A long contiguous stretches of homozygosity disclosed a novel biallelic pathogenic variant in STAG3 causing Primary Ovarian Insufficiency.

Simona Mellone¹, Marco Zavattaro¹, Denise Vurchio³, Sara Ronzani³, Marina Caputo²,³, Ilaria Leone², Flavia Prodam²,³ and Mara Giordano*¹,³

*These authors contributed equally to this work and share first authorship

1- Laboratory of Genetics, SCDU Biochimica Clinica, Ospedale Maggiore della Carità, Novara

2- SCDU Endocrinology, Department of Translational Medicine, Università del Piemonte Orientale

3- Department of Health Sciences, Università del Piemonte Orientale

*Corresponding author:
Mara Giordano
mara.giordano@med.uniupo.it
ABSTRACT

Background: Primary ovarian insufficiency (POI) refers to an etiologically heterogeneous disorder characterized by hypergonadotropic hypogonadism that represents a major cause of infertility in women under 40 years. Most cases of isolated POI still appear sporadically, but ∼10–15% have an affected first-degree relative, indicating a significant genetic etiology. Several genes implicated in development, meiosis, hormonal signaling and metabolism are involved in the genetic form of the disorder in both syndromic and isolated POI. However, most cases of POI remain unsolved even after exhaustive investigation.

Results: Here is reported a 19-year-old Senegalese female affected by non-syndromic POI showing primary amenorrhea, who well answered to the hormonal induction of puberty reaching a complete sexual maturation in two years. In order to investigate the presence of a genetic defect, aCGH-SNP analysis was performed. A 13.5 Mb long contiguous stretch of homozygosity (LCSH) region on chromosome 7q21.13-q22.1 was identified where the exome sequencing revealed a novel homozygous 4-bp deletion (c.3381_3384delAGAA) in STAG3. Pathogenic variants in this gene, encoding for a meiosis-specific protein, have been previously reported as cause POI in only seven families and recently as cause of infertility in a male. The here identified mutation leads to the truncation of the last 55 aminoacids, confirming the important role in meiosis of the STAG3 C-terminal domain.

Conclusions: In conclusion we identified a loss of function variant in STAG3 in a Senegalese woman with POI reinforcing the role the cohesin complex in the genetic etiology of this disorder. This gene should be included in the screening of POI to offer a better genetic counseling and long term follow-up considering the risk of ovarian tumor in woman carrying pathogenic variants in genes involved in germ cell differentiation.

Key words: POI, STAG3, SNP-array, clinical exome sequencing
Background

Primary ovarian insufficiency (POI) is characterized by depletion of ovary follicles, leading to hypoestrogenism and hypergonadotropic hypogonadism, infertility and amenorrhea in women younger than 40 years (1,2). This disorder represents one of the main causes of infertility affecting approximately 1% of women (3,4) with a worldwide prevalence of 3.7% (5). Women with POI may show a wide range of clinical phenotypes, from primary (PA) to secondary amenorrhea (SA) and other congenital or acquired abnormalities and present either syndromic or isolated forms of the disease (6,7). PA is usually diagnosed in adolescence in patients showing delayed puberty and absence of secondary sex characteristic development, while SA, which represents the most frequent POI phenotype, may occur at any age after menarche and is characterized by normal pubertal development (8,9). Although POI is a heterogeneous disorder caused by iatrogenic, viral, or autoimmune factors, more than 70% of cases remain idiopathic (1). Chromosomal aberrations have long been recognized as a cause of POI and at least 10-13% of the syndromic cases present anomalies revealed by standard karyotype (10). About 10-15% of women with POI have an affected first-degree relative and different modes of inheritance can be observed in families, suggesting the presence of several monogenic causes in the etiology of the disease (7). To date the use of different genomic approaches, including linkage studies, sequencing of candidate genes and whole exome sequencing (WES) allowed the identification of pathogenic alterations in more than 60 genes implicated in both syndromic and isolated POI (1,10). Additionally, submicroscopic copy number variations (CNVs) encompassing genes potentially implicated in reproductive function have emerged as an important genetic determinant in POI (11). The genes altered in POI encode for protein involved in DNA repair and meiosis (1) and pathogenic variants also predispose to different forms of cancer (6,12,13,14). Thus the early diagnosis of the molecular mechanism
underlying POI is essential to develop strategies for preventing the irreversible consequences on fertility, to improve clinical management and to perform genetic counseling with a long-term follow-up also considering tumor susceptibility.

Mutations in STAG3 (stromal antigen 3), which encodes a subunit of the cohesin complex participating to sister chromatid pairing during meiosis, have been identified as a rare POI monogenic cause. To date STAG3 biallelic variants have been reported in seven families worldwide, five of which were consanguineous pedigree (6,14-19). All the affected women had isolated POI except for a patient belonging to a Palestinian family (6) who presented simultaneous bilateral ovarian tumors.

Here we report the identification of a novel biallelic pathogenic variation in STAG3 through a combined approach of CGH-SNP microarray and Clinical Exome Sequencing (CES) in a young Senegalese woman affected by non-syndromic POI.

MATERIALS AND METHODS

Patient description

The patient was a 19-year-old female from Senegal, referring to the Emergency Room Department for dyspareunia and pelvic pain after her first sexual intercourse. She was also suffering from primary amenorrhea and anosmia. Familiar anamnesis revealed the absence of parental consanguinity. Computerized tomography (CT) investigation showed vaginal stenosis and failure to display ovaries. Physical examination showed absence of breast and pubic hair development, consistent with Tanner stage 1.

Hormonal profile was indicative of primary hypogonadism (17β-estradiol 20 pg/ml, FSH 88.1 mU/ml, LH 28.7 mU/ml), while the remaining pituitary function was preserved (TSH 1.4 µU/ml, free-T₄ 1.33 ng/dl, prolactin 9.78 ng/ml, GH 1.99 ng/ml, IGF-I 271.8 ng/ml, ACTH 22.4 pg/ml,
8.00 a.m. cortisol 11.7 µg/dl). Further investigations revealed reduced bone mineral density at lumbar spine (Z-score -3.6) and delayed hand bone age (14 years). Moreover, pelvic magnetic resonance (MR) confirmed the presence of a pre-pubertal uterus, characterized by hypoplasia (38x11x19 mm) and a fundus-to-cervix ratio of 1:1 (Figure 1, panel a,b). The cytogenetic analysis revealed a normal karyotype (46,XX).

After hormonal and radiological diagnostic work-up, a puberty induction therapy using transdermal 17β-oestradiol (oestradiol hemyhidrate 25 µg/3 days for 2 months, then 50 µg/3 days) was started. After 3 months of treatment the presence of pubic hair (P2) and breast development (B2), consistent with Tanner stage 2, were appreciable at physical examination. Six months since the beginning of puberty induction therapy, breast size and pubic hair further increased (B3 P3), and menarche occurred. Hormonal evaluation showed adequate 17β-oestradiol levels (75 pg/ml) and consensual reduction in gonadotropin concentration (FSH 4.7 mU/ml, LH 0.4 mU/ml). Treatment with transdermal estrogen led to oligomenorrhea and progressive pubertal development over the following months. After 24 months of therapy complete pubertal (B5 P4) and uterine development occurred, and combined therapy with progesterone was started (Tab. 1). Pelvic magnetic resonance performed at that time revealed an anteverted-antiflex uterus of normal size and regular vaginal morphology; endometrial thickness was normal according to menstrual phase (Fig. 1 - panel c,d).

**CGH-SNP microarray analysis**

Following written informed consent, the genomic DNA of the proband was extracted from peripheral blood through the ReliaPrep Blood gDNA Miniprep System (Promega), according to the manufacturer's recommendation. The GenetiSure Dx Postnatal Array 4x180K + SNP (Agilent Technologies, USA) was used following standard protocols. These platform contains about 59000 SNP probes and ~107000 oligonucleotide probes (60-mer) with a resolution of 5 ~ 10 Mb for ROH detection.
The slides were scanned through the Agilent SureScan Dx Microarray Scanner System (Agilent Technologies, USA) and data were analyzed through the Agilent Feature Extraction Software. Copy Number Variations (CNVs) were annotated by the Agilent CytoDx Software 1.1.1.0 using tiff images from data.

**Clinical exome sequencing (CES)**

The Agilent SureSelect Custom Constitutional Panel 17 Mb (Agilent, Santa Clara, CA, USA) which provides comprehensive coverage of 5227 clinically relevant genes was used for the preparation of the library. Exon-enriched library was subjected to a 150 bp paired-end sequencing on the Illumina MiSeq platform (Illumina, San Diego, California, USA). Sequencing reads passing quality filters were aligned to the human reference genome build (GRGh37/hg19) and variant calling was performed using the SureCall v3.5 software (Agilent Technologies, Santa Clara CA). Then VCF files were annotated with the wANNOVAR tool. Finally the variants were filtered and prioritized using a personalized bioinformatics pipeline: variants with a read coverage of less than 5x and a Q score below 20 were filtered out; allelic frequency in public databases variants <0.1% (1000 Genomes, the ESP cohort data set, GnomAD, Exome Aggregation Consortium) were considered; synonymous variants were excluded; for the missense variants, at least four in-silico prediction tools (SIFT, CADD, Polyphen, MutationTaster) were used.

**Sanger sequencing**

The \textit{STAG3} variant (NM_001282716.1:c.3381_3384delAGAA; p.Glu1128MetfsTer42) identified through clinical-exome sequencing was validated by PCR and Sanger Sequencing using the following PCR primers: forward:5'-TTGGAAAGAGAGCACACCTG-3' and reverse : 5'-TGGTGTATAATGGGGAGAAAA-3'. The Sequencing reaction was performed in both directions using the PCR primers on a SeqStudio Genetic AnalyzerSequencer (Thermo Fisher Scientific, UK)
Results

An array CGH-SNP assay was used as first tier test to investigate the presence of both DNA copy number variations (CNVs) and Loss of heterozygosity (LOH) regions. No chromosomal unbalance was identified. However the aCGH-SNP revealed a LCSH (Long contiguous stretches of homozygosity) region of 13.5 Mb (from rs2374083 to rs1990167) on chromosome 7q21.13-q22.1 (Fig 2A). This region contained 127 OMIM genes among which the best candidate to explain the patient’s phenotype was STAG3, whose mutations had been previously detected in patients with autosomal recessive POI (6,14-19).

Due to the large size of STAG3 (34 exons) and eventually to undisclosed pathogenic variants in other genes a NGS platform including 5227 genes (Agilent SureSelect Custom Constitutional Panel 17) involved in human genetic disorders (Clinical Exome) was used to identify the causative molecular defect. The average sequencing depth was 77x with 96.93% of the target sequence covered at ≥20x, 89.27% covered at ≥50x and 70.52% covered at ≥100x. The variants obtained from CES were filtered and confirmed the presence of a LCSH on chromosome 7q21.13-q22.1. A variant at the homozygous state in exon 30 of the STAG3 gene was identified, namely c.3381_3384delAGAA (NM_001282716.1; Fig. 2B), then validated by Sanger sequencing. This 4-bp deletion causes a frameshift with the introduction of a premature stop codon (p.E1128Mfs*42) resulting in a predicted truncated protein of 1,170 amino acids (wild-type STAG3 protein: 1225 amino acids). The variant was not present in the public databases and was classified as pathogenic according to the American College of Medical Genetics and Genomics (ACMG) standards and guidelines for the interpretation of variations (criteria: PVS1, PM2, PP3) (20). We did not find any other mutations in genes known to cause POI and consequently STAG3 p.E1128Mfs*42 was considered the best disease-causing candidate variant in this patient.

Discussion
POI is an extremely heterogeneous disorder caused by pathogenic variations in genes involved in germ cell development, oogenesis, folliculogenesis, steroidogenesis and hormone signaling (9). Chromosomal unbalances detected through standard karyotype as well as submicroscopic alterations are involved in syndromic forms of POI (10). Leding et al. (2010) (21) reported the identification of several micro-rearrangements detected through aCGH in 74 German POI patients encompassing several genes involved in meiosis, DNA repair and folliculogenesis.

In the present patient a platform combining aCGH and SNP array has been used as the first-tier genetic test. In addition to CNVs this platform, detects the presence of homozygosity regions through SNP genotyping as the consequence of i) hemizygosity caused by a CNV, that can be simultaneously confirmed by CGH probes ii) autozygosity of identity-by-descent regions underlying the presence of a putative recessive disorder iii) uniparental disomy (UPD) when the homozygosity is confined to a chromosomal region (or to entire chromosome) underlying an imprinting disorder. In the here described patient the CGH-SNP array revealed a normal female karyotype and the presence of a unique interstitial long contiguous stretches of homozygosity region (LCSH) of 13.5 Mb on chromosome 7q21.13-q22.1. Besides the 7q LCSH, no other large regions (>5 Mb) of homozygosity were detected, indicating either segmental UPD (22) or a distant shared parental ancestry. Unfortunately the parent’s DNA was not available to investigate the origin of this LCSH.

The best candidate in the LCSH region was STAG3, that has been previously associated with a recessive form of POI consistent with the patient’s phenotype. STAG3 encodes a subunit of the multiprotein cohesin complex required during meiosis I for homologous chromosome pairing, correct synapsis and segregation of chromosomes, proper recombination and DNA repair (23). In mammals, meiosis-specific cohesin subunits includes a SMC1 subunit (SMC1β), two additional α-kleisins (RAD21L and REC8) and a stromal antigen protein (STAG3) that are specifically
expressed in meiosis. Moreover, Stag3 specifically localizes on the spindle apparatus and is required for microtubule stability and spindle assembly to preserve the euploidy in the mouse eggs (24).

Females \textit{Stag3\textsuperscript{−/−}} mice lack ovarian follicles indicating a severe ovarian dysgenesis (25) and also develop ovarian tumors as observed in humans (6). It is likely that homozygous or compound heterozygous variants in \textit{STAG3} may also affect male fertility due to the early prophase I arrest and apoptosis in both sex germ cells. Indeed, all \textit{Stag3\textsuperscript{−/−}} male mice described to date are infertile (26,27) and a biallelic loss-of-function \textit{STAG3} variant has been recently reported in an infertile male (28) affected by complete bilateral meiotic arrest.

Up to date, only ten variants in \textit{STAG3} have been reported as genetic causes of POI in seven pedigrees (Fig 3) (6,14-19) with a predicted loss of function effect in most cases. The here identified \textit{STAG3} frameshift variant is a 4-bp deletion in exon 30 and the resulting transcript would either undergo nonsense-mediated decay (29) or result in a truncated protein devoided of the C-terminal domain similarly to a previously described frameshift mutation truncating the last 206 C-terminal residues, in exon 28 (14). The C-terminal domain is specific of \textit{STAG3} as it is not conserved neither in \textit{STAG1} nor in \textit{STAG2}, that belong to the same family but are involved in somatic cell division (30). The findings of two POI patients with pathogenic variants truncating the C-terminal domain support an important specific role of this \textit{STAG3} domain during meiosis.

In the present patient, hormone replacement therapy (HRT) allowed to induce puberty and achieve complete secondary sexual characteristics, adequate growth, uterine development, and menstrual fluxes. HRT currently represent the mainstay therapy for women with POI and plays a pivotal role in reducing long-term comorbidities (e.g. osteoporosis and cardiovascular disease) and improving sexual health by restoring normal serum estrogen concentrations according to age (3,8,31).

Taken together our observations strength the importance of a correct diagnosis and clinical management in women with POI. The diagnostic workflow should include the molecular test with
multi-gene panel including STAG3 to allow an appropriate ovarian monitoring in the long term follow-up, considering the high risk to develop ovarian tumors.

**Conflict of Interest**
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Author Contributions**
MG and FP: Conception and coordination of the work and drafted the article. SM: acquisition, analysis, interpretation of data. DV, SR: performing of genetic tests. MZ, MC, IL: clinical assessment and management of the patient. All authors contributed to the article and approved the submitted version.

**Availability of data and materials**
All data generated in this study are included in the published article

**Funding**
Fondi Vari Prof. Giordano Dipartimento di Scienze della Salute, Università del Piemonte Orientale.

**Acknowledgments**
We acknowledge the technical staff of the laboratory of Genetics (SCDU Biochimica Clinica, Ospedale Maggiore della Carità)

**References**


Legends to figures

Figure 1: Magnetic resonance before and after oestrogen replacement therapy. Panels “a” and “b” show uterine hypoplasia (continuous white arrows) in sagittal (panel “a”) and transversal plane (panel “b”). Reduced uterine thickness, consistent with pre-pubertal morphology, is evidenced in panel “a” (dashed arrow). In panels “c” and “d” continuous white arrows indicate the uterine development after 24 months of oestrogen therapy both in sagittal (panel “c”) and transversal (panel “d”) planes. Endometrial development and increase in uterine thickness are evident in panel “c” (dashed arrow).

Figure 2: CGH-SNP array and sequencing results. (A) Array-CGH-SNP results: the LCSH on chromosome 7q21.13-q22.1 is indicated by a blue arrow (B) In the upper part is shown the partial sequencing electropherogram including the novel homozygous mutation identified in STAG3 (c.3381_3384delAGAA; NM_001282716.1); the corresponding wild-type sequence is showed below with a dashed rectangle showing the deleted bases.

Figure 3: Schematic representation of STAG3 and protein structure with pathogenic variants reported in POI patients. The exon-intron structure of STAG3 (ENSEMBL: transcript ID ENST00000426455.5) is displayed in the lower part of the figure. The scheme of STAG3 protein consisting of 1225 amino acids (aa) (Protein domains for ENSP00000400359.1) is reported in the middle part of the figure with colored rectangles indicating: the STAG domain (174-283 aa; PF08514 Pfam database), an Armadillo-type fold domain (ARM-type fold) (303-813 aa; SSF48371 Superfamily database) and a Stromalin conservative domain (SCD) (309-394 aa; PS51425 Prosite profiles). The pathogenetic variants identified in previous studies are shown in blue boxes whereas the variant identified this study is shown in the red box.