted using aroma.affymetrix [23].

Multiplex Ligation-dependent Probe Amplification (MLPA)

CNVs identified by microarray analysis were validated by Multiplex Ligation-dependent Probe Amplification (MLPA). Probes were designed according to previously described criteria [24]. Oligonucleotides were ordered from Sigma Genosys (Australia), and were desalted without further purification. The right hand oligonucleotide of each pair was 5’ phosphorylated to allow ligation to occur. Probe mixes were prepared by combining each oligonucleotide so that all were present at a final concentration of 4 fmol/µl in TE −4. MLPA reagents were purchased from Fisher Biotec (Australia). MLPA reactions and data analysis were performed as previously described [25,26].

DNA Sequencing

All primers used in sequencing were purchased from Sigma Genosys (Australia). Sequencing was conducted at either the Brisbane node of the Australian Genome Research Facility, or at the Department of Pathology, University of Melbourne.

Putative SOX9 Enhancer

Putative SOX9 and SRY binding sites were identified using the HMR Conserved Transcription Factor Binding Sites track on the UCSC Genome Browser (hg18) with the default settings. PCR primers were designed to amplify seven conserved genomic regions (~0.5–1.5 kb in size) containing putative SOX9 and SRY binding sites (table S1). The PCR fragments were cloned into pGL4.10 minSOX9 (SOX9 minimal promoter) [27] driving the expression of firefly luciferase upstream of the promoter.

Dual Luciferase Assays

Luciferase assays were performed in a recently developed human cell line that expresses several markers of Sertoli cells [28], e.g. SOX9, GATA4 and AMH, with 104 cells plated in 24-well plates. After 24 hours, the cells were transfected with the relevant combination of plasmids using Lipofectamine 2000 (Invitrogen) according to the manufacturer’s instructions. Briefly, cells in each well received 500 ng of luciferase reporter construct (pGL4-minSOX9 with or without putative enhancer fragments), 200 ng of transcription factor expression plasmid (pcDNA-SF1, SOX9, SRY or empty pcDNA), and 50 ng of pRL-TK as transfection control. After 48 hours, cells were harvested and luciferase activity was measured using the Dual Luciferase Assay Kit (Promega) in a TD-20/20 Luminometer (Turner Designs). Values were normalized for transfection efficiency with respect to the effects of the expression constructs on the empty pGL4-minSOX9 reporter plasmid. Relative light units shown represent mean values ± SEM obtained from at least four independent experiments performed in duplicate.

RNA Isolation and Amplification

Gonads were dissected from E12.5-E15.5 Oct4-GFP mouse embryos, dissociated with trypsin and FACS sorted into germ and somatic fractions. Total RNA was isolated and amplified as previously described [29].

Reverse Transcription and Real-Time PCR Analysis

Amplified RNA was reverse transcribed using the Transcriptor High Fidelity cDNA synthesis kit (Roche, Mannheim, Germany). Briefly, 100 ng amplified RNA was mixed with 2 µl of Random Hexamer Primers, denatured at 65°C for 10 minutes and immediately cooled on ice. Samples were reverse transcribed following addition of 4 µl of 5x Transcriptor High Fidelity Reverse Transcritpase Reaction Buffer, 0.5 µl of Protecto RNase Inhibitor, 2 µl of Deoxynucleotide Mix, 1 µl of DTT and 1 µl of Transcriptor High Fidelity Reverse Transcritpase at 30°C for 30 minutes and inactivated at 85°C for 5 minutes.

Real-time PCR was performed in triplicate using 1 ng cDNA in each 10 µl reaction, using the mouse Universal Probe Library system (Roche), LightCycler 480 Probe Master mix (Roche) and a LightCycler 480 instrument (Roche). All primers were designed using the UPL Assay Design Centre (https://www.roche-applied-science.com). Relative expression was determined using the comparative Ct method (ΔΔCt), with samples normalised against Silh and expressed relative to the sample showing the lowest level of expression for each experiment. In addition, the efficiency of each primer/probe combination was determined using standard curves.

Results

Microarray analysis

Twenty-three unrelated patients were diagnosed as having 46,XY GD. No mutation was detected in the SRY coding sequence in any of the patients. We screened the genomic DNA of these patients using the Affymetrix Genome-Wide Human SNP Array 6.0. A stringent threshold of at least 10 consecutive probes was set for a CNV to be called. Using this criterion 1,498 high-confidence CNV regions were identified, with no difference in the average number or size of CNVs detected between DNA isolated from cell lines or lymphocytes (Table 2). Of the CNVs detected, 91% overlapped for at least 50% with previously reported CNVs in the database of genomic variants (DGV). To prioritise regions for further analysis we focussed on rearrangements that a) covered or were within 500 kb of genes known to be involved in gonad development; or b) affected the coding region of RefSeq genes and were not listed in the DGV.
Rearrangements affecting known gonadal genes

Three cases had rearrangements that affected genes known to play a role in sex determination or gonad development (for all CNVs identified in these cases see table S2). Firstly, a 708 kb duplication on the X chromosome was identified in case 13 (Figure 1a). This 46,XY DSD patient was diagnosed with complete gonadal dysgenesis and no other clinical features were reported. One of the seven genes contained within the duplicated region was \textit{DAX1} (\textit{NR0B1}), which has been previously shown to be duplicated in 46,XY complete gonadal dysgenesis.

Secondly, a deletion of 1.193 Mb approximately 300 kb upstream of the \textit{SOX9} gene was identified in case 10 (Figure 1b). This 46,XY individual had been diagnosed with complete gonadal dysgenesis. The only other clinical features reported were a cleft palate and height at the 3rd centile. This is very unusual as rearrangements and mutations affecting \textit{SOX9} normally result in patients with campomelic dysplasia (deformity of chondrogenesis) and in some 46,XY cases, gonadal dysgenesis. \textit{SOX9} is known to play a critical role in testis development, and other rearrangements upstream of this gene have been described in 46,XY DSD (or CGD) patients. The deletion described here did not cover the orthologous sequence of the recently identified TESCO element in mouse, which was proposed to act as an enhancer in regulating \textit{SOX9} expression in the gonad [14]. The orthologous human TESCO region was sequenced in the index patient, and no mutations or polymorphisms were identified (data not shown). The gonadal dysgenesis observed in this patient suggests that one or more testis-specific regulatory elements of \textit{SOX9} are contained within the 1.2 Mb deletion. In an attempt to identify such regions we performed a bioinformatic analysis of the 1.2 Mb region, searching for potential SRY/SOX9 binding sites that were conserved in human, mouse and rat. In total we identified seven such regions, hereafter referred to as enh1–7 (Figure 2a). Each of these regions was PCR amplified and cloned into a pGL4 vector containing the human \textit{SOX9} minimal promoter. Following transfection into a human Sertoli cell line, Addition of SF1, SRY or \textit{SOX9} led to a statistically higher level of enhancer activity for enh3, 4, 5, 6 and 7 for at least one of the transcription factors, with enh5 affected by all three (Figure 2b).

Thirdly, a 35 kb deletion including \textit{NEIL2} was identified in case 14 (Figure 1c and Figure 3). Although there is no known role for \textit{NEIL2} in sex determination or differentiation, the deletion was immediately downstream of \textit{GATA4}, a gene previously implicated in gonad development in the mouse. Low microarray probe density around the deletion breakpoint made it unclear whether the \textit{GATA4} 3’ UTR was included in the deletion. Several MLPA probes were designed in the intervening space, and analysis showed that the deletion did not extend into \textit{GATA4} 3’ UTR sequence (data not shown). The deletion was not present in the father or unaffected sister (DNA from the mother was not available for analysis), and sequence analysis did not reveal any variants in either \textit{NEIL2} or \textit{GATA4} in the index patient.

Rearrangements affecting candidate gonadal genes

Several previously unreported CNVs were identified that affected the coding region of genes not currently associated with gonad development. The genes affected were \textit{GPR83}, \textit{ACBD7}, \textit{CAMK1D}, \textit{OLA1H}, \textit{NEIL2} and \textit{DNAJC15} (table S3). To assess the potential of these genes to be involved in DSD, expression analysis

| Cases 1–8 (derived from cell lines) | Cases 9–23 (derived from lymphocytes) |
|-----------------------------------|--------------------------------------|
| Number of samples                | 8                                    | 15                                  |
| CNVs/genome                      | 64                                   | 66                                  |
| Min/max/median CNV size (kb)      | 0.4/1577/20                          | 0.3/1778/19                         |
| >50% overlap with known CNVs     | 92%                                  | 90%                                  |

doi:10.1371/journal.pone.0017793.t002

Figure 1. CNV analysis of three 46,XY GD cases using the AROMA algorithm. Data are plotted along each chromosome, with each point represents the copy number estimate of an individual probe. The horizontal solid black line denotes the predicted copy number of the genomic region. For each panel coverage of 2 Mb is shown, with numbers on the horizontal axis corresponding to the March 2006 human reference sequence (hg18). A) Duplication of ~708 kb on the X-chromosome in case 13. B) Deletion of 1.193 Mb on chromosome 17 in case 10. C) Deletion of 35 kb on chromosome 8 in case 14.
doi:10.1371/journal.pone.0017793.g001
of the orthologous mouse genes was performed on cDNA from purified somatic cells isolated from male and female mouse gonads at the time of sex differentiation. Two genes (Dnajc15 and Camk1d) showed sexually dimorphic expression in the somatic cells (Figure 4). Dnajc15 was expressed more highly in the female gonad than the male (p, 0.005 at E13.5, E14.5, E15.5). Conversely, Camk1d expression was significantly higher in the developing male gonad when compared to the female gonad (p, 0.05 at E13.5, p, 0.005 at E15.5).

Discussion

We have used the Affymetrix Genome-Wide Human SNP Array 6.0 to identify deletions and duplications in a cohort of 23 patients with 46,XY GD. A relatively stringent threshold of 10 consecutive probes was set before a CNV was considered genuine. However, despite this stringent criterion over 65 CNVs per sample were detected by the microarray. Discriminating causative CNVs from benign polymorphisms was complicated by the incomplete knowledge of polymorphic CNVs [30]. When comparing our findings with those reported in the database of genomic variants [31], over 90% overlapped with previously identified rearrangements for at least half the affected region. A confounding issue associated with any mutation screening study, but particularly relevant for DSD, is parent of origin effect. As gonad development in males and females involves distinct genes and pathways, a given variant might disrupt one gonad specific pathway, leading to DSD, whilst having no effect on the development of the other gonad type. It is therefore not possible to completely exclude CNVs that have been inherited (particularly from the parent with the opposite sex chromosome composition or that have previously been identified in apparently healthy individuals. Therefore, in addition to previously undescribed CNVs affecting RefSeq genes we have also chosen to examine rearrangements overlapping or in

Figure 2. Deletion (1.193 Mb) on chromosome 17, upstream of SOX9 in a patient with 46,XY GD. a) The location and extent of the 1.193 Mb deletion on chromosome 17, upstream of SOX9 identified in case 10. The numbers at the top of the figure correspond to nucleotide position based on the March 2006 human reference sequence (hg18). Also shown are structural variants within this region that are listed in the database of genomic variants. (http://projects.tcag.ca/variation/). The numbered arrows indicate the positions of the seven potential gonad specific regulatory elements (enh1–7) that were cloned into reporter constructs. The position of the orthologous sequence corresponding to the mouse TESCO sequence is indicated by an asterisk (*). b). Reporter construct analysis of SOX9 regulatory regions. Effect of selected transcription factors on luciferase activity driven by putative gonad regulatory regions (enh1–7) inserted upstream of the minimal SOX9 promoter (sox prom). Results are given as relative activation of the reporter by the expression constructs (SF1, SOX9, SRY) compared with the empty vector (pcDNA3). Data represent mean values ±SEM obtained from at least four independent experiments. Statistical analysis was performed with a 2-tailed t-test. ** p<0.005; * p<0.05.

doi:10.1371/journal.pone.0017793.g002
proximity to genes known to be involved in gonad development, irrespective of whether or not they overlap with previously reported CNVs.

**DAX1 (NROB1)**

The first rearrangement was a 708 kb duplication containing DAX1 in case 13. Other 46,XY GD patients have been described with duplications overlapping the same region of the X chromosome [19,32]. DAX1 shows sexually dimorphic gonadal expression, being downregulated in the developing testis [33]. Studies in mouse models support Dax1 as having both pro-testis [34,35] and anti-testis functions [33]. These studies indicate that a critical level of Dax1 is required for proper testis cord formation – levels above or below this level interfere with the early stages of testicular differentiation (e.g., a duplication interferes with SRY function, while a loss of function retards early cord organisation on a susceptible genetic background). As testis development proceeds, DAX1 is down-regulated, that is it antagonises the SFI/WT1/SOX9 activation of AMH. Hence, DAX1 has dynamic effects on the developing testis that changes over time.[36]

**SOX9**

The second rearrangement affecting a known testis gene was a 1.2 Mb deletion, approximately 300 kb upstream of SOX9 in case 10. Several studies have demonstrated that correct expression of SOX9 is critical for appropriate development of chondrogenic tissue [37] and testes [38]. A large duplication encompassing SOX9 was reported in a single, mosaic case of 46,XX testicular DSD [39]. This rearrangement presumably led to increased levels of SOX9 expression and testis development. Mutations within and outside of SOX9 are known to cause the bone disorder campomelic dysplasia (CD), which in approximately 70% of cases is associated with 46,XY GD. Non-overlapping deletions and translocation breakpoints up to 1 Mb from the SOX9 coding sequence have been identified in CD patients, with and without 46,XY GD, suggesting that there are multiple regulatory elements controlling SOX9 expression[40,41,42,43,44, 45,46,47,48]. Case 10 is remarkable as it is the first reported case of a rearrangement affecting the SOX9 locus in a 46,XY GD individual in the absence of campomelic or acampomelic dysplasia. The only clinical features reported other than gonadal dysgenesis were short height (3rd centile) and a cleft palate. The deletion in case 10 does not overlap with the cis-regulatory element identified in a study of Pierre-Robin sequence patients [49], a disorder that presents with several clinical features including cleft palate. The close proximity (~13 kb from) however, suggests that a position effect altering the function of this putative enhancer may underlie the cleft palate observed in case 10.

The lack of any bone pathology consistent with CD would suggest that the regulatory element(s) necessary for appropriate SOX9 expression during skeletal development are present outside the deleted region in this patient. Furthermore, it suggests that mutations affecting SOX9 regulation in the gonad may have been underestimated as a cause of gonadal dysgenesis in the absence of CD. However, we cannot exclude that the deletion has created a new regulatory element, or relocated an existing sequence that can function as a tissue-specific silencer. This would be analogous to a previously described insertion/deletion upstream of the Sox9 locus in a XX sex reversed transgenic mouse [50]. In that instance it was originally thought that the deletion removed a gonad-specific suppressive element, but subsequent studies showed that the inserted material was capable of driving Sox9 expression in the mouse gonad [51].

In case 10 the deletion upstream of SOX9 did not extend into the orthologous human sequence of a recently identified enhancer that has been proposed to regulate Sox9 expression in the developing mouse gonad [14]. In mouse a 3.2 kb sequence (TES – testis-specific enhancer) ~10 kb upstream of Sox9 was defined, with 1.4 kb of this (TESCO – TES core element) as the critical region. Functional studies showed that the TESCO sequence was capable of acting as an enhancer in an Sry- and Sfi1-dependent manner, consistent with Sox9 being a direct downstream target of Sry. Sfi1 and Sry were shown to be required for enhancer activity, as mutating their binding sites within the TESCO region eliminated enhancer activity. However, it has not yet been demonstrated that deletion of the TESCO region is sufficient by itself to cause male to female sex reversal. Indeed, a recent study of 66 46,XY gonadal dysgenesis patients did not identify any mutations in the genomic region orthologous to TESCO [52]. We sequenced this region in case 10, and did not identify any sequence variants. As no point mutations have been identified in the TESCO region of 46,XY patients, and loss-of-function rearrangements outside this region have been identified in other 46,XY GD patients, regulatory elements other than or in addition to TESCO are likely to be involved in gonad-specific SOX9 expression in humans.

Common bioinformatic approaches to identify potential regulatory elements include examining sequences that show significant interspecies conservation, or sequences containing consensus binding site motifs for transcription factors. Previous
studies have used these approaches to identify putative Sox9 enhancers in mice [53,54]. We selected seven human DNA sequences within the region deleted in case 10 that are conserved between human, mouse and rat and contain putative SRY/SOX9 binding sites. Co-transfecting reporter constructs containing each of these sequences together with SRY, SF1 or SOX9 into a human Sertoli cell line revealed that five of the seven regions showed significantly increased luciferase expression when compared to the SOX9 minimal promoter only. Enhancer 5 (enh5) was the only one to show a significant increase in luciferase expression for each of the transcription factors, and was the only one to respond to SF1. Enh6 responded only to SRY, whereas each of enh3, enh4 and enh7 responded to both SOX9 and SRY. These DNA sequences are the first putative SOX9 gonad enhancers identified in humans, and are targets for mutation screening in a larger cohort of 46,XY DSD GD patients.

**Figure 4. Expression analysis of Camk1d and Dnajc15 in developing mouse gonads.** Expression analysis was performed on cDNA from sorted GFP- (somatic) cells derived from male and female mouse gonads, embryonic days 12.5–15.5. The data is normalised such that the lowest expression is 1.0, and represent mean values ±SEM from three independent experiments. Shown here are the expression patterns for a) Camk1d and b) Dnajc15. Comparisons between male and female expression for Camk1d and Dnajc15 were performed with a 2-tailed t-test. **p<0.005; * p<0.05. doi:10.1371/journal.pone.0017793.g004**
NEIL2 and GATA4

We identified a 35 kb deletion that completely removed NEIL2 in case 14. Although there is no evidence supporting a role for NEIL2 in testis determination, the deletion is immediately 3' of the neighbouring GATA4 gene. In mice, mutations of Gata4 that interfere with Gata4-Fog2 interaction lead to gonadal abnormalities in specific strains of XY mice [53]. It has been shown that Gata4 also acts synergistically with Wt1 in activating both the Sry and Amh promoters.[56] Amh (anti-Mullerian hormone, also known as Mullerian inhibiting substance) is responsible for the regression of the Mullerian structures that would otherwise form the female internal reproductive tract. Mutations affecting the coding sequence of GATA4 in humans are associated with cardiac malformations, with no mutations as yet associated with gonadal dysgenesis. As the rearrangement we have identified does not affect GATA4 coding sequence, we would not necessarily expect cardiac malformations as a consequence of the deletion. It is possible that a regulatory element controlling GATA4 expression in the developing gonad is either removed or relocated by the deletion. This is analogous to the situation we recently described for rearrangements nearby the SOX3 coding sequence in 46,XX males which affected gene regulation [57].

While this manuscript was under revision there was a report describing a family where a GATA4 missense mutation segregated with congenital heart disease and 46,XY DSD[58]. There were three affected 46,XY males with DSD. All three patients had inguinal gonads (in one individual the gonads were removed surgically as they were dysgenic) and two had hypospadias. In addition one of the males had a dystrophic murmur without any atrial septal defect. Two 46,XX females within the family had congenital heart conditions, although there was no sign of gonadal dysfunction.

The authors proposed that the DSD observed in 46,XY carriers of the GATA4 mutation was due to a loss-of-function effect on GATA4 in the developing gonad, interfering with its ability to interact with FOG2 and disturbing activation of the AMH promoter together with NR5A1. This finding further strengthens the case for the deletion we observed downstream of GATA4 in case 14 as being causative for DSD.

Other CNVs

We identified several other previously unreported CNVs that affected the coding region of genes. To assess the potential role of any of these genes in gonadal development we examined their expression levels in FACS sorted somatic cells from E12.5 to E15.5 mouse gonads. The advantage of using a purified cell population for expression analysis is that testis determining genes are expected to be enriched specifically in the somatic component [1]. Two genes (Dnajc15 and Camkld) showed sexually dimorphic expression in somatic cells. It is interesting to note that the two genes show opposite patterns of expression, as Camkld is higher in the male gonad vs the female whereas Dnajc15 is higher in the female. Although this is apparently contradictory in patients with similar clinical features, it may be that the genes have opposite roles in testis development i.e. one has a positive effect and the other has an inhibitory function.

Dnajc15 has been suggested to be involved in pronephros development in Xenopus embryos [59], and Camkld has been associated with type 2 diabetes [60]. Further analysis needs to be conducted on these genes to define their potential role in gonad development.

Our analysis of 23 patients identified a large number of previously unidentified CNVs outside coding sequence. Unfortu-
References

1. Wilhelm D, Palmer S, Koopman P (2007) Sex determination and gonadal development in mammals. Physiol Rev 87: 1–28.
2. Lee PA, Housp CP, Ahmed SF, Hughes IA (2006) Consensus statement on management of intersex disorders. International Consensus Conference on Intersex. Pediatrics 118: e488–500.
3. Warne GL, Hewitt JK (2009) Disorders of sex development: current understanding and continuing controversy. Med J Aust 190: 612–613.
4. Cameron EJ, Sinclair AH (1997) Mutations in SRY and SOX9: testis-determining genes. Hum Mutat 9: 388–395.
5. Wilhelm D, Koopman P (2006) The making of maleness: towards an integrated view of male sexual development. Nat Rev Genet 7: 620–631.
6. Swain A, Narvaez V, Burgoyne P, Camerino G, Lovell-Badge R (1998) Dax1 antagonizes Sox9 in mammalian sex determination. Nature 391: 761–767.
7. Luo X, Reza Y, Parker KL (1994) A cell-specific nuclear receptor is essential for adren al and gonadal development and sexual differentiation. Cell 77: 481–490.
8. Jordan BK, Mohammed M, Ching ST, Delot E, Chen NX, et al. (2001) Up-regulation of WNT-1 signaling and dosage-sensitive sex reversal in humans. Am J Hum Genet 68: 1102–1109.
9. Canto P, Soderlund D, Reyes E, Mendez JP (2004) Mutations in the desert hedgehog (DHH) gene in patients with 46,XY complete pure gonadal dysgenesis. J Clin Endocrinol Metab 89: 4400–4403.
10. Pearlman A, Loke J, Le Caignec C, White S, Chin L, et al. (2010) Mutations in MAP1K1 cause 46,XY disorders of sex development and implicate a common signal transduction pathway in human testis determination. Am J Hum Genet 87: 799–804.
11. Domenice S, Correa RV, Costa EM, Nishi MY, Vilain E, et al. (2004) Mutations in SRY and SOX9 in Frasier syndrome. Mol Reprod Dev 75: 1484–1494.
12. Ledig S, Hiort O, Scherer G, Hoffmann M, Wolff G, et al. (2010) Array-CGH analysis in patients with syndromic and non-syndromic XY gonadal dysgenesis: detection of large-scale variation in the human genome. Nat Genet 36: 949–952.
13. Barbaro M, Cisognani A, Balsamo A, Baldazzi L, et al. (2008) Gene dosage imbalances in patients with 46,XY gonadal DSD detected by an in-house-designed synthetic probe set for multiple ligation-dependent probe amplification analysis. Clin Genet 75: 435–464.
14. Sekido R, Lovell-Badge R (2008) Sex determination involves synergistic action of SRY and SOX9 on a specific Sox9 enhancer. Nature 453: 930–934.
15. Pop R, Conz C, Lindenberg KS, Blesson S, Schmalenberger B, et al. (2004) Fine mapping of chromosome 17 translocation breakpoints > 900 Kb upstream of SOX9 in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. Am J Hum Genet 76: 663–671.
16. Vraka-Greven GV, Bierer-Gilber GA, Northup JK, Lockhart LH, Hawkins JC, et al. (2005) Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of SOX9 in two patients with campomelic dysplasia. Am J Hum Genet 76: 632–642.
17. Pop R, Gona C, Lindemeyer M, Steegers-Theunissen PR, Schmalenberger B, et al. (2004) Screening of the 1 Mb SOX9 5′ control region by array CGH identifies a large deletion in a case of campomelic dysplasia with XX sex reversal. J Med Genet 41: e47.
18. Pfeifer D, Kist R, Dewar K, Devon L, Kander E, et al. (1999) Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to SOX9 evidence for an extended control region. Am J Hum Genet 65: 111–124.
19. Wunderle VM, Crittcher R, Haste N, Goodfellow PN, Schelling A (1998) Deletion of long-range regulatory elements upstream of SOX9 causes campomelic dysplasia. Proc Natl Acad Sci U S A 95: 10649–10654.
20. Foster JW, Dominguez-Steglich MA, Guirard S, Kwok C, Weller PA, et al. (1994) Campomelic dysplasia and autosomal sex reversal caused by mutations in an SKY-related gene. Nature 372: 529–530.
21. Wright E, Hargrave MR, Christiansen J, Cooper L, Kun J, et al. (1995) The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. Nat Genet 9: 15–20.
22. Plis F, Chaboinier MC, de Roosig DG, Scheller A (2001) Sox9 induces sex determination in XX transgenic mice. Nat Genet 28: 216–217.
23. Huang B, Wang S, Ning Y, Lamb AN, Bartley J (1999) Autosomal XX sex reversal caused by duplication of SOX9. Am J Hum Genet 87: 349–353.
24. Hill-Crichter HJ, Kaplan LR, Pop R, et al. (2002) Fine mapping of chromosome 17 translocation breakpoints > 900 Kb upstream of SOX9 in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. Am J Hum Genet 76: 663–671.
25. Vraka-Greven GV, Bierer-Gilber GA, Northup JK, Lockhart LH, Hawkins JC, et al. (2005) Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of SOX9 in two patients with campomelic dysplasia. Am J Hum Genet 76: 632–642.
26. Pop R, Gona C, Lindemeyer M, Steegers-Theunissen PR, Schmalenberger B, et al. (2004) Screening of the 1 Mb SOX9 5′ control region by array CGH identifies a large deletion in a case of campomelic dysplasia with XX sex reversal. J Med Genet 41: e47.
27. Pfeifer D, Kist R, Dewar K, Devon L, Kander E, et al. (1999) Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to SOX9 evidence for an extended control region. Am J Hum Genet 65: 111–124.
28. Wunderle VM, Crittcher R, Haste N, Goodfellow PN, Schelling A (1998) Deletion of long-range regulatory elements upstream of SOX9 causes campomelic dysplasia. Proc Natl Acad Sci U S A 95: 10649–10654.
29. Foster JW, Dominguez-Steglich MA, Guirard S, Kwok C, Weller PA, et al. (1994) Campomelic dysplasia and autosomal sex reversal caused by mutations in an SKY-related gene. Nature 372: 529–530.
30. Wagner T, Wirth J, Meyer A, Hengel A, Hugel S, et al. (1999) Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SKY-related gene SOX9. Cell 79: 1111–1129.
31. Kwok C, Weller PA, Guirard S, Foster JW, Mansour S, et al. (1995) Mutations in SOX9, the gene responsible for Campomelic dysplasia and autosomal sex reversal. Am J Hum Genet 57: 1029–1036.
32. Lecoindre C, Pichon O, Hamel A, Helou Y, Michel-Calemand L, et al. (2009) Familial acampomelic form of campomelic dysplasia caused by a 960 kb deletion upstream of SOX9. Am J Hum Genet 145A: 1183–1189.
33. Benko S, Fantes JA, Amiel J, Klinefelter DJ, Thomas S, et al. (2009) Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. Nat Genet 41: 359–364.
34. Bishop CE, Whitworth DJ, Qin Y, Agoulnik AI, Agoulnik IU, et al. (2000) A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. Nat Genet 26: 490–494.
35. Qin Y, Kong LK, Poirier C, Truong C, Overbeek PA, et al. (2004) Long-range activation of Sox9 in Odd Sex (Ods) mice. Hum Mol Genet 13: 1213–1218.
36. Geoc D, Bagheri-Fam S, Knover KC, Weller P, Scherger B, et al. (2010) Mutations of the SRY-Responsive Enhancer of SOX9 Are Uncommon in XY Gonadal Dysgenesis. Sex Dev 4: 321–325.
37. Bagheri-Fam S, Barreirobuco F, Dohrmann U, Gunther T, Schale R, et al. (2006) Long-range upstream and downstream enhancers control distinct subsets of SOX9 expression pattern. Dev Biol 291: 302–317.
38. Bagheri-Fam S, Ferraz C, Demaille J, Scherger G, Pfeifer D (2001) Comparative genomics of the SOX9 region in human and Fugu rubripes: conservation of short regulatory sequence elements within large intergenic regions. Genomics 78: 73–82.
39. Bouma GJ, Washburn LL, Albrecht KH, Eicher EM (2007) Correct dosage of Fog2 and Gata4 transcription factors is critical for fetal testis development in mice. Proc Natl Acad Sci U S A 104: 14994–14999.
56. Miyamoto Y, Taniguchi H, Hamel F, Silversides DW, Viger RS (2008) A GATA4/WT1 cooperation regulates transcription of genes required for mammalian sex determination and differentiation. BMC Mol Biol 9: 44.
57. Sutton E, Hughes J, White S, Sekido R, Tan J, et al. (2011) Identification of SOX3 as an XX male sex reversal gene in mice and humans. J Clin Invest 121: 328–341.
58. Lourenco D, Brauner R, Rybczynska M, Nihoul-Fekete C, McElreavey K, et al. (2011) Loss-of-function mutation in GATA4 causes anomalies of human testicular development. Proc Natl Acad Sci U S A 108(4): 1597–602.
59. Kyuno J, Masse K, Jones EA (2008) A functional screen for genes involved in Xenopus pronephros development. Mech Dev 125: 571–586.
60. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, et al. (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 40: 630–645.
61. Heintzman ND, Hon GC, Hawkins RD, Kheradpisour P, Stark A, et al. (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459: 108–112.
62. Crawford GE, Holt IE, Whittle J, Webb BD, Tai D, et al. (2006) Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing MPSS. Genome Res 16: 123–131.
63. Ohnesorg T, Eggers S, Leonhard WN, Sinclair AH, White SJ (2009) Rapid high-throughput analysis of DNase hypersensitive sites using a modified MLPA approach. BMC Genomics 10: 412.
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
White, S; Ohnesorg, T; Notini, A; Roeszler, K; Hewitt, J; Daggag, H; Smith, C; Turbitt, E; Gustin, S; van den Bergen, J; Miles, D; Western, P; Arboleda, V; Schumacher, V; Gordon, L; Bell, K; Bengtsson, H; Speed, T; Hutson, J; Warne, G; Harley, V; Koopman, P; Vilain, E; Sinclair, A

Title:
Copy Number Variation in Patients with Disorders of Sex Development Due to 46,XY Gonadal Dysgenesis

Date:
2011-03-07

Citation:
White, S., Ohnesorg, T., Notini, A., Roeszler, K., Hewitt, J., Daggag, H., Smith, C., Turbitt, E., Gustin, S., van den Bergen, J., Miles, D., Western, P., Arboleda, V., Schumacher, V., Gordon, L., Bell, K., Bengtsson, H., Speed, T., Hutson, J., ... Sinclair, A. (2011). Copy Number Variation in Patients with Disorders of Sex Development Due to 46,XY Gonadal Dysgenesis. PLOS ONE, 6 (3), https://doi.org/10.1371/journal.pone.0017793.

Persistent Link:
http://hdl.handle.net/11343/263935

File Description:
Published version

License:
CC BY