GMR
The Original

REPLICATION STUDY REVEALS ASSOCIATION OF SELECTED GWAS GLOBAL VARIANTS WITH PCOS IN INDIAN WOMEN

Kuppachi Himaja[1], Janani Dakshinamoorthy[1], Pavithra Selvanathan[1], Preethi Sundar[1], Poornima Gopalakrishnan[1], Bhuvaneswari Karthik Ganesh[1], Usha Balasundaram[1*]

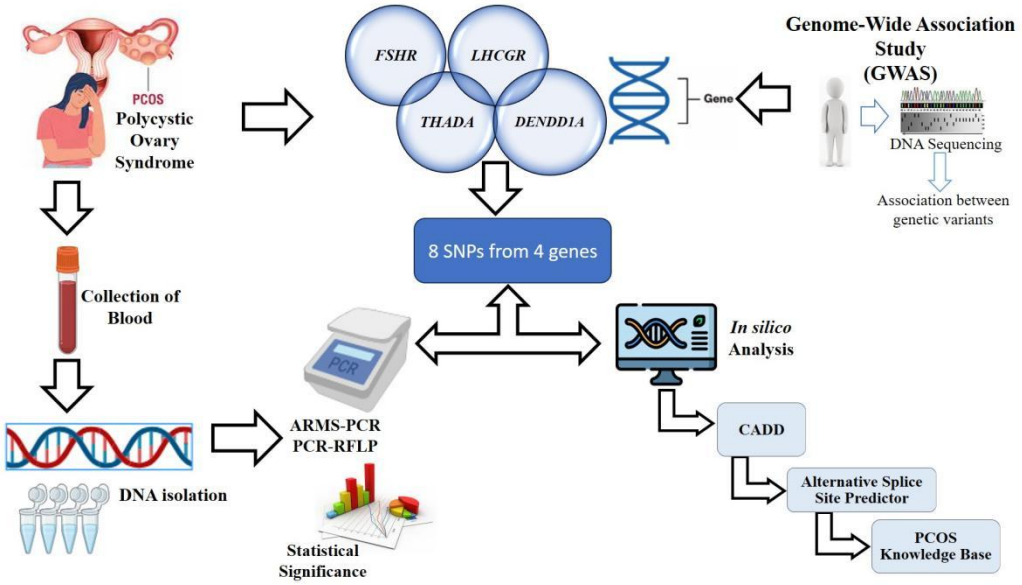
[1]Medicinal Plants and PCOS Research Lab, Department of Genetic Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India.
[*]**Corresponding author**: Usha Balasundaram, E-mail: sundaram.usha@gmail.com, ushab@srmist.edu.in, Orcid ID: https://orcid.org/0000-0003-3039-9131

ABSTRACT

Polycystic ovary syndrome (PCOS) is a complex metabolic and endocrine disorder affecting reproductive-age women. Diagnosis is challenging in adolescents due to characteristic symptoms after menarche. Hence, it is important to identify early markers for an accurate diagnosis of PCOS in teenage girls. This study identifies eight variants from four GWAS genes associated with PCOS in the Indian population. About 250 PCOS and 300 control Indian women aged 18-34 were recruited, and their blood samples, written informed consent, and sociodemographic data were obtained. Genomic DNA was isolated from peripheral vein blood and genotyped for variants from *FSHR*, *LHCGR*, and *THADA* genes by ARMS-PCR and *DENND1A* by PCR-RFLP. Odds ratios (OR) were calculated to ascertain the association of selected variants with PCOS, with confidence intervals maintained at 95%. Of the eight intronic GWAS variants, five variants, rs11692782, rs2349415 from *FSHR*, rs10818854 from *DENND1A*, rs12478601, rs13429458 from *THADA* genes were found to be associated with PCOS women in India. In-silico study showed rs13429458 of *THADA* being CADD-tolerated and others are non-pathogenic. This study identifies five SNPs associated with PCOS in Indian women, which were previously reported to show association in other ethnic populations, and thus may be useful in early diagnosis of PCOS.

KEYWORDS: Polycystic ovary syndrome, SNPs, variant, ARMS-PCR, PCR-RFLP.

Graphical Abstract



22Proof 22 , 22

Article Highlights

- PCOS is one of the endocrinological and gynecological disorders that affect women of child-bearing age and GWAS has previously reported many susceptibility loci associated with PCOS.
- The genes selected for this study were based on the pathophysiological importance of PCOS and some of the selected variants were studied in Asian and/or European populations.
- Genotyping the selected variants and logistic regression analysis revealed that rs11692782, rs2349415 from *FSHR*, rs12478601, rs13429458 from *THADA*, and rs10818854 from *DENND1A* are associated with PCOS in Indian women.

Abbreviations

95% CI- 95% Confidence Interval; A- Adenine; ARMS-PCR- Amplificatory Refractory Mutation System- Polymerase Chain Reaction; ASSP- Alternative Splice Site Predictor tool; C- Cytosine; C11orf41- Chromosome 11 open reading

frame 41; CADD- Combined Annotation Dependent Depletion; CYP1A1- cytochrome P1A1; CYP17A1- CYtochrome P17A1, CYP19- CYtochrome P 19; DENND1A- Differentially Expressed in Normal Neoplastic Domain- containing 1; DNA- Deoxyribo Nucleic Acid; dbSNP- Single Nucleotide Polymorphism Database; FSH- Follicle Stimulating Hormone; FSHR- Follicle Stimulating Hormone Receptor; G- Guanine; GITC- Guanidinium Thiocyanante; GWAS- Genome-Wide Association Studies; HSD17B6- 17β-Hydroxy Steroid Dehydrogenase; HMGA2- High-Mobility Group AT-hook 2; HWE- Hardy-Weinberg Equilibrium; INSR- INsulin Receptor; INS- Insulin; ICMR- Indian Council of Medical Research; LH- Luteinizing Hormone; LHCGR- Luteinizing Hormone/ChorioGonadotropin Receptor; NCBI- National Centre for Biotechnological Information; NIH- National Institutes of Health; OR- Odds Ratio; PCOS- Polycystic Ovary Syndrome; PCR-RFLP- Polymerase Chain Reaction- Restriction Fragment Length Polymorphism; PEX5L- Peroxisomal Biogenesis Factor 5-like; PPAR-γ- Peroxisome Proliferator-activated Receptor Gamma; RAB5B- Ras-related protein Rab-5B; RBC- Red Blood Cells; SDS- Sodium Dodecyl Sulphate; SHBG- Sex Hormone-Binding Globulin; SNP- Single Nucleotide Polymorphism; SUOX- Sulfite Oxidase; SUMO1P1- Small Ubiquitin-related MOdifier 5; SOX8- Sex Determining Region Y-box 8; T- Thymine; TE- Tris-EDTA; THADA- THyroid ADenoma Associated gene; TOX3- TOX High Mobility Group box family member 3; TRIML2- TRIpartite motif family-like 2; WBC - White Blood Cells; YAP1- Yes-Associated Protein 1.

1.INTRODUCTION

Polycystic ovary syndrome (PCOS) is a multifactorial endocrinological and gynaecological disorder among women of reproductive age. PCOS is diagnosed by one of the three criteria; Rotterdam criteria, AE-PCOS and NIH, 1990. According to the widely followed Rotterdam criteria, women with two of the three symptoms - oligomenorrhoea, hyperandrogenism and PCO morphology on an ultrasound imaging - are diagnosed to have PCOS. Other manifestations are hyperinsulinemia and metabolic changes, further leading to type 2 diabetes mellitus, hirsutism, acne, acanthosis nigricans, infertility, cardiovascular disease, and sometimes gonadal cancers. Genetic and lifestyle factors are known to influence the severity of the syndrome [1].

The evidence that PCOS is a polygenic disorder involving dysregulated genetic factors comes from several case-control and genome-wide association studies. Candidate gene approaches in cohorts from different parts of the world have put forth genetic variants in more than hundred genes to be associated with PCOS. Some of these include *PPAR-γ, CYP1A1, CYP17A1, CYP19, HSD17B6, SHBG, FSHR, LHCGR, THADA, INS, INSR,* etc [2]. However, the majority of these results were not replicated in other ethnicities.

Genome-wide association studies (GWAS) have the advantage of studying numerous variants in one go and thus help in identifying novel variants. GWAS was first done in the Han-Chinese PCOS population and replicated in a larger cohort. Combined results of the two studies in a total of 8,226 cases and 7,578 age-matched controls revealed 11 susceptibility loci 2p16.3, 2p21, 9q33.3, 9q22.32, 11q22.1, 12q13.2, 12q14.3, 16q12.1, 19p13.3, 20q13.2 and 2p16.3 in the genes, *LHCGR, THADA, DENND1A, C9orf3, YAP1, RAB5B, SUOX, HMGA2, TOX3, INSR, SUMO1P1* and *FSHR* respectively [3,4]. Some of these variants in *THADA* and *DENDD1A* were also confirmed to be associated with PCOS in European women [5]. In a recent study, GWAS of Korean patients identified one novel locus 8q24.2 which showed a strong association and seven other weakly associated loci, 1p22, 3q26.33, 4p12, 4q35.2, 9q21.32, 11p13 and 16p13.3 in the genes *LINC01709, PEX5L, GABRA4, TRIML2, RASEF, C11orf41* and *SOX8* respectively [6].

Replication studies carried out in Indian, Han-Chinese, European, Arabian, Korean, Asian, and Iranians for one or other variants from GWAS revealed that the genes *DENND1A, THADA, FSHR, LHCGR,* and *INSR* are associated with PCOS [7]. However, the results are contradictory in different populations, indicating that differences in ethnicity may be a causal factor for variation in gene susceptibility.

In India, association studies for *THADA* and *LHCGR* variants- rs13429458 and rs13405728 respectively, were done in PCOS women in this, *THADA* variant rs13429458 was not statistically significant in the study done by Dadachanji et al., while Vishnubotla et al., showed significant association [8,9]. Replication studies in other ethnic populations will support the use of these variants as global markers for the early diagnosis of PCOS. Also, personalized treatment may be warranted for the women at risk and the SNPs can be used for wellness gene panels.

In our previous study, we have shown a significant association of *INSR* GWAS variants rs1799817 and rs2059807 to Indian PCOS women [10]. In this study, eight GWAS variants from four genes *FSHR, LHCGR, DENND1A* and *THADA*, were genotyped in Indian women.

2. MATERIALS AND METHODS

2.1 Participants

For this case-control study, 250 cases and 300 age and ethnicity-matched controls were recruited from two hospitals in the metro city, Chennai, Tamil Nadu, India. Only women in the reproductive age group between 18 and 34 years were recruited for this prospective analysis. Diagnosis for PCOS was made based on the Rotterdam criteria as mentioned previously. Women with hyper/hypothyroidism, Cushing syndrome, diabetes, and exogenous androgen exposure were excluded from the study. Control samples were collected from the women who visited the hospital for regular health checkups. The inclusion criteria for choosing control samples were (i) should not have any history of irregular periods (ii) no history of PCOS in the family members and iii) women with no hypo/hyperthyroidism.

2.2 Ethics Statement and blood collection

Approval for the study was obtained from the Institutional Ethical Committee (No. 955/IEC/2016) to draw the peripheral blood from the patients. According to the "National Ethical Guidelines for Biomedical & Health Research Involving Human Participants (2017)" put forward by the Indian Council of Medical Research (ICMR), the Government of India and consent from the participants to enroll in the study was obtained in a written format. Blood was collected on day 2 of the last menstrual cycle of the cohort. Also, the socio-demographic data were obtained from the participants using a questionnaire. Biochemical data were collected from their clinical reports.

2.3 DNA Isolation

Genomic DNA from the blood samples was isolated by the salting-out method or by modified Miller's method [11]. In short, RBCs were lysed using lysis buffer following disruption of WBCs with 10% SDS and GITC buffer. After the degradation of proteins with proteinase K, the tubes were centrifuged to collect the aqueous supernatant containing the DNA. Following further purification steps, the DNA was precipitated with isopropyl alcohol, and washed with 70% ethanol. The pellet obtained after centrifugation was air-dried and resuspended in TE buffer and stored at -20°C until further use. The quality of the isolated DNA was assessed in agarose gel electrophoresis and quantified using a nano-spectrophotometer (Thermo Fisher Ltd., USA).

2.4 Genotyping

Variants rs11692782, rs2268361, rs2349415 of *FSHR*, rs13405728 of *LHCGR*, rs1247860, rs13429458 of *THADA* were genotyped by ARMS-PCR while *DENDD1A* variants rs2479106, rs10818854 were analyzed using PCR-RFLP. For genotyping, about 50-100 ng of the total genomic DNA was used as a template for amplification of the target region using primers. Primers and enzymes used in ARMS-PCR and PCR-RFLP are given in Table 1 along with the annealing temperatures of each primer set. Tetra ARMS-PCR primers were designed using Primer 1 software (http://primer1.soton.ac.uk/primer1.html) and the enzymes for RFLP were selected using enzyme finder tool (https://enzymefinder.neb.com/#!#nebheader). The NCBI primer blast software was used to check the best primer pairs; the parameters checked were specificity, annealing temperature, GC content, and the complementarity of primers. The primers were synthesized at Bioserve Biotechnologies India Pvt. Ltd, reconstituted in TE buffer and stored at -20°C. Genotyping data was validated by sequencing the randomly picked samples on SeqStudio Genetic Analyser, Thermo Fisher Scientific, USA.

Table 1: Genotyping method and primer sequences for the selected variants.

S. No.	SNP ID	Method	Primer Sequence (5'-3')	Annealing temperature
1	rs11692782 (*FSHR*)	ARMS-PCR	Outer FP:5'TTTGGAGTTTCAAGTCTCCATAACTG3' Outer RP:5'AAAATTTCACTTCATAAAATCCTGC3' Inner FP:5'TTCTACATTGTGATCTTGGCTTTGAT3' Inner RP:5'AGTTTCTCAGATCCCTTGGTTATCCT3'	54°C
2	rs2268361 (*FSHR*)	ARMS-PCR	Outer FP:5'CAAAATTTGGAAGTCTTGACAGG3' Outer RP:5'TGTTTAGGTGTTAGATTGGGAAG3' Inner FP:5'TGAAAGATCTCAATCCCAGGGCATTA3' Inner RP:5'GATGCTGTGAGACGAAGGCATCTTATG3'	54°C
3	rs2349415 (*FSHR*)	ARMS-PCR	Outer FP:5'TAAATACTTTTTTGCCCTTGCTGTGG3' Outer RP:5'TATAAGTGAGAATATGCGGTGTTTGG3' Inner FP:5'TAAAAACAGGTGTCAGGCTGGATTTTAC3' Inner RP:5'CAGTGTCACTGAACTACAGCCAATTGA3'	56°C
4	rs13405728 (*LHCGR*)	ARMS-PCR	Outer FP:5'CTGGGAACCCTGAGTTGAATGACATTT3' Outer RP:5'ATTAGGAGTCACTTCCCCATGAAGCCTT3' Inner FP:5'TGCTCTGGCAGAAGAGGCACATGTCGG3' Inner RP:5'ATATCTCACCATAATGCAGCCATTTATT3'	54°C
5	rs12478601 (*THADA*)	ARMS-PCR	Outer FP:5' TAGAAGATGCACCATGTACAAGAC3' Outer RP:5' TCCAACTCCAGAATGTGTGCTATTG3' Inner FP:5' CATTCCTGCTGGTCTTGGTTAGTACAAC3' Inner RP:5' CATTCCTGCTGGTCTTGGTTAGTACAAC3'	50°C
6	rs13429458 (*THADA*)	ARMS-PCR	Outer FP:5'TCAGCGGTATGATTTCGTAGTGG3' Outer RP:5'AAGACTTGAAGGCAATGTGATTC3' Inner FP:5'GCTGTGCAAAGTTAGAAGATGAAGCA3' InnerRP:5'GGCAGGGTATAGGTGTATGTAATCAGTCTG3'	54°C
7	rs2479106 (*DENDD1A*)	PCR-RFLP	FP: 5'CCTACATGAGTTAAGCCACCA3' RP: 5'AAGCCACAGACCTAAACCTTC3' Enzyme for RFLP: *Spe* 1	52.8°C

8	rs10818854 (*DENDD1A*)	PCR-RFLP	FP: 5' TTGGTGCTAAGGTTGTGGTTT3' RP:5'CCTCTAATCTGCAGTGCCACA3' Enzyme for RFLP: *Dde* 1	54.7°C

FP- Forward Primer, RP-Reverse Primer

2.5 In-silico pathogenicity scoring of SNPs

Combined Annotation Dependent Depletion (CADD) (https://cadd.gs.washington.edu/) is a tool used for scoring the deleteriousness of single nucleotide variants as well as insertion or deletion variants in the human genome. CADD can assess the pathogenicity of multiple variants at a time. Input for CADD is chromosome number, SNP position, and the output is predicted as the CADD score. The CADD score is inversely proportional to the pathogenicity: CADD score 10 implies 10% deleterious and CADD score 20 means 1% deleterious and so on.

Alternative Splice Site predictor (ASSP) tool (http://wangcomputing.com/assp/index.html) is an online free-access tool for predicting the splice sites near the variant of interest. The input for this tool is the FASTA sequence flanking the SNP from dbSNP–NCBI and the output will give the predicted splice sites present in the sequence and the position of the bases in the sequence.

PCOS knowledge base (PCOSKB-http://pcoskb.bicnirrh.res.in/) is a manually curated knowledge base that gathers information about the PCOS-related genes, SNPs and pathways etc. with the supporting literature. The previously reported studies done on each SNP were retrieved by searching against their respective identifier.

2.6 Statistical analysis and power calculation

Sample size was calculated for a power of 85% at a 95% confidence interval with an estimated value of $P < 0.05$ based on the reference from Dakshinamoorthy et al., 2020 [10]. The parameters were taken as P1 = 44.87 which is the proportion of control volunteers having mutant alleles; while P2=58 which is the proportion of PCOS candidates with mutant allele and alpha = 0.05. Hardy-Weinberg equilibrium (HWE) was evaluated with a goodness of fit Chi-square test with one degree of freedom with a significance level held at P values < 0.05. Allelic frequencies and different genotyping models of polymorphisms were checked with the Chi-square Pearson or Fisher exact test. For statistical significance, the Mann-Whitney U test and Student's T test were used. A two-tailed type I error rate of 5% was considered statistically significant. All the analyses were carried out using the Statistical Package for Social Sciences (SPSS version 17) software.

3. RESULTS AND DISCUSSION

3.1 Clinical characteristics of the patients

In this study, women of age 18 to 34 who had clear manifestations of PCOS were chosen. Clinical parameters including follicle-stimulating hormone, luteinizing hormone, testosterone, estradiol, blood sugar, insulin fasting, and cholesterol were checked for both control and PCOS women. PCOS patients showed higher metabolic parameters like fasting blood sugar, insulin and cholesterol as compared to control women (p-value < 0.0001). Also, PCOS cohorts showed high levels of luteinizing hormone, testosterone, estradiol and thyroid-stimulating hormone when compared to the control group. Table 2 shows the elevated metabolic and hormonal parameters among PCOS women.

Table 2: Clinical parameters of the study participants.

S.No.	Variable	Normal Range	PCOS n=250	Control n=300
1	Age	18-34 years	25.58±4.88	26.248±3.86
2	Blood Sugar	> 70 to 150 mg/dL	125.17±64.87	93.08±21.31
3	HbA1C	< 5.7%	5.65±0.81	5.38±0.96
4	Insulin Fasting	< 25 mIU/L	16.81±10.61	8.03±4.18
5	Total Cholesterol	125 to 200 mg/dL	154.43±29.83	137.72±22.53
6	LH	5-20 IU/L	5.56±3.15	4.78±2.4
7	FSH	4.7-21.5 mIU/L	4.94±3.63	4.89±4.6

8	Estradiol	15-250 pg/ml	109.15±68.94	41.96±16.84
9	Testosterone	0.5-2.4 nmol/l	0.489± 0.19	0.309± 0.13
10	TSH	0.5-5 mIU/L	5.41±5.31	3.009±2.49

*HbA1C- Glycated Hemoglobin, LH- Luteinizing Hormone, FSH- Follicle-stimulating hormone, TSH- Thyroid-stimulating hormone.

3.2 Genotyping

Eight GWAS variants (rs11692782, rs2268361, rs2349415, rs13405728, rs1247860, rs13429458, rs2479106, rs10818854) were selected for the study and their details such as chromosome position, region, allele variation including the in-silico results are summarized in Table 3.

Table 3: Information of selected gene variants and their predicted pathogenicity.

S.No.	Gene	SNP ID	Position	Region	Allele	CADD Prediction	Alternative Splice Site Prediction
1	*FSHR*	rs11692782	chr2:49064754	Intron	T>A	Non-pathogenic	No splice site
2	*FSHR*	rs2268361	chr2:48974473	Intron	C>T	Non-pathogenic	Splice site occurs at SNP position
3	*FSHR*	rs2349415	chr2:49020693	Intron	T>C	Non-pathogenic	No splice site
4	*LHCGR*	rs13405728	chr2:48751020	Intron	A>G	Non-pathogenic	No splice site
5	*THADA*	rs12478601	chr2:43494369	Intron	C>T	Non-pathogenic	Splice site occurs next to the SNP
6	*THADA*	rs13429458	chr2:43411699	Intron	A>C	Tolerated	No splice site
7	*DENND1A*	rs2479106	chr9:123762933	Intron	A>G	Non-pathogenic	No splice site
8	*DENND1A*	rs10818854	chr9:123684499	Intron	G>A	Non-pathogenic	No splice site

3.3 *FSHR* variants and PCOS susceptibility

Follicle-stimulating hormone receptor is a G-protein coupled receptor that binds FSH (follicle-stimulating hormone) and aids in the conversion of testosterone to estradiol in the ovaries. Three intronic variants (rs11692782, rs2268361, rs2349415) from the *FSHR* gene were selected to find their association with PCOS in the Indian population. Two variants (rs11692782, rs2349415) showed an association with PCOS susceptibility with a statistical significance. rs11692782 showed significance for the genotypic model of wild type and homozygous mutant allele (AA vs TT) with an Odds ratio of 2.0952, 95% CI= 1.1921 to 3.6826 and p-value = 0.0102. The other SNP, rs2349415 showed significance for the genotypic model of wild type and heterozygous mutant allele (CC vs CT) with OR= 1.9375, 95% CI= 1.2279 to 3.0573 and p-value = 0.0045. rs2268361 variant did not show statistical significance for any of the genotypic models (p-value >0.05), thus not associating with PCOS in Indian women (Table 4). In-silico tools predicted these three SNPs to be non-pathogenic and rs2268361 variant to have a splice site near the SNP.

3.4 *LHCGR* variant and PCOS association

Luteinizing hormone/chorionic gonadotropin receptor (LHCGR) belongs to the G-protein coupled receptor 1 family. Luteinizing hormone (LH) induces ovulation; chorionic gonadotropin helps maintain conditions necessary for pregnancy. One of the GWAS identified, intronic SNPs (rs13405728) of the *LHCGR* gene was genotyped in Indian PCOS women. Statistical analysis for the genotypic model and allelic model showed a higher p-value and a lower Odds ratio, thus showing no association with PCOS. Allelic model A vs G showed p-value of 0.8956 and genotypic models showed a p-value of 0.2313 for AA vs GG and 0.1879 for AA vs AG (Table 4). In-silico analysis predicted rs13405728 to be non-pathogenic and no splice site near this SNP.

3.5 *DENND1A* variants and PCOS susceptibility

DENND1A encodes DENN (differentially expressed in normal and neoplastic cells) Domain-containing 1A protein. Overexpression of *DENND1A* can increase the possibility of a woman's susceptibility to PCOS by increasing steroid

hormone biosynthesis [12]. *DENND1A* SNPs were identified as potentially pathogenic for PCOS by three of the GWAS [13]. *DENND1A* gene variants rs2479106, and rs10818854 were included in this study to test their associated risk for PCOS. In the present study, rs2479106 polymorphism did not show any association with PCOS as per the statistical models with an Odds ratio less than 1 and a p-value greater than 0.05. In-silico analysis predicted a non-pathogenic nature of this SNP with no nearby splice site. For rs10818854, all the genotypic models showed highly significant association; GG vs AA (p-value= 0.0336, OR= 22.25, 95% CI= 1.2725 to 389.0715) and GG vs GA (p-value= <0.0001, OR= 5.0220, 95% CI= 1.2725 to 389.0715) (Table 4). Hence, this could be one of the SNPs implicated in early diagnosis of PCOS.

Table 4: Association of selected SNPs with PCOS in Indian population.

S.No.	Gene	SNP ID	Genotyping Model	p- Value, Odds ratio (95%CI*)
1	*FSHR*	rs11692782	A vs T 1 AA vs TT 2 AA vs AT 3	0.0628, 1.3768 (0.9830 to 1.9283) 0.0102, 2.0952 (1.1921 to 3.6826) 0.5067, 1.2065 (0.6932 to 2.0999)
		rs2268361	G vs A GG vs AA GG vs GA	0.3906, 0.8444 (0.5740 to 1.2422) 0.7709, 0.9180 (0.5161 to 1.6328) 0.3128, 1.4075 (0.7247 to 2.7337)
		rs2349415	C vs T CC vs TT CC vs CT	0.8279, 1.0541 (0.6558 to 1.6942) 0.1632, 1.4531 (0.8593 to 2.4572) 0.0045, 1.9375 (1.2279 to 3.0573)
2	*LHCGR*	rs13405728	A vs G AA vs GG AA vs AG	0.8956, 0.9661 (0.5776 to 1.6161) 0.2313, 3.8450 (0.4238 to 34.8806) 0.1879, 4.5161 (0.4788 to 42.5929)
3	*THADA*	rs12478601	T vs C TT vs CC TT vs CT	<0.0001, 5.8 (4.2067 to 7.9907) <0.0001, 54.47 (18.3 to 156.7429) 0.2101, 1.38 (0.8340 to 2.2828)
		rs13429458	A vs C AA vs CC AA vs AC	0.0001, 2.3672 (1.5222- 3.6815) 0.0008, 8.2907 (2.4031- 28.6025) 0.4822, 1.2306 (0.6899- 2.1949)
4	*DENND1A*	rs2479106	G vs A GG vs AA GG vs GA	0.5468, 0.8987 (0.6348 to 1.2722) 0.6656, 0.8980 (0.5512 to 1.4629) 0.2070, 0.7463 (0.4737 to 1.1758)
		rs10818854	G vs A GG vs AA GG vs GA	<0.0001, 3.6388 (2.2391 to 5.9133) 0.0336, 22.25 (1.2725 to 389.0715) <0.0001, 5.0220 (1.28 to 389.0715)

*95% CI= 95% confidence interval, 1-allelic model, 2- homozygous genotypic model, 3- heterozygous genotypic model

3.6 *THADA* variants and PCOS association

Thyroid adenoma-associated protein (THADA) is expressed in the pancreas, thyroid, testes, adrenal gland, small intestine, and stomach [8]. *THADA* polymorphisms are associated with insulin resistance pathophysiology in PCOS onset. In our study, variant rs12478601 showed strong association with PCOS in the homozygous mutant allele model (TT vs CC) with higher significance (p-value of <0.0001, OR of 54.47 and 95% CI of 18.93 to 156.75). The wild type vs homozygous mutant allele (AA vs CC) genotypic model of rs13429458 showed a higher odds ratio of 8.2907 and p-value of 0.0008 exhibiting a significant association with PCOS (Table 4). In-silico analysis of the chosen *THADA* variants revealed a splice site near the SNP, rs12478601, and thus predicted that this variant is likely to affect alternative splicing.

We compared our results with other similar association studies carried out for the 10 GWAS-identified SNPs in Asian and Caucasian populations and tabulated the same (Table 5). It was interesting to find that our results are in comparison with that from Chinese population populations [7]. Of all the SNPs, only rs2349415 of *FSHR* revealed dissimilarity in association with PCOS between the Indian and Chinese populations.

Table 5: Association of selected SNPs with PCOS different ethnicities.

S. No.	Gene and SNP ID	Association in Indian Population (This study)	Association in Indian Population (Other Study)	Association in Chinese Population	Association in Korean Population	Association in Caucasian Population	Association in Arab Population	Association in European Population

1	*FSHR* rs11692782	No	-	-	-	-	Yes [14]	-
2	*FSHR* rs2268361	No	-	Yes [15]	Yes [16]	-	No [17]	Yes [18]
3	*FSHR* rs2349415	Yes	-	No [15]	-	-	-	-
4	*LHCGR* rs13405728	Yes	Yes [9]	Yes [19]	-	No [20]	-	No [21]
5	*DENND1A* rs2479106	Yes	No [9] Yes [22]	Yes [23]	No [16]	No [24]	No [25]	No[26]
6	*DENND1A* rs10818854	Yes	-	Yes [23]	-	-	Yes [25]	Yes [26]
7	*THADA* rs12478601	Yes	Yes [22]	Yes [27]	-	-		No[26]
8	*THADA* rs13429458	Yes	No [8] Yes [9]	Yes [27]	No [16]	No [24]	-	No[26]
9	*INSR* rs1799817	-	Yes [28] No [22]	-	No [29]	-	Yes [30]	-
10	*INSR* rs2059807	-	-	Yes [27]	No [16]	-	-	-

In India, three of the eight SNPs including rs2479106 of *DENND1A*, rs13405728 of *LHCGR*, and rs13429458 of *THADA* were genotyped previously [8,9]. Vishnubotla et al., genotyped 15 variants identified from GWAS and case-control studies in individuals with PCOS of Indian ethnicity. Among these, rs13405728 of *LHCGR*, and rs13429458 of *THADA* were found to be associated while rs2479106 of *DENND1A* was not. In another similar study, Dadachanji et al., genotyped rs13429458 of *THADA* and found that genotypic and allele frequencies were not different between controls and PCOS. However, in our study, we found that the genotypic models for rs13429458 mutant allele C showed strong association with PCOS in Indian women, while rs2479106 of *DENND1A*, rs13405728 of *LHCGR* showed no association. Such differences could be due to the variation in the sample size and the choice of the location where samples were collected.

4. CONCLUSION

Of the eight GWAS variants genotyped in this study, five variants, rs11692782, rs2349415 from *FSHR*, rs10818854 from *DENND1A* and rs12478601, rs13429458 from *THADA* show strong association to PCOS in Indian women. The results are comparable to that of the Chinese population as can be observed from Table 5. However, only rs10818854 of *DENND1A* shows an association to PCOS in many ethnicities including Indian, Chinese, Arab, and European populations (Table 5). These variants should be assessed in a larger population across different ethnicities to exploit them as potential early diagnosable markers for PCOS.

Authors Contribution

Kuppachi Himaja conducted the experiments for *FSHR*, *LHCGR* SNP analysis, project execution, data curation and manuscript drafting, Janani Dakshina Moorthy conducted the methodology design, statistical analysis, and manuscript review, Pavithra Selvanathan conducted the *FSHR's* one SNP analysis, Poornima Gopalakrishnan and Bhuvaneswari Karthik Ganesh conducted the *DENND1A* SNP analysis, Preethi Sundar conducted the *THADA* SNP analysis, Usha Balasundaram responsible for conceptualization, project administration, supervision, writing - review and editing.

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Ethical Approval

This project has been evaluated and approved by the Institutional Ethics and Biosafety Committee.

Consent to Participate

All the participants involved in the study provided their consent for providing blood samples.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGMENT

The authors sincerely thank the members of the PCOS Research Foundation, India for their support in collecting blood samples from PCOS patients.

REFERENCES

1. Goodarzi MO, Dumesic DA, Chazenbalk G, Azziz R. Polycystic ovary syndrome: Etiology, pathogenesis and diagnosis. Nat Rev Endocrinol. 2011;7(4):219–31.
2. Ganie M, Vasudevan V, Wani I, Baba M, Arif T, Rashid A. Epidemiology, pathogenesis, genetics & management of polycystic ovary syndrome in India. Indian J Med Res 2019;150:333–44.
3. Chen ZJ, Zhao H, He L, Shi Y, Qin Y, Shi Y, 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):55–9.
4. Shi Y, Zhao H, Shi Y, Cao Y, Yang D, Li Z, et al. Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. Nat Genet. 2012;44(9):1020–5.
5. Goodarzi MO, Jones MR, Li X, Chua AK, Garcia OA, Chen YDI, et al. Replication of association of DENND1A and THADA variants with polycystic ovary syndrome in European cohorts. J Med Genet. 2012;49(2):90–5.
6. Lee H, Oh JY, Sung YA, Chung H, Kim HL, Kim GS, et al. Genome-wide association study identified new susceptibility loci for polycystic ovary syndrome. Hum Reprod 2015;30(3):723–31.
7. Wan P, Meng L, Huang C, Dai B, Jin Y, Chai L, et al. Replication study and meta-analysis of selected genetic variants and polycystic ovary syndrome susceptibility in Asian population. J Assist Reprod Genet. 2021;38(10):2781-9.
8. Dadachanji R, Sawant D, Patil A, Mukherjee S. Replication study of THADA rs13429458 variant with PCOS susceptibility and its related traits in Indian women. Gynecol Endocrinol 2021;37(8):716–20.
9. Vishnubotla DS, Shek AP, Madireddi S. Pooled genetic analysis identifies variants that confer enhanced susceptibility to PCOS in Indian ethnicity. Gene 2020;752:144760.
10. Dakshinamoorthy J, Jain PR, Ramamoorthy T, Ayyappan R, Balasundaram U. Association of GWAS identified INSR variants (rs2059807 & rs1799817) with polycystic ovarian syndrome in Indian women. Int J Biol Macromol 2020;144:663–70.
11. Miller SA, Gallie DR, Sleat DE, Watts JW, Turner PC, Wilson TM a. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 1988;16(3):883–93.
12. McAllister JM, Modi B, Miller BA, Biegler J, Bruggeman R, Legro RS, et al. Overexpression of a DENND1A isoform produces a polycystic ovary syndrome theca phenotype. Proc Natl Acad Sci U S A. 2014;111:1519–27.
13. Janani DM, Usha B. In silico analysis of functional non-synonymous and intronic variants found in a polycystic ovarian syndrome (PCOS) candidate gene : DENND1A. Indian J Biochem Biophys 2020;57:584–93.
14. Almawi WY, Hubail B, Arekat DZ, Al-Farsi SM, Al-Kindi SK, Arekat MR, et al. Leutinizing hormone/choriogonadotropin receptor and follicle stimulating hormone receptor gene variants in polycystic ovary syndrome. J Assist Reprod Genet. 2015;32(4):607–14.
15. Du T, Duan Y, Li K, Zhao X, Ni R, Li Y, et al. Statistical genomic approach identifies association between FSHR polymorphisms and polycystic ovary morphology in women with polycystic ovary syndrome. BioMed Res Int. 2015; 483726.
16. Hong S hyeon, Hong YS, Jeong K, Chung H, Lee H, Sung YA. Relationship between the characteristic traits of polycystic ovary syndrome and susceptibility genes. Sci Rep 2020;10(1):1–8.
17. Bakhashab S, Ahmed N. Genotype based risk predictors for Polycystic Ovary Syndrome in Western Saudi Arabia. Bioinformation 2019;15:812-9.
18. Saxena R, Georgopoulos NA, Braaten TJ, Bjonnes AC, Koika V, Panidis D, et al. Han Chinese polycystic ovary syndrome risk variants in women of European ancestry: Relationship to FSH levels and glucose tolerance. Hum Reprod 2015;30(6):1454–9.
19. Ha L, Shi Y, Zhao J, Li T, Chen ZJ. Association study between polycystic ovarian syndrome and the susceptibility genes polymorphisms in Hui Chinese women. PLoS One 2015;10(5):1–10.
20. Lerchbaum E, Trummer O, Giuliani A, Gruber HJ, Pieber T, Obermayer-Pietsch B. Susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21, and 9q33.3 in a cohort of Caucasian women. Horm Metab Res 2011;43(11):743–7.
21. Mutharasan P, Galdones E, Bernabe BP, et al. Evidence for Chromosome 2p16.3 Polycystic Ovary Syndrome Susceptibility Locus in Affected Women of European Ancestry. J Clin Endocrinol Metab 2012; 98:185-90.
22. Irgam K, Reddy BM, Annamaneni S, Rayabarapu P. The genetic susceptibility profile 13 of the South Indian women with polycystic ovary syndrome and the universality of the lack of association of type 2 diabetes genes. Gene 2019;701:113–20.
23. Bao S, Cai JH, Yang SY, Ren Y, Feng T, Jin T, et al. Association of DENND1A gene polymorphisms with polycystic ovary syndrome: A meta-analysis. J Clin Res Pediatr Endocrinol 2016;8:135–43.
24. Eriksen MB, Brusgaard K, Andersen M, Tan Q, Altinok ML, Gaster M, et al. Association of polycystic ovary syndrome susceptibility single nucleotide polymorphism rs2479106 and PCOS in Caucasian patients with PCOS or hirsutism as referral diagnosis. Eur J Obstet Gynecol Reprod Biol 2012;163:39–42.
25. Dallel M, Sarray S, Douma Z, Hachani F, Al-Ansari AK, Letaifa DB, et al. Differential association of DENND1A genetic variants with polycystic ovary syndrome in Tunisian but not Bahraini Arab women. Gene 2018;647:79–84.
26. Welt CK, Styrkarsdottir U, Ehrmann DA, Thorleifsson G, Arason G, Gudmundsson JA, et al. Variants in DENND1A are associated with polycystic ovary syndrome in women of European ancestry. J Clin Endocrinol Metab 2012;97:1342–7.
27. Cui L, Li G, Zhong W, Bian Y, Su S, Sheng Y, et al. Polycystic ovary syndrome susceptibility single nucleotide polymorphisms in women with a single PCOS clinical feature. Hum Reprod 2015;30:732–6.

28. Mukherjee S, Shaikh N, Khavale S, Shinde G, Meherji P, Shah N, et al. Genetic variation in exon 17 of INSR is associated with insulin resistance and 14 hyperandrogenemia among lean Indian women with polycystic ovary syndrome. Eur J Endocrinol 2009;160:855–62.
29. Lee EJ, Yoo KJ, Kim SJ, Lee SH, Cha KY, Baek KH. Single nucleotide polymorphism in exon 17 of the insulin receptor gene is not associated with polycystic ovary syndrome in a Korean population. Fertil Steril 2006;86:380–4.
30. Daghestani MH. Rs1799817 in INSR associates with susceptibility to polycystic ovary syndrome. J Med Biochem 2019;39.