Genetic evidence of the association of DEAH-box helicase 37 defects with 46,XY gonadal dysgenesis spectrum

Citation for published version:

Digital Object Identifier (DOI):
10.1210/jc.2019-00984

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published in:
Journal of Clinical Endocrinology & Metabolism

Publisher Rights Statement:
This is a pre-copyedited, author-produced version of an article accepted for publication in [insert journal title] following peer review. The version of record is available online at: 10.1210/jc.2019-00984

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Genetic evidence of the association of DEAH-box helicase 37 defects as a novel cause of 46,XY gonadal dysgenesis spectrum

Thatiana Evilen da Silva PhD\textsuperscript{1*}, Nathalia Lisboa Gomes MD\textsuperscript{1*}, Antonio Marcondes Lerário MD,\textsuperscript{2,3} Catherine Elizabeth Keegan MD,\textsuperscript{4,5} Mirian Yumi Nishi PhD,\textsuperscript{1} Filomena Marino Carvalho MD,\textsuperscript{6} Eric Vilain MD,\textsuperscript{7} Hayk Barseghyanm PhD,\textsuperscript{7} Alejandro Martinez-Aguayo MD,\textsuperscript{8} María Verónica Forclaz MD,\textsuperscript{9} Regina Papazian MD\textsuperscript{9}, Leila Cristina Pedroso de Paula MD\textsuperscript{10}, Eduardo Corrêa Costa MD,\textsuperscript{10} Luciani Renata Carvalho MD,\textsuperscript{1} Alexander A Jorge MD\textsuperscript{1}, Felipe Elias MS\textsuperscript{1}, Rod Mitchell, MBChB, PhD\textsuperscript{11}, Elaine Maria Frade Costa MD,\textsuperscript{1} Berenice Bilharinho Mendonca MD,\textsuperscript{1,2*} Sorahia Domenice MD,\textsuperscript{1*}

\textsuperscript{1}Unidade de Endocrinologia do Desenvolvimento, Laboratório de Hormônios e Genética Molecular (LIM/42) da Disciplina de Endocrinologia e Metabologia do Hospital das Clínicas da Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brasil

\textsuperscript{2}Laboratório de Sequenciamento em Larga Escala (SELA), Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brasil

\textsuperscript{3}Division of Metabolism, Endocrinology and Diabetes, Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA

\textsuperscript{4}Department of Pediatric Genetics, University of Michigan Medical School, Ann Arbor, MI, USA

\textsuperscript{5}Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI, USA

\textsuperscript{6}Departamento de Patologia, Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP, Brasil
7Center for Genetic Medicine Research, Children’s National Health System, Children’s Research Institute, Children’s National Medical Center, Washington, DC, USA
8Division de Pediatria, Escuela de Medicina, Pontificia Universidad Catolica de Chile, Santiago, Chile
9Servicio de Pediatría, Hospital Nacional Prof. Dr. A. Posadas, Buenos Aires, Argentina
10Hospital de Clínicas de Porto Alegre, Programa de Atendimento às Desordens do Desenvolvimento Sexual (PADS), Brasil, UFRGS
11MRC Centre for Reproductive Health, Queens Medical Research Institute, Edinburgh, UK

*First authors da Silva, TE and Gomes, NL and last authors Mendonca, BB and Domenice, S contributed equally to this article.

All authors have nothing to disclose.

Short title: DHX37 defects in 46,XY Gonadal Dysgenesis

**Key words:** DHX37; 46,XY gonadal dysgenesis; embryonic testicular regression syndrome

Corresponding author: Berenice B Mendonca

Hospital das Clínicas, Faculdade de Medicina da Universidade de São Paulo, Disciplina de Endocrinologia e Metabologia.

Av. Dr. Enéas de Carvalho Aguiar, 155, 2º andar, bloco 6

CEP: 05403-900, São Paulo, Brasil.

Email: beremen@usp; sorahiad@gmail.com

Phone number: 55 11 2661-7564; 55 11 2661-7512
Acknowledgements

This work was supported by grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Grant No. 305743/2011-2), the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Grants No. 05/04726-0, 07/512156, 10/51102-0, 2013/02162-8 and 2014/50137-5), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/ PNPD; institucional fellowship Grant). The authors are very grateful to Dr. Frederico Moraes Ferreira for his technical assistance with the in silico prediction analysis, Dr. Beverly M. Yashar and Dr. John Park for medical care of patient F6:II-1, and Dr Ivo Jorge Prado Arnhold for suggestions and significant review of the manuscript.
ABSTRACT

Context: 46,XY gonadal dysgenesis (GD) is a heterogeneous group of disorders with a wide phenotypic spectrum, including embryonic testicular regression syndrome (ETRS). Most patients with GD remain without a molecular diagnosis. Objective: To report a novel gene for 46,XY GD etiology, especially for ETRS. Design: Screening of familial cases of 46,XY GD using whole exome sequencing and sporadic cases by target gene panel sequencing. Setting: Tertiary referral center for Differences/Disorders of sex Development (DSD). Patients and methods: We selected 87 patients with 46,XY DSD (17 familial cases from eight unrelated families and 70 sporadic cases); 55 patients had GD (among them, ten patients from five families and eight sporadic cases had ETRS) and 32 patients had 46,XY DSD of unknown etiology. Results: We identified four heterozygous missense rare variants classified as pathogenic or likely pathogenic in DEAH-box helicase 37 (DHX37) gene in five families (n=11 patients) and in six sporadic cases. Two variants were recurrent: the p.Arg308Gln (in two families and in three sporadic cases) and the p.Arg674Trp (in two families and in two sporadic cases). The variants were specifically associated with ETRS (7/14 index cases; 50%). The frequency of rare, predicted to be deleterious DHX37 variants in this cohort (0.14) is significantly higher than that observed in gnomAD population database (0.004; p<0.001). Immunohistochemistry analysis in human testis showed that DHX37 is mainly expressed in germ cells, at different stages of testis maturation, in Leydig cells and rarely in Sertoli cells. Conclusion: This strong genetic evidence identifies DHX37 as a new player in the complex cascade of male gonadal differentiation and maintenance.
Introduction

46,XY gonadal dysgenesis (GD) represents a heterogeneous group of disorders/differences of sex development (DSD) characterized by abnormal gonadal development leading to a wide phenotypic spectrum. Variable degrees of external genitalia undervirilisation are observed, ranging from micropenis to female-like genitalia and partially- or fully-developed Mullerian derivatives. The gonads from these patients display a wide spectrum of histological abnormalities, ranging from ovarian-like stroma with disorganized seminiferous tubules to complete absence of gonadal tissue (1). Embryonic testicular regression syndrome (ETRS) is considered a part of the clinical spectrum of 46,XY gonadal dysgenesis (2). Most individuals with ETRS present with micropenis or atypical genitalia and lack of gonadal tissue on one or both sides (2). Partial or complete Mullerian duct regression associated with micropenis suggests an intrinsically functional testis in the first months of fetal life subsequent loss of testicular function before the last trimester of gestation, when the increase in penile length occurs. Numerous genes are known to be involved in the process of gonadal determination (3). However, a genetic diagnosis is identified in less than 40% of the patients with 46,XY GD (4). Moreover, few patients with ETRS were included in large cohorts of 46,XY DSD previously studied (4). However, the fact that some familial cases of ETRS were reported indicates a genetic etiology (5,6).

In the present work, high throughput parallel sequencing methods, including whole-exome sequencing (WES) and targeted DSD-gene panels, were used to investigate the underlying genetic etiology in a large cohort of 46,XY patients with GD and 46,XY DSD patients with unknown etiological cause.
We identified recurrent rare variants in DEAH (Asp-Glu-Ala-His) box polypeptide 37 (DHX37) in several affected individuals from distinct families, establishing a novel genetic cause for 46,XY gonadal dysgenesis spectrum, including ETRS.

**Ethics**

This study was approved by the Ethics Committee of the Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, the Institutional Review Board of the University of Michigan Medical School, the Hospital de Garrahan Escuela de Medicina, Pontificia Universidad Católica de Chile, and the Hospital Nacional Prof. Dr. A. Posadas, Buenos Aires, Argentina. Written informed consent was obtained from all patients, their parents or legal guardians.

**Subjects and Methods**

We studied eighty-seven 46,XY DSD patients without previous molecular diagnosis, including 17 familial cases of 46,XY GD from 8 non-consanguineous families and 70 sporadic cases (38 with GD and 32 with 46,XY DSD of unknown etiology). Out of the 55 patients with GD, 10 patients from 5 families and 8 sporadic cases had an ETRS phenotype. The patients had different nationalities: Brazilian (81 patients), Argentinian (three siblings), Chilean (two siblings) and Chinese-American (one patient). All patients have a normal GTG-banded metaphases 46,XY karyotype.

The 46,XY DSD patients were classified as having complete GD (CGD) if they had female external genitalia, Mullerian derivatives and streak gonads; as partial GD (PGD) if they had atypical external genitalia, Mullerian derivatives and at least one gonad with histopathological features of dysgenetic testis; as ETRS if they had micropenis, partially developed Mullerian derivatives and no gonadal tissue or small area of gonadal stroma; and as 46,XY DSD of unknown etiology if hormonal profile was not conclusive or not available due to previous gonadectomy. In this latter group, molecular defects of LHCG
and androgen receptors, CYP17A1, HSD17B3, HSD3B2, and 5ARD2 genes were ruled out by DNA sequencing.

**Genomic DNA**

For molecular diagnosis, genomic DNA was extracted from peripheral blood leukocytes by the proteinase K-SDS salting-out method (7).

**Genetic study**

Whole exome sequencing (WES) was performed in 14 familial cases from 7 families. In all but one family, the probands and their first-degree relatives and other affected family members were studied.

Sixty-eight sporadic cases were studied by targeted massively parallel sequencing. DHX37 was studied by Sanger sequencing in two sporadic cases and in three patients from Family 2 (Figure 1). (Supplementary information, in DOI: 10.13140/RG.2.2.35903.76968).

Enrichment for massively parallel sequencing was performed with Nextera Exome Enrichment Kit (Illumina, San Diego, CA), followed by paired-end sequencing on the Illumina HiSeq 2500 System (Illumina, San Diego, CA).

For target sequencing, we designed an amplicon-based capture panel against exonic regions of 63 genes, including 43 genes known to be associated with human DSDs and 20 candidate genes, including DHX37 (3) (see Table S1, Supplementary information in DOI: 10.13140/RG.2.2.35903.76968). Target sequences were captured using a custom Sure Select Target Enrichment System Kit (Agilent Technologies, Santa Clara, CA, USA) and sequencing was performed on the Illumina MiSeq platform (San Diego, CA, USA).

Sanger sequencing was used to confirm the potentially pathogenic variants identified by massively parallel sequencing and for segregation analysis. Sequencing was performed
on the ABI 3730XL DNA Analyzer (Applied Biosystems) using the BigDye (Applied
Biosystems), followed by data analysis using a Genetic Analyzer (ThermoFisher Scientific).

The identified variants were classified according to American College of Medical Genetics (ACMG) criteria (8).

**Data analysis**

The exome and the targeted panel sequencing data were screened for rare variants (minor allele frequency < 0.1% in the public databases: Genome Aggregation Database (gnomAD) (9), 1000 Genomes (10), and in the Brazilian population database (ABraOM) (11), located in exonic and consensus splice site regions. Subsequently, the filtration pipeline prioritized potentially pathogenic candidate variants (loss of function variants and variants classified as pathogenic by multiple *in silico* programs). For variants identified by WES, we selected variants that fitted an autosomal-dominant model. The sequencing reads carrying candidate variants were visually confirmed using the Integrative Genomics Viewer (Broad Institute, Cambridge, MA). Candidate variants were segregated in the family members by Sanger method. The filtering of the variants is provided in Supplementary data (Figure S1) DOI: 10.13140/RG.2.2.35903.76968

**Histological analysis**

**Immunohistochemical staining**

Eight formalin-fixed paraffin-embedded testicular autopsy samples from 46,XY individuals with different chronological ages (27 and 33 weeks gestational age, 1, 53, and 180 days of age, 13, 23 and 53 years of age) were collected and used for DHX37 expression analysis by immunohistochemistry. All samples were sliced into 3-μm-thick sections using an automatic Leica RM2255 microtome (Leica Biosystems, Nussloch,
Germany). The sections were briefly stretched in xylol at 600°C for 20 min, cooled in xylol, and dried in an incubator (Fanem Orion 515, São Paulo, Brazil) at 600°C. Sections were subjected to hematoxylin-eosin (HE) staining for histological analysis. For immunohistochemical study, slides were deparaffinized with xylene, hydrated in ethanol, washed in phosphate-buffered saline (0.01 M/pH 7.4), and blocked using methanol and hydrogen peroxide. Epitope exposure was carried out by placing the slide in boiling 10 mM citric acid (pH 6) or 100 mM EDTA (pH 9), followed by blocking non-specific protein. Rabbit polyclonal anti-DHX37 antibody (NB110-40581; Novus Biologicals, USA) was added at a dilution factor of 1:50. Dilution was standardized after testing on ovarian and skin tissues where protein expression was identified in cytoplasm of oocytes and nuclear membranes of ovarian stromal and squamous cells. The samples were incubated with universal secondary antibodies using the Novo Link Detection Systems kit (Leica Biosystems, USA) according to the manufacturer’s instruction.

**Statistical analysis**

To test the genetic evidence for the association between DHX37 and GD phenotype, we performed aggregate variant analyses comparing allele frequencies among our 46,XY DSD cohort and public databases [gnomAD and ABraOM]. Variants with similar characteristics of the DHX37 variants observed in our cohort (rare nonsynonymous variants with a minor allele frequency of 0.01 and located in the two highly conserved protein (ATP-binding and Helicase C-terminal domains) that are predicted to be pathogenic by at least four in silico tools (Mutation Taster, SIFT, PolyPhen-2, Mutation Assessor and PROVEAN) were selected. Allele frequency differences between groups were analyzed by \( \chi^2 \) test, and statistical significance was set at \( p<0.05 \). Statistical analyses were performed using SIGMAstat statistical software package (Windows version 3.5; SPSS Inc., San Rafael, CA).
Results

Patient phenotype and DHX37 variants

Firstly, WES identified the same DHX37 variant p.Arg308Gln (c.923G>A) (GenBank: NM_032656.3) in heterozygous state in two unrelated Brazilian families with ETRS (Families 1 and 2). All the affected individuals have the same phenotype (micropenis and absence or bilateral rudimentary gonadal tissue) (Figure 1, Table 1). A founder effect for p.Arg308Gln variant was ruled out in Families 1 and 2.

The p.Arg308Gln variant was also identified by WES in a Chinese-American sporadic case of ETRS from Michigan University performed in Eric Vilain’s laboratory (sporadic case F6:II-1, Figure 1, Table 2).

As a novel candidate gene for 46,XY DSD, DHX37 was included in our target DSD-panel. The same p.Arg308Gln variant was identified in another two sporadic cases: one had ETRS (sporadic case F7:II-1) and the other had PGD (sporadic case F8:II-1) (Figure 1; Table 2).

A further three different heterozygous DHX37 missense variants (the p.Arg674Trp, p.Ser595Phe and p.Thr304Met) were identified in seven affected members from three families and in three sporadic cases (Figure 2).

All of these four variants are predicted to be pathogenic by at least four in-silico prediction tools (Table 3) and are absent in genomic population databases, except for the p.Arg308Gln, which has a very low allele frequency (0.00003) in the gnomAD database (Tables 4-5).

The p.Arg674Trp (c.2020C>T) variant was identified in the two Chilean brothers, both with ETRS (cases F3:II-1 and F3:II-2, Family 3), and also in the three Argentinian affected members (two brothers with ETRS and their uncle with PGD; cases F4:III-1, F4:III-2 and F4:II-4, respectively, Family 4) (Figure 1, Table 1). In addition, the
The p.Arg674Trp variant was also identified in another two Brazilian sporadic cases, one patient with ETRS (sporadic case F10:II-1) and the other with PGD (sporadic case F11:II-1) (Figure 1, Table 2).

The p.Ser595Phe (c.1784C>T) variant was identified in two affected individuals from the same Brazilian family (Family 5). The proband had PGD and her nephew had ETRS (F5:II-6 and F5:III-1, respectively) (Figure 1, Table 1).

The p.Thr304Met (c.911C>T) was identified in a Brazilian female (sporadic case F9:II-5), who had previously undergone bilateral gonadectomy and genitoplasty (Figure 1, Table 2).

The p.Arg308Gln variant is classified as pathogenic and the other three variants, p.Arg674Trp, p.Ser595Phe and p.Thr304Met, are classified as likely pathogenic accordingly the ACMG criteria (Tables 4-5).

**Segregation analysis of DHX37 variants**

Segregation analysis of the DHX37 variants in eight families displayed a sex-limited autosomal dominant pattern, maternally inherited in five families (F2, F3, F4, F5, F11).

In the Family 1, the presence of the p.Arg308Gln variant in the asymptomatic father suggests an autosomal dominant pattern of inheritance with incomplete penetrance (Figure 1). In two sporadic cases (F6. II-1 and F8.II-1), the confirmed paternity displayed a de novo status of the p.Arg308Gln DHX37 variant.

**DHX37 gene and its protein structure**

DHX37 is located in the 12q24.31 region. It is a member of the large DEAH family of proteins and encodes an RNA helicase (12). The DHX37 protein (NP_116045) comprises 1157 amino acids and four main domains. The conserved motifs of the helicase core region contain the Helicase ATP-binding domain (position 262-429) and the Helicase superfamily c-terminal domain (position 585-674); the two other domains are the helicase...
DHX37 protein was identified in different testicular cells

DHX37 expression was characterized in testes from newborns, children and adults using immunohistochemistry. DHX37 was expressed in fibroblasts, endothelial cells and epithelial cells of epididymis. These cells were used as internal positive controls for immunohistochemistry. We found DHX37 expression in Leydig cell cytoplasm and in germ cells at different stages of maturation. Our analysis indicates that DHX37 expression in spermatogonia is characterized by a regular perinuclear halo pattern in both newborns (five samples) and adults (three samples). This pattern of staining differs from that seen in Leydig cells (granular cytoplasmatic) and during other stages of maturation of germ cells. A progressive condensation of protein around the nucleus was observed as cells differentiate from spermatocytes to spermatids, generating a localized paranuclear dot-like pattern. There was no staining in spermatozoa. Rare Sertoli cells displayed a weak and focal cytoplasmatic stain (Figure 3).

Frequency of the DHX37 variants in our 46,XY DSD cohort

The allele frequency of rare and predicted to be deleterious DHX37 variants identified in our cohort of 46,XY DSD patients [11/78 index cases (0.14)] was markedly higher than that observed in individuals from gnomAD [568 /141456 individuals (0.004; p<0.001)] and from a Brazilian cohort [1/609 individuals (0.002); p<0.001].

Discussion

The present study analyzed a large cohort of 46,XY DSD patients without a molecular diagnosis, most of whom had a GD phenotype, including a large number of familial and
sporadic cases with ETRS.

Pathogenic or likely pathogenic allelic variants in the DHX37 were identified in 11 familial cases from 5 unrelated families and in six sporadic cases. Deleterious variants are recurrent in familial and sporadic cases of 46,XY GD in patients of different nationalities.

The DHX37 gene has never been directly associated with gonadal development, but deletions or rearrangements of the 12q24 chromosomal region, which contains DHX37 gene, have been associated with atypical genital development (13). Four syndromic patients with micropenis or perineal hypospadias, and/or hypergonadotropic hypogonadism are reported to have deletions or rearrangements involving the 12q24 region (13-15).

The DHX37 gene encodes a RNA helicase protein which is involved in RNA-related processes, including transcription, splicing, ribosome biogenesis (16), translation and degradation (12,17). DHX37 is required for maturation of the small ribosomal subunit in human cells, through its catalytic activity, required for dissociation of the U3 snoRNA from pre-ribosomal complexes (18). Disturbance of human ribosome production is associated with cancer and genetic diseases known as ribosomopathies (19).

Disease-causing variants in the DExH-box helicase 30 (DHX30), were previously described in syndromic patients with global developmental delay, intellectual disability, severe speech impairment and gait abnormalities. Functional studies of allelic variants in DHX30 demonstrated that they affect protein folding or stability interfering with the RNA binding (mutations located in Motif Ia) or with ATPase activity (mutations located in Motif II and VI) (17,18). Two DHX37 allelic variants found in the present study are located in the same motifs.

Despite lack of experimental evidence to formally demonstrate the deleterious effects of
the four variants identified in the present study, they are located in the highly conserved helicase core region of the DHX37 protein. The spontaneous p.Leu489Pro Dhx37 pathogenic variant was identified in Zebrafish in association with behavior scape defects (20). This study demonstrated that Dhx37 is involved in pre-mRNA splicing reinforcing the role of Dhx37 in RNA-related processes.

Although there is no direct evidence of DHX37 being involved in mRNA processing during gonadal development, DExD/H-box RNA helicase genes are differentially expressed between males and females during the critical period of male sex differentiation in channel catfish (21).

Further, we show population evidence that the DHX37 variants are enriched among the 46,XY GD patients in comparison with the population database. The statistical analysis confirmed that the predicted deleterious DHX37 variants located in the helicase core region are more frequently identified in our 46,XY DSD cohort than in the public databases, emphasizing that this finding was not by chance (p<0.01).

Therefore, in vitro and in vivo studies on DHX37 mechanism have demonstrated a role of DHX37 in ribosome biogenesis (23). Based on this new knowledge, 46,XY gonadal dysgenesis could be classified as a ribosomopathy, expanding the etiological mechanisms of dysgenetic 46,XY DSD spectrum.

Since the discovery of the sex-determining region Y (SRY) variants in patients with GD in 1990 (22), several genes have been associated with the molecular etiology of this disorder. The nuclear receptor subfamily 5 group A member 1 (NR5A1) and Mitogen-Activated Kinase Kinase Kinase 1 (MAP3K1) variants are the most frequent causes of 46,XY gonadal dysgenesis identified to date (23-26).

In this study we found pathogenic/likely pathogenic variants in DHX37 in patients with 46,XY GD at a frequency of 14%, which is slightly higher than the frequency of NR5A1
defects (11%) in our whole cohort (24,27). Considering only the ETRS phenotype (micropenis and absence of uni or bilateral testicular tissue) this frequency increases to 50% (7/14 families).

In the literature, different inheritance patterns have already been described in 46,XY gonadal dysgenesis kindreds (28), including the description of asymptomatic male carriers of proven pathogenic variants of genes involved in testicular determination, such as \textit{SRY} and \textit{NR5A1} genes (29,30). Uncertain mechanisms might prevent the appearance of the phenotype in asymptomatic 46,XY carriers.

Maternal inheritance was observed in all familial cases with pathogenic/likely pathogenic variants in \textit{DHX37} with the exception of family 1, where the variant was inherited from a seemingly unaffected father carrier.

In adult humans, the DHX37 protein is expressed in the ovarian stroma and in the cells within seminiferous tubules (Human Protein Atlas database) (31-33). In our study, the immunohistochemistry analysis of normal testicular tissue from newborn, pubertal and adult males revealed that DHX37 is expressed during specific stages of germ cell maturation, in Leydig cells and rarely in Sertoli cells.

An elaborate paracrine cell-cell network transporting signaling molecules between germ cells and Sertoli cells has been described (34). Indeed, \textit{in vitro} studies have shown that there is a bidirectional trafficking between Sertoli and germ cells, and that each cell type regulates the function of the other (35-38). In addition, RNA expression profiles of DHX37 in human testicular cancer cells are higher than in other tissues (The Human Protein Atlas – Pathology), suggesting that DHX37 may be involved in the regulatory process of the cell proliferation in the testis (31-33).

The present study provides several lines of genetic evidence to indicate that defects in \textit{DHX37} are associated with 46,XY GD spectrum, mainly with ETRS. First, we observed
that the variants segregate with the DSD phenotype in a dominant inheritance pattern in most of the families and that two \textit{de novo} variants were identified. Second, we provide statistical evidence that rare \textit{DHX37} variants are enriched in the analyzed 46,XY DSD cohort in comparison with public databases involving a large number of individuals not selected by this phenotype.

In conclusion, our findings indicate that \textit{DHX37} is a new player in the complex cascade of male gonadal differentiation and maintenance, thus establishing a novel and frequent molecular etiology for 46,XY gonadal dysgenesis spectrum, which includes a high proportion of individuals with embryonic testicular regression syndrome.

\textbf{Supplementary information:} displayed in DOI: 10.13140/RG.2.2.35903.76968).
References


<table>
<thead>
<tr>
<th>Nationality</th>
<th>Brazilian</th>
<th>Brazilian</th>
<th>Chilean</th>
<th>Argentinian</th>
<th>Brazilian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex of rearing</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
</tr>
<tr>
<td>Age at presentation (yrs)</td>
<td>2.2</td>
<td>1.8</td>
<td>14.0</td>
<td>0.6</td>
<td>24 days</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>ETRS</td>
<td>ETRS</td>
<td>ETRS</td>
<td>ETRS</td>
<td>ETRS</td>
</tr>
<tr>
<td>External genitalia</td>
<td>Micropenis</td>
<td>Micropenis</td>
<td>Micropenis</td>
<td>Micropenis</td>
<td>Micropenis</td>
</tr>
<tr>
<td>Gonads</td>
<td>Non-palpable</td>
<td>Non-palpable</td>
<td>Non-palpable</td>
<td>Non-palpable</td>
<td>Non-palpable</td>
</tr>
<tr>
<td>Histologic analysis</td>
<td>Small bilateral dysgenetic gonads</td>
<td>Left gonad not found. Right dysgenetic gonad</td>
<td>No gonadal tissue found</td>
<td>Left gonad not found. Small right dysgenetic gonad</td>
<td>Small bilateral dysgenetic gonads</td>
</tr>
<tr>
<td>Wolffian derivatives</td>
<td>Present</td>
<td>Present</td>
<td>Present</td>
<td>Present</td>
<td>Present</td>
</tr>
<tr>
<td>Mullerian Derivatives</td>
<td>Tubes: Present&lt;sup&gt;y&lt;/sup&gt;</td>
<td>Absent</td>
<td>Present&lt;sup&gt;y&lt;/sup&gt;</td>
<td>Absent</td>
<td>Present&lt;sup&gt;y&lt;/sup&gt;</td>
</tr>
<tr>
<td>Uterus</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
</tr>
<tr>
<td>LH (IU/L)</td>
<td>14.5</td>
<td>12</td>
<td>3.5</td>
<td>1.9</td>
<td>&lt;0.5</td>
</tr>
<tr>
<td>FSH (IU/L)</td>
<td>117</td>
<td>133</td>
<td>87</td>
<td>56</td>
<td>10.9</td>
</tr>
<tr>
<td>Basal Testosterone (ng/dL)</td>
<td>&lt;10</td>
<td>&lt;10</td>
<td>&lt;10</td>
<td>NA</td>
<td>16</td>
</tr>
<tr>
<td>Testosterone after hCG test (ng/dL)</td>
<td>&lt;10</td>
<td>NA</td>
<td>29</td>
<td>26</td>
<td>18</td>
</tr>
<tr>
<td>Variant state</td>
<td>Heterozygous</td>
<td>Heterozygous</td>
<td>Heterozygous</td>
<td>Heterozygous</td>
<td>Heterozygous</td>
</tr>
</tbody>
</table>

NA: not available; PGD: partial gonadal dysgenesis; GCNIS: germ cell neoplasia in-situ; *: testicular biopsy; ¥: Rudimentary Fallopian tubes; Conversion factors to SI units: T, ng/dL to nmol/L, multiply by 0.0347.
Table 2. Phenotype of 46,XY DSD patients with sporadic gonadal dysgenesis spectrum and heterozygous rare pathogenic or likely pathogenic *DHX37* variants

<table>
<thead>
<tr>
<th>Nationality Variables</th>
<th>Chinese-American</th>
<th>Brazilian</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patient</td>
<td>Social sex</td>
</tr>
<tr>
<td></td>
<td>F6:II-1</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>F7:II-1</td>
<td>Male to Female</td>
</tr>
<tr>
<td></td>
<td>F8:II-1</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>F9:II-5</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>F10:II-1</td>
<td>Male to female</td>
</tr>
<tr>
<td></td>
<td>F11:II-1</td>
<td>Female</td>
</tr>
</tbody>
</table>

NA: not available; PGD Partial gonadal dysgenesis; Conversion factors to SI units: T, ng/dL to nmol/L, multiply by 0.0347.
Table 3. *In silico* prediction analysis of *DHX37* allelic variants identified in 46,XY DSD patients

<table>
<thead>
<tr>
<th>Families</th>
<th>Nucleotide changed</th>
<th>AA changed</th>
<th>Functional domain</th>
<th><em>In silico</em> prediction tools</th>
<th>Mutation Taster</th>
<th>Mutation Assessor</th>
<th>SIFT</th>
<th>Polyphen-2</th>
<th>PROVEAN</th>
<th>CAD D</th>
<th>GERP</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1, F2, F6, F7, F8</td>
<td>c.923G&gt;A</td>
<td>p.Arg308Gln</td>
<td>Helicase ATP-binding</td>
<td>Disease Cause (score: 0.999)</td>
<td>High functional impact (score: 4.38)</td>
<td>Protein function affected (score 0.001)</td>
<td>Probably damaging (score 1.000)</td>
<td>Deleterious (score -3.93)</td>
<td>35</td>
<td>5.3</td>
<td></td>
</tr>
<tr>
<td>F3, F4, F10, F11</td>
<td>c.2020C&gt;T</td>
<td>p.Arg674Trp</td>
<td>Helicase superfamily C-terminal domain</td>
<td>Disease Cause (score 1.000)</td>
<td>Middle functional impact (score: 4.83)</td>
<td>Protein function affected (score 0.001)</td>
<td>Probably damaging (score 1.000)</td>
<td>Deleterious (score -7.42)</td>
<td>33</td>
<td>4.2</td>
<td></td>
</tr>
<tr>
<td>F5</td>
<td>c.1784C&gt;T</td>
<td>p.Ser595Phe</td>
<td>Helicase superfamily C-terminal domain</td>
<td>Disease Cause (score 1.000)</td>
<td>High functional impact (score: 4.26)</td>
<td>Protein function affected (score 0.001)</td>
<td>Benign (score 0.24)</td>
<td>Deleterious (score -5.57)</td>
<td>24.4</td>
<td>4.13</td>
<td></td>
</tr>
<tr>
<td>F9</td>
<td>c.911C&gt;T</td>
<td>p.Thr304Met</td>
<td>Helicase ATP-binding</td>
<td>Disease Cause (score 1.000)</td>
<td>High functional impact (score: 4.45)</td>
<td>Protein function affected (score 0.001)</td>
<td>Probably damaging (score 1.000)</td>
<td>Deleterious (score -5.89)</td>
<td>29.8</td>
<td>5.3</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. *DHX37* missense allelic variants identified in 46,XY DSD patients and their frequency in population databases

<table>
<thead>
<tr>
<th>Families</th>
<th>cDNA position</th>
<th>AA change</th>
<th>Phylogenetic Conservation</th>
<th>State</th>
<th>dbSNP</th>
<th>MAFs in population databases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1000 Genomes</td>
</tr>
<tr>
<td>F1, F2,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absent</td>
</tr>
<tr>
<td>F6, F7,</td>
<td>c.923 G&gt;A</td>
<td>p.Arg308Gln</td>
<td>Highly conserved</td>
<td>Heterozygous</td>
<td>Not available</td>
<td>Absent</td>
</tr>
<tr>
<td>F8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absent</td>
</tr>
<tr>
<td>F3, F4,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absent</td>
</tr>
<tr>
<td>F10,</td>
<td>c.2020C&gt;T</td>
<td>p.Arg674Trp</td>
<td>Highly conserved</td>
<td>Heterozygous</td>
<td>Not available</td>
<td>Absent</td>
</tr>
<tr>
<td>F11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absent</td>
</tr>
<tr>
<td>F5</td>
<td>c.1784C&gt;T</td>
<td>p.Ser595Phe</td>
<td>Highly conserved</td>
<td>Heterozygous</td>
<td>Not available</td>
<td>Absent</td>
</tr>
<tr>
<td>F9</td>
<td>c.911C&gt;T</td>
<td>p.Thr304Met</td>
<td>Highly conserved</td>
<td>Heterozygous</td>
<td>Not available</td>
<td>Absent</td>
</tr>
</tbody>
</table>
**Table 5.** Pathogenicity classification of DHX37 variants according to the American College of Medical Genetics and Genomics guidelines

<table>
<thead>
<tr>
<th>Families</th>
<th>Nucleotide changed</th>
<th>AA changed</th>
<th>Population Data</th>
<th>Computational and prediction data</th>
<th>De novo data</th>
<th>Other data</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1, F2, F6, F7, F8</td>
<td>c.923G&gt;A</td>
<td>p.Arg308Gln</td>
<td>PM2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>PP2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PS2&lt;sup&gt;d&lt;/sup&gt;</td>
<td>PM1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Pathogenic</td>
</tr>
<tr>
<td>F3, F4, F10, F11</td>
<td>c.2020C&gt;T</td>
<td>p.Arg674Trp</td>
<td>PM2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>PP2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PP3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>PM1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Likely</td>
</tr>
<tr>
<td>F5</td>
<td>c.1784C&gt;T</td>
<td>p.Ser595Phe</td>
<td>PM2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>PP2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PP3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>PM1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Pathogenic</td>
</tr>
<tr>
<td>F9</td>
<td>c.911C&gt;T</td>
<td>p.Thr304Met</td>
<td>PM2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>PP2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>PP3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>PM1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Likely</td>
</tr>
</tbody>
</table>

PM2: moderate piece of evidence for pathogenicity; PP3: supporting evidence for pathogenicity by computational (*in silico*) data;

PS2: strong support for pathogenicity when the variants are *de novo*; PP4: supporting evidence using phenotype; PM1: pathogenic moderate;

VUS: Variant of Uncertain Significance.

a Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium.
b Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease
c Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.)
d De novo (both maternity and paternity confirmed) in a patient with the disease and no family history.
e Located in a mutational hot spot and/or critical and well-established functional domain without benign variation.
Figure 1. Pedigrees of the eleven families with potential disease-causing DHX37 variants. Filled symbols represent affected individuals. The affected males (46,XY males) are indicated by filled squares and the affected individuals raised as females (46,XY females) are shown by large dark dots within the squares. Symbols with a diagonal line represent deceased individuals. The DHX37 genotype is shown for the individuals whose DNA sample was available; +/− indicates a heterozygous state and −/− indicates a homozygous state for wild-type allele. NA- DNA not available. Paternity and maternity was confirmed in families 6 and 8.
The identified variants are localized within conserved helicase domains of DHX37. Top: Schematic protein structure of DHX37 showing conserved motifs of the helicase core region, the helicase associated domain (HA2) and the oligonucleotide/oligosaccharide-binding-fold. Middle: Nucleotide-interacting motifs (I, II, and VI), nucleic acid-binding motifs (Ia, Ib, and IV), motif V, which binds nucleic acid and interacts with nucleotides, and motif III, which couples ATP hydrolysis to RNA unwinding (N- N terminus; C- C terminus). Bottom: Amino acids within conserved motifs of the helicase core region. The position of the first and last amino acid within each motif is denoted below left and right, respectively. The position of the allelic variants identified in this study are indicated with vertical arrows and shown in bold in the different species sequence.
Figure 3. Immunoexpression patterns of DHX37 in testis tissues. A- Newborn testis showing strongly positive staining in occasional spermatogonia (arrow) among numerous Sertoli cells, some of which show weak cytoplasmic staining (original magnification 100X). B- Seminiferous tubules of a 13 year old boy demonstrating tubules with predominance of Sertoli cells, all of them negative for DHX37. Note some positive stromal cells (arrow) (original magnification 100X). C- Adult testis of a 54 year old man showing positive Leydig (arrow) and germ cell staining. D- Detail of (C) showing the different pattern of stain in different stages of germ cells. Note strong perinuclear halo in spermatogonia, progressive paranuclear condensation in spermatocytes and spermatids, and absence of DHX37 expression in spermatozoa (original magnification ×20X).