Next generation sequencing in women affected by nonsyndromic premature ovarian failure displays new potential causative genes and mutations

Dora Janeth Fonseca, Ph.D.,[a] Liliana Catherine Patiño, M.Sc.,[a] Yohjana Carolina Suárez, M.Sc.,[a] Asid de Jesús Rodríguez, M.Sc.,[a] Heidi Eliana Mateus, M.D., M.Sc.,[a] Karen Marcela Jiménez, M.Sc.,[a] Oscar Ortega-Recalde, M.D., M.Sc.,[a] Ivonne Díaz-Yamal, M.D.,[b] and Paul Laissue, M.D., Ph.D.[a]

[a] Unidad de Genética, Grupo GENIUROS, Escuela de Medicina y Ciencias de la Salud, Universidad del Rosario; and [b] Unidad de Fertilidad PMA, Clínica Marly, Bogota, Colombia

Objective: To identify new molecular actors involved in nonsyndromic premature ovarian failure (POF) etiology.
Design: This is a retrospective case-control cohort study.
Setting: University research group and IVF medical center.
Patient(s): Twelve women affected by nonsyndromic POF. The control group included 176 women whose menopause had occurred after age 50 and had no antecedents regarding gynecological disease. A further 345 women from the same ethnic origin (general population group) were also recruited to assess allele frequency for potentially deleterious sequence variants.
Intervention(s): Next generation sequencing (NGS), Sanger sequencing, and bioinformatics analysis.
Main Outcome Measure(s): The complete coding regions of 70 candidate genes were massively sequenced, via NGS, in POF patients. Bioinformatics and genetics were used to confirm NGS results and to identify potential sequence variants related to the disease pathogenesis.
Result(s): We have identified mutations in two novel genes, *ADAMTS19* and *BMPR2*, that are potentially related to POF origin. *LHCGR* mutations, which might have contributed to the phenotype, were also detected.
Conclusion(s): We thus recommend NGS as a powerful tool for identifying new molecular actors in POF and for future diagnostic/prognostic purposes. (Fertil Steril® 2015;104:154–62. ©2015 by American Society for Reproductive Medicine.)
Key Words: Premature ovarian failure, POF, next generation sequencing, *ADAMTS19*, *BMPR2*

Discuss: You can discuss this article with its authors and with other ASRM members at http://fertstertforum.com/fonsecad-ngs-pof-potential-genes/

Use your smartphone to scan this QR code and connect to the discussion forum for this article now.*

* Download a free QR code scanner by searching for "QR scanner" in your smartphone's app store or app marketplace.

Primary ovarian insufficiency is a heterogeneous pathology consisting of irregular ovulation secondary to distinct functional defects (1). This condition may evolve to premature ovarian failure (POF), a frequent condition characterized by the cessation of menses before age 40 and increased plasma gonadotropin levels (e.g., FSH >40 IU/L) (2). Women with POF may have primary or secondary amenorrhea depending on whether the menarche has occurred (secondary) or not (primary) (3). Although most POF cases are considered idiopathic, distinct etiologies have been described, such as iatrogenic events, infectious agents, autoimmune conditions, environmental factors, and metabolic disorders. Most nonsyndromic POF cases that are considered idiopathic have pointed to the suspicion of a role for genetic and epigenetic etiological factors. Syndromic and nonsyndromic forms of POF have indeed been related to differing types of X chromosome and

Received February 28, 2015; revised April 7, 2015; accepted April 12, 2015; published online May 16, 2015.
D.J.F. has nothing to disclose. L.C.P. has nothing to disclose. Y.C.S. has nothing to disclose. A.d.J.R. has nothing to disclose. H.E.M. has nothing to disclose. K.M.J. has nothing to disclose. O.O.-R. has nothing to disclose. I.D.-Y. has nothing to disclose. P.L. has nothing to disclose.
D.J.F. and L.C.P. should be considered similar in author order.
This study was supported by the Universidad del Rosario (grant no. CS/ABN062/Genetics 015).
Reprint requests: Professor Paul Laissue, M.D., Ph.D., Unidad de Genética, Grupo GENIUROS, Escuela de Medicina y Ciencias de la Salud, Universidad del Rosario, Bogotá, Colombia (E-mail: paul.laissue@urosario.edu.co).

Fertility and Sterility® Vol. 104, No. 1, July 2015 0015-0282/$36.00

http://dx.doi.org/10.1016/j.fertnstert.2015.04.016

autosomal anomalies. For instance, it has been postulated that the hypergonadotropic hypogonadism observed in individuals with Turner's syndrome is due to the haploinsufficiency of genes located on the critical regions of the X chromosome that escape genetic inactivation (4). Fragile X syndrome has been associated with *FMR1* premutations and *FMR2* microdeletions (5). X chromosome deletions and translocations have led to identifying loci on the X chromosome (POF1-POF2 and POF3) that are related to the disease pathogenesis (6 and references therein).

Considerable efforts have been made during the last 20 years to identify sequence variants mainly located in POF coherent candidate genes' coding regions. However, less than 50 mutations have been validated in only nine genes (*FSHR*, *LHCGR*, *NR5A1*, *NOBOX*, *FOXL2*, *FIGLA*, *BMP15*, *NANOS3*, and *STAG3*) via functional assays as causative of nonsyndromic POF (7–20). Such apparent failure to identify POF's molecular origin has been partly due to the fact that very few family-related cases allow gene mapping by linkage analysis. In fact, POF etiological mutations are under strong negative selection because they participate in the molecular process responsible for reproductive fitness. Moreover, ovarian development and physiology imply fine-tuned quantitative molecular events that are regulated by hundreds of genes, thereby hampering the selection of coherent candidates for direct sequencing. It should be noted here that, although Sanger sequencing has been widely used for studying POF candidate gene-coding regions, it does suffer from technical limitations as it only allows up to ~700 base pairs (bp) to be analyzed per reaction. Studying large candidate genomic regions is therefore especially challenging since it involves numerous assays in setting up polymerase chain reaction (PCR) and sequencing experimental conditions. The next generation sequencing (NGS) approach (first used in 2007) has been widely used as an efficient alternative for detecting novel monogenic disease-related genes (21–23). This technique has been used more recently for studying, simultaneously, coding regions of thousands of genes (e.g., exome sequencing) that participate in the pathophysiology of multigene complex diseases (24–27).

The present study involved, for the first time, performing NGS assays on women affected by nonsyndromic POF. The complete coding regions of 70 candidate genes were sequenced in 12 women affected by POF. Bioinformatics and genetic analysis led to identifying mutations in *ADAMTS19* and *BMPR2* that are potentially related to POF origin. *LHCGR* mutations that might have contributed to the phenotype were also detected. NGS would thus seem to be a powerful tool for identifying new molecular actors involved in POF and for diagnostic/prognostic purposes.

MATERIALS AND METHODS

Patients and Controls

Patients and controls were of Colombian origin. They were attending the Clínica Marly's PMA Fertility Unit (Bogotá, Colombia) and/or the Genetics Unit at the Universidad del Rosario (Bogotá, Colombia). Twelve POF-affected women (POF3, -4, -7, -9, -10, -11, -13, -17, -19, -23, and -136) were included in the present study. All had at least 6 months of amenorrhea before age 40, high FSH plasma levels (>40 IU/mL), and a normal 46,XX karyotype. Women having a background of anticancer treatment, pelvic surgery, ovarian infection, and/or autoimmune disease were excluded from the study.

The control group (CG) included 176 women whose menopause had occurred after age 50 and had no antecedents regarding gynecological disease. These individuals reported having had regular menstrual cycles during their reproductive life spans and having had at least one child. A further 345 women (general population group [GP]; ages ranging from 19 to 63) from the same ethnic origin were also recruited to assess allele frequency for potentially deleterious alleles. Depending on availability, DNA samples were collected from family members of patients affected by potential pathogenic variants. These samples were used for studying candidate sequence variants potentially segregating with the phenotype. All of this study's clinical and experimental steps were approved by the Institutional Ethics Committee of both participating institutions. All individuals signed a written informed consent form.

NGS and Bioinformatics Analysis

Human genome sequences (Build Hg19) for the coding regions of the 70 genes of interest were retrieved from the UCSC Genome Browser (Table 1). These genes were selected after exploring public Internet databases, such as PubMed, Highwire, Geoprofiles, and MGI (Jackson Laboratory). For Internet exploration of the literature we included keywords, such as "genetics of sex determination," "genetics of folliculogenesis," "genetics of ovulation," "gene expression and ovary," "ovary and transcriptomics," "mouse models of premature ovarian failure," "premature ovarian failure," "premature ovarian failure genetics," and "primary ovarian insufficiency." In addition, we revised previous reports on gene sequencing involving women with POF to identify those displaying positive functional tests. The MGI database was use to specifically explore the POF phenotype. The Geoprofiles tool of the National Center for Biotechnology Information (NCBI) was used to evaluate expression patterns of selected genes.

Around 168,118 bases in length transcript sequences were flanked by 100 bp in the upstream and downstream regions, resulting in 247 kb target sequences. The 602 target coordinates were generated by merging the overlapping 706 exon coordinates in the selected genes' coding regions. These target sequences were masked for 3 kb length repeats, and a final set of 244 Kb sequences was considered for array design. Dense tiling of the target sequences was done to generate 60mer oligos regarding sense orientation. A final set of 241,399 probes was designed to fit into Agilent's 1X244K array, consisting of 241,399 experimental features and 2,105 control features. Total genomic DNA was isolated from leukocytes by use of the standard phenol/choloroform procedure. The DNA from the 12 POF samples was prepared and used for genome capture, as described elsewhere (28, 29). An Illumina GAII sequencer was used for high-throughput sequencing. The human

TABLE 1

Genes included in the sequencing microarray.

Ensembl ID	Gene name	Description
ENSG00000154734	*ADAMTS1*	A disintegrin and metalloproteinase with thrombospondin motifs 1 precursor
ENSG00000145808	*ADAMTS19*	A disintegrin and metalloproteinase with thrombospondin motifs 19 precursor
ENSG00000125378	*BMP4*	Bone morphogenetic protein 4 precursor
ENSG00000112175	*BMP5*	Bone morphogenetic protein 5 precursor
ENSG00000153162	*BMP6*	Bone morphogenetic protein 6 precursor
ENSG00000101144	*BMP7*	Bone morphogenetic protein 7 precursor
ENSG00000130385	*BMP15*	Bone morphogenetic protein 15 precursor
ENSG00000138696	*BMPR1B*	Bone morphogenetic protein receptor type-1B precursor
ENSG00000204217	*BMPR2*	Bone morphogenetic protein receptor type-2 precursor
ENSG00000111276	*CDKN1B*	Cyclin-dependent kinase inhibitor 1B
ENSG00000164442	*CITED2*	Cbp/p300-interacting transactivator 2
ENSG00000137869	*CYP19A1*	Cytochrome P450 19A1
ENSG00000100206	*DMC1*	Meiotic recombination protein DMC1/LIM15 homolog
ENSG00000091831	*ESR1*	Estrogen receptor
ENSG00000140009	*ESR2*	Estrogen receptor beta
ENSG00000183733	*FIGLA*	Factor in the germline alpha
ENSG00000102081	*FMR1*	Fragile X mental retardation 1 protein
ENSG00000183770	*FOXL2*	Forkhead box protein L2
ENSG00000118689	*FOXO3*	Forkhead box protein O3
ENSG00000184481	*FOXO4*	Forkhead box protein O4
ENSG00000131808	*FSHB*	Follitropin subunit beta precursor
ENSG00000170820	*FSHR*	Follicle-stimulating hormone receptor precursor
ENSG00000134363	*FST*	Follistatin precursor
ENSG00000174804	*FZD4*	Frizzled-4 precursor
ENSG00000164404	*GDF9*	Growth/differentiation factor 9 precursor
ENSG00000152661	*GJA1*	Gap junction alpha-1 protein
ENSG00000187513	*GJA4*	Gap junction alpha-4 protein
ENSG00000147437	*GNRH1*	Progonadoliberin-1 precursor
ENSG00000109163	*GNRHR*	Gonadotropin-releasing hormone receptor
ENSG00000123999	*INHA*	Inhibin alpha chain precursor
ENSG00000122641	*INHBA*	Inhibin beta A chain precursor
ENSG00000163083	*INHBB*	Inhibin beta B chain precursor
ENSG00000175189	*INHBC*	Inhibin beta C chain precursor
ENSG00000170498	*KISS1*	Metastasis-suppressor KiSS-1 precursor
ENSG00000116014	*KISS1R*	KiSS-1 receptor
ENSG00000157404	*KIT*	Mast/stem cell growth factor receptor precursor
ENSG00000049130	*KITLG*	Kit ligand precursor
ENSG00000116678	*LEPR*	Leptin receptor precursor
ENSG00000104826	*LHB*	Lutropin subunit beta precursor
ENSG00000138039	*LHCGR*	Lutropin-choriogonadotropic hormone receptor precursor
ENSG00000162624	*LHX8*	LIM/homeobox protein Lhx8
ENSG00000057468	*MSH4*	MutS protein homolog 4
ENSG00000204410	*MSH5*	MutS protein homolog 5
ENSG00000171487	*NALP5*	NACHT, LRR and PYD domains-containing protein 5
ENSG00000188425	*NANOS2*	Nanos homolog 2
ENSG00000187556	*NANOS3*	Nanos homolog 3
ENSG00000104320	*NBN*	Nibrin
ENSG00000106410	*NOBOX*	NOBOX_HUMAN Isoform 2
ENSG00000136931	*NR5A1*	Steroidogenic factor 1
ENSG00000167744	*NTF4*	Neurotrophin-4 precursor
ENSG00000082175	*PGR*	Progesterone receptor (PR)
ENSG00000101856	*PGRMC1*	Membrane-associated progesterone receptor component 1
ENSG00000121879	*PIK3CA*	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha isoform
ENSG00000127948	*POR*	NADPH-cytochrome P450 reductase
ENSG00000064835	*POU1F1*	Pituitary-specific positive transcription factor 1
ENSG00000204531	*POU5F1*	POU domain, class 5, transcription factor 1
ENSG00000172179	*PRL*	Prolactin precursor
ENSG00000113494	*PRLR*	Prolactin receptor precursor
ENSG00000175325	*PROP1*	Homeobox protein prophet of Pit-1
ENSG00000171862	*PTEN*	Phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase and dual-specificity protein phosphatase
ENSG00000125384	*PTGER2*	Prostaglandin E_2 receptor EP2 subtype
ENSG00000073756	*PTGS2*	Prostaglandin G/H synthase 2 precursor
ENSG00000163661	*PTX3*	Pentraxin-related protein PTX3 precursor
ENSG00000169218	*RSPO1*	R-spondin-1 precursor
ENSG00000170365	*SMAD1*	Mothers against decapentaplegic homolog 1
ENSG00000175387	*SMAD2*	Mothers against decapentaplegic homolog 2

Fonseca. NGS and potential new POF genes. Fertil Steril 2015.

TABLE 1

Continued.

Ensembl ID	Gene name	Description
ENSG00000166949	*SMAD3*	Mothers against decapentaplegic homolog 3
ENSG00000113658	*SMAD5*	Mothers against decapentaplegic homolog 5
ENSG00000142168	*SOD1*	Superoxide dismutase
ENSG00000147465	*STAR*	Steroidogenic acute regulatory protein, mitochondrial precursor

Note: Ensembl ID refers to the genomic sequence code from the Ensembl database (http://www.ensembl.org).

Fonseca. NGS and potential new POF genes. Fertil Steril 2015.

reference genome GRCh37/hg19 was used for mapping exome sequencing (Exome-Seq). The sequence database downloaded from the ensembl website (http://www.ensembl.org/Homo_sapiens/Info/Index) was used as our gene model and for determining amino acid substitutions. Raw data obtained from Illumina GAII and SeqQC (in-house tool) were used for checking data quality. Low quality and adaptor contamination were filtered using BLAT software (in-house tool). The high-quality filtered data were further aligned with the human reference genome using BWA 0.5.9rcl software (http://bio-bwa.sourceforge.net/), and variants were identified using samtools-0.1.7a (http://samtools.sourceforge.net/). The variants were then filtered on the basis of read depth and mapping quality. The filtered variations were annotated, along with their potential effect, by using snpEFF 3.4i software (http://snpeff.sourceforge.net/). The results were compared with dbSNP database data (NCBI version db137). This led to identifying novel variations and their effect. Annotated variations were further reviewed manually. Genotypic Technology did the exome sequencing and primary data analysis. Excel functions were used for subsequent filtering.

ADAMTS19 (gi|112789555), BMPR2 (gi|15451916), and LHCGR (gi|106067657) human wild type sequences were used for multiple protein alignment (ClustalW software) for studying conservation during the evolution of interchanged residues (candidate mutations). These sequences were aligned with those from distinct vertebrate species (Supplemental Table 1).

PolyPhen2 and SIFT bioinformatics tools were used for assessing potentially damaging effects. PolyPhen2 prediction values resulted from an algorithm using comparative analysis of protein sequences from different species, the exchanged amino acids physicochemical characteristics, and mapping residue replacement regarding available three-dimensional structures. The results are assessed as benign, probably damaging, or possibly damaging (30). The SIFT program uses amino acid hydrophobic features and conservation among species (protein sequence similarity) to establish the probability of deleterious effects caused by missense mutations (31). A score under 0.05 would argue in favor of pathogenicity.

Sanger Sequencing

ADAMTS19 (ENST00000274487) exon 19, *BMPR2* (ENST00000374580) exon 13, and *LHCGR* (ENST00000294954) exons 3 and 6 were amplified by PCR from relevant patients to confirm potentially deleterious candidate sequence variants obtained from NGS assays (see below and Table 2) by using exon-flanking primers. Amplicons were purified using shrimp alkaline phosphatase and exonuclease I, following the manufacturer's instructions. Direct sequencing was performed with internal primers using a capillary sequencer (Applied Biosystems). Primer sequences and PCR conditions are available upon request. An identical experimental setup was used for screening these sequence variants in control individuals and in the GP group. Familial segregation of *BMPR2* and *LHCGR* mutations found in the POF19 patient was screened in some members of her family (II:1, II:2, II:3, III:3, from Supplemental Fig. 1).

RESULTS

NGS and in Silico Analysis

NGS assays produced ~2.5 GB data for 12 samples for each individual as pair-end reads, having up to 97 bp mean read length, and about 90% (0.2 Mb length) of the targeted bases were covered, thereby sufficiently passing our thresholds for calling single nucleotide polymorphisms (SNPs) and short insertions or deletions (indels). Excel filtering of sequence variants revealed five heterozygous coding variants: *ADAMTS19* c.2828C→T (p.Thr943Ile), *LHCGR* c.296A→G (p.Asn99Ser), *LHCGR* c.526C→T (p.Ser176Pro), *BMPR2* c.2960C→T (p.Ser987Phe), and *NR5A1* c.1195G→T (p.Ala399Ser).

The POF7 patient had the ADAMTS19-Thr943Ile mutation, the POF19 patient had LHCGR-Asn99Ser and BMPR2-Ser987Phe mutations, and the POF23 and POF136 women were carriers of the LHCGR-Ser176Pro and NR5A1-Ala399Ser mutations, respectively.

Multiple protein alignment showed that all interchanged amino acids were conserved among vertebrate species (data not shown). SIFT and PolyPhen predictions have been included in Table 2.

Direct Sequencing of Candidate Mutations

The confirmation step for NGS results (via Sanger sequencing) revealed that four out of five variants (*ADAMTS19* c.2828C→T, *LHCGR* c.296A→G, *LHCGR* c.526C→T, and *BMPR2* c.2960C→T) were definitely present in the corresponding patients (Fig. 1 and Table 2). The *NR5A1* c.1195G→T sequence variant was not present in the POF136 patient, thus displaying an NGS artifact. The aforementioned positive variants were not found in the CG and GP groups. Therefore, the allele frequency for each variant was as follows: POF group, 1/24; CG, 0/352; GP group,

TABLE 2

Clinical features and molecular findings.

Patient	POF7	POF19	POF23
Clinical features			
Actual age	29	36	30
Amenorrhea	Secondary	Secondary	Secondary
FSH, IU/L	96.2	130	90.4
LH, IU/L	58.1	19.6	20.8
E_2 pg/mL	38	9.8	16
Molecular findings			
Gene	*ADAMTS 19*	*BMPR2* *LHCGR*	*LHCGR*
Sequence variation	c.2828C>T	*BMPR2*-c.2960C>T *LHCGR*-c.296A>G	c.526T>C
AA change	Thr943Ile	BMPR2-Ser987Phe LHCGR-Asn99Ser	Ser176Pro
Pathogenicity prediction			
SIFT	0,32	BMPR2-Ser987Phe: 0,00 LHCGR-Asn99Ser: 0,26	0,03
PolyPhen	Possibly damaging	BMPR2-Ser987Phe: Possibly damaging LHCGR-Asn99Ser: Bening	Bening

Note: Clinical features include details on reproductive conditions of POF patients presenting potential pathogenic mutations. Molecular and in silico results suggest deleterious effects of missense mutations. AA = amino acid.

Fonseca. NGS and potential new POF genes. Fertil Steril 2015.

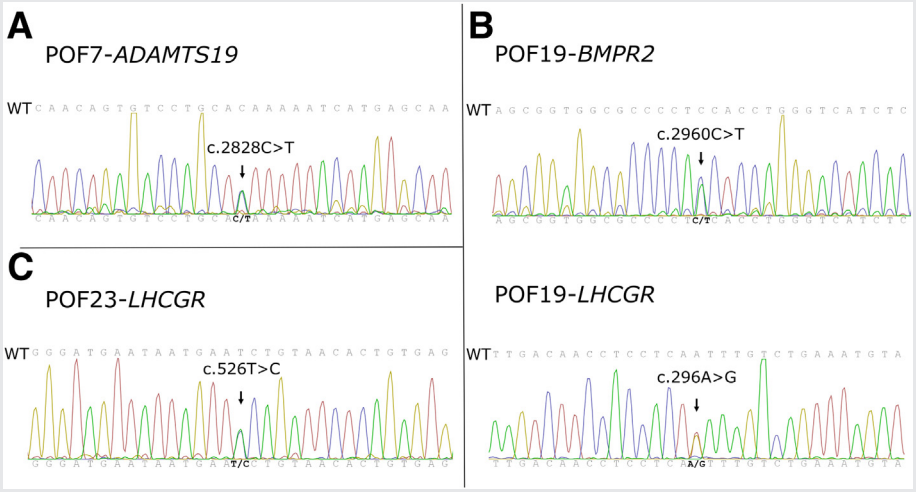
0/690. Family members of POF19 patient were carriers of the *BMPR2* c.2960A→G (II:3, III:3) or *LHCGR* c.296A→G (II:1, II:2) sequence variant. Only the POF19 individual (III:5) displayed both variants at the heterozygous state (Supplemental Fig. 1).

Patients' Phenotype

ADAMTS19- p.Thr943Ile mutation. The POF7 woman had experienced menarche at 12 years of age and had normal menstrual cycles (25 days) until the age of 28. Then she had irregular cycles for 1 year, rapidly evolving to amenorrhea. Transvaginal ultrasound reported normal uterus and ovaries. FSH, LH, and E_2 plasma levels at 29 years of age were, respectively, 96.2 IU/L, 58.1 IU/L, and 38 pg/mL (normal ranges: FSH, 4–13 IU/L; LH, 2–15 IU/L; E_2, 20–145 pg/mL). Antithyroid and anti-DNA serum antibodies were undetectable. She did not refer to a family background of fertility disorders.

BMPR2-Ser987Phe and LHCGR-Asn99Ser mutations. The POF19 patient was a 37-year-old woman who was affected by amenorrhea, hot flushes, and dizziness. Menarche had occurred at 11 years of age with regular menstrual cycles

FIGURE 1



Chromatograms of POF patients presenting *ADAMTS19*, *BMPR2*, and *LHCGR* mutations. (**A**) POF7 patient: *ADAMTS19*-c.2828C>T mutation. (**B**) POF19 patient: *BMPR2* c.2960C>T and *LHCGR*-c.296A>G. (**C**) POF23 patient: c.526T>C. WT = wild type sequences. *Arrows* indicate positions of heterozygous mutations.

Fonseca. NGS and potential new POF genes. Fertil Steril 2015.

(every 28 days). When the menses stopped, her FSH, LH and E_2 levels were, respectively, 130 IU/L, 19.6 IU/L, and 9.8 pg/mL. Antithyroid and anti-DNA serum antibodies were undetectable. She did not mention symptoms evoking cardiac or pulmonary dysfunction. Physical examination did not reveal any clinical signs of pulmonary hypertension. Chest radiograph and transthoracic echocardiogram gave normal results. No family background of hypofertility or further diseases was recorded.

LHCGR-p.Ser176Pro mutation. The patient carrying the LHCGR-p.Ser176Pro was a 30-year-old woman who presented with secondary amenorrhea. She had experienced menarche at the age of 12, with irregular menstrual cycles. However, she did become pregnant and had one child. She had no family background of reproductive disease. Transvaginal ultrasound did not reveal any morphological abnormalities of the reproductive tract. Paraclinical tests showed increased FSH plasma levels (90.4 IU/L), as well as low LH (20.8 IU/L) and E_2 levels (16 pg/mL). Antithyroid and anti-DNA serum antibodies were undetectable.

DISCUSSION

Human infertility can be considered to be a public health concern since up to 15% of couples are affected by this disorder. Distinct conditions, such as tubal dysfunction, endometriosis, and ovarian pathology, have been recorded as exclusively female etiological factors. It has been estimated that ~15% of women have dysfunctions leading to reduced ovarian reserve and, eventually, to POF (32). To date, mutations in only a few genes have been definitely related to POF pathogenesis, owing to Sanger sequencing's technical limitations regarding the polygenic nature of reproductive phenotypes.

To overcome such drawbacks in the present study, the complete coding regions of 70 candidate genes were massively sequenced, via NGS, in 12 patients with POF (Table 1). Several characteristics were considered when including specific genes in the sequencing microarray, such as previous reported tests validating a functional role, coherent spatiotemporal expression pattern, positive genetic association studies, and animal models having similar to POF phenotypes. These genes belong to molecular pathways that are involved in distinct steps regarding gonad development and ovarian physiology, such as sex determination, meiosis, folliculogenesis, and ovulation. It should be noted that, since we used a 244K microarray, our approach probably did not encompass all POF candidate genes (e.g., more than 350) (33–37). However, our results indicated that 25% (3/12) of patients (POF7, POF19, and POF23) carried mutations that potentially contribute to disease pathogenesis.

The POF7 patient, who was affected by secondary amenorrhea, was a carrier of the *ADAMTS19*-c.2828C→T heterozygous sequence variant (Fig. 1). Human *ADAMTS19*, which is located on chromosome 5, encodes a zinc metalloprotease from the 19-member ADAMTS (a disintegrin and metalloproteinase having thrombospondin motifs) family of proteins (38). At the structural level, ADAMTS factors have distinct domains, motifs, and modules (38–40). These include (from the N-ter to the C-Ter regions) a signal peptide, a prodomain, a metalloproteinase domain, a disintegrin-like domain, a central thrombospondin type I–like (TS) repeat, a cysteine-rich domain, a spacer region, and a C-terminal region having a variable number of TS repeats (39). ADAMTS19, like other members of the family, includes an additional region, the PLAC (protease and lacunin) domain, located in the protein's C-terminal region. It has been shown (by phylogenetic analysis) that ADAMTS17 is the closest related member to ADAMTS19 (38, 39, 41).

Although specific *ADAMTS19* expression has not been widely studied, it has been detected in an osteosarcoma cDNA library and in fetal lung (38). Low expression signals have been reported in developing heart, kidney, skeletal muscle, and testis (42, 43). *Adamts19* expression in mouse XX gonads has been observed at E10.5 and from E12.5 to E15.5 as well as in postnatal life (8 days postpartum) (42, 44). It is interesting that ovarian somatic cells expressing specific markers (e.g., *Follistatin* and *Foxl2*) also express *Adamts19* (42, 43). ADAMTS19's function remains unknown, but, similar to other family members, it might contribute to extracellular matrix remodeling by proteolytic activity on specific substrates, such as proteoglycans and fibrillar collagens (45). Concerning POF, it has been reported (using genome-wide SNP genotyping) that *ADAMTS19* might be associated with the phenotype (46). Furthermore, some *ADAMTS19* sequence variants seem to be associated with *ACVR2B* and *IGF2R* SNPs, which might confer increased susceptibility to POF (47, 48).

The *ADAMTS19* mutation (c.2828C→T, p.Thr943Ile mutation) found in our patient is novel, since it has not been described in public SNP databases and it was not present in 176 women from the CG. Furthermore, we could not find it in 345 women (690 alleles) from the GP, suggesting that it might be a rare causative variant. In silico analysis at the protein level showed that the interchanged threonine residue in position 943 was highly conserved among vertebrate species, thereby underlining a putative functional role. Regarding physicochemical properties, the presence of an isoleucine instead of a threonine is predicted to modify protein structure and hydrophobicity. In fact, Thr is a small, polar amino acid having hydrophilic properties, while Ile is a large, nonpolar hydrophobic residue. Such differences might alter protein folding and its function. Accordingly, PolyPhen2 predicted a deleterious effect ("possibly damaging" score). SIFT did not suggest a pathogenic consequence (score, 0.32), and it should be stressed that both PolyPhen2 and SIFT bioinformatics tools estimate only predictions regarding missense variants' potentially deleterious effects. PolyPhen2 and SIFT involve using some similar features, such as alignments methods, to propose potential deleterious effects. However, they automatically explore distinct Internet databases containing differing quality sequences for a specific protein. This leads, in some cases, to contradictory results, as well as to imprecise predictions. For instance, if, for a specific protein, a significant number of noncurated sequences (background noise) from other species are taken for analysis, it is highly probable that negative (harmless) predictions will be obtained for a particular missense mutation. On the contrary, if only a

few well-described sequences from closely related species (e.g., humans and chimpanzees) are used for calculations, then false-positive predictions could be found. Positive or negative results would certainly be considered as a part of a more comprehensive analysis including protein alignment, mutation localization in a protein's functional domains, and examining mutation-induced physicochemical changes. The p.Thr943Ile mutation is located in the first TS repeat in the protein's C-terminal region. Interestingly, it has been shown that the C-terminal region of some ADAMTS proteins (e.g., ADAMTS1, ADAMTS4) is necessary for both substrate specificity and localization (39 and references therein, 40). We thus hypothesized that the p.Thr943Ile mutation might alter ADAMTS19 substrate-binding properties and modify enzyme selectivity. The mutant protein could also display abnormal cell/tissue localization. Such features regarding an ovarian context might be linked to abnormal somatic cells' extracellular matrix turnover, leading to defects in follicular development and thus to POF. It would be interesting to perform in vitro functional tests to validate this hypothesis as well as to create pertinent *Adamts19* mouse models (e.g., knockout, knock-in).

The POF19 patient, who was affected by secondary amenorrhea, had two heterozygous coding variants in *BMPR2* and *LHCGR* genes (Fig. 1). *BMPR2* (bone morphogenetic protein receptor type II) encodes a serine-threonine kinase, which is responsible for binding bone morphogenetic proteins (BMP) (49, 50). BMP factors belong to the TGF-β superfamily of growth factors, which are involved in a wide variety of biological processes, including reproduction (51). After secretion to the extracellular space, BMP ligands bind to transmembrane type II receptors, which in turn transphosphorylate type I receptors (52). Type I receptor activation leads to SMAD protein phosphorylation and then to nuclear translocation of signaling for regulating target genes (53). BMP15 and GDF9 ligands, two key actors in ovarian physiology, have been related to BMPR2 activation in mice and humans (54–56). In these species, BMP15:GDF9 heterodimers (produced by oocytes) have higher biological activity than homodimers (BMP15:BMP15 or GDF9:GDF9). It has been shown that BMP15:GDF9 heterodimers act via the BMPR2-ALK4/5/7-ALK6 receptor complex located in granulosa cells (56). Since *BMP15* sequence variants have been related to POF etiology, we estimate that *BMPR2* coding mutations might also cause the disease (6, 17, 57). The heterozygous mutation found in our patient strongly suggested a deleterious effect since it was not present in our control population groups and it displays a minimal allele frequency $<.01$ (rs150642992). In silico analysis has shown that the BMPR2 p.Ser987Phe mutation affects a highly evolutionary conserved residue during evolution. SIFT and PolyPhen2 predictions agreed with such an assumption (SIFT, 0.00; PolyPhen2, probably damaging; Table 2). At the physicochemical level, the p.Ser987Phe mutation implies a drastic effect because serine is a small, polar amino acid, while phenylalanine is a large, aromatic residue. This mutation affects the transmembrane receptor's intracellular C-terminal tail region, which interacts with proteins involved in cellular features such as ion transport, transcription, cytoskeleton structure/modulation, and trafficking (58–60). Furthermore, this region might affect SMAD intracellular signaling, and we thus hypothesize that the BMPR2-p.Ser987Phe mutation in granulosa cells might interfere with BMP15 downstream signaling, thereby contributing to POF etiology. It should be noted that *BMPR2* encoding mutations (most being nonsense mutations) have been related to primary pulmonary hypertension (PPH), a disease having an autosomal dominant inheritance pattern (61–65). Our patient had no clinical signs or background of PPH, which might be explained by the disease's low penetrance (20%–30%) (64). Concerning individuals with PPH, an association with POF has not been reported to date, probably because it has not been systematically explored or reported.

POF19 patient is also affected by the heterozygous *LHCGR* c.296A→G (p.Asn99Ser) mutation. As for the ADAMTS19 and BMPR2 mutations described above, control population screening and bioinformatics studies suggested a potentially deleterious effect. This mutation, located in the protein's extracellular domain, is novel and affects an amino acid that is conserved during vertebrate species' evolution. *LHCGR* and *FSHR* were the first genes for which coding mutations were related to POF (7, 8). Only homozygous mutations in these genes in XX women have been definitely linked to POF etiology. However, this feature does not rule out heterozygous variants contributing to the phenotype. A recent report described an association of a heterozygote insertion of six nucleotides in *LHCGR* with amenorrhea and poor oocyte recovery during IVF (66). In our patient, both *ADAMTS19* and *LHCGR* mutations may have had additive pathogenic effects. Family segregation of mutant alleles would support such a hypothesis (Supplemental Fig. 1). Nevertheless, both mutations deserve functional tests to clarify their contribution to the phenotype.

POF23 was affected by the heterozygous LHCGR-p.Ser176Pro mutation (Fig. 1). As this mutation was absent from our control populations and had computational predictions (protein alignments; SIFT, 0.03) compatible with a deleterious effect, this variant might contribute to patient phenotype (Table 2).

CONCLUSIONS

Taken together, our results lead to mutations in two novel genes (*ADAMTS19* and *BMPR2*) being potentially related to POF pathogenesis. We have also proposed that heterozygous *LHCGR* mutations might contribute, with variants in further POF candidate genes, to premature menopause. These assumptions must be validated by further sequencing projects and functional tests. We thus recommend NGS as a powerful tool for identifying new molecular actors in POF and for future diagnostic/prognostic purposes.

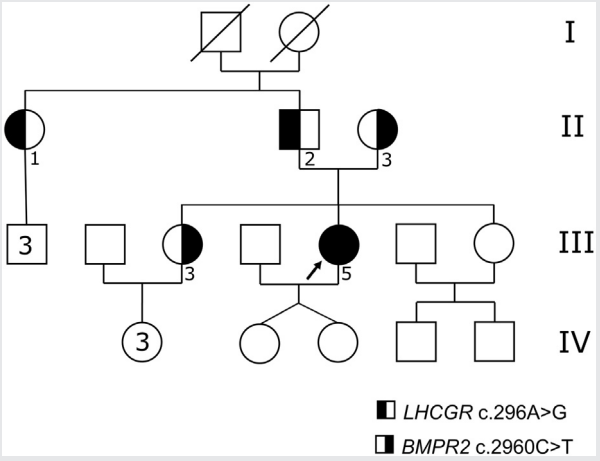
REFERENCES

1. Cox L, Liu JH. Primary ovarian insufficiency: an update. Int J Womens Health 2014;6:235–43.
2. Coulam CB. Premature gonadal failure. Fertil Steril 1982;38:645–55.

3. Timmreck LS, Reindollar RH. Contemporary issues in primary amenorrhea. Obstet Gynecol Clin North Am 2003;30:287–302.
4. Zinn AR, Ross JL. Turner syndrome and haploinsufficiency. Curr Opin Genet Dev 1998;8:322–7.
5. Sherman SL. Premature ovarian failure in the fragile X syndrome. Am J Med Genet 2000;97:189–94.
6. Laissue P, Vinci G, Veitia RA, Fellous M. Recent advances in the study of genes involved in non-syndromic premature ovarian failure. Mol Cell Endocrinol 2008;282:101–11.
7. Aittomäki K, Lucena JL, Pakarinen P, Sistonen P, Tapanainen J, Gromoll J, et al. Mutation in the follicle-stimulating hormone receptor gene causes hereditary hypergonadotropic ovarian failure. Cell 1995;82:959–68.
8. Latronico AC, Anasti J, Arnhold IJ, Rapaport R, Mendonca BB, Bloise W, et al. Brief report: testicular and ovarian resistance to luteinizing hormone caused by inactivating mutations of the luteinizing hormone-receptor gene. N Engl J Med 1996;334:507–12.
9. Beau I, Touraine P, Meduri G, Gougeon A, Desroches A, Matuchansky C, et al. A novel phenotype related to partial loss of function mutations of the follicle stimulating hormone receptor. J Clin Invest 1998;102:1352–9.
10. Touraine P, Beau I, Gougeon A, Meduri G, Desroches A, Pichard C, et al. New natural inactivating mutations of the follicle-stimulating hormone receptor: correlations between receptor function and phenotype. Mol Endocrinol 1999;13:1844–54.
11. Rannikko A, Pakarinen P, Manna PR, Beau I, Misrahi M, Aittomäki K, et al. Functional characterization of the human FSH receptor with an inactivating Ala189Val mutation. Mol Hum Reprod 2002;8:311–7.
12. Doherty E, Pakarinen P, Tiitinen A, Kiilavuori A, Huhtaniemi I, Forrest S, et al. A novel mutation in the FSH receptor inhibiting signal transduction and causing primary ovarian failure. J Clin Endocrinol Metab 2002;87:1151–5.
13. Di Pasquale E, Beck-Peccoz P, Persani L. Hypergonadotropic ovarian failure associated with an inherited mutation of human bone morphogenetic protein-15 (BMP15) gene. Am J Hum Genet 2004;75:106–11.
14. Qin Y, Choi Y, Zhao H, Simpson JL, Chen Z-J, Rajkovic A. NOBOX homeobox mutation causes premature ovarian failure. Am J Hum Genet 2007;81:576–81.
15. Zhao H, Chen Z-J, Qin Y, Shi Y, Wang S, Choi Y, et al. Transcription factor FIGLA is mutated in patients with premature ovarian failure. Am J Hum Genet 2008;82:1342–8.
16. 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.
17. Rossetti R, Di Pasquale E, Marozzi A, Bione S, Toniolo D, Grammatico P, et al. BMP15 mutations associated with primary ovarian insufficiency cause a defective production of bioactive protein. Hum Mutat 2009;30:804–10.
18. Wu X, Wang B, Dong Z, Zhou S, Liu Z, Shi G, et al. A NANOS3 mutation linked to protein degradation causes premature ovarian insufficiency. Cell Death Dis 2013;4:e825.
19. Caburet S, Arboleda VA, Llano E, Overbeek PA, Barbero JL, Oka K, et al. Mutant cohesin in premature ovarian failure. N Engl J Med 2014;370:943–9.
20. Santos MG, Machado AZ, Martins CN, Domenice S, Costa EM, Nishi MY, et al. Homozygous inactivating mutation in NANOS3 in two sisters with primary ovarian insufficiency. Biomed Res Int 2014;2014:787465.
21. Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, Smith SW, et al. Genome-wide in situ exon capture for selective resequencing. Nat Genet 2007;39:1522–7.
22. Albert TJ, Molla MN, Muzny DM, Nazareth L, Wheeler D, Song X, et al. Direct selection of human genomic loci by microarray hybridization. Nat Methods 2007;4:903–5.
23. Boycott KM, Vanstone MR, Bulman DE, MacKenzie AE. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. Nat Rev Genet 2013;14:681–91.
24. Kiezun A, Garimella K, Do R, Stitziel NO, Neale BM, McLaren PJ, et al. Exome sequencing and the genetic basis of complex traits. Nat Genet 2012;44:623–30.
25. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. Cell 2013;155:27–38.
26. Shyr D, Liu Q. Next generation sequencing in cancer research and clinical application. Biol Proced Online 2013;15:4.
27. Lange LA, Hu Y, Zhang H, Xue C, Schmidt EM, Tang Z-Z, et al. Whole-exome sequencing identifies rare and low-frequency coding variants associated with LDL cholesterol. Am J Hum Genet 2014;94:233–45.
28. Ortega-Recalde O, Vergara JI, Fonseca DJ, Ríos X, Mosquera H, Bermúdez OM, et al. Whole-exome sequencing enables rapid determination of xeroderma pigmentosum molecular etiology. PLoS One 2013;8:e64692.
29. Patiño LC, Battu R, Ortega-Recalde O, Nallathambi J, Anandula VR, Renukaradhya U, et al. Exome sequencing is an efficient tool for variant late-infantile neuronal ceroid lipofuscinosis molecular diagnosis. PLoS One 2014;9:3–7.
30. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–9.
31. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res 2003;31:3812–4.
32. Santoro N. Update in hyper- and hypogonadotropic amenorrhea. J Clin Endocrinol Metab 2011;96:3281–8.
33. Matzuk MM, Lamb DJ. Genetic dissection of mammalian fertility pathways. Nat Cell Biol 2002;4(Suppl):s41–9.
34. Roy A, Matzuk MM. Deconstructing mammalian reproduction: using knockouts to define fertility pathways. Reproduction 2006;131:207–19.
35. Matzuk MM, Lamb DJ. The biology of infertility: research advances and clinical challenges. Nat Med 2008;14:1197–213.
36. Jagarlamudi K, Reddy P, Adhikari D, Liu K. Genetically modified mouse models for premature ovarian failure (POF). Mol Cell Endocrinol 2010;315:1–10.
37. Jagarlamudi K, Rajkovic A. Oogenesis: transcriptional regulators and mouse models. Mol Cell Endocrinol 2012;356:31–9.
38. Cal S, Obaya AJ, Llamazares M, Garabaya C, Quesada V, López-Otín C. Cloning, expression analysis, and structural characterization of seven novel human ADAMTSs, a family of metalloproteinases with disintegrin and thrombospondin-1 domains. Gene 2002;283:49–62.
39. Porter S, Clark IM, Kevorkian L, Edwards DR. The ADAMTS metalloproteinases. Biochem J 2005;386:15–27.
40. Apte SS. A disintegrin-like and metalloprotease (reprolysin-type) with thrombospondin type 1 motif (ADAMTS) superfamily: functions and mechanisms. J Biol Chem 2009;284:31493–7.
41. Nicholson AC, Malik S-B, Logsdon JM, Van Meir EG. Functional evolution of ADAMTS genes: evidence from analyses of phylogeny and gene organization. BMC Evol Biol 2005;5:11.
42. Menke DB, Page DC. Sexually dimorphic gene expression in the developing mouse gonad. Gene Expr Patterns 2002;2:359–67.
43. McClelland KS, Wainwright EN, Bowles J, Koopman P. Rapid screening of gene function by systemic delivery of morpholino oligonucleotides to live mouse embryos. PLoS One 2015;10:e0114932.
44. Bouma GJ, Hart GT, Washburn LL, Recknagel AK, Eicher EM. Using real time RT-PCR analysis to determine multiple gene expression patterns during XX and XY mouse fetal gonad development. Gene Expr Patterns 2004;5:141–9.
45. Lu P, Takai K, Weaver VM, Werb Z. Extracellular matrix degradation and remodeling in development and disease. Cold Spring Harb Perspect Biol 2011;3.
46. Knauff EA, Franke L, van Es MA, van den Berg LH, van der Schouw YT, Laven JS, et al. Genome-wide association study in premature ovarian failure patients suggests ADAMTS19 as a possible candidate gene. Hum Reprod 2009;24:2372–8.
47. Pyun JA, Kim S, Cha DH, Kwack K. Epistasis between IGF2R and ADAMTS19 polymorphisms associates with premature ovarian failure. Hum Reprod 2013;28:3146–54.
48. Pyun JA, Kim S, Kwack K. Epistasis between polymorphisms in ACVR2B and ADAMTS19 is associated with premature ovarian failure. Menopause 2015;22:212–6.
49. Nohno T, Ishikawa T, Saito T, Hosokawa K, Noji S, Wolsing DH, et al. Identification of a human type II receptor for bone morphogenetic protein-4 that forms differential heteromeric complexes with bone morphogenetic protein type I receptors. J Biol Chem 1995;270:22522–6.

50. Liu F, Ventura F, Doody J, Massagué J. Human type II receptor for bone morphogenic proteins (BMPs): extension of the two-kinase receptor model to the BMPs. Mol Cell Biol 1995;15:3479–86.
51. Chang H, Brown CW, Matzuk MM. Genetic analysis of the mammalian transforming growth factor-beta superfamily. Endocr Rev 2002;23: 787–823.
52. Shimasaki S, Moore RK, Otsuka F, Erickson GF. The bone morphogenetic protein system in mammalian reproduction. Endocr Rev 2004;25:72–101.
53. Massagué J, Seoane J, Wotton D. Smad transcription factors. Genes Dev 2005;19:2783–810.
54. Moore RK, Otsuka F, Shimasaki S. Molecular basis of bone morphogenetic protein-15 signaling in granulosa cells. J Biol Chem 2003;278: 304–10.
55. Pulkki MM, Mottershead DG, Pasternack AH, Muggalla P, Ludlow H, van Dinther M, et al. A covalently dimerized recombinant human bone morphogenetic protein-15 variant identifies bone morphogenetic protein receptor type 1B as a key cell surface receptor on ovarian granulosa cells. Endocrinology 2012;153:1509–18.
56. Peng J, Li Q, Wigglesworth K, Rangarajan A, Kattamuri C, Peterson RT, et al. Growth differentiation factor 9: bone morphogenetic protein 15 heterodimers are potent regulators of ovarian functions. Proc Natl Acad Sci U S A 2013;110:E776–85.
57. Laissue P, Christin-Maitre S, Touraine P, Kuttenn F, Ritvos O, Aittomaki K, et al. Mutations and sequence variants in GDF9 and BMP15 in patients with premature ovarian failure. Eur J Endocrinol 2006;154:739–44.
58. Machado RD, Rudarakanchana N, Atkinson C, Flanagan JA, Harrison R, Morrell NW, et al. Functional interaction between BMPR-II and Tctex-1, a light chain of dynein, is isoform-specific and disrupted by mutations underlying primary pulmonary hypertension. Hum Mol Genet 2003;12:3277–86.
59. Foletta VC, Lim MA, Soosairaiah J, Kelly AP, Stanley EG, Shannon M, et al. Direct signaling by the BMP type II receptor via the cytoskeletal regulator LIMK1. J Cell Biol 2003;162:1089–98.
60. Hassel S, Eichner A, Yakymovych M, Hellman U, Knaus P, Souchelnytskyi S. Proteins associated with type II bone morphogenetic protein receptor (BMPR-II) and identified by two-dimensional gel electrophoresis and mass spectrometry. Proteomics 2004;4:1346–58.
61. Lane KB, Machado RD, Pauciulo MW, Thomson JR, Phillips JA, Loyd JE, et al. Heterozygous germline mutations in BMPR2, encoding a TGF-beta receptor, cause familial primary pulmonary hypertension. Nat Genet 2000;26:81–4.
62. Deng Z, Haghighi F, Helleby L, Vanterpool K, Horn EM, Barst RJ, et al. Fine mapping of PPH1, a gene for familial primary pulmonary hypertension, to a 3-cM region on chromosome 2q33. Am J Respir Crit Care Med 2000; 161:1055–9.
63. Machado RD, Aldred MA, James V, Harrison RE, Patel B, Schwalbe EC, et al. Mutations of the TGF-beta type II receptor BMPR2 in pulmonary arterial hypertension. Hum Mutat 2006;27:121–32.
64. Li W, Dunmore BJ, Morrell NW. Bone morphogenetic protein type II receptor mutations causing protein misfolding in heritable pulmonary arterial hypertension. Proc Am Thorac Soc 2010;7:395–8.
65. Ma L, Chung WK. The genetic basis of pulmonary arterial hypertension. Hum Genet 2014;133:471–9.
66. Bentov Y, Kenigsberg S, Casper RF. A novel luteinizing hormone/chorionic gonadotropin receptor mutation associated with amenorrhea, low oocyte yield, and recurrent pregnancy loss. Fertil Steril 2012;97:1165–8.

SUPPLEMENTAL FIGURE 1



Pedigree of POF19 family. *Completely dashed symbols* refer to the affected individual (POF19) carrying both *BMPR2* and *LHCGR* mutations. *Half dashed symbols* refer to heterozygous carriers (*BMPR2* or *LHCGR* mutations).

Fonseca. NGS and potential new POF genes. Fertil Steril 2015.

SUPPLEMENTAL TABLE 1

Sequences used for multiple protein alignments.

Protein	Species	Protein sequence (GI)
ADAMTS19	*Homo sapiens*	112789555
	Bos taurus	194668875
	Bison bison	742173366
	Pygoscelis adeliae	690436459
	Gallus gallus	513231250
	Xenopus tropicalis	512817668
	Rattus norvegicus	564393555
	Mus musculus	28202057
BMPR2	*Homo sapiens*	15451916
	Macaca mulatta	383409919
	Bos taurus	296490468
	Sus scrofa	325974528
	Gallus gallus	47825379
	Xenopus laevis	147901570
	Rattus norvegicus	281371326
	Mus musculus	6680804
LHCGR	*Homo sapiens*	106067657
	Bos taurus	41386703
	Sus scrofa	47523950
	Gallus gallus	45384388
	Xenopus tropicalis	512844990
	Rattus norvegicus	6981160
	Mus musculus	7305233
	Danio Rerio	190336595

Note: GI refers to the protein code from the NCBI database (http://www.ncbi.nlm.nih.gov/protein/).

Fonseca. NGS and potential new POF genes. Fertil Steril 2015.