Journal of the Endocrine Society, 2022, **6**, 1–10
https://doi.org/10.1210/jendso/bvac106
Advance access publication 05 July 2022
Clinical Research Article

High-throughput Sequencing to Identify Monogenic Etiologies in a Preselected Polycystic Ovary Syndrome Cohort

Raiane P. Crespo,[1] Thais P. Rocha,[1] Luciana R. Montenegro,[1,2] Mirian Y. Nishi,[1,2] Alexander A.L. Jorge,[3] Gustavo A.R. Maciel,[4] Edmund Baracat,[4] Ana Claudia Latronico,[1] Berenice B. Mendonca,[1,2] and Larissa G. Gomes[1]

[1]Unidade de Endocrinologia do Desenvolvimento, Laboratório de Hormônios e Genética Molecular/LIM42, Hospital das Clínicas, Disciplina de Endocrinologia e Metabologia, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil
[2]Laboratório de Sequenciamento em Larga Escala (SELA), Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil
[3]Unidade de Endocrinologia Genética (LIM 25), Hospital das Clínicas, Disciplina de Endocrinologia, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil
[4]Disciplina de Ginecologia, Faculdade de Medicina da Universidade de São Paulo, Brazil

Correspondence: Larissa Garcia Gomes. Avenida Doutor Eneas de Carvalho Aguiar, 155, 2º andar, Bloco 6, CEP 05403-900, São Paulo, SP, Brazil. Email: larissa.gomes@hc.fm.usp.br

Abstract

Context: Polycystic ovary syndrome (PCOS) etiology remains to be elucidated, but familial clustering and twin studies have shown a strong heritable component.

Objective: The purpose of this study was to identify rare genetic variants that are associated with the etiology of PCOS in a preselected cohort.

Methods: This prospective study was conducted among a selected group of women with PCOS. The study's inclusion criteria were patients with PCOS diagnosed by the Rotterdam criteria with the following phenotypes: severe insulin resistance (IR), normoandrogenic–normometabolic phenotype, adrenal hyperandrogenism, primary amenorrhea, and familial PCOS. Forty-five patients were studied by target sequencing, while 8 familial cases were studied by whole exome sequencing.

Results: Patients were grouped according to the inclusion criteria with the following distribution: 22 (41.5%) with severe IR, 13 (24.5%) with adrenal hyperandrogenism, 7 (13.2%) with normoandrogenic phenotype, 3 (5.7%) with primary amenorrhea, and 8 (15.1%) familial cases. DNA sequencing analysis identified 1 pathogenic variant in *LMNA*, 3 likely pathogenic variants in *INSR*, *PIK3R1*, and *DLK1*, and 6 variants of uncertain significance level with interesting biologic rationale in 5 genes (*LMNA*, *GATA4*, *NR5A1*, *BMP15*, and *FSHR*). *LMNA* was the most prevalent affected gene in this cohort (3 variants).

Conclusion: Several rare variants in genes related to IR were identified in women with PCOS. Although IR is a common feature of PCOS, patients with extreme or atypical phenotype should be carefully evaluated to rule out monogenic conditions.

Key Words: polycystic ovary syndrome, genetics, insulin resistance, target sequencing, exome sequencing

Abbreviations: ACGM, American College of Medical Genetics and Genomics; BMI, body mass index; DHEAS, dehydroepiandrosterone sulfate; FSH, follicle stimulating hormone; GWAS, genome-wide association studies; HbA1C, glycated hemoglobin; HTS, high-throughput sequencing; IR, insulin resistance; OGTT, oral glucose tolerance test; PCOS, polycystic ovary syndrome; POI, premature ovarian insufficiency; SHORT, short stature, joint hyperextensibility, ocular depression, Rieger anomaly (developmental defect in the iris), and teething delay; VUSp, variants of uncertain significance with potential impact at the protein level; WES, whole exome sequencing.

Polycystic ovary syndrome (PCOS) is a common disorder affecting 5% to 20% of women of reproductive age worldwide [1-3]. It is characterized by hyperandrogenism, ovulatory dysfunction, and polycystic ovary morphology [1-3]. The clinical presentation is heterogeneous varying from the normoandrogenic–normometabolic profile to the classical PCOS phenotype characterized by hyperandrogenism and metabolic abnormalities [3]. The broad clinical spectrum probably reflects different pathophysiological components of the syndrome and different genetic backgrounds.

Different pathophysiological components such as dysfunction in ovarian steroidogenesis and folliculogenesis, insulin resistance (IR), altered gonadotropin secretion, and adrenal hyperactivity could vary in intensity and interact with each other, perpetuating or enhancing PCOS clinical features [3]. Notably, genome-wide association studies (GWAS) in PCOS cohorts have mapped PCOS susceptibility loci and identified candidate genes, especially those related to ovarian androgen biosynthesis (*DENND1A*, *GATA4*), IR (*INSR*), and gonadotropin secretion and action (*LHCGR*, *FSHR*, *FSHB*), indicating the role of different genetic mechanisms in this disease [4-9].

Twin studies have estimated that as much as 70% to 80% of PCOS risk may be explained by genetic factors [10].

Received: 5 January 2022. **Editorial Decision:** 1 July 2022. **Corrected and Typeset:** 25 July 2022
© The Author(s) 2022. Published by Oxford University Press on behalf of the Endocrine Society.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

However, the GWAS common susceptibility loci account for a small proportion of the estimated genetic heritability of PCOS [11]. This missing heritability could be caused by the presence of rare variants with larger biological effects that are poorly detected by GWAS [12]. Then, the advent of high-throughput sequencing (HTS) has allowed the simultaneous genotyping of several regions in the genome making the identification of rare pathogenic variants possible [13].

The search for causal rare genetic variants may be enriched by using 2 established strategies: studying individuals at the extremes of the phenotype distribution or by incorporating families with multiple affected individuals [14].

Therefore, the aim of our study was to search for rare genetic coding variants in genes related to the main pathophysiological components of PCOS, through a target panel, in women with PCOS in the extreme of the phenotype distribution and/or presenting an atypical noncommon characteristic of the syndrome such as severe adrenal hyperandrogenism and/or primary amenorrhea. Familial cases, with multiple affected individuals, were studied through whole exome sequencing (WES).

Materials and Methods

Study Design

This was a prospective cohort study.

Study Population

The cohort was selected from women with PCOS who were routinely attending the Endocrinology clinic at our tertiary center, Hospital das Clinical da Faculdade de Medicina da Universidade de São Paulo, from 2015 to 2019. PCOS diagnosis was confirmed according to the Rotterdam criteria, characterized by the presence of at least 2 out of 3 criteria: clinical and/or biochemical androgen excess, ovulatory dysfunctions, and polycystic ovarian morphology; and by the exclusion of differential diagnosis; this included the evaluation of thyroid disease, hyperprolactinemia, androgen-secreting tumors, and nonclassic 21-hydroxylase deficiency. Values of 2 early morning 17-hydroxyprogesterone measurements, collected in the follicular phase and performed using liquid chromatography—tandem mass spectrometry, under 2 ng/mL excluded this condition [15]. Subsequently, patients who met 1 or more of the following criteria, severe IR characterized by severe hyperinsulinemia (basal insulin >50 mUi/mL and/or after oral glucose tolerance test [OGTT] insulin >300 mUI/mL) or early-onset diabetes mellitus (type 2 diabetes in patients under 30 years of age), severe adrenal hyperandrogenism (dehydroepiandrosterone sulfate [DHEAS] levels at least 1.5 higher than the upper limit of the normal range), primary amenorrhea, or normoandrogenic–normometabolic phenotype, were selected for target HTS. Patients in the normoandrogenic–normometabolic phenotype group had only 2 PCOS Rotterdam diagnostic criteria, ovulatory dysfunctions, and polycystic ovarian morphology, with no clinical signs of hyperandrogenism (hirsutism, acne, or alopecia) and with normal serum androgen levels, and normal glycemic and lipid parameters. The cut-off levels for the severe hyperinsulinemia group were established in order to select individuals with hyperinsulinemia of at least 3 times the average expected for the population [16].

For cases who fulfilled more than 1 of the selection criteria, we prioritized allocation by laboratory characteristics (severe IR, adrenal, and/or normoandrogenic), followed by clinical characteristics (primary amenorrhea), because the former consisted of a larger and more representative group of patients in contrast to primary amenorrhea.

Selected familial cases had at least 2 members affected by PCOS, with complete clinical, biochemical, and endocrine testing performed by our team of endocrinologists.

Since our focus was identifying rare genetic variants, patients with additional medical conditions or syndromic features were included in the study: 1 patient with muscular dystrophy and 3 patients with intellectual disability were included.

Data Collection

The initial clinical and anthropometric evaluation included clinical signs of hyperandrogenism (hirsutism, female pattern androgenic alopecia, and acne), weight, height, body mass index (BMI), and blood pressure. Hirsutism was defined as a modified Ferriman–Gallwey score of ≥8 [17]. Androgenic alopecia and acne were classified as present or absent. Acanthosis nigricans and signs of virilization were consistently evaluated. Biochemical evaluation included 8 hours of fasting glucose, insulin, glycated hemoglobin (HbA1C), lipid profile, and OGTT. Hormonal profiles were evaluated during the follicular phase, without any hormonal contraceptive for at least 3 months and included 8 hours of fasting luteinizing hormone, follicle-stimulating hormone (FSH), total testosterone, free testosterone, sex hormone–binding globulin, DHEAS, and androstenedione. Ovarian morphology was evaluated by transvaginal ultrasound, and polycystic ovary morphology was defined according to previously published criteria [18].

Hormonal Measurements

Luteinizing hormone and FSH were measured by immunofluorometric assay (AutoDELFIA, Turku, Finland). Total testosterone levels and DHEAS were measured using an electrochemiluminescence assay (Cobas e601, Roche Diagnostics, Indianapolis, IN). Androstenedione was measured using liquid chromatography–mass spectrometry. Free testosterone was calculated from total testosterone, and sex hormone–binding globulin was determined by immunoassay. Age- and method-specific reference ranges were used to define biochemical hyperandrogenism.

High-throughput Sequencing

The DNA from peripheral blood leukocytes was extracted according to a standardized procedure adapted from a previously described method [19]. Cases with nonfamilial or familial PCOS whose familial DNA samples were unavailable (n = 45) were studied by target panel sequencing, and familial cases (n = 8) by WES. The panel was composed of genes related to ovarian folliculogenesis, gonadotropin action, steroidogenesis, and the insulin-signaling pathway (Table 1). Target panel and WES were performed according to previously published protocols [20]. The libraries were constructed with the SureSelect Target Enrichment System (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. The sequences were generated in the Illumina NextSeq 500 for target panel sequencing, and

Table 1. Genes included in the target gene panel classified according to their functions.

Ovarian folliculogenesis	Gonadotropin action	Steroid hormone synthesis	Insulin signaling pathway
FOXO3	*AMHRII*	*CYP11A1*	*INSR*
GDF9	*SRD5A2*	*HSD3B2*	*POMC*
AMH	*AR*	*POR*	*PPARG*
BMP15	*AKR1C3*	*CYP17A1*	*LMNA*
BMP2	*DLK1*	*CYP19A1*	*FST*
BMP4	*PGR*	*H6PD*	*AKT2*
	ESR1	*HSD11B1*	*TNFRSF1B*
	FSHR	*SULT2A1*	*LMNB2*
	LHCGR	*PAPSS2*	*IRS2*
		CYP21A2	*IRS1*
		NR3C1	*GHR*
		GATA4	*CIDEC*
		GATA6	*BSCL2*
			GCK
			LEPR

Table 2. In silico predictions of missense variants

Gene	cDNA	Protein	In silico prediction					
			Mutation taster[a]	Mutation assessor[b]	FATHMM[c]	SIFT[d]	Polyphen[e]	CADD[f]
LMNA	c.746G>A	p.Arg249Gln	D	M	D	D	D	34
LMNA	c.1912G>A	p.Gly638Arg	D	N	D	D	D	27.2
LMNA	c.1930C>T	p.Arg644Cys	D	N	D	D	D	34
FSHR	c.847C>T	p.Arg283Trp	D	M	D	D	B	27.2
PIK3R1	c.78G>T	p.Leu26Phe	D	N	D	T	P	16.2
GATA4	c.793C>T	p.Arg265Cys	D	M	D	D	D	35
NR5A1	c.386C>T	p.Pro129Leu	A	M	D	D	B	12
INSR	c.3568T>C	p.Tyr1190His	D	M	D	D	D	28.9
BMP15	c.202C>T	p.Arg68Trp	A	M	D	D	D	24.1

[a]Mutation taster prediction: D (disease causing—ie, probably deleterious), A (disease causing automatic—ie, known to be deleterious).
[b]Mutation assessor prediction: L (low), M (moderate), H (high), N (neutral).
[c]FATHMM prediction: T, tolerated; D, deleterious.
[d]SIFT output prediction: D, damaging; T, tolerated.
[e]PolyPhen prediction: B, benign; D, probably damaging; P, possibly damaging.
[f]CADD: considered pathogenic if ≥ 15.

HiSEQ 2500 (Illumina, Inc, San Diego, CA) platforms for WES, both running on the paired-end mode. The sequences were aligned with the human reference assembly (GRCh37/hg19).

Data analysis

HTS data were screened for rare variants (minor allele frequency of 1%) in public databases such as gnomAD v2.1.1 (http://gnomad.broadinstitute.org/), 1000 genome, and ABraOM (http://abraom.ib.usp.br/) [21], and also in-house databases [22], located in exonic regions and consensus splice site sequences. Subsequently, variant filtration prioritized genes based on their potential pathogenicity: loss-of-function variants (stop-gain, splice site disrupting, and frameshift variants) and missense variants predicted to be pathogenic by multiple in silico programs (Mutation taster, Mutation assessor, FATHMM, SIFT, PolyPhen, CADD) (Table 2). For variants identified by WES, we selected variants that fit based on different modes of inheritance (de novo, autosomal dominant, autosomal recessive, and X-linked). The sequencing reads carrying candidate variants were inspected visually using the Integrative Genomics Viewer to decrease false-positive calls. The assessment of gene function was performed using the OMIM, VarElect, and the PubMed databases. Sanger sequencing for segregation was performed to identify candidate variants in family members. All variants were classified according to the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology variant pathogenicity guidelines [23].

Statistical Analysis

Comparisons of quantitative variables among groups were performed by analysis of variance or the Kruskal–Wallis test for numerical variables as appropriate using software R (https://www.r-project.org/) and IBM Statistical Package for the Social Sciences—version 25. Correlations between

variables were determined using the Spearman correlation coefficient. The level of statistical significance was set at $P < .05$.

Ethical Consideration

Ethics approval to conduct the study was obtained from the Ethics Committee of Hospital das Clinicas, Faculty of Medicine, Sao Paulo University, Brazil. Patients and/or family members provided written informed consent (CAPPesq 15688) before participating in the study. Written informed consent for publication of their clinical details and clinical images was obtained from the patients.

Results

Description of Cohort

We selected 53 women with PCOS who met the inclusion criteria: 22 (41.5%) with severe IR, of whom 17 had severe hyperinsulinemia and 5 early-onset diabetes mellitus, 13 (24.5%) with adrenal hyperandrogenism, 7 (13.2%) with normoandrogenic-normometabolic phenotype, 3 (5.7%) with primary amenorrhea, and 8 (15.1%) familial cases. Primary amenorrhea was a very rare finding in our cohort and comparisons involving this group were not possible due to the small sample size.

Table 3 shows the wide spectrum of clinical and metabolic characteristics of the selected patients. The mean age of initial evaluation was 26 ± 8 years. Biochemical hyperandrogenism was present in this cohort, with mean testosterone levels elevated in all groups, except in the normoandrogenic group. As expected, in the severe adrenal hyperandrogenism group, the DHEAS levels were at least twice higher than the other groups. Compared with the adrenal and normoandrogenic groups, patients in the severe IR group presented an unfavorable metabolic profile with significantly higher BMI, 120-minute glucose, HbA1c, 120-minute insulin, and triglycerides (Table 3). This group also had lower high-density lipoprotein levels than the normoandrogenic group (Table 3).

Genetic analysis by HTS identified rare molecular variants that are potentially associated with the phenotype in 10 out of 53 cases (18.8%) (see Table 4). According to the ACMG classification, this analysis identified 1 pathogenic variant in *LMNA*, 3 likely pathogenic variants in *INSR*, *PIK3R1*, and *DLK1*, and 6 variants of uncertain significance with potential impact at the protein level (VUSp) in 5 genes (*LMNA*, *GATA4*, *NR5A1*, *BMP15*, and *FSHR*). The variants in *DLK1* and *NR5A1* were identified by exome sequencing and the others by a target gene sequencing panel.

Clinical Features of PCOS Patients With Candidate Pathogenic Variants and Discussion About The Genes

Variants associated with insulin signaling and glucose metabolism

INSR. A heterozygous missense variant in *INSR* (c.T3568C; p.Tyr1190His) was identified by target sequencing in a patient with PCOS and severe IR. The patient presented with normal BMI (23.9 kg/m^2), hirsutism (Ferriman score of 24), and acanthosis nigricans. The hormonal profile revealed marked hyperandrogenism (total testosterone of 296 ng/dL, reference range < 49 ng/dL) and significant hyperinsulinemia on OGTT (120 minutes of insulin > 1000 µU/mL) with normal glucose and HbA1c levels. Pelvic magnetic resonance imaging detected enlarged ovaries (34 mL and 20 mL) (Fig. 1A).

Familial segregation analysis revealed that the mother and 2 sisters were also heterozygous for the same rare variant. They were asymptomatic but presented with hyperinsulinemia on OGTT (Fig. 1B and 1C).

The p.Tyr1190His variant is very rare in public databases and it is predicted to be pathogenic by all evaluated in silico prediction sites. It is located at the tyrosine kinase site of the insulin receptor protein, which increases its potential for pathogenicity.

LMNA. Three heterozygous *LMNA* variants were identified. The variant c.G746A (p.Arg249Gln) was found in a patient with PCOS, muscular dystrophy, and severe IR. This pathogenic variant was previously associated with muscular dystrophy and functional studies confirmed its pathogenicity [24].

Additionally, another 2 variants, classified as VUSp, were identified in PCOS patients with IR and/or familial component. The variant c.C1930T (p.Arg644Cys) was identified in a patient with precocious diabetes mellitus and family history of arrhythmia and cardiomyopathy (her father used an external defibrillator). Unfortunately, the patient was lost to follow-up and cardiologic studies could not be done. This variant c.C1930T was frequently reported (49 times) in the UMD-LMNA mutations database (http://www.umd.be/LMNA/) and was associated with multiple laminopathy phenotypes, varying from asymptomatic to lipodystrophy and cardiomyopathy.

The other *LMNA* variant, c.G1912A (p.Gly638Arg), was identified in a slightly overweight patient with glucose intolerance and family history of PCOS. This variant c.1912A has no reports in the UMD-LMNA mutations database.

PIK3R1. The heterozygous missense variant in the *PIK3R1* gene (c.G78T; p.Leu26Phe) was identified in a patient with severe hyperinsulinemia, obesity, hypercholesterolemia, and central precocious puberty. The patient's sister also had the same phenotype (Fig. 2A and 2B). Sanger sequencing confirmed the same variant in the affected sister. The *PIK3R1* variant c.G78T is absent in all public databases and is located at an important functional domain of the protein PIK3R1, which plays a critical role in PI3K recruitment to phosphorylated tyrosine kinase receptors and insulin receptor substrate proteins, a crucial step in the insulin pathway [25]. The ACMG classified this variant as likely pathogenic.

DLK1. The heterozygous frameshift variant (c.594_594delC; p.Gly199Alafs*11) in the *DLK1* gene was identified by WES in 2 siblings with severe IR. Both sisters exhibited truncal obesity, precocious type 2 diabetes mellitus, hepatic steatosis, and central precocious puberty. This family had been reported previously [26].

Table 5 summarizes clinical features of the patients with variants related to IR.

Variants in genes related to adrenal hyperandrogenism

GATA 4. A heterozygous missense variant in *GATA4* (c.C793T; p.Arg265Cys) was found by target sequencing in a patient with adrenal hyperandrogenism (DHEAS 8800 ng/mL; reference range < 4070 ng/mL). The patient presented

Table 3. Description of the metabolic and hormonal features of the cohort according to their clinical phenotype, and comparisons among groups

	Adrenal	Primary Amenorrhea	Severe IR	Familial	Normoandrogenic	*P*	*P*[1]	*P*[2]
Number of patients	13	3	22	8	7			
Menarche (years)	12.0 ± 2.0	—	12.0 ± 2.0	12.0 ± 1.0	12.0 ± 1.0			
Age (years)	25.0 ± 10.0	25.0 ± 5.0	29.0 ± 8.0	24.0 ± 5.0	25.0 ± 8.0			
BMI (kg/m^2)	27.0 ± 3.7	30.4 ± 8.6	33.8 ± 6.8	27.9 ± 3.3	25.5 ± 4.9	.007	.023	.045
Total testosterone (ng/dL)	75.3 ± 24.9	48.3 ± 24.1	92.7 ± 62.3	51.2 ± 25.5	21.7 ± 9.6	<.001	>.05	<.001
Free testosterone (pmol/L)	49.3 ± 24.5	27.7 ± 19.5	62.2 ± 40.1	34.2 ± 26.0	11.3 ± 7.5	<.001	>.05	<.001
DHEAS (ng/mL)	6164.4 ± 1675.2	1897.3 ± 807.6	2018.9 ± 1275.2	2851.7 ± 1144.8	1355.3 ± 671.5	<.001	<.001	<.001
Androstenedione (ng/mL)	2.6 ± 1.5	1.8 ± 0.9	2.0 ± 0.7	1.4 ± 0.8	1.2 ± 0.6	ns		
HbA1c (%)	5.7 ± 0.5	5.1 ± 0.3	6.6 ± 1.9	5.2 ± 0.4	5.0 ± 0.2	<.001	.045	<.001
0-minute glucose (mg/dL)	88.0 ± 7.0	83.0 ± 3.0	102.0 ± 42.0	88.0 ± 10.0	84.0 ± 8.0	ns		
120-minute glucose (mg/dL)	122.0 ± 35.0	105.0 ± 22.0	177.0 ± 44.0	141.0 ± 23.0	100.0 ± 14.0	<.001	.008	<.001
0-minute insulin (μUI/mL)	23.2 ± 15.6	13.9 ± 4.9	41.2 ± 19.7	15.0 ± 3.2	11.6 ± 3.4	.001	>.05	.003
120-minute insulin (μUI/mL)	141.3 ± 99.1	91.4 ± 13.4	600.1 ± 281.1	153.3 ± 96.1	67.1 ± 31.0	.001	.007	.002
Cholesterol (mg/dL)	177.0 ± 39.0	173.0 ± 17.0	192.0 ± 37.0	194.0 ± 52.0	192.0 ± 41.0	ns		
Low-density lipoprotein levels (mg/dL)	102.0 ± 30.0	90.0 ± 10.0	118.0 ± 35.0	110.0 ± 33.0	108.0 ± 42.0	ns		
High-density lipoprotein levels (mg/dL)	51.0 ± 13.0	70.0 ± 27.0	40.0 ± 9.0	55.0 ± 11.0	57.0 ± 15.0	.03	>.05	.03
Triglycerides (mg/dL)	106.0 ± 58.0	86.0 ± 29.0	214.0 ± 130.0	175.0 ± 162.0	155.0 ± 143.0	.029	.05	>.05
Systolic blood pressure (mmHg	119.0 ± 17.0	108.0 ± 10.0	124.0 ± 13.0	114.0 ± 13.0	103.0 ± 15.0	ns		
Diastolic blood pressure (mmHg	82.0 ± 15.0	77.0 ± 6.0	83.0 ± 10.0	77.0 ± 9.0	68.0 ± 5.0	.046	>.05	.034
Waist circumference (cm)	91.9 ± 14.0	96.0 ± 13.3	106.3 ± 12.9	91.8 ± 11.8	86.1 ± 6.2	.005	>.05	.011
Ferriman score	9.0 ± 7.0	11.0 ± 9.0	13.0 ± 7.0	10.0 ± 7.0	2.0 ± 1.0	.009	>.05	.003

Abbreviations: BMI, body mass index; *P*, comparison among all groups; *P*[1], comparison between severe IR vs adrenal group; *P*[2], comparison between severe IR vs normoandrogenic phenotype group.

Table 4. Rare genetic variants identified by HTS in 10 PCOS patients

Group	Gene	Genomic coordinate (hg19)	cDNA	Protein	ID	Prevalence		ACMG	
						GnomAD	AbraOM	Criteria	Classification
Severe IR	LMNA	Chr1: 156104702	c.746G>A	p.Arg249Gln	rs59332535	0	0	PS3, PM2, PM1, PP2, PP3, PP4	Pathogenic
Familial	LMNA	Chr1: 156108492	c.1912G>A	p.Gly638Arg	rs144851946	0.02%	0.08%	PP2, PP3, PM1	VUSp
Severe IR	LMNA	Chr1: 156108510	c.1930C>T	p.Arg644Cys	rs142000963	0.12%	0.08%	PP2, PP3, PP5, PM1	VUSp
Severe IR	FSHR	Chr2: 49195844	c.847C>T	p.Arg283Trp	rs200328782	0.006%	0	PP3	VUSp
Severe IR	PIK3R1	Chr5: 67522581	c.78G>T	p.Leu26Phe	NA	0	0	PM1, PM2, PP2, PP3	Likely Pathogenic
Adrenal	GATA4	Chr8: 11607629	c.793C>T	p.Arg265Cys	rs776523140	0.001%	0	PP2, PP3, PM1	VUSp
Familial	NR5A1	Chr9: 127262853	c.386C>T	p.Pro129Leu	rs200749741	0.02%	0	PP3	VUSp
Severe IR	DLK1	Chr14: 101200674	c.594delC	p.(Gly199Alafs*11)	NA	0	0	PM2, PM4, PP1, PP3	Likely Pathogenic
Severe IR	INSR	Chr19: 7120722	c.3568T>C	p.Tyr1190His	rs1448499462	0.003%	0	PM1, PM2, PP2, PP3, PP4	Likely Pathogenic
Normoandrogenic	BMP15	ChrX: 50653985	c.202C>T	p.Arg68Trp	rs104894763	0.06%	0.09%	PP3	VUSp

Abbreviations: NA, not available; VUS, variant of uncertain significance.
*gnomAD version 2.1.1—checked in December, 2020.

with a basal total testosterone of 131 ng/dL. After dexamethasone suppression, total testosterone decreased to 14 ng/dL, but after gonadotropin-releasing hormone analog suppression testosterone remained elevated (158 ng/dL), suggesting that the adrenal would be the exclusive androgen source in this case. Interestingly, the variant in *GATA4* is located in the zinc finger domain of the protein. The zinc finger domain is also termed the GATA-binding transcription factor domain because it is the motif of interaction for multiple transcription factors.

Variants in genes previously related to ovary insufficiency

NR5A1. A double NR5A1 variant (p.Gly123Ala/p.Pro129Leu) was identified by WES in a patient with familial PCOS. Her mother, affected by PCOS and hyperthecosis, carried the same variants. This variant has been previously described in a child with clitoromegaly and premature ovarian insufficiency (POI) [27]. Functional assays of NR5A1 activity show severe loss of activity of p.Pro129Leu compared with the wild-type protein [27].

BMP15. Heterozygous missense variant in *BMP15* (c.C202T; p.Arg68Trp) was identified in a patient with the normoandrogenic phenotype. In vitro studies showed that this variant leads to marked reduction in mature BMP15 protein secretion, which is reflected in defective stimulation of the BMP15 pathway in granulosa cells [28]. Normally, BMP15 synergizes with GDF9 to stimulate granulosa cell proliferation [29]. An impairment in BMP15 activity could contribute to a disorder in folliculogenesis. Until now, this variant has been reported only in patients with POI [28, 30].

FSHR. We identified a heterozygous variant (c.C847T; p.Arg283Trp) in *FSHR* in a patient with PCOS and early-onset diabetes mellitus. This variant c.C847T is located in the receptor N-terminal extracellular domain, which is responsible for the recognition and binding of the receptor to its ligand. Decreased FSH–FSH receptor binding could affect follicular development up to the early antral stage. Mutation in this gene is a cause of ovarian dysgenesis and POI [31]. The locus that contains *FSHR* has been identified in Chinese and European PCOS GWAS.

Discussion

PCOS is a common complex disorder with well-known reproductive and metabolic aspects, but with much uncertainty regarding its etiology. PCOS follows a non-Mendelian pattern of inheritance consistent with a complex genetic disorder, but, in rare cases, it can follow a Mendelian pattern associated with extreme phenotypes ([32]. Performing large-scale sequencing in a preselected PCOS population, we identified rare genetic variants potentially associated with the phenotype in 18.8% of cases. The study detected variants associated with already known monogenic conditions that mimic the PCOS phenotype, such as the ones caused by the *INSR* gene and the *LMNA* gene, but also raised the possibility of new candidate genes associated with the PCOS phenotype, such as the ones caused by rare genetic variants in the *PIK3R1* and *DLK1* genes.

Mutations in the *INSR* gene have been described as a rare cause of an autosomal dominant disorder known as type A IR syndrome [33]. Additionally, the locus on chromosome

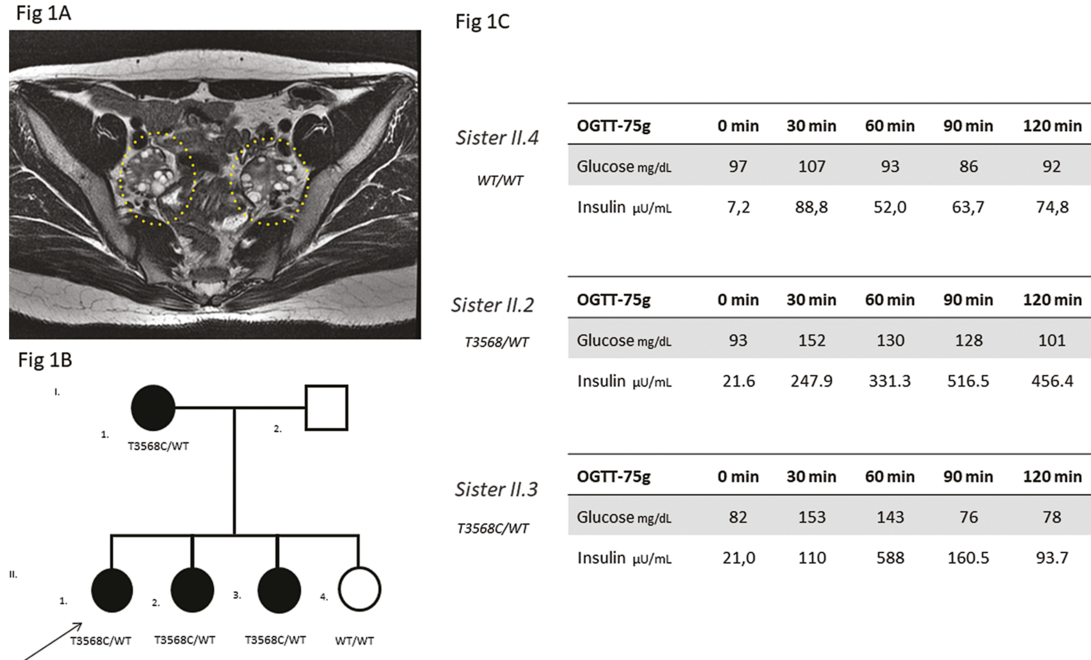

Figure 1. (A) Enlarged ovaries detected by pelvic magnetic resonance imaging of the patient with *INSR* variant. Micropolycystic ovaries are indistinguishable from the appearance of primary PCOS. (B) Pedigree of the family with the *INSR* variant showing that 2 sisters and the mother harbor the same variant as the index case. They are all asymptomatic. (C) OGTT result of the family with the *INSR* variant. The 2 affected sisters are asymptomatic, but presented hyperinsulinemia on OGTT.

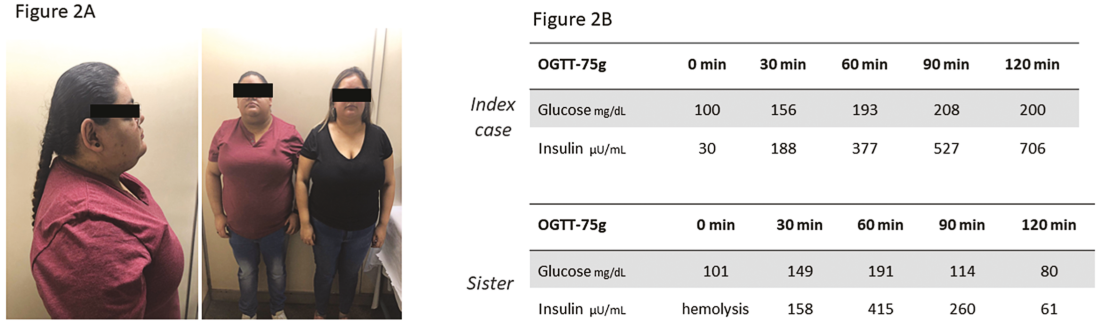

Figure 2. (A) Two sisters with *PIK3R1* mutations. Both of them present with PCOS and insulin resistance. (B) OGTT of the patients with PIK3R1 mutation demonstrating hyperinsulinemia.

19q13.3 that contains *INSR* has been identified by GWAS as a susceptibility locus for PCOS [11]. In the current cohort, we identified a rare variant in *INSR* in 1 patient with the classical phenotype of PCOS associated with severe IR, mainly detected after measurement of insulin at 120 minutes on the OGTT, since her glucose and HbA1C were normal at baseline. Notably, 2 affected sisters had no symptoms of hyperandrogenism and demonstrated hyperinsulinemia only at 120 minutes on the OGTT (Fig. 1C), indicating potential phenotypic expression variability of this variant. Although, IR can be present in women with PCOS [34], certain patients, especially lean ones, may benefit from careful evaluation by OGTT or genetic testing to rule out this monogenic condition.

The most common findings, the LMNA mutations, can cause several conditions, including disorders of adipose and muscular tissue, cardiomyopathies, and neurological disorders [35]. Other features include IR and hyperandrogenism, also mimicking PCOS [36]. These conditions are frequent in PCOS patients and may be misdiagnosed in the absence of a genetic evaluation. As cited above, we identified the variant c.G746A classified as pathogenic in *LMNA* in a patient with muscular dystrophy and arrhythmia, conditions that made us suspect this patient had laminopathy even before the genetic study. However, in other situations, the clinical distinction between PCOS and laminopathies can be challenging. The other 2 LMNA variants were classified as VUSp: the variant c.C1930T was identified in a patient with early-onset type 2 diabetes mellitus and family history of cardiomyopathy; in addition, this variant was frequently reported to be associated with multiple laminopathy phenotypes, varying from asymptomatic to lipodystrophy and cardiomyopathy; the other variant c.G1912A was identified in a patient with glucose intolerance and family history of PCOS. This variant is located in the prelamin A tail. Prelamin A undergoes a multistep maturation that includes cleavage of the C-terminus 15 amino acids of the prelamin A tail. Failure to cleave prelamin A causes progeroid disorders. Barrowman et al demonstrated that the minimal region of the prelamin A C-terminus that

Table 5. Features of patients harboring variants associated with severe IR

Gene	*LMNA*	*LMNA*	*LMNA*	*INSR*	*PIK3R1*	*DLK1*
Variant	p.Gly638Arg	p.Arg644Cys	p.Arg249Gln	p.Tyr1190His	p.Leu26Phe	p.Gly199Alafs*11
Clinical features	Glucose intolerance	Precocious diabetes	Muscular dystrophy	Insulin resistance	Insulin resistance	Precocious diabetes, central precocious puberty
BMI (kg/m^2)	25.9	32.0	25.3	22.7	40.5	29.2
Total testosterone (reference value (RV) < 48 ng/dL)	37	103	136	306	67	32
Glucose 0 minutes (mg/dL)	106	268	72	84	100	—
Glucose 120 minutes (mg/dL)	150	310	141	136	200	—
Insulin 0 minutes (µUI/mL)	14.5	44.6	Hemolysis	35.7	30	—
Insulin 120 minutes (µUI/mL)	72.3	59.2	507.2	>1000	706.4	—
HbA1c %	5.8	10.8	5.0	5.3	5.6	9.5
Ferriman score	12	19	18	24	11	16
Liver disease	Not evaluated	Steatosis Fibroscan: fibrosis score F3	Steatosis	Absent	Steatosis	Steatosis
Familial history	Mother with PCOS	Father died due to arrythmia	—	Mother: gestational diabetes	Sister with similar phenotype	Sister with similar phenotype

Abbreviations: BMI, body mass index; HbA1c, glycated hemoglobin.

contains critical sites for the ZMPSTE24-mediated cleavage of the prolamin A tail were the last 31 amino acid positions (positions 634-664) [37]. Therefore, we speculate that the G1912A variant could reduce the efficiency of prelamin A cleavage.

PIK3R1 mutations are an established cause of SHORT syndrome, denoting short stature, joint hyperextensibility, ocular depression, Rieger anomaly (developmental defect in the iris), and teething delay [25]. Huang-Doran et al described a phenotype of severe IR in a cohort of 5 patients with SHORT syndrome with the heterozygous hot spot mutation p.Tyr657X [25]. In our study, a new variant located at an important functional domain of the protein PIK3R1 was identified in 2 sisters with a classic PCOS phenotype, severe IR, obesity, hypercholesterolemia, and central precocious puberty without features of SHORT syndrome. The ACMG classified this variant as likely pathogenic. This association of central precocious puberty, severe IR, and PCOS was also present in another 2 sisters, in which WES identified the loss-of-function variant in the *DLK1* gene, initially described by our group [26, 38]. The DLK1 protein is an adipogenesis gatekeeper and the loss of protein function promotes truncal obesity and IR in mice and humans [39].

In the adrenal component group, identification of a rare variant in the *GATA4* gene located in the important zinc finger motif, an essential domain for interaction with other transcription factor, supports the potential role of causality of this gene. The *GATA4* was identified by GWAS and plays an essential role in regulating steroidogenesis in the adrenal and gonads [40]. Several steroidogenic enzymes contain 1 or more GATA regulatory motif, and transactivation experiments showed that GATA factors regulate steroidogenic enzyme promoters [41]. These interactions of the GATA factors to the DNA motif occur via a highly conserved zinc finger domains, in which the zinc ion is coordinated by 4 cysteine residues. Two GATA zinc fingers are found in the GATA4 transcription factor, and Arg265Cys is located in 1 of these motifs according to the Human Reference Protein Database (https://hprd.org/index_html). Therefore, we speculate that the p.Arg265Cys variant could modify steroidogenesis regulation. Besides that, other genetic and epigenetic factors might be implicated in adrenal steroidogenesis dysfunction in this group of patients with PCOS with the severe adrenal component.

We identified 3 variants in genes previously associated with POI: *NR5A1*, *BMP15*, and *FSHR*. Notably, Mendelian randomization of POCS suggests shared genetic architecture with menopause timing, suggesting correlation between these 2 conditions [8].

All 3 genes are involved in ovarian folliculogenesis. The *NR5A1* gene encodes protein steroidogenic factor 1, a transcriptional regulator involved in adrenal and gonadal development and function. Mutations in *NR5A1* have been associated with a wide phenotypic spectrum varying from gonadal development disorders to primary ovarian insufficiency. The variant found in our study was previously described in a child with clitoromegaly, a clinical sign of hyperandrogenism [27]. The BMP15 protein stimulates granulosa cell proliferation [29]. An impairment in BMP15 activity could contribute to disordered folliculogenesis, a well-known feature of PCOS. In our study, this variant was found in a patient with the normoandrogenic phenotype, demonstrating mainly a dysfunction in folliculogenesis. Regarding *FSHR*, it has an

important role in ovarian follicular maturation, particularly during early antral development. Loss-of-function FSHR mutations, in the homozygous state, cause ovarian dysgenesis and/or primary ovarian insufficiency [42]. We speculate that the FSHR variant in the heterozygous state has a dominant-negative effect as occurred in the thyrotropin receptor [43]. Different variants in the same gene causing different conditions have been described in others hypothalamic–pituitary–ovarian conditions such as precocious puberty and hypogonadotropic hypogonadism [44]. However, additional studies are needed to elucidate the real role of these variants in the PCOS phenotype.

Although in rare disease, aggregating probands with similar phenotype have been crucial for identifying rare causal variants, and HTS is now the standard of care for evaluating these patients, the impact of rare variants on common diseases and on selected phenotypes has been less explored [45].

In recent years, some advances in the genetic basis of PCOS have been achieved. Gorsic et al [4]. identified rare genetic variants in the gene of antimullerian hormone in 3% of their PCOS cohort by target sequencing. The same authors detected noncoding variants in *DENND1A* by whole genome sequencing that were significantly associated with reproductive and metabolic traits [46].

Our study is a pioneer in aggregating probands with PCOS and similar phenotypes to identify groups of rare variants associated with a specific subphenotype. In this context, searching for rare variants in PCOS cohorts with severe IR seems to be more promising.

Our study has some limitations. We had a small cohort and limited access to proband family members for segregation analysis. Despite the small sample, our inclusion criteria allowed the identification of relevant findings in our cohort. The replication of this study design in larger PCOS cohorts will help to understand the real impact of rare variants in different phenotypes and the group of patients who would best benefit from genetic testing.

The characterization of rare coding variants will improve our understanding of the complex genetic architecture of PCOS, and, on an individual level, may refine the clinical management and choice of medication of these patients and their relatives.

Financial Support

This work was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (2015/17350-0 FAPESP to L.G.G.) and Coordenação de Aperfeiçoamento de Pessoas de Nível Superior (2014/1459789 CAPES to L.G.G.); Fundação de Amparo à Pesquisa do Estado de São Paulo (2019/27631-7 FAPESP to A.C.L), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (153237/2016-3 CNPq to R.P.C.).

Conflict of Interest

The authors have nothing to disclose

Data Availability

The data sets generated during and/or analyzed during the present study are not publicly available but are available from the corresponding author on reasonable request.

References

1. Azziz R, Marin C, Hoq L, Badamgarav E, Song P. Health care-related economic burden of the polycystic ovary syndrome during the reproductive life span. *J Clin Endocrinol Metab.* 2005;90(8):4650-4658.
2. Lizneva D, Suturina L, Walker W, Brakta S, Gavrilova-Jordan L, Azziz R. Criteria, prevalence, and phenotypes of polycystic ovary syndrome. *Fertil Steril.* 2016;106(1):6-15.
3. Azziz R, Carmina E, Chen Z, *et al*. Polycystic ovary syndrome. *Nat Rev Dis Primers*. 2016;2(1):16057.
4. Gorsic LK, Kosova G, Werstein B, *et al*. Pathogenic anti-Müllerian hormone variants in polycystic ovary syndrome. *J Clin Endocrinol Metab*. 2017;102(8):2862-2872.
5. Chen ZJ, Zhao H, He L, *et al*. Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. *Nat Genet.* 2011;43(1):55-59.
6. Shi Y, Zhao H, Cao Y, *et al*. Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. *Nat Genet.* 2012;44(9):1020-1025.
7. Hayes MG, Urbanek M, Ehrmann DA, *et al*. Genome-wide association of polycystic ovary syndrome implicates alterations in gonadotropin secretion in European ancestry populations. *Nat Commun*. 2015;6(1):7502.
8. Day F, Karaderi T, Jones MR, *et al*. Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. *PLoS Genet.* 2018;14(12):e1007813.
9. Zhang Y, Ho K, Keaton JM, *et al*. A genome-wide association study of polycystic ovary syndrome identified from electronic health records. *Am J Obstet Gynecol.* 2020;223(4):559.e1-55.e21.
10. Vink JM, Sadrzadeh S, Lambalk CB, Boomsma DI. Heritability of polycystic ovary syndrome in a Dutch twin-family study. *J Clin Endocrinol Metab.* 2006;91(6):2100-2104.
11. Dunaif A. Perspectives in polycystic ovary syndrome: from hair to eternity. *J Clin Endocrinol Metab.* 2016;101(3):759-768.
12. Manolio TA, Collins FS, Cox NJ, *et al*. Finding the missing heritability of complex diseases. *Nature* 2009;461(7265):747-753.
13. Agrawal N, Frederick MJ, Pickering CR, *et al*. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* 2011;333(6046):1154-1157.
14. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet.* 2010;11(6):415-425.
15. Speiser PW, Arlt W, Auchus RJ, *et al*. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab.* 2018;103(11):4043-4088.
16. Kulshreshtha B, Ganie MA, Praveen EP, *et al*. Insulin response to oral glucose in healthy, lean young women and patients with polycystic ovary syndrome. *Gynecol Endocrinol.* 2008;24(11):637-643.
17. Hatch R, Rosenfield RL, Kim MH, Tredway D. Hirsutism: implications, etiology, and management. *Am J Obstet Gynecol.* 1981;140(7):815-830.
18. Balen AH, Laven JS, Tan SL, Dewailly D. Ultrasound assessment of the polycystic ovary: international consensus definitions. *Hum Reprod Update.* 2003;9(6):505-514.
19. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988;16(3):1215.
20. de Bruin C, Finlayson C, Funari MF, *et al*. Two patients with severe short stature due to a FBN1 mutation (p.Ala1728Val) with a mild form of acromicric dysplasia. *Horm Res Paediatr* 2016;86(5):342-348.
21. Naslavsky MS, Yamamoto GL, de Almeida TF, *et al*. Exomic variants of an elderly cohort of Brazilians in the ABraOM database. *Hum Mutat.* 2017;38(7):751-763.
22. Lerario AM, Mohan DR, Montenegro LR, *et al*. SELAdb: a database of exonic variants in a Brazilian population referred to

a quaternary medical center in São Paulo. *Clinics (Sao Paulo)* 2020;75:e1913.
23. Richards S, Aziz N, Bale S, *et al*. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424.
24. Emerson LJ, Holt MR, Wheeler MA, Wehnert M, Parsons M, Ellis JA. Defects in cell spreading and ERK1/2 activation in fibroblasts with lamin A/C mutations. *Biochim Biophys Acta*. 2009;1792(8):810-821.
25. Huang-Doran I, Tomlinson P, Payne F, *et al*. Insulin resistance uncoupled from dyslipidemia due to C-terminal PIK3R1 mutations. *JCI Insight* 2016;1(17):e88766.
26. Gomes LG, Cunha-Silva M, Crespo RP, *et al*. DLK1 Is a novel link between reproduction and metabolism. *J Clin Endocrinol Metab*. 2019;104(6):2112-2120.
27. Lourenço D, Brauner R, Lin L, *et al*. Mutations in NR5A1 associated with ovarian insufficiency. *N Engl J Med*. 2009;360(12):1200-1210.
28. Rossetti R, Di Pasquale E, Marozzi A, *et al*. BMP15 mutations associated with primary ovarian insufficiency cause a defective production of bioactive protein. *Hum Mutat*. 2009;30(5):804-810.
29. Persani L, Rossetti R, Di Pasquale E, Cacciatore C, Fabre S. The fundamental role of bone morphogenetic protein 15 in ovarian function and its involvement in female fertility disorders. *Hum Reprod Update*. 2014;20(6):869-883.
30. Di Pasquale E, Rossetti R, Marozzi A, *et al*. Identification of new variants of human BMP15 gene in a large cohort of women with premature ovarian failure. *J Clin Endocrinol Metab*. 2006;91(5):1976-1979.
31. Gromoll J, Simoni M, Nordhoff V, Behre HM, De Geyter C, Nieschlag E. Functional and clinical consequences of mutations in the FSH receptor. *Mol Cell Endocrinol*. 1996;125(1-2):177-182.
32. Dapas M, Dunaif A. Deconstructing a syndrome: genomic insights into PCOS causal mechanisms and classification. *Endocr Rev*. 2022.
33. Moller DE, Cohen O, Yamaguchi Y, *et al*. Prevalence of mutations in the insulin receptor gene in subjects with features of the type A syndrome of insulin resistance. *Diabetes*. 1994;43(2):247-255.
34. Dunaif A, Segal KR, Futterweit W, Dobrjansky A. Profound peripheral insulin resistance, independent of obesity, in polycystic ovary syndrome. *Diabetes*. 1989;38(9):1165-1174.
35. Rankin J, Auer-Grumbach M, Bagg W, *et al*. Extreme phenotypic diversity and nonpenetrance in families with the LMNA gene mutation R644C. *Am J Med Genet A*. 2008;146A(12):1530-1542.
36. Gambineri A, Zanotti L. Polycystic ovary syndrome in familial partial lipodystrophy type 2 (FPLD2): basic and clinical aspects. *Nucleus* 2018;9(1):392-397.
37. Barrowman J, Hamblet C, Kane MS, Michaelis S. Requirements for efficient proteolytic cleavage of prelamin A by ZMPSTE24. *PLoS One*. 2012;7(2):e32120.
38. Dauber A, Cunha-Silva M, Macedo DB, *et al*. Paternally inherited DLK1 deletion associated with familial central precocious puberty. *J Clin Endocrinol Metab*. 2017;102(5):1557-1567.
39. Moon YS, Smas CM, Lee K, *et al*. Mice lacking paternally expressed Pref-1/Dlk1 display growth retardation and accelerated adiposity. *Mol Cell Biol*. 2002;22(15):5585-5592.
40. Schrade A, Kyrönlahti A, Akinrinade O, *et al*. GATA4 is a key regulator of steroidogenesis and glycolysis in mouse Leydig cells. *Endocrinology* 2015;156(5):1860-1872.
41. Tremblay JJ, Viger RS. Novel roles for GATA transcription factors in the regulation of steroidogenesis. *J Steroid Biochem Mol Biol*. 2003;85(2-5):291-298.
42. Meduri G, Bachelot A, Cocca MP, *et al*. Molecular pathology of the FSH receptor: new insights into FSH physiology. *Mol Cell Endocrinol*. 2008;282(1-2):130-142.
43. Beck-Peccoz P, Persani L, Calebiro D, Bonomi M, Mannavola D, Campi I. Syndromes of hormone resistance in the hypothalamic-pituitary-thyroid axis. *Best Pract Res Clin Endocrinol Metab*. 2006;20(4):529-546.
44. Chambers AE, Banerjee S. Natural antisense LHCGR could make sense of hypogonadism, male-limited precocious puberty and pre-eclampsia. *Mol Cell Endocrinol*. 2005;241(1-2):1-9.
45. Cirulli ET, White S, Read RW, *et al*. Genome-wide rare variant analysis for thousands of phenotypes in over 70,000 exomes from two cohorts. *Nat Commun*. 2020;11(1):542.
46. Dapas M, Sisk R, Legro RS, Urbanek M, Dunaif A, Hayes MG. Family-based quantitative trait meta-analysis implicates rare noncoding variants in DENND1A in polycystic ovary syndrome. *J Clin Endocrinol Metab*. 2019;104(9):3835-3850.