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

Sylvie Jaillard, R. Sreenivasan, M. Beaumont, G. Robevska, C. Dubourg, I.M. Knarston, L. Akloul, J. van den Bergen, S. Odent, B. Croft, et al.

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

Sylvie Jaillard[1,2,3*],Rajini Sreenivasan[1], Marion Beaumont[3], Gorjana Robevska[1], Christèle Dubourg[4], Ingrid M Knarston[1], Linda Akloul[5], Jocelyn van den Bergen[1], Sylvie Odent[5], Brittany Croft[1], Guilhem Jouve[6], Sonia R Grover[1,8], Solène Duros[7], Céline Pimentel[7], Marc-Antoine Belaud-Rotureau[2,3,6], Katie L Ayers[1,9], Célia Ravel[2,6], Elena J Tucker[1,9#], Andrew H Sinclair[1,9#]

1- Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, VIC Australia, 3052
2- Univ Rennes, CHU Rennes, INSERM, EHESP, IRSET (Institut de recherche en santé, environnement et travail) – UMR_S 1085, F-35000 Rennes, France
3- CHU Rennes, Service de Cytogénétique et Biologie Cellulaire, F-35033 Rennes, France
4- CHU Rennes, Service de Génétique Moléculaire, F-35033 Rennes, France
5- CHU Rennes, Service de Génétique Clinique, CLAD Ouest, F-35033 Rennes, France
6- CHU Rennes, Service de Biologie de la Reproduction, F-35033 Rennes, France
7- CHU Rennes, Département de Gynécologie Obstétrique et Reproduction Humaine, F-35033 Rennes, France
8- Department of Paediatric and Adolescent Gynaecology, Royal Children's Hospital, Melbourne, VIC Australia, 3052
9- Department of Paediatrics, University of Melbourne, Melbourne, VIC Australia, 3052

*Corresponding author: sylvie.jaillard@chu-rennes.fr

[#] Co-senior authors of this study

Abstract

Ovarian deficiency, including diminished ovarian reserve and premature ovarian insufficiency, represents one of the main causes of female infertility. Little is known of the genetic basis of diminished ovarian reserve, while premature ovarian insufficiency often has a genetic basis, with genes affecting various processes. *NR5A1* is a key gene required for gonadal function, and variants are associated with a wide phenotypic spectrum of disorders of sexual development, and are found in 0.26-8% of patients with premature ovarian insufficiency. As there is some debate about the extent of involvement of *NR5A1* in the pathogenesis of ovarian deficiency, we performed an in-depth analysis of *NR5A1* variants detected in a cohort of 142 patients with premature ovarian insufficiency, diminished ovarian reserve, or unexplained infertility associated with normal ovarian function. We identified rare non-synonymous protein-altering variants in 2.8% of women with ovarian deficiency and no such variants in our small cohort of women with infertility but normal ovarian function. We observed previously reported variants associated with premature ovarian insufficiency in patients with diminished ovarian reserve, highlighting a genetic relationship between these conditions. We confirmed functional impairment resulting from a p.Val15Met variant, detected for the first time in a patient with premature ovarian insufficiency. The remaining variants were associated with preserved transcriptional activity and localization of *NR5A1*, indicating that rare *NR5A1* variants may be incorrectly curated if functional studies are not undertaken, and/or that *NR5A1* variants may have only a subtle impact on protein function and/or confer risk of ovarian deficiency via oligogenic inheritance.

Keywords

Premature ovarian insufficiency; diminished ovarian reserve; *NR5A1*; whole-exome sequencing

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 it 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 (European Society of Human Reproduction and Embryology https://www.eshre.eu/). There is no evidence that DOR is a precursor of POI and women under 40 can have a DOR without POI. Nevertheless, both conditions are likely manifestations along a continuum of ovarian deficiency. This loss of ovarian function represents one of the main causes of female infertility. POI often has a genetic basis, with many genes causing POI in a monogenic manner, but it is likely that the reproductive phenotype can also result from interactions of many genes in a polygenic manner. These genes can affect various processes, such as gonadal development, meiosis, DNA repair, folliculogenesis, steroidogenesis, apoptosis and mitochondrial function. Although our understanding of the genetic basis of POI has grown, particularly due the implementation of whole-exome sequencing, the majority of patients remain without a genetic diagnosis [7]. Genes associated with DOR, like *GDF9*, have also been reported [8,9] but the genetic basis of DOR has not been explored in depth. An early molecular diagnosis of

these conditions could lead to the planning of pregnancies, the cryopreservation of eggs, the management of comorbidities in patients, and accurate genetic counselling for their families. Furthermore, the understanding of the molecular pathways involved in ovarian development and function could highlight novel targets for treatment. Due to the advances in fertility management and the possibility of preservation of female fertility, the identification of causative genetic variants is now a major focus.

NR5A1 (Nuclear Receptor subfamily 5, group A, member 1, OMIM 184757) encodes SF1 (Steroidogenic Factor-1), a member of the orphan nuclear receptor family of transcription factors involved in reproduction, steroidogenesis, adrenal 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 infertile men [15]. Variants within this gene have been found in 0.26% to 8% of sporadic POI patients [11,16] and in a few cases of 46,XX ovotesticular DSD [17]. According to the HMGD, ClinVar and the literature, 22 variants have been associated 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 seminogram, 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.43G>A / p.Val15Met, c.368G>C / p.Gly123Ala, c.386C>T / p.Pro129Leu, c.[368G>C;386C>T] / p.[Gly123Ala;Pro129Leu], c.407C>T / p.Pro136Leu, c.1198G>A / p.Ala400Thr) 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 in a PCMV6 expression vector as a template pCMV6-Entry-*hNR5A1* (RC207577, OriGene Technologies Inc., Rockville, MD) (for primer sequences see Supp. data Table 2). The coding sequence of all mutant plasmids was confirmed by Sanger sequencing using vector primers prior to functional studies. The previously characterized c.43G>A / p.Val15Met variant acted as a positive control.

2.3.2. Transcriptional activity

The transcriptional activity of the protein-altering *NR5A1* variants was evaluated with Dual-Luciferase reporter assays (Dual-Luciferase Reporter 1000 Assay System Kit, Promega, Fitchburg, WI, USA). The *NR5A1* expression plasmids (WT and variants, 30 ng/well) were transfected into HEK293T cells with the following luciferase reporter constructs: pGL4-mTesco, pGL4.10-hCYP11A1, pGL4.10-hStAR, pGL3-hCYP17A1, pGL3-HSD3β2 (80 ng/well)

and a *Renilla* control plasmid (p*RL*-TK, 10 ng/well) corresponding to a marker for transfection efficiency. Cells were harvested 24 hours post-transfection and luciferase activity was measured on an Infinite M200 Pro plate reader (Tecan, Männedorf, Zürich, Switzerland). The analysis was performed by calculation of the relative fold change, compared with WT (100% activity) for three technical replicates. These fold changes were averaged across six to thirteen biological replicates (mTesco N=11, STAR N= 6, CYP11A1 N=10, CYP17A1 N=11, HSD3β2 N=13), and standard error of the mean (SEM) was calculated. The ANOVA test was used for the statistical analysis (GraphPad Prism Software).

2.3.3. Sub-cellular localization of SF1

HEK293T cells were seeded on eight-well chamber slides (Lab-Tek® Chamber Slides, Sigma-Aldrich, St Louis, MO, USA), and transfected with the *NR5A1* expression plasmids (WT and variants) using Lipofectamine 2000 Transfection Reagent (Invitrogen, Carlsbad, CA, USA). Twenty-four hours after transfection, cells were washed with PBS, then fixed with Paraformaldehyde 4%, permeabilized with Triton 1% and blocked using BSA 2% in PBS. Cells were incubated overnight with a primary monoclonal SF1 (A-1) mouse antibody (1:200; Santa Cruz Biotechnology, Dallas, TX, USA) in 1% BSA. Cytoskeleton was stained with a primary actin antibody (1:200; Sigma-Aldrich 5060, St Louis, MO, USA). Cells were washed three times with PBS then incubated with secondary antibodies: donkey anti-mouse Alexa 488 (1:1 000, Alexa Fluor 488) and donkey anti-rabbit (1:1 000, Alexa Fluor 594) (Invitrogen - Thermo Fisher Scientific, Waltham, MA, USA) in 1% BSA/PBS. Nuclear counterstaining was performed with DAPI. Cells were imaged on a Zeiss LSM 780 confocal microscope. Images were taken at 20x magnification.

2.3.4. Protein structural analysis (*in silico*)

The crystal structures were visualized and compared using SF1 DBD (PDB 2FF0) and LBD (PDB 1ZDT) 3D crystal structures with the PyMOL Molecular Graphics System v1.7.6 (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].

2.3.5. Variant phasing (c.368G>C and c.386C>T)

To assess the presence of the two variants c.368G>C and c.386C>T on the same allele as previously described [11,22], DNA from the patient was cloned using pGEM-T Easy Vector System (Promega Corporation, Madison, WI, USA) (see Supp. data for the method, the results, and Figure).

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 to 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.

<u>Four synonymous heterozygous variants</u>

c.1020G>A / p.=(Ala340Ala) was observed in a POI patient (Patient 3) whereas c.225G>C / p.=(Thr75Thr) and c.594G>A / p.=(pro198Pro) were observed in a single patient presenting with DOR (Patient 6). c.375G>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.225G>C, c.594G>A, c.375G>A; activation of an exonic cryptic acceptor site for c.594G>A, c.1020G>A). The fact that the rare synonymous variants are predicted to impact splicing means their potential for pathogenicity cannot be excluded, although further work would be required to validate their involvement.

<u>Six missense variants</u>

The well-described c.437G>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.43G>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.1198G>A / p.Ala400Thr was similarly identified in a POI patient (Patient 2).

Two variants, c.368G>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 #612964) [11] and spermatogenic failure 8 (SPGF8 OMIM #613957) [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 Figure 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), HSD3β2 (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] (Figure 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] (Figure 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 (Figure 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% to 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 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 (0.002496 and 0.002474

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 penetrate 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 [6]. 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 with primary amenorrhea, 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 trans-activation 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 Akloul 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.

Brittany 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.

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 in France, a Peter Doherty Early Career Fellowship (1054432; to E.J.T.), a National Health and Medical Research Council program grant (1074258; to A.H.S.), a fellowship (1062854; to A.H.S.) from the Australian National Health and Medical Research Council, and the Victorian Government's Operational Infrastructure Support Program.

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.

Provenance and peer review

This article has undergone peer review.

Research data (data sharing and collaboration)

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

Acknowledgments

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

References

[1] Gleicher N, Kim A, Weghofer A, Kushnir VA, Shohat-Tal A, Lazzaroni E, et al. Hypoandrogenism in association with diminished functional ovarian reserve. Hum Reprod Oxf Engl 2013;28:1084–91. doi:10.1093/humrep/det033.

[2] Iwase A, Osuka S, Goto M, Murase T, Nakamura T, Takikawa S, et al. Clinical application of serum anti-Müllerian hormone as an ovarian reserve marker: A review of recent studies. J Obstet Gynaecol Res 2018;44:998–1006. doi:10.1111/jog.13633.

[3] Mutlu MF, Erdem M, Erdem A, Yildiz S, Mutlu I, Arisoy O, et al. Antral follicle count determines poor ovarian response better than anti-Müllerian hormone but age is the only predictor for live birth in in vitro fertilization cycles. J Assist Reprod Genet 2013;30:657–65. doi:10.1007/s10815-013-9975-3.

[4] Victoria M, Labrosse J, Krief F, Cédrin-Durnerin I, Comtet M, Grynberg M. Anti Müllerian Hormone: More than a biomarker of female reproductive function. J Gynecol Obstet Hum Reprod 2019;48:19–24. doi:10.1016/j.jogoh.2018.10.015.

[5] Cohen J, Chabbert-Buffet N, Darai E. Diminished ovarian reserve, premature ovarian failure, poor ovarian responder--a plea for universal definitions. J Assist Reprod Genet 2015;32:1709–12. doi:10.1007/s10815-015-0595-y.
[6] Pastore LM, Christianson MS, Stelling J, Kearns WG, Segars JH. Reproductive ovarian testing and the alphabet soup of diagnoses: DOR, POI, POF, POR, and FOR. J Assist Reprod Genet 2018;35:17–23. doi:10.1007/s10815-017-1058-4.
[7] Tucker EJ, Grover SR, Bachelot A, Touraine P, Sinclair AH. Premature Ovarian Insufficiency: New Perspectives on Genetic Cause and Phenotypic Spectrum. Endocr Rev 2016;37:609–35. doi:10.1210/er.2016-1047.
[8] Wang T-T, Ke Z-H, Song Y, Chen L-T, Chen X-J, Feng C, et al. Identification of a mutation in GDF9 as a novel cause of diminished ovarian reserve in young women. Hum Reprod 2013;28:2473–81. doi:10.1093/humrep/det291.
[9] Pashaiasl M, Ebrahimi M, Ebrahimie E. Identification of the key regulating genes of diminished ovarian reserve (DOR) by network and gene ontology analysis. Mol Biol Rep 2016;43:923–37. doi:10.1007/s11033-016-4025-8.
[10] Bashamboo A, McElreavey K. NR5A1/SF-1 and development and function of the ovary. Ann Endocrinol 2010;71:177–82. doi:10.1016/j.ando.2010.02.013.
[11] Lourenço D, Brauner R, Lin L, De Perdigo A, Weryha G, Muresan M, et al. Mutations in NR5A1 associated with ovarian insufficiency. N Engl J Med 2009;360:1200–10. doi:10.1056/NEJMoa0806228.
[12] Jeyasuria P, Ikeda Y, Jamin SP, Zhao L, De Rooij DG, Themmen APN, et al. Cell-specific knockout of steroidogenic factor 1 reveals its essential roles in gonadal function. Mol Endocrinol Baltim Md 2004;18:1610–9. doi:10.1210/me.2003-0404.
[13] Camats N, Fernández-Cancio M, Audí L, Schaller A, Flück CE. Broad phenotypes in heterozygous NR5A1 46,XY patients with a disorder of sex development: an oligogenic origin? Eur J Hum Genet EJHG 2018;26:1329–38. doi:10.1038/s41431-018-0202-7.
[14] Robevska G, van den Bergen JA, Ohnesorg T, Eggers S, Hanna C, Hersmus R, et al. Functional characterization of novel NR5A1 variants reveals multiple complex roles in disorders of sex development. Hum Mutat 2018;39:124–39. doi:10.1002/humu.23354.
[15] Bashamboo A, Ferraz-de-Souza B, Lourenço D, Lin L, Sebire NJ, Montjean D, et al. Human male infertility associated with mutations in NR5A1 encoding steroidogenic factor 1. Am J Hum Genet 2010;87:505–12. doi:10.1016/j.ajhg.2010.09.009.
[16] Jiao X, Qin Y, Li G, Zhao S, You L, Ma J, et al. Novel NR5A1 missense mutation in premature ovarian failure: detection in han chinese indicates causation in different ethnic groups. PloS One 2013;8:e74759. doi:10.1371/journal.pone.0074759.
[17] Knarston IM, Robevska G, van den Bergen JA, Eggers S, Croft B, Yates J, et al. NR5A1 gene variants repress the ovarian-specific WNT signaling pathway in 46,XX disorders of sex development patients. Hum Mutat 2019;40:207–16. doi:10.1002/humu.23672.
[18] Stenson PD, Ball EV, Mort M, Phillips AD, Shiel JA, Thomas NST, et al. Human Gene Mutation Database (HGMD): 2003 update. Hum Mutat 2003;21:577–81. doi:10.1002/humu.10212.
[19] Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res 2018;46:D1062–7. doi:10.1093/nar/gkx1153.

[20] Sadedin SP, Dashnow H, James PA, Bahlo M, Bauer DC, Lonie A, et al. Cpipe: a shared variant detection pipeline designed for diagnostic settings. Genome Med 2015;7:68. doi:10.1186/s13073-015-0191-x.
[21] Venselaar H, Te Beek TAH, Kuipers RKP, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. BMC Bioinformatics 2010;11:548. doi:10.1186/1471-2105-11-548.
[22] Voican A, Bachelot A, Bouligand J, Francou B, Dulon J, Lombès M, et al. NR5A1 (SF-1) mutations are not a major cause of primary ovarian insufficiency. J Clin Endocrinol Metab 2013;98:E1017-1021. doi:10.1210/jc.2012-4111.
[23] Tucker EJ, Jaillard S, Grover SR, den Bergen J, Robevska G, Bell KM, et al. TP63-truncating variants cause isolated premature ovarian insufficiency. Hum Mutat 2019. doi:10.1002/humu.23744.
[24] Camats N, Pandey AV, Fernández-Cancio M, Andaluz P, Janner M, Torán N, et al. Ten novel mutations in the NR5A1 gene cause disordered sex development in 46,XY and ovarian insufficiency in 46,XX individuals. J Clin Endocrinol Metab 2012;97:E1294-1306. doi:10.1210/jc.2011-3169.
[25] Lin L, Philibert P, Ferraz-de-Souza B, Kelberman D, Homfray T, Albanese A, et al. Heterozygous missense mutations in steroidogenic factor 1 (SF1/Ad4BP, NR5A1) are associated with 46,XY disorders of sex development with normal adrenal function. J Clin Endocrinol Metab 2007;92:991–9. doi:10.1210/jc.2006-1672.
[26] Janse F, de With LM, Duran KJ, Kloosterman WP, Goverde AJ, Lambalk CB, et al. Limited contribution of NR5A1 (SF-1) mutations in women with primary ovarian insufficiency (POI). Fertil Steril 2012;97:141-146.e2. doi:10.1016/j.fertnstert.2011.10.032.
[27] Sreenivasan R, Ludbrook L, Fisher B, Declosmenil F, Knower KC, Croft B, et al. Mutant NR5A1/SF-1 in patients with disorders of sex development shows defective activation of the SOX9 TESCO enhancer. Hum Mutat 2018;39:1861–74. doi:10.1002/humu.23603.
[28] Eggers S, Smith KR, Bahlo M, Looijenga LHJ, Drop SLS, Juniarto ZA, et al. Whole exome sequencing combined with linkage analysis identifies a novel 3 bp deletion in NR5A1. Eur J Hum Genet EJHG 2015;23:486–93. doi:10.1038/ejhg.2014.130.

Legend to Figures and Tables

Figure 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 non-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.

Figure 2: Trans-activational activity of the protein-altering *NR5A1* variants. The trans-activational activity of each variant compared with wild-type SF1 was tested using a dual-Luciferase reporter assay in HEK293T cells. Luciferase reporter constructs used were pGL4-mTesco, pGL4.10-hCYP11A1, pGL4.10-hStAR, pGL3-hCYP17A1, pGL3-HSD3β2. 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 - $p<0.001$).

Figure 3: **Sub-cellular localization of the protein-altering *NR5A1* variants.** Protein expression of each variant and wild-type (WT) SF1 was assessed in HEK293T 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.

Figure 4: **Protein structural analysis and conformational changes in SF1 variant proteins.** To investigate the potential impact of each variant on protein conformation, we performed an in silico prediction using PyMol modeling software and HOPE database. **4a**: SF1 DBD bound to the inhibin α-subunit promoter shows the close proximity of Val15 (pink and surronded) to the Zn finger ("Zn" / green). **4b**: substitution of Val15 (green / dot arrow) with Met (red / full arrow). **4c**: Ala400 (pink and surrounded) is located far from the ligand binding pocket ("Ligand" / green) and the co-factor interaction domain ("Co-factor NCOA2" / brown). **4d**: substitution of hydrophobic Ala400 (green / dot arrow) with polar/hydrophilic Thr (red / full arrow) is predicted to affect protein folding and conformation.

Table 1: Clinical data of the patients with rare *NR5A1* variants.

Table 2: Molecular data of all *NR5A1* variants.

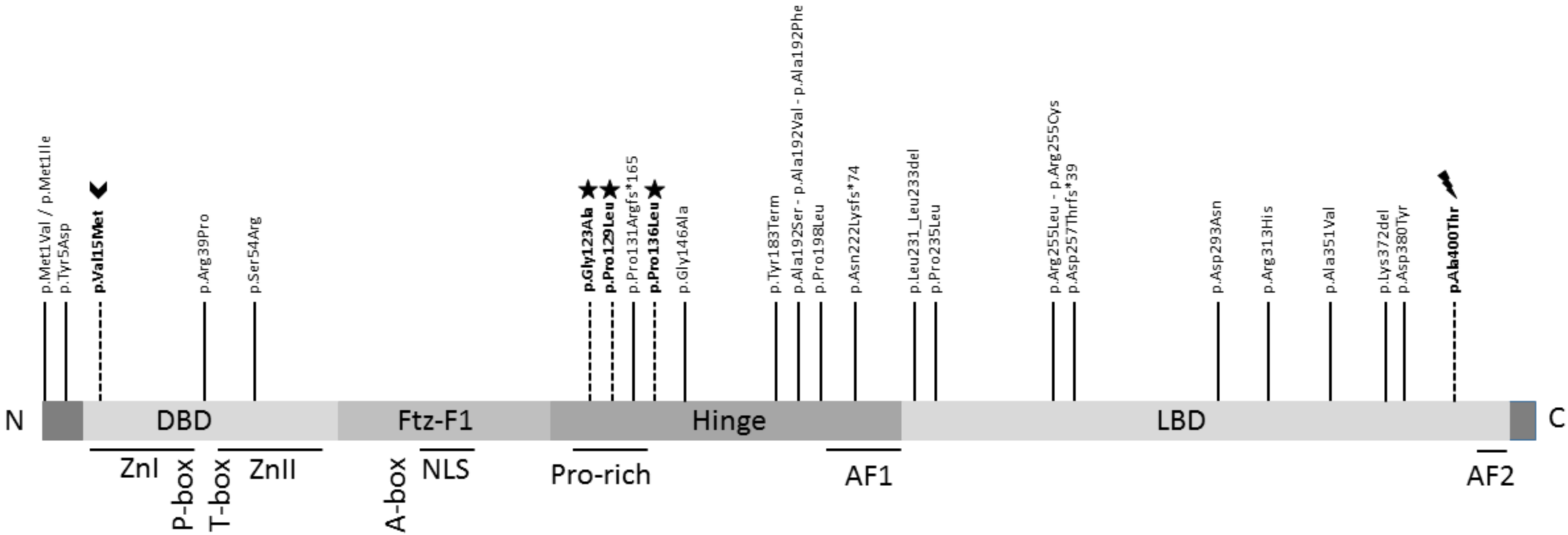

p.Met1Val / p.Met1Ile
p.Tyr5Asp
p.Val15Met
p.Arg39Pro
p.Ser54Arg
p.Gly123Ala
p.Pro129Leu
p.Pro131Argfs*165
p.Pro136Leu
p.Gly146Ala
p.Tyr183Term
p.Ala192Ser - p.Ala192Val - p.Ala192Phe
p.Pro198Leu
p.Asn222Lysfs*74
p.Leu231_Leu233del
p.Pro235Leu
p.Arg255Leu - p.Arg255Cys
p.Asp257Thrfs*39
p.Asp293Asn
p.Arg313His
p.Ala351Val
p.Lys372del
p.Asp380Tyr
p.Ala400Thr
N
DBD
Ftz-F1
Hinge
LBD
C
ZnI
P-box
T-box
ZnII
A-box
NLS
Pro-rich
AF1
AF2

Figure 1

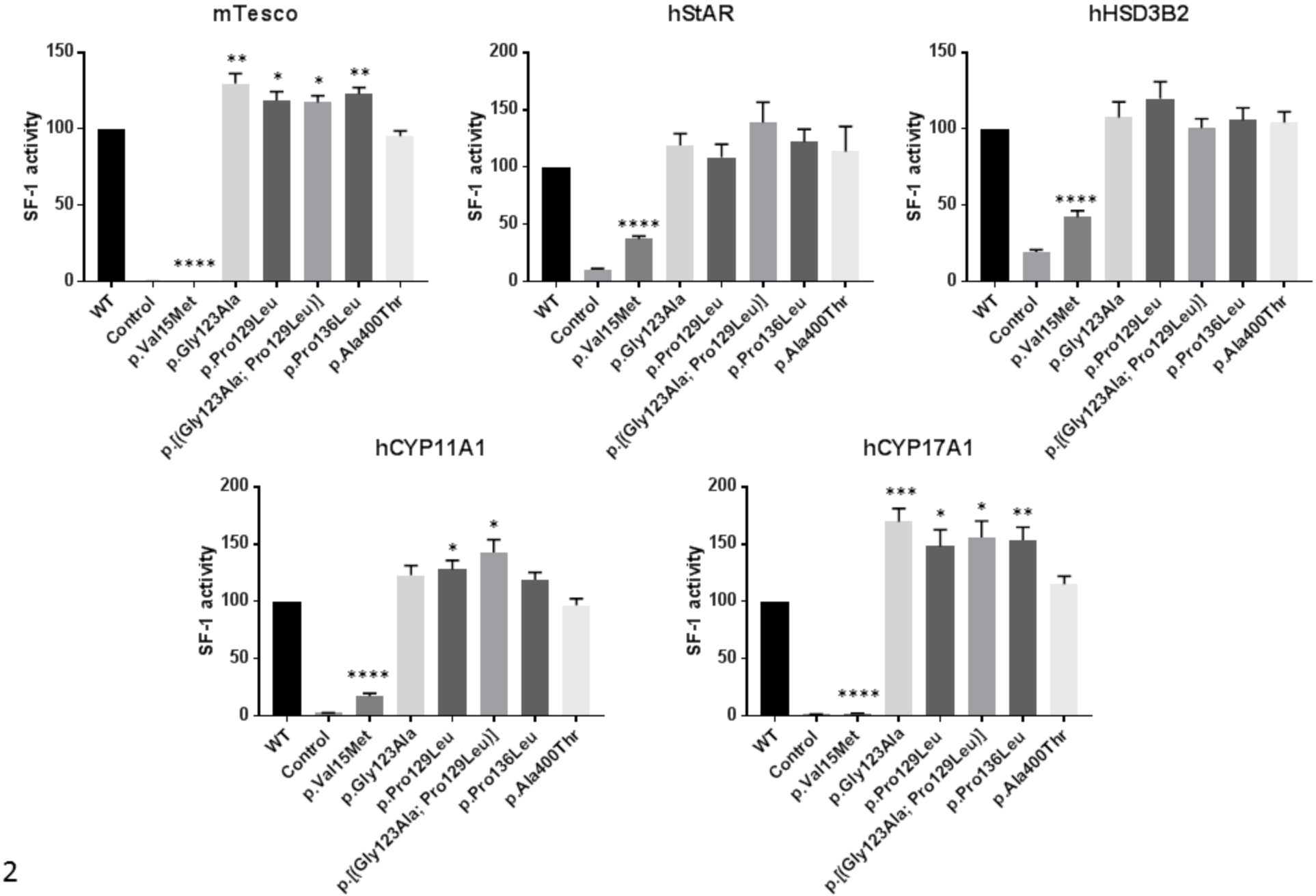

mTesco
hStAR
hHSD3B2
hCYP11A1
hCYP17A1
SF-1 activity
WT
Control
p.Val15Met
p.Gly123Ala
p.Pro129Leu
p.[(Gly123Ala; Pro129Leu)]
p.Pro136Leu
p.Ala400Thr

Figure 2

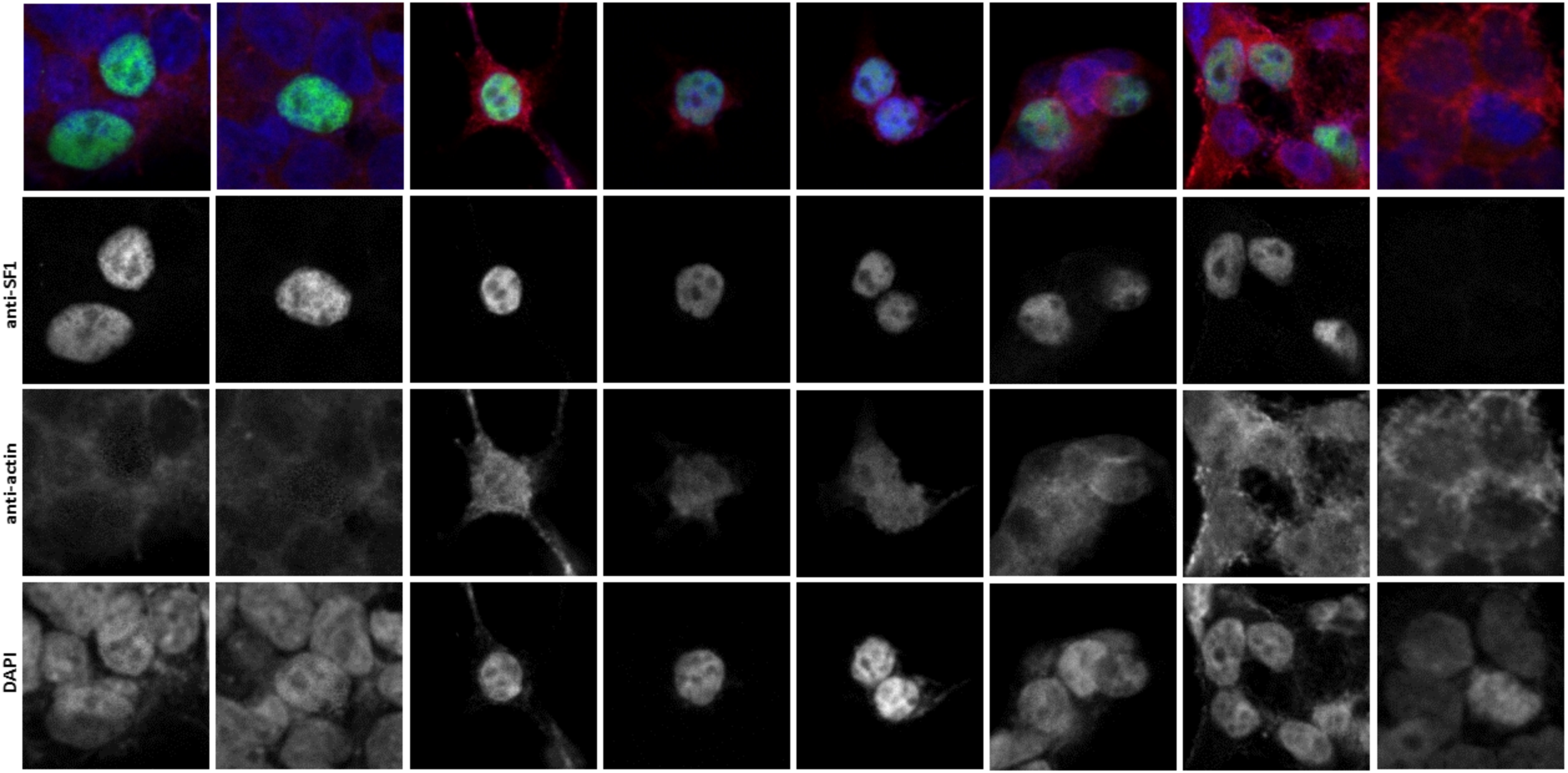
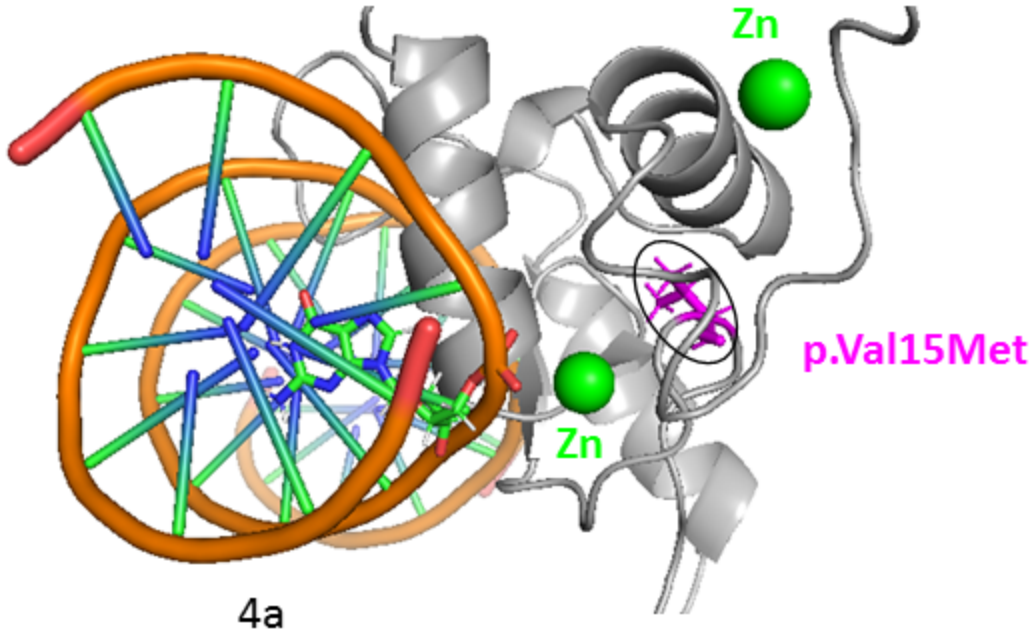
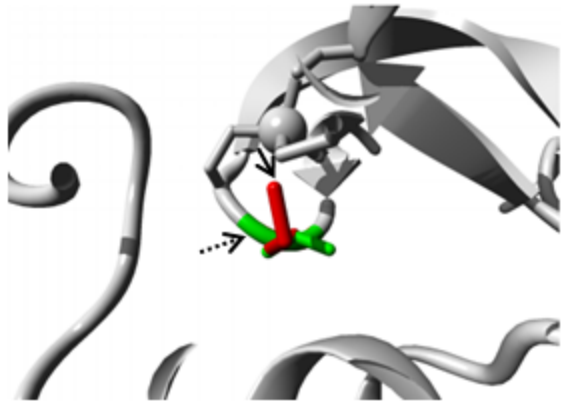

anti-SF1
anti-actin
DAPI
WT
p.Val15Met
p.Gly123Ala
p.Pro129Leu
p.[(Gly123Ala; Pro129Leu)]
p.Pro136Leu
p.Ala400Thr
Control

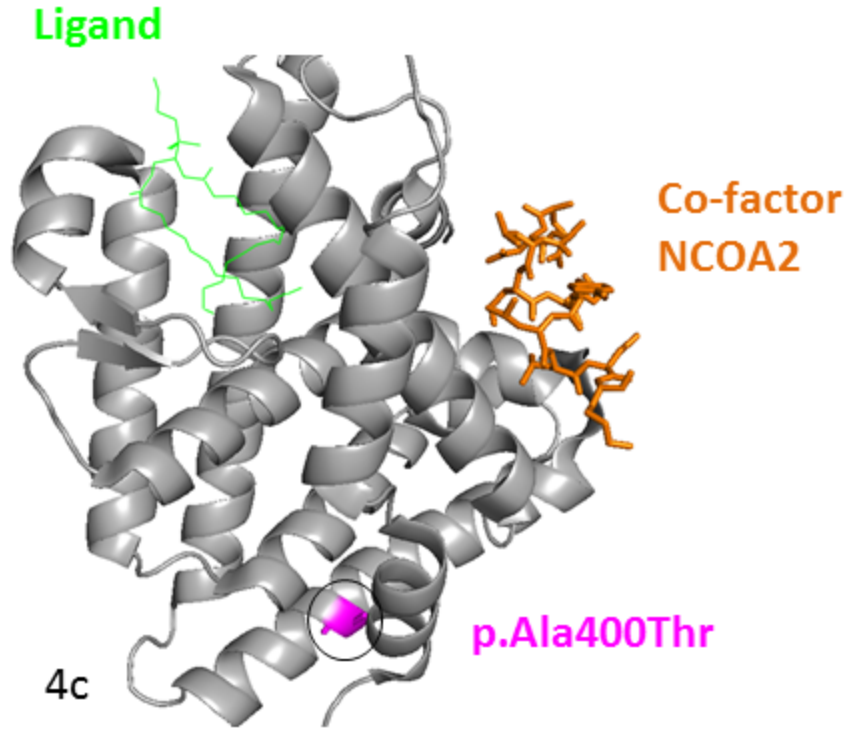
Figure 3



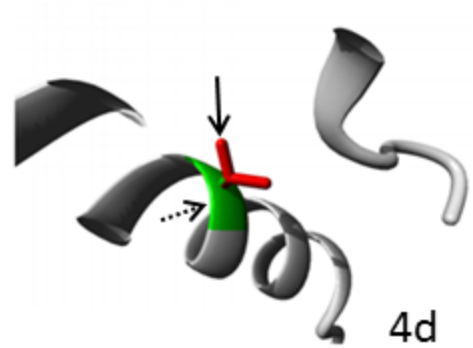
4a



4b



4c



4d

Figure 4

	Patient 1 (POI)	Patient 2 (POI)	Patient 3 (POI)	Patient 4 (DOR)	Patient 5 (DOR)	Patient 6 (DOR)
	c.43G>A	c.1198G>A	c.1020G>A	c.368G>C + c.386C>T (cis)	c.407C>T	c.225G>C + c.594G>A
Age of OD	28	31	35	30	32	30
Menstruation	Secondary amenorrhea	Secondary amenorrhea	Irregular	Regular	Regular	Regular
FSH IU/l	29	25	26	6.2	4.9	11.9
LH IU/l	10.7	ND	11	4.4	2.7	4.6
Estradiol pg/ml	26	ND	22.6	34.9	117	35
AMH ng/ml	0.4	ND	<0.4	1.1	0.8	<0.4
AFC	0	ND	3	4	1	1
Karyotype	46,XX	46,XX	46,XX	46,XX	46,XX	46,XX
FMR1	N	N	N	N	N	N
Other	European ethnicity Omphalocele at birth	African ethnicity	European ethnicity One miscarriage	North African ethnicity Pregnancy with oocyte donation in 2016	European ethnicity Failure of seven intra-uterine inseminations	African ethnicity Bilateral hydrosalpinx, endometrial hyperplasia One child, one VTP

Table 1

OD: ovarian deficiency, POI: premature ovarian insufficiency, DOR: diminished ovarian reserve, FSH: follicle stimulating hormone, LH: luteinizing hormone, AMH: anti-Müllerian hormone, AFC: antral follicular count, N: normal, ND: not determined, VTP: voluntary termination of pregnancy

Variant description NM_004959.4	Chromosomal variant NC_000009.11	Exon	Amino acid change NP_004950.2	MAF (gnomAD)	In silico prediction (Mutation Taster, SIFT, PolyPhen-2, DANN)	CADD score (Phred)	Nucleotide conservation GERP score	Amino-acid conservation (UCSC Multiz Alignments of 100 Vertebrates)	ClinVar	American College of Medical Genetics ACMG
c.43G>A	g.127265632C>T rs104894124	2	p.Val15Met	0	MT: disease causing SIFT: damaging PP-2: probably damaging DANN: 0.9965	Phred: 25 (pathogenic)	Moderate (GERP RS: 2.79)	High	46,XY Sex Reversal 3	Pathogenic
c.368G>C	g.127262871C>G rs200163795	4	p.Gly123Ala	T: 0.0002782 AP: 0.002496	MT: disease causing SIFT: tolerated PP-2: benign DANN: 0.7211	Phred: 8.465 (benign)	Moderate (GERP RS: 2.49)	High in mammals Low in other species	Premature ovarian failure 7 Spermatogenic failure 8	Uncertain significance
c.386C>T	g.127262853G>A rs200749741	4	p.Pro129Leu	T: 0.0002754 AP: 0.002474	MT : disease causing SIFT : damaging PP-2 : benign DANN: 0.9983	Phred: 16.71 (likely pathogenic)	Moderate (GERP RS: 3.82)	High in mammals Low in other species	Premature ovarian failure 7 Spermatogenic failure 8	Uncertain significance
c.407C>T	g.127262832G>A rs780022827	4	p.Pro136Leu	T: 0.00004825 EP: 0.00003309	MT: polymorphism SIFT: tolerated PP-2: benign DANN: 0.9868	Phred: 12.98 (likely benign)	Moderate (GERP RS: 3.55)	High in mammals Low in other species	Unknown	Uncertain significance
c.437G>C	g.127262802C>A rs1110061	4	p.Gly146Ala	T: 0.1165	MT: polymorphism SIFT: tolerated PP-2: benign DANN: 0.06778	Phred: 3.668 (benign)	Low (GERP RS: -0.606)	Low	Benign	Benign
c.1198G>A	g.127245225C>T rs761421822	7	p.Ala400Thr	T: 0.000003984 AP: 0	MT: disease causing SIFT: tolerated PP-2: probably damaging DANN: 0.9994	Phred: 26 (pathogenic)	High (GERP RS: 5.17)	High	Unknown	Likely pathogenic
c.225G>C	g.127265377C>G rs138805488	3	p.=(Thr75Thr)	T: 0.006440 AP: 0.06680	MT: disease causing DANN: 0.8345 HSF: potential alteration of splicing	Phred: 15.47 (likely pathogenic)	Low (GERP RS: -0.911)	High in mammals Low in other species	Benign	Uncertain significance
c.375G>A	g.127262864C>T rs1110062	4	p.=(Pro125Pro)	T: 0.02887 AP: 0.3040 EP: 0.0008281	MT: polymorphism DANN: 0.7006 HSF: potential alteration of splicing	Phred: 0.527 (benign)	Low (GERP RS: -9.08)	High in mammals Low in other species	Benign	Uncertain significance
c.594G>A	g.127262645C>T rs142414614	4	p.=(pro198Pro)	T: 0.004125 AP: 0.04125	MT: disease causing DANN: 0.7777 HSF: potential alteration of splicing	Phred: 1.057 (benign)	Low (GERP RS: -9.68)	High	Benign	Uncertain significance
c.1020G>A	g.127253478C>T rs758624846	6	p.=(Ala340Ala)	T: 0.00002263 EP: 0.00001685	MT: disease causing DANN: 0.772 HSF: potential alteration of splicing	Phred: 5.965 (benign)	Low (GERP RS: -10.19)	High	Unknown	Uncertain significance

Table 2

AP: African population, EP: European population, MT: Mutation Taster, PP-2: PolyPhen-2, DANN: Deleterious Annotation of genetic variants using Neural Networks (0 to 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_04959.4, 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.