The Polycystic Ovary Syndrome Evolutionary Paradox: a Genome-Wide Association Studies–Based, in silico, Evolutionary Explanation

Livio Casarini and Giulia Brigante

Unit of Endocrinology, Department of Biomedical, Metabolic, and Neural Sciences (L.C., G.B.), and Center for Genomic Research (L.C.), University of Modena and Reggio Emilia, 41121 Modena, Italy

Context: Polycystic ovary syndrome (PCOS) is a common female endocrine disorder characterized by phenotypes ranging from hyperandrogenism to metabolic disorders, more prevalent in people of African/Caucasian and Asian ancestry. Because PCOS impairs fertility without diminishing in prevalence, it was considered an evolutionary paradox. Genome-Wide Association Studies identified 17 single nucleotide polymorphisms (SNPs) associated with PCOS, with different allele frequencies, ethnicity-related, in 11 susceptibility loci.

Objective: In this study we analyze the PCOS phenotype-genotype relationship in silico, using SNPs of representative genes for analysis of genetic clustering and distance, to evaluate the degree of genetic similarity.

Data Source: 1000 Genomes, HapMap, and Human Genome Diversity Project databases were used as source of allele frequencies of the SNPs, using data from male and female individuals grouped according to their geographical ancestry.

Setting and Design: Genetic clustering was calculated from SNPs data by Bayesian inference. The inferred ancestry of individuals was matched with PCOS phenotype data, extracted from a previous meta-analysis. The measure of genetic distance was plotted against the geographic distance between the populations.

Results: The individuals were assigned to five genetic clusters, matching with different world regions (Kruskal-Wallis/Dunn's post test; $P < .0001$), and converging in two main PCOS phenotypes in different degrees of affinity. The overall genetic distance increased with the geographic distance among the populations (linear regression; $R^2 = 0.21$; $P < .0001$), in a phenotype-unrelated manner.

Conclusions: Phenotype-genotype correlations were demonstrated, suggesting that PCOS genetic gradient results from genetic drift due to a serial founder effect occurred during ancient human migrations. The overall prevalence of the disease supports intralocus sexual conflict as alternative to the natural selection of phenotypic traits in females. **(*J Clin Endocrinol Metab* 99: E2412–E2420, 2014)**

Polycystic ovary syndrome (PCOS) is the most common endocrinopathy affecting 5–10% of women in reproductive age worldwide. It is a familial, polygenic condition associated with infertility, irregular menstrual cycles, anovulation, hyperandrogenism, as well as nonreproductive health problems depending on genetic background and affected by lifestyle (1–3).

PCOS phenotypic features

Even if the disease displays a wide variety of characteristics, it is widely accepted that PCOS features decrease within two main phenotypes. According to the 2003 Rotterdam criteria (4), the prevalent clinical symptoms define the hyperandrogenic or metabolic phenotype (5–10). The hyperandrogenic PCOS phenotype is defined mainly by

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in U.S.A.

Received June 19, 2014. Accepted July 29, 2014.
First Published Online August 5, 2014

Abbreviations: GWAS, Genome-Wide Association Studies; PCOS, polycystic ovary syndrome; SNP, single nucleotide polymorphis.

hirsutism, androgenic alopecia, and relatively high androgen levels, whereas the metabolic phenotype is characterized by metabolic syndrome, insulin resistance, increased risk for type 2 diabetes, and high body mass index or central obesity (5–10). It can be approximately established that the metabolic phenotype is prevalent in Central Asians and Americans whereas the hyperandrogenic phenotype is prevalent in the other world regions. However, the wide spectrum of secondary disorders results in an overlap of PCOS characteristics among human populations, reflecting the polygenic condition of the disease and genetic admixture. Curiously, no clear differences in the prevalence of the disease among different ethnic groups has been identified so far (10–15).

Evolutionary origin of PCOS

Given that the disease impairs fertility without diminishing its high global prevalence, it was extensively discussed as an evolutionary paradox. Previous studies attempted to explain how a genetic pattern linked to metabolic or reproductive disadvantages spread across continents, generating a dozen theories suggesting different explanations for the evolutionary origin of PCOS (16). Most these hypotheses prompt a balancing mechanism between viability selection and metabolic thrift against fertility disadvantages associated with this condition. For example, androgenisation and insulin resistance may confer survival benefit to females and improve the glucose availability for ovulatory functions in hunter-gatherer societies (17). Moreover, metabolic thrift and increased fat storage are advantages for mother and fetus under low food conditions (18). However, the effect on the individual fitness of the PCOS phenotype during the evolution of humans is not understood, and no evolutionary advantage for the PCOS genotype carriers has been proven. Surprisingly, all previous theories about PCOS consider evolutionary dynamics involving only females, not considering the contribution of the male in the genotype-phenotype inheritance and evolution. All the genetic and evolutionary analyses of PCOS were carried out on a sample of female individuals, presumably resulting in biased evaluations and in an overall loss of genetic information.

Genetic markers of the disease

Previous studies identified a hundred candidate genes associated with PCOS and several genetic markers affecting the pathogenesis, phenotype, and prevalence of the disease have been proposed (19). Recently, two Genome-Wide Association Studies (GWAS) performed in Han Chinese women identified 11 new risk loci for PCOS (20, 21), which count 17 single nucleotide polymorphisms (SNPs) leading to genetic variants strongly associated with the disease. The gene sequences located within the PCOS susceptibility loci are involved in the ovarian response to the gonadotropic hormones, in the metabolism of glucose and lipids, and in cell cycle regulation. This finding is corroborated by other studies showing the association between these markers and the disease (22–27). The results of these two statistically powerful analyses were confirmed by other works performed in populations of non-Chinese ancestry (19, 28, 29), showing a common genetic risk profile across human populations.

All the genes falling within the susceptibility loci identified by GWAS may potentially be implicated in the modulation of the PCOS phenotype and its severity. These genes are *FSHR*, *LHCGR*, *DENND1A*, *THADA*, *C9orf3*, *YAP1*, *HMGA2*, *RAB5B/SUOX*, *INSR*, *TOX3* and *SUMO1P1*; their potential relation with PCOS was described separately (Supplemental Discussion).

Geography of PCOS genetic markers

The distribution of the allelic variants associated to PCOS is different among the populations worldwide, as observable by Human Genome Diversity Project (HGDP) selection browser (http://hgdp.uchicago.edu/cgi-bin/gbrowse/HGDP) (30–32), a web-based software, which calculates the geographic distributions of user-selected markers from Stanford SNP genotyping data (33, 34). A different genetic pattern distribution of PCOS markers could be reflected in different phenotypic features of the disease, resulting from adaptive evolution (19, 28, 35) or from genetic drift generated by a serial founder effect occurring during the ancient human migrations out of Africa (36). Accordingly, the decay of expected heterozygosity as measure of genetic variation accompanies the increase of genetic and geographic distance from Africa (33, 34, 37). But the overall constant prevalence of the disease remains unexplained.

The determination of the genetic background and its relationship with the phenotype may be relevant to optimize the pharmacological treatment of the disease and protocols for assisted reproduction. To define the degree of similarity of the PCOS genotypes, we show a population genetics analysis by Bayesian clustering and an evaluation of pairwise genetic distance using SNPs data from different populations, available in online databases. The genotype-phenotype link and the correlation between genetic and geographic data are discussed from an evolutionary point of view.

Materials and Methods

A detailed description of the methods is available as Supplemental Materials and Methods.

SNP selection

The genetic analyses were performed using the frequencies data of the 17 PCOS-related SNPs of individuals from human populations sampled in different world regions. The SNP panel (Supplemental Materials and Methods) was taken from two GWAS (20, 21), which found a strong association between PCOS and the 17 genetic markers. The SNP data were obtained from the HGDP-CEPH (Centre d'Etude du Polymorphisme Humain) Stanford (33, 34), 1000 Genomes (The 1000 Genomes Project Consortium, 2010) (38) and in part from the HapMap (The International HapMap Consortium, 2003) (39) panels, which provide SNP frequencies and geographical coordinates of a wide number of human populations worldwide, to ensure a good coverage in the territorial distribution among the continents. Based on their geographic coordinates, the populations were grouped by continent (Supplemental Table 1). All the genetic data are from both male and female individuals unselected for PCOS, ensuring that the analysis takes into account males as carrier of a PCOS-linked genotype and avoiding the bias arising from the use of only PCOS patients.

Selection of PCOS phenotypes

In order to evaluate the link between genotype and ethnicity, the geographical distribution of the different PCOS phenotypes was evaluated by analyzing the clinical data registered in the scientific literature (Supplemental Materials and Methods). Phenotypic data are shown in a world map (Figure 1).

Human genetic clustering analysis

Genetic clustering analysis assigns the individuals to the group (cluster) that best represents their genetic background, calculated using the frequencies of PCOS markers. To this end, SNPs data from individuals were used and the genetic population stratification was inferred by the Bayesian analysis implemented in STRUCTURE 2.3.4 software (Pritchard Lab) (40). The number of subpopulations (K) in which the individuals were assigned was selected using the ΔK method (41) (Supplemental Figure 1). The degree of affinity of the populations to the resulting genetic clusters is expressed as a numeric value (Q value) by the software. Thus, Q values define an estimation of ancestry, inferred by the SNP frequencies. Then, Q values were grouped for world area, and used for a graphical representation together with geographical and clinical data.

The genotype-phenotype link was obtained from the analysis of geographical data and genetic clustering. It was confirmed by principal component analysis implemented in National Institute on Aging (NIA) Array Analysis software (42) using the SNP frequency data.

Evaluation of the genetic drift

To evaluate the contribution of the genetic drift in the establishment of the modern PCOS markers distribution, a linear regression of the expected/observed heterozygosity and genetic against geographic distance was performed (36). The SNPs panel was used to obtain the heterozygosity data and to calculate the

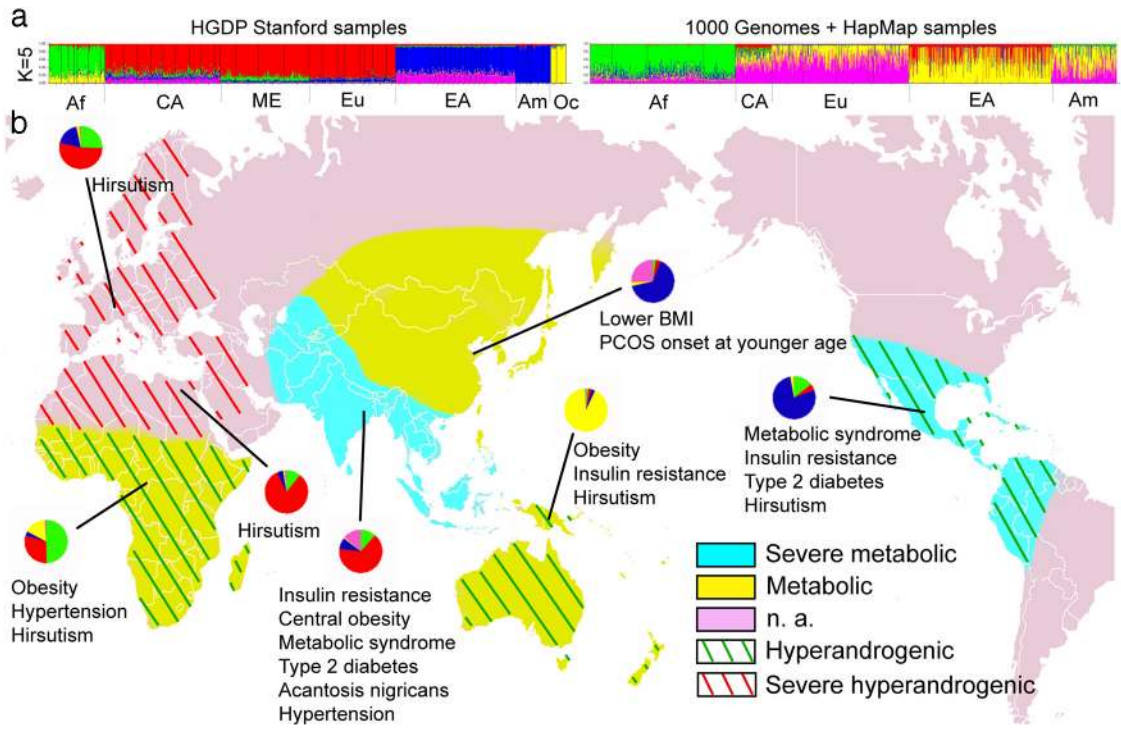

Figure 1. World distribution of the affinity to the genetic clusters and PCOS phenotypes prevalence. A, Bar plots of individual Q values calculated by the STRUCTURE software assign each individual to different subpopulations that matches the main world areas, with a certain degree of admixture. The analysis was performed differentially for the HGDP and the 1000 Genomes merged together with the HapMap samples. Each color indicates the membership of individuals in a genetic cluster (K = 5); Af, African; CA, Central Asian; ME, Mediterranean/Middle Eastern; Eu, European; Am, American; Oc, Oceanian. B, Pie chart of cluster affinity among continents, indicating the frequencies of the PCOS susceptibility markers. The charts were obtained as the means of the Q values by merging the HGDP, 1000 Genomes, and HapMap populations for each genetic cluster (colors of the pie charts do not refer to panel A). The overall prevalence of a genetic cluster is different between the geographic area, suggesting a link with the corresponding PCOS phenotype and clinic features, which were obtained by a review of the literature; green, cluster 1; red, cluster 2; blue, cluster3; yellow, cluster 4; magenta, cluster 5; n.a., data not available or not assessed.

fixation index (Fst) as a measure of pairwise genetic distance (43) resulting by comparing each world population vs Africans. A scatter plot of Fst and heterozygosity against the geographic distance was illustrated by the waypoints-method previously described (36, 44). Briefly, the geographic distance between populations was calculated using Addis Ababa in Ethiopia as the starting point and taking into account the kilometers covered by humans during their expansion worldwide through migratory waypoints. The waypoints were chosen simulating the migratory routes (Supplemental Table 2), assuming that humans bypassed the natural obstacles during migrations, such as oceans and high mountains. The radius of the Earth was also considered for the calculation of the geographic distance.

Statistical analysis and image software

ANOVA, *t* tests, or linear regression analysis was applied as appropriate and indicated in the figure legends. The statistical analysis was performed by GraphPad Prism software (GraphPad).

Results

Geographical distribution of PCOS genotypes and phenotypes

Bayesian genetic clustering analysis assigned the individuals to five subpopulations (K = 5) obtained from the estimation of the proportion of ancestry inferred by the SNP frequencies. It can be defined by the Q value as the degree of affinity of the population to each genetic cluster by the SNPs combination of the individuals within the indicated geographic area. Thus, the degree of prevalence of each cluster is variable among the world continents, revealing that human populations could be divided into five groups with geographically different, non-homogeneous genetic background calculated using the frequency of PCOS markers, though a degree of admixture exists (Figure 1A). The proportion of affinity of the populations to each genetic cluster is also illustrated for each world area, which is represented by the prevalent PCOS genotype and phenotype (Figures 1B and 2). The metabolic phenotype characterized by insulin resistance, metabolic syndrome, type 2 diabetes, hypertension, and acanthosis nigricans is dominant in Asia, especially Central Asia, and in America. The hyperandrogenic phenotype predominates in European, Mediterranean, and Middle Eastern patients, who show the most severe hirsutism. Considering African and Oceanian PCOS–affected women, the phenotype is characterized by both hirsutism and insulin resistance with a predominant role of the latter feature. Indeed, a certain degree of mixed PCOS phenotypes co-

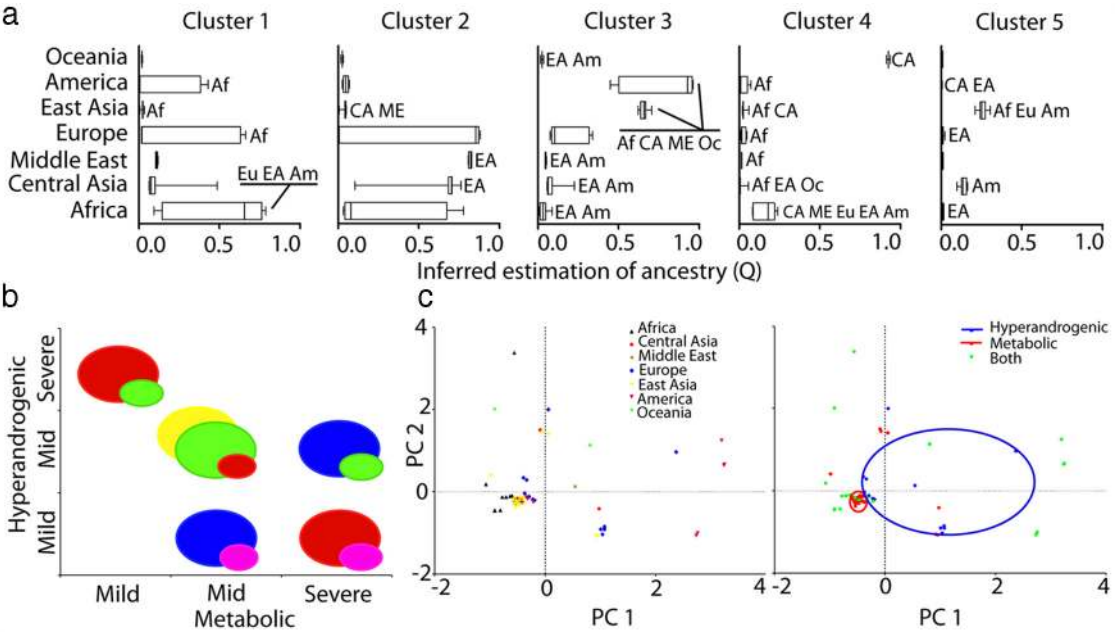

Figure 2. Box and whiskers plot of the overall continental membership in each genetic cluster. The continents are represented by the distribution of the Q values calculated by the STRUCTURE software. The charts, A, were obtained as the means of the Q values by merging the HGDP, 1000 Genomes, and HapMap populations for each genetic cluster. Each cluster is peculiar for a specific geographic area because of the nonhomogeneous distribution of PCOS susceptibility marker frequencies. The acronym above the bars indicates a significant difference vs Africa, Af; America, Am; Central Asia, CA; Europe, Eu; East Asia, EA; Middle East/Mediterranean area, ME; and Oceania, Oc. Kruskal-Wallis and Dunn's post-test ($P < .0001$). B, Relationship between genetic background and PCOS phenotype. The genetic clusters are represented by colored ovals and located in the position of the corresponding phenotype (Figure 1). Prevalence of each cluster in each world area is proportional to the area of the oval, calculated using Q values; green, cluster 1; red, cluster 2; blue, cluster3; yellow, cluster 4; magenta, cluster 5. C, Clustering by principal component (PC) analysis performed using SNP genotypes in relation to the metabolic and hyperandrogenic PCOS phenotypes (cumulative percentage of data set coverage from PC1 and PC2 = 54.187%). The populations are represented by points colored depending on their geographical origin (left panel) or prevalent phenotype (right panel).

exists in each world population, as a result of genetic admixture (Figure 1).

Genotype-phenotype link

The analysis of the geographical distribution of the merged Q values reveals that the affinity to a predominant, peculiar genetic cluster is typical for each continent (Figure 2). Thus, the genetic background inferred by PCOS markers is different and nonhomogeneous among the human populations from diverse world areas (Figure 2A). Cluster 1 is predominant in Africans. Cluster 2 is typical in Central Asian, European, and Mediterranean/Middle Eastern people, but with a lesser extent of membership to the cluster 1, revealing that these populations share a similar genetic background, calculated using PCOS markers. However, Central Asians share a certain degree of membership also to Cluster 5 together with East Asians. Cluster 3 is typical of Americans and Eastern Asians. Cluster 4 is typical of Oceanians and partly of Africans. The degree of severity of the two main PCOS phenotypes results from the contribution of the genetic background. Each PCOS genotype is linked to its geographic area with a peculiar combination of prevalent PCOS features. This is shown by cluster memberships, calculated as the Q value of the prevalent clusters in each geographic area (Figure 2B). Indeed, cluster 1 is well represented in European, Mediterranean/Middle Eastern, African, and American people, where hyperandrogenic hirsutism is mid/severe. Independently of the PCOS features, cluster 2 is prevalent where the phenotype is more severe. Clusters 3 and 5 are represented in association with the metabolic phenotype, which mainly differs for the overall severity and type of its features. African and Oceanian share a similar phenotype (cluster 4) characterized by mid degrees of metabolic risk and hyperandrogenism, suggesting a genetic similarity for PCOS markers between these populations, according to the result of genetic clustering at K = 4 (HGDP samples, Supplemental Figure 1). The link between PCOS genotype and phenotype was confirmed by principal component analysis (Figure 2C), given that the population characterized by the metabolic or hyperandrogenic phenotype shares peculiar graph areas. However, all clusters are represented in each geographic area (Figures 1 and 2), probably contributing to the PCOS features and its severity (Figure 2).

Genetic diversity resulting from PCOS markers

The measure of the genetic distance, Fst, calculated for the panel of PCOS susceptibility markers reveals a strong diversity in the genetic background among human populations, grouped by continents (Figure 3). The pairwise Fst of the non-Africans vs Africans increases together with the geographic distance from Africa, producing $R^2 = 0.21$

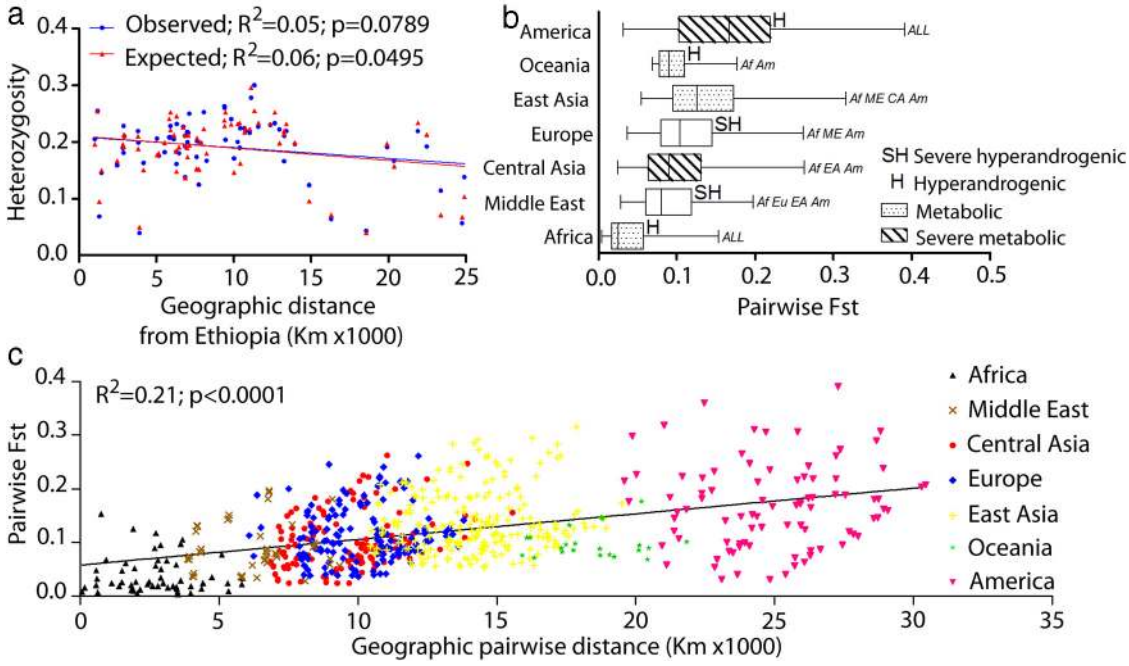

Figure 3. Analysis of the genetic distance (Fst) and decay of heterozygosity between continents. The charts were obtained by merging the data from HGDP, 1000 Genomes, and HapMap populations. A, Scatterplot of heterozygosity and geographic distance. The expected heterozygosity calculated for PCOS susceptibility loci decreases together with geographic distance, evaluated by linear regression. B, Box and whiskers plot of Fst distribution grouped for continents. The acronym above the bars indicates a significant difference vs Africa, Af; America, Am; Central Asia, CA; Europe, Eu; East Asia, EA; Middle East/Mediterranean area, ME; and all other continents, ALL. Kruskal-Wallis and Dunn's post-test ($P < .0001$). C, Scatterplot of Fst and geographic distance calculated using the waypoints. Each point indicates the comparison between African vs other populations, as specified in the figure legend. Fst calculated for PCOS susceptibility loci increases together with geographic distance, evaluated by linear regression.

($P < .0001$) calculated by linear regression (Figure 3C) and indicating the strong contribution of genetic drift in the establishment of PCOS markers distribution. The decay of the expected heterozygosity (Figure 3A) strengthens this observation. The result is corroborated by the increase of Fst and decay of heterozygosity, calculated for a wider range of genetic markers in a previous study, when plotted against the distance from the putative starting point of human migrations, in Ethiopia (36, 37). Differently from what expected, the observed heterozygosity does not decay with distance, suggesting a relatively recent genetic admixture. Moreover, the Fst distributions in Africans vs other populations (Figure 3A) is different among continents, indicating strongly diverse, nonhomogeneous allele frequencies in the distribution of PCOS markers worldwide (Figure 3).

Discussion

Previous stratification analyses using a wide number of genetic markers showed that the modern human population comprises six main genetic clusters depending on the ethnic background, reflecting with surprising accuracy the ethnicity and admixture degree of ancestry (45). Using the PCOS markers human population was stratified into five different genetic clusters falling within two main PCOS phenotypic groups. Thus, PCOS results in a hyperandrogenic and in a metabolic phenotype, reflecting the world distribution of the degree of affinity to genetic clusters. This analysis provides evidence that PCOS ethnic variations are strongly determined by the genetic background in humans (27, 34, 46–48), as already demonstrated by a comparative experiment between different PCOS mouse strains (49). Genetic cluster analysis relies on the simultaneous combination of different SNPs, providing a higher level of accuracy than case-control studies, which, in fact, yielded conflicting results (50–53), not resulting in any clear cause-effect indication related to ethnicity and providing a hundred putative markers not independently confirmed (19, 51). The study of ethnic variations of PCOS genotype-phenotype link may be a useful approach for the pharmacological treatment of the disease and during infertility treatment.

PCOS phenotypes

The severity of the PCOS phenotypes may result from different combinations of SNPs represented by the clusters (Figure 1 and Figure 2), eg, Americans and Asians with a prevalence of the metabolic phenotype, belonging to different prevalent clusters (3 and 2, respectively; Figures 1B and 2). Conversely, European and Mediterranean/Middle Eastern people share the hirsute-hyperandrogenic phenotype and a high affinity to cluster 2, but also the affinity to clusters 1 and 4 is high among populations with a mid hyperandrogenic phenotype, although characterized by metabolic features rather than hirsutism. Thus, the ancient humans' migratory routes do not completely reflect the current distribution of the PCOS phenotypes. Additional considerations regarding the distribution of the genotypes and phenotypes are available separately (Supplemental Discussion).

PCOS and genetic diversity among humans

Although other alternatives were proposed and discussed (54, 55), Fst remains a widely used measure of genetic distance. The pairwise Fst calculated for the various populations vs Africans increases together with the geographic distance from Africa. This result is consistent with previous data showing that genetic diversity is determined by a serial founder effect that occurred during the ancient human migrations across the continents (36). The increase of Fst with the geographic distance calculated for PCOS markers in humans is the same of that previously observed for a wider set of genetic markers from different organisms (36, 56–59), suggesting that its distribution is the result of the random genetic drift. Nevertheless, the variations of PCOS phenotypes among continents is displayed through various characteristics and symptoms apparently incoherent with the geographic distance (Figure 3), probably as a result of genetic admixture rather than natural selection. In fact, each world population shares a different degree of affinity with the five genetic clusters, and therefore it may differently contribute to phenotype determination. Limitations due to genetic admixture were discussed separately (Supplemental Discussion).

All things considered, the overall continental variability of the two main PCOS phenotypes is clearly linked to a peculiar genetic background, resulting from genetic drift and indicating that different genetic markers may reflect convergent phenotypic features.

Natural selection or genetic drift?

These data suggest that PCOS genotype and phenotype may be not strongly affected by natural selection during human evolution. Nevertheless, the reason why the prevalence of PCOS is similar among the different world continents remains to be demonstrated. A large number of evolution-based theories were produced and extensively discussed (16), providing rational evaluations for the evolution of different PCOS phenotypes in females, especially the metabolic phenotype, but not for the overall constant prevalence of the disease. Surprisingly, none of these theories considers the male as the carrier of an hyperandro-

genic trait, which may reasonably provide an improvement of his individual fitness, likely emphasizing the male secondary sex characteristics. Indeed, both sexes share most of their genomes and express the same traits, however resulting in antagonistic selection (60, 61). From this point of view, the prevalence of PCOS among different human populations may be the result of the balance between a positive selection in males against a negative selection in females. Our results are strengthened by the demonstration that natural selection favors in similar proportions both protective and risk alleles for type 2 diabetes (62), suggesting that the phenotype linked to the disease results in opposite effects on the fitness of both sexes. Previous observations in other mammals support this hypothesis, since the selection for the sex hormone testosterone leads to antagonistic reproductive fitness between parents and their opposite-sex progeny (63). On the other hand, it is well known that evolution is often sex-dependent in different species (64–67), given that alleles can have positive effects on fitness in one sex and negative in the other, resulting in intralocus sexual conflict (61, 68). Different genetic background for PCOS converging in two main phenotypes together with overall constant prevalence of the disease support the presence of intralocus sexual conflict, which may have affected the decay of observed heterozygosity. Even if speculative, in humans this mechanism seems to be a "bug" inherited from admixture with different hominids or from an ancient genome evolved in environmental and social conditions, strongly different from those in which *Homo sapiens* lived in the last 100 000 years (69, 70), but every hypothesis in this regard must be demonstrated.

Summary

The phenotypic expression of PCOS varies among human populations, depending on ethnicity. The distribution of previously identified susceptibility disease markers results in different, nonhomogeneous continental genetic backgrounds by Bayesian clustering, reflecting the ethnic distribution of the main PCOS phenotypes. Thus, a clear indication for PCOS ethnicity is shown, taking into account a certain degree of genetic admixture between human populations. The genetic distance increases together with the distance from Africa, suggesting that the modern distribution of PCOS susceptibility markers is the result of genetic drift likely due to a serial founder effect occurred during the ancient human migrations as alternative to the natural selection theory. Intralocus sexual conflict may contribute to the maintenance of an overall constant prevalence of PCOS measured in females. The analysis of the genetic background may lead to important implications for the pharmacological approach to the disease.

Acknowledgments

We thank Professor Manuela Simoni for her commitment, support, and guidance in the field of endocrinology.

Address all correspondence and requests for reprints to: Livio Casarini, PhD, Unit of Endocrinology, Nuovo Ospedale Civile Sant'Agostino Estense (NOCSAE), Via P. Giardini 1355, 41126 Modena, Italy. E-mail: livio.casarini@unimore.it.

This work was supported by a grant of the Italian Ministry of Education, University and Research, No. PRIN 2010C8ERKX.

Disclosure Summary: The authors have nothing to disclose.

References

1. **Domecq JP, Prutsky G, Mullan RJ, et al.** Lifestyle modification programs in polycystic ovary syndrome: Systematic review and meta-analysis. *J Clin Endocrinol Metab*. 2013;98:4655–4663.
2. **Legro RS, Arslanian SA, Ehrmann DA, et al.** Diagnosis and treatment of polycystic ovary syndrome: An Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*. 2013;98:4565–4592.
3. **Moran LJ, Ranasinha S, Zoungas S, McNaughton SA, Brown WJ, Teede HJ.** The contribution of diet, physical activity and sedentary behaviour to body mass index in women with and without polycystic ovary syndrome. *Hum Reprod*. 2013;28:2276–2283.
4. **Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group.** Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod*. 2004;19:41–47.
5. **Kauffman RP, Baker TE, Baker VM, DiMarino P, Castracane VD.** Endocrine and metabolic differences among phenotypic expressions of polycystic ovary syndrome according to the 2003 Rotterdam consensus criteria. *Am J Obstet Gynecol*. 2008;198:670.e1–7.
6. **Mojiminiyi OA, Safar FH, Al Rumaih H, Diejomaoh M.** Variations in alanine aminotransferase levels within the normal range predict metabolic and androgenic phenotypes in women of reproductive age. *Scand J Clin Lab Invest*. 2010;70:554–560.
7. **Wijeyaratne CN, Seneviratne Rde A, Dahanayake S, et al.** Phenotype and metabolic profile of South Asian women with polycystic ovary syndrome (PCOS): Results of a large database from a specialist Endocrine Clinic. *Hum Reprod*. 2011;26:202–213.
8. **Dunaif A, Fauser BC.** Renaming PCOS—a two-state solution. *J Clin Endocrinol Metab*. 2013;98:4325–4328.
9. **Zhang HY, Guo CX, Zhu FF, Qu PP, Lin WJ, Xiong J.** Clinical characteristics, metabolic features, and phenotype of Chinese women with polycystic ovary syndrome: A large-scale case-control study. *Arch Gynecol Obstet*. 2013;287:525–531.
10. **Zhao Y, Qiao J.** Ethnic differences in the phenotypic expression of polycystic ovary syndrome. *Steroids*. 2013;78:755–760.
11. **Diamanti-Kandarakis E, Kouli CR, Bergiele AT, et al.** A survey of the polycystic ovary syndrome in the Greek island of Lesbos: Hormonal and metabolic profile. *J Clin Endocrinol Metab*. 1999;84:4006–4011.
12. **Asunción M, Calvo RM, San Millán 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;85:2434–2438.
13. **Kumarapeli V, Seneviratne Rde A, Wijeyaratne CN, Yapa RM, Dodampahala SH.** A simple screening approach for assessing community prevalence and phenotype of polycystic ovary syndrome in a semi-urban population in Sri Lanka. *Am J Epidemiol*. 2008;168:321–328.
14. **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.
15. **Wang S, Alvero R.** Racial and ethnic differences in physiology and clinical symptoms of polycystic ovary syndrome. *Semin Reprod Med*. 2013;31:365–369.
16. **Corbett S, Morin-Papunen L.** The polycystic ovary syndrome and recent human evolution. *Mol Cell Endocrinol*. 2013;373:39–50.
17. **Azziz R, Dumesic DA, Goodarzi MO.** Polycystic ovary syndrome: An ancient disorder? *Fertil Steril*. 2011;95:1544–1548.
18. **Holte J, Bergh T, Berne C, Berglund L, Lithell H.** Enhanced early insulin response to glucose in relation to insulin resistance in women with polycystic ovary syndrome and normal glucose tolerance. *J Clin Endocrinol Metab*. 1994;79:1052–1058.
19. **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.
20. **Chen ZJ, Zhao H, He L, et al.** Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. *Nat Genet*. 2011;43:55–59.
21. **Shi Y, Zhao H, Shi Y, et al.** Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. *Nat Genet*. 2012;44:1020–1025.
22. **Simoni M, Tempfer CB, Destenaves B, Fauser BC.** Functional genetic polymorphisms and female reproductive disorders: Part I: Polycystic ovary syndrome and ovarian response. *Hum Reprod Update*. 2008;14:459–484.
23. **Unsal T, Konac E, Yesilkaya E, et al.** Genetic polymorphisms of *FSHR*, *CYP17*, *CYP1A1*, *CAPN10*, *INSR*, *SERPINE1* genes in adolescent girls with polycystic ovary syndrome. *J Assist Reprod Genet*. 2009;26:205–216.
24. **Gu BH, Park JM, Baek KH.** Genetic variations of follicle stimulating hormone receptor are associated with polycystic ovary syndrome. *Int J Mol Med*. 2010;26:107–112.
25. **Eriksen MB, Brusgaard K, Andersen 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.
26. **Goodarzi MO, Jones MR, Li X, et al.** Replication of association of DENND1A and THADA variants with polycystic ovary syndrome in European cohorts. *J Med Genet*. 2012;49:90–95.
27. **Li T, Zhao H, Zhao X, et al.** Identification of *YAP1* as a novel susceptibility gene for polycystic ovary syndrome. *J Med Genet*. 2012;49:254–257.
28. **Welt CK, Styrkarsdottir U, Ehrmann DA, et al.** Variants in *DENND1A* are associated with polycystic ovary syndrome in women of European ancestry. *J Clin Endocrinol Metab*. 2012;97:E1342–1347.
29. **Mutharasan P, Galdones E, Peñalver Bernabé B, et al.** Evidence for chromosome 2p16.3 polycystic ovary syndrome susceptibility locus in affected women of European ancestry. *J Clin Endocrinol Metab*. 2013;98:E185–190.
30. **Coop G, Pickrell JK, Novembre J, et al.** The role of geography in human adaptation. *PLoS Genet*. 2009; 5:e1000500.
31. **Pickrell JK, Coop G, Novembre J, et al.** Signals of recent positive selection in a worldwide sample of human populations. *Genome Research*. 2009;19:826–837.
32. **Pritchard JK, Pickrell JK, Coop G.** The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. *Curr Biol*. 2010;20:R208–215.
33. **Cann HM, de Toma C, Cazes L, et al.** A human genome diversity cell line panel. *Science*. 2002;296:261–262.
34. **Cavalli-Sforza LL.** The human genome diversity project: Past, present and future. *Nat Rev Genet*. 2005;6:333–340.
35. **Cui L, Zhao H, Zhang B, et al.** Genotype-phenotype correlations of PCOS susceptibility SNPs identified by GWAS in a large cohort of Han Chinese women. *Hum Reprod*. 2013;28:538–544.
36. **Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, Cavalli-Sforza LL.** Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proc Natl Acad Sci U S A*. 2005;102:15942–15947.
37. **Li JZ, Absher DM, Tang H, et al.** Worldwide human relationships inferred from genome-wide patterns of variation. *Science*. 2008;319:1100–1104.
38. **Abecasis GR, Altshuler D, Auton A, et al.** A map of human genome variation from population-scale sequencing. *Nature*. 2010;467:1061–1073.
39. **The International HapMap Consortium.** The International HapMap Project. *Nature*. 2003;426:789–796.
40. **Pritchard JK, Stephens M, Donnelly P.** Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–959.
41. **Evanno G, Regnaut S, Goudet J.** Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14:2611–2620.
42. **Sharov AA, Dudekula DB, Ko MS.** A web-based tool for principal component and significance analysis of microarray data. *Bioinformatics*. 2005;21:2548–2549.
43. **Wright S.** Isolation by distance. *Genetics*. 1943;28:114–138.
44. **Henn BM, Cavalli-Sforza LL, Feldman MW.** The great human expansion. *Proc Natl Acad Sci U S A*. 2012;109:17758–17764.
45. **Rosenberg NA, Pritchard JK, Weber JL, et al.** Genetic structure of human populations. *Science*. 2002;298:2381–2385.
46. **Valkenburg O, Uitterlinden AG, Piersma D, et al.** Genetic polymorphisms of GnRH and gonadotrophic hormone receptors affect the phenotype of polycystic ovary syndrome. *Hum Reprod*. 2009;24:2014–2022.
47. **Hwang JY, Lee EJ, Jin Go M, et al.** Genome-wide association study identifies GYS2 as a novel genetic factor for polycystic ovary syndrome through obesity-related condition. *J Hum Genet*. 2012;57:660–664.
48. **Zadeh-Vakili A, Ramezani Tehrani F, Daneshpour MS, Zarkesh M, Saadat N, Azizi F.** Genetic polymorphism of vitamin D receptor gene affects the phenotype of PCOS. *Gene*. 2013;515:193–196.
49. **Dowling AR, Nedorezov LB, Qiu X, Marino JS, Hill JW.** Genetic factors modulate the impact of pubertal androgen excess on insulin sensitivity and fertility. *PLoS One*. 2013;8:e79849.
50. **Fu L, Zhang Z, Zhang A, et al.** Association study between FSHR Ala307Thr and Ser680Asn variants and polycystic ovary syndrome (PCOS) in Northern Chinese Han women. *J Assist Reprod Genet*. 2013;30:717–721.
51. **Pau C, Saxena R, Welt CK.** Evaluating reported candidate gene associations with polycystic ovary syndrome. *Fertil Steril*. 2013;99:1774–1778.
52. **Skrgati L, Baldani DP, Gersak K, Cerne JZ, Ferk P, Cori M.** Genetic polymorphisms of INS, INSR and IRS-1 genes are not associated with polycystic ovary syndrome in Croatian women. *Coll Antropol*. 2013;37:141–146.
53. **Xu P, Shen SM, Zhang XL, et al.** Haplotype analysis of single nucleotide polymorphisms in anti-Müllerian hormone gene in Chinese PCOS women. *Arch Gynecol Obstet*. 2013;288:125–130.
54. **Nei M.** Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci U S A*. 1973;70:3321–3323.
55. **Excoffier L, Smouse PE, Quattro JM.** Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*. 1992;131:479–491.
56. **Linz B, Balloux F, Moodley Y, et al.** An African origin for the intimate association between humans and Helicobacter pylori. *Nature*. 2007;445:915–918.
57. **Stewart JF, Liu Y, Tauer CG, Nelson CD.** Microsatellite versus AFLP analyses of pre-management introgression levels in loblolly pine (*Pinus taeda L.*) and shortleaf pine. (*P. echinata* Mill.). *Tree Genetics, Genomes*. 2010;6:853–862.
58. **Legrand D, Vautrin D, Lachaise D, Cariou ML.** Microsatellite variation suggests a recent fine-scale population structure of *Drosophila*

sechellia, a species endemic of the Seychelles archipelago. *Genetica*. 2011;139:909–919.
59. **Taylor SM, Antonia AL, Parobek CM, et al.** Plasmodium falciparum sulfadoxine resistance is geographically and genetically clustered within the DR Congo. *Sci Rep*. 2013;3:1165.
60. **Cox RM, Calsbeek R.** Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *Am Nat*. 2009;173:176–187.
61. **Pennell TM, Morrow EH.** Two sexes, one genome: The evolutionary dynamics of intralocus sexual conflict. *Ecol Evol*. 2013;3:1819–1834.
62. **Ayub Q, Moutsianas L, Chen Y, et al.** Revisiting the thrifty gene hypothesis via 65 loci associated with susceptibility to type 2 diabetes. *Am J Hum Genet*. 2014;94:176–185.
63. **Mills SC, Koskela E, Mappes T.** Intralocus sexual conflict for fitness: Sexually antagonistic alleles for testosterone. *Proc Biol Sci*. 2012; 279:1889–1895.
64. **Mainguy J, Côté SD, Festa-Bianchet M, Coltman DW.** Father-offspring phenotypic correlations suggest intralocus sexual conflict for a fitness-linked trait in a wild sexually dimorphic mammal. *Proc Biol Sci*. 2009;276:4067–4075.
65. **Delph LF, Andicoechea J, Steven JC, Herlihy CR, Scarpino SV, Bell DL.** Environment-dependent intralocus sexual conflict in a dioecious plant. *New Phytol*. 2011;192:542–552.
66. **Lewis Z, Wedell N, Hunt J.** Evidence for strong intralocus sexual conflict in the Indian meal moth, *Plodia interpunctella*. *Evolution*. 2011;65:2085–2097.
67. **Postma E, Spyrou N, Rollins LA, Brooks RC.** Sex-dependent selection differentially shapes genetic variation on and off the guppy Y chromosome. *Evolution*. 2011;65:2145–2156.
68. **Stearns SC, Govindaraju DR, Ewbank D, Byars SG.** Constraints on the coevolution of contemporary human males and females. *Proc Biol Sci*. 2012;279:4836–4844.
69. **Vasseur E, Quintana-Murci L.** The impact of natural selection on health and disease: Uses of the population genetics approach in humans. *Evol Appl*. 2013;6:596–607.
70. **Wang S, Lachance J, Tishkoff SA, Hey J, Xing J.** Apparent Variation in Neanderthal Admixture among African Populations is Consistent with Gene Flow from non-African Populations. *Genome Biol Evol*. 2013;5:2075–2081.