European Review for Medical and Pharmacological Sciences 2023; 27: 3719-3732

# Genome-wide linkage and association study identifies novel genes and pathways implicated in polycystic ovarian syndrome

M. AMIN[1,2], C. GRAGNOLI[3,4,5]

[1]INSERM, US14-Orphanet, Paris, France
[2]Department of Biochemistry and Molecular Biology, Faculty of Medicine, Al-Neelain University, Khartoum, Sudan
[3]Department of Medicine, Division of Endocrinology, Creighton University School of Medicine, Omaha, NE, USA
[4]Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, USA
[5]Molecular Biology Laboratory, Bios Biotech Multi-Diagnostic Health Center, Rome, Italy

**Abstract. – OBJECTIVE: Polycystic ovarian syndrome (PCOS) is a complex heterogeneous condition that affects women of reproductive age, conferring increased cardiovascular morbidity and mortality. The syndrome is characterized by oligomenorrhea, hyperandrogenism, and/or polycystic ovaries and is often associated with obesity and type 2 diabetes. Individuals are predisposed to PCOS by environmental factors and risk variants in genes mostly involved in ovarian steroidogenesis and/or insulin resistance. Genetic risk factors have been identified by both familial and genome-wide (GW) association studies. However, most genetic components are still unknown and missing heritability needs to be elucidated. To learn more about the genetic determinants of PCOS, we performed a GW study in genetically highly homogeneous peninsular families.**

**PATIENTS AND METHODS: We conducted the first GW-linkage and linkage disequilibrium (i.e., linkage + association) study in Italian families with PCOS.**

**RESULTS: We identified several novel risk variants, genes, and pathways potentially implicated in the pathogenesis of PCOS. Specifically, we detected 79 novel variants with significant GW-linkage and/or -association with PCOS across 4 inheritance models ($p < 0.00005$), of which 50 variants were within 45 novel PCOS-risk genes.**

**CONCLUSIONS: This is the first GW-linkage and linkage disequilibrium study performed in peninsular Italian families and reporting novel genes in PCOS.**

*Key Words:*
Polycystic ovary syndrome, Pathways, Gene, Variant, Risk, Linkage, Parametric analysis, Linkage disequilibrium, Association, Cortisol, Metabolic, Insulin resistance, Obesity, Type 2 diabetes, Families, Familial, Italy, Italian, Single nucleotide polymorphisms, SNP, Hyperandrogenism, Hyperandrogenemia, Irregular menses, Cycles, Amenorrhea, Oligomenorrhea, Anovulation, Subfertility, Infertility, Folliculogenesis, Follicle maturation, Fat metabolism, Ethnic group, inflammation, Cardiovascular morbidity, Mortality, PPP1, Protein phosphatase 1, PPP1R12B, Protein phosphatase 1 regulatory subunit 12B, NEGR1, Neuronal growth regulator 1, SESTD1, SEC14 And spectrin domain containing 1, ATL2, Atlastin GTPase 2, C1QTNF7, C1q and TNF related 7, PROM1, Prominin 1, FGF1, Fibroblast growth factor 1, LMBR1, Limb development membrane protein 1, DGKB, Diacylglycerol kinase beta, LINC02476, Long intergenic non-protein coding RNA 2476, PTK2B, Protein tyrosine kinase 2 beta, MLLT3, MLLT3 Super elongation complex subunit, KCNT1, Potassium sodium-activated channel subfamily T member 1, ADAM12, ADAM metallopeptidase domain 12, HK1, hexokinase 1, CASP5, Caspase 5, CDCA5, Cell division cycle associated 5, SPPL3, Signal peptide peptidase like 3, PCDH9, Protocadherin 9, DCAF4, DDB1 and CUL4 associated factor 4, TPM1, Tropomyosin 1, CFAP52, Cilia and flagella associated protein 52, CDH7, Cadherin 7, PTPRM, Protein tyrosine phosphatase receptor type M, LRRC74B, Leucine rich repeat containing 74B, bHLH2, Basic helix-loop-helix transcription factor, CCDC54, Coiled-coil domain containing 54, PDZD2, PDZ domain containing 2, FARS2, Phenylalanyl-tRNA synthetase 2, Mitochondrial, CHRNB3, Cholinergic receptor nicotinic beta 3 subunit, BEND7, BEN domain containing 7, LAMTOR1, Late endosomal/lysosomal adaptor, MAPK and MTOR activator 1, LRTOMT, Leucine rich transmembrane and O-methyltransferase domain containing, B3GAT1-DT, B3GAT1 divergent transcript, MYCBP2, MYC binding protein 2, LINC02341, Long intergenic non-protein coding RNA 2341, UNC13C, unc-13 homolog C, PLA2G4E, Phospholipase A2 group IVE, ALOX12P2, Arachidonate 12-lipoxygenase pseudogene 2, CACNG5, Calcium voltage-gated channel auxiliary subunit gamma 5.

*Corresponding Author:* Claudia Gragnoli, MD, Ph.D; e-mail: claudia.gragnoli@gmail.com

## Introduction

Polycystic ovarian syndrome (PCOS) is a common condition affecting up to 10% of worldwide women who present with two of three of the following: oligo- and/or anovulation, clinical and/or biochemical signs of hyperandrogenism, and polycystic ovaries[1]. Psychological aspects include sexual dysfunction, depression, anxiety, and decreased self-esteem and quality of life[2,3].

Hyperandrogenism is a characteristic feature of PCOS, and it is aggravated by metabolic and immune impairments associated with PCOS manifestation and exacerbation[4]. These include a pro-inflammatory state, impaired glucose metabolism, insulin resistance (IR)[5,6], obesity, and type 2 diabetes (T2D)[7,8]. IR and low-grade chronic inflammation underlie PCOS phenotypes and traits[4], such as obesity[9], elevated circulating androgens[10] and hyperandrogenic anovulation[10,11]. Insulin-sensitizing treatments, namely metformin, diet, and exercise, improve PCOS symptomatology, contribute to ovulation resumption, and reduce inflammatory markers[12-15].

Several genetic, epigenetic, and environmental risk factors have been associated with the risk of PCOS[16-18]. A Dutch twin study demonstrated the genetic component of PCOS[19]. Two GWAS studies performed in Han Chinese women with PCOS identified 11 susceptibility loci[20,21.] Due to the vulnerability of the case-control design to population admixture, causing false-positive results, a Chinese family-based study of 321 trios (parents and PCOS-affected proband) was performed using the transmission disequilibrium test (TDT) for 10 previously detected GWAS signals: significant allelic transmission differences were found at rs2349415 (*FSHR* gene, $p=0.0001$) and rs3802457 (*C9orf3* gene, $p=0.0001$)[22]. Several PCOS-risk variants are in genes involved in ovarian steroidogenesis, such as *CYP11A*, *CYP17*, and *CYP19*[23], or hormonal folliculogenesis, such as *AMH*[24]. Other genetic risk factors have been identified within the insulin-signaling pathway (e.g., insulin receptor, insulin receptor substrate-1)[25-28]. Most genes and genes' variants conferring risk for PCOS are, however, still unknown. To detect novel risk variants, genes, and pathways potentially implicated in the pathogenesis of PCOS, we conducted the first GW-linkage and -association study in peninsular Italian families with PCOS. This is, therefore, the second familial study in PCOS, but, to the best of our knowledge, the first GW-linkage and linkage disequilibrium study including families and affected sib-pairs, and the first PCOS GW study performed in a genetically homogenous peninsular population with an increased detection power.

## Patients and Methods

We studied a cohort of 212 families originating from the Italian peninsula with multigenerational cases of type 2 diabetes and diagnosed for PCOS according to the Rotterdam diagnostic criteria (presence of at least two of the following three characteristics: chronic anovulation or oligomenorrhea, clinical or biological hyperandrogenism, and/or polycystic ovaries)[29]. Subjects were Italian from at least 3 generations. We excluded subjects with uncertain paternity or identical twins. We used familial GW-data, derived from the UK Biobank Axiom Array platform, which had passed stringent quality control (QC, $\geq 0.96$) and random genomic replication checks and kinship correlation verification. PLINK tool was used to exclude Mendelian and genotyping errors[30,31]. The study was conducted following the Helsinki declaration guidelines and approved by the Bios Ethical Committee. Informed written consent was obtained from each participant before the start of the study.

### In Silico Analysis

We analyzed the GW-significant variants with *in silico* tools that predict the pathogenicity of nonsynonymous variants with Sorting Intolerant From Tolerant (SIFT)[32], Polymorphism Phenotyping v2[33] (PolyPhen-2), and MutationTaster[34]. The non-coding variants were analysed using tools that predict splicing (SpliceAI)[35], transcription-factor (TF) binding (SNP Function Prediction)[36], regulatory potential (RegulomeDB)[37], and miRNA binding (mirSNP)[38]. To investigate genes-related pathways, we analyzed gene sets using PANTHER[39].

### Statistical Analysis

The PCOS-related informatic genetic variants ($\geq$ 600 k) were tested for parametric linkage to and/or linkage disequilibrium (LD, linkage + association) with PCOS via the dominant models with complete (D1) and incomplete penetrance (D2) and the recessive models with complete penetrance (R1) and incomplete penetrance (R2), using Pseudomarker[31]. Variants with $p <$ 0.00005 were considered statistically significant at GW-significance level.

## Results

We detected a total of 79 novel variants significantly GW-linked and/or -associated with PCOS across the 4 inheritance models (26/D1, 20/D2, 44/R1 and 13/R2) ($p < 0.00005$) (Figures 1-5). Fifty (63%) variants were located within 45 novel PCOS-risk genes (8 of which encode non-coding RNAs) (Table I). The most significantly associated risk variant was rs75798356 ($p < 0.000001$, R1), which is an intronic SNP lying within the gene named myeloid/lymphoid or mixed-lineage leukemia translocated to chromosome 3 protein (MLLT3) super elongation complex subunit (*MLLT3*). The results are detailed in Table I.

### *In Silico Findings*

We found that the variant rs7486605, which is located in the signal peptide peptidase like 3 (*SPPL3*) gene, affects the binding of the neuronal specific basic helix-loop-helix 1 (bHLH1) TF, which belongs to a family of TFs involved in neural growth and development[40]. Furthermore, the 3'-UTR variant rs2240688, located in the prominin 1 (*PROM1*) gene, creates a new binding site for the miRNA hsa-miR-135a.

The GW-significant genes detected in our study were involved in 15 sets of pathways: 1. Wingless-related integration site (Wnt)-signaling pathway (11%); 2. Cadherin-signaling pathway (11%); 3. Fructose-galactose metabolism (6%); 4. Glycolysis (6%); 5. Pentose-phosphate pathway (6%); 6. Angiogenesis (6%); 7. Fibroblast-growth factor (FGF) signaling pathway (6%); 8. Cholecystokinin-receptor (CCKR) signaling map (6%); 9. Gonadotropin-releasing hormone receptor (GnRHR) pathway (6%); 10. Inflammation mediated by chemokine- and cytokine-signaling pathway (6%); 11. Ionotropic glutamate-receptor pathway (6%); 12. Nicotine-pharmacodynamics pathway (6%); 13. Nicotinic acetylcholine-receptor signaling pathway (6%); 14. Synaptic vesicle trafficking (6%) and, 15. Integrin-signaling pathway (6%) (Figure 5).

## Discussion

PCOS is a complexly inherited condition predisposed by environmental factors and genetic risk variants[18] in genes mostly involved in ovarian steroidogenesis and/or insulin resistance[23,25-28]. In this

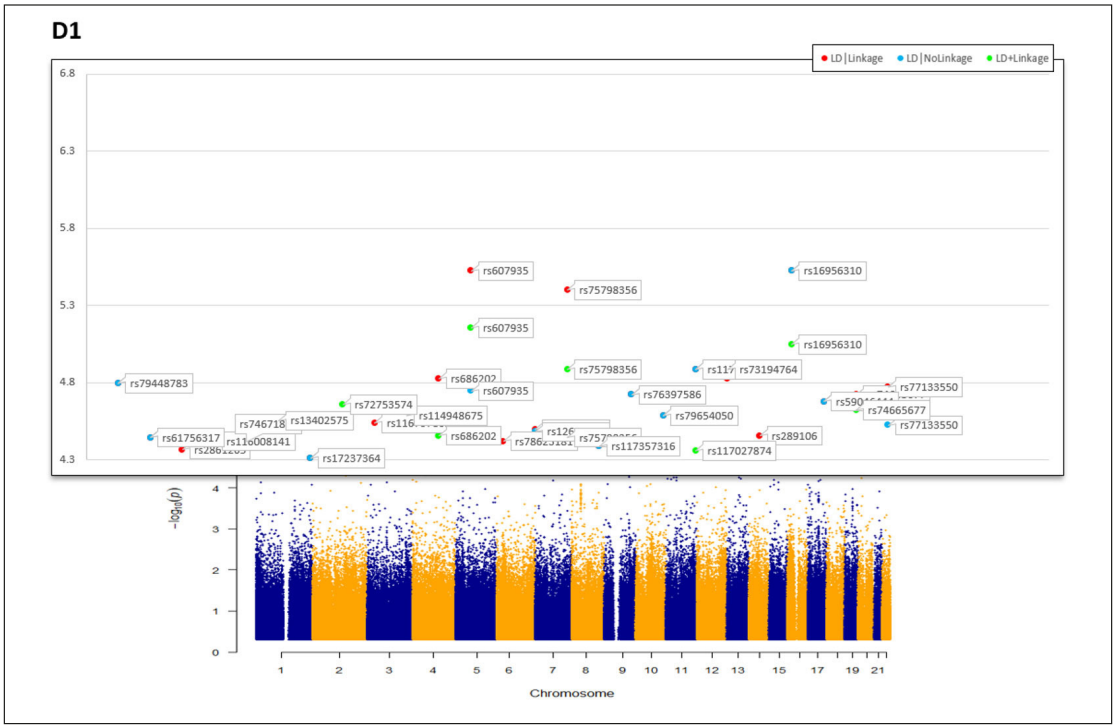


**Figure 1.** Manhattan plot for genome-wide linkage and linkage disequilibrium with polycystic ovarian syndrome under the dominant with complete penetrance model (D1). For each genomic-wide significant SNP in polycystic ovarian syndrome, we present the − log10(P) as a function of the test statistics and highlight above the significant ($p < 0.00005$) test statistics [(Linkage disequilibrium (LD)|Linkage, LD|NoLinkage, and LD + Linkage)] per inheritance model D1: dominant, complete penetrance.

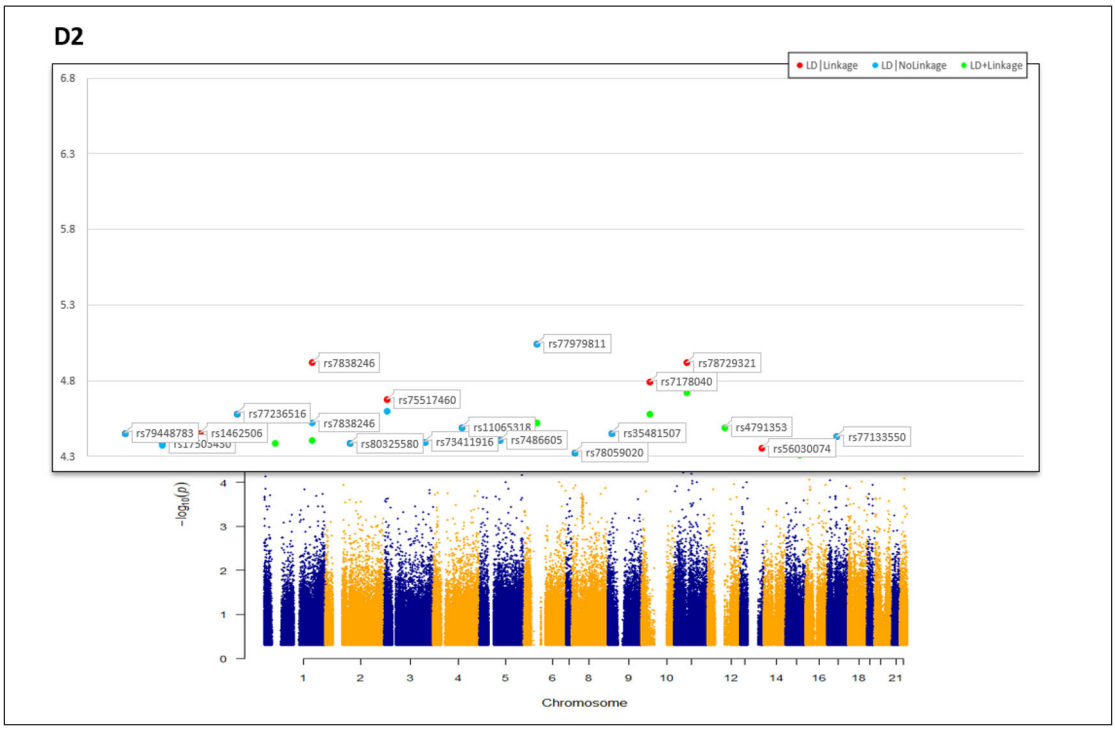


**Figure 2.** Manhattan plot for genome-wide linkage and linkage disequilibrium with polycystic ovarian syndrome under the dominant with incomplete penetrance model (D2). For each genomic-wide significant SNP in polycystic ovarian syndrome, we present the −log10(P) as a function of the test statistics and highlight above the significant ($p < 0.00005$) test statistics [(Linkage disequilibrium (LD)|Linkage, LD|NoLinkage, and LD + Linkage)] per inheritance model D2: dominant, incomplete penetrance.

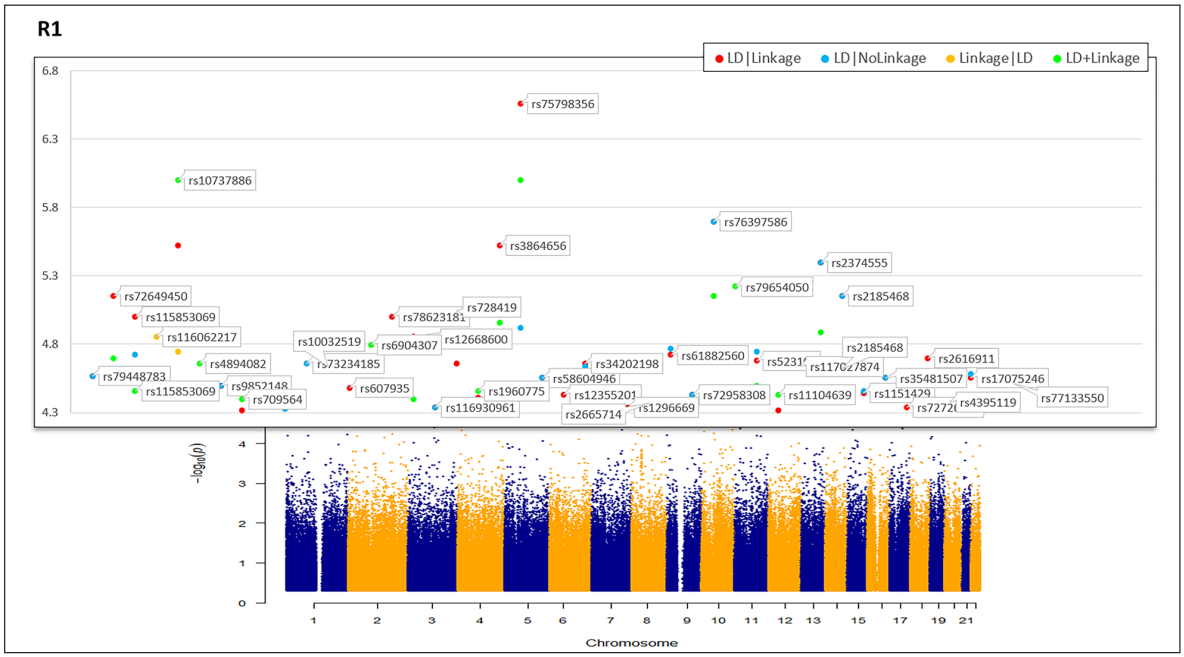


**Figure 3.** Manhattan plot for genome-wide linkage and linkage disequilibrium with polycystic ovarian syndrome under the recessive with complete penetrance model (R1). For each genomic-wide significant SNP in polycystic ovarian syndrome, we present the −log10(P) as a function of the test statistics and highlight above the significant ($p < 0.00005$) test statistics [(Linkage disequilibrium (LD)|Linkage, LD|NoLinkage, Linkage|LD, and LD + Linkage)] per inheritance model: R1: recessive, complete penetrance.

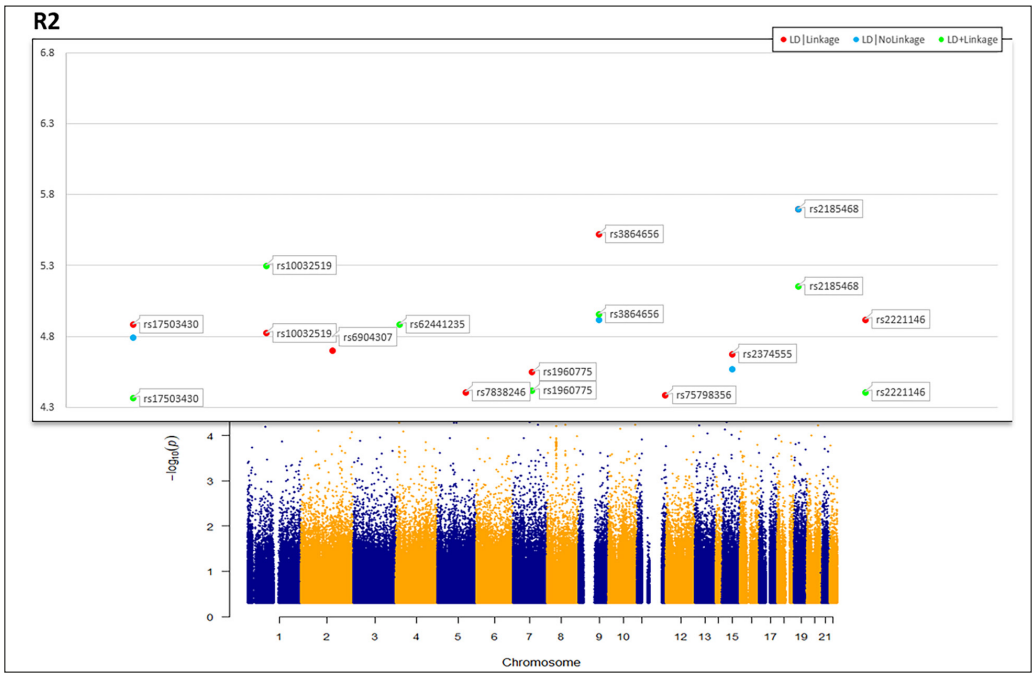


**Figure 4.** Manhattan plot for genome-wide linkage and linkage disequilibrium with polycystic ovarian syndrome under the recessive with incomplete penetrance model (R2). For each genomic-wide significant SNP in polycystic ovarian syndrome, we present the −log10(P) as a function of the test statistics and highlight above the significant ($p < 0.00005$) test statistics [(Linkage disequilibrium (LD)|Linkage, LD|NoLinkage, and LD + Linkage)] per inheritance model R2: recessive, incomplete penetrance.

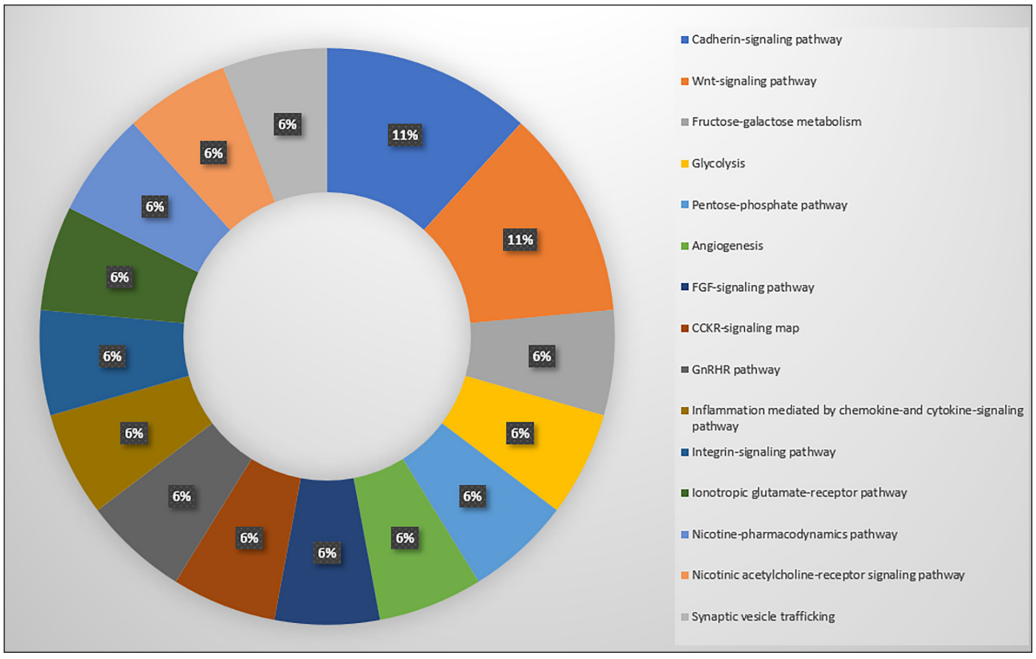


**Figure 5.** Pathways analysis for genome-wide significantly genes linked to and/or associated with polycystic ovarian syndrome ($p < 0.00005$). We present the significant pathways for the genome-wide genes significantly linked to and/or associated with polycystic ovarian syndrome ($p < 0.00005$) analyzed by PANTHER. Wnt = Wingless-related integration site, FGF = Fibroblast growth factor, CCKR = cholecystokinin receptor, GnRHR = Gonadotropin-releasing hormone receptor. The indicated percentage refers to the genes reaching genome-wide significance within a specific pathway across the total of genome-wide significant genes.

**Table I.** Risk variants and genes significantly linked to and/or in linkage disequilibrium with polycystic ovarian syndrome ($p < 0.00005$).

| Model[1] | Chr | Position | SNP | Ref/Alt | Risk Allele | Gene | Reported in PCOS or Related Phenotype(s)? |
|---|---|---|---|---|---|---|---|
| R1 | Chr01 | 242932849 | rs10737886 | G/A | G | - | Novel |
| R1 | Chr01 | 22362851 | rs72649450 | G/T | T | - | Novel |
| R1 | Chr01 | 202462315 | rs116062217 | C/T | C | *PPP1R12B* | Novel |
| R1 | Chr01 | 83144353 | rs115853069 | G/T | *T* | *LOC107985037* | Novel |
| D1, D2, R1 | Chr01 | 5218182 | rs79448783 | C/T | T | - | Novel |
| D2, R2 | Chr01 | 71527609 | rs17503430 | C/T | T | *NEGR1* | BMI42, obesity43 and age of menarche[44] |
| R1 | Chr02 | 179175932 | rs4894082 | G/T | T | *SESTD1* | Novel |
| D1 | Chr02 | 234633786 | rs13402575 | C/T | T | 2 kb upstream of *LOC105373936* | Novel |
| D1 | Chr02 | 38298465 | rs61756317 | T/C | C | *ATL2* | Novel |
| D1 | Chr02 | 226493033 | rs116008141 | C/T | T | - | Novel |
| D1 | Chr02 | 226536603 | rs74671850 | C/T | T | - | Novel |
| D1 | Chr02 | 78139065 | rs2861265 | C/T | T | *LOC101927967* | Novel |
| R1 | Chr03 | 59433128 | rs9852148 | C/T | T | *LOC105377110* | Novel |
| R1 | Chr03 | 107377700 | rs709564 | G/A | G | *CCDC54* | Novel |
| R1 | Chr04 | 19004092 | rs73234185 | A/G | G | *LOC107986263* | Novel |
| R1, R2 | Chr04 | 138604261 | rs10032519 | T/C | T | - | Novel |
| R1 | Chr04 | 15381910 | rs4388081 | T/C | C | *C1QTNF7* | Insulin resistance[49] |
| D2 | Chr04 | 178088516 | rs1462506 | C/T | C | - | Novel |
| R1, R2 | Chr04 | 15968726 | rs2240688 | T/G | G | *PROM1* | Novel |
| D1 | Chr04 | 57952725 | rs17237364 | C/T | C | - | Novel |
| D1 | Chr05 | 31665181 | rs72753574 | G/A | A | *PDZD2* | Upregulated in T2D |
| D1 | Chr05 | 142626129 | rs114948675 | T/C | C | *FGF1* | Serum FGF1 levels are abnormal in PCOS[62] |
| D1 | Chr05 | 64386935 | rs116737809 | G/A | A | - | Novel |
| D1, R1 | Chr06 | 56425611 | rs607935 | C/A | A | - | Novel |
| R2 | Chr06 | 139349999 | rs62441235 | C/T | C | - | Novel |
| D1 | Chr06 | 56416128 | rs686202 | G/A | A | - | Novel |
| R1, R2 | Chr06 | 57256886 | rs6904307 | A/G | G | - | Novel |
| D2 | Chr06 | 5571777 | rs77236516 | A/C | C | *FARS2* | Insulin resistance in mice[56] |
| R1 | Chr07 | 156884573 | rs728419 | C/T | C | *LMBR1* | Novel |
| R1 | Chr07 | 79823526 | rs12668600 | T/C | C | *LOC105375371* | Novel |
| D1, R1 | Chr07 | 79784784 | rs78623181 | C/T | T | *LOC105375371* | Novel |
| D2 | Chr07 | 14224564 | rs75019655 | C/A | A | *DGKB* | T2D42,48 |
| R1 | Chr07 | 119901960 | rs116930961 | A/G | G | *LINC02476* | Novel |
| R1, R2 | Chr08 | 95858411 | rs3864656 | T/C | C | - | Novel |
| D2, R2 | Chr08 | 42730154 | rs7838246 | A/G | G | *CHRNB3* | Waist circumference[58] and insulin resistance[59] |
| D1 | Chr08 | 27311717 | rs12679570 | C/T | T | *PTK2B* | Novel |
| R1, R2 | Chr08 | 83106775 | rs1960775 | G/A | G | - | Novel |
| D1, R1, R2 | Chr09 | 20412030 | rs75798356 | T/C | C | *MLLT3* | Novel |
| R1 | Chr09 | 135765207 | rs58604946 | T/G | T | *KCNT1* | Novel |
| R1 | Chr10 | 126102428 | rs34202198 | C/T | T | *ADAM12* | LH levels in PCOS[65] |
| R1 | Chr10 | 126116507 | rs1278389 | A/G | A | *ADAM12* | LH levels in PCOS[65] |
| R1 | Chr10 | 126124614 | rs1296669 | T/C | T | *ADAM12* | LH levels in PCOS[65] |
| R1 | Chr10 | 69329923 | rs12355201 | A/G | A | *HK1* | HbA1C level[50] |
| D2 | Chr10 | 13522299 | rs80325580 | T/C | C | *BEND7* | Novel |

*Continued*

**Table I (continued).** Risk variants and genes significantly linked to and/or in linkage disequilibrium with polycystic ovarian syndrome ($p < 0.00005$).

| Model[1] | Chr | Position | SNP | Ref/Alt | Risk Allele | Gene | Reported in PCOS or Related Phenotype(s)? |
|---|---|---|---|---|---|---|---|
| D1, R1 | Chr11 | 72100511 | rs76397586 | G/C | C | *LAMTOR1 LRTOMT* | Insulin resistance in mice[57] Upregulated in T2D[55] |
| R1 | Chr11 | 45351795 | rs61882560 | G/C | C | - | Novel |
| R1 | Chr11 | 104998981 | rs523104 | G/C | C | *CASP5* | Novel |
| D2 | Chr11 | 134477799 | rs75517460 | A/G | G | *B3GAT1-DT* | HbA1C levels[51] |
| D1, R1 | Chr11 | 72279338 | rs79654050 | G/A | A | - | Novel |
| R1 | Chr11 | 71530961 | rs72958308 | A/G | G | - | Novel |
| D1 | Chr11 | 65078680 | rs117357316 | C/T | T | *CDCA5* | Novel |
| R1 | Chr11 | 22397463 | rs2665714 | A/G | A | - | Novel |
| R1, R2 | Chr12 | 105878162 | rs2374555 | G/A | A | - | Novel |
| D1, R1 | Chr12 | 92443974 | rs117027874 | G/A | A | - | Novel |
| D2 | Chr12 | 120770435 | rs73411916 | T/C | C | *SPPL3* | Fat deposition[60] |
| D2 | Chr12 | 120896330 | rs11065318 | C/A | A | *SPPL3* | Fat deposition[60] |
| D2 | Chr12 | 120903362 | rs7486605 | C/T | T | *SPPL3* | Fat deposition[60] |
| R1 | Chr12 | 87844628 | rs11104639 | G/A | G | - | Novel |
| R1, R2 | Chr13 | 77092041 | rs2185468 | C/A | A | *MYCBP2* | Novel |
| D2 | Chr13 | 42402021 | rs77979811 | C/T | T | - | Novel |
| D1 | Chr13 | 66600253 | rs73194764 | A/G | G | *PCDH9* | Novel |
| R1 | Chr13 | 111449939 | rs1151429 | C/T | T | - | Novel |
| D2 | Chr13 | 42444658 | rs78059020 | T/G | G | *LINC02341* | Novel |
| D2, R1 | Chr14 | 72958656 | rs35481507 | C/T | T | *DCAF4* | Novel |
| D2 | Chr15 | 63073961 | rs78729321 | A/G | G | - | Novel |
| D2 | Chr15 | 63071899 | rs7178040 | G/T | T | *TPM1* | Novel |
| R1 | Chr15 | 53983656 | rs2616911 | G/A | G | *UNC13C* | FSH levels in PCOS[65] |
| D1 | Chr15 | 62339215 | rs289106 | C/T | T | - | Novel |
| R1 | Chr15 | 42040739 | rs72726102 | C/A | A | *PLA2G4E* | Downregulated in T2D[54] |
| D1 | Chr17 | 6865000 | rs16956310 | T/C | C | *ALOX12P2* | Novel |
| D1 | Chr17 | 6873148 | rs59046444 | T/C | C | *ALOX12P2* | Novel |
| D2 | Chr17 | 9601625 | rs4791353 | G/A | A | *CFAP52* | Novel |
| R1 | Chr17 | 52770506 | rs4395119 | G/T | G | - | Novel |
| D2 | Chr17 | 66857694 | rs56030074 | C/T | T | *CACNG5* | Novel |
| D1 | Chr18 | 73545694 | rs74665677 | G/A | A | - | Novel |
| R1 | Chr18 | 65814857 | rs17075246 | A/G | G | *CDH7* | Novel |
| D2 | Chr18 | 8257258 | rs8094332 | A/G | G | *PTPRM* | Age of first birth & number of children[66] |
| R1, R2 | Chr20 | 55003965 | rs2221146 | G/C | C | - | Novel |
| D1, D2, R1 | Chr22 | 21048083 | rs77133550 | G/A | A | *LRRC74B* | Novel |

[1]Models: D1: dominant, complete penetrance, D2: dominant, incomplete penetrance, R1: recessive, complete penetrance, R2: recessive, incomplete penetrance; - = intergenic; BMI: body mass index, T2D: type 2 diabetes; FGF1: fibroblast growth factor 1; PCOS: polycystic ovarian syndrome; LH: luteinizing hormone; Hba1c: glycated hemoglobin; FSH: follicle-stimulating hormone; kb: Kilobase.

familial study, we reported several novel variants, genes, and pathways potentially implicated in the pathogenesis of PCOS. PCOS-family based studies are lacking; and the one prior family study reporting loci associated with the risk of PCOS had only the probands affected and tested only 10 variants[22].

We conducted the first PCOS-related family-based GW-linkage and -association study in multigenerational Italian families and identified a total of 79 novel variants, of which 50 are in 45 new risk genes and 29 variants are intergenic (Table I) linked to and/or in linkage disequilibrium and thus associated with

PCOS. None of the genes reported in our study has been previously reported as a risk gene in PCOS. Several genes were, however, implicated in one or more of PCOS-related phenotypes (i.e., BMI, obesity, insulin resistance, PCOS abnormal serum FGF1 levels, glycated hemoglobin [Hba1c] levels, luteinizing hormone [LH] or follicle-stimulating hormone [FSH] levels, fat deposition in PCOS, T2D, age of menarche, subfertility, reproductive behavior, and number of children [Table I]) and could therefore have served as candidate genes. Variants in the neuronal growth regulator 1 (*NEGR1*) gene have previously been studied in PCOS, but no association has been found[41]. The *NEGR1* gene is, however, associated with BMI[42], obesity[43] and age of menarche[44] and has been reported as a PCOS-candidate gene[45]. Our study is, therefore, the first to confirm the association of *NEGR1* with PCOS.

Several genes reported in our study are implicated in metabolic phenotypes and traits, such as obesity, glucose metabolism, and/or insulin resistance (Table I), which are all essentially related to the pathogenesis of PCOS[6,46,47]. The diacylglycerol kinase beta (*DGKB*) and the C1q and tumor necrosis factor (TNF) related 7 (*C1QTNF7*) genes reported in our study are respectively associated with T2D[48] and IR[49] and variants in the hexokinase 1 (*HK1*) and the B3GAT1 divergent transcript (*B3GAT1-DT*) genes are associated with HbA1C levels[50,51]. The latter gene (*B3GAT1-DT*) is also associated with insulin secretion[52]. The three genes (PDZ domain containing 2 [*PDZD2*], leucine rich transmembrane and O-methyltransferase domain containing [*LRTOMT*], and phospholipase A2 group IVE [*PLA2G4E*]) are differentially expressed in pancreatic islets of patients with T2D (*PDZD2)*[53] or without T2D (*PLA2G4E*) [54] or in the peripheral blood of patients with T2D and fatigue (*LRTOMT*)[55].

High expression of phenylalanyl-tRNA synthetase 2, mitochondrial (*FARS2)* in mice is associated with features of T2D[56], and late endosomal/lysosomal adaptor, MAPK and MTOR activator 1 (*LAMTOR1)* [57] knockout-mice are protected from insulin resistance when fed a high fat diet[57]. The cholinergic receptor nicotinic beta 3 subunit (*CHRNB3*) gene is associated with waist circumference[58] and insulin resistance[59]. A variant near the *SPPL3* gene was previously found to be associated with body fats deposition[60]. One of *SPPL3* PCOS-risk variants reported in our study was predicted to disrupt the binding of the neuronal specific bHLH1 TF, which belongs to a family of TFs involved in neural growth and development[40]. Interestingly, defects in the mouse homologous of the human basic helix-loop-helix 2 (bHLH2) TF cause in mice disruption of the hypothalamic-pituitary axis, obesity, hypogonadism, and infertility[61]. On the other hand, the *FGF1* gene, which we detected as a PCOS-risk gene, encodes the fibroblast growth factor 1 whose serum levels are either increased or decreased in patients with PCOS[62]. In mice, the FGF1 protein lowers hepatic glucose production[63] and *FGF1*-knockout mice have insulin resistance[64].

Three genes in our study (a disintegrin and metalloprotease domain [ADAM] metallopeptidase domain 12 [*ADAM12*], unc-13 homolog C [*UNC13C*] and protein tyrosine phosphatase receptor type M [*PTPRM*]) have roles in other PCOS-related phenotypes: *ADAM12* and *UNC13C* genes are, respectively, associated with serum LH and FSH levels in PCOS[65], and the *PTPRM* gene has a role in the age at first birth and number of born children[66].

The remaining PCOS-risk genes detected in our study are all newly implicated in PCOS and were never reported in any PCOS-related phenotype (i.e., irregular menses, anovulation, infertility, oligomenorrhea, obesity, insulin resistance, T2D, hyperandrogenism, hirsutism, male-pattern baldness) and are the following: protein phosphatase 1 regulatory subunit 12B (*PPP1R12B*), SEC14 and spectrin domain containing 1 (*SESTD1*), atlastin GTPase 2 (*ATL2*), *PROM1*, limb development membrane protein 1 (*LMBR1*), long intergenic non-protein coding RNA 2476 (*LINC02476*), protein tyrosine kinase 2 beta (*PTK2B*), *MLLT3,* potassium sodium-activated channel subfamily T member 1 (*KCNT1*), caspase 5 (*CASP5*), cell division cycle associated 5 (*CDCA5*), protocadherin 9 (*PCDH9*), DDB1 and CUL4 associated factor 4 (*DCAF4*), tropomyosin 1 (*TPM1*), cilia and flagella associated protein 52 (CFAP52), cadherin 7 (CDH7), MYC binding protein 2 (*MYCBP2*), arachidonate 12-lipoxygenase pseudogene 2 (*ALOX12P2*), calcium voltage-gated channel auxiliary subunit gamma 5 (*CACNG5*), coiled-coil domain containing 54 (*CCDC54*), BEN domain containing 7 (*BEND7*), and leucine rich repeat containing 74B (*LRRC74B*). All genes identified in our study are novel and involved in 15 genes' sets pathways, which all have been associated with PCOS[67-74], with the exception of the Fructose-galactose metabolism, Ionotropic glutamate-receptor pathway, Nicotinic acetylcholine-receptor signaling pathway, Synaptic vesicle trafficking, and CCKR-signaling map, which thus are five novel pathways identified in our study. Interestingly, 6% of genes in our study were implicated in the GnRHR pathway, which is directly involved in PCOS[75]. Of further interest, the *PROM1* gene en-

codes CD133 whose circulating level is decreased in amenorrhoeic subjects compared to healthy controls[76]. Specifically, circulating levels of CD133(+) bone marrow-derived stem cells increased with glucose load in healthy females but were significantly reduced in amenorrhoeic women. Oral glucose-induced increase in circulating CD133(+) bone marrow-derived stem cells and endothelial differentiation potential of peripheral blood-derived endothelial progenitor cells were attenuated in insulin resistant amenorrhoeic women[76]. This could explain a reduced potential of cell regeneration in women with oligo- or amenorrhoeic cycles. The *PROM1* gene 3'-UTR variant rs2240688, found to confer PCOS risk in our study, was predicted to create a new binding site for the miRNA hsa-miR-135a (SNP Function Prediction) [36]. miRNAs are known to be involved in the pathogenesis of PCOS[77] and hsa-miR-135a in particular is elevated in patients with endometriosis[78]. Endometriosis is a common condition, representing per a novel hypothesis the other extreme phenotype of PCOS and a PCOS-related diametric (opposite) outcome of variation in the hypothalamic-pituitary gonadal axis development and activity, for which while endometriosis is mediated by low prenatal and postnatal testosterone levels, PCOS is mediated by high prenatal testosterone levels[79].

## Conclusions

This is the first GW-linkage and LD study performed in families with PCOS, and namely from the Italian peninsula. The study reports novel risk variants, genes, and pathways implicated in the risk of PCOS at least in Italian families. Functional and replication studies in other ethnicities are needed.

### Acknowledgments
We thank the families who participated in the study, and we thank Bios Biotech Multi-Diagnostic Health Center, Rome, Italy, for data access and for financial, medical, and laboratory staff support.

### Funding
This publication was supported in part by the funds received under Nebraska Laws 2021, LB 380, Section 109 awarded to C.G. (PI), Creighton University School of Medicine, through the Nebraska Department of Health & Human Services (DHHS). Its contents represent the views of the authors and do not necessarily represent the official views of the State of Nebraska or DHHS.

### Authors' Contribution
C.G. conceived, performed, and supervised the project, including statistical analysis and manuscript drafting and revision. M.A. helped with the bioinformatic analysis, literature search, and manuscript drafting. Both authors have approved the final manuscript.

### Informed Consent
Families were recruited following the Helsinki Declaration guidelines, and individuals provided written informed consent prior to participation.

### Ethics Approval
The Bios Ethical Committee approved this study (Prot. PR/Mg/Cg/311708).

### Availability of Data and Materials
The study data are available on reasonable request, and due to lacking specific patients' consent and privacy restrictions, they are not publicly available.

### Conflict of Interests
The authors have declared that they have no conflicts of interest.

### Authors' Information
C.G. is Professor of Medicine, Chief of Endocrinology, and Endowed Puller Chair at Creighton University School of Medicine, Omaha, NE, and Adjunct Professor of Public Health Sciences, Penn State University College of Medicine, Hershey, PA; M.A. has an MD and a PhD in Genetics (University of Paris) and is a geneticist at Orphanet | INSERM-US14, Paris, France, and member of the executive committee, European Reference Network for the Intellectual Disabilities (ERN-ITHACA) and the International Consortium of Gene Curation Coalition (GenCC).

### ORCID ID
C.G.: 0000-0002-3873-6617
M.A.: 0000-0003-2876-0784

## References

1) Rotterdam EA-SPCWG. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil Steril 2004; 81: 19-25.
2) Coffey S, Mason H. The effect of polycystic ovary syndrome on health-related quality of life. Gynecol Endocrinol 2003; 17: 379-386.
3) Deeks AA, Gibson-Helm ME, Teede HJ. Anxiety and depression in polycystic ovary syndrome: a comprehensive investigation. Fertil Steril 2010; 93: 2421-2423.

4) Gonzalez F. Inflammation in polycystic ovary syndrome: underpinning of insulin resistance and ovarian dysfunction. Steroids 2012; 77: 300-305.
5) Gonzalez F, Rote NS, Minium J, Kirwan JP. Increased activation of nuclear factor kappaB triggers inflammation and insulin resistance in polycystic ovary syndrome. J Clin Endocrinol Metab 2006; 91: 1508-1512.
6) Legro RS, Castracane VD, Kauffman RP. Detecting insulin resistance in polycystic ovary syndrome: purposes and pitfalls. Obstet Gynecol Surv 2004; 59: 141-154.
7) Ehrmann DA, Barnes RB, Rosenfield RL, Cavaghan MK, Imperial J. Prevalence of impaired glucose tolerance and diabetes in women with polycystic ovary syndrome. Diabetes Care 1999; 22: 141-146.
8) Naderpoor N, Shorakae S, Joham A, Boyle J, De Courten B, Teede HJ. Obesity and polycystic ovary syndrome. Minerva Endocrinol 2015; 40: 37-51.
9) Hotamisligil GS. Inflammation and metabolic disorders. Nature 2006; 444: 860-867.
10) Zeng X, Xie YJ, Liu YT, Long SL, Mo ZC. Polycystic ovarian syndrome: Correlation between hyperandrogenism, insulin resistance and obesity. Clin Chim Acta 2020; 502: 214-221.
11) Azziz R. Polycystic Ovary Syndrome. Obstet Gynecol 2018; 132: 321-336.
12) Moran LJ, Noakes M, Clifton PM, Tomlinson L, Galletly C, Norman RJ. Dietary composition in restoring reproductive and metabolic physiology in overweight women with polycystic ovary syndrome. J Clin Endocrinol Metab 2003; 88: 812-819.
13) Nestler JE, Jakubowicz DJ, Evans WS, Pasquali R. Effects of metformin on spontaneous and clomiphene-induced ovulation in the polycystic ovary syndrome. N Engl J Med 1998; 338: 1876-1880.
14) Giallauria F, Palomba S, Maresca L, Vuolo L, Tafuri D, Lombardi G, Colao A, Vigorito C, Francesco O. Exercise training improves autonomic function and inflammatory pattern in women with polycystic ovary syndrome (PCOS). Clin Endocrinol (Oxf) 2008; 69: 792-798.
15) Tan BK, Adya R, Shan X, Aghilla M, Lehnert H, Keay SD, Randeva HS. The anti-atherogenic aspect of metformin treatment in insulin resistant women with the polycystic ovary syndrome: role of the newly established pro-inflammatory adipokine Acute-phase Serum Amyloid A; evidence of an adipose tissue-monocyte axis. Atherosclerosis 2011; 216: 402-408.
16) De Leo V, Musacchio MC, Cappelli V, Massaro MG, Morgante G, Petraglia F. Genetic, hormonal and metabolic aspects of PCOS: an update. Reprod Biol Endocrinol 2016; 14: 38.
17) Stener-Victorin E, Deng Q. Epigenetic inheritance of polycystic ovary syndrome: Challenges and opportunities for treatment. Nat Rev Endocrinol 2021; 17: 521-533.
18) Bruni V, Capozzi A, Lello S. The role of genetics, epigenetics and lifestyle in polycystic ovary syndrome development: The state of the art. Reprod Sci 2022; 29: 668-679.
19) 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: 2100-2104.
20) Chen ZJ, Zhao H, He L, Shi Y, Qin Y, Shi Y, Li Z, You L, Zhao J, Liu J, Liang X, Zhao X, Zhao J, Sun Y, Zhang B, Jiang H, Zhao D, Bian Y, Gao X, Geng L, Li Y, Zhu D, Sun X, Xu JE, Hao C, Ren CE, Zhang Y, Chen S, Zhang W, Yang A, Yan J, Li Y, Ma J, Zhao Y. Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. Nat Genet 2011; 43: 55-59.
21) Shi Y, Zhao H, Shi Y, Cao Y, Yang D, Li Z, Zhang B, Liang X, Li T, Chen J, Shen J, Zhao J, You L, Gao X, Zhu D, Zhao X, Yan Y, Qin Y, Li W, Yan J, Wang Q, Zhao J, Geng L, Ma J, Zhao Y, He G, Zhang A, Zou S, Yang A, Liu J, Li W, Li B, Wan C, Qin Y, Shi J, Yang J, Jiang H, Xu JE, Qi X, Sun Y, Zhang Y, Hao C, Ju X, Zhao D, Ren CE, Li X, Zhang W, Zhang Y, Zhang J, Wu D, Zhang C, He L, Chen ZJ. Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. Nat Genet 2012; 44: 1020-1025.
22) Zhao S, Tian Y, Gao X, Zhang X, Liu H, You L, Cao Y, Su S, Chan W-Y, Sun Y, Zhao H, Chen Z-J. Family-based analysis of eight susceptibility loci in polycystic ovary syndrome. Sci Rep 2015; 5: 12619.
23) Chaudhary H, Patel J, Jain NK, Joshi R. The role of polymorphism in various potential genes on polycystic ovary syndrome susceptibility and pathogenesis. J Ovarian Res 2021; 14: 125.
24) Gorsic LK, Kosova G, Werstein B, Sisk R, Legro RS, Hayes MG, Teixeira JM, Dunaif A, Urbanek M. Pathogenic anti-mullerian hormone variants in polycystic ovary syndrome. J Clin Endocrinol Metab 2017; 102: 2862-2872.
25) Siegel S, Futterweit W, Davies TF, Concepcion ES, Greenberg DA, Villanueva R, Tomer Y. A C/T single nucleotide polymorphism at the tyrosine kinase domain of the insulin receptor gene is associated with polycystic ovary syndrome. Fertil Steril 2002; 78: 1240-1243.
26) Chen ZJ, Shi YH, Zhao YR, Li Y, Tang R, Zhao LX, Chang ZH. [Correlation between single nucleotide polymorphism of insulin receptor gene with polycystic ovary syndrome]. Zhonghua Fu Chan Ke Za Zhi 2004; 39: 582-585.
27) Jin L, Zhu XM, Luo Q, Qian Y, Jin F, Huang HF. A novel SNP at exon 17 of INSR is associated with decreased insulin sensitivity in Chinese women with PCOS. Mol Hum Reprod 2006; 12: 151-155.
28) Ruan Y, Ma J, Xie X. Association of IRS-1 and IRS-2 genes polymorphisms with polycystic ovary syndrome: a meta-analysis. Endocr J 2012; 59: 601-609.
29) Rotterdam EA-SPcwg. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). Hum Reprod 2004; 19: 41-47.
30) Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81: 559-575.

31) Hiekkalinna T, Schaffer AA, Lambert B, Norrgrann P, Goring HH, Terwilliger JD. PSEUDOMARKER: A powerful program for joint linkage and/or linkage disequilibrium analysis on mixtures of singletons and related individuals. Hum Hered 2011; 71: 256-266.
32) Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic acids research 2003; 31: 3812-3814.
33) Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013; Chapter 7:Unit7.20. doi: 10.1002/0471142905.hg0720s76.
34) Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nature Methods 2010; 7: 575-576.
35) Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, Kosmicki JA, Arbelaez J, Cui W, Schwartz GB, Chow ED, Kanterakis E, Gao H, Kia A, Batzoglou S, Sanders SJ, Farh KKH. Predicting Splicing from Primary Sequence with Deep Learning. Cell 2019; 176: 535-548.e24.
36) Xu Z, Taylor JA. SNPinfo: Integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. Nucleic Acids Res 2009; 37: W600-W605.
37) Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012; 22: 1790-1797.
38) Liu C, Zhang F, Li T, Lu M, Wang L, Yue W, Zhang D. MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs. BMC Genomics 2012; 13: 1-10.
39) Mi H, Thomas P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. Methods Mol Biol 2009; 563: 123-140.
40) Dennis DJ, Han S, Schuurmans C. bHLH transcription factors in neural development, disease, and reprogramming. Brain Res 2019; 1705: 48-65.
41) Patel S. Polycystic ovary syndrome (PCOS), an inflammatory, systemic, lifestyle endocrinopathy. J Steroid Biochem Mol Biol 2018; 182: 27-36.
42) McCarthy MI. Genomics, type 2 diabetes, and obesity. N Engl J Med 2010; 363: 2339-2350.
43) Wheeler E, Huang N, Bochukova EG, Keogh JM, Lindsay S, Garg S, Henning E, Blackburn H, Loos RJ, Wareham NJ, O'Rahilly S, Hurles ME, Barroso I, Farooqi IS. Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. Nat Genet 2013; 45: 513-517.
44) Perry JR, Day F, Elks CE, Sulem P, Thompson DJ, Ferreira T, He C, Chasman DI, Esko T, Thorleifsson G, Albrecht E, Ang WQ, Corre T, Cousminer DL, Feenstra B, Franceschini N, Ganna A, Johnson AD, Kjellqvist S, Lunetta KL, McMahon G, Nolte IM, Paternoster L, Porcu E, Smith AV, Stolk L, Teumer A, Tšernikova N, Tikkanen E, Ulivi S, Wagner EK, Amin N, Bierut LJ, Byrne EM, Hottenga JJ, Koller DL, Mangino M, Pers TH, Yerges-Armstrong LM, Zhao JH, Andrulis IL, Anton-Culver H, Atsma F, Bandinelli S, Beckmann MW, Benitez J, Blomqvist C, Bojesen SE, Bolla MK, Bonanni B, Brauch H, Brenner H, Buring JE, Chang-Claude J, Chanock S, Chen J, Chenevix-Trench G, Collée JM, Couch FJ, Couper D, Coveillo AD, Cox A, Czene K, D'Adamo A P, Smith GD, De Vivo I, Demerath EW, Dennis J, Devilee P, Dieffenbach AK, Dunning AM, Eiriksdottir G, Eriksson JG, Fasching PA, Ferrucci L, Flesch-Janys D, Flyger H, Foroud T, Franke L, Garcia ME, García-Closas M, Geller F, de Geus EE, Giles GG, Gudbjartsson DF, Gudnason V, Guénel P, Guo S, Hall P, Hamann U, Haring R, Hartman CA, Heath AC, Hofman A, Hooning MJ, Hopper JL, Hu FB, Hunter DJ, Karasik D, Kiel DP, Knight JA, Kosma VM, Kutalik Z, Lai S, Lambrechts D, Lindblom A, Mägi R, Magnusson PK, Mannermaa A, Martin NG, Masson G, McArdle PF, McArdle WL, Melbye M, Michailidou K, Mihailov E, Milani L, Milne RL, Nevanlinna H, Neven P, Nohr EA, Oldehinkel AJ, Oostra BA, Palotie A, Peacock M, Pedersen NL, Peterlongo P, Peto J, Pharoah PD, Postma DS, Pouta A, Pylkäs K, Radice P, Ring S, Rivadeneira F, Robino A, Rose LM, Rudolph A, Salomaa V, Sanna S, Schlessinger D, Schmidt MK, Southey MC, Sovio U, Stampfer MJ, Stöckl D, Storniolo AM, Timpson NJ, Tyrer J, Visser JA, Vollenweider P, Völzke H, Waeber G, Waldenberger M, Wallaschofski H, Wang Q, Willemsen G, Winqvist R, Wolffenbuttel BH, Wright MJ, Boomsma DI, Econs MJ, Khaw KT, Loos RJ, McCarthy MI, Montgomery GW, Rice JP, Streeten EA, Thorsteinsdottir U, van Duijn CM, Alizadeh BZ, Bergmann S, Boerwinkle E, Boyd HA, Crisponi L, Gasparini P, Gieger C, Harris TB, Ingelsson E, Järvelin MR, Kraft P, Lawlor D, Metspalu A, Pennell CE, Ridker PM, Snieder H, Sørensen TI, Spector TD, Strachan DP, Uitterlinden AG, Wareham NJ, Widen E, Zygmunt M, Murray A, Easton DF, Stefansson K, Murabito JM, Ong KK. Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature 2014; 514: 92-97.
45) Seow KM, Chang YW, Chen KH, Juan CC, Huang CY, Lin LT, Tsui KH, Chen YJ, Lee WL, Wang PH. Molecular mechanisms of laparoscopic ovarian drilling and its therapeutic effects in polycystic ovary syndrome. Int J Mol Sci 2020; 21: 8147.
46) Kiddy DS, Sharp PS, White DM, Scanlon MF, Mason HD, Bray CS, Polson DW, Reed MJ, Franks S. Differences in clinical and endocrine features between obese and non-obese subjects with polycystic ovary syndrome: an analysis of 263 consecutive cases. Clin Endocrinol (Oxf) 1990; 32: 213-220.
47) Lim SS, Norman RJ, Davies MJ, Moran LJ. The effect of obesity on polycystic ovary syndrome: a systematic review and meta-analysis. Obes Rev 2013; 14: 95-109.

48) Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Mägi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JR, Egan JM, Lajunen T, Grarup N, Sparsø T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P, Cavalcanti-Proença C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G, Ingelsson E, An P, O'Connell J, Luan J, Elliott A, McCarroll SA, Payne F, Roccasecca RM, Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y, Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Böttcher Y, Brunner E, Bumpstead SJ, Charpentier G, Chen YD, Chines P, Clarke R, Coin LJ, Cooper MN, Cornelis M, Crawford G, Crisponi L, Day IN, de Geus EJ, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A, Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ, Grundy S, Gwilliam R, Gyllensten U, Hadjadj S, Hallmans G, Hammond N, Han X, Hartikainen AL, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA, Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PR, Jørgensen T, Jula A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimaki M, Knight B, Koskinen S, Kovacs P, Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y, Lyssenko V, Mahley R, Mangino M, Manning AK, Martínez-Larrad MT, McAteer JB, McCulloch LJ, McPherson R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S, Narisu N, Neville MJ, Oostra BA, Orrù M, Pakyz R, Palmer CN, Paolisso G, Pattaro C, Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AF, Pichler I, Polasek O, Posthuma D, Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S, Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet P, Scott LJ, Seedorf U, Sharp SJ, Shields B, Sigurethsson G, Sijbrands EJ, Silveira A, Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvänen AC, Tanaka T, Thorand B, Tichet J, Tönjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M, Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A, Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemsen G, Witteman JC, Yarnell JW, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC, Borecki IB, Loos RJ, Meneton P, Magnusson PK, Nathan DM, Williams GH, Hattersley AT, Silander K, Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper C, Dedoussis GV, Serrano-Ríos M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW, Ebrahim S, Marmot M, Kao WH, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T, Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M, Campbell H, Wilson JF, Bergman RN, Buchanan TA, Collins FS, Mohlke KL, Tuomilehto J, Valle TT, Altshuler D, Rotter JI, Siscovick DS, Penninx BW, Boomsma DI, Deloukas P, Spector TD, Frayling TM, Ferrucci L, Kong A, Thorsteinsdottir U, Stefansson K, van Duijn CM, Aulchenko YS, Cao A, Scuteri A, Schlessinger D, Uda M, Ruokonen A, Jarvelin MR, Waterworth DM, Vollenweider P, Peltonen L, Mooser V, Abecasis GR, Wareham NJ, Sladek R, Froguel P, Watanabe RM, Meigs JB, Groop L, Boehnke M, McCarthy MI, Florez JC, Barroso I. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 2010; 42: 105-116.

49) Arner P, Sahlqvist AS, Sinha I, Xu H, Yao X, Waterworth D, Rajpal D, Loomis AK, Freudenberg JM, Johnson T, Thorell A, Näslund E, Ryden M, Dahlman I. The epigenetic signature of systemic insulin resistance in obese women. Diabetologia 2016; 59: 2393-2405.

50) Gjesing AP, Nielsen AA, Brandslund I, Christensen C, Sandbæk A, Jørgensen T, Witte D, Bonnefond A, Froguel P, Hansen T, Pedersen O. Studies of a genetic variant in HK1 in relation to quantitative metabolic traits and to the prevalence of type 2 diabetes. BMC Med Genet 2011; 12: 99.

51) Fadista J, Vikman P, Laakso EO, Mollet IG, Esguerra JL, Taneera J, Storm P, Osmark P, Ladenvall C, Prasad RB, Hansson KB, Finotello F, Uvebrant K, Ofori JK, Di Camillo B, Krus U, Cilio CM, Hansson O, Eliasson L, Rosengren AH, Renström E, Wollheim CB, Groop L. Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. Proc Natl Acad Sci U S A 2014; 111: 13924-13929.

52) Saeedi Borujeni MJ, Esfandiary E, Baradaran A, Valiani A, Ghanadian M, Codoñer-Franch P, Basirat R, Alonso-Iglesias E, Mirzaei H, Yazdani A. Molecular aspects of pancreatic β-cell dysfunction: Oxidative stress, microRNA, and long noncoding RNA. J Cell Physiol 2019; 234: 8411-8425.

53) Li L, Pan Z, Yang S, Shan W, Yang Y. Identification of key gene pathways and coexpression networks of islets in human type 2 diabetes. Diabetes Metab Syndr Obes 2018; 11: 553-563.

54) Dayeh T, Volkov P, Salö S, Hall E, Nilsson E, Olsson AH, Kirkpatrick CL, Wollheim CB, Eliasson L, Rönn T, Bacos K, Ling C. Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion. PLoS Genet 2014; 10: e1004160.

55) Lv B, Yang X, An T, Wu Y, He Z, Li B, Wang Y, Tan F, Wang T, Zhu J, Hu Y, Liu X, Jiang G. Combined analysis of whole-exome sequencing and RNA sequencing in type 2 diabetes mellitus patients with thirst and fatigue. Diabetol Metab Syndr 2022; 14: 111.

56) Zhou Q, Sun WW, Chen JC, Zhang HL, Liu J, Lin Y, Lin PC, Wu BX, An YP, Huang L, Sun WX, Zhou XW, Li YM, Yuan YY, Zhao JY, Xu W, Zhao SM. Phenylalanine impairs insulin signaling and inhibits glucose uptake through modification of IRβ. Nat Commun 2022; 13: 4291.

57) Ying L, Zhang M, Ma X, Si Y, Li X, Su J, Yin J, Bao Y. Macrophage LAMTOR1 deficiency prevents dietary obesity and insulin resistance through inflammation-induced energy expenditure. Front Cell Dev Biol 2021; 9: 672032.
58) Zhu Y, Yang J, Yeh F, Cole SA, Haack K, Lee ET, Howard BV, Zhao J. Joint association of nicotinic acetylcholine receptor variants with abdominal obesity in American Indians: The Strong Heart Family Study. PLoS One 2014; 9: e102220.
59) Yang J, Zhu Y, Cole SA, Haack K, Zhang Y, Beebe LA, Howard BV, Best LG, Devereux RB, Henderson JA, Henderson P, Lee ET, Zhao J. A gene-family analysis of 61 genetic variants in the nicotinic acetylcholine receptor genes for insulin resistance and type 2 diabetes in American Indians. Diabetes 2012; 61: 1888-1894.
60) Agrawal S, Wang M, Klarqvist MDR, Smith K, Shin J, Dashti H, Diamant N, Choi SH, Jurgens SJ, Ellinor PT, Philippakis A, Claussnitzer M, Ng K, Udler MS, Batra P, Khera AV. Inherited basis of visceral, abdominal subcutaneous and gluteofemoral fat depots. Nat Commun 2022; 13: 3771.
61) Good DJ, Porter FD, Mahon KA, Parlow AF, Westphal H, Kirsch IR. Hypogonadism and obesity in mice with a targeted deletion of the Nhlh2 gene. Nat Genet 1997; 15: 397-401.
62) Hart RJ. Physiological Aspects of Female Fertility: Role of the environment, modern lifestyle, and genetics. Physiol Rev 2016; 96: 873-909.
63) Sancar G, Liu S, Gasser E, Alvarez JG, Moutos C, Kim K, van Zutphen T, Wang Y, Huddy TF, Ross B, Dai Y, Zepeda D, Collins B, Tilley E, Kolar MJ, Yu RT, Atkins AR, van Dijk TH, Saghatelian A, Jonker JW, Downes M, Evans RM. FGF1 and insulin control lipolysis by convergent pathways. Cell Metab 2022; 34: 171-183.e6.
64) Suh JM, Jonker JW, Ahmadian M, Goetz R, Lackey D, Osborn O, Huang Z, Liu W, Yoshihara E, van Dijk TH, Havinga R, Fan W, Yin YQ, Yu RT, Liddle C, Atkins AR, Olefsky JM, Mohammadi M, Downes M, Evans RM. Endocrinization of FGF1 produces a neomorphic and potent insulin sensitizer. Nature 2014; 513: 436-439.
65) Hayes MG, Urbanek M, Ehrmann DA, Armstrong LL, Lee JY, Sisk R, Karaderi T, Barber TM, McCarthy MI, Franks S, Lindgren CM, Welt CK, Diamanti-Kandarakis E, Panidis D, Goodarzi MO, Azziz R, Zhang Y, James RG, Olivier M, Kissebah AH, Stener-Victorin E, Legro RS, Dunaif A. Genome-wide association of polycystic ovary syndrome implicates alterations in gonadotropin secretion in European ancestry populations. Nat Commun 2015; 6: 7502.
66) Barban N, Jansen R, de Vlaming R, Vaez A, Mandemakers JJ, Tropf FC, Shen X, Wilson JF, Chasman DI, Nolte IM, Tragante V, van der Laan SW, Perry JR, Kong A, Ahluwalia TS, Albrecht E, Yerges-Armstrong L, Atzmon G, Auro K, Ayers K, Bakshi A, Ben-Avraham D, Berger K, Bergman A, Bertram L, Bielak LF, Bjornsdottir G, Bonder MJ, Broer L, Bui M, Barbieri C, Cavadino A, Chavarro JE, Turman C, Concas MP, Cordell HJ, Davies G, Eibich P, Eriksson N, Esko T, Eriksson J, Falahi F, Felix JF, Fontana MA, Franke L, Gandin I, Gaskins AJ, Gieger C, Gunderson EP, Guo X, Hayward C, He C, Hofer E, Huang H, Joshi PK, Kanoni S, Karlsson R, Kiechl S, Kifley A, Kluttig A, Kraft P, Lagou V, Lecoeur C, Lahti J, Li-Gao R, Lind PA, Liu T, Makalic E, Mamasoula C, Matteson L, Mbarek H, McArdle PF, McMahon G, Meddens SF, Mihailov E, Miller M, Missmer SA, Monnereau C, van der Most PJ, Myhre R, Nalls MA, Nutile T, Kalafati IP, Porcu E, Prokopenko I, Rajan KB, Rich-Edwards J, Rietveld CA, Robino A, Rose LM, Rueedi R, Ryan KA, Saba Y, Schmidt D, Smith JA, Stolk L, Streeten E, Tönjes A, Thorleifsson G, Ulivi S, Wedenoja J, Wellmann J, Willeit P, Yao J, Yengo L, Zhao JH, Zhao W, Zhernakova DV, Amin N, Andrews H, Balkau B, Barzilai N, Bergmann S, Biino G, Bisgaard H, Bønnelykke K, Boomsma DI, Buring JE, Campbell H, Cappellani S, Ciullo M, Cox SR, Cucca F, Toniolo D, Davey-Smith G, Deary IJ, Dedoussis G, Deloukas P, van Duijn CM, de Geus EJ, Eriksson JG, Evans DA, Faul JD, Sala CF, Froguel P, Gasparini P, Girotto G, Grabe HJ, Greiser KH, Groenen PJ, de Haan HG, Haerting J, Harris TB, Heath AC, Heikkilä K, Hofman A, Homuth G, Holliday EG, Hopper J, Hyppönen E, Jacobsson B, Jaddoe VW, Johannesson M, Jugessur A, Kähönen M, Kajantie E, Kardia SL, Keavney B, Kolcic I, Koponen P, Kovacs P, Kronenberg F, Kutalik Z, La Bianca M, Lachance G, Iacono WG, Lai S, Lehtimäki T, Liewald DC, Lindgren CM, Liu Y, Luben R, Lucht M, Luoto R, Magnus P, Magnusson PK, Martin NG, McGue M, McQuillan R, Medland SE, Meisinger C, Mellström D, Metspalu A, Traglia M, Milani L, Mitchell P, Montgomery GW, Mook-Kanamori D, de Mutsert R, Nohr EA, Ohlsson C, Olsen J, Ong KK, Paternoster L, Pattie A, Penninx BW, Perola M, Peyser PA, Pirastu M, Polasek O, Power C, Kaprio J, Raffel LJ, Räikkönen K, Raitakari O, Ridker PM, Ring SM, Roll K, Rudan I, Ruggiero D, Rujescu D, Salomaa V, Schlessinger D, Schmidt H, Schmidt R, Schupf N, Smit J, Sorice R, Spector TD, Starr JM, Stöckl D, Strauch K, Stumvoll M, Swertz MA, Thorsteinsdottir U, Thurik AR, Timpson NJ, Tung JY, Uitterlinden AG, Vaccargiu S, Viikari J, Vitart V, Völzke H, Vollenweider P, Vuckovic D, Waage J, Wagner GG, Wang JJ, Wareham NJ, Weir DR, Willemsen G, Willeit J, Wright AF, Zondervan KT, Stefansson K, Krueger RF, Lee JJ, Benjamin DJ, Cesarini D, Koellinger PD, den Hoed M, Snieder H, Mills MC. Genome-wide analysis identifies 12 loci influencing human reproductive behavior. Nat Genet 2016; 48: 1462-1472.
67) Cao J, Huo P, Cui K, Wei H, Cao J, Wang J, Liu Q, Lei X, Zhang S. Follicular fluid-derived exosomal miR-143-3p/miR-155-5p regulate follicular dysplasia by modulating glycolysis in granulosa cells in polycystic ovary syndrome. Cell Commun Signal 2022; 20: 61.
68) Jin L, Ren L, Lu J, Wen X, Zhuang S, Geng T, Zhang Y. CXCL12 and its receptors regulate granulosa cell apoptosis in PCOS rats and human KGN tumor cells. Reproduction 2021; 161: 145-157.

69) Aydos A, Gurel A, Oztemur Islakoglu Y, Noyan S, Gokce B, Ecemis T, Kaya C, Aksu AT, Gur Dedeoglu B. Identification of polycystic ovary syndrome (PCOS) specific genes in cumulus and mural granulosa cells. PLoS One 2016; 11: e0168875.
70) Hu M, Zhang Y, Li X, Cui P, Li J, Brännström M, Shao LR, Billig H. Alterations of endometrial epithelial-mesenchymal transition and MAPK signalling components in women with PCOS are partially modulated by metformin in vitro. Mol Hum Reprod 2020; 26: 312-326.
71) Laczko R, Csiszar K. Lysyl Oxidase (LOX): Functional contributions to signaling pathways. Biomolecules 2020; 10: 1093.
72) Naseri R, Alimoradi Y, Sohrabi M, Cheraghian Fard M, Barzingarosi E, Abdolmaleki A, Jalili C. H6PD gene polymorphisms (R453Q and D151A) and polycystic ovary syndrome: A case-control study in a population of Iranian Kurdish Women. Int J Fertil Steril 2022; 16: 180-183.
73) Patil K, Joseph S, Shah J, Mukherjee S. An integrated in silico analysis highlighted angiogenesis regulating miRNA-mRNA network in PCOS pathophysiology. J Assist Reprod Genet 2022; 39: 427-440.
74) Rehman K, Haider K, Akash MSH. Cigarette smoking and nicotine exposure contributes for aberrant insulin signaling and cardiometabolic disorders. Eur J Pharmacol 2021; 909: 174410.
75) Tata B, Mimouni NEH, Barbotin AL, Malone SA, Loyens A, Pigny P, Dewailly D, Catteau-Jonard S, Sundström-Poromaa I, Piltonen TT, Dal Bello F, Medana C, Prevot V, Clasadonte J, Giacobini P. Elevated prenatal anti-Müllerian hormone reprograms the fetus and induces polycystic ovary syndrome in adulthood. Nat Med 2018; 24: 834-846.
76) Bairagi S, Gopal J, Nathan AA, Babu SS, Kumar NP, Dixit M. Glucose-induced increase in circulating progenitor cells is blunted in polycystic amenorrhoeic subjects. Hum Reprod 2012; 27: 844-853.
77) Chen B, Xu P, Wang J, Zhang C. The role of MiRNA in polycystic ovary syndrome (PCOS). Gene 2019; 706: 91-96.
78) Perricos A, Proestling K, Husslein H, Kuessel L, Hudson QJ, Wenzl R, Yotova I. Hsa-mir-135a shows potential as a putative diagnostic biomarker in saliva and plasma for endometriosis. Biomolecules 2022; 12: 1144.
79) Dinsdale NL, Crespi BJ. Endometriosis and polycystic ovary syndrome are diametric disorders. Evol Appl 2021; 14: 1693-1715.