20

The Genetic Basis of the Polycystic Ovary Syndrome

Mark O. Goodarzi

SUMMARY

Polycystic ovary syndrome (PCOS) is a common, complex genetic disorder. Its inherited basis was established by studies demonstrating an increased prevalence of PCOS, hyperandrogenemia, insulin resistance, and disordered insulin secretion in relatives of women with PCOS. To date, efforts to elucidate the genetic basis of PCOS have focused on candidate genes chosen from logical pathways, including steroid synthesis and action, insulin sensitivity and secretion, obesity and fuel regulation, gonadotropin production and action, and, most recently, cardiovascular risk modifiers. Although several positive results have been reported, no gene or genes are universally accepted as important in PCOS pathogenesis, largely because of lack of replication of positive results. This has resulted, in part, from various factors, most importantly inadequate coverage of genes by the analysis of only one or two variants and of small study cohorts in many studies. In the future, optimal application of the candidate gene approach using haplotype-based analyses, intermediate phenotypes, and internal replication of positive results will enhance gene discovery in PCOS, as will the application of pharmacogenetics and whole-genome analysis.

Key Words: Polycystic ovary syndrome; candidate gene; genetic association; genetic linkage; single nucleotide polymorphism; haplotype; pharmacogenetics.

1. INTRODUCTION

Polycystic ovary syndrome (PCOS) is considered a common, complex genetic disorder, as are conditions such as schizophrenia, asthma, and type 2 diabetes *(1)*. Such common diseases, including PCOS, appear to have a complex, multifactorial etiology, in which a variety of predisposing genes, not just one gene, interact with environmental factors to produce disease. Studies in families have demonstrated the heritable nature of PCOS itself as well as the component phenotypes of PCOS. Subsequently, a large number of population studies have attempted to discover genes that influence PCOS using the candidate gene approach.

2. BACKGROUND

2.1. Heritability of PCOS and Hyperandrogenemia

Family studies demonstrate that PCOS is significantly more prevalent among family members than in the general population (2). Among first-degree female relatives, of 93 patients with PCOS, 35% of premenopausal mothers and 40% of sisters on no hormonal therapy were also affected with the disorder (3). These affection rates are significantly higher than the 6–7% observed in the general population (4). In another study, 115 sisters of 80 women with PCOS were evaluated. Of these, 22% met the criteria for PCOS (5). An additional 24% of sisters had hyperandrogenemia with normal

From: *Contemporary Endocrinology: Androgen Excess Disorders in Women: Polycystic Ovary Syndrome and Other Disorders, Second Edition*
Edited by: R. Azziz et al. © Humana Press Inc., Totowa, NJ

menses. Total and free testosterone levels were similar between the sisters with hyperandrogenemia only and the sisters and probands with PCOS. A bimodal distribution of testosterone levels in the sisters of women with PCOS was observed, suggesting a major genetic component to hyperandrogenemia *(5)*. Brothers of women with PCOS also display abnormal androgens: a study of such brothers found them to have elevated levels of dehydroepiandrosterone sulfate (DHEAS) *(6)*.

Physiological evidence that hyperandrogenemia in PCOS is genetically controlled is supported by the observation that ovarian theca cells removed from women with PCOS and propagated in culture display persistently elevated testosterone secretion compared to cells from unaffected women *(7)*. Because removal from the body and multiplication in culture removes the influence of an abnormal hormonal milieu from these cells, it is likely that persistent differences from normal represent intrinsic (genetic) defects.

2.2. *Heritability of Insulin-Related Traits in PCOS*

Not only is PCOS itself a heritable condition, but within PCOS insulin resistance and insulin secretion also appear to be under significant genetic control. Among sisters of women with PCOS, those who had PCOS or hyperandrogenemia with regular menses had lower insulin sensitivity than unaffected sisters, assessed by fasting insulin and glucose measurements *(8)*. Likewise, in families of Australian patients with PCOS, hyperinsulinemia was found to occur in 69% of all family members, suggesting that this trait was inherited *(9)*. One hundred and two relatives of 52 Turkish women with PCOS underwent assessment of insulin resistance in the fasting and post-glucose-challenge states; compared to population controls, mothers, sisters, and brothers of PCOS subjects had greater insulin resistance *(10)*. Brothers of Indian women with PCOS had insulin resistance and endothelial dysfunction *(11)*. In studies of families of women with PCOS, insulin secretion levels, quantified directly by the frequently sampled intravenous glucose tolerance test, displayed significant heritability, suggesting a genetic component to β-cell dysfunction in PCOS *(12)*. The most abnormal insulin secretion was observed in women with PCOS and a history of type 2 diabetes in a first-degree relative *(13)*.

Abnormal responses to insulin in cells that have been removed from women with PCOS suggest the presence of intrinsic cellular defects because these cells have been removed from the hormonal milieu present in these women. Lower insulin receptor substrate (IRS)-1-associated phosphatidylinositol 3-kinase (PI3K) activity and higher IRS-2 content were observed in the myocytes of PCOS patients compared to controls, in the face of similar amounts of IRS-1 and the p85 subunit of PI3K *(14)*. In PCOS adipocytes, the maximum glucose uptake stimulated by insulin was found by some investigators *(15)* to be lower compared to controls. Although others *(16)* found the maximum response to be normal, these latter investigators found that the sensitivity of the response to insulin was deficient in PCOS. The amount of glucose transporter-4 in adipocytes was lower in PCOS (on either a membrane protein or cell surface basis) than in controls *(15)*. Such defects are likely caused by abnormal expression or function of genes encoding products of the insulin-signaling pathway.

2.3. *The Principle of Population Genetics*

Given that investigators often start with no knowledge of genetic variants that lead to disease, they must take advantage of chromosomal markers. Markers, such as microsatellites and single-nucleotide polymorphisms (SNPs), are polymorphic variants interspersed throughout the genome. Microsatellites are tandem repeats of short nucleotide sequences, occurring with variable numbers of the repeated unit; SNPs are changes at a single genomic base pair, comprising two possible alleles. These markers are used as tags to track disease-causing variants or mutations. The underlying principle is that markers that are close to disease-causing variants tend to be inherited on the same chromosomes. Linkage refers to the situation wherein markers in a region of the genome are inherited in families in a nonrandom fashion in relation to a particular phenotype. Association refers to the situation wherein a particular allele of a marker is found with greater frequency in those with a particular phenotype.

2.4. Overview of the Candidate Gene Approach

One method of identifying disease genes is the candidate gene approach, in which common polymorphic genetic markers within a gene of interest, selected based on its hypothesized role in the disease, are evaluated to determine whether the polymorphisms are associated with the phenotype in populations or in families. This approach works because about 90% of variation is a result of common polymorphisms, most of which arose from single historical mutation events *(17)*. Each variant is associated with nearby variants that were present on the ancestral chromosome (or haplotype, the collection of variants existing together) on which the mutation occurred. These associations (referred to as linkage disequilibrium) allow the use of known markers to track down unknown disease mutations. In recent years, the availability of large numbers of DNA polymorphisms, both microsatellite repeats and SNPs, the development of the technology to perform relatively inexpensive, high-throughout genotyping, and improved statistical analysis approaches have made it feasible to comprehensively investigate the role of multiple candidate genes in disease susceptibility.

The candidate gene approach has resulted in the identification of a significant number of genes that contribute to susceptibility to various diseases. In type 2 diabetes mellitus (DM), for example, positive associations have been reported with the genes for the Kir6.2 subunit of the adenosine triphosphate (ATP)-sensitive potassium (K_{ATP}) channels of pancreatic βcells (*KCNJ11*), glycogen synthetase, IRS-1, peroxisome proliferator-activated receptor-γ (PPAR-γ 2), and PPAR-γ coactivator 1, among others *(18)*. Although none of these genes has been singly demonstrated to account for a large portion of diabetes susceptibility, replication in several studies suggests that these genes are indeed playing important roles in this disease.

2.5. Candidate Genes in PCOS: Progress and Challenges

To date, efforts to identify genes that influence PCOS susceptibility have largely utilized the candidate gene approach. Candidate genes studied in PCOS have generally targeted loci regulating five areas: (1) steroid biosynthesis and action (Table 1), (2) gonadotropic action (Table 2), (3) insulin secretion and action (Table 3), (4) weight and energy regulation (Table 3), and (5) cardiovascular factors (Table 4). The field has been comprehensively reviewed *(19)*. Several provocative genetic associations with PCOS have been reported that are slowly starting to illuminate the underlying causes of PCOS.

Despite repeated attempts to identify the putative gene or genes responsible for this disorder, the PCOS gene(s) remain elusive. As is evident in Table 4, despite many positive results, no gene or genes has clearly emerged as most important in PCOS, and many positive results were not confirmed in subsequent studies. Studies of the genetic etiology of PCOS have been hampered by various limitations, including (1) only one or two variants genotyped in each gene; (2) incomplete characterization of the phenotype in family members; (3) inability to assign a PCOS phenotype to prepubertal girls, postmenopausal women, and men; (4) possible inclusion of patients with nonclassic adrenal hyperplasia; (5) lack of appropriate controls; (6) unclear ethnic/racial composition; (7) varying criteria used to diagnose PCOS in different studies (in part because of the lack of universally accepted diagnostic criteria); and (8) small numbers of subjects in most studies. Regarding the latter, many of the studies in Tables 1–4 report results on fewer than 100 women with PCOS. Therefore, it is likely that many underpowered studies resulted in false-negative reports and that several small studies produced false-positive results.

The power issue is particularly relevant to common disease genetics. Validated genetic determinants of type 2 DM, such as the Pro12Ala variant of the *PPARG* gene and the Glu23Lys variant of the *KCNJ11* gene, only modestly alter risk for type 2 DM, on the order of 10–20% *(18)*. If genes with similar magnitude of effect influence PCOS risk, then many of the studies to date were seriously underpowered and inadequate to detect genetic variants predisposing to PCOS.

Table 1
Published Candidate Gene Studies in PCOS: Concerning Steroid Metabolism and Action

Genes	Comments
CYP17 (17α-hydroxylase/17,20-lyase)	A promoter polymorphism was associted with PCOS and hyperandrogenism in PCOS; however, this was not confirmed in several other studies. A microsatellite near *CYP17* was not associated with PCOS. No coding variants in *CYP17* were found in PCOS.
CYP11A (cholesterol side-chain cleavage enzyme)	A pentanucleotide VNTR was associated with PCOS in independent studies; however, other studies did not confirm this association, including one with a very large sample size.
AR (androgen receptor)	Decreased length of a polymorphic CAG repeat in exon 1 was found to be associated with hirsutism in Hispanic women and anovulation/infertility in women with PCOS. However, this polymorphism was not linked or associated with PCOS in other studies.
CYP21 (21-hydroxylase)	Heterozygosity for *CYP21* mutations was found more frequently in hyperandrogenic adolescents than controls. *CYP21* genotype was not found to be predictive of ovarian or adrenal hyperandrogenism in PCOS.
CYP19 (aromatase)	A linkage study and mutation screening were negative; however, an association study utilizing haplotypes showed association with precocious pubarche, PCOS, symptom score and testosterone level.
SHBG	Longer alleles of a promoter VNTR were associated with PCOS and lower SHBG levels.
H6PD (hexose-6-phosphate dehydrogenase)	Associated with PCOS and increased cortisol and 17-hydroxyprogesterone levels in one study.

Negative studies for linkage or association with PCOS or androgen levels have also been published regarding *HSD3B2*, *HSD17B3*, *UGT2B15*, SF-1, *StAR*, *DAX-1*, and *HSD11B1*.

VNTR, variable number tandem repeat; PCOS, polycystic ovary syndrome; SHBG, sex hormone-binding globulin.

Additional challenges face PCOS genetics as well as the genetics of other common, complex genetic disorders. Heterogeneity—in terms of multiple susceptibility genes, multiple alleles at a given gene, and possibly multiple subsets of PCOS with differing underlying pathophysiologies—confounds gene discovery. Gene-by-gene interactions that are likely to be important are difficult to model in genetic analyses of limited sample size. Environmental influences and gene-by-environment interactions are even more difficult to factor into genetic analyses. Local environmental and genetic differences between populations may explain why many results were found in one population but not another. For example, women with PCOS in the United States tend to have a greater body mass index than women with PCOS in Europe *(20)*. Because adiposity likely influences the expression and impact of various genes, this difference in body mass may have important implications for comparison of genetic results from the different regions.

Finally, inherent to the candidate gene approach are assumptions regarding the underlying pathophysiology of PCOS. This is a particular problem in PCOS, as the underlying causes are still fundamentally unknown. Candidate genes evaluated so far were selected from pathways affecting components of PCOS. Genes coding for transcription factors or signaling pathway components that may globally affect the organs involved in PCOS (pituitary, ovaries, adrenals, pancreatic β cells, insulin-responsive tissues) are unlikely to be selected as candidate genes by this approach.

Table 2
Published Candidate Gene Studies in PCOS: Gonadotropin Action and Regulation

Genes	Comments
Follistatin	Linkage of follistatin to PCOS has been demonstrated. Subsequent studies of follistatin in PCOS failed to discover any associated functional variants.
LH β-subunit	Coding variants were associated with elevated testosterone in healthy women and protection from developing PCOS in obese women. Three other studies showed no association of these variants with PCOS.
FSH β-subunit	A polymorphism in exon 3 was associated with PCOS; however, this variant does not alter the amino acid sequence.
Dopamine D3 receptor	Association with PCOS was found in a Hispanic population, but not in non-Hispanic white women.
FSH receptor	Several negative association studies have been reported. A positive study demonstrated association of coding variants with PCOS and responses to fertility treatment.

Studies have found no evidence for linkage and association of PCOS with several activin receptor subtypes, inhibins, the LH receptor, and the GnRH receptor.

PCOS, polycystic ovary syndrome; LH, luteinizing hormone; FSH, follicle-stimulating hormone, GnRH, gonadotropin-releasing hormone.

3. CONCLUSIONS

The inherited nature of PCOS, as well as hyperandrogenism and abnormal insulin action and metabolism in PCOS, has been firmly established by family studies. This led to over 100 publications focusing on logically chosen candidate genes. Unfortunately, most of these studies were in small cohorts. Additional issues such as only one or two variants genotyped for each candidate gene, phenotypic misclassification, genetic and phenotypic heterogeneity, and ignorance of the underlying causes of PCOS have confounded many genetic studies of PCOS. As a result, despite a large number of positive reports, no particular gene is universally recognized as importantly contributing to PCOS risk.

4. FUTURE AVENUES OF INVESTIGATION

Various new approaches may assist in the elucidation of involved genes, including haplotype-based analyses, inclusion of intermediate phenotypes, internal replication, whole-genome scan, and pharmacogenetics.

4.1. Advantages of the Haplotype-Based Approach to Genetic Association Studies

Almost every candidate gene study in PCOS has assessed the effect of one or two variants in each gene. This provides only partial information on whether a gene is associated with PCOS. There is increasing evidence that genetic variation is best described by groups of associated polymorphisms (inherited together on the same chromosome), referred to as haplotypes. Haplotypes reflect global gene structure, encompassing chromosomal blocks that have remained unbroken by recombination during the population history of the gene. These haplotypes encompass both coding and regulatory elements that individually or in combination could be responsible for the differences in phenotype and therefore can serve as important surrogates in the early stages of gene finding. Identification of a haplotype associated with increased or decreased disease risk should facilitate identification of the actual functional variant that affects disease risk, because this variant should lie on chromosomes identified by that haplotype.

Haplotypes carry more information than the individual SNPs that comprise them. Haplotypes capture the majority of common variation in a gene; consequently, the use of haplotypes is more likely to

Table 3
Published Candidate Gene Studies in PCOS: Insulin Secretion and Action, Obesity, and Fuel Metabolism

Genes	Comments
Insulin	A promoter VNTR was linked and associated with PCOS and/or insulin sensitivity in women with polycystic ovaries in some studies but not others. Analysis of multiple data sets showed no association of the VNTR with PCOS or testosterone levels.
Insulin receptor	An SNP in the tyrosine kinase domain was associated with PCOS, particularly in lean women. Other studies of the insulin receptor in women with PCOS have identified only common, silent polymorphisms.
Microsatellite D19S884 (located 1 cM from the insulin receptor gene)	Two independent studies have reported linkage and association of the microsatellite marker D19S884 with PCOS. A study of Italian and Spanish women with PCOS failed to confirm this association.
IRS1 and *IRS2* (insulin receptor substrates-1 and -2)	Variants of both *IRS1* and *IRS2* were found to influence fasting insulin and postload glucose levels, respectively, in women with PCOS. *IRS1* was associated with PCOS and with adolescent hyperandrogenism and obese insulin-resistant PCOS. *IRS1* genotype influenced response to metformin in PCOS. *IRS1* was not found to be linked or associated with PCOS or hyperandrogenemia in other studies.
CAPN10 (calpain-10)	One study found the 112/121 haplotype combination associated with a 2-fold increased risk of PCOS, but a larger study showed no such association. Two different studies from Spain produced conflicting results regarding association of SNP-44 with PCOS.
PPARG (PPAR-γ)	In Caucasian but not African American women with PCOS, the Pro12Ala variant was found to be associated with less insulin resistance. A marker near *PPARG* was not linked or associated with PCOS. Pro12Ala influenced body mass index in hyperandrogenic adolescents. Pro12Ala was associated with insulin sensitivity and lower hirsutism in German PCOS patients. Pro12Ala associated with obesity and increased insulin sensitivity in Turkish PCOS cases and controls. Pro12Ala associated with PCOS in a Finnish study. Negative association studies of Pro12Ala were also reported.
Resistin	A promoter variant was not associated with PCOS; however, it was associated with body mass index in PCOS.
IGF2 (insulin-like growth factor-2	An SNP in *IGF2* was associated with PCOS in a small study.
PPP1R3	A single variant was associated with insulin response to glucose challenge and hyperandrogenemia in PCOS.
Plasma cell membrane glycoprotein-1 (PC-1)	A coding SNP was associated with PCOS.

Negative studies have also been published concerning IGF-1, IGF-1 receptor, IGFBP1, IGFBP3, Leydig insulin-like protein 3, *SORBS1*, IGF-2 receptor, and *PTP1B*.

Negative linkage and association studies of genes relating to obesity and fuel metabolism in PCOS have been published concerning leptin and leptin receptor, glucocorticoid receptor, glycogen synthase, melanocortin 4 receptor, proopiomelanocortin, and uncoupling proteins 2 and 3.

PCOS, polycystic ovary syndrome; *VNTR*, variable number tandem repeat; SNP, single nucleotide polymorphism.

Table 4
Published Candidate Gene Studies in PCOS: Cardiovascular Disease

Genes	Comments
Paraoxonase	An SNP in paraoxonase was associated with PCOS in a small study.
PAI-1 (plasminogen activator inhibitor-1)	A promoter variant was associated with PCOS and increased PAI-1 levels. Negative association of PAI-1 with PCOS was also reported.
IL-6 (interleukin-6)	Two promoter SNPs were associated with PCOS. In another study, one of these SNPs was not associated with PCOS but was associated with BMI, testosterone, and response to glucose challenge.
IL-6 receptor complex	A polymorphism in the gp130 subunit was associated with hyperandrogenism.
Adiponectin	Positive associations with PCOS and fasting insulin in PCOS and negative association have been reported. Not associated with PCOS but associated with body mass index and hyperinsulinemia within PCOS.
EPHX (microsomal epoxide hydrolase)	A haplotype made up of only two coding SNPs was associated with PCOS.
Aldosterone synthatase	A promoter SNP was associated with PCOS and increased renin-angiotensin system activity.
Tumor necrosis factor receptor-2	A coding SNP was associated with PCOS in a small study.
Matrix metalloproteinase-1	Associated with PCOS in a small study.
Factor V	Factor V Leiden mutation found associated and not associated with PCOS in small studies from the same center.

Negative association studied have been reported for the β-adrenergic receptor, TNF-α (tumor necrosis factor-α), methylene tetrahydrofolate reductase, prothrombin, and apolipoprotein E genes.

PCOS, polycystic ovary syndrome; SNP, single nucleotide polymorphism.

identify gene variations than is the use of random SNPs *(21)*. Gabriel et al. *(22)* sequenced 13 megabases across the genome in subjects from Africa, Europe, and Asia. They showed that the human genome is organized in haplotype blocks (most of which are longer than 10 kilobases), with three to five commonly occurring (>5%) haplotypes per block. Only six to eight variants were sufficient to define the most common haplotypes in each block. Thus, a manageable number of appropriately chosen SNPs (termed haplotype-tagging SNPs [htSNPs]) can be genotyped to identify the most common haplotypes in a population, providing critical tools for association studies. The goal of the International HapMap Project is to delineate haplotype-tagging SNPs in all human genes *(23)*, which will greatly facilitate future haplotype-based association studies.

As discussed above, prior candidate gene studies in PCOS genotyped only one or a few SNPs per gene, indicating only incomplete coverage of each candidate gene. This is particularly true for larger genes that may contain multiple haplotype blocks. Application of the haplotype approach to PCOS genetics, particularly for genes wherein functional variants are unknown, should reduce the number of false-negative studies and may allow more positive findings to be replicated. To date, only studies of the calpain-10 gene in PCOS utilized haplotypes, based on the htSNPs from the original report associating calpain-10 with type 2 diabetes *(24)*. A few other PCOS studies constructed haplotypes from only two variants, unlikely to fully characterize haplotype blocks. Only recently did another gene, *CYP19* (aromatase), undergo haplotype analysis in PCOS. Previously, a study utilizing microsatellites in and around *CYP19* did not show linkage with PCOS *(25)*, and mutation screening in the exons and promoter failed to discover any variants *(26)*. In contrast, the association study utilizing haplotypes showed association with precocious pubarche (an antecedent of PCOS) as well

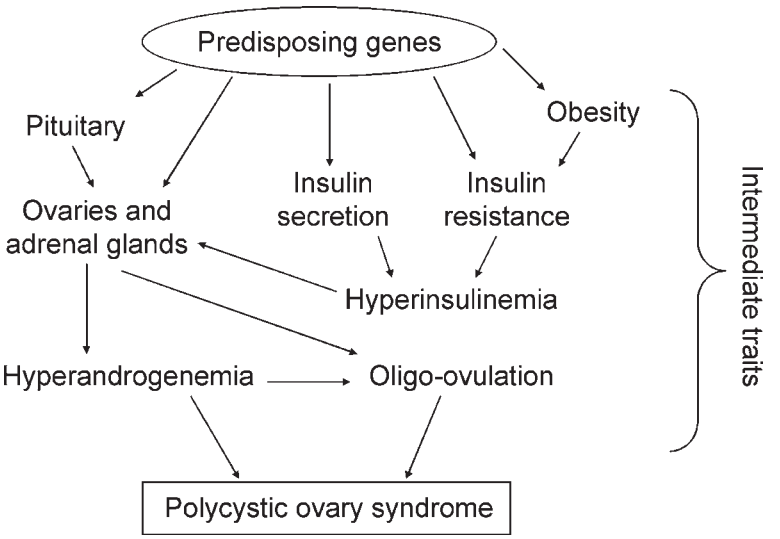

Fig. 1. Component (intermediate) traits in polycystic ovary syndrome (PCOS). In this conception of the pathophysiology of PCOS, it is apparent that intermediate traits are closer in the pathway to causative genes. In addition, these are quantitative variables that avoid the arbitrary nature of the qualitative variable of presence or absence of PCOS. Thus, genetic analyses utilizing intermediate traits may be powerful tools in gene discovery in PCOS.

as PCOS *(27)*. This demonstrates the potentially greater ability to discover genetic effects when using haplotypes to capture global variation across a gene.

4.2. Intermediate Phenotypes as the Focus of Genetic Study

Given that there are no universally accepted diagnostic criteria for PCOS, it is prudent for genetic analyses to focus on component (intermediate) phenotypes such as androgen levels, insulin resistance, and their related traits. Such intermediate phenotypes are likely to be closer to the basic abnormality, that is, action of predisposing genes, than the end-point disease (*see* Fig. 1). Another advantage of using intermediate phenotypes is that they can be analyzed as continuous variables. Thus, there is no arbitrary cut point to be imposed, and everyone with the quantitative trait measures (including healthy relatives and controls and males for insulin-related traits) can be used in the analyses. The latter increases the statistical power in identifying susceptibility genes in association studies. An advantage to measuring intermediate phenotypes in healthy subjects, especially at-risk relatives, is that it avoids secondary changes in phenotype caused by the disease itself or its therapy, which can obscure the detection of genetic influences.

Androgen-related intermediate phenotypes in PCOS include free or bioavailable testosterone, dehydroepiandrosterone (DHEA) and DHEAS, androstenedione and other androgens, sex hormone-binding globulin, and the modified Ferriman–Gallwey hirsutism score. Insulin-related phenotypes include fasting and postload glucose and insulin measurements, insulin sensitivity, and insulin secretion indices. Obesity is quantified by waist-to-hip ratio, body mass index, and body fat distribution measured by dual x-ray absorptiometry or computed tomography scanning. Of note, the most accurate phenotyping (but also the most expensive and labor-intensive) is performed with detailed physiological studies such as the euglycemic hyperinsulinemic clamp or frequently sampled intravenous glucose tolerance test for insulin-related traits (as opposed to fasting or oral glucose tolerance test measurements). There is a clear trend in the more recent candidate gene studies that association with PCOS is evaluated along with association with component phenotypes within PCOS (and sometimes also within controls).

4.3. Internal Replication as Validation of Positive Results

Genetic epidemiology is often criticized because positive reports of linkage or association are not subsequently confirmed by others. Besides false-positive and false-negative studies, such lack of replication may be a result of the study of different ethnic groups: a certain genetic variant may interact with other variants and local environmental influences such that it alters phenotype only in a particular group. Thus, replication studies should first be carried out in the same ethnic group, with the goal of validating the initial result. Subsequent replication attempts in other populations would serve to determine whether the particular genetic variant affects disease susceptibility universally. However, even when ostensibly the same ethnic group is studied, subtle differences in the history of the population may lead to ethnic differences such that two nominally similar cohorts are sufficiently different in genetic background to limit replication.

Additionally, a valid confirmation test should be carried out with the same phenotypes in a cohort from the same population, ascertained by identical criteria. This has been done only rarely in the field of common metabolic disorders. One example is the FUSION (Finland–United States Investigation of NIDDM Genetics) study, wherein an initial genome scan for type 2 diabetes was carried out in almost 500 affected sib-pair Finnish families (FUSION 1) *(28)* with the identification of 10 suggestive regions linked to type 2 diabetes. In FUSION 2, an independent set of 242 families ascertained in the same way *(29)*, four regions (on chromosomes 6, 11, 14, X) replicated loci found in FUSION 1, providing compelling evidence that these regions harbor genes for type 2 diabetes. Genetic studies in PCOS have generally not been replicated in this way, largely because of sample size limitations. With time and increasing subject recruitment, high-volume centers and/or collaborative consortia will be able to reassess their own positive results, providing critical validation of these results and prioritizing them for study by other groups.

4.4. The Whole-Genome Scan Approach to Gene Discovery

Linkage scans of the entire genome are often carried out using a panel of microsatellites (~400) that cover the whole genome. The key benefit of whole-genome approaches is that no prior knowledge or assumption regarding the underlying pathogenetic mechanism of disease is required (in contrast to the candidate gene approach). A challenge in conducting whole-genome linkage analyses is the need to recruit large numbers of families. In addition, linkage analyses are statistically less sensitive than association analyses. Thus, if genetic variants have small to modest effects on disease risk, they may not be detected by linkage analyses. Furthermore, positive linkage signals identify large chromosomal regions that often contain hundreds of genes. Fine-mapping with additional markers or selection of positional candidate genes is often performed to pursue positive linkage signals.

To date, no genome-wide linkage analysis has been published for PCOS. The main challenge has been recruiting enough families to conduct a linkage analysis with sufficient power. Investigators are actively pursuing this goal; genome-wide scans in PCOS are anticipated in the near future.

With further advances in high-throughput genotyping and statistical analysis, another option in the future will be whole-genome association studies, combining an assumption-free systematic search of the whole genome with the increased power of linkage disequilibrium (association) over allele sharing (linkage) for gene discovery.

4.5. Pharmacogenetics as a Probe to Understanding Disease Pathophysiology

Pharmacogenetics seeks to determine whether genetic variation influences response to drug therapy. Variants in drug-metabolizing enzymes may alter therapeutic response. Alternatively, variants in genes coding for components of key biological pathways may affect drug response. Knowledge of such variants allows prediction of therapeutic response and adverse effects. Pharmacogenetics may also be used as a probe to understanding disease pathophysiology. Drug therapy is a controlled environmental stimulus designed to provoke a genetically determined response. Without this physi-

ological challenge, these genetic variants may go undetected. In PCOS, the first pharmacogenetic study published observed that the Gly972Arg variant in the insulin receptor substrate-1 (*IRS1*) gene influenced response to metformin in terms of insulin and androgen levels *(30)*. This implicates IRS-1 as an important factor in the pathogenesis of PCOS, particularly as a participant in insulin resistance in PCOS. Future application of pharmacogenetics to PCOS will shed further light on its genetic determinants and may in the future lead to clinical genotyping to assist in decisions as to whom to treat and what agent(s) to use.

KEY POINTS

- PCOS clearly runs in families.
- Within PCOS, hyperandrogenemia and abnormalities in insulin action and secretion are also heritable.
- The candidate gene approach has been extensively applied to PCOS, with many intriguing positive results; however, to date, no genes have emerged as predominant.
- In the future, application of haplotypes, intermediate phenotypes, internal replication, genome-wide scans, and pharmacogenetics will significantly assist gene-discovery efforts in PCOS.

ACKNOWLEDGMENTS

Supported in part by an endowment from the Helping Hand of Los Angeles and by grants RO1-HD29364 and K24-D01346 from the National Institutes of Health.

REFERENCES

1. King RA, Rotter JI, Motulsky AG. The Genetic Basis of Common Diseases, 2nd ed. New York: Oxford University Press, 2002.
2. Legro RS, Strauss JF. Molecular progress in infertility: polycystic ovary syndrome. Fertil Steril 2002;78:569–576.
3. Kahsar-Miller MD, Nixon C, Boots LR, Go RC, Azziz R. Prevalence of polycystic ovary syndrome (PCOS) in first-degree relatives of patients with PCOS. Fertil Steril 2001;75:53–58.
4. Azziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The prevalence and features of the polycystic ovary syndrome in an unselected population. J Clin Endocrinol Metab 2004;89:2745–2749.
5. Legro RS, Driscoll D, Strauss JF 3rd, Fox J, Dunaif A. Evidence for a genetic basis for hyperandrogenemia in polycystic ovary syndrome. Proc Natl Acad Sci USA 1998;95:14956–14960.
6. Legro RS, Kunselman AR, Demers L, Wang SC, Bentley-Lewis R, Dunaif A. Elevated dehydroepiandrosterone sulfate levels as the reproductive phenotype in the brothers of women with polycystic ovary syndrome. J Clin Endocrinol Metab 2002;87:2134–2138.
7. Nelson VL, Legro RS, Strauss JF, McAllister JM. Augmented androgen production is a stable steroidogenic phenotype of propagated theca cells from polycystic ovaries. Mol Endocrinol 1999;13:946–957.
8. Legro RS, Bentley-Lewis R, Driscoll D, Wang SC, Dunaif A. Insulin resistance in the sisters of women with polycystic ovary syndrome: association with hyperandrogenemia rather than menstrual irregularity. J Clin Endocrinol Metab 2002;87:2128–2133.
9. Norman RJ, Masters S, Hague W. Hyperinsulinemia is common in family members of women with polycystic ovary syndrome. Fertil Steril 1996;66:942–947.
10. Yildiz BO, Yarali H, Oguz H, Bayraktar M. Glucose intolerance, insulin resistance, and hyperandrogenemia in first degree relatives of women with polycystic ovary syndrome. J Clin Endocrinol Metab 2003;88:2031–2036.
11. Kaushal R, Parchure N, Bano G, Kaski JC, Nussey SS. Insulin resistance and endothelial dysfunction in the brothers of Indian subcontinent Asian women with polycystic ovaries. Clin Endocrinol (Oxf) 2004;60:322–328.
12. Colilla S, Cox NJ, Ehrmann DA. Heritability of insulin secretion and insulin action in women with polycystic ovary syndrome and their first degree relatives. J Clin Endocrinol Metab 2001;86:2027–2031.
13. Ehrmann DA, Sturis J, Byrne MM, Karrison T, Rosenfield RL, Polonsky KS. Insulin secretory defects in polycystic ovary syndrome. Relationship to insulin sensitivity and family history of non-insulin-dependent diabetes mellitus. J Clin Invest 1995;96:520–527.
14. Dunaif A, Wu X, Lee A, Diamanti-Kandarakis E. Defects in insulin receptor signaling in vivo in the polycystic ovary syndrome (PCOS). Am J Physiol Endocrinol Metab 2001;281:E392–E399.
15. Rosenbaum D, Haber RS, Dunaif A. Insulin resistance in polycystic ovary syndrome: decreased expression of GLUT-4 glucose transporters in adipocytes. Am J Physiol 1993;264:E197–E202.
16. Ciaraldi TP, el-Roeiy A, Madar Z, Reichart D, Olefsky JM, Yen SS. Cellular mechanisms of insulin resistance in polycystic ovarian syndrome. J Clin Endocrinol Metab 1992;75:577–583.

17. Kruglyak L, Nickerson DA. Variation is the spice of life. Nat Genet 2001;27:234–236.
18. Parikh H, Groop L. Candidate genes for type 2 diabetes. Rev Endocr Metab Disord 2004;5:151–176.
19. Escobar-Morreale HF, Luque-Ramirez M, San Millan JL. The molecular-genetic basis of functional hyperandrogenism and the polycystic ovary syndrome. Endocr Rev 2005;26:251–282.
20. Carmina E, Legro RS, Stamets K, Lowell J, Lobo RA. Difference in body weight between American and Italian women with polycystic ovary syndrome: influence of the diet. Hum Reprod 2003;18:2289–2293.
21. Judson R, Salisbury B, Schneider J, Windemuth A, Stephens JC. How many SNPs does a genome-wide haplotype map require? Pharmacogenomics 2002;3:379–391.
22. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science 2002;296:2225–2229.
23. The International HapMap Consortium. The International HapMap Project. Nature 2003;426:789–796.
24. Horikawa Y, Oda N, Cox NJ, et al. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet 2000;26:163–175.
25. Gharani N, Waterworth DM, Batty S, et al. Association of the steroid synthesis gene CYP11a with polycystic ovary syndrome and hyperandrogenism. Hum Mol Genet 1997;6:397–402.
26. Soderlund D, Canto P, Carranza-Lira S, Mendez JP. No evidence of mutations in the P450 aromatase gene in patients with polycystic ovary syndrome. Hum Reprod 2005;20:965–969.
27. Petry CJ, Ong KK, Michelmore KF, et al. Association of aromatase (CYP 19) gene variation with features of hyperandrogenism in two populations of young women. Hum Reprod 2005;20:1837–1843.
28. Ghosh S, Watanabe RM, Valle TT, et al. The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes. Am J Hum Genet 2000;67:1174–1185.
29. Silander K, Scott LJ, Valle TT, et al. A large set of Finnish affected sibling pair families with type 2 diabetes suggests susceptibility loci on chromosomes 6, 11, and 14. Diabetes 2004;53:821–829.
30. Ertunc D, Tok EC, Aktas A, Erdal EM, Dilek S. The importance of IRS-1 Gly972Arg polymorphism in evaluating the response to metformin treatment in polycystic ovary syndrome. Hum Reprod 2005;20:1207–1212.