Analysis of NR5A1 in 142 patients with premature ovarian insufficiency, diminished ovarian reserve, or unexplained infertility

Sylvie Jaillarda,b,c,⁎, Rajini Sreenivasana, Marion Beaumontb, Gorjana Robevskaa, Christèle Dubourgd, Ingrid M. Knarstona, Linda Akloule, Jocelyn van den Bergena, Sylvie Odentb, Brittany Crofta, Guilhem Jouvef, Sonia R. Grovera,h, Solène Durosg, Céline Pimentelg, Marc-Antoine Belaud-Rotureaub,c,f, Katie L. Ayersa,i, Célia Ravelb,f, Elena J. Tuckera,i,1, Brittany Crofta, Guilhem Jouvef, Sonia R. Grovera,h, Solène Durosg, Céline Pimentelg, Marc-Antoine Belaud-Rotureaub,c,f, Katie L. Ayersa,i, Célia Ravelb,f, Elena J. Tuckera,i,1, Andrew H. Sinclaira,i,1

⁎ Corresponding author at: Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, VIC, 3052, Australia.
E-mail address: sylvie.jaillard@chu-rennes.fr (S. Jaillard).
1 Co-senior authors of this study.

1. Introduction

Ovarian reserve represents the number of available follicles/oocytes within ovaries and this can be assessed by measuring follicle stimulating hormone (FSH) levels, anti-Müllerian hormone (AMH) levels, and/or antral follicle count (AFC) determined by ultrasound [1]. AMH is produced in women by the granulosa cells of developing follicles and represents a marker of ovarian reserve that can be used in assisted reproduction and for prediction of menopausal age, probability of natural conception, assessment of ovarian function after gonadotoxic treatment or diagnosis of polycystic ovarian syndrome [2], although its reliability in these roles is limited by a lack of evidence that it predicts...
implantation, pregnancy or live birth [3,4]. A diminished ovarian reserve (DOR) remains difficult to define but can be diagnosed by testing that shows low AMH (< 0.5-1.1 ng/ml) and/or low AFC (< 5-7) and/or elevated FSH but lower than the post-menopausal level (> 10 UI/L), without menstrual irregularities [5,6]. This condition is mostly the consequence of age; it normally occurs after 40 years of age and is considered abnormal if it occurs before this time (10% of women). DOR differs as a clinical diagnosis from premature ovarian insufficiency (POI), which occurs in 1% of women under 40 and is characterized by high FSH levels > 25 UI/L (post-menopausal level) on two separate occasions, low estradiol levels and irregular or absent menstruation for at least four months before the age of 40 [40].

A key role in ovarian development and function in the mouse has been demonstrated, with inactivation causing infertility and hypoplastic ovaries [12]. According to the Human Gene Variant Database (HMGD), approximately 160 variants in genes such as NR5A1 (Nuclear Receptor subfamily 5, group A, member 1, OMIM 184757) encode SF1 (Steroidogenic Factor-1), a member of the orphan nuclear receptor family of transcription factors involved in reproductive, adrenial and gonadal development in both males and females. NR5A1 expression occurs in early ovarian development and during sex differentiation phases leading to normal ovarian morphogenesis. SF1 protein expression in the ovary (granulosa and theca cells) is associated with follicular development and is important for proper ovarian functioning and regulation of ovarian steroidogenesis [10]. It regulates the expression of reproductive and steroidogenic genes such as STAR, HSD3B2, CYP19A1, CYP17A1, and CYP11A1 [11]. A key role in ovarian development and function in the mouse has been demonstrated, with inactivation causing infertility and hypoplastic ovaries [12]. According to the Human Gene Variant Database (HMGD), approximately 160 variants in NR5A1 have been described, most of which cause a loss of or reduction in function, and are associated with a wide phenotypic spectrum of disorders of sex development (DSD) due to a variable expressivity and incomplete penetrance. The broad range of DSD phenotypes is proposed to be in part caused by an oligogenic mode of inheritance [13]. Missense variants are the most common, with pathogenic variants dispersed throughout the coding sequence, showing no evident clustering. Frame-shift, nonsense, in-frame deletions and splice site variants have also been described [14]. Pathogenic variants in NR5A1 account for 18–20% of 46,XY DSD cases [14] and are also reported in 46,XY fertile men [15]. Variants within this gene have been found in 0.26–8% of sporadic POI patients [11,16] and in a few cases of 46,XX ovotoxicol DSD [17]. According to the HMGD, ClinVar and the literature, 22 variants have been described with POI, to date [18,19].

The aim of the study was to identify and clarify functional activity, and thus involvement, of NR5A1 variants in conditions with ovarian deficiency such as non-syndromic POI and DOR, and contrast this with unexplained infertility without ovarian deficiency.

2. Material and methods

2.1. Patients

142 women with sporadic non-syndromic ovarian deficiency (POI or DOR, which are likely manifestations along a continuum of ovarian deficiency) or unexplained infertility without ovarian deficiency, and without adrenal insufficiency, were included in this study. 127 patients enrolled in an oocyte donation program were recruited from Rennes Hospital, France, for targeted sequencing of NR5A1, including 25 with POI, 85 with DOR and 17 with unexplained infertility, demonstrated after genetic, hormonal and ultrasonographic assessment. The diagnosis of ovarian deficiency was established when a woman aged under 40 years had, for POI, elevated FSH (≥ 25 IU/L), or, for DOR, FSH between 10 and 25 IU/L and/or low AMH (< 1.1 ng/ml) (both measured by immunochemiluminometric assay) and/or low AFC (< 7) after ultrasonographic assessment, according to the thresholds defined by Cohen et al. without age-based cutoff [5]. For unexplained infertility, these values were within the normal reference range, as was the partner’s semenogram, and the couple had experienced repeated failure of in vitro fertilization (IVF) despite these normal parameters. All included cases were 46,XX and negative for FMR1 premutation. Whole-exome sequencing (WES) was performed for 15 Australian and French patients, 14 of whom had POI and one of whom had DOR. Parental DNA was available for only three patients tested by WES (DNA from the mother for two patients, and both parents for one patient). All participants gave their written informed consent for genetic investigations according to the law of their country.

2.2. Detection of NR5A1 variants by sequencing

2.2.1. Sanger sequencing

Samples were analysed by direct sequencing using the ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction kit and the ABI 3100 Genetic Analyser (Applied Biosystems, CA, USA) for primer sequences see Supp. data Table 1).

2.2.2. Whole-exome sequencing WES

DNA from 15 patients underwent WES at the Australian Genome Research Facility (AGRF). Library preparation was performed with Agilent SureSelect Human All Exon V6 (Agilent Technologies, Santa Clara, CA, USA) and sequencing with the NovaSeq™ 6000 Sequencing System (Illumina Inc., San Diego, CA, USA). All WES data were processed using the Cpipe pipeline [20] and deposited into SeqR for analysis (https://seqr.broadinstitute.org/).

NR5A1 variants observed by sequencing were described based on the Genbank NM_004959.4 cDNA reference sequence and NP_004950.2 protein reference sequence, according to the Sequence Variant Nomenclature for the Human Genome (http://varnomen.hgvs.org/). For greater clarity, the variants are reported as c. instead of NM_004959.4:c. and p. instead of NP_004950.2:p and parentheses are not used to describe protein substitution without experimental evidence.

2.3. Functional studies

2.3.1. Mutant NR5A1 expression vectors

Expression plasmids with NR5A1 variants (c.43 G > A / p.Val15Met, c.368 G > C / p.Gly123Ala, c.386 C > T / p.Pro129Leu, c.[368 G > C;386 C > T] / p.[Gly123Ala;Pro129Leu], c.407 C > T / p.Pro136Leu, c.1198 G > A / p.Alu400Thr) 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, using the wild-type (WT) human NR5A1 cDNA
Table 1
Clinical data of the patients with rare NR5A1 variants.

<table>
<thead>
<tr>
<th>Age of OD</th>
<th>Menstruation</th>
<th>Phenotype</th>
<th>Mutation</th>
<th>Age</th>
<th>Duration</th>
<th>FSH IU/l</th>
<th>LH IU/l</th>
<th>Estradiol pg/ml</th>
<th>AMH ng/ml</th>
<th>AFC</th>
<th>AMH/FSH</th>
<th>FMR1</th>
<th>Other</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>Secondary</td>
<td>amenorrhea*</td>
<td>c.43 G &gt; A</td>
<td>28</td>
<td>31</td>
<td>6.2</td>
<td>4.4</td>
<td>26.6</td>
<td>0.4</td>
<td>ND</td>
<td>3</td>
<td>N</td>
<td>N</td>
<td>amenorrhea</td>
</tr>
<tr>
<td>31</td>
<td>Secondary</td>
<td>amenorrhea*</td>
<td>c.1198 G &gt; A</td>
<td>31</td>
<td>35</td>
<td>4.9</td>
<td>2.7</td>
<td>117</td>
<td>1</td>
<td>ND</td>
<td>4</td>
<td>N</td>
<td>N</td>
<td>amenorrhea</td>
</tr>
<tr>
<td>35</td>
<td>Irregular</td>
<td></td>
<td>c.1020 G &gt; A</td>
<td>35</td>
<td>30</td>
<td>11.9</td>
<td>4.6</td>
<td>35</td>
<td>4.9</td>
<td>ND</td>
<td>3</td>
<td>N</td>
<td>N</td>
<td>amenorrhea</td>
</tr>
<tr>
<td>30</td>
<td>Regular</td>
<td></td>
<td>c.368 G &gt; C / c.386 C &gt; T (cis)</td>
<td>30</td>
<td>32</td>
<td>117</td>
<td>35</td>
<td>35</td>
<td>4.9</td>
<td>ND</td>
<td>3</td>
<td>N</td>
<td>N</td>
<td>amenorrhea</td>
</tr>
<tr>
<td>32</td>
<td>Regular</td>
<td></td>
<td>c.407 C &gt; T</td>
<td>32</td>
<td>30</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>4.9</td>
<td>ND</td>
<td>3</td>
<td>N</td>
<td>N</td>
<td>amenorrhea</td>
</tr>
</tbody>
</table>


2.3.4. Protein structural analysis (in silico)

The crystal structures were visualized and compared using SFI DBD (PDB 2DF0) and LBD (PDB 1ZDT) 3D crystal structures with the PyMOL Molecular Graphics System v1.7.6.6 (https://www.pymol.org) for p.Val15Met and p.Ala400Thr. The effects of the missense mutations identified were assessed using the HOPE database (Q13285) (http://www.cmbi.ru.nl/hope/) [21].

3. Results

3.1. Identification of NR5A1 variants in 46,XX patients with ovarian deficiency and/or unexplained infertility without ovarian deficiency

We found 10 NR5A1 exonic variants scattered throughout exons 2–6 (four synonymous in six patients, one common missense in fourteen patients, five rare missense in four patients) in a group of 142 46,XX patients presenting with ovarian deficiency (39 POI, 86 DOR) or unexplained infertility without ovarian deficiency (N = 17). Of these 142 patients, 15 had WES (one variant identified in a singleton) and 127 were Sanger sequenced for NR5A1 (nine variants identified). Additional candidate genes identified by WES are discussed in other manuscripts published [23] or in preparation. Clinical data of the patients bearing the rare variants are summarized in Table 1. Molecular data of all NR5A1 variants are summarized in Table 2.

3.1.1. Four synonymous heterozygous variants

c.1020G > A / p.=(Ala340Ala) was observed in a POI patient (Patient 3) whereas c.225 G > C / p.=(Thr75Thr) and c.594 G > A / p.=(Pro198Pro) were observed in a single patient presenting with DOR (Patient 6). c.375 G > A / p.=(Pro125Pro) with minor allele frequency of 0.02887 (gnomAD https://gnomad.broadinstitute.org/) was more frequent, being observed in four patients (three DOR, one infertile). Variants were analysed with regard to their possible impact on splicing (Human Splicing Finder www.umd.be/HSF/) and were all predicted to cause a potential alteration of splicing (alteration of an exonic splicing enhancer site for c.225 G > C, c.594 G > A, c.375 G > A; activation of an exonic cryptic acceptor site for c.594 G > A, c.1020 G > A). The fact that the rare synonymous
<table>
<thead>
<tr>
<th>Variant description</th>
<th>Chromosomal variant</th>
<th>Exon</th>
<th>Amino acid change</th>
<th>MAF (gnomAD)</th>
<th>In silico prediction (Mutation Taster, SIFT, PolyPhen-2, DANN)</th>
<th>CADD score (Phred)</th>
<th>Nucleotide conservation (UCSC Multiz Alignments of 100 Vertebrates)</th>
<th>Amino-acid conservation</th>
<th>ClinVar</th>
<th>American College of Medical Genetics ACMG</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.43 G &gt; A</td>
<td>g.127265632C &gt; T</td>
<td>2</td>
<td>p.Val15Met</td>
<td>0</td>
<td>MT: disease causing SIFT: damaging PP-2: probably damaging DANN: 0.9965</td>
<td>Phred: 25 (pathogenic)</td>
<td>Moderate (GERP RS: 2.79)</td>
<td>High</td>
<td>46;XY Sex Reversal 3</td>
<td>Pathogenic</td>
</tr>
<tr>
<td>c.396C &gt; T</td>
<td>g.127262853G &gt; A</td>
<td>4</td>
<td>p.Pro129Leu</td>
<td>T: 0.002474</td>
<td>MT: disease causing SIFT: damaging PP-2: benign DANN: 0.9983</td>
<td>Phred: 16.71 (likely pathogenic)</td>
<td>Moderate (GERP RS: 3.82)</td>
<td>High in mammals Low in other species</td>
<td>Prematurity ovarian failure 7</td>
<td>Uncertain significance</td>
</tr>
<tr>
<td>c.407C &gt; T</td>
<td>g.127262832G &gt; A</td>
<td>4</td>
<td>p.Pro136Leu</td>
<td>T: 0.0025252</td>
<td>MT: polymorphism SIFT: tolerated PP-2: benign DANN: 0.9868</td>
<td>Phred: 12.98 (likely benign)</td>
<td>Moderate (GERP RS: 3.55)</td>
<td>High in mammals Low in other species</td>
<td>Unknown</td>
<td>Uncertain significance</td>
</tr>
<tr>
<td>c.43G &gt; C</td>
<td>g.127262802C &gt; A</td>
<td>4</td>
<td>p.Gly146Ala</td>
<td>T: 0.1165</td>
<td>MT: polymorphism SIFT: tolerated PP-2: benign DANN: 0.9678</td>
<td>Phred: 3.668 (benign)</td>
<td>Low (GERP RS: -0.606)</td>
<td>Low</td>
<td>Benign</td>
<td>Benign</td>
</tr>
<tr>
<td>c.1198G &gt; A</td>
<td>g.127245225C &gt; T</td>
<td>7</td>
<td>p.Ala400Thr</td>
<td>T: 0.00043984</td>
<td>MT: disease causing SIFT: tolerated PP-2: probably damaging DANN: 0.9994</td>
<td>Phred: 26 (pathogenic)</td>
<td>High (GERP RS: 5.17)</td>
<td>High</td>
<td>Unknown</td>
<td>Likely pathogenic</td>
</tr>
<tr>
<td>c.225 G &gt; C</td>
<td>g.127265377C &gt; G</td>
<td>3</td>
<td>p.=(Thr75Thr)</td>
<td>T: 0.006440</td>
<td>MT: disease causing SIFT: DANN: 0.8345 HSF: potential alteration of splicing</td>
<td>Phred: 15.47 (likely pathogenic)</td>
<td>Low (GERP RS: -0.911)</td>
<td>High in mammals Low in other species</td>
<td>Benign</td>
<td>Uncertain significance</td>
</tr>
<tr>
<td>c.375 G &gt; A</td>
<td>g.127262864C &gt; T</td>
<td>4</td>
<td>p.=(Pro125Pro)</td>
<td>T: 0.028807</td>
<td>MT: polymorphism SIFT: DANN: 0.7006 HSF: potential alteration of splicing</td>
<td>Phred: 0.527 (benign)</td>
<td>Low (GERP RS: -9.08)</td>
<td>High in mammals Low in other species</td>
<td>Benign</td>
<td>Uncertain significance</td>
</tr>
<tr>
<td>c.1020 G &gt; A</td>
<td>g.127253478C &gt; T</td>
<td>6</td>
<td>p.=(Ala340Ala)</td>
<td>T: 0.0002263</td>
<td>MT: disease causing SIFT: DANN: 0.772 HSF: potential alteration of splicing</td>
<td>Phred: 5.965 (benign)</td>
<td>Low (GERP RS: -10.19)</td>
<td>High</td>
<td>Unknown</td>
<td>Uncertain significance</td>
</tr>
</tbody>
</table>

AP: African population, EP: European population, MT: Mutation Taster, PP-2: PolyPhen-2, DANN: Deleterious Annotation of genetic variants using Neural Networks (0–1, 1 = most damaging), CADD: Combined Annotation-Dependent Depletion (> 30: highly pathogenic, 20–30: pathogenic, 15–20: likely pathogenic, 10–15: likely benign, < 10: benign), GERP RS: Genomic Evolutionary Rate Profiling Rejected Substitutions (low conservation: -12.3 to high conservation: 6.17), HSF: Human splicing Finder. DNA variant numbering is based on GenBank reference DNA sequence NM_004959.4, with the A of the ATG initiation codon designated +1. Chromosomal variants are described according to the nucleotide sequence NC_000009.11. Predicted protein annotations are based on NP_004950.2.
variants are predicted to impact splicing means their potential for pathogenicity cannot be excluded, although further work would be required to validate their involvement.

### 3.1.2. Six missense variants

The well-described c.437 G > C / p.Gly146Ala polymorphism [24], which has a minor allele frequency of 0.1165 (gnomAD), was identified in fourteen patients (9.6%). It was homozygous in five patients (one POI, four DOR) and heterozygous in nine patients (one POI, six DOR, two infertile).

Of the remaining five missense variants, which were all heterozygous, four have been previously described in the literature, and one is novel. The c.437 G > A / p.Val15Met variant has previously been reported in association with 46,XY sex reversal, type 3 [25], but has not previously been described in POI (Patient 1). The novel c.1198 G > A / p.Ala400Thr was similarly identified in a POI patient (Patient 2).

Two variants, c.368 G > C / p.Gly123Ala and c.386C > T / p.Pro129Leu, have previously been found in cis in individuals with premature ovarian failure 7 (POF 7 OMIM #612,964) [11] and spermatogenic failure 8 (SPGF8 OMIM #613,957) [15], and were observed on the same allele in one DOR patient (Patient 4). The last exonic variant c.407C > T / p.Pro136Leu, previously reported in POI patients but without functional studies [26], was observed in a DOR patient (Patient 5).

The detection rate of rare non-synonymous protein-altering variants in women with POI was 5.1% and in women with DOR was 2.3%. The overall detection rate in patients with ovarian deficiency was 2.8%. No potential causative variant was observed in the infertile women.

### 3.2. Location of the non-synonymous variants with respect to SF1 protein structure

The human SF1 protein spans 461 amino acids and contains an amino terminal DNA binding domain (DBD) with two zinc finger motifs (ZnI and ZnII), a P-box (proximal box), and a T-box (terminal box), a Ftz-F1/A-box (accessory box) which aids in DNA binding and contains the nuclear localization signal (NLS), a carboxy terminal ligand binding domain (LBD) which interacts with various cofactors, and a hinge region crucial for stabilizing the interactions of the LBD. Two activation domains (AF1 and AF2) are also present, with AF2 important for transcriptional activity and interaction with cofactors. Variants identified in our study are mainly observed in the hinge region (p.Gly123Ala, p.Pro129Leu, p.Pro136Leu), one is in the LBD (p.Ala400Thr), and one in the DBD (p.Val15Met). These variants and those previously observed in POI are shown in Fig. 1.

### 3.3. In vitro functional studies of the non-synonymous protein-altering variants

To assess the activity of the protein-altering NR5A1 variants, we used a well-established Dual-Luciferase reporter assay. WT and variant human SF1 activity was assessed using various constructs, including the mouse Tesco-Sox9 (mTesco) reporter, which responds strongly to SF1 in vitro, and constructs with promoters for human enzymes involved in the steroidogenic pathway. STAR (Steroidogenic Acute Regulatory Protein), CYP11A1 (Cholesterol Desmolase), both involved in the conversion of cholesterol to pregnenolone, CYP17A1 (Steroid 17-Alpha-Hydroxylase/17,20 Lyase converting the pregnenolone and progesterone to their 17-alpha-hydroxylated products and subsequently to dehydroepiandrosterone (DHEA) and androstenedione), HSD3B2 (Hydroxy-Delta-5-Steroid Dehydrogenase, 3 Beta- And Steroid Delta-isomerase 2 converting DHEA to androstenedione and androstenediol to testosterone) were tested. For p.Val15Met, a significant impairment of the transcriptional activity on all promoters tested was evident. In contrast, we did not observe any defect in the transcriptional activity of SF1 harbouring the variants in the LBD and the hinge region. Moreover, for some SF1-responsive elements (mTesco, hCYP11A1, hCYP17A1), a significant increase in transcriptional activity was observed, with similar results previously observed for variants located in these regions [24] (Fig. 2). We also assessed whether the mutant SF1 proteins were correctly translocated into the nucleus. Using immunofluorescence, we confirmed that the variant SF1 proteins all showed uniform localization in the nucleus except for p.Val15Met, which showed an atypical staining pattern as previously described [27] (Fig. 3).

### 3.4. In silico prediction for non-synonymous protein-altering variants

For p.Gly123Ala, p.Pro129Leu, and p.Pro136Leu, the mutant residue is larger, potentially leading to disturbance of the local protein structure. The flexibility of glycine is necessary to make torsion angles and could be disrupted by the substitution with alanine. Proline residues are known to form a rigid structure, which could be altered by substitution with leucine. Modelling of the SF1 DBD bound to the inhibin α-subunit promoter showed likely disruption of DNA binding by the p.Val15Met variant, due to the close proximity of Val15 to the Zn finger. Impaired DNA binding due to this variant has been previously shown by electrophoretic mobility shift assay [25]. For p.Ala400Thr, modelling of the SF1 LBD bound to the NCOA2 co-factor and a phospholipid ligand showed that Ala400 is located far from the ligand binding pocket and the co-factor interaction domain, but the substitution of hydrophobic Ala400 with polar/hydrophilic Thr may affect protein folding and conformation, and abolish the function of the domain (Fig. 4).

### 4. Discussion

#### 4.1. Prevalence of NR5A1 variants in patients with ovarian deficiency

Previous studies have reported a discrepancy in the frequency of NR5A1 variants in sporadic or familial cases of 46,XX POI (frequency varying from 0.26%–8%). While this may be explained, at least partially, by differences between populations [11,16,22,26], we set out in this study to further examine the frequency of NR5A1 variants in women with POI. We also aimed to study the involvement of NR5A1 variants in POI/DOR patients. Variants reported in the literature in POI patients are unbold, variants observed in our study are bold. Star: known variant previously associated with POI, Arrow: known variant not previously associated with POI, Lightning: novel variant. N: Amino terminal domain, DBD: DNA binding domain Zn: Zinc finger motifs, Ftz-F1: Fushi Tarazu factor 1 region, NLS: Nuclear Localization Signal, Pro-rich: Proline-rich region, AF: Activation domains, LBD: Ligand Binding Domain, C: Carboxy terminal domain.

![Fig. 1. NR5A1 variants identified in POI/DOR patients. Variants reported in the literature in POI patients are unbold, variants observed in our study are bold. Star: known variant previously associated with POI, Arrow: known variant not previously associated with POI, Lightning: novel variant. N: Amino terminal domain, DBD: DNA binding domain Zn: Zinc finger motifs, Ftz-F1: Fushi Tarazu factor 1 region, NLS: Nuclear Localization Signal, Pro-rich: Proline-rich region, AF: Activation domains, LBD: Ligand Binding Domain, C: Carboxy terminal domain.](image-url)
The trans-activational activity of each variant compared with wild-type (WT) SF1 was tested using a dual-Luciferase reporter assay in HEK293 T cells. Luciferase reporter constructs used were pGL4-mTesco, pGL4.10-hCYP11A1, pGL4.10-hStAR, pGL3-hCYP17A1, pGL3-HSD3B2. For p.Val15Met, a significant impairment of the trans-activational activity on all promoters tested was evident (four asterisks - p < 0.0001). For the other variants, no reduction in trans-activational activity was observed and was even elevated for some SF1-responsive elements (mTesco, hCYP11A1, hCYP17A1 indicated by one asterik - p < 0.05, two asteriks - p < 0.01 or three asterisks - p < 0.001).

Wild-type SF1 showed strong nuclear staining with nucleolar exclusions. SF1 protein variants all showed uniform localization in the nucleus except for p.Val15Met, which showed an atypical staining pattern with SF1-immunoreactive aggregates within sub-nuclear foci (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).

Fig. 2. Trans-activational activity of the protein-altering NR5A1 variants. The trans-activational activity of each variant compared with wild-type (WT) SF1 was tested using a dual-Luciferase reporter assay in HEK293 T cells. Luciferase reporter constructs used were pGL4-mTesco, pGL4.10-hCYP11A1, pGL4.10-hStAR, pGL3-hCYP17A1, pGL3-HSD3B2. For p.Val15Met, a significant impairment of the trans-activational activity on all promoters tested was evident (four asterisks - p < 0.0001). For the other variants, no reduction in trans-activational activity was observed and was even elevated for some SF1-responsive elements (mTesco, hCYP11A1, hCYP17A1 indicated by one asterik - p < 0.05, two asteriks - p < 0.01 or three asterisks - p < 0.001).

Fig. 3. Sub-cellular localization of the protein-altering NR5A1 variants. Protein expression of each variant and wild-type (WT) SF1 was assessed in HEK293 T cells with an SF1 antibody (anti SF1 / green). Nuclear counterstaining was performed with DAPI (blue) and the cytoskeleton was stained with actin (anti-actin / red). Wild-type SF1 showed strong nuclear staining with nucleolar exclusions. SF1 protein variants all showed uniform localization in the nucleus except for p.Val15Met, which showed an atypical staining pattern with SF1-immunoreactive aggregates within sub-nuclear foci (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.).
variants in DOR, because less data are available regarding the genetic basis of this condition. We found five rare non-synonymous protein-altering heterozygous missense variants located across the protein, leading to a detection rate of 5.1 % in POI and 2.3 % in DOR (overall detection rate of 2.8 % in patients with ovarian deficiency). This frequency of NR5A1 variants in sporadic POI is between the extreme values previously reported and is consistent with an intermediate overall prevalence [16,22,26]. We also identified rare synonymous variants which potentially impact transcript splicing, suggesting the failure to diagnose the genetic cause of POI may well be because synonymous variants in established POI genes are overlooked.

4.2. The p.Val15Met NR5A1 variant is responsible for sporadic POI

Due to the variable reported prevalence of NR5A1 variants in POI populations, as well as the failure to detect a functional consequence of some POI-related NR5A1 variants, the involvement of NR5A1 in POI pathogenesis has been controversial, with some authors suggesting that it does not play a major role in the pathogenesis of the condition [22]. We therefore tried to clarify the functional consequences of NR5A1 variants. We first demonstrated a clear and consistent functional impairment of SF1 harbouring the p.Val15Met variant that lies within the DBD. This is the first report of this pathogenic variant in 46,XX sporadic POI, having previously been associated only with 46,XY sex reversal [25]. It is known that the two DSD conditions, 46,XY DSD and 46,XX POI, can be caused by the same variant, as observed in some families, due to the different roles of SF1 in the developing testis and ovary [28]. Our data demonstrate the clear involvement of p.Val15Met in POI pathogenicity. We suggest in this case that POI associated with evidently deleterious variants would have a more severe presentation with early and/or primary amenorrhea (likely hidden in patient 1 by oral contraceptive use).

4.3. Discrepancy between in silico predictions and functional assays

Based on in silico predictions of pathogenicity, and an extremely low minor allele frequency in databases of healthy controls, the novel p.Ala400Thr detected in a POI patient and located in the LBD, was a particularly strong candidate. Despite the strong in silico indication for the involvement of this variant, we failed to detect any major changes in activity, which suggests that it may not be pathogenic. Previously reported variants in the LBD have been shown to have variable functional consequences based on their location and their effect on ligand specificity and recognition [24]. At least nine other variants within the LBD of NR5A1 have been described in 46,XX POI, with a variable impact on transcriptional activity (no or moderate impact for p.Pro235Leu and p.Asp293Asn, and severe impact for p.Leu231_Leu233 and p.Lys372del), depending on the structural importance of the affected amino acid [28]. The p.Ala400Thr variant was detected by WES, and analysis of the data reveals candidate variants in other genes that may be involved in the development of POI [23].

We also identified rare synonymous variants which potentially impact transcript splicing based on in silico analysis. The failure to diagnose the genetic cause of POI may well be because synonymous variants in established POI genes are overlooked; however, the discrepancy between in silico predictions and functional reality means splicing assays would be required to establish the involvement of these rare synonymous variants.

4.4. Recurrence of an NR5A1 variant in African POI patients

The p.Gly123Ala and p.Pro129Leu variants have been previously detected in five patients with a reproductive phenotype (two with ovarian deficiency and three with spermatogenic failure). Interestingly, all reported patients, including ours, are of African origin, raising the possibility of a founder event [15,22]. Importantly, the frequency of the p.Gly123Ala - p.Pro129Leu variants is low in healthy controls (0.0002782 and 0.0002754 respectively in gnomAD), but is higher in the African healthy population (0.0002496 and 0.0002474 respectively in gnomAD, considering more than 20 000 African individuals). They were not found in 388 control men from West and North Africa in the study of Bashamboo et al., which is in keeping with the global African population frequency of ~3/1000. The frequency may then be considered too high for the variants to cause POI in isolation in a fully penetrant autosomal dominant manner. Reporter assays investigating the impact of the p.Pro129Leu variant have been inconsistent, while for p.Gly123Ala the absence of functional impact has been consistently observed [11,22]. We performed a more in-depth analysis, using a myriad of promoters, with each variant alone as well as in combination on one allele as found in the patients. Our study did not reveal any defect in transcriptional activation, adding further evidence that these variants are either benign or impart a subtle effect on SF1 function.

4.5. The discovery of NR5A1 variants with no detectable functional impact in patients with DOR

We report, for the first time, NR5A1 variants in DOR (detection rate of 2.3 % in DOR). DOR and POI differ in their clinical and biological definition but are likely manifestations along a continuum of ovarian deficiency, suggesting some common underlying genetic connection.
Our results further demonstrate the potential for a genetic continuum underpinning the ovarian deficiency continuum, with the same genes involved (e.g. NR5A1, GDF9) but milder variants potentially contributing to a milder phenotype condition, as discussed below. In keeping with NR5A1 variants having a role in ovarian dysfunction, no rare non-synonymous variants were detected in our small cohort of women with infertility but normal ovarian function. A much larger cohort needs to be studied to establish the significance of this finding. Further investigation is also required to identify any mild impact of the NR5A1 variants detected in POI/DOR patients on NR5A1 function, which would strengthen the evidence for an integral role of NR5A1 in normal ovarian function.

4.6. The potential impact of NR5A1 variants with no detectable functional consequence

Our results indicate that the p.Ala400Thr, p.Gly123Ala, p.Pro129Leu, p.Pro136Leu NR5A1 variants are not pathogenic, unlikely to explain the phenotype in isolation or may have a mild effect on protein function that escaped detection by our assays. The latter hypothesis could explain a potential association with relatively mild clinical conditions [22] such as DOR or secondary amenorrhea. We indeed observed NR5A1 missense variants in DOR with a frequency of 2.3 %, consisting of variants with no detectable consequence, reinforcing their potential association with milder clinical conditions. In contrast, more severe phenotypes (gonadal dysgenesis, ambiguous genitalia, with or without adrenal failure, primary amenorrhea) are likely observed in association with more severe NR5A1 variants that have a proven functional consequence [14]. The pathogenicity of the rare variants with apparent preservation of transcriptional activity cannot be ruled out, due to the potential interaction of SF1 with other cofactors in the real ovarian environment, which could modify the transcriptional capacity [22]. There is also a possibility that SF1 variants cause dysregulation of target genes not tested here. The NR5A1 variants could alter protein folding or translation, impacting the function of SF1 in other processes, such as its binding to cofactors. Furthermore, environmental or genetic modifying factors may contribute to SF1-related patient phenotypes. The hypothesis of oligogenic inheritance has been recently discussed [13,14], and is supported by our study, where the patient with the p.Ala400Thr also bears other variants in interesting candidate genes that need further investigation [23]. Such inheritance is difficult to demonstrate conclusively and may rely on further studies of large POI cohorts to determine the frequency and consequence of NR5A1 variants, as well as their co-inheritance with variants in other genes. As such, a whole-exome sequencing approach is more informative than candidate gene sequencing. For this reason, our most recently recruited patients have undergone this method of analysis. It would be informative to perform WES on the entire cohort to gain further insight into oligogenic inheritance, and to investigate the contribution of variants in genes other than NR5A1; however, at the time of patient enrolment, the cost of WES in such a large cohort was prohibitive. Despite the shortcomings of candidate gene sequencing, this in-depth study of the involvement of NR5A1 in ovarian deficiency has provided a number of important insights, such as 1) the frequency of NR5A1 variants in POI is in keeping with prior studies, 2) the p.Val15Met variant is a cause of sporadic POI, 3) in silico predictions can contribute to incorrect variant curation, 4) a recurrent allele in African women with POI has no detectable functional impact, 5) women with DOR can have rare non-synonymous NR5A1 variants, albeit with no detectable functional consequence and 6) although further functional investigations are required, we raise the possibility of NR5A1 variants having a subtle impact on protein function and causing mild ovarian deficiency.

5. Conclusions

In conclusion, our study investigates the prevalence of NR5A1 variants in a cohort of 142 women with POI, DOR or unexplained infertility without ovarian deficiency, and provides in-depth analysis of the functional consequence of these variants. We identified rare non-synonymous protein-altering variants in 2.8 % of women with ovarian deficiency (POI or DOR), and demonstrate for the first time the association of a known deleterious variant (p.Val15Met) with sporadic POI. The remaining four rare non-synonymous variants, some of which have repeatedly been found in POI patients, had no detectable loss of transcription and had normal localization. Their recurrence in POI and/or DOR patients indicates they may subtly impair protein function, leading to a milder clinical presentation or conferring risk for ovarian deficiency. Further studies are required to confirm the involvement of these variants in ovarian deficiency, which may aid preservation of fertility for young women carrying these variants.

Contributors

Sylvie Jaillard conceived and designed the study, contributed to data acquisition, data analysis and interpretation, and drafted the manuscript. Rajini Sreenivasan contributed to data analysis and interpretation. Marion Beaumont contributed to data acquisition, and critically reviewed the manuscript. Gorjana Robevska provided technical support, and critically reviewed the manuscript. Christèle Dubourg contributed to data acquisition. Ingrid M Knarston provided technical support. Linda Akioul was involved in patient care and evaluation. Jocelyn van den Bergen provided technical support. Sylvie Odent was involved in patient care and evaluation, and critically reviewed the manuscript. Britanny Croft provided technical support. Guilhem Jouve was involved in patient care and evaluation. Sonia R Grover was involved in patient care and evaluation, and critically reviewed the manuscript. Solène Duros was involved in patient care and evaluation. Céline Pimentel was involved in patient care and evaluation. Marc-Antoine Belaud-Rotureau participated in project supervision. Katie L Ayers contributed to data analysis and interpretation, and critically reviewed the manuscript. Célia Ravel conceived and designed the study, was involved in patient care and evaluation, and critically reviewed the manuscript. Elena J Tucker conceived and designed the study, contributed to data acquisition, data analysis and interpretation, and drafted the manuscript. Andrew H Sinclair participated in the project supervision, and critically reviewed the manuscript. Jocelyn van den Bergen provided technical support. Ingrid M Knarston provided technical support. Rajini Sreenivasan contributed to data analysis and interpretation. All authors saw and approved the final manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

This work was supported by CHU Rennes and Rennes 1 University, Faculty of Medicine are from France. Peter Doherty Early Career Fellowship, National Health and Medical Research Council program, and Victorian Government’s Operational Infrastructure Support Program are from Australia.
Ethical approval

All participants gave their written informed consent for genetic investigations according to the law of their country (France and Australia).

All procedures were in accordance with the ethical standards of the Ethics Committee of Rennes University Hospital and the French law.

All procedures were in accordance with the ethical standards of the Human Research Ethics Committee of the Royal Children’s Hospital, Melbourne.

Research data (data sharing and collaboration)

There are no linked research data sets for this paper. Data will be made available on request.

Provenance and peer review

This article has undergone peer review.

Acknowledgment

We would like to thank all the patients and families for taking part in our research.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.maturitas.2019.10.011.

References


