



Published in final edited form as:

*Fertil Steril.* 2016 December ; 106(7): 1772–1777. doi:10.1016/j.fertnstert.2016.08.033.

## The Relationship Between Polycystic Ovary Syndrome and Ancestry in European Americans

Andrew C. Bjornes, MS<sup>a</sup>, Richa Saxena, PhD<sup>a</sup>, and Corrine K. Welt, MD<sup>b,1</sup>

<sup>a</sup>Department of Anaesthesia, Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, 02114 USA

<sup>b</sup>Division of Endocrinology, Metabolism and Diabetes, University of Utah, Salt Lake City, UT, 84112 USA

### Abstract

**Objective**—To determine whether European Americans with PCOS would exhibit genetic differences associated with PCOS status and phenotypic features.

**Design**—The study was a case-control association study in European Americans.

**Setting**—Subjects were studied in an academic center.

**Subjects**—Women with PCOS diagnosed using the NIH criteria (n=532) and controls with regular menstrual cycles and no evidence of hyperandrogenism (n=432) were studied.

**Interventions**—Blood was drawn for measurement of sex steroids, metabolic parameters and genotyping.

**Main outcome measure**—Associations were identified between PCOS status, phenotype and genetic background determined using principal components.

**Results**—Principal component analysis identified 5 principal components (PCs). PC1 captured northwest to southeast European genetic variation and was associated with PCOS status. Acanthosis was associated with southern European ancestry, while larger waist:hip ratio was associated with northern European ancestry. PC2 was associated with east to west European genetic variation and cholesterol levels.

**Conclusions**—These data provide evidence for genetic influence based on European ethnicity in women with PCOS. There is also evidence for a genetic component in the phenotypic features of PCOS within a mixed European population. The data point to the need to control for population stratification in genetic studies in women of mixed European ethnicity. They also emphasize the need for better studies of PCOS prevalence and phenotype as a function of genetic background.

<sup>1</sup>Corresponding author and person to who reprint requests should be addressed: Corrine Welt, Division of Endocrinology, Metabolism and Diabetes, University of Utah, EIHG, 15 N 2030 E, Salt Lake City, UT 84112, USA, Phone: 801-585-1875, cwelt@genetics.utah.edu.

C.K.W. has received royalties from UptoDate unrelated to the current topic. There are no other conflicts to report.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Keywords

European; Genetics; Population stratification; waist:hip ratio

---

## Introduction

It is important to control for ancestry in genetic association studies to avoid false positive results from confounding population differences between cases and controls. Genetic association studies in Americans of European descent may be particularly prone to population stratification (1). While it is often accepted that Europeans are genetically homogeneous, there is distinct genetic substructure in the European population (2–5). Three unique populations can be identified, with individuals of northwest European, southeast European and Ashkenazi Jewish ancestry sharing genetic substructure (1, 2, 6). The genetic substructure can affect studies in European countries with multiple immigrant populations, and countries such as the United States, in which ethnic groups from these European regions have distinct immigration patterns (2, 5). Thus, population stratification may occur if the ethnic populations are not carefully matched. Genome-wide association studies employ a predetermined set of variants that can be used to examine, and control for, population substructure and stratification. However, these European population markers may not be taken into account in candidate association and replication studies. They also fail to prevent spurious associations from occurring when using next generation sequencing to study associations between disease and rare exome variants (<5%)(7).

Previous studies have examined population stratification among ethnically diverse women with polycystic ovary syndrome. A multiethnic group of women with PCOS from the Netherlands (8) was examined using a genome-wide panel of ancestry informative markers that distinguished women of African, Southeast Asian, Hindustani and European ethnicity. Six distinct clusters were identified representing the distinct ethnic subgroups of African, Surinam Creole, Asian and Caribbean and African ethnicities. Importantly, women of northern European and Turkish ethnicity were clustered into two distinct groups. In addition, the genetic ancestral background accounted for a proportion of the phenotypic features of PCOS (9), with previous work by the same group demonstrating that the subset of women from Mediterranean Europe manifested greater obesity and hyperandrogenism compared to other ethnic groups in that study (10). In addition, a mixed European group from Boston manifested greater hyperandrogenism than women from Iceland (11). Therefore, the PCOS phenotype may include distinctive features depending on European ethnic origin (10, 11), and these may be partially determined by differences in genetics. Taken together, the genetic and phenotypic distinctions among Europeans may result in differences in ascertainment of PCOS or in expression of its features in distinct ethnic groups. These studies make it important to examine a broad population of European women, and compare genetic stratification in women with PCOS and controls.

We hypothesized that European population stratification would be present in association studies of PCOS in European Americans. Based on data from the Netherlands, we also hypothesized that phenotypic features in women with PCOS of European ancestry would

exhibit differences based on the southeast to northwest population substructure of European Americans (1, 2, 6). To test these hypotheses, we analyzed PCOS status and phenotype as a function of the principal components of population structure, using markers informative for European ancestry, in our cohort from a genome-wide association study (GWAS) of women with PCOS and controls of European ethnicities (12, 13).

## Research Design and Methods

### Subjects

All subjects were U.S. women of reported European ethnicity, between the ages of 18 and 45 years, and recruited at Massachusetts General Hospital in Boston, MA. Subjects with PCOS (n=532) had oligomenorrhea (< 9 menstrual periods/yr) and clinical and/or biochemical evidence of hyperandrogenism, fulfilling the NIH criteria (11). Clinical hyperandrogenism was defined by: 1) an elevated Ferriman Gallwey score > 9 (14); or 2) acne on the face or back. Biochemical hyperandrogenism was defined as testosterone >63 ng/dL (2.8 nmol/L), DHEAS >430 µg/dL (1.16 µmol/L) or androstenedione levels >3.8 ng/mL (13.3 nmol/L) (11). Control subjects had regular menstrual cycles, 21 to 35 days, and no physical exam or biochemical evidence of hyperandrogenism (n=432).

Subjects were excluded for a personal history or biochemical evidence of late onset congenital adrenal hyperplasia (11). All subjects had normal thyroid function and prolactin levels and a follicular phase FSH level in the premenopausal range. Subjects were on no hormonal medication, except for stable thyroid hormone replacement.

### Protocol

The study was approved by the Institutional Review Board of the Massachusetts General Hospital, and all subjects gave written informed consent. All PCOS subjects were studied >10 days after their last menstrual period and after a 12 hour fast (11). Subjects underwent a detailed history; physical exam; a pelvic ultrasound (ATL HDI 1500, 5 MHz convex array transducer); and blood samples for lipids, glucose, insulin, gonadotropin and sex-steroid levels. An oral glucose tolerance test was performed, with blood sampling 2 hours after a 75 gram glucose load.

### Genotyping

Patient DNA was isolated from whole blood and genotyped using the OmniHumanExpress Bead Chip (Illumina, San Diego, CA) with 951,117 SNPs. Subjects were removed for inbreeding (n=16) and for population stratification after analysis using Eigenstrat for subjects failing to cluster with European cohorts (n=60), with some samples excluded for both (n=15). Single nucleotide polymorphisms with more than 5% missing genotype were excluded.

### Statistical Analysis

**PCOS Status**—A subset of 240,000 markers informative for European, African American and Latin ancestry was used for analysis (2, 15, 16). These variants were used to determine the genetic variability mathematically, by structuring the data into principal components.

Principal components analysis was performed using Eigensoft for cases, controls and the combined groups with age and BMI as a covariates (17, 18).

**Phenotype**—Quantitative traits were log transformed for analysis. Logistic or linear regression analysis was used to examine associations between the five principal components identified and PCOS status and 17 log-transformed quantitative traits in the combined sample of PCOS cases and controls and the cases and controls as separate groups, adjusting for PCOS status, age and BMI. A  $p$  value  $< 0.007$  was considered significant after Benjamini and Hochberg False Discovery Rate correction for 5 PCAs and 10 independent traits, with other variables highly correlated (trait family [correlated measurements]; gonadotropins [LH, FSH], 17-OH progesterone, testosterone [androstenedione, DHEAS, SHBG], cholesterol [LDL, HDL], acanthosis nigricans, blood pressure [systolic blood pressure, diastolic blood pressure], body mass index (BMI, waist:hip ratio), fasting glucose, fasting insulin and ovarian volume). Data were plotted against that obtained from European-based Human Genome Diversity Project datasets, including Italian from Bergamo, Tuscan (Central Italy), Russian, Orcadian (Orkney Islands, Scotland), French, Basque (Northern Spain and Southern France), Sardinian (Autonomous Italian Island) and Adygei (Republic of Russia, Caucasian) (19–21).

## Results

Principal component analysis identified 5 principal components (PCAs). Principal component 1 correlated with the first principal component in the Human Genome Diversity Project ( $r=0.75$ )(19–21), which captured northwest to southeast European genetic variation. Principal component 2 correlated with east to west genetic variation from Basque to Adygei, near Turkey. Principal components 3 through 5 were not discernible when plotted with previously genotyped European populations.

There was a relationship between PCOS status and principal component 1 (Figure 1). PCOS cases were stratified with a predominant south to north European cline, while controls were stratified with a north to south European cline ( $p<0.002$ ). When specific populations were graphed against the current data, PCOS cases clustered statistically closer the Tuscan Italian and other Italian populations, whereas the controls clustered closer to the French and Orcadian populations.

In cases alone, acanthosis nigricans was associated with principal component 1, with acanthosis nigricans more prevalent in southern European compared to northern European ancestry (Table 1). The relationship was not apparent after controlling for BMI. In contrast, the waist:hip ratio was higher in northern compared to southern European ancestry in the combined group after controlling for BMI.

Principal component 2 represented an east to west European population cline. It was not associated with PCOS status, but it was associated with total cholesterol and LDL cholesterol levels in the controls.

Three additional principal components were identified. There was no precise relationship between principal components 3 through 5 and ethnicity. In PCOS cases, principal

component 3 was associated with FSH levels and principal component 4 was associated with 17-OH progesterone levels (Table 1).

## Discussion

Principal component analysis using 240,000 genetic variants demonstrates an association between PCOS status and principal component 1, which identifies the southeast to northwest genetic cline in the European population. PCOS cases demonstrated a greater association with southeast European populations and a greater superimposition with available Italian population data. The controls exhibited greater association with the northwest European populations and greater superimposition with available French and Orcadian population data. There was no obvious phenotype that might account for the stratification, with the exception of acanthosis nigricans, which was also associated with southeast ethnic populations. The data emphasize the need for studies examining individual ethnic groups simultaneously to determine phenotypic and prevalence differences in PCOS. They also highlight the need to control for population stratification in genetic association studies of European women with PCOS from the U.S.

Principal component analysis condenses information from genetic variants into groups using a multivariate analysis (22). Previous studies demonstrate that in European Americans, two major axes of variation exist; one for southeast to northwest European origin and the second representing southeast Europeans and Ashkenazi Jewish populations (2). In these studies, southeastern Europeans were represented by Greeks, Ashkenazi Jewish individuals, some Italians and Armenians, whereas northwestern Europeans were represented by Polish, Irish, English, German, Swedish and some Italian populations (1, 6). The gradient may represent ancestral migration patterns or population blocks that are isolated by distance from other blocks (22, 23), although the cause of the genetic cline remains controversial (6). Individuals of Spanish and Portuguese origin have been demonstrated to cluster with the southeast European group or form their own cluster, depending on the variants used to genotype (1, 6). In the current study, the principal component analysis did not include an Ashkenazi Jewish population. Therefore, the east to west variation represents the population cline from the Adygei to the Basque. These clusters are important in genetic studies based in countries with distinct European population migration, such as the U.S. (2, 5).

A study of height in European Americans illustrated the problem of population stratification (3). In this study a variant in the lactase gene (*LCT*), which exhibits a wide variation in allele frequency across European populations with a similar distribution to the cline in height, demonstrated a false positive association with height. A large number of ancestry informative markers are included in genome-wide association studies, allowing correction for this type of population stratification. However, a panel of ancestry informative markers should be employed to control for population stratification in PCOS case-control association and replication studies in European Americans and other European countries with multiple immigrant groups when a genome-wide panel is not used (3). Rare variants (<5% frequency) are also subject to spurious association if case and control populations are not properly matched because rare variants are more likely to originate in a differentiated subpopulation (7). Previous data suggest that a panel of 100 markers was sufficient to control for

populations stratification from northwest to southeast Europe, while 300 markers were necessary to control for population stratification in height from subjects predominantly of Ashkenazi Jewish descent (2, 18).

PCOS is not the only disorder in which a potential relationship to genetic ancestry exists. In IgA nephropathy, a genetic risk score demonstrated higher genetic risk in northern compared to southern Europeans (24). Further, among patients with end-stage renal disease who underwent biopsies, the prevalence of IgA nephropathy exhibited a north-to-south gradient that mirrored the genetic risk. These gradients continue to exist in population groups that have immigrated to the U.S., supporting a genetic rather than an environmental effect.

If there is genetic influence in the PCOS diagnosis, it might be expected that disease prevalence also varies by European ethnicity. For IgA nephropathy, the prevalence reflects the genetic risk, with the prevalence lowest in the Southern Europeans (Spain, Italy and Greece) and the highest in Northern Europeans (Sweden, Finland and Iceland). When defined using the NIH criteria, the prevalence of PCOS is 6.8% in Greece (25), 5.1–6.5% in Spain (26, 27), 6% in Italy (26) and 6.1% in Turkey (28). There are no true prevalence studies in women of purely Northern European ethnicity, but estimates suggest that the prevalence is 4% in women from the U.K. recruited in a study requesting participation in health screening (29). In contrast, the prevalence was 8.7% in Australian women of mixed European ethnicity (1/3 southern European) (30), and 4.8% in white women from the southeast U.S. (31). Taken together, the prevalence estimates are similar but not identical in distinct populations. Importantly, some of the studies were not true population prevalence studies and all were performed by distinct investigative groups and may therefore not be directly comparable. A multi-ethnic prevalence study performed by a team of investigators using identical assays and physical exams is needed.

In addition to potentially affecting the prevalence of a disorder, ancestry differences could account for differences in phenotypic expression, possibly independent of the underlying disorder (6). A recent study using the allele frequency of previously identified genetic variants that confer risk for PCOS and published phenotypes suggests stratification into 5 population groups with two main phenotypes; metabolic and hyperandrogenic (32). Phenotypic data in the current study were associated with genetic population clusters. In particular, acanthosis nigricans and waist:hip ratio were associated with the south to north European population clusters. Similarly, a study from the Netherlands also suggested a strong relationship between genetic background and phenotype (8). Women from Mediterranean Europe manifested greater obesity and hyperandrogenism than other ethnic groups (10), and the genetic cluster including Turkish women was associated with higher free androgen index and insulin levels and lower SHBG (8). Taken together, the PCOS phenotype may vary among ethnic groups of European genetic origin (8, 10, 11).

We also demonstrated an association between cholesterol and LDL levels and east to west population variation. Previous studies have demonstrated differences in cholesterol levels among men in European countries (Finland, Netherlands, Italy, Crete and Serbia)(33). In addition, we previously demonstrated lower HDL levels in women with PCOS from Iceland compared to women of mixed ethnicity from Boston (11). Lower cholesterol levels were



found in Mediterranean countries compared to those of northern Europe. Similarly, the east to west PC2 was associated with lower LDL levels in the east, encompassing countries bordering the Mediterranean Sea. It is likely that diet played a large role in the differences between cholesterol levels in previous studies, although genetic differences were not excluded (33). These dietary differences would not be expected to be as marked in European women from the U.S., as in the current study. Thus, genetic substructure may also explain the differences in LDL levels in European women with PCOS living in Boston.

The underlying cause for the difference in the principal component structure between PCOS cases and controls is not clear. Results point to the possibility that phenotypic features vary on the basis of genetic background for both PCOS cases and controls. It is possible that subjects with southern European ancestry may be more likely to be diagnosed with PCOS based on their darker hair color. Conversely, the subjects with dark hair may have been excluded from the controls because of hirsutism. However, the fact that the phenotypic features were not stratified by north-to-south make the possibility unlikely. In particular, Ferriman Gallwey score was not stratified by the first principal component and one previous study comparing Ferriman Gallwey scores in European and Asian women in San Francisco did not see a difference in the scores between these two ethnic groups (34). The results may be related to underlying genetic factors that are not reflected in the measured phenotypic parameters, but that make PCOS more prominent in the southeast Europeans. Future studies will be able to examine the overall frequency of PCOS risk alleles in distinct ethnic populations.

## Conclusions

These data provide evidence for population stratification in women of European ethnicity with PCOS. In addition, there is evidence for a genetic component in the phenotypic features of PCOS within a mixed European population. The data point to the need to control for population stratification in genetic studies in women of mixed European ethnicity. They also emphasize the need for better studies of PCOS prevalence and phenotype as a function of genetic background.

## Acknowledgments

The project described was supported by Award Number R01HD065029 from the Eunice Kennedy Shriver National Institute Of Child Health & Human Development, Award Number 1 UL1 RR025758, Harvard Clinical and Translational Science Center, from the National Center for Research Resources and award 1-10-CT-57 from the American Diabetes Association. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute Of Child Health & Human Development, National Center For Research Resources, the National Institutes of Health or the American Diabetes Association.

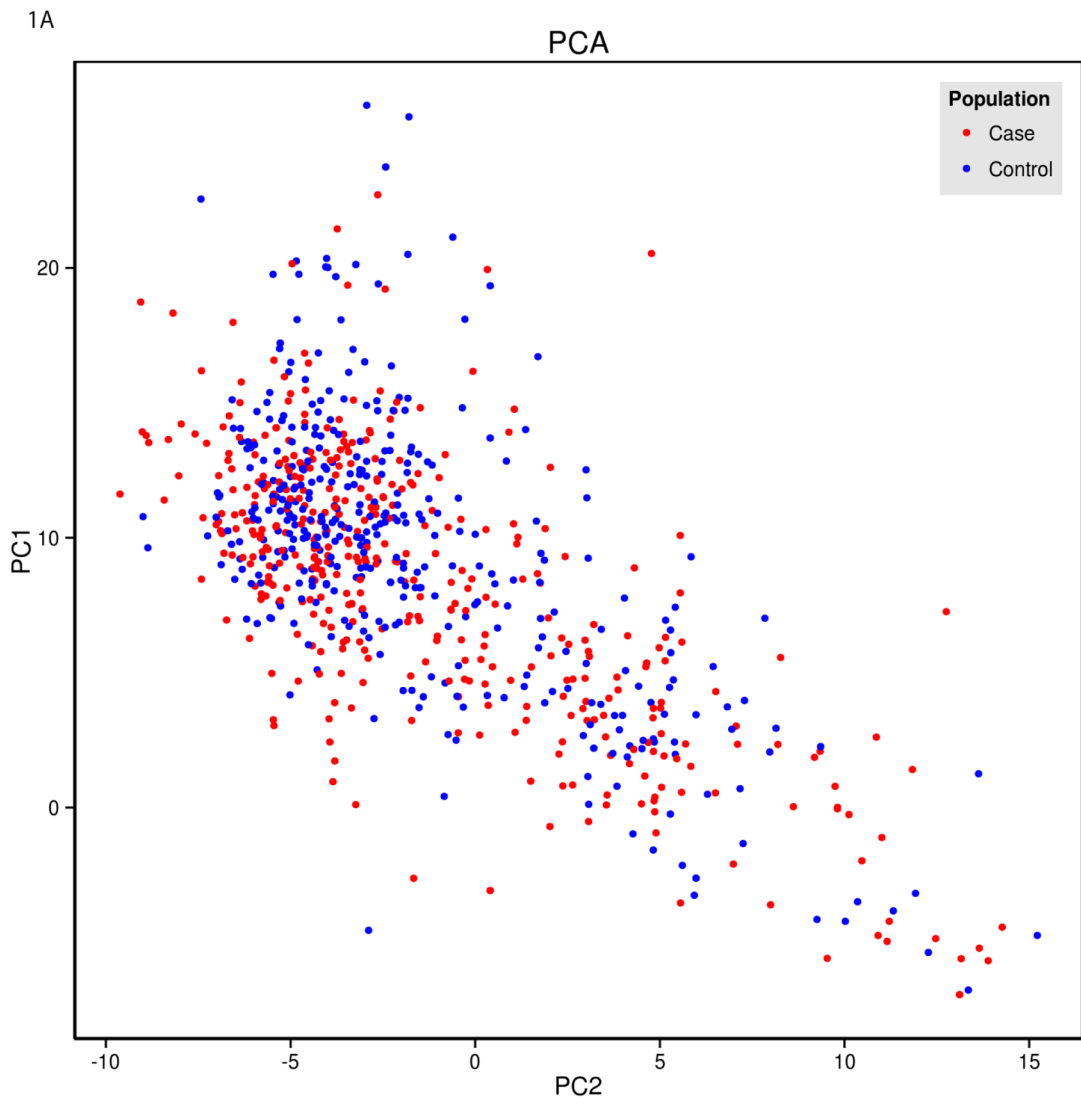
## References

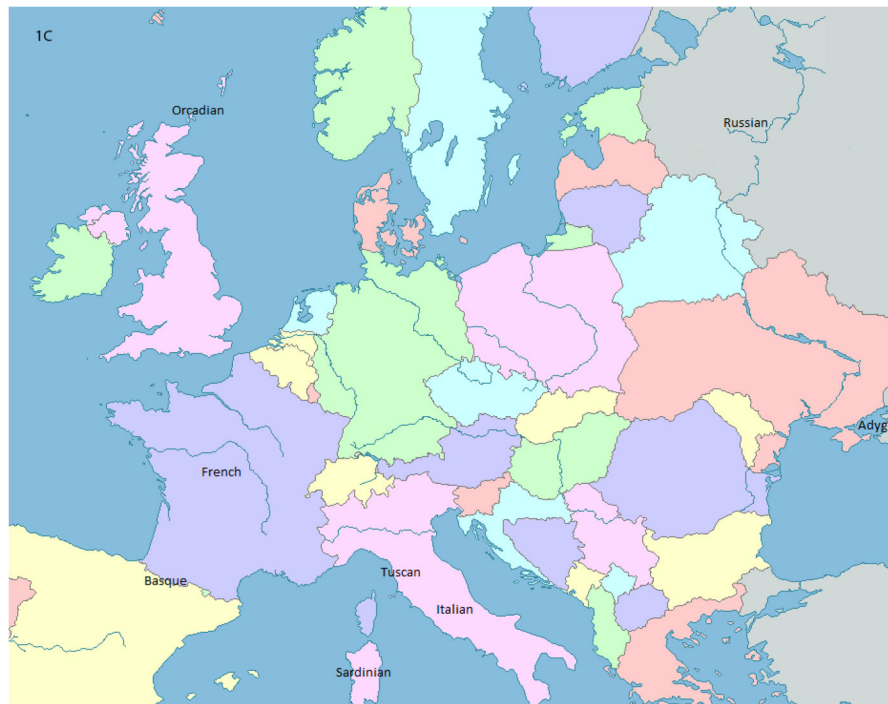
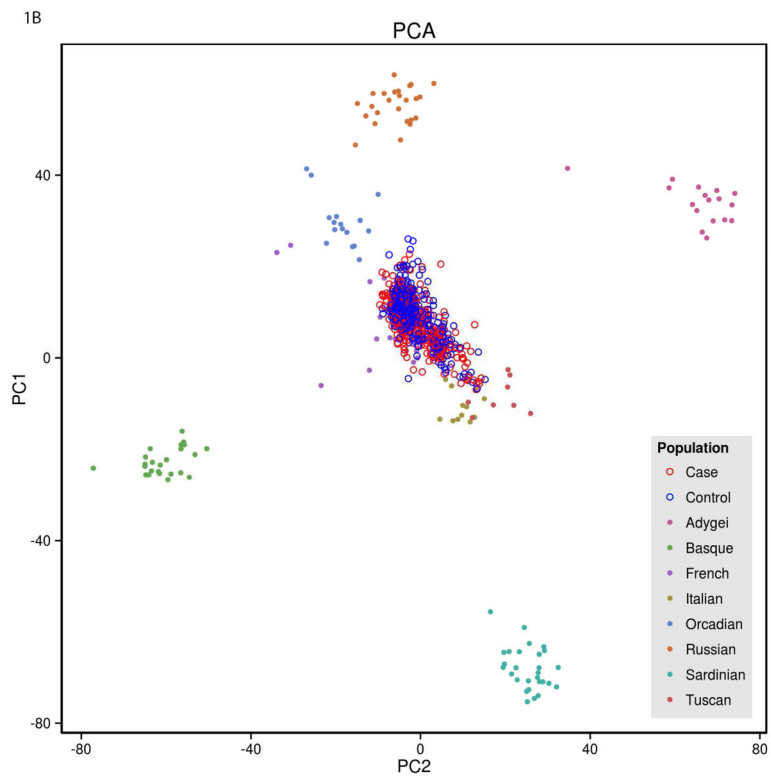
1. Bauchet M, McEvoy B, Pearson LN, Quillen EE, Sarkisian T, Hovhannesyan K, et al. Measuring European Population Stratification with Microarray Genotype Data. *Am J Hum Genet.* 2007; 80:948–56. [PubMed: 17436249]
2. Price AL, Butler J, Patterson N, Capelli C, Pascali VL, Scarnicci F, et al. Discerning the Ancestry of European Americans in Genetic Association Studies. *PLoS Genet.* 2008; 4:e236. [PubMed: 18208327]

3. Campbell CD, Ogburn EL, Lunetta KL, Lyon HN, Freedman ML, Groop LC, et al. Demonstrating Stratification in a European American Population. *Nat Genet.* 2005; 37:868–72. [PubMed: 16041375]
4. Bernardi F, Arcieri P, Bertina RM, Chiarotti F, Corral J, Pinotti M, et al. Contribution of Factor Vii Genotype to Activated Fvii Levels. Differences in Genotype Frequencies between Northern and Southern European Populations. *Arterioscler Thromb Vasc Biol.* 1997; 17:2548–53. [PubMed: 9409226]
5. Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL. The Genetic Ancestry of African Americans, Latinos, and European Americans across the United States. *Am J Hum Genet.* 2015; 96:37–53. [PubMed: 25529636]
6. Seldin MF, Shigeta R, Villoslada P, Selmi C, Tuomilehto J, Silva G, et al. European Population Substructure: Clustering of Northern and Southern Populations. *PLoS Genet.* 2006; 2:e143. [PubMed: 17044734]
7. O'Connor TD, Kiezun A, Bamshad M, Rich SS, Smith JD, Turner E, et al. Fine-Scale Patterns of Population Stratification Confound Rare Variant Association Tests. *PLoS One.* 2013; 8:e65834. [PubMed: 23861739]
8. Louwers YV, Lao O, Fauser BC, Kayser M, Laven JS. The Impact of Self-Reported Ethnicity Versus Genetic Ancestry on Phenotypic Characteristics of Polycystic Ovary Syndrome (PCOS). *J Clin Endocrinol Metab.* 2014; 99:E2107–16. [PubMed: 24960542]
9. Louwers YV, Stolk L, Uitterlinden AG, Laven JS. Cross-Ethnic Meta-Analysis of Genetic Variants for Polycystic Ovary Syndrome. *J Clin Endocrinol Metab.* 2013; 98:E2006–12. [PubMed: 24106282]
10. Valkenburg O, Lao O, Schipper I, Louwers Y, Uitterlinden AG, Kayser M, et al. Genetic Ancestry Affects the Phenotype of Normogonadotropic Anovulatory (WHOII) Subfertility. *J Clin Endocrinol Metab.* 2011; 96:E1181–7. [PubMed: 21525156]
11. Welt CK, Arason G, Gudmundsson JA, Adams J, Palsdottir H, Gudlaugsdottir G, et al. Defining Constant Versus Variable Phenotypic Features of Women with Polycystic Ovary Syndrome Using Different Ethnic Groups and Populations. *J Clin Endocrinol Metab.* 2006; 91:4361–8. [PubMed: 16940441]
12. Day FR, Hinds DA, Tung JY, Stolk L, Styrkarsdottir U, Saxena R, et al. Causal Mechanisms and Balancing Selection Inferred from Genetic Associations with Polycystic Ovary Syndrome. *Nat Commun.* 2015; 6:8464. [PubMed: 26416764]
13. Hayes MG, Urbanek M, Ehrmann DA, Armstrong LL, Lee JY, Sisk R, et al. Genome-Wide Association of Polycystic Ovary Syndrome Implicates Alterations in Gonadotropin Secretion in European Ancestry Populations. *Nat Commun.* 2015; 6:7502. [PubMed: 26284813]
14. Ferriman D, Gallwey JD. Clinical Assessment of Body Hair Growth in Women. *J Clin Endocrinol Metab.* 1961; 21:1440–7. [PubMed: 13892577]
15. Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, Waliszewska A, et al. A High-Density Admixture Map for Disease Gene Discovery in African Americans. *Am J Hum Genet.* 2004; 74:1001–13. [PubMed: 15088270]
16. Price AL, Patterson N, Yu F, Cox DR, Waliszewska A, McDonald GJ, et al. A Genomewide Admixture Map for Latino Populations. *Am J Hum Genet.* 2007; 80:1024–36. [PubMed: 17503322]
17. Patterson N, Price AL, Reich D. Population Structure and Eigenanalysis. *PLoS Genet.* 2006; 2:e190. [PubMed: 17194218]
18. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal Components Analysis Corrects for Stratification in Genome-Wide Association Studies. *Nat Genet.* 2006; 38:904–9. [PubMed: 16862161]
19. Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, et al. Genetic Structure of Human Populations. *Science.* 2002; 298:2381–5. [PubMed: 12493913]
20. Rosenberg NA, Mahajan S, Ramachandran S, Zhao C, Pritchard JK, Feldman MW. Clines, Clusters, and the Effect of Study Design on the Inference of Human Population Structure. *PLoS Genet.* 2005; 1:e70. [PubMed: 16355252]



21. Cann HM, de Toma C, Cazes L, Legrand MF, Morel V, Piouffre L, et al. A Human Genome Diversity Cell Line Panel. *Science*. 2002; 296:261–2. [PubMed: 11954565]
22. Menozzi P, Piazza A, Cavalli-Sforza L. Synthetic Maps of Human Gene Frequencies in Europeans. *Science*. 1978; 201:786–92. [PubMed: 356262]
23. Reich D, Price AL, Patterson N. Principal Component Analysis of Genetic Data. *Nat Genet*. 2008; 40:491–2. [PubMed: 18443580]
24. Kiryluk K, Novak J, Gharavi AG. Pathogenesis of Immunoglobulin a Nephropathy: Recent Insight from Genetic Studies. *Annu Rev Med*. 2013; 64:339–56. [PubMed: 23072577]
25. Diamanti-Kandarakis E, Kouli CR, Bergiele AT, Filandra FA, Tsianateli TC, Spina GG, et al. A Survey of the Polycystic Ovary Syndrome in the Greek Island of Lesbos: Hormonal and Metabolic Profile. *J Clin Endocrinol Metab*. 1999; 84:4006–11. [PubMed: 10566641]
26. Sanchon R, Gambineri A, Alpanes M, Martinez-Garcia MA, Pasquali R, Escobar-Morreale HF. Prevalence of Functional Disorders of Androgen Excess in Unselected Premenopausal Women: A Study in Blood Donors. *Hum Reprod*. 2012; 27:1209–16. [PubMed: 22343706]
27. Asuncion M, Calvo RM, San Millan JL, Sancho J, Avila S, Escobar-Morreale HF. A Prospective Study of the Prevalence of the Polycystic Ovary Syndrome in Unselected Caucasian Women from Spain. *J Clin Endocrinol Metab*. 2000; 85:2434–8. [PubMed: 10902790]
28. Yildiz BO, Bozdag G, Yapici Z, Esinler I, Yarali H. Prevalence, Phenotype and Cardiometabolic Risk of Polycystic Ovary Syndrome under Different Diagnostic Criteria. *Hum Reprod*. 2012; 27:3067–73. [PubMed: 22777527]
29. Michelmore KF, Balen AH, Dunger DB, Vessey MP. Polycystic Ovaries and Associated Clinical and Biochemical Features in Young Women. *Clin Endocrinol*. 1999; 51:779–86.
30. March WA, Moore VM, Willson KJ, Phillips DI, Norman RJ, Davies MJ. The Prevalence of Polycystic Ovary Syndrome in a Community Sample Assessed under Contrasting Diagnostic Criteria. *Hum Reprod*. 2010; 25:544–51. [PubMed: 19910321]
31. 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; 89:2745–9. [PubMed: 15181052]
32. Casarini L, Brigante G. The Polycystic Ovary Syndrome Evolutionary Paradox: A Genome-Wide Association Studies-Based, in Silico, Evolutionary Explanation. *J Clin Endocrinol Metab*. 2014; 99:E2412–20. [PubMed: 25093623]
33. Verschuren WM, Jacobs DR, Bloemberg BP, Kromhout D, Menotti A, Aravanis C, et al. Serum Total Cholesterol and Long-Term Coronary Heart Disease Mortality in Different Cultures. Twenty-Five-Year Follow-up of the Seven Countries Study. *JAMA*. 1995; 274:131–6. [PubMed: 7596000]
34. Wang ET, Kao CN, Shinkai K, Pasch L, Cedars MI, Huddleston HG. Phenotypic Comparison of Caucasian and Asian Women with Polycystic Ovary Syndrome: A Cross-Sectional Study. *Fertil Steril*. 2013; 100:214–8. [PubMed: 23557763]





**Figure 1.** Principal Components 1 and 2 in PCOS and Control Subjects Compared to Subjects in Datasets of Known Ethnicity

A) Individual PCOS subjects (Case, red circles) and controls (blue circles) plotted as a function principal components 1 and 2. B) Data from women with PCOS (Case, red open circles) and controls (blue open circles) plotted as a function of principal components 1 and 2 and superimposed on data from populations of known ethnicity. The PCOS cases cluster statistically closer to the Tuscan and Italian populations, whereas the controls cluster closer to the French and Orcadian populations ( $p < 0.002$ ). C) Map indicating location of the populations of known ethnicity.

**Table 1**

Association of principal components with PCOS status and phenotypic variables.

Phenotype	Cohort	Principal Component	BETA± SE	p-value	<i>I</i> BETA± SE <sub>BMI</sub>	p-value <sub>BMI</sub>
PCOS	<sup>2</sup> combined	PC1	9.30±2.29	0.000049	8.04±2.59	0.0019
Waist:Hip Ratio	combined	PC1	<sup>3</sup> NS	NS	-0.24±0.086	0.0048
Acanthosis Nigrans	cases	PC1	9.02±3.00	0.0027	8.23±3.36	0.014
LDL Cholesterol Levels	controls	PC2	-1.34±0.43	0.0020	-1.39±0.43	0.0013
Cholesterol Levels	controls	PC2	-0.76±0.27	0.0049	-0.78±0.27	0.0040
FSH Levels	cases	PC3	11.40±3.26	0.00054	11.40±3.26	0.00051
17-OH progesterone Levels	cases	PC4	-10.88±3.73	0.0038	-10.77±3.64	0.0033

<sup>1</sup> Controlled for body mass index (BMI)

<sup>2</sup> Cases = subjects with PCOS, Controls = subjects without PCOS in the control group, Combined = both PCOS and controls together

<sup>3</sup> NS = not significant