Survey based on classification of PCOS using gene expression

B. Poorani
Research Scholar, Department of Computer Science, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Thandalam, Chennai

A. Kalaivani
Professor, Department of Information Technology, School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University Thandalam, Chennai

Abstract---Polycystic ovary syndrome (PCOS) is a complicated endocrine ailment affecting 5–10 % of girls of reproductive age. It is generally has complications such as irregular cycles, hirsutism and polycystic ovaries, collectively with a considerable incidence of insulin resistance. PCOS is taken into consideration a multifactorial ailment with numerous genetic, endocrine and environmental abnormalities. Eventhough, women who are suffered with PCOS are also have chances for cause of metabolic and cardiovascular illnesses and their associated diseases or disorders which are compared to the overall population. PCOS classification done on both phenotype and genotype, based on phenotype it is segmented into four phenotypes as A, B, C, D. This phenotype shows the association between hyperandrogenisms, structure or shape polycystic ovary and immature ovary formation. Phenotypes can be visualized through ultrasound images. Based on genotype, PCOS are categorized into three based on the hormonal factors. By analyze PCOS genetically we early diagnose the syndrome and take necessary remedial actions and also prevent from inheritance to next generations. In this paper, surveys the different genes involved in causes of PCOS and hormone factors associated with the PCOS.

Keywords---polycystic ovary syndrome, phenotype, genotype, insulin resistance, antrogen based, inflammatory PCOS.
**Introduction**

DNA is basic functional units for storing heritance information as Gene Expression. Gene Expression analysis is the process of analyzing or calculating the multiple levels of DNA functional units on gene expression until to predicate particular information which is also termed as gene responsibility. Using a gene expression analysis we can predict the particular dysfunctions on specific organs, hormones, etc. Polycystic ovary syndrome (PCOS) is the heterogenous and complex diseases raised in one of fifth of the women in India at the reproductive ages. PCOS is intersection on both phenotypic and genotypic changes. Genetically PCOS is calculated by evaluating the hormonal changes such that thyroid, androgens, insulin, etc. Genes is analyzed by gene expression data taken from the gene expression omnibus datasets (GEO). PCOS causes various impacts on women’s health related to infertility, type 2 diabetes, Acne, irregular periods, cardiac complications.

![Figure 1](image.png)

**Fig. 1. Past Year Statistics various disease contributed with PCOS [20]**

Figure 1 estimates that, over the year statistics on the various disease contribution with PCOS [20]. About 70 percent of women suffer with PCOS are affected with hirsutism and 60-80% women have androgen levels and merely 80% metabolic disorders [1]. Long term health implications are due to insulin resistance and leads to diseases such as high blood pressure, pre-diabetes, sleep apnea, greater impact on the cardiac complications. Hyper-insulinemia causes symptoms such as weight loss and associated with regulated menstrual cycle. Figure 2 shows the general architecture to diagnose PCOS using gene expression.
Fig. 2. General Architecture for Diagnose PCOS using Gene Expression

The National Center for Biotechnology Information manages the Gene Expression Omnibus database [GEO] for gene expression profiling and RNA methylation profiling. Extract the gene expression from the gene profiling for analyze the PCOS gene from that sequence [6-10]. Various analysis technique to diagnose the PCOS gene from the gene expression /gene sequence such that, Genome Wide Associations Study [GWAS], Next-generation Sequence [NES], Polymerase Chain Reaction [PCR], Machine learning Techniques, etc.[2-10]. using above mention techniques we can identified the PCOS gene. The PCOS gene is classified based on phenotypic and genotypic alterations that are linked. Phenotype and genotype changes is a statistical link between genotype and phenotype predicts a physical characteristic in a person or abnormalities in a patient with a specific mutation or a combination of mutations[18]. Multiple of phenotype changes occur due to PCOS and it is heritable also, such type are listed below [8, 11, 15]. The classifications of PCOS are shown in Figure 3.

- Phenotype A: called classic PCOS, which cause Hyperandrogenism, oligoovalution, and polycystic ovary morphology.
- Phenotype B: also called classic PCOS which causes the Hyperandrogenism and polycystic ovary morphology.
- Phenotype C: called ovulation PCOS which causes the Hyperandrogenism and oligoovalution.
- Phenotype D: called Non-Hyperandrogenism which causes polycystic ovary morphology and oligoovalution.

Figure 3 shows that, PCOS classification based on the Phenotype-genotype correlations.
These types are identified by the clinical evolution but we can't predict the future impact of the particular gene. To evaluate the PCOS, standard analyses methods are defined in table 1 and the evaluation criteria for PCOS [15-18].

### Table 1
Evaluation Criteria for PCOS [15-18]

<table>
<thead>
<tr>
<th>Evaluation Criteria</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIC</td>
<td>Both hyperandrogenism and irregulation in periods</td>
</tr>
<tr>
<td>Rotterdam</td>
<td>Hyperandrogenism, irregulation, and different sized ovaries</td>
</tr>
<tr>
<td>Androgen Excess Society</td>
<td>Hyperandrogenism + ovarian dysfunction indicated by irregulation and/or PCOS</td>
</tr>
</tbody>
</table>

All of the PCOS diagnostic criteria were based on expert opinion, which is the weakest degree of evidence. There was no formal consensus procedure for any of these criteria. The National Institutes of Health (NIH) Consensus Development Program (http://consensus.nih.gov/) is a widely accepted consensus process in the United States, administered by the Office of Medical Applications of Research, which has recently become part of the Office of Disease Prevention (http://consensus.nih.gov/). These meetings follow a "court" paradigm, with evidence being presented to a panel that acts as a jury (http://consensus.nih.gov/FAQs.htm#whatisetheCDP). The panel is made up of people who are specialists in their disciplines but aren't really interested in the issue. As a result, these Consensus Development Conferences allow an unbiased panel to analyze the challenges in the sector independently. In December 2012, an NIH symposium on PCOS will be held utilizing this court concept [19-20]. Various PCOS genes are identified in prior research using different techniques and also classify them based the phenotype correlated with the genotype of PCOS. We review the prior research works for identification of the PCOS gene and hereditability of PCOS.

### Related Works

PCOS classified based on the statistical relation between phenotype and genotype. PCOS in women cause various complications on three phases they are adult phase (18-20yrs), reproductive phase (during pregnancy), menopause (<50) [18-
During adult phase, major complications are irregular periods, pimples on face and unwanted hair growth. This is due to decrease in insulin secretions and imbalance secretions of androgens [male hormone] [8-15]. In reproductive phase, infertility of ovary is the major cause occurred due to oligovalution (mismatch menstrual cycle) [9-12]. In menopause phase, overflow of bleeding without correct intervals is another major complications. This is due to insulin resistance, Hyperandrogenism and oligovalution. This can be predicting early we can avoid such complications. Many researches are done on the PCOS and PCOS subtypes genes which are highly impact on the long-term complications are reviewed. Figure 4 shows that PCOS varies on different stages of women [6].

![PCOS: a life long disease](image)

**Fig. 4.** PCOS varies on different stages of women [6]

**Insulin resistance and Hyperinsulinemia PCOS**

Past proof recommends that insulin opposition and compensatory hyperinsulinemia are key highlights of PCOS [14]. In many investigations, its pervasiveness was somewhere in the range of 44% and 75% [15, 16, 17], a lot higher than the 10-25% saw in youthful sound people [18]. PCOS patients with and without weight have higher paces of insulin obstruction than sound controls [19], in spite of the fact that insulin opposition is more serious in large subjects [20]. According to our data, there is also a difference in the diagnostic value of different insulin resistance diagnostic methods between lean and obese PCOS subjects [20]. Imbalance secretion of insulin in women leads to PCOS. Insulin resistance implicates lots of long-term diseases such as hypertension, glucose intolerance, obesity, etc. Insulin secretions are directly interacting with the Luteinizing hormone (LH) and Follicular Stimulating hormone (FSH). LH and FSH are present in Theca cell and Granulosa cell respectively, which are responsible for growth of ovary in the uterus and helps for ovulation process during menstrual cycle. Figure 2 role insulin in PCOS [20].
Increased insulin resistance plays important drivers for PCOS. Hyperinsulinemia is nothing but increase in resistance of insulin which leads to affect the hyperandrogenemia, imbalance LH/FSH secretions. Genes for insulin resistance PCOS CAPN10 is protein-coded gene belongs to calcium-dependent family. This gene associated with diabetes mellitus 2 and PCOS. Patients who suffer with PCOS has highly associated with type II diabetes (CAPN10) has multiple single polymorphism (SNP) such that, UCSNP-63 and UCSNP-19 have been associated with PCOS[12]. Gly972Arg shows more insulin resistance and it more obese and it also shows higher for PCOS. Cytochrome family p450, Insulin gene, AR, FTO, FSHR genes are highly associated with PCOS[12,11]. The phosphoinositide3kinase/protein kinase B signaling pathway is activated in ovarian Schwann cells. PCOS patients participate in insulin activation during this process [13]. Excess insulin is associated with high androgen production [13, 15]. The insulin gene is a gene sandwiched between insulin growth factor 11p15.5 (IGFII) and tyrosine hydroxylase. Location [11]. The variable number of repeats (VNTR) occupies 50, untranslated regions [10-8]. VNTR polymorphisms participate in insulin gene (INS) regulation. And IGFII transcription rates which are associated with PCOS. In past research, the insulin SNP patterns are identified by various comparisons of different PCOS datasets. Table 1 SNP of Insulin gene responsible for PCOS [18].

Table 2
SNP of Insulin gene responsible for PCOS [18]

<table>
<thead>
<tr>
<th>S.No</th>
<th>Gene</th>
<th>Polymorphism type</th>
<th>Gene Marker</th>
<th>External Function</th>
<th>Case study on various countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C</td>
<td>UCSNP-43</td>
<td>UCSNP-43,</td>
<td>Calcium</td>
<td>Chilean women</td>
</tr>
</tbody>
</table>
Hyperandrogenism

Hyperandrogenism is an another important driver for PCOS complications. Hyperandrogenism divers male hormone in women. Androgen hormone are secreted from the gland called adrenal gland. Adrenal androgens are released in an unbound condition by the adrenal cortex. Binding steroids have no biological effect. Albumin is mostly bound by androstenedione, DHEA, and DHEAS [2]. Approximately 90% of adrenal androgens are linked to albumin, with just 3% bound to sex hormone binding globulin (SHBG). Sex hormones, such as testosterone, are androgens. Androgens aid physical maturation and assist humans reach puberty, Acne, facial hair, and other problems can occur in women with high androgen levels [5]. These hormones affect the LH and FSH secretions in the theca cell in zona reticular Figure 6 explains about the directly relationship between hyperinsulineamia and Hyperandrogenism[12-15].

Fig. 6. Interconnection Between Insulin And Hyperandrogenism on PCOS[12-15]

Various genes have been linked to PCOS hyperandrogenism during the assessment and identification phases. The gene CYP19 is found on the 15q21.2
chromosome and is responsible for aromatase p450 activities, which are essential for oestrogen synthesis. Aromatase activity is decreased in lean and obese PCOS individuals [9]. Pregnenolone is converted to 17-hydroxypregnenolone and progesterone is converted to 17-hydroxyprogesterone by the enzyme P450c17, which is encoded by CYP17. Overexpression of CYP17 in theca cells, as well as a polymorphism in the promoter area, has been linked to PCOS [11,8,9]. The CYP21 gene codes for an enzyme that converts 17-hydroxyprogesterone to 11-deoxycortisol, which is a stage in the steroid hormone production process. Ineffective anabolism of steroids is caused by the enzyme’s inactivity, which is also responsible for PCOS. CYP11a encodes the enzyme that is involved in the rate-limiting phase of cholesterol conversion to progesterone [8-11]. CYP11a polymorphism and variation have been observed in CYP11a, and several investigations have established a link between CYP11a and PCOS [12-15].

**Ovarian and Adrenal Steroid genesis**

Various genes were discovered to be associated with PCOS Hyperandrogenism during the assessment and identification stages.[8]. CYP19 is the gene that controls aromatase p450 activities, which are necessary for. It is found on the 15q21.2 chromosome and is involved in estrogen formation. PCOS in the fat and lean Aromatase activity is decreased in patients [8-6]. The enzyme P450c17, which is encoded by CYP17, catalyzes the conversion of Pregnenolone to testosterone. Progesterone and 17-hydroxypregnenolone are converted to 17-hydroxyprogesterone. Overexpression of CYP17 in theca cells was discovered, as well as polymorphism in the promoter region. PCOS is related with [7] CYP21 gene is involved in the encoding of an enzyme that converts stage in the manufacture of steroid hormones is the conversion of 17-hydroxyprogesterone to 11-deoxycortisol. Inactivity of the enzyme results in poor steroid anabolism, which is also responsible for PCOS [11-15].

**Inflammatory PCOS**

Inflammatory PCOS is composed of behavior (stress) of person phenotypic factors such the BMI, age and obesity of women. This type incorporates the hormonal issues address under BMI, age and obesity factors. Inflammatory stage is highly complication stage.[1][6]. Negative mutual causes of obesity and insulin resistance leads to increased inflammation and that probability of initiating influence PCOS pathogenesis [8]. Inflammation of PCOS pathway leads to cause infertility of women and menopause problems [9]. Inflammatory PCOS biomarkers Glycoproteins are important one in the inflammation process. Amid aggravation, a few proteins experience basic changes and glycan chains adjustments. Such glycoprotein profiles can act as irritation markers. Glyc A and Glyc B biomarkers of irritation, IR, affront discharge, and corpulence have moreover been as of connected to PCOS[6-7]. C-Receptive Protein Progressed Glycation End-Products and Oxidative Stretch (AGE), Pro-Inflammatory Cytokines and Chemokine’s are another biomarker for inflammation PCOS. [11]

Various implications of inflammatory PCOS Obesity, Insulin Resistance and Inflammatory Process are led by pro-inflammation on PCOS. Obesity, insulin resistance are correlates for high impact on the inflammation PCOS. Obesity and
insulin are pre-dominant on ovary dysfunction. Hyperglycemia can play a part in the aggravation preparation in PCOS women. Glucose may be a primary redox substrate of menopause.[10-15]. Endothelial Inflammation and Risk of Cardiovascular Disease – endothelial cell dysfunction are independent of age and weight factors but it shows wide association with vascular disorders. Vascular disorders are aided due to arterial stiffness. Hyperandrogenisms and hyperglycemia and hypertension are increasing the coronary heart diseases and directly connected to vascular disorders. Table 3 represents the various gene involved in causes of PCOS[15].

Table 3
Genes Involved For the Cause of PCOS Based On Various Factors [15]

<table>
<thead>
<tr>
<th>S No</th>
<th>Various factors of PCOS</th>
<th>Gene Involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Adrenal based gene</td>
<td>i. CYP19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii. CYP17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iii. CYP21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iv. CYP11a</td>
</tr>
<tr>
<td>2</td>
<td>Insulin secretion and insulin action</td>
<td>i. CAPN10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii. IRS-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iii. IRS-2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iv. INS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>v. INS</td>
</tr>
<tr>
<td>3</td>
<td>Ovarian and Adrenal Steroidogenesis</td>
<td>i. LH 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii. AMH</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iii. FSHR</td>
</tr>
<tr>
<td>4</td>
<td>Inflammatory</td>
<td>i. FTO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ii. PCO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iii. SRD5A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>iv. SRD5B</td>
</tr>
<tr>
<td></td>
<td></td>
<td>v. NCOR1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vi. PPARG1</td>
</tr>
</tbody>
</table>

Inference from survey

Gene expressions are fetched from the Gene Expression Omnibus database. Genome Wide Association a study (GWAS) is one approach is utilized to analyze the hereditary varieties that happened for specific illnesses. Common loci are distinguished to assess hereditary heritability. Minor allele frequencies were planned to get to the common allelic variations [11]. Next-Generation Sequencing (NGS) approaches ought to recognize uncommon variations which are contributed to complex malady pathogenesis. Affiliation consider on the regenerative, quantitative, and metabolic characteristics, Linkage disequilibrium (LD) is distinguished between uncommon variations and chance alleles fetch from GWAS [5]. Another way to analyze the gene expression called Polymerized chain reaction (PCR). Single nuclei variant (SNP), chromosome loci of gene is analyzed by polymer of the gene sequences [2-5].

More research needed to analyzed the pathophysiology of PCOS and determine the preventative risk factors and treatment should be predicate for particular
classifications [1]. Gene Pathway prediction should be analyzed for particular type PCOS [2]. Genetic architecture should create for analyses the various associated gene in PCOS.[10]. Genotype and phenotype change are identified and classify according to environment.[12]. Various therapeutic methods are diagnosed according to the genetic evolution. Heredity of PCOS is predicated to avoid the long-term diseases [15].

**Conclusion**

In the existing researches confirmed that, various gene responsible for different types of PCOS. PCOS-related hormone imbalance, insulin dysfunction, ovary dysfunction, infertility gene are identified .since in advancement the classification of PCOS made easy to predicate the PCOS type and may provide the recommended treatments. By classify the PCOS, the heritable genes are found early to avoid negative impacts on next generations and also avoid PCOS as the long-term diseases to further generations.

**References**


12. Linlin Cui1,2,3,4,†, Han Zhao1,2,3,4,†, Bo Zhang5,†, Zhongyu Qu6, Jiayin Liu7, Xiaoyan Liang8, Xiaoming Zhao9, Junli Zhao10, Yingpu Sun11, Peng Wang1,2,3,4, Tao Li1,2,3,4, Yuhua Shi1,2,3,4, And Zi-Jiang Chen1,2,3,4,:Genotype –Phenotype Correlations Of Pcos Susceptibility Snps Identified By Gwas In A Large Cohort Of Han Chinese Women. Human Reproduction, Vol.28, No.2 Pp. 538–544, 2013 Advanced Access Publication On December 2, 2012 Doi:10.1093/Humrep/Des424.


