RESEARCH

Next generation sequencing (NGS) to improve the diagnosis and management of patients with disorders of sex development (DSD)

L A Hughes[1], K McKay-Bounford[1], E A Webb[2], P Dasani[1], S Clokie[1], H Chandran[2], L McCarthy[2], Z Mohamed[2], J M W Kirk[2], N P Krone[2], S Allen[1] and T R P Cole[1]

[1]West Midlands Regional Genetics Service, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK
[2]Department of Endocrinology & Diabetes, Birmingham Women's and Children's Hospital, Birmingham, UK

Correspondence should be addressed to T R P Cole: trevor.cole1@nhs.net

Abstract

Disorders of sex development (DSDs) are a diverse group of conditions where the chromosomal, gonadal or anatomical sex can be atypical. The highly heterogeneous nature of this group of conditions often makes determining a genetic diagnosis challenging. Prior to next generation sequencing (NGS) technologies, genetic diagnostic tests were only available for a few of the many DSD-associated genes, which consequently had to be tested sequentially. Genetic testing is key in establishing the diagnosis, allowing for personalised management of these patients. Pinpointing the molecular cause of a patient's DSD can significantly impact patient management by informing future development needs, altering management strategies and identifying correct inheritance pattern when counselling family members. We have developed a 30-gene NGS panel, designed to be used as a frontline test for all suspected cases of DSD (both 46,XX and 46,XY cases). We have confirmed a diagnosis in 25 of the 80 patients tested to date. Confirmed diagnoses were linked to mutations in *AMH*, *AMHR2*, *AR*, *HSD17B3*, *HSD3B2*, *MAMLD1*, *NR5A1*, *SRD5A2* and *WT1* which have resulted in changes to patient management. The minimum diagnostic yield for patients with 46,XY DSD is 25/73. In 34/80 patients, only benign or likely benign variants were identified, and in 21/80 patients only variants of uncertain significance (VOUS) were identified, resulting in a diagnosis not being confirmed in these individuals. Our data support previous studies that an NGS panel approach is a clinically useful and cost-effective frontline test for patients with DSDs.

Introduction

Disorders of sex development (DSD) encompass a wide range of conditions with diverse clinical features, pathophysiology and clinical management (1, 2, 3). The recently revised stratified DSD diagnostic pathway consists of clinical examination, biochemical investigations and karyotype determination (4, 5). Once a presumptive diagnosis has been made, targeted sequencing of candidate genes may then be performed at a later stage

(5). Whilst reaching the correct diagnosis has a significant impact on management decisions, determining the aetiology of genital ambiguity in patients with DSD on the basis of clinical and biochemical assessment remains challenging (6). In DSD where no clear abnormality in the steroidogenesis pathway is present, the yield from genetic testing had historically remained low, and with single-gene sequencing, was both costly and time consuming (4).

Improvements in gene sequencing technology in conjunction with rapidly falling costs have led to the use of targeted next generation sequencing (NGS) assays. These enable multiple known disease-causative genes to be sequenced in parallel alongside initial clinical assessment and biochemical investigations, potentially avoiding the need for additional expensive biochemical and radiological investigations (7). Reaching a timely diagnosis is extremely important as it ends diagnostic uncertainty, avoids further unnecessary investigations, enables appropriate disease-specific counselling (including assessment of future fertility potential and malignancy risk) and implementation of personalised medical plans in accordance with current disease-specific consensus guidelines (8). Additionally, in the long term, accurate early diagnosis will support the development of better-designed outcome studies.

The use of targeted NGS panels for molecular diagnosis of DSD patients has already been reported successfully in several previous publications (9, 10, 11, 12). These publications show diagnostic yield and clinical utility in predominantly 46,XY DSD cohorts using panels of 64–219 genes. We present our data showing sequence analysis of a cohort of 80 DSD patients using a 30-gene panel.

Materials and methods

Patients

Eighty patients with a DSD were referred to the West Midlands Regional Genetics Laboratory (WMRGL) at the Birmingham Women's and Children's NHS Foundation Trust for diagnostic DSD testing between March 2014 and March 2017, comprising 73 patients with 46,XY DSD and 7 with 46,XX DSD. Referrals were from Clinical Genetics, Urology or Endocrinology specialists. Karyotyping and/or microarray results were typically available to confirm the patient's karyotype. Single-gene testing may have also been performed in advance of the NGS screen but a pathogenic mutation had not been identified. Where DNA was available, cascade testing was performed on parental or sibling samples to confirm segregation or to confirm a diagnosis in the proband's similarly affected siblings. Data from cascade testing are not included here; all figures therefore represent only probands referred for diagnostic testing who underwent analysis via the NGS panel. Consent was obtained for clinical testing from all patients in this study. Patients undertaking routine clinical testing in this report are not identifiable.

This report has been registered with the audit committee at the Birmingham Women's and Children' Hospital NHS Foundation Trust (CARMS-30120).

Gene selection

Thirty genes with a reported clinical association with a DSD were selected following discussion between the WMRGL and clinical specialists in Genetics and Endocrinology (Table 1). Genes include those thought to be involved in 46,XY DSD and 46,XX DSD and are tested as a single panel pipeline covering both these groups of patients.

Table 1 Genes included in the DSD panel.

	Gene name	Location
(A) 46,XY DSD		
Disorders of testicular development	*ARX*	Xp22.13
	ATRX	Xq13.3
	CBX2	17q25
	DHH	12q13.1
	DMRT1	9p24.3
	MAMLD1	Xq28
	NR0B1	Xp21.3
	NR5A1	9q33
	SOX9	17q24–q25
	SRY	Yp11.3
	TSPYL1	6q22–23
	WNT4	1p35
	WT1	11p13
Disorders of hormone synthesis or action	*AMH*	19p13.3–p13.2
	AMHR2	12q13
	AR	Xq11–q12
	CYB5A	18q23
	CYP11A1	15q23–24
	CYP17A1	10q24.3
	DHCR7	11q12–q13
	HSD3B2	1p13.1
	HSD17B3	9q22
	LHCGR	2p21
	POR	7q11.2
	SRD5A2	2p23
	StAR	8p11.2
(B) 46,XX DSD		
Disorders of ovarian development	*RSPO1*	1p34.3
	SOX9	17q24
	SRY	Yp11.3
	WNT4	1p35
Androgen excess	*CYP11B1*	8q21–q22
	CYP19A1	15q21
	HSD3B2	1p13
	NR3C1	5q31
	POR	7q11.2

A summary of all genes and their chromosomal location which are included in the panel for both 46,XY DSD (A) and 46,XX DSD (B). CAG repeat in *AR* not analysed to avoid incidental diagnosis of spinal bulbar muscular atrophy.

Endocrine
CONNECTIONS

L A Hughes *et al.*

NGS improves DSD diagnosis and management

8:2

102

The *CYP21A2* gene associated with 95% of cases of congenital adrenal hyperplasia (CAH) is not included in this panel. This is because this patient group typically has a clinical diagnosis prior to genetic testing and also the *CYP21A2* pseudogene makes accurate mapping of short reads to the functional gene very difficult.

Sample preparation

Genomic DNA was typically extracted from peripheral blood samples using Qiasymphony technology (Qiagen) following the manufacturer's instructions. The concentration of all genomic DNA samples was assessed using a Qubit (Life Technologies) prior to sequencing.

Next generation sequencing

Library preparation was initially performed by a customised TruSeq Custom Amplicon (TSCA, Illumina Inc) 30-gene panel run on the MiSeq (Illumina Inc). Exons were targeted with 25 base pairs of padding on either side, resulting in 431 amplicons of 425 bp. Enrichment was performed on 250 ng of genomic DNA and sequencing was done using 250 base paired-end reads. TruSight One (TSO) technology was then used to capture exonic regions of 4813 genes, sequencing 24 samples on a HiSeq 2500 (Illumina Inc). The same 30 DSD genes were analysed as a virtual panel using TSO. A depth of coverage of 20× was considered sufficient for either approach, and a technical report was generated indicating the proportion of each gene covered to this level. All patients that were initially run by TSCA but where a diagnosis was not confirmed were subsequently retested using TSO. NGS was performed on probands only and analysis for copy number variation in these genes was not performed.

Bioinformatic analysis

Bioinformatic analysis was performed using an in-house pipeline where sequence reads were mapped to the human genome hg19 reference. Several programmes are incorporated in the pipeline; Trimmomatic (quality trimming of reads), BWA mem (alignment to hg19), Samblaster (duplicate marking), Abra (realigning), Platypus (variant calling), Annovar (variant annotation) PLINK (IBS calculation) and Picard (calculating hybridisation and mapping metrics). Custom Python code and bedtools were used to calculate coverage and the Python module pandas to produce patient-specific Excel files.

Variant interpretation and reporting

Variants were classified following the Association of Clinical Genetic Science (ACGS, www.acgs.uk.com) best practice guidelines, based on the American College of Medical Genetics and genomics recommendations (13). This included utilisation of in-house frequency data, population frequency data (dbSNP, 1000 genomes and EXAC), *in silico* tools including Polyphen, Align GVGD and splice tools (searched through the alamut interface), the Human Gene Mutation Database (HGMD Professional, Biobase Corporation) and evidence from peer-reviewed literature. The five classes are described in Table 2. For suspected compound heterozygous mutations, parental samples were requested to confirm that the mutations were on opposite alleles (in trans). Regions of interest were all exonic regions plus 30 bp upstream and 10 bp downstream of each exon. Intronic variants outside of these regions were considered as deep intronic variants (DIVs), and no further investigation was undertaken. In some cases, Sanger sequencing was performed to complete gene coverage to a depth of 20×. For example, where a single heterozygous mutation in a likely candidate gene associated with a recessive condition had been identified.

Results

The DSD NGS panel provided a molecular diagnosis in 25 out of the 80 patients tested (Table 3). A diagnosis was deemed as confirmed where variants of class 4 or

Table 2 Classification and follow-up of variants.

Variant class	Variant description	Confirmation by Sanger	Reported
1	Clearly not pathogenic	No	No
2	Unlikely to be pathogenic	No	Yes
3	Variant of uncertain significance (VOUS)	Yes	Yes
4	Likely to be pathogenic	Yes	Yes
5	Clearly pathogenic	Yes	Yes

All variants determined to be class 3–5 were confirmed using Sanger sequencing and all were included in the clinical reports. Variants considered to be unlikely to be pathogenic (class 2) were not confirmed by Sanger sequencing and were recorded for information only in the clinical report. Class 1 variants were not reported.

Table 3 Patient details with a confirmed molecular diagnosis.

Patient	Sex	Reason for referral	Gene	Allele 1	Allele 2	Results reported and follow-up
1	XY M	?PAIS	AR	c.2402C>T p.(Thr801Ile)[LP]		Confirmed diagnosis of PAIS
2	XY M	?PMDS	AMH	c.283C>T p.(Arg95*)[P]	c.905G>A p.(Arg302Gln)[V]	Consistent with diagnosis of PMDS
3	XY F	?46,XY DSD	NR5A1	c.1171A>T p.(Lys391*)[LP]	Normal[a]	Consistent with a diagnosis of a 46,XY DSD. Novel mutation
4	XY[p] F	?46,XY DSD	HSD17B3	c.614T>A p.(Val205Glu)[P]	c.645A>T p.(Glu215Asp)[P]	Supports diagnosis of 46,XY DSD due to HSD17B3 deficiency
5	XY[p] F	?46,XY DSD	HSD17B3	c.194C>T p.(Ser65Leu)[P]	c.729_735del7 p.(Ile244Argfs*11)[P]	Supports diagnosis of 46,XY DSD due to HSD17B3 deficiency
6	XY[p] M	?46,XY DSD	SRD5A2	c.698+1G>T[LP]	c.698+1G>T[LP]	Consistent with diagnosis of 46,XY due to SRD5A2 deficiency
7	XY[p] M	X-linked hypospadias	MAMLD1	c.1366C>T p.(Arg456*)[LP]		Consistent with MAMLD1 associated hypospadias. Confirmed in two affected brothers and mother (carrier). Carrier of PMDS (c.35T>G p.(Val12Gly) in AMH)
8	XY[p] M	?46,XY DSD	AR	c.2391G>A p.(Trp797*)[LP]		Mosaic (70% of reads). Likely causally related to clinical features
9	XY F	?XY DSD	NR5A1	c.69 C>A p.(Tyr23*)[P]	Normal[a]	Consistent with diagnosis of 46,XY DSD. Novel mutation
10	XY F	?46,XY DSD	HSD17B3	c.695C>T p.(Ser232Leu)[P]	c.695C>T p.(Ser232Leu)[P]	Confirms diagnosis 46,XY DSD due to HSD17B3 deficiency
11	XY[p] M	?PMDS	AMHR2	c.813_817delGCTCT, p.(Leu272Trpfs*24)[P]	c.931G>A, p.(Gly311Ser)[V]	Consistent with features of PMDS. Novel mutation and novel variant
12	XY F	?46,XY DSD	SRD5A2	c.737G>A, p.(Arg246Gln)[P]	c.737G>A, p.(Arg246Gln)[P]	Consistent with diagnosis of SRD5A2 deficiency
13	XY[p] M	Penoscrotal hypospadias	SRD5A2	c.586G>A, p.(Gly196Ser)[P]	c.586G>A, p.(Gly196Ser)[P]	Consistent with diagnosis of SRD5A2 deficiency assuming XY
14	XY M	Gynaecomastia, Hypospadias, micropenis	AR	c.2057T>C p.(Val686Ala)[LP]		Consistent with clinical features. Confirmed inherited from mother
15	XY M	Ambiguous genitalia	HSD17B3	c.277+4A>T[P]	c.133C>T p.(Arg45Trp)[V]	Consistent with clinical features. c.13C>T p.(Arg45Trp) is novel variant
16	XY F	?AIS	AR	c.2343G>A p.(Met 781Ile)[P]		Consistent with diagnosis of AIS
17	XY M	?PMDS	AMHR2	c.289C>T p.(Arg97*)[P]	c.289C>T p.(Arg97*)[P]	Confirms diagnosis of PMDS. Both parents are carriers. Also had another child affected child who was homozygous for the same mutation (detected in neonatal period)

(*Continued*)

Table 3 Continued.

Patient	Sex	Reason for referral	Gene	Allele 1	Allele 2	Results reported and follow-up
18	XY M	?XY DSD	*HSD3B2*	c.518T>G p.(Leu173Arg)[P]	c.518T>G p.(Leu173Arg)[P]	Confirms diagnosis of CAH due to HSD3B2 deficiency. Both parents are carriers. Patient also a carrier of the c.964-1G>C splice mutation in *DHCR7*
19	XY[p] F	?46, XY DSD	*HSD17B3*	c.277+4A>T[P]	c.645A>T p.(Glu215Asp)[P]	Confirmed diagnosis of 46, XY DSD due to 17-beta hydroxysteroid dehydrogenase deficiency. Each parent carries 1 mutation
20	XY M	Undervirilised male	*SRD5A2*	c.307C>T p.(Arg103*)[P]	c.107A>G, p.(His36Arg)[V]	Consistent with clinical features. Parental samples confirmed compound heterozygous. Follow-up biochemical testing confirmed SRD5A2 deficiency. Novel variant
21	XY F	?XY DSD	*AR*	c.2407dupC p.(Gln803Profs*27)[P]		Confirms diagnosis of AIS. Two affected siblings also have mutation Novel mutation. Also heterozygous for HSD17B3 familial mutation c.803G>A p.(Cys268Tyr)
22	XY M	Ambiguous genitalia	*WT1*	c.1087A>T p.(Arg363*)[P]	Normal[a]	May be contributing to features. Confirmed de novo. Tumour screening initiated
23	XY M	Severe hypospadias	*AR*	c.2384T>A p.(Phe795Tyr)[LP]		Initially reported as VOUS. Once identified in affected (milder) brother who's karyotype was 47,XXY more confident that linked to features. Novel variant
24	XY M	Severe hypospadias and penile transposition	*AR*	c.2645T>C p.(Leu882Pro)[LP]		Mosaic (30% of reads). Likely causally related to phenotype
25	XY M	?PMDS	*AMH*	c.649C>T p.(Gln217*)[P]	c.649C>T p.(Gln217*)[P]	Consistent with diagnosis of PMDS. Novel mutation

Details of mutations and variants found in patient reported with a confirmed molecular diagnosis. Details of the karyotypic and phenotypic sex are in the second column with M and F representing phenotypic sex. P indicates presumed karyotype (reports not seen) from SRY sequence reads. Pathogenic mutations linked to the diagnoses listed in 'gene column'. Allele 1 and 2 describe the mutations in the different alleles.
[a]Normal (WT); [V]class 3 variant (variant of uncertain significance (VOUS)); [LP]class 4 variant (likely pathogenic); [P]class 5 variant (clearly pathogenic).
AIS, androgen insensitivity syndrome; CAH, congenital adrenal hyperplasia; PAIS, partial androgen insensitivity syndrome; PMDS, persistent mullerian duct syndrome.

class 5 were identified which were consistent with the inheritance pattern for that gene, and where disruption of the gene was in keeping with the patient's phenotype. On four occasions, a class 3 variant was found in combination with a class 5 mutation (patients 2, 11, 15 and 20).

Although technically class 3 variants are of uncertain significance, their presence in combination with a class 5 variant in these patients, when considered with the clinical information provided led us to believe that these findings were causally related to the clinical features.

The overall diagnostic yield for this panel is therefore currently 25/80 for all samples and 25/73 for 46,XY DSD. This figure represents the likely minimum detection rate of this panel as some samples were received with an initial request for Sanger sequencing of a specific gene, and typically only included for panel testing if negative on Sanger sequencing.

Class 3 variants (listed in Table 4) were typically missense mutations which had not been previously reported, and therefore, no clinical information was available. These were included in the clinical report with a statement that a diagnosis had not been confirmed due to the uncertainty around the pathogenicity of such variants. Where only class 1 and/or 2 variants were identified, patient reports stated that no evidence of a pathogenic mutation had been identified. A summary of the findings can be seen in Fig. 1. Both previously reported and novel pathogenic mutations and variants were identified in *AMH*, *AMHR2*, *AR*, *DHCR7*, *HSD17B3*, *HSD3B2*, *LHCGR*, *MAMLD1*, *NR5A1*, *SRD5A2* and *WT1* (Table 5). Diagnosis due to mutations in the *AR* gene (seven patients) were the most commonly observed (Table 3) followed by diagnosis due to *HSD17B3* (five patients) and *SRD5A2* (four patients).

Discussion

DSDs, estimated to be present in 1.7% of live births (14), are a diagnostic challenge due to variable expressivity and pleiotrophy, clinical overlap of the different DSD and their significant aetiological heterogeneity. Historically a genetic diagnosis was made in as few as 13% of cases (15). We present data from 80 patients who underwent routine diagnostic testing for DSD using a 30-gene NGS panel. This diagnostic DSD panel was utilised irrespective of clinical and biochemical features, unless a specific single Sanger sequence request was made based on phenotypic assessment. The diagnostic yield of this DSD panel was shown to be 25/80 for all DSDs, higher for 46,XY DSD (25/73) and would have been higher in this cohort if all cases with a suspected diagnosis (all subsequently confirmed on Sanger sequencing) had not been filtered out prior to implementation of the panel test. Pathogenic (or likely pathogenic) mutations in the *AMH*, *AMHR2*, *AR*, *HSD17B3*, *HSD3B2*, *MAMLD1*, *NR5A1*, *SRD5A2* and *WT1* genes were identified. Our detection rate and findings are similar to those in previous studies and are summarised in Table 6. Dong *et al.* demonstrated an increased detection rate of 9/13 in 46,XY DSD patients using a panel of 219 genes (10); however, the study included small patient numbers and so may not be representative.

Separating out analysis of 46,XY DSD from those with 46,XX DSD results in an improved 46,XY yield to 25/73 but highlights the 0/7 diagnostic yield of individuals with a 46,XX DSD. Difficulty in confirming a molecular diagnosis in those with an 46,XX DSD has also been seen in other studies (11, 12). Sample numbers for those with 46,XX DSD are very small in this study and therefore are unlikely to be representative of the true diagnostic capability of the panel for these patients. In addition, other causes of 46,XX DSD such as translocation of *SRY* to the X chromosome, duplications of *SOX9* or CAH due to CYP21A2 deficiency are not detectable by this method. Patients would typically have had karyotype and/or microarray prior to testing on the panel and would have been tested separately for CYP21A2 deficiency if CAH was suspected. Increased sample numbers and incorporation of more 46,XX DSD-associated genes as they are identified may allow a more accurate estimate of the panel's usefulness for those with 46,XX DSD. It will also be important to include new 46,XX DSD genes that are likely to be identified in current international exome/genome sequencing projects such as the 100,000 genome project in England (16).

Novel mutations and variants in several genes were identified where functional studies were not available. Variants were considered likely to be causative if they were observed in trans (on opposite chromosome alleles) with a known pathogenic mutation, in a disease gene showing autosomal recessive inheritance (four patients in our cohort). In some cases, segregation studies confirming the bi-allelic nature of the findings also supported a likely pathogenic role. This information has expanded our knowledge of likely diagnostic DSD variants for future investigation of DSD patients. Novel VOUS in the absence of a confirmed diagnosis were also seen. Whilst their significance currently remains uncertain, wider data sharing through publication of studies such as this is crucial to further our understanding of such variants. The large number of VOUS in this cohort is predominantly due to limited clinical information related to some DSD-related genes, for example, only VOUS were detected in the *CBX2* gene. Clinicians will have to manage any patient confusion or anxiety within the current uncertainty until more data are available. When designing future NGS DSD panels, it will remain important to recognise that increasing the number of genes, especially those with limited data, will generate greater numbers of VOUS with

Table 4 Details of variants of uncertain significance (VOUS) identified.

Patient	Sex	Reason for referral	Gene	Allele 1	Allele 2	Results reported and follow-up
(A) Patients with variants of uncertain significance (VOUS) where a diagnosis was not confirmed						
26	XY ?	N/A	POR	c.948-30G>A[V]	Normal[a]	A molecular diagnosis has not been confirmed
27	XY[p] M	1° gonadal failure, short stature	WT1	c.11C>G, p.(Pro4Arg)[V]	Normal[a]	A molecular diagnosis has not been confirmed
28	XY M	Mullerian resistant disorder	HSD17B3	c.133C>T p.(Arg45Trp)[V]	Normal[a]	A molecular diagnosis has not been confirmed
29	XY M	?46,XY DSD	RSPO1	c.658C>T p.(Arg220Trp)[V]	Normal[a]	A molecular diagnosis has not been confirmed
30	XY M	Hypogonadism	HSD3B2	c.809T>C p.(Ile270Thr)[V]	Normal[a]	A molecular diagnosis has not been confirmed
31	XY F	Facial dysmorphism, gastric motility issues, undescended testes, adrenal insufficiency	CBX2 CBX2#	c.1411C>G p.(Pro471Ala)[V] c.616C>T p.(Gln206*)[V]	Normal[a] Normal[a]	A molecular diagnosis has not been confirmed. #Variant found in alternative transcript
32	XX M	Hypospadias	NR5A1	c.275G>A p.(Arg92Gln)[V]	Normal[a]	A molecular diagnosis has not been confirmed
33	XY F	? Gonadal dysgenesis	CYP11A1	c.1250T>G p.(Val417Gly)[V]	Normal[a]	A molecular diagnosis has not been confirmed
34	XY F	? CYP11A1 imbalance	CYP11A1 MAMLDI	c.989C>T p.(Thr330Met)[V] c.2009C>T p.(Thr670Ile)[V]	Normal[a]	A molecular diagnosis has not been confirmed
35	XY F	Primary ovarian failure	WT1	c.1493A>G p.(Glu498Gly)[V]	Normal[a]	A molecular diagnosis has not been confirmed
36	XY M	Severe penoscrotal hypospadias	CBX2 HSD3B2	c.1416C>G p.(Asp472Glu)[V] c.500C>T p.(Ala167Val)[V]	Normal[a] c.500C>T p.(Ala167Val)[V]	A molecular diagnosis has not been confirmed
37	XY F	Tall stature, uterus present, no obvious ovaries	CBX2 AMH AMH	c.1411C>G p.Pro471Ala[V] c.53C>T p.(Ala18Val)[V] c.1556C>T p.(Ala519Val)[V]	Normal[a] Normal[a] Normal[a]	A molecular diagnosis has not been confirmed
38	XX F	Premature ovarian failure	CYP11B1	c.1451T>A p.(Val484Asp)[V]	Normal[a]	A molecular diagnosis has not been confirmed
39	XY F	Clitoromegaly, no vaginal opening	AR	c.1174C>T p.(Pro392Ser)[V]		Pathogenicity of variant uncertain due to conflicting evidence
40	XY M	Penoscrotal hypospadias, micropenis and undescended testes	CBX2	c.785G>A p.(Arg262Gln)[V]	Normal[a]	A molecular diagnosis has not been confirmed
41	XY M	Hypospadias and penoscrotal transposition	CYP11A1 HSD17B3	c.940G>A p.(Glu314Lys)[V] c.133C>T p.(Arg45Trp)[V]	Normal[a] c.133C>T p.(Arg45Trp)[V]	A molecular diagnosis has not been confirmed
42	XY M	?46,XY DSD	LHCGR CBX2	c.828delC p.(Ser277Alafs*32)[P] c.785G>A p.(Arg262Gln)[V]	Normal[a] Normal[a]	A molecular diagnosis has not been confirmed. Variant likely to be pathogenic but absence of second mutation leads to uncertain significance

(*Continued*)

Table 4 Continued

Patient	Sex	Reason for referral	Gene	Allele 1	Allele 2	Results reported and follow-up
43	XY M	Ambiguous genitalia	*NR5A1*	c.146G>A p.(Cys49Tyr)[V]	Normal[a]	A molecular diagnosis has not been confirmed. De novo variant
44	XY[p] F	?46,XY DSD	*NR5A1*	c.1019C>T p.(Ala340Val)[V]	Normal[a]	A molecular diagnosis has not been confirmed. Maternally inherited
45	XY M	Ambiguous genitalia	*HSD17B3*	c.202-22G>A[V]	c.202-22G>A[V]	A molecular diagnosis has not been confirmed. Recommend biochemical testing
46	XX F	Ambiguous genitalia, complete labial fusion	*LHCGR* *NR5A1*	c.458+3A>G[V] c.486C>T p.(=)[V]	Normal[a] Normal[a]	A molecular diagnosis has not been confirmed
(B) Patients with variants of uncertain significance where a diagnosis has been confirmed						
3	XY F	?46,XY DSD	*ATRX*	c.2595C>G p.(His865Gln)[V]		ATRX and AMH variants found in in addition to NR5A1 class 4 mutation (Table 3)
			AMH	c.-2C>T[V]	Normal[a]	
4	XY F	? 46,XY DSD	*MAMLD1*	c.2744A>C p.(Asp915Ala)[V]		MAMLD1 variant found in addition to HSD17B3 mutations (Table 3)
7	XY M	X-linked hypospadias	*AMH*	c.35T>G p.(Val12Gly)[P]	Normal[a]	AMH and CBX2 variants found in addition to MAMLD1 class 4 mutation (Table 3)
			AMH *CBX2*	c.-74C>G[V] c.565G>A p.(Ala189Thr)[V]	Normal[a] Normal[a]	
19	XY F	?46,XY DSD	*MAMLD1*	c.728G>A p.(Cys243Tyr)[V]	Normal[a]	MAMLD1 variant found in addition to HSD17B3 mutations (Table 3)
22	XY M	Ambiguous genitalia	*ATRX*	c.546A>G p.(=)[V]	Normal[a]	ATRX variant found in addition to WT1 mutation (Table 3)

Details of all VOUS found in this study. (A) Indicates patients where only VOUS were found that is no pathogenic or likely pathogenic variants were identified. As such, a diagnosis could not be confirmed in these patients. (B) Indicates patients where VOUS were found in addition to the pathogenic/likely pathogenic mutations which were believed to be causative of the patients phenotype. Details of the karyotypic and phenotypic sex are in the second column with M and F representing phenotypic sex. P indicates presumed karyotype (reports not seen) from SRY sequence reads. Allele 1 and 2 describe the variants in the different alleles.

[a]Normal (wildtype); [V]class 3 variant (Variant Of Uncertain Significance (VOUS)); [P]class 5 variant (clearly pathogenic).

increased cost and complexity of analysis. This should be balanced against the potential for gene discovery.

Segregation studies have been helpful in the determination of pathogenicity in several families, as illustrated by patient 23, who had an *AR* variant initially reported as a VOUS. The patient's older brother, previously shown to have a 47,XXY karyotype and a 'milder' phenotype, was subsequently shown to have the same *AR* variant. Given his additional X chromosome, without skewed X inactivation a milder phenotype would be expected. This increased our confidence in calling the variant a 'likely pathogenic' mutation and highlights the

importance of reviewing interpretation of variants when new information becomes available.

NGS technology has also allowed the identification of mosaic mutations in the *AR* gene, which may have gone undetected by other methods. Identification of an accurate number of reads enhances our knowledge of the level of mosaicism present. It should be noted, however, that the results represent the mutation load in peripheral blood (70% patient 8 and 30% patient 24) and not necessarily other relevant tissues. Of the mutations identified, both had been previously reported in non-mosaic form in the literature (17, 18). Mosaic mutations
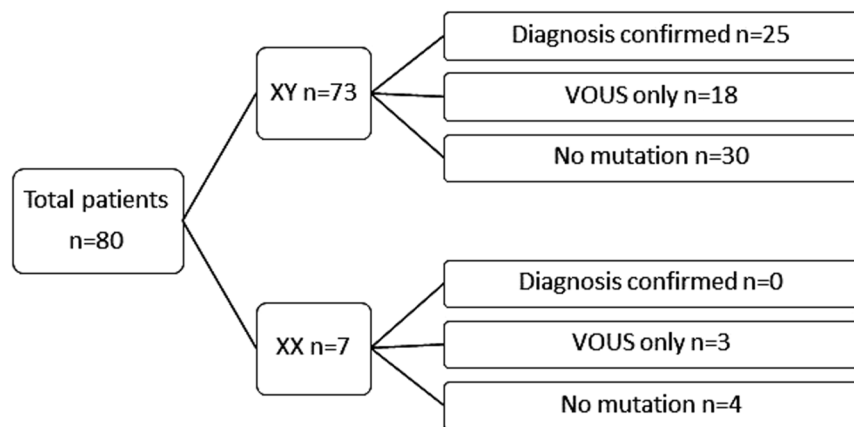
Figure 1
Summary of findings of DSD panel. Results of the panel are separated by karyotype (XX or XY) and by result. 'Diagnosis confirmed' indicates patients where a pathogenic variant was detected compatible with the patient's phenotype. 'VOUS (variant of uncertain significance) only' indicates solely class 3 variants were detected and therefore a diagnosis could not be confirmed. 'No mutation' indicates only class 1 (clearly not pathogenic) or class 2 variants (unlikely to be pathogenic) were detected.

have however been reported in the *AR* gene (19, 20) and taken together with the clinical features, allowed us to conclude that these mutations in mosaic form were likely to be contributing to the phenotype in these patients. Importantly for patient management, *AR* mutations in mosaic form are believed to pose a risk of virilisation in patients due to the presence of wildtype androgen receptor (19). Distinguishing mosaic from non-mosaic forms therefore can have significant consequences for patient management and genetic recurrence risk.

Identifying the correct genetic diagnosis modifies the patient management and impacts on the accuracy of information and choices available to family members.

Table 5 Summary of the frequency of mutations and VOUS found for each gene.

Gene	Pathogenic (class 4 or 5)	VOUS (class 3)	Gene	Pathogenic (class 4 or 5)	VOUS
AMH	3	3	HSD3B2	2	3
AMHR2	3	1	LHCGR	1	1
AR	7	1	MAMLD1	1	1
ARX	0	0	NR0B1	0	0
ATRX	0	0	NR3C1	0	0
CBX2	0	6	NR5A1	2	4
CYB5A	0	0	POR	0	1
CYP11A1	0	3	RSPO1	0	1
CYP11B1	0	1	SOX9	0	0
CYP17A1	0	0	SRD5A2	7	1
CYP19A1	0	0	SRY	0	0
DHCR7	1	0	STAR	0	0
DHH	0	0	TSPYL1	0	0
DMRT1	0	0	WNT4	0	0
HSD17B3	9	6	WT1	1	2

The total number of pathogenic mutations (class 4 and 5 variants) and variants of uncertain significance (VOUS, class 3) identified in the patients tested. Each mutated allele is given a score of 1 therefore a patient homozygous for a pathogenic mutation (score = 2) would be equally represented in the table as a patient who is compound heterozygous for two pathogenic mutations.

The former is clearly illustrated in case 21 where mutations in two different DSD-related genes were present in the same family – *HSD17B3* and *AR*. The family were requesting gonadectomy pre-adolescence in a 46,XY DSD female, on the basis that her cousins who were 46,XY DSD females, due to a homozygous *HSD17B3* mutations, had virilised and been managed with gonadectomy in another centre. Identification of a previously unidentified mutation in *AR* in this family meant three girls with 46,XY changed their subsequent management. In four cases, the diagnosis in affected siblings has been confirmed (listed in Table 3 patients 7, 17, 21 and 23) including cases enabling early prenatal (patient 20) or neonatal (patient 17) diagnosis, and thus, implementation of appropriate management from birth. This highlights the importance of identifying the molecular diagnosis not only for the proband but also for the wider family.

The panel also identified pathogenic mutations which were thought to be co-incidental and not to be related to the initial clinical presentation in the proband. In patient 18 with CAH due to HSD3B2 deficiency, the patient was also shown to be a carrier for the common splice mutation c.964-1G>C in the *DHCR7* gene which is linked to Smith Lemli Opitz syndrome (SLOS). Whilst such incidental findings can be challenging for patient counselling, the information provided may also be of great significance. This will be especially true when previously undetected autosomal recessive mutations are uncovered in highly consanguineous families, which are common within many DSD cohorts.

Where a clinician has a strong suspicion of the involvement of a specific gene, Sanger sequencing may be more cost effective, especially where the number of amplicons is relatively small. For example, pathogenic mutations in the *AMH*, *AMHR2*, *AR* and *SRD5A2* gene were identified in this way. When the original clinical diagnosis

Table 6 Summary of previous studies using NGS analysis for DSDs.

Paper	Number of genes in panel	Number of patients tested and karyotypic sex	Diagnostic yield overall	Diagnostic yield 46,XY	Diagnostic yield 46,XX
Baxter 2015 (9)	64	40 XY	–	14/40	–
Dong 2016 (10)	219	13 XY, 8 XX	8/21*	6/13	2/8
Eggers 2016 (11)	64	278 XY, 48 XX	126/326	118/278	8/48
Kim 2017 (12)	67	37 XY, 7 XX	13/44	13/37	0/7
This study	30	73 XY, 7 XX	25/80	25/73	0/7

*Dong *et al*. report their detection rate of 46,XY DSD as 9/13 however 3 of these patients had variants of uncertain significance (VOUS) only and therefore using the same parameters as in this study a diagnosis would not be confirmed. Numbers modified above to allow more accurate comparison with this study.

is incorrect, however, the potential cost savings rapidly disappear if sequential Sanger sequencing is required, and therefore, any benefit is highly dependent on the clinical expertise and the specificity of the additional non-genetic investigations. There may also be atypical presentations not yet recognised for mutations in some genes, and therefore, a wider panel approach has the potential to address this.

Despite the diagnostic rate of 25/73 for 46,XY DSD, the failure to achieve a diagnosis in 55 patients confirms the need for further development. Of note, the panel cannot currently detect copy number changes (CNVs), and so further development should include the detection of CNVs that are below the resolution of the current chromosomal microarray assay. The 30 genes in our panel were selected due to published evidence of their involvement in DSD, but new gene discoveries in studies such as the 100,000 genome project (16) should enhance the diagnostic utility, especially for 46,XX DSD. Eggers *et al*. demonstrated a 118/278 diagnostic rate for 46,XY DSD patients by including 64 DSD genes (11) and Dong *et al*. reported a 6/13 rate by including 219 genes (10); thus, diagnostic capability may be improved by increasing the gene number. These benefits may be marginal however as only one likely pathogenic mutation in the Dong series was in the additional 179 genes not included in this 30-gene panel. Comparison of these studies could indicate that a greater impact on the detection rate may be due to patient selection. It will be important, therefore, that further candidate gene inclusion is critically evaluated as the addition of genes without clear clinical utility will likely result in increasing cost and numbers of VOUS without necessarily increasing diagnostic capability. The future of this investigative pathway may well be transformed by implementation of whole exome or whole genome sequencing, but any benefits of diagnostic detection will have to be weighed against increased cost and clinical complexities resulting from VOUS and co-incidental findings. It therefore remains important to optimise such

NGS panels for DSD so that a valid comparison can be made in future.

These data demonstrate clear advantages of an NGS panel approach for highly heterogenous conditions such as DSD. Despite the limitations of the panel including incomplete coverage and inability to detect copy number changes, the results presented here demonstrate that an NGS-based panel approach is a useful frontline tool for diagnosing DSDs. In addition to a diagnostic yield of at least 25/80 we have shown examples of cases where the information provided from the panel has identified diagnoses in complex families with the potential for multiple aetiologies, cases where panel findings have significantly impacted management and treatment decisions and examples of novel variants being identified, thus expanding our current knowledge. As more and more patients are tested, the information provided by such panels will continue to grow and improve our understanding of these complex conditions and hopefully improve the diagnostic capability of such tests. Despite its limitations, the clinical benefit of this approach is clearly demonstrated for DSD patients allowing for timely accurate diagnoses, more informed management strategies and improved counselling for patients and their families.

References

1 Biason-Lauber A. Control of sex development. *Best Practice and Research: Clinical Endocrinology and Metabolism* 2010 **24** 163–186. (https://doi.org/10.1016/j.beem.2009.12.002)

2 Higashi Y, Yoshioka H, Yamane M, Gotoh O & Fujii-Kuriyama Y. Complete nucleotide sequence of two steroid 21-hydroxylase genes

tandemly arranged in human chromosome: a pseudogene and a genuine gene. *PNAS* 1986 **83** 2841–2845. (https://doi.org/10.1073/pnas.83.9.2841)

3 Arboleda VA, Sandberg DE & Vilain E. DSDs: genetics, underlying pathologies and psychosexual differentiation. *Nature Reviews Endocrinology* 2014 **10** 603–615. (https://doi.org/10.1038/nrendo.2014.130)

4 Ahmed SF, Achermann JC, Arlt W, Balen AH, Conway G, Edwards ZL, Elford S, Hughes IA, Izatt L, Krone N, *et al.* UK guidance on the initial evaluation of an infant or an adolescent with a suspected disorder of sex development. *Clinical Endocrinology* 2011 **75** 12–26. (https://doi.org/10.1111/j.1365-2265.2011.04076.x)

5 Lee PA, Nordenström A, Houk CP, Ahmed SF, Auchus R, Baratz A, Baratz Dalke K, Liao LM, Lin-Su K, Looijenga LHJ, *et al.* Global disorders of sex development update since 2006: perceptions, approach and care. *Hormone Research in Paediatrics* 2016 **85** 158–180. (https://doi.org/10.1159/000442975)

6 Lee YS, Kirk JM, Stanhope RG, Johnston DI, Harland S, Auchus RJ, Andersson S & Hughes IA. Phenotypic variability in 17beta-hydroxysteroid dehydrogenase-3 deficiency and diagnostic pitfalls. *Clinical Endocrinology* 2007 **67** 20–28. (https://doi.org/10.1111/j.1365-2265.2007.02829.x)

7 Ahmed SF, Bashamboo A, Lucas-Herald A & McElreavey K. Understanding the genetic aetiology in patients with XY DSD. *British Medical Bulletin* 2013 **106** 67–89. (https://doi.org/10.1093/bmb/ldt008)

8 Lee PA, Houk CP, Ahmed SF, Hughes IA & International Consensus Conference on Intersex organized by the Lawson Wilkins Pediatric Endocrine Society and the European Society for Paediatric Endocrinology. Consensus statement on management of intersex disorders. International Consensus Conference on Intersex. *Pediatrics* 2006 **118** e488–e500. (https://doi.org/10.1542/peds.2006-0738)

9 Baxter RM, Arboleda VA, Lee H, Barseghyan H, Adam MP, Fechner PY, Bargman R, Keegan C, Travers S, Schelley S, *et al.* Exome sequencing for the diagnosis of 46,XY disorders of sex development. *Journal of Clinical Endocrinology and Metabolism* 2015 **100** E333–E344. (https://doi.org/10.1210/jc.2014-2605)

10 Dong Y, Yi Y, Yao H, Yang Z, Hu H, Liu J, Gao C, Zhang M, Zhou L, Asan, Yi X & Liang Z. Targeted next-generation sequencing identification of mutations in patients with disorders of sex development. *BMC Medical Genetics* 2016 **17** 23. (https://doi.org/10.1186/s12881-016-0286-2)

11 Eggers S, Sadedin S, van den Bergen JA, Robevska G, Ohnesorg T, Hewitt J, Lambeth L, Bouty A, Knarston IM, Tan TY, *et al.* Disorders of sex development: insights from targeted gene sequencing of a large international patient cohort. *Genome Biology* 2016 **17** 243. (https://doi.org/10.1186/s13059-016-1105-y)

12 Kim JH, Kang E, Heo SH, Kim GH, Jang JH, Cho EH, Lee BH, Yoo HW & Choi JH. Diagnostic yield of targeted gene panel sequencing to identify the genetic etiology of disorders of sex development. *Molecular and Cellular Endocrinology* 2017 **444** 19–25. (https://doi.org/10.1016/j.mce.2017.01.037)

13 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine* 2015 **17** 405–424. (https://doi.org/10.1038/gim.2015.30)

14 Blackless M, Charuvastra A, Derryck A, Fausto-Sterling A, Lauzanne K & Lee E. How sexually dimorphic are we? Review and synthesis. *American Journal of Human Biology* 2000 **12** 151–166. (https://doi.org/10.1002/(SICI)1520-6300(200003/04)12:2<151::AID-AJHB1>3.0.CO;2-F)

15 Arboleda VA, Lee H, Sánchez FJ, Délot EC, Sandberg DE, Grody WW, Nelson SF & Vilain E. Targeted massively parallel sequencing provides comprehensive genetic diagnosis for patients with disorders of sex development. *Clinical Genetics* 2013 **83** 35–43. (https://doi.org/10.1111/j.1399-0004.2012.01879.x)

16 Siva N. UK gears up to decode 100,000 genomes from NHS patients. *Lancet* 2015 **385** 103–104. (https://doi.org/10.1016/S0140-6736(14)62453-3)

17 Galani A, Sofocleous C, Karahaliou F, Papathanasiou A, Kitsiou-Tzeli S & Kalpini-Mavrou A. Sex-reversed phenotype in association with two novel mutations c.2494delA and c.T3004C in the ligand-binding domain of the androgen receptor gene. *Fertility and Sterility* 2008 **90** 2008.e1–2008.e4. (https://doi.org/10.1016/j.fertnstert.2008.01.030)

18 Marcelli M, Tilley WD, Wilson CM, Wilson JD, Griffin JE & McPhaul MJ. A single nucleotide substitution introduces a premature termination codon into the androgen receptor gene of a patient with receptor-negative androgen resistance. *Journal of Clinical Investigation* 1990 **85** 1522–1528. (https://doi.org/10.1172/JCI114599)

19 Holterhus PM, Wiebel J, Sinnecker GH, Brüggenwirth HT, Sippell WG, Brinkmann AO, Kruse K & Hiort O. Clinical and molecular spectrum of somatic mosaicism in androgen insensitivity syndrome. *Pediatric Research* 1999 **46** 684–690. (https://doi.org/10.1203/00006450-199912000-00009)

20 Köhler B, Lumbroso S, Leger J, Audran F, Grau ES, Kurtz F, Pinto G, Salerno M, Semitcheva T, Czernichow P, *et al.* Androgen insensitivity syndrome: somatic mosaicism of the androgen receptor in seven families and consequences for sex assignment and genetic counseling. *Journal of Clinical Endocrinology and Metabolism* 2005 **90** 106–111. (https://doi.org/10.1210/jc.2004-0462)