Received: 26 July 2017 | Revised: 5 October 2017 | Accepted: 9 October 2017

DOI: 10.1002/humu.23354

RESEARCH ARTICLE

WILEY HGVS HUMAN GENOME VARIATION SOCIETY

Functional characterization of novel *NR5A1* variants reveals multiple complex roles in disorders of sex development

Gorjana Robevska[1] | Jocelyn A. van den Bergen[1] | Thomas Ohnesorg[1] | Stefanie Eggers[1] | Chloe Hanna[1,2] | Remko Hersmus[3] | Elizabeth M. Thompson[4,5] | Anne Baxendale[4] | Charles F. Verge[6,7] | Antony R. Lafferty[8,9] | Nanis S. Marzuki[10] | Ardy Santosa[11] | Nurin A. Listyasari[12] | Stefan Riedl[13,14] | Garry Warne[1,2,15] | Leendert Looijenga[3] | Sultana Faradz[12] | Katie L. Ayers[1,15]* | Andrew H. Sinclair[1,15]*

[1]Murdoch Children's Research Institute, Melbourne, Australia

[2]Royal Children's Hospital, Melbourne, Australia

[3]Department of Pathology, Josephine Nefkens Institute, Erasmus University Medical Centre, Rotterdam, The Netherlands

[4]SA Clinical Genetics Service, SA Pathology at the Women's and Children's Hospital, Adelaide, Australia

[5]School of Medicine, University of Adelaide, Adelaide, Australia

[6]Sydney Children's Hospital, Sydney, Australia

[7]School of Women's and Children's Health, UNSW, Sydney, Australia

[8]Centenary Hospital for Women and Children, Canberra, Australia

[9]ANU Medical School, Canberra, Australia

[10]Eijkman Institute for Molecular Biology, Jakarta, Indonesia

[11]Division of Urology, Department of Surgery, Dr. Kariadi Hospital, Semarang, Indonesia

[12]Division of Human Genetics, Centre for Biomedical Research Faculty of Medicine Diponegoro University (FMDU), Semarang, Indonesia

[13]St Anna Children's Hospital, Department of Paediatrics, Medical University of Vienna, Wien, Austria

[14]Division of Paediatric Pulmology, Allergology, and Endocrinology, Department of Paediatrics, Medical University of Vienna, Wien, Austria

[15]Department of Paediatrics, University of Melbourne, Melbourne, Australia

Correspondence
Katie L. Ayers, Murdoch Children's Research Institute, Parkville, Australia.
Email: katie.ayers@mcri.edu.au

*Katie L. Ayers and Andrew H. Sinclair contributed equally to this work.

Funding information
Contract Grant Sponsors: The National Health and Medical Research Council, Australia (546517); the Ian Potter Centre for Genomics and Personalised Medicine; Victorian Government's Operational Infrastructure Support Program, and the National Health and Medical Research Council Australia Independent Medical Research Institutes Infrastructure Support Scheme.

Communicated by Sergio Ottolenghi

Abstract

Variants in the *NR5A1* gene encoding SF1 have been described in a diverse spectrum of disorders of sex development (DSD). Recently, we reported the use of a targeted gene panel for DSD where we identified 15 individuals with a variant in *NR5A1*, nine of which are novel. Here, we examine the functional effect of these changes in relation to the patient phenotype. All novel variants tested had reduced trans-activational activity, while several had altered protein level, localization, or conformation. In addition, we found evidence of new roles for SF1 protein domains including a region within the ligand binding domain that appears to contribute to SF1 regulation of Müllerian development. There was little correlation between the severity of the phenotype and the nature of the *NR5A1* variant. We report two familial cases of *NR5A1* deficiency with evidence of variable expressivity; we also report on individuals with oligogenic inheritance. Finally, we found that the nature of the *NR5A1* variant does not inform patient outcomes (including pubertal androgenization and malignancy risk). This study adds nine novel pathogenic *NR5A1* variants to the pool of diagnostic variants. It highlights a greater need for understanding the complexity of SF1 function and the additional factors that contribute.

KEYWORDS
disorders of sex development, genotype–phenotype correlation, mutation, NR5A1, oligogenic, variable expressivity

1 | INTRODUCTION

Steroidogenic Factor 1 (SF1), encoded by the gene *NR5A1* (Ad4BP; MIM# 184757), is a member of the orphan nuclear receptor family of transcription factors and plays a central role in sex development, steroidogenesis, and reproduction in both males and females (reviewed in Ferraz-de-Souza, Lin, & Achermann, 2011; Lin & Achermann, 2008). SF1 protein expression is observed in both the developing ovaries and testes from the time of the bipotential gonad and all through sex determination (Hanley et al., 1999). In the developing testis, SF1 is expressed by fetal Sertoli cells (Hanley et al., 1999) where it upregulates the expression of two genes crucial for male sex determination and differentiation, *SOX9* and *Anti Müllerian Hormone (AMH)* (De Santa Barbara et al., 1998a), through synergistic interactions with SRY (Sekido & Lovell-Badge, 2008) and GATA4 (Shen, Moore, Ikeda, Parker, & Ingraham, 1994), respectively. SF1 is also expressed within the Leydig cells, where it controls various factors involved in steroidogenesis ultimately leading to male virilization in utero and in the pre-pubertal gonad (Hoivik, Lewis, Aumo, & Bakke, 2010).

Pathogenic variants in the *NR5A1* gene are thought to account for 10%–15% of non-syndromic 46,XY disorders of sex development (DSD) (El-Khairi & Achermann, 2012). Given SF1 plays multiple roles in both testicular development and steroidogenesis it is not surprising that variants in *NR5A1* have been associated with a diverse and ever growing spectrum of 46,XY DSD phenotypes. These include gonadal and testicular dysgenesis with or without Müllerian remnants (Lin et al., 2007), ambiguous genitalia, mild and severe forms of hypospadias (Allali et al., 2011; Köhler et al., 2009), varying degrees of undervirilization; such as micropenis (Wada, Okada, Hasegawa, & Ogata, 2005), cryptorchidism (Ferlin et al., 2015; Wada, Okada, Fukami, Sasagawa, & Ogata, 2006), as well as anorchia (Philibert et al., 2007) and male infertility (Röpke et al., 2013). These clinical phenotypes often manifest in conjunction with normal adrenal function (Pedace et al., 2014).

Most *NR5A1* variants described are heterozygous and either de novo or maternally inherited in a sex-limited manner, where the mother may develop premature ovarian insufficiency or remain asymptomatic (Eggers et al., 2015; Fabbri et al., 2014; Köhler et al., 2008). To complicate matters, paternal inheritance of *NR5A1* variants has also been reported, where the father may be unaffected (Swartz et al. 2017; Yagi et al., 2015) or may present with a milder phenotype and spontaneous fertility (Baetens et al., 2014; Ciaccio et al., 2012). The varying degrees of expressivity and incomplete penetrance of *NR5A1* variants, even within a single pedigree, mean that genotype-phenotype correlations are difficult to establish.

Upwards of 130 genetic variations in *NR5A1* have been described [HGMD], most of which cause a loss or reduction in function. Missense variants are common and often dispersed throughout the coding sequence, with no evident clustering. However, a substantial number of frame-shift, nonsense, and in-frame deletions have been described (Barbaro, Cools, Looijenga, Drop, & Wedell, 2011). Recently, several splice site variants have also been reported in the literature (Fabbri, Ribeiro de Andrade, Maciel-Guerra, Guerra-Júnior, & de Mello, 2016; Hussain et al. 2016; Swartz et al. 2017).

We recently reported the use of a massively parallel sequencing (MPS) targeted gene panel on a large cohort of DSD patients (Eggers et al., 2016). During the course of this study, we identified 15 individuals with variants in the *NR5A1* gene, including nine novel previously unreported variants. Here, we have assessed the impact of these novel variants in *NR5A1* on the function of SF1.

2 | MATERIALS AND METHODS

2.1 | Study subjects and clinical evaluation

Patients were recruited by collaborating clinicians and EDTA bloods for DNA extraction were collected after obtaining written informed consent from the patient (as detailed in Eggers et al., 2016). Approval for this study was obtained from the Human Ethics Committee of the Faculty of Medicine at the Royal Children's Hospital, Melbourne, Victoria, Australia (HREC22073).

2.2 | Targeted gene capture, MPS, molecular, and biochemical data analysis

Genomic analysis was carried out as previously described (Eggers et al., 2016). DNA mutation numbering is based on GenBank reference DNA sequence NM_04959.4, with the A of the ATG initiation codon designated +1. Annotations were initially created by our custom pipeline (Eggers et al., 2016; Sadedin et al., 2015) and were also processed through Mutalyser name checker (https://mutalyzer.nl) (Wildeman, van Ophuizen, Dunnen den, & Taschner, 2008). The splice site mutation annotations are provided in relation to NC_000009.12 as well as cDNA reference sequences. Predicted protein annotations are based on NP_004950. The sequencing data for each patient variant are available from the Sequencing Read Archive using reference numbers SRP092281 and project PRJNA350857. All biochemical profiling, where available, was performed by our collaborators/pathology labs.

2.3 | Variant confirmation using Sanger sequencing

NR5A1 variants were visually inspected using Integrative Genome Viewer and in cases where coverage was low or of reduced quality Sanger sequencing was used for variant confirmation. Primers for each affected exon were manually designed (Supp. Table S1). PCR was carried out using the Phusion High Fidelity DNA polymerase (NEB, Ipswich, Massachusets, USA) or AmpliTaq Gold DNA polymerase

(ThermoFisher Scientific, Waltham, Massachusetts, USA). PCR clean-up was performed with ExoSAP (ThermoFisher) according to the manufacturer's instructions and BigDye v3.1 Terminators sequencing was performed at the Australian Genomics Research Facility (AGRF).

2.4 | Mutant NR5A1 expression vectors

The mutant *NR5A1* expression vectors (p.Gly35Asp, p.Arg84His, p.His310Asp, p.Asp364Tyr, p.47_54del, p.Arg89Glyfs17*, p.Leu209Cysfs87*, p.[Pro210Gln;Tyr211*]) were created by site-directed mutagenesis (QuickChange II XL Site-directed Mutagenesis Kit; Agilent Technologies Inc., Santa Clara, CA) according to the manufacturer's instructions (for primer sequences see Supp. Table S2) using the mammalian expression vector pCMV6-Entry-*hNR5A1* (RC207577; OriGene Technologies Inc., Rockville, MD) containing the human cDNA ORF of *NR5A1*. Sanger sequencing using vector primers was used to confirm the presence of the mutation in the vectors.

2.5 | TESCO luciferase assay

Cells were co-transfected with reporter construct, pGL4-mTesco (75 ng/well), *Renilla* (pRL-TK, 20 ng/well) as a marker for transfection efficiency and SF1 expression vectors (WT, p.Gly35Asp, p.Arg84His, p.His310Asp, p.Asp364Tyr, p.47_54del, p.Arg89Glyfs17*, p.Leu209Cysfs87*, p.[Pro210Gln;Tyr211*]) either (i) alone or in conjunction with (ii) SOX9 (50 ng/well) or (iii) SRY (50 ng/well). Negative controls were included with mTesco and either SOX9 or SRY (without SF1). Empty pcDNA was used to adjust total DNA amounts to 200 ng for each reaction. Cells were harvested 24 hr post-transfection and luciferase activity was measured using (Dual-Luciferase Reporter 1000 Assay System Kit; Promega; Fitchburg, Wisconsin, USA) on an Infinite M200 Pro plate reader (Tecan; Männedorf, Zürich, Switzerland). Data represent the mean with standard error of four independent experiments performed in duplicate transfections. The relative fold change (compared with negative control—without SF1) was calculated for each technical replicate. These fold changes were then averaged across biological replicates, and standard error calculated.

2.6 | Splice site variants and in silico protein structure analysis

Human Splicing Finder Version 3.0 – UMD (https://www.umd.be/HSF3/index.html, Desmet et al., 2009) and SpliceAid2 (https://193.206.120.249/splicing_tissue.html, Piva, Giulietti, Burini, & Principato, 2012) were used to analyse the consequences of the two splice site acceptor mutations (Parameters used can be found in Supp. Table S3.). In silico protein structure was predicted for SF1 variants using the online protein modeling I-TASSER server (https://zhanglab.ccmb.med.umich.edu/I-TASSER/, Roy, Kucukural, & Zhang, 2010; Zhang, 2008). The resultant crystal structures were visualized and compared using the PyMOL Molecular Graphics System v1.7.6.6 Enhanced for Mac OS X (https://www.pymol.org). We also used HOPE to analyze the functional and structural effects of the missense mutations identified (https://www.cmbi.ru.nl/hope/, Venselaar, Beek Te, Kuipers, Hekkelman, & Vriend, 2010).

2.7 | Protein immunofluorescence and confocal analysis

COS-7 cells were seeded on eight-well chamber slides (Lab-Tek; Brendale, QLD, AUS), and transfected with SF1 expression vectors (WT, p.Gly35Asp, p.Arg84His, p.His310Asp, p.Asp364Tyr, p.47_54del, p.Arg89Glyfs17*, p.Leu209Cysfs87*, p.[Pro210Gln;Tyr211*]) using Lipofectamine 2000 (Invitrogen). Twenty-four hours after transfection, cells were washed once with PBS, then fixed, permeabilized with 1% Triton and blocked using 2%BSA in PBS. Cells were incubated overnight with a polyclonal SF1 (E18) antibody (1:300; Santa Cruz SC10976) in 1% BSA. Cytoskeleton was stained with actin (1:200; 5060 Sigma; St Louis, Missouri, USA). Cells were washed three times with PBS then incubated with secondary antibodies Alexa 488 (1:1,000; green, Invitrogen; Waltham, Massachusetts, USA) and Alexa 594 (1:1500; Invitrogen, red) in 1% BSA/PBS. Nuclear counterstaining was performed with DAPI (blue). Cells were imaged on a Zeiss LSM 780 confocal microscope. Images were taken at 10× and 40× magnification. Two images at 10× were taken per well and CellProfiler (Jones et al., 2008) was used to quantify number of expressing SF1. To quantify protein expression, we assessed the number of cells co-expressing SF1 and DAPI above a baseline threshold, compared with the total number of cells (assessed by DAPI), in two different quadrants within each transfection.

3 | RESULTS

3.1 | MPS identifies known and novel *NR5A1* variants in patients with 46,XY DSD

We have developed a MPS-targeted DSD gene panel that allowed us to screen 64 diagnostic DSD genes in affected patients (see Eggers et al., 2016). From a total of 279 patients with a 46,XY DSD sequenced using this panel, 15 individuals were found to harbor a rare change within *NR5A1*. All *NR5A1* variants identified within our cohort were heterozygous and only reported if they were rare (present in less than 0.01 of the population in ExAC). Detailed clinical information was collected for these 15 patients where available (Table 1, biochemical profiling values can be found in Supp. Table S4). The majority of the patients presented with some genital ambiguity at birth such as clitorohypertrophy, micropenis, or hypospadias, while two had completely female external genitalia (Patients 6 and 7; Table 1). At least seven of the individuals were assigned as females at birth; Müllerian structures were reported in three individuals; one patient has a rudimentary uterus and streak gonads, while a pair of siblings with normal AMH levels was found to have retained Müllerian remnants (Table 1). Nine of the individuals presented with hypospadias, some of which are also noted to have micropenis. Where gonadal histology was available, variable degrees of gonadal dysgenesis were noted. Gonadoblastomas were identified in two individuals (Patients 3 and 6; Table 1). Basal testosterone (T) levels varied between low to normal with some patients exhibiting high T/dihydrotestosterone (DHT) ratios (Table 1). Human chorionic gonadotropin (hCG) test results were available for nine patients who showed mostly a minimal response. AMH levels were available for

TABLE 1 Clinical, anatomical, and biochemical characteristics in 15 46,XY individuals in whom NR5A1 mutations were identified

Patient ID	Karyotype	Sex of rearing	Country of recruitment	External genitalia	Gonadal location (R/L)	Müllerian structures	Gonadal histology (R/L)	Basal gonadotrophins	Basal testosterone	T resp. hCG	Adrenal function	AMH	Additional information
1	46,XY	Female	NLD	Atypical, clitorohypertrophy, single perineal opening	Inguinal bilat.	Not present	LCH (bilaterally)	ND	Decreased for males Decreased to normal for females	No response	ND	ND	
2	46,XY	Female	IDN	Clitoromegaly, vaginal pouch	Nonpalpable/inguinal	ND	ND	LH&FSH-markedly elevated	T-markedly elevated DHT- low but within normal range	ND	Normal	ND	
3	46,XY	Female	IDN	Clitoromegaly, one perineal opening	Inguinal bilat.	Not present	GD	FSH&LH-elevated	T&DHT-markedly decreased	No response	Normal	Decreased	Seminoma & carcinoma in situ
4	46,XY	Male	IDN	Proximal hypospadias & chordee	Scrotal bilat.	Not present	Testes	FSH&LH: not elevated	T-markedly elevated DHT-normal	Minimal response	Normal	Normal	AR gene mutation was ruled out
5.1	46,XY	Male	AUS	Proximal hypospadias & micropenis	Palpable bilat.	Müllerian remnants	Suspected to be PGD	**1yo**: FSH&LH-not elevated **14yo**: FSH&LH-normal	**1yo**: T-normal for age **14yo**: T-normal	Minimal response	Normal	Normal	
5.2	46,XY	Male	AUS	Proximal hypospadias & micropenis	Palpable bilat.	Müllerian remnants	Suspected to be PGD	**1yo**: FSH&LH-not elevated **10yo**: FSH-normal; LH-decreased	**1yo**: T-low but within normal range & DHT: normal for age	ND	Normal	Normal	
6	46,XY	Female	AUS	Female	Intra-abdominal bilat.	Rudimentary uterus	Dysgenetic testis/streak gonad	ND	ND	ND	ND	ND	R gonad: gonadoblastoma at one pole
7	46,XY	Female	NLD	Female	ND	ND	CGD	ND	ND	ND	ND	ND	
8.1	46,XY	Male	AUS	Proximal hypospadias	Descended bilat.	Not present	ND	LH-elevated; FSH-markedly elevated	Decreased	Minimal response	Normal	ND	

(Continues)

TABLE 1 (Continued)

Patient ID	Karyotype	Sex of rearing	Country of recruitment	External genitalia	Gonadal location (R/L)	Müllerian structures	Gonadal histology (R/L)	Basal gonadotrophins	Basal testosterone	T resp. hCG	Adrenal function	AMH	Additional information
8.2	46,XY	Male	AUS	Male with hypospadias	ND	ND	ND	ND	ND	ND	ND	ND	
9	46,XY	Female	IDN	Proximal hypospadias & micropenis	-/Scrotal high position	ND	Suspected to be GD	LH&FSH not applicable	T-elevated DHT - low but within the normal range	No response	Normal	Markedly decreased	
10	46,XY	Male	AUT	Penoscrotal hypospadias, small penisoid (8 mm)	Inguinal bilat.	Not present	PGD	**3mo**: LH&FSH-normal **Puberty**: normal LH, FSH	**3mo**: T-decreased **Puberty**: T-normal	**3mo**: minimal response **6yo**: minimal response	Normal	Normal	Normal pubertal development
11	46,XY	Male	IDN	Proximal hypospadias	Scrotal bilat.	Not present	ND	LH-normal; FSH-elevated	T-elevated; DHT-normal	Elevated	Normal	Decreased	
12	46,XY	Female	BEL	Clitoromegaly, synechia vulvae, shallow vagina	Abdominal bilat.	Not present	Bilat. testes with spermatogonia no OCT+ cells	ND	ND	Minimal response	ND	ND	
13	46,XY	Male	AUS	Proximal hypospadias & chordee	Scrotal bilat.	Not present	Not biopsied US & clinically testes	**1y9mo**: LH & FSH-not elevated	**1y9mo**: T-normal	ND	**1y9mo**: T - Normal	**7do**: Normal	

AMH, anti Müllerian hormone; CGD, complete gonadal dysgenesis; DHT, dihydrotestosterone; FSH, follicle stimulating hormone; GCNIS, Germ Cell Neoplasia In Situ; LCH, Leydig cell hypoplasia; LH, luteinizing hormone; mo, months old; ND, not done; PGD, partial gonadal dysgenesis; T, testosterone; yo, years old

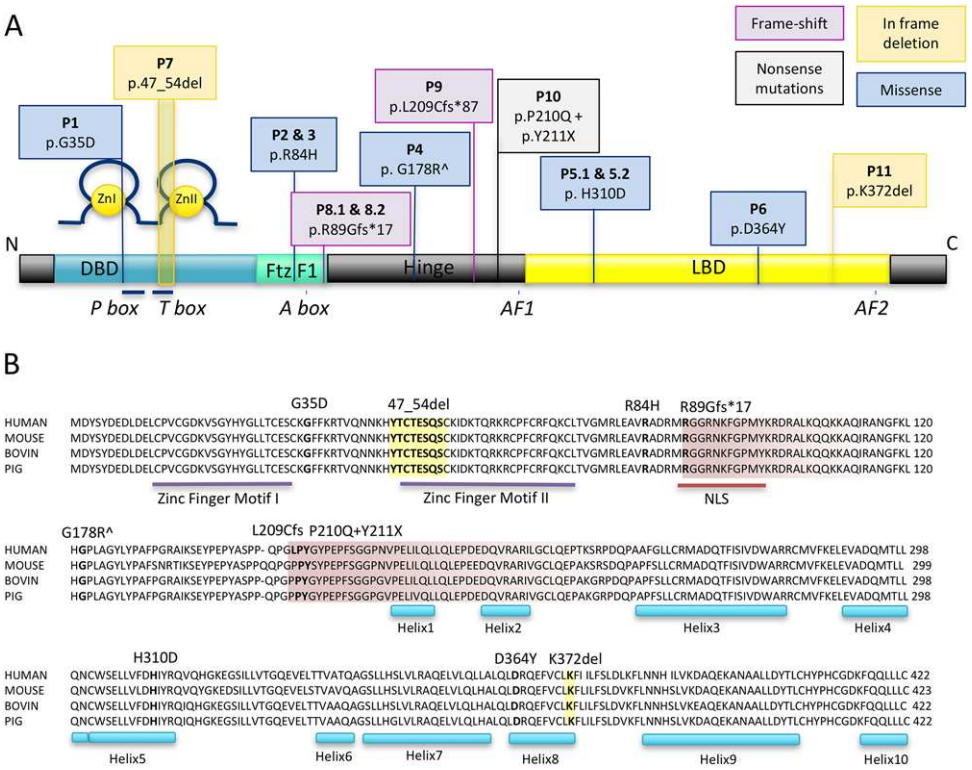

FIGURE 1 Variants in SF1 identified in patients with disorders of sex development using massively parallel sequencing. **A**: A schematic representation of the predicted protein structure of SF1 showing the approximate location of the variants identified in a cohort of DSD patients. The protein domains are as follows: DNA binding domain (DBD) containing two zinc finger motifs (Zn) and the Fushi-tarazu factor 1 box (Ftz-F1), the hinge region and ligand binding domain (LBD). P Box, T-box, A-box, as well as two activational domains—AF1, AF2. Patient variants (with patient number denoted by P#) are shown. Five missense mutations were identified in our cohort of 46,XY DSD patients, some of which were recurrent. Two fell within the DBD (p.G35D, p.R84H) one in the hinge region (p.G178R), two were in the LBD (p.H310D and p.D364Y). Two in-frame deletions were found, one within the second zinc finger motif (p.47_54del) and the other in the LBD (p.K372del); two frame-shifts were identified (p.R89Gfs*17 and p.L209Cfs*87) as well as a nonsense mutation at position 211. Blue boxes denote single nucleotide variants, yellow are in-frame deletions and pink are variants that cause a frame-shift and the black box is a nonsense mutation. **B**: Evolutionary conservation of the SF1 protein (the last 40 amino acids are not shown). All variants are indicated, in-frame deletions are shaded in yellow while frame shifts are indicated with pink shading at the point of the new transcript, in addition the zinc finger motifs, nuclear localization signal (NLS), and the conserved helices which are in the LBD are shown

eight patients, five of whom had normal levels, while three were considered low (Table 1). All of the individuals were found to have normal adrenal function at the time of sample collection, based on hormonal profiling and clinician interpretation.

3.2 | Variants in *NR5A1* affect conserved residues in important SF1 protein domains

SF1 protein comprises a DNA binding domain (DBD) containing two zinc finger motifs (Zn I and II), a ligand binding domain (LBD) made up of 12 helices, which interacts with various cofactors, as well as a hinge region that is crucial for stabilizing the interactions of the LBD, controlling SF1 transcriptional activity and is subject to post-translational modification (Figure 1A). SF1 also contains a Fushi-tarazu factor 1 box (Ftz-F1), which aids in DNA anchoring; as well as P, T, and A boxes, which assist with DNA binding. Two activation functional domains (AF1 and 2) are also present and AF2 is crucial for trans-activational activity (Hoivik et al., 2010) (Figure 1A).

Together our 15 patients had 12 unique *NR5A1* variants, nine of which are novel (Table 2). Missense variants were identified in seven individuals who presented with a range of phenotypes (Tables 1 and 2). Further investigation of the variants showed that each change affected highly conserved amino acid(s) (Figure 1B) and all changes were deemed to be damaging by the in silico algorithms used (PolyPhen2, MutationTaster, SIFT, LRT; numerical output from the in silico analyses can be found in Supp. Table S5). The only exception was the known c.G532A (p.Gly178Arg, rs543895681:G>A) variant, which was predicted damaging only by PolyPhen2. We mapped these variants on a schematic representation of the predicted SF1 protein (see Figure 1A). Patient 1 had a variant in the P-box of the first zinc finger motif within DBD (c.G104A, p.Gly35Asp). Two unrelated individuals (Patients 2 and 3) both harbored a previously described change c.G251A (p.Arg84His, rs548473217:G>A), which falls within the Ftz-F1 box. This variant was previously found in a 46,XY patient with severe hypo-androgenization (Köhler et al., 2008) and acts as a positive control in our study. A rare variant c.G532A (p.Gly178Arg ExAC all population total frequency: 4.544e−05) was identified in the hinge region of Patient 4 who was initially suspected of having androgen insensitivity. An androgen receptor mutation was ruled out using our targeted MPS panel (Table 1). Three variants were identified in the LBD. In a pair of

TABLE 2 Molecular characteristics of the *NR5A1* variations identified in our cohort

Patient ID	DNA change	Protein Change	Zygosity	Inheritance	Novel or known	In silico
1	c.G104A	p.Gly35Asp	Heterozygous	N/A	Novel	Damaging
2	c.G251A	p.Arg84His	Heterozygous	N/A	Köhler 2008 rs548473217 ExAC all pop.: 4.262e-05	Damaging
3	c.G251A	p.Arg84His	Heterozygous	N/A	Köhler 2008 rs548473217 ExAC all pop.: 4.262e-05	Damaging
4	c.G532A^	p.Gly178Arg^	Heterozygous	N/A	rs543895681 ExAC all pop.: 4.544e-05	1 of 4
5.1	c.C928G	p.His310Asp	Heterozygous	N/A	Novel	Damaging
5.2	c.C928G	p.His310Asp	Heterozygous	N/A	Novel	Damaging
6	c.G1090T	p.Asp364Tyr	Heterozygous	N/A	Novel	Damaging
7	c.140_163del	p.47_54del	Heterozygous	N/A	Novel	Loss of Zn II finger
8.1	c.265delA	p.Arg89Glyfs*17	Heterozygous	Paternal	Novel	Frameshift in DBD
8.2	c.265delA	p.Arg89Glyfs*17	Heterozygous	Father of 8.1	Novel	Frameshift in DBD
9	c.624delG	p.Leu209Cysfs*87	Heterozygous	N/A	Novel	Frameshift in Hinge
10	c.[C629A;C633A]	p.[Pro210Gln;Tyr211*]	Heterozygous	N/A	Novel	Termination in hinge
11	c.1114_1116del^	p.Lys372del^	Heterozygous	N/A	Eggers 2015	In-frame del in LBD
12	c.G991-1C splice site acceptor^	Intron 5	Heterozygous	N/A	Novel	Loss of WT splicing
13	c.G1139-1T splice site acceptor^	Intron 6	Heterozygous	De novo	Novel	Loss of WT splicing

In silico (only available for missense changes): PolyPhen2, MutationTaster, SIFT, LRT; damaging: deleterious or possibly deleterious in 4/4 predictors—numerical output can be found in Supp. Table S5; DBD, DNA Binding domain; LBD, ligand binding domain; ^, not included in functional analyses; N/A, not available. Known refers to either or previous publication and ExAC; All pop., all populations total. DNA mutation numbering is based on GenBank reference DNA sequence NM_04959.4, with the A of the ATG initiation codon designated +1. Predicted protein annotations are based on NP_004950.

siblings (Patients 5.1 and 5.2) with suspected PGD and retained Müllerian remnants, we identified a c.C928G change, predicted to lead to a histamine to aspartic acid transition at residue 310, which falls within the highly conserved alpha helix 5 of the LBD. Patient 6 was found to have c.G1090T, predicted to lead to a aspartic acid to tyrosine substitution adjacent to predicted alpha helix 8 at position 364 (Figure 1A).

Six patients had in-frame deletions or nonsense mutations (Table 2). We identified an in-frame deletion of eight amino acids (c.140_163del, p.47_54del) in the second zinc finger motif of the DBD in a female with CGD (Patient 7). Another in-frame deletion of a single codon (c.1114_1116del, p.Lys372del) was identified in Patient 11 who has proximal hypospadias. The same variant has been previously published by us in an unrelated Indonesian family (Eggers et al., 2015). A paternally inherited frame-shift with a premature stop codon insertion after 17 amino acids (c.265Adel, p.Arg89Glyfs17*), truncating the protein from the accessory DBD, was identified in a Patient 8.1 who presented with penoscrotal hypospadias and a hormonal profile suggestive of gonadal dysgenesis. The patient's father, Patient 8.2 (also included in the study) had hypospadias as a child. Additionally, Patient 9 was found to have a single base pair deletion at coding position 624 that lead to a frame-shift midway through the hinge region of the protein and premature stop codon at position 87 of the new transcript (c.624delG, p.Leu209Cysfs87*). Lastly, we identified a nonsense mutation (c.[C629A;C633A]:p.[P210Q;p.Y211*] on the same allele) in an individual with penoscrotal hypospadias (Patient 10 who had micropenis with testicular tissue, minimal T response following hCG), most intriguingly this individual has entered spontaneous and progressive puberty (Table 1).

3.3 | *NR5A1* splice site variants and oligogenecity in DSD

Two splice site acceptor mutations in *NR5A1* were identified within our cohort (Table 2 and Figure 2). Patient 12 who presented with clitoromegaly and no internal female structures (Table 1) had a heterozygous splice site mutation in intron 5 (c.G991-1C, NC_000009.11:g.127253508C>G) predicted to affect the splice site acceptor for exon 6 (Table 2 and Figure 2A). Patient 13, who had proximal hypospadias, bifid scrotum, and chordee, was found to have a de novo splice site mutation in intron 6 (c.G1139-1T, NC_000009.11:g.127245285C>A)

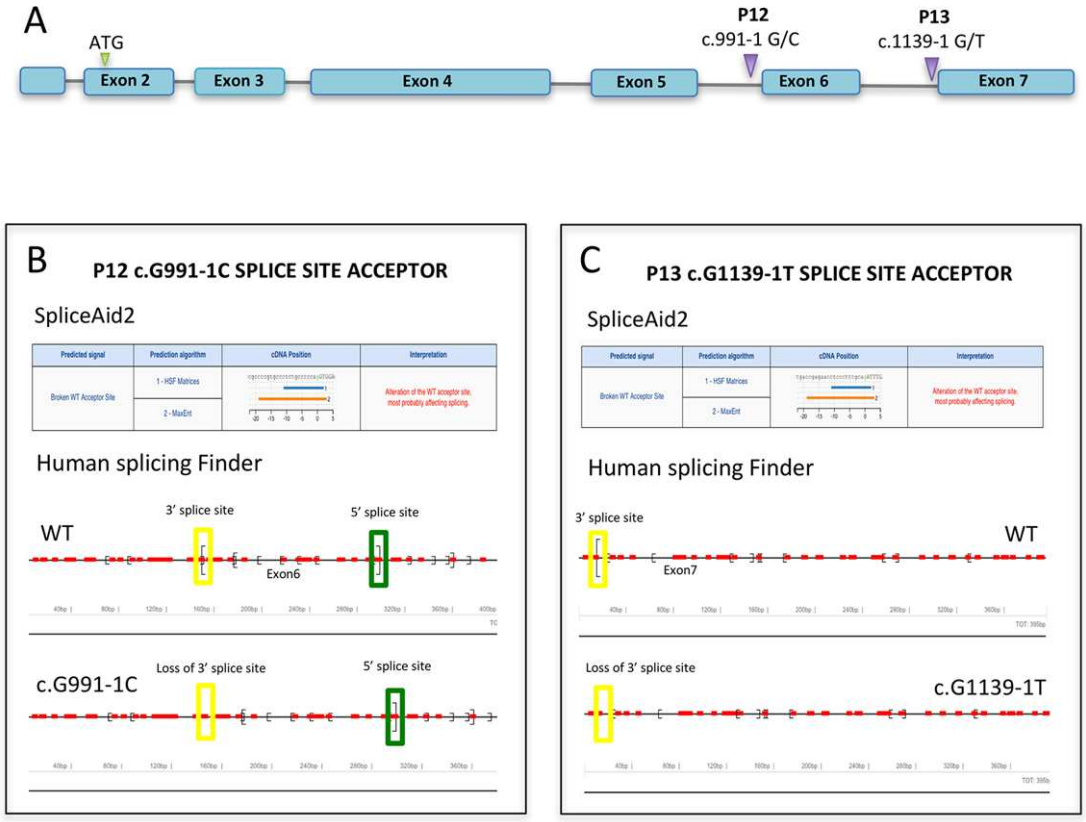

FIGURE 2 Splice site mutations identified in 46,XY individuals. **A**: A schematic representation of *NR5A1*, showing the six coding exons and the position of the two splice site acceptor mutations identified. **B and C**:Human Splicing Finder (Desmet et al., 2009) and SpliceAid 2 (Piva et al., 2012) were used to analyze the consequences of the two splice site acceptor mutations. Both of which suggest that wild-type splicing would be affected

predicted to be affecting the splice site acceptor for exon 7 (Table 2 and Figure 2A). Publicly available prediction programs Human Splicing Finder (Desmet et al., 2009) and SpliceAid 2 (Piva et al., 2012) were used to analyze the consequences of the two splice site acceptor mutations. Both programs predicted that each mutation has a high chance of loss or alteration of wild-type splicing of the acceptor site. In the instance of c.G991-1C, a new splice acceptor site is predicted (exon length variation -6), which would skip bases 991–996 and may lead to the in-frame deletion of residues 330 and 331, which fall within the LBD (p.330_331del, according to the HSF Matrices algorithm) (Figure 2B). While the c.G1139-1T change is predicted to lead to a new splice acceptor site eight base pairs into exon 7, which may lead to a frame-shift at position p.Asp380Alafs (Figure 2C).

A recent study (Mazen et al., 2016) identified pathogenic variants in *NR5A1* and *MAP3K1* in an individual with 46,XY gonadal dysgenesis, highlighting the possibility that digenic inheritance may play a role in the large phenotypic spectrum associated with *NR5A1* variants. We also identified two patients with multiple affected diagnostic DSD genes (Table 3). Patient 3 was found to harbor a previously reported p.Arg84His variant in *NR5A1*, alongside a rare variant in *ZFPM2* (c.A2107C, p.Met703Leu, rs121908603:A>C), which has been previously reported in individuals with a diaphragmatic hernia 9 (Bleyl et al., 2007) (Table 3). We also identified a monoallelic change in *SRD5A2* (c.G680A, p.Arg227Gln, rs9332964:G>A) in Patient 11, who also harbored a single codon deletion at position 372 of *NR5A1* (Table 3).

3.4 | All novel SF1 variants have reduced trans-activational activity

To assess the activity of the seven novel protein-affecting *NR5A1* variants, we used a well-established Dual-Luciferase reporter assay. Wild-type and variant human SF1 activity was assessed in this assay using the mouse Tesco-Sox9 (mTesco) reporter, a construct that responds strongly to SF1 with its co-activators in vitro. *hNR5A1*-expression plasmids (either wild-type or novel variant) were transfected into COS-7 cells with the mTesco-Luciferase reporter and a *Renilla* control plasmid. hSF1 was transfected either alone, with hSOX9 or with hSRY (Figure 3). A previously reported pathogenic variant—p.Arg84His (Köhler et al., 2008)—was used as a positive control. Negative controls (no SF1) were included for each condition. Consistent with previous reports (Sekido & Lovell-Badge, 2008), SF1 (wild-type or variant) alone showed little activation of the mTesco reporter (Figure 3). In contrast, when wild-type SF1 was co-transfected with hSOX9 expression vector a significant increase in transactivation activity is seen (four times that seen for wild-type SF1 alone) (Figure 3). A complete loss of this trans-activational activity was noted for the proteins with missense mutations located in the LBD (p.His310Asp and p.Asp364Tyr), the SF1 p.47_54del, as well as both frame-shift mutations assayed (p.Arg89Glyfs*17 and p.Leu209Cysfs*87) (Figure 3). Intriguingly, a nonsense SF1 variant (p.[Pro210Gln;Tyr211*]) seemed to retain a low level of activity (Figure 3). A similar pattern was seen

TABLE 3 Oligogenic inheritance of DSD-related genes

Patient ID	Gene	Coding change	Protein change	Zygosity	Novel or known	In silico
3	*NR5A1*	c.G251A	p.Arg84His	Heterozygous	Köhler et al., 2008	Damaging
	ZFPM2	c.A2107C	p.Met703Leu	Heterozygous	Bleyl et al. (2007) rs121908603	3 of 4
11	*NR5A1*	c.1114_1116del^	p.Lys372del^	Heterozygous	Eggers et al. (2015)	In-frame del
	SRD5A2	c.G680A	p.Arg227Gln	Heterozygous	Sasaki et al. (2003) rs9332964 as homoz.	N/A

In silico (only available for missense *changes*): PolyPhen2, MutationTaster, SIFT, LRT; damaging: deleterious or possibly deleterious in 4/4 predictors—numerical output can be found in Supp. Table S5; ^, not included in functional analyses. Known refers to both literature and ExAC. *NR5A1* DNA mutation numbering is based on GenBank reference DNA sequence NM_04959.4, with the A of the ATG initiation codon designated +1. Predicted protein annotations are based on NP_004950. *ZFPM2* DNA mutation numbering is based on the GenBank reference DNA sequence NM_012082.

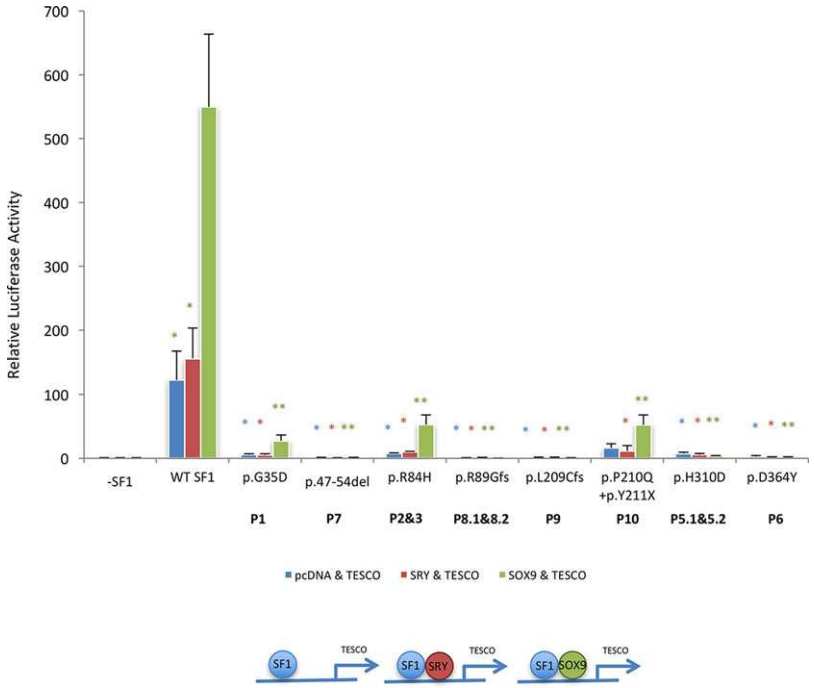

FIGURE 3 Reduced trans-activational activity of SF1 variants. The trans-activational activity of each variant compared with wild-type SF1 was tested using a dual-luciferase reporter assay in COS-7 cells. Each SF1 variant was transfected either alone (blue bars), with hSRY (red bars) or with hSOX9 (green bars) and transcriptional activity was measured by activation of the mTesco promoter driven luciferase reporter (p.GL4). Empty vector, in place of SF1 was run in each condition as a negative control. Briefly wild-type SF1 alone shows only low activation of mTesco, however trans-activation activity increased around four times with the addition of SOX9. All SF1 variants tested show a significant decrease in trans-activational activity with SOX9 and with SRY. Complete loss of trans-activational activity was noted for the proteins with missense mutations located in the LBD (p.H310D and p.D364Y), the p.47_54 in-frame deletion, as well as both frame-shift mutations assayed (p.R89Gfs*17 and p.L209Cfs*87). The non-sense mutation SF1 (p.[P210Q;Y211*]) seemed to retain a low level of activity. Data represent the mean with standard error of four independent experiments performed in duplicate transfections. Unpaired *t*-test was applied and for $^{**}P$ value < 0.005; $^{*}P$ value < 0.05

with the SRY/SF1 transfected cells; however, the magnitude of activation was in general lower than that of the SOX9/SF1 transfection (Figure 3). All the SF1 variants identified in our DSD patients showed reduced transactivation activity in vitro when co-transfected with SRY or SOX9. This suggests that the reason these variants are pathogenic is because they result in a dramatic reduction or loss of SF1 transactivation activity in these patients.

3.5 | Protein expression is affected in some SF1 variants

A reduction in SF1 activity could be due to several factors; unstable mRNA or protein, protein conformational changes or a reduction in DNA or co-activator binding. To assess whether the SF1 variants affected the levels or localization of the protein, we used immunofluorescence (Figure 4). Wild-type SF1 shows strong nuclear localization with nucleolar exclusions as previously described (Köhler et al., 2008) (Figure 4A-b and C-b). Missense variants p.His310Asp (Figure 4A-x) and p.Asp364Tyr (Figure 4A-aa) were found to be localized within the nucleus, albeit with lower expression levels than that of the wild-type protein. The p.Arg84His variant used as a control (Köhler et al., 2008) seemed to be highly concentrated at the nuclear edge (Figure 4A-k&l). This was not the case in the previously published study and may be cell type dependent. SF1 p.Gly35Asp was dispersed within the cell cytoplasm with low levels of staining in the nucleus (Figure 4C-d). Of note, SF1 p.47_54del was localized in sub-nuclear aggregates. This has been described previously for other mutations in SF1, that is, p.Cys33Ser (Köhler et al., 2008), p.Val15Met, p.Met78Ile (Lin et al.,

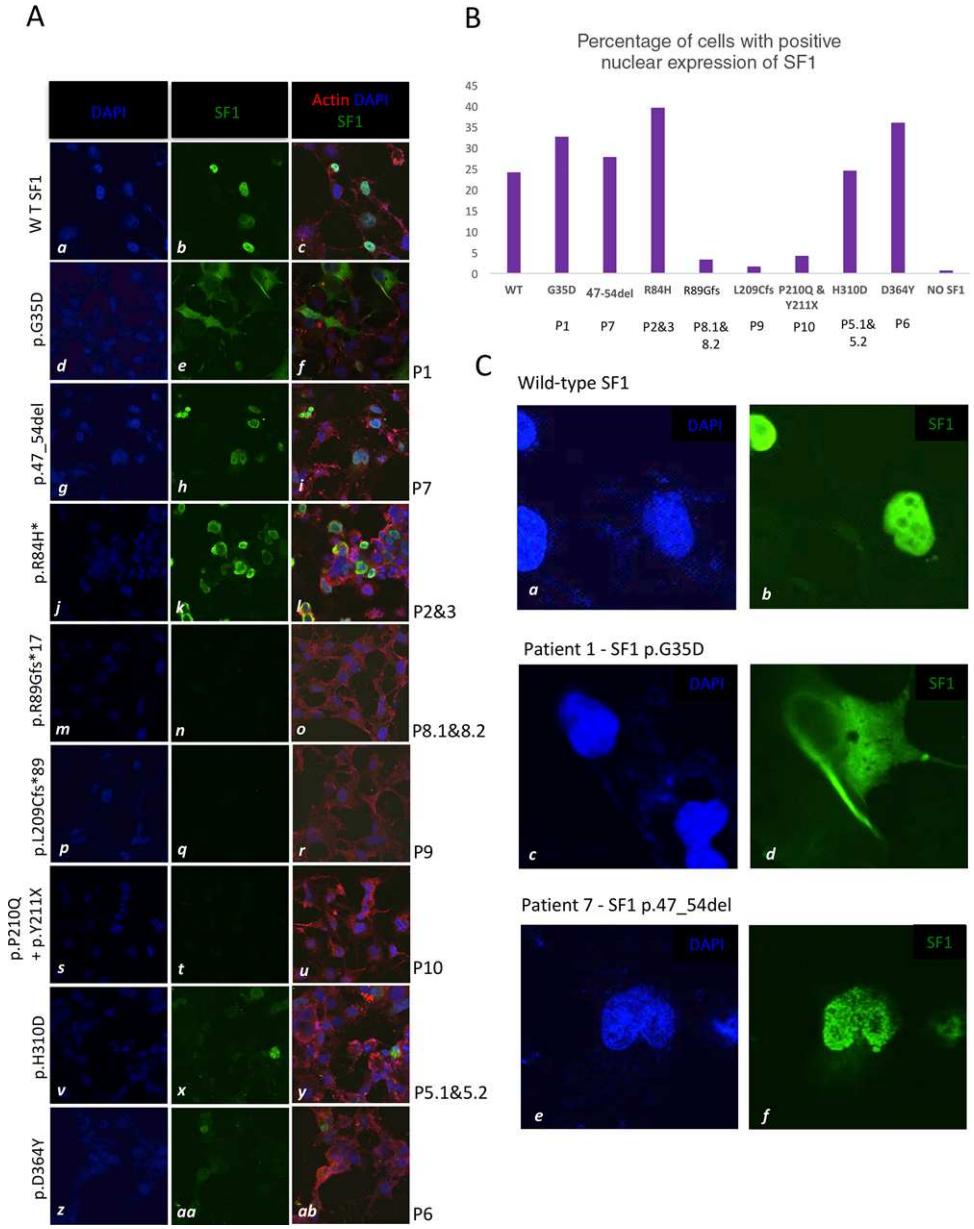

FIGURE 4 **A:** Protein expression is affected in some SF1 variants. Protein expression of each variant and wild-type SF1 was assessed in COS-7 cells with an SF1 antibody (green). Cells were transfected with an equal amount of SF1 expression vector (wild-type or variant). Nuclear counterstaining was performed with DAPI (blue) and the cytoskeleton was stained with actin (red). Wild-type SF1 showed strong nuclear staining with nucleolar exclusions. All SF1 expression vectors with missense mutations and in-frame deletion were found to be expressed while none of the truncated proteins were detected (p.R89Gfs*17, p.L209Cfs*87, and p.[P210Q;Y211*]). **B:** This was quantified as the number of SF1 expressing nuclei per image (2–10×, two cropped areas c.f. 40× displayed in the image). **C:** Mutant protein p.G35D was dispersed through the cytosol; while the in-frame deletion affecting the Zn II motif was clumped in sub-nuclear aggregates; and p.R84H seemed to be concentrated on the nuclear border

2007), and they are thought to represent nuclear domain bodies which inhibit the transcription factors' ability to enter in the nuclear cytosol thereby affecting trans-activational activity. The proteins predicted to result in frame-shifts (p.Arg89Glyfs*17 and p.Leu209Cysfs*87) as well as the nonsense change (p.[Pro210Gln;Tyr211*]) were not expressed at all. This may be due to RNA mediated decay or unstable protein being produced. Alternatively, even though a polyclonal antibody was used in the staining of SF1, the antibody may not be able to bind to the altered epitopes of the mutated protein. We found all missense changes to be expressed at levels comparable, if not higher to that of wild-type (Figure 4B).

3.6 | Protein conformational changes are seen in some SF1 variants

To investigate the potential impact of each variant on protein conformation we performed an in silico analysis. Results are shown for affected residues in individuals with gonadal dysgenesis, some with retained Müllerian remnants (Figure 5). The truncated proteins have also been imaged and can be found in Supp. Figure S1. In the wild-type SF1, the glycine at position 35 falls within a highly specific ^{34}KGFFK^{38} motif, where both lysine residues are thought to be subjected to post-translational acetylation. In Patient 1, the glycine residue which is a

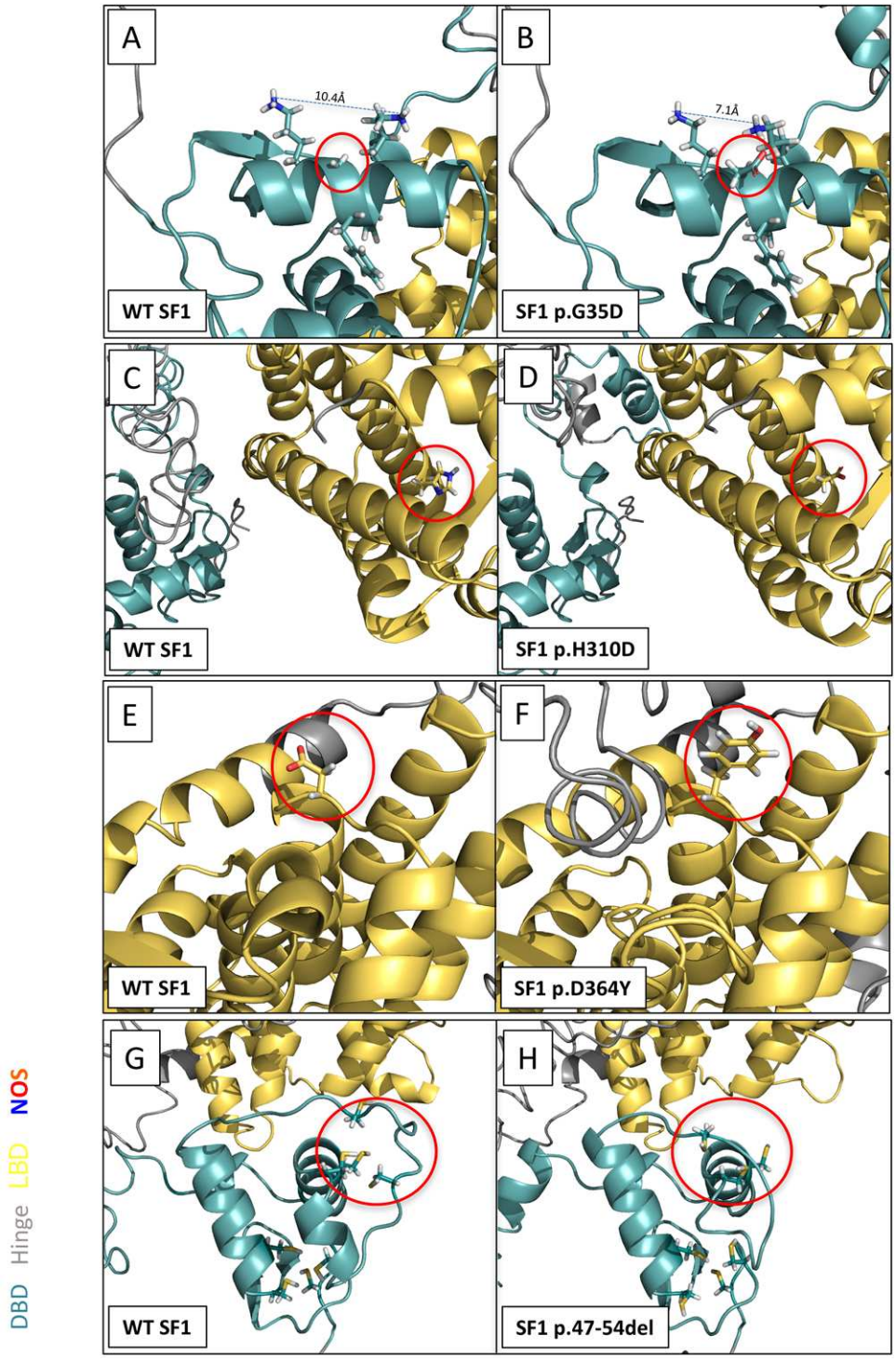

FIGURE 5 Conformational changes in SF1 variant proteins. To investigate the potential impact of each variant on protein conformation, we performed an in silico prediction with the WT SF1 and each variant using I-Tasser and PyMol modeling software. For variants identified in patients with CGD or PGD, the side chains are shown on the implicated residues. **A**: Wild-type Gly at position 35 falls within a highly specific ^{34}KGFFK^{38} motif (motif side chains shown), where both lysine residues are thought to be subjected to post translational acetylation. The glycine residue is substitute with a larger, negatively charged aspartic acid, **(B)** decreasing the distance between the two lysine residues at position 35 and 38, from 10.4 to 7.1 A. **C**: Wild-type His residue at position 310 falls within the highly conserved alpha helix 5 of LBD (circled) and is a large, neutral amino acid commonly involved in stacking, **(D)** while mutant Asp is negatively charged and much smaller. **E**: The Asp to Tyr transition at position 364, which falls within the LBD, adjacent to a highly conserved alpha helix 8. The wild-type Asp is quite small and secondary structure is predicted be at a turn, **(F)** unlike the bulky Tyr, which would not have that secondary structure. **G**: Eight amino acid deletion in the second Zn finger motif of the DNA binding domain, **(H)** results in a loss of one of the cysteine residues, which is crucial for the Zn2+ interaction and possible loss of stability and ability to bind conical 6-bp HRE

very small, neutral, and highly flexible amino acid is substituted with a larger, negatively charged aspartic acid. The resultant change appears to decrease the distance between the two lysine residues at position 34 and 38, from 10.4 Å to 7.1 Å. (Figure 5A and B), which may affect the cyclical acetylation and deacetylation of the motif.

In the pair of siblings who presented with PGD with retained Müllerian remnants, a histidine residue at position 310 has been replaced with aspartic acid. This His310 is in the highly conserved alpha helix 5 of the LBD and is believed to be involved in ligand interaction (Venselaar et al., 2010). The large and neutral side chain of histidine is commonly involved in stacking (Betts and Russell 2003) and may play a role in the stability of the alpha helix pocket that is crucial for activity (Hoivik et al., 2010). The mutant aspartic acid is negatively charged and much smaller, and this substitution is predicted to lead to loss of external interactions (Venselaar et al., 2010) (Figure 5C and D).

We also found that replacing the aspartic acid at position 364 with a bulky and highly reactive tyrosine, as is observed in Patient 6 had a strong effect on SF1 conformation (Figure 5E and F). This residue is adjacent to the evolutionarily conserved alpha helix 8. The wild-type aspartic acid is predicted to form hydrogen bonds with the glutamic acid at position 367 and tyrosine at position 404, as well as a salt bridge with the serine at position 218 and the lysine at 415 (Venselaar et al., 2010); the difference in hydrophobicity and charge of the mutant residue will thereby affect hydrogen bond formation as well as ionic interactions, respectively.

Lastly, we modeled an in-frame deletion of eight residues located in the second zinc finger motif of the DBD which was identified in a patient with CGD (Patient 7) (Figure 5G and H). The deletion results in a loss of one of the cysteine residues which is crucial for Zn^{2+} interaction. Without the second zinc finger motif, the stability of the mutant protein may be compromised and its ability to bind the zinc metal ion, and in turn the conical 6-bp sequence hormone response element is likely affected (Little et al., 2006).

4 | DISCUSSION

SF1 has highly specific interactions with a multitude of co-factors at various time points throughout development. Variants in *NR5A1* have been shown to be causative in a substantial number of 46,XY DSD patients and are now associated with an ever increasing DSD phenotypic spectrum. Yet, to date, a clear correlation between genotype and phenotype remains elusive, plagued with complicated heritability, varying degrees of expressivity, incomplete penetrance and limited understanding of the molecular function of the SF1 protein.

In a previous study, we found that pathogenic variants in *NR5A1* accounted for 15%–20% of 46,XY DSD cases and we identified numerous novel, uncharacterized variants in this gene (Eggers et al., 2016). In this study, we set out to examine these newly identified variants by assessing their pathogenicity using a variety of in vitro and in silico methods. Our studies show that all these novel *NR5A1* variants have reduced SF1 function, suggesting they contribute to the patients' phenotypes. Nevertheless, we found little correlation between the severity of the phenotype and the nature of the change or its associated reduction in SF1 function. This may be due to a variety of factors, some of which we discuss below.

It is known that the same genetic changes in *NR5A1* can display markedly different clinical phenotypes (variable expressivity). One explanation for this may be oligogenecity, highlighted by a recent study in which pathogenic variants in *NR5A1* and *MAP3K1* were identified in an individual with 46,XY gonadal dysgenesis (Mazen et al., 2016). We found two additional examples of this. Patients 2 and 3, in whom we identified the known p.Arg84His change, have greatly differing clinical phenotypes. This p.Arg84His change was first identified in a severely under-virilized child with inguinal testes and no Müllerian structures (Köhler et al., 2008). In keeping with this, our Patient 2 had an inguinal gonad with marked testosterone production (although low levels of DHT), which is suggestive of some testicular function. However, Patient 3 with the same variant had a more severe clinical presentation with atrophic testes, no testosterone and seminoma and carcinoma in situ, currently known as Germ Cell Neoplasia In Situ (GCNIS) (Berney et al., 2016). Our functional analyses showed that the altered protein retains a low level of trans-activational activity, while protein expression and localization was not affected. Most intriguingly, the more severely affected of the two patients, Patient 3, was found to harbor a known heterozygous variant in *ZFPM2* (c.A2107C, p.Met703Leu, rs121908603:A>C), a testes determining gene associated with heart anomalies, and this variant has been previously reported in an individual with diaphragmatic hernia (Bleyl et al., 2007). We postulate that the cumulative effect of these changes in two different genes may be contributing to the patient's more severe phenotype.

In a second example, we identified a monoallelic change in *SRD5A2* (c.G680A, p.Arg227Gln, rs9332964:G>A), in conjunction with the single amino acid deletion at position 372 of SF1. Monoallelic inheritance of *SRD5A2*, although uncommon, has been reported in a severely under-virilized individual with hypospadias and bilateral inguinal testes (Chávez, Ramos, Gómez, & Vilchis, 2014). Additionally, the p.Arg227Gln *SRD5A2* change has been previously found to be causative of micropenis, where it was found in compound heterozygosity or homozygosity in three individuals (Sasaki et al., 2003). Our findings further highlight the possibility that oligogenic inheritance may play a role in the large phenotypic spectrum associated with *NR5A1* variants, and this may involve numerous genes associated with gonadal development or steroidogenesis.

In addition to known DSD genes, other as yet uncharacterized genes may also contribute to the vast spectrum of *NR5A1*-associated phenotypes. Environmental or epigenetic changes may also play a role (Gunes, Metin Mahmutoglu, & Agarwal, 2016). Another example of variable expressivity in our cohort was found in Patient 8.1, a severely under-virilized child with a hormonal profile suggestive of gonadal dysgenesis. This patient has a paternally inherited frame-shift in *NR5A1*. Our in vitro analyses of the truncated protein showed complete loss of activity for the p.Arg89Glyfs17* variant; yet the father presented with a much milder phenotype of hypospadias and retained spontaneous fertility. Paternally inherited cases of *NR5A1* insufficiency have been previously reported where the father has been asymptomatic

(Philibert et al., 2011; Swartz et al. 2017; Yagi et al., 2015) or affected with retained spontaneous fertility (Baetens et al., 2014; Fabbri et al., 2016). Furthermore, there have been reports of siblings with the same mutation who present with vastly different effects on Leydig and Sertoli cell function (Coutant et al., 2007; Philibert et al. 2007). Such variability between individuals harboring identical genetic changes, even within a single pedigree may be a result of a difference in expression of wild-type and mutant allele (allelic expression imbalance) or even a potential mosaicism. Although beyond the scope of this paper, analysis of allelic expression levels in patient cells will be important to assess this effect in the future.

Although we did not find a strong correlation between the protein changes and patient phenotypes, our analysis does reveal several interesting observations about SF1 form and function. Much is known about the protein structure of SF1, yet analysis of mutations in specific regions or amino acids highlight their importance in different biological functions, including how they may affect certain essential post-translational modifications (reviewed in Gunes et al., 2016). A ^{34}KGFFK^{38} motif that resides within the DBD is known to be important for SF1 function, as the post-translational acetylation of lysine residues at position 34 and 38 regulates trans-activational activity and sequence specific DNA binding (Jacob, Lund, Martinez, & Hedin, 2001). Indeed, previous studies have shown that inhibition of deacetylation led to the accumulation of SF1 in the cytosol (Chen, Weng, Huang, & Chung, 2007). One of the first mutations characterised in SF1 was a glycine to glutamic acid substitution at codon 35 in an individual with complete gonadal dysgenesis and adrenal insufficiency (Achermann, Ito, Ito, Hindmarsh, & Jameson, 1999). In the current study, we identified and characterized a glycine to aspartic acid transition at the same position (amino acid 35), enforcing the importance of this residue. We found that the mutant protein localized in the cytosol (compared to wild-type nuclear expression). Our in silico modeling suggests that the glycine to aspartic acid change may alter the structure and the resultant interactions of the motif as it introduces a negatively charged, hydrophobic residue which changes the conformation of the motif and decreases the distance between the two lysine residues. Inhibition of deacetylation has been shown to reduce steroidogenesis through ubiquitination and degradation of *NR5A1* (Chen et al., 2007). This is consistent with the phenotype of our patient, who was found to have Leydig cell hypoplasia and no detectable testosterone production.

In addition to its role in steroidogenesis, another key function of SF1 is the induction of AMH (Shen et al., 1994). This process is thought to be reliant on a multi protein complex, which includes the likes of GATA4, WT1, SOX9, and NR0B1 (Arango, Lovell-Badge, & Behringer, 1999). AMH is responsible for the repression of Müllerian duct development, and Müllerian remnants are absent in 75% of affected individuals (Pedace et al., 2014) suggesting that in these cases a certain amount of SF1 activity (and therefore AMH production) has been retained. However, the AMH receptor, AMHR2, is also thought to be upregulated through synergistic interactions between β-catenin and SF1 (De Santa Barbara et al. 1998b). Little is known about this role of SF1 and the site of protein–protein interaction has yet to be mapped. Interestingly, we observed that in a pair of siblings with suspected PGD and retained Müllerian remnants, normal AMH levels were recorded. Siblings 5.1 and 5.2 had a single amino acid change at residue 310 in SF1, which falls within the highly conserved helix 5 of the LBD, and is thought to be involved in ligand interactions (Vanselaar et al., 2010). While the mutant SF1 protein was still expressed, it had a complete loss of trans-activational activity in the Tesco luciferase assay. However, given the normal levels of AMH, these results suggest that AMHR2 expression may have been hampered in these individuals. Further validation of the role of the SF1 alpha helix 5 and the possible interaction with β-catenin and the *AMHR2* promoter are warranted.

Outcomes can differ vastly in DSD patients with *NR5A1* mutations. Some individuals with *NR5A1* variants present with under-virilized genitalia and a hormonal profile resembling disorders of androgen synthesis or action, such as androgen insensitivity or 5-alpha reductase deficiency, where serum testosterone at the neonatal stage can be normal or even high (Fabbri et al., 2014; Swartz et al. 2017; Wu et al., 2013), and there have been reports of pubertal androgen production and virilization in a number of patients (Pedace et al., 2014). In keeping with previous findings, three of our patients (siblings 5.1 and 5.2 and Patient 10) who were severely under-virilized in the neonatal period, have had progressive and spontaneous androgenization and pubertal development, suggesting partial rescue of Leydig cell function in the pubertal gonads. As mentioned above, siblings 5.1 and 5.2 shared a single amino acid change within the LBD but by contrast, Patient 10 had a premature termination at residue 211 in the hinge, and yet retained some trans-activational activity. Therefore, the genetic etiology of a DSD patient does not allow us to predict the eventual outcome for that individual. Clearly, further studies are needed to address this issue.

Patients with a DSD can also have an increased risk of gonadal germ cell tumors (Hersmus et al., 2017). Although only a limited number of cases have been reported in individuals with gonadal dysgenesis and *NR5A1* mutations, they do appear to have a heightened cancer risk (Barbaro et al., 2011; Cools et al., 2012; Cools, Looijenga, Wolffenbuttel, & T'Sjoen, 2014). Indeed, Patient 3 was found to have seminoma and GCNIS while Patient 6 was found to have gonadoblastoma. These patients were both 46,XY DSD (females) with gonadal dysgenesis. While they both have variants affecting different SF1 protein domains, with different loss of SF1 activity, they share the risk of development of the precursor of germ cell tumors, either GCNIS or gonadoblastoma, which appears to be higher in the more severely affected patient. Surveillance in these patients is highly recommended, irrespective of mutation type.

Undoubtedly there are significant implications for patients and families diagnosed with a *NR5A1*-associated DSD. Given the known incomplete penetrance and variable expressivity of these changes, one cannot assume these changes are de novo. Consequently, if a variant in *NR5A1* is identified in a DSD patient, the family of the index patient ought to be screened for the change. Where applicable, fertility preservation should be considered as there are reports of progressive decline in testicular function (Philibert et al., 2011) as well as the possibility of ovarian dysfunction in 46,XX carriers (Baetens et al., 2014; Fabbri et al. 2014).

5 | CONCLUSIONS

Phenotypic variability among individuals with *NR5A1* mutations is remarkable. Our study and others (Camats et al., 2012; Swartz et al. 2017) have suggested that neither the mutation in *NR5A1* nor its reduction in SF1 activity is a good indicator of a patient's phenotype or clinical outcome. Inherited mutations in DSD tend to be rare as fertility is often affected; however, familial cases of *NR5A1* including those shown here are shifting this paradigm (Brauner et al., 2016). Our studies suggest that oligogenecity may be a contributing factor in the variable expressivity and incomplete penetrance is often observed with *NR5A1* variants in DSD patients. However, neither of these factors explains all phenotypic differences observed. Additional genes, polymorphisms, and contributing environmental factors may also play a role. Only very large-scale studies that compare a range of *NR5A1* mutations and DSD patient phenotypes, as well as investigations into patient cell specific effects, will likely reveal the complexities of these interactions.

ACKNOWLEDGMENTS

We would like to thank all the patients and families for taking part in our research as well as all the clinicians involved who help us in recruiting patients. We thank The Translational Genomics Unit at the MCRI/VCGS and the AGRF for sequencing.

DISCLOSURE STATEMENT

The authors declare no conflict of interest.

AUTHOR CONTRIBUTION

G.R., K.A., and A.S. compiled and wrote the manuscript. Library preparation and patient sample management was performed by G.R., JvdB, and S.E. Functional in vitro analysis was performed by G.R. K.A. and A.S. provided overall project supervision. All other authors were involved in recruitment of patients and critically reviewed the manuscript.

ORCID

Katie L. Ayers http://orcid.org/0000-0002-6840-3186

REFERENCES

Achermann, J. C., Ito, M., Ito, M., Hindmarsh, P., & Jameson, J. L. 1999. A mutation in the gene encoding steroidogenic factor-1 causes XY sex reversal and adrenal failure in humans. *Nature Genetics*, *22*, 125–126.

Allali, S., Muller, J-B., Brauner, R., Lourenço, D., Boudjenah, R., Karageorgou, V., … Bashamboo, A. (2011). Mutation analysis of NR5A1 encoding steroidogenic factor 1 in 77 patients with 46, XY disorders of sex development (DSD) including hypospadias. *PLoS One*, *6*, e24117–e2418.

Arango, N. A., Lovell-Badge, R., & Behringer, R. R. (1999). Targeted mutagenesis of the endogenous mouse. *Cell*, *99*, 409–419.

Baetens, D., Mladenov, W., Delle Chiaie, B., Menten, B., Desloovere, A., Iotova, V., … Cools, M. (2014). Extensive clinical, hormonal and genetic screening in a large consecutive series of 46,XY neonates and infants with atypical sexual development. *Orphanet Journal of Rare Diseases*, *9*, 209.

Barbaro, M., Cools, M., Looijenga, L. H. J., Drop, S. L. S., & Wedell, A. (2011). Partial deletion of the NR5A1 (SF1) gene detected by synthetic probe MLPA in a patient with XY gonadal disorder of sex development. *Sexual Development*, *5*, 181–187.

Berney, D. M., Looijenga, L. H. J., Idrees, M., Oosterhuis, J. W., Rajpert-De Meyts, E., Ulbright, T. M., & Skakkebaek, N. E. (2016). Germ cell neoplasia in situ (GCNIS): Evolution of the current nomenclature for testicular pre-invasive germ cell malignancy. *Histopathology*, *69*, 7–10.

Betts, M. J., & Russell, R. B. (2003). Amino acid properties and consequences of substitutions. *Bioinformatics for geneticists*, *317*, 290–316.

Bleyl, S. B., Moshrefi, A., Shaw, G. M., Saijoh, Y., Schoenwolf, G. C., Pennacchio, L. A., & Slavotinek, A. M. (2007). Candidate genes for congenital diaphragmatic hernia from animal models: Sequencing of FOG2 and PDGFRα reveals rare variants in diaphragmatic hernia patients. *European Journal of Human Genetics*, *15*, 950–958.

Brauner, R., Picard-Dieval, F., Lottmann, H., Rouget, S., Bignon-Topalovic, J., Bashamboo, A., & McElreavey, K. (2016). Familial forms of disorders of sex development may be common if infertility is considered a comorbidity. *BMC Pediatrics*, *16*, 195.

Camats, N., Pandey, A. V., Fernández-Cancio, M., Andaluz, P., Janner, M., Torán, N.,… Flück, C. E. (2012). Ten novel mutations in the NR5A1 gene cause disordered sex development in 46,XY and ovarian insufficiency in 46,XX individuals. *The Journal of Clinical Endocrinology & Metabolism*, *97*, E1294–E1306.

Chávez, B., Ramos, L., Gómez, R., & Vilchis, F. (2014). 46,XY disorder of sexual development resulting from a novel monoallelic mutation (p.Ser31Phe) in the steroid 5 α-reductase type-2 (SRD5A2) gene. *Molecular Genetics and Genomic Medcine*, *2*, 292–296.

Chen, W. Y., Weng, J. H., Huang, C. C., & Chung, B. C. (2007). Histone deacetylase inhibitors reduce steroidogenesis through SCF-mediated ubiquitination and degradation of steroidogenic factor 1 (NR5A1). *Molecular and Cellular Biology*, *27*, 7284–7290.

Ciaccio, M., Costanzo, M., Guercio, G., De Dona, V., Marino, R., Ramirez, P. C.,… Belgorosky, A. (2012). Preserved fertility in a patient with a 46,XY disorder of sex development due to a new heterozygous mutation in the NR5A1/SF-1 gene: Evidence of 46,XY and 46,XX gonadal dysgenesis phenotype variability in multiple members of an affected kindred. *Hormone Research in Paediatrics*, *78*, 119–126.

Cools, M., Hoebeke, P., Wolffenbuttel, K. P., Stoop, H., Hersmus, R., Barbaro, M., … Drop, S. L. S. (2012). Pubertal androgenization and gonadal histology in two 46,XY adolescents with NR5A1 mutations and predominantly female phenotype at birth. *European Journal of Endocrinology*, *166*, 341–349.

Cools, M., Looijenga, L. H. J., Wolffenbuttel, K. P., & T'Sjoen, G. (2014). Managing the risk of germ cell tumourigenesis in disorders of sex development patients. *Endocrine Development*, *27*, 185–196.

Coutant, R., Mallet, D., Lahlou, N., Bouhours-Nouet, N., Guichet, A., Coupris, L., … Morel, Y. (2007). Heterozygous mutation of steroidogenic factor-1 in 46,XY subjects may mimic partial androgen insensitivity syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *92*, 2868–2873.

De Santa Barbara, P., Bonneaud, N., Boizet, B., Desclozeaux, M., Moniot, B., Sudbeck, P., … Berta, P. (1998a). Direct interaction of SRY-related protein SOX9 and steroidogenic factor 1 regulates transcription of the human anti-Müllerian hormone gene. *Molecular and Cellular Biology*, *18*, 6653–6665.

De Santa Barbara P, Moniot B., Poulat, F., Boizet, B., & Berta, P. (1998b). Steroidogenic factor-1 regulates transcription of the human anti-müllerian hormone receptor. *Journal of Biological Chemistry*, *273*, 29654–29660.

Desmet, F-O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M., & Béroud, C. (2009). Human Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Research*, *37*, e67–e67.

Eggers, S., Sadedin, S., van den Bergen, J. A., Robevska, G., Ohnesorg, T., Hewitt, J.,… Sinclair, A. H. (2016). Disorders of sex development: Insights from targeted gene sequencing of a large international patient cohort. *Genome Biology*, *17*, 243.

Eggers, S., Smith, K. R., Bahlo, M., Looijenga, L. H. J., Drop, S. L. S., Juniarto, Z. A., … Sinclair, A. H. (2015). Whole exome sequencing combined with linkage analysis identifies a novel 3 bp deletion in NR5A1. *European Journal of Human Genetics*, *23*, 486–493.

El-Khairi, R., & Achermann, J. C. (2012). Steroidogenic factor-1 and human disease. *Seminars in Reproductive Medicine*, *30*, 374–381.

Fabbri, H. C., de Andrade, J. G. R., Soardi, F. C., de Calais, F. L., Petroli, R. J., Maciel-Guerra, A. T., … de Mello, M. P. (2014). The novel p.Cys65Tyr mutation in NR5A1 gene in three 46,XY siblings with normal testosterone levels and their mother with primary ovarian insufficiency. *BMC Medical Genetics*, *15*, 7.

Fabbri, H. C., Ribeiro de Andrade, J. G., Maciel-Guerra, A. T., Guerra-Júnior, G., & de Mello, M. P. (2016). ***NR5A1*** loss-of-function mutations lead to 46,XY partial gonadal dysgenesis phenotype: Report of three novel mutations. *Sexual Development*, *10*, 191–199.

Ferlin, A., Rocca, M. S., Vinanzi, C., Ghezzi, M., Di Nisio, A., & Foresta, C. (2015). Mutational screening of NR5A1 gene encoding steroidogenic factor 1 in cryptorchidism and male factor infertility and functional analysis of seven undescribed mutations. *Fertility and Sterility*, *104*, 163–169.e1.

Ferraz-de-Souza, B., Lin, L., & Achermann, J. C. (2011). Steroidogenic factor-1 (SF-1, NR5A1) and human disease. *Molecular and Cellular Endocrinology*, *336*, 198–205.

Gunes, S. O., Metin Mahmutoglu, A., & Agarwal, A. (2016). Genetic and epigenetic effects in sex determination. *Birth Defects Research Part C: Embryo Today*, *108*, 321–336.

Hanley, N. A., Ball, S. G., Clement-Jones, M., Hagan, D. M., Strachan, T., Lindsay, S., … Wilson, D. I. (1999). Expression of steroidogenic factor 1 and Wilms' tumour 1 during early human gonadal development and sex determination. *Mechanisms of Development*, *87*, 175–180.

Hersmus, R., van Bever, Y., Wolffenbuttel, K. P., Biermann, K., Cools, M., & Looijenga, L. H. J. (2017). The biology of germ cell tumors in disorders of sex development. *Clinical Genetics*, *91*, 292–301.

Hoivik, E. A., Lewis, A. E., Aumo, L., & Bakke, M. (2010). Molecular aspects of steroidogenic factor 1 (SF-1). *Molecular and Cellular Endocrinology*, *315*, 27–39.

Hussain, S., Amar, A., Najeeb, M. N., & Khaliq, S. (2016). Two novel mutations in the NR5A1 gene as a cause of disorders of sex development in a Pakistani cohort of 46,XY patients. *Andrologia*, *48*, 509–517.

Jacob, A. L., Lund, J., Martinez, P., & Hedin, L. (2001). Acetylation of steroidogenic factor 1 protein regulates its transcriptional activity and recruits the coactivator GCN5. *Journal of Biological Chemistry*, *276*, 37659–37664.

Jones, T. R., Kang, I. H., Wheeler, D. B., Lindquist, R. A., Papallo, A., Sabatini, D. M., … Carpenter, A. E. (2008). CellProfiler Analyst: Data exploration and analysis software for complex image-based screens. *BMC Bioinformatics*, *9*, 482.

Köhler, B., Lin, L., Ferraz-de-Souza, B., Wieacker, P., Heidemann, P., Schröder, V., … Achermann, J. C. (2008). Five novel mutations in steroidogenic factor 1 (SF1, NR5A1) in 46,XY patients with severe underandrogenization but without adrenal insufficiency. *Human Mutation*, *29*, 59–64.

Köhler, B., Lin, L., Mazen, I., Cetindag, C., Biebermann, H., Akkurt, I., … Achermann, J. C. (2009). The spectrum of phenotypes associated with mutations in steroidogenic factor 1 (SF-1, NR5A1, Ad4BP) includes severe penoscrotal hypospadias in 46,XY males without adrenal insufficiency. *European Journal of Endocrinology*, *161*, 237–242.

Lin, L., & Achermann, J. C. (2008). Steroidogenic factor-1 (SF-1, Ad4BP, NR5A1) and disorders of testis development. *Sexual Development*, *2*, 200–209.

Lin, L., Philibert, P., Ferraz-de-Souza, B., Kelberman, D., Homfray, T., Albanese, A., … Achermann, J. C. (2007). Heterozygous missense mutations in steroidogenic factor 1 (SF1/Ad4BP, NR5A1) are associated with 46,XY disorders of sex development with normal adrenal function. *The Journal of Clinical Endocrinology & Metabolism*, *92*, 991–999.

Little, T. H., Zhang, Y., Matulis, C. K., Weck, J., Zhang, Z., Ramachandran, A., … Radhakrishnan, I. (2006). Sequence-specific deoxyribonucleic acid (DNA) recognition by steroidogenic factor 1: A helix at the carboxy terminus of the DNA binding domain is necessary for complex stability. *Molecular Endocrinology*, *20*, 831–843.

Mazen, I., Abdel-Hamid, M., Mekkawy, M., Bignon-Topalovic, J., Boudjenah, R., Gammal El, M., … McElreavey, K. (2016). Identification of ***NR5A1*** mutations and possible digenic inheritance in 46,XY gonadal dysgenesis. *Sexual Development*, *10*, 147–151.

Pedace, L., Laino, L., Preziosi, N., Valentini, M. S., Scommegna, S., Rapone, A. M., … Grammatico, P. (2014). Longitudinal hormonal evaluation in a patient with disorder of sexual development, 46,XY karyotype and one NR5A1mutation. *American Journal of Medical Genetics*, *164*, 2938–2946.

Philibert, P., Polak, M., Colmenares, A., Lortat-Jacob, S., Audran, F., Poulat, F., & Sultan, C. (2011). Predominant Sertoli cell deficiency in a 46,XY disorders of sex development patient with a new NR5A1/SF-1 mutation transmitted by his unaffected father. *Fertility and Sterility*, *95*, 1788.e5–9.

Philibert, P., Zenaty, D., Lin, L., Soskin, S., Audran, F., Léger, J., … Sultan, C. (2007). Mutational analysis of steroidogenic factor 1 (NR5a1) in 24 boys with bilateral anorchia: A French collaborative study. *Human Reproduction*, *22*, 3255–3261.

Piva, F., Giulietti, M., Burini, A. B., & Principato, G. (2012). SpliceAid 2: A database of human splicing factors expression data and RNA target motifs. *Human Mutation*, *33*, 81–85.

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocol*, *5*, 725–738.

Röpke, A., Tewes, A-C., Gromoll, J., Kliesch, S., Wieacker, P., & Tüttelmann, F. (2013). Comprehensive sequence analysis of the NR5A1 gene encoding steroidogenic factor 1 in a large group of infertile males. *European Journal of Human Genetics*, *21*, 1012–1015.

Sadedin, S. P., Dashnow, H., James, P. A., Bahlo, M., Bauer, D. C., Lonie, A., … Thorne, N.P. (2015). Cpipe: A shared variant detection pipeline designed for diagnostic settings. *Genome Medicine*, *7*, 68.

Sasaki, G., Ogata, T., Ishii, T., Kosaki, K., Sato, S., Homma, K., … Matsuo, N. (2003). Micropenis and the 5α-reductase-2 (SRD5A2) gene: Mutation and V89L polymorphism analysis in 81 Japanese patients. *The Journal of Clinical Endocrinology & Metabolism*, *88*, 3431–3436.

Sekido, R., & Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature*, *456*, 824–824.

Shen, W. H., Moore, C. C., Ikeda, Y., Parker, K. L., & Ingraham, H. A. (1994). Nuclear receptor steroidogenic factor 1 regulates the müllerian inhibiting substance gene: A link to the sex determination cascade. *Cell*, *77*, 651–661.

Swartz, J. M., Ciarlo, R., Guo, M. H., Abrha, A., Diamond, D. A., Chan, Y-M., & Hirschhorn, J. N. 2017. Two unrelated undervirilized 46,XY males with inherited ***NR5A1*** variants identified by whole-exome sequencing. *Hormone Research in Paediatrics*, *87*, 264–270.

Venselaar, H., Beek Te, T. A. H., Kuipers, R. K. P., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics*, *11*, 548.

Wada, Y., Okada, M., Fukami, M., Sasagawa, I., & Ogata, T. (2006). Association of cryptorchidism with Gly146Ala polymorphism in the gene for steroidogenic factor-1. *Fertility and Sterility*, *85*, 787–790.

Wada, Y., Okada, M., Hasegawa, T., & Ogata, T. (2005). Association of severe micropenis with Gly146Ala polymorphism in the gene for steroidogenic factor-1. *Endocrine Journal*, *52*, 445–448.

Wildeman, M., van Ophuizen, E., Dunnen den, J. T., & Taschner, P. E. M. (2008). Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. *Human Mutation*, *29*, 6–13.

Wu, J. Y., McGown, I. N., Lin, L., Achermann, J. C., Harris, M., Cowley, D. M., … Cotterill, A. M. (2013). A novel NR5A1variant in an infant with elevated testosterone from an Australasian cohort of 46,XY patients with disorders of sex development. *Clinical Endocrinology*, *78*, 545–550.

Yagi, H., Takagi, M., Kon, M., Igarashi, M., Fukami, M., & Hasegawa, Y. (2015). Fertility preservation in a family with a novel NR5A1 mutation. *Endocrine Journal*, *62*, 289–295.

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, *9*, 40–48.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Robevska G, van den Bergen JA, Ohnesorg T, et al. Functional characterisation of novel *NR5A1* variants reveals multiple complex roles in Disorders of Sex Development. *Human Mutation*. 2017;1–16. https://doi.org/10.1002/humu.23354