PORTABILITY OF GWAS RESULTS BETWEEN ETHNIC POPULATIONS: GENETIC MARKERS FOR POLYCYSTIC OVARY SYNDROME (PCOS) IN MEDITERRANEAN AREA

Z. Douma¹, C. Lautier², S. Haydar², T. Mahjoub¹, F. Grigorescu^{2,3,*}

¹University of Monastir, Faculty of Pharmacy of Monastir - Laboratory of Human Genome and Multifactorial Diseases (LR12ES07), Monastir, Tunisia, ²University of Montpellier - Molecular Endocrinology - IURC Montpellier, ³Collège de France - Institut Convergences Migrations, Paris, France

Abstract

Genome Wide Association Studies (GWAS) are excellent opportunities to define culprit genes in complex disorders such as the polycystic ovary syndrome (PCOS). PCOS is a prevalent disorder characterized by anovulation, hyperandrogenism and polycystic ovaries, which benefitted from several GWASs in Asians and Europeans revealing more than 20 potential culprit genes near associated single nucleotide variations (SNV). Translation of these findings into the clinical practice raises difficulties since positive hits are surrogate SNVs linked with causative mutations by linkage disequilibrium (LD). Studies in Mediterranean populations (e.g. Southern Europe and North Africa) raise supplementary problems because of a different LD-pattern, which may disrupt the link with causative mutations. Our experience in MEDIGENE program between Tunisia and France enforces the necessity of genetic anthropology studies before translating GWAS data. Tunisians are a heterogeneous population with ancestral Berbers, European, Arab and Sub-Saharan African components while South Europeans display a high level of genetic diversity, partially explained by gene flow from North Africa. Human diversity studies require sampling from Middle East and North Africa (MENA) region that will help to understand genetic factors in complex diseases.

Key words: Genome Wide Association Studies, polycystic ovary syndrome, population genetics, Maghreb, admixture, evolution, single nucleotide variations.

INTRODUCTION

One of the major goals of the MEDIGENE European project (FP7-279171) was to study genetic markers for insulin resistance in the context of the genetic landscape of Mediterranean populations and considering both native (autochthonous) and migrant populations in Europe. Beside the metabolic syndrome (MetS) as a typical example of insulin resistance,

the polycystic ovary syndrome (PCOS) in women also appeared as an excellent model. The interest in PCOS is explained by the strong association of this syndrome with metabolic abnormalities, including insulin resistance, type-2 diabetes (T2D) and MetS (1). Tempting to translate genetic discoveries into the clinical practice, practitioners are watching for potential use of Genome Wide Association Studies (GWAS) data for diagnosis of PCOS. Among other complex disorders, PCOS benefitted in the last years from several GWAS on thousands of individuals of Asian or European ancestry (2-8). These studies revealed more than 20 genes, all with potential applications. One usual way to validate these discoveries is to replicate results in ethnic groups. The underlying hypothesis is that, once data are replicated, it would exist some universal significance of single nucleotide variations (SNVs), and thus, more chances to be used as clinical tools. Although many aspects of translation of GWAS discoveries deserve discussion, one aspect that we would like to stress here is the portability of GWAS data on ethnic populations. Indeed, major discoveries were obtained in Chinese and more recently in Europeans, but portability of results to Mediterranean populations, including North Africa raises difficulties (2, 5). This is explained by dissimilarity in the linkage disequilibrium (LD) patterns in North African populations. LD may be eroded during generations by recombination or in the case of a heterogeneous population, by admixture of people of different ancestries and both can affect the link between surrogate SNVs (used in GWAS) and causative variants. In this article we intend to reveal some aspects of the ethnic diversity among neighbor populations in the Mediterranean area taking as example the variability of PCOS in France (Europe) and Tunisia (North Africa), two geographical regions in the focus of the MEDIGENE program (<https://cordis.europa>.

*Correspondence to: Florin Grigorescu PhD, IURC, Molecular - Endocrinology, 641 Av Du Doyen Gaston Giraud, Montpellier, 34093, France, E-mail: florin.grigorescu@inserm.fr

eu/result/rcn/195753_fr.html). While epidemiological data indicated some variability in the PCOS phenotype, we would like to show how anthropological knowledge on North Africa might influence genetics of Europeans, aspects in population genetics to which practitioners are less accustomed.

Variability of PCOS phenotype in populations

PCOS is a common endocrine disorder in women at reproductive age (9-11). Prevalence of PCOS is evaluated between 3-12% in populations, with regional [geographic] and ethnic differences (1). Evaluation of PCOS prevalence was hampered by its variable phenotype, diagnosis criteria and ethnic variability. Cardinal features of PCOS are anovulation (AO), hyperandrogenism (HA) and polycystic ovaries at ultrasound scan (1, 9, 10). PCOS is accompanied by various degrees of obesity, dyslipidemia and intrinsic insulin resistance, which expose the affected women to the risk of developing T2D, MetS and cardiovascular (CV) complications later in life (1). Despite these clear statements, PCOS definition varied with time. In 1990, the definition formulated by NIH considered PCOS in the presence of 2 criteria: HA and chronic anovulation (11). Patients with Cushing's syndrome, non-classical adrenal 21-hydroxylase deficiency, hyperprolactinaemia and androgen-secreting neoplasms were excluded. The definition was revised in 2003 by Rotterdam ESHRE/ASRM - Sponsored PCOS Consensus Workshop Group (commonly referred as Rotterdam criteria), which considered PCOS in the presence of 2 out of 3 criteria: oligo-ovulation or AO, presence of polycystic ovaries at ultrasound scan, and clinical and/or biochemical signs of HA (9). In 2006, Androgen Excess-PCOS society revised Rotterdam definition excluding women with PCO and AO in the absence of HA (12). In 2018, the most recent NIH guidelines considered PCOS in the presence of 2 of 3 cardinal criteria, but encouraged the definition of PCOS sub-phenotypes (13). Obviously these variable definitions influenced the estimation of prevalence rates of PCOS as recently reviewed (12). In general, Rotterdam definition yielded higher prevalence values in all populations. The study of Okoroh *et al.*, 2012, reported the prevalence of 1.6% using all 3 criteria and 18% using the Rotterdam definition (14). Diagnosis criteria influence also the evaluation of PCOS in lean *versus* obese women, since prevalence rates from 5% to 15% were reported in lean and obese women, respectively (15). Ethnic differences were also reported in the clinical manifestation of the syndrome

(10). Therefore, different modified Ferriman-Gallwey scales (mFG) and cut-off values for testosterone levels should be used in different countries. The same study of Okoroh *et al.*, 2012 reported variation of PCOS prevalence in various regions of US. The values were as high as 47.5% in southern regions but only 10.3% in North Eastern regions (14). Although some studies identified higher prevalence in African Americans, others did not find major differences, at least in southern regions (around 4%) (reviewed in Ref 10). Mexican Americans, by contrast, display almost 13% prevalence of PCOS accompanied by more severe HA, obesity and MetS. In this ethnic group, women with insulin resistance and T2D showed a 2-fold increase in prevalence of PCOS (16). In Caribbean Hispanics PCOS was found at similar prevalence to Caucasians or Africans (3.4 and 4.7%), but the affected women display more severe insulin resistance and a 2-fold increased risk for T2D (17). In Europe, in Mediterranean populations the prevalence rates were reported in Greece (6.8%) and Spain (6.5%) (18). In Asians, the prevalence of PCOS was estimated around 6.3% and 5.6% in Chinese women and they tend to be less hirsute (19). Clinical manifestations of PCOS appear to vary as function of geographical regions. Thus, by exploring distribution of allele associated with PCOS from GWAS in the HGDP database (<http://hgdp.uchicago.edu/cgi-bin/gbrowse/HGDP/>), Casarini *et al.*, 2014 found that metabolic abnormalities, including insulin resistance and Acanthosis Nigricans (AN) or MetS were dominant in Central Asia and America while hyperandrogenic PCOS (hirsutism) would predominate in Europe, Mediterranean regions and Middle East. Africa would be dominated by obesity, high blood pressure (HBP) and hirsutism and Indian sub-continent by central obesity, insulin resistance, T2D, AN and HBP (20). One of the most striking data was obtained in immigrant Pakistani women in London who display 52% PCOS (21). The same population displays 26% Acanthosis Nigricans, 61% consanguinity and 61% antecedents for T2D in the 1st degree relatives (21). We also note that infertility rates (at age between 20 to 44) in Pakistan were reported at 2.7%, much higher than in Europeans (22).

In the MEDIGENE program we proposed to study PCOS in Mediterranean populations, such as Southern France (Languedoc-Roussillon region) and North Africa (Tunisia). The Middle East and North Africa (MENA) geographic region is important because many of the emigrants in Europe are coming from this region. Studies on the prevalence of PCOS in North

Africa are scarce, particularly as randomized studies in the general population. We can obtain some information from recent studies providing data on infertility rate in women between 20 and 44 years in the MENA region (29). Infertility varies from 2.4% in Morocco, 1.9% in Algeria, 2.3% in Tunisia, 2.3% in Egypt and 2.1% in Saudi Arabia (22). These values are much higher than in other countries in Europe (1.1% in France and 0.9 in Spain). The highest rate (upper confidence interval) was reported in Iran (3.6%) and Saudi Arabia (3.7%). Therefore, it appears that it would be a West-East gradient of infertility rate in MENA region. Compared to indigenous population of Central Europe, a study in Austria (23) reported that immigrants from MENA region displayed a higher degree of obesity (20.5 *versus* 7.2%), and more polycystic ovaries (6.4 *versus* 3.5%). Although infertility is a complex condition and due to multiple factors (e.g. male infertility and in women genital infections), higher prevalence in MENA region raises interest for North Africa. In Palestinians, a cross-sectional study of university students (137 females) of 18-24 years detected a PCOS prevalence rate of 7.3% (24). A proportion of 70% of women displayed hirsutism while acne (7.3%) was a major risk factor with OR of 8.4. In Iran, among women of 18-45 years, there was 7.1% PCOS by NIH criteria and 14.6% by Rotterdam criteria (25). In El-Minia region in Egypt, PCO diagnosis was done in 37.5% of women with primary or secondary infertility and 14.0% in fertile women with an intrauterine contraceptive device (26). All these epidemiological studies should be placed in the context of anthropological diversity of populations, which present without doubt great genetic variation. Finally, in evolutionary perspective, PCOS was considered as a very ancient condition and exposed as a paradox. It is expected that the genetic susceptibility for PCOS would represent some evolutionary advantages for survival in human populations, hypothesis similar to that of thrifty genotype supposed for occurrence of obesity and diabetes. It is worth indicating that a simple examination of frequency of surrogate SNVs in ethnic populations is insufficient for drawing conclusions since SNVs described in GWAS (contrary to Mendelian mutations) are surrogate SNVs and being bi-allelic are non-informative. The most interesting conclusion from these studies for an anthropologist is that prevalence of PCOS might be related in fact to that of obesity and T2D in a given population. Indeed several studies in Sri Lanka and Pima Indians in US suggested this aspect (38) and this might be the case of the MENA region.

Genes from GWAS of PCOS

As previously indicated, several GWAS were performed in PCOS from women of Asian and European ancestry. The first study was performed in Han Chinese on 744 cases and 895 controls and then replicated in 2,840 cases and 5,012 controls (2). It should be indicated that *stricto sensu* replication means application of the same methods of genotyping in another sample of the same ethnic cohort (internal validation). However, the extended term is often used as replication in another ethnic group (external validation). In Han Chinese, among 29 positive hits in the 1st stage of GWAS, 28 were replicated on the larger samples. The paper concluded on the potential implication of 24 genes, among which only 7 were new, including LHCGR, GTF2A1L, FSHR, ZFP36L2, THADA, LOC100129726 and DENND1A. The remaining genes were already studied in PCOS by the candidate gene approach, including insulin receptor (INSR), insulin receptor substrate 2 (IRS-2) and adiponectin (ADIPOQ). A second study was performed in the same Chinese population on 8,226 cases and 7,578 controls, which reported hits in several new loci: 9q22.32, 11q22.1, 12q13.2, 12q14.3, 16q12.1, 19p13.3, 20q13.2 (3). An independent signal was found at Chr2p16.3, near the follicle stimulating hormone receptor (FSHR) gene. Lee *et al.*, 2015, reported a GWAS in Korean population on 976 PCOS cases and 946 controls with replication on 249 PCOS cases and 778 controls (4). The strongest association was found at the locus on Chr8q24.2. The nearest gene was KH RNA binding domain containing, signal transduction associated 3 (KHDRBS3) related with telomerase activity. Two novel loci on Chr 8p23.1 and Chr11p14.1 were found by Hayes *et al.*, 2015 in population of European ancestry (5). In this study the locus at Chr9q22.32 was confirmed (DENND1A) and also revealed the importance of the positive association of SNV rs11031006 close to FSHB (follicle stimulating hormone subunit beta) gene and correlated to the levels of luteinizing hormone (LH). A more recent and larger study in Europeans by Day *et al.*, in 2015 revealed in 5,184 self-reported cases of PCOS and 82,759 controls statistical significance for ERBB4/HER4, YAP1, THADA, FSHB, RAD50 and KRR1 genes (5). Some association was suggested for epidermal growth factor receptor ERBB2-4 genes. The investigation also performed a Mendelian Randomization (MR) study indicating causal role of these variants on BMI (body mass index), insulin resistance and sex hormone binding globulin (SHBG) levels. Interestingly, the study of Day *et al.* also indicated a relation with later

menopause and higher anti-Müllerian (AMH) hormone levels. Taken together these previous studies identified 16 loci for PCOS. In 2019 a meta-analysis on 10,074 PCOS cases and 103,164 controls of European ancestry replicated 11 loci and found 3 novel loci near *PLGRKT*, *ZBTB16* and *MAPRE1* genes (7). MR studies also detected a relation with depression and male-pattern balding in PCOS (7). The genes near to the most robust hits in PCOS previously identified by GWASs are indicated in Table 1. A series of studies replicated or reviewed these genetic findings in PCOS, leading to the conclusion that indeed there would be loci robustly confirmed in ethnic population and with biological importance. This is the case, for instance, for *LHCGR*, *FSHR* and *FSHB* genes related to hormonal alterations in PCOS or that of *INSR* and *HMGA2* involved in metabolic feature of PCOS (27). A recent study reviewed these loci together with those found by narrative gene candidate approaches (28) and Table 2 indicated some of the most significant SNVs identified by GWAS in PCOS with clinical implications.

PCOS studies in North Africa

It is believed that replication (external validation) of association in ethnic groups would represent some universal feature and thus a potential clinical application. Indeed, with some exceptions, these SNVs used in GWAS and available on gene chips are common variants and found in almost all

populations. It should be stressed that SNVs detected in GWAS even at high frequency, remained surrogate markers linked with causative mutations by linkage disequilibrium (LD). These effective mutations can be located near the association signals or far away from SNVs detected in GWAS. Their potential use as genetic markers is pending on the LD in various ethnic populations.

In the MEDIGENE program we were interested in investigating candidate genes from GWAS in Tunisian population from North Africa in comparison with populations of Europe, as for instance Southern France (29-33). We paid special attention in recruitment of controls for PCOS patients by selection from the same geographical regions of Tunisia or France. Several genes that entered in our focus, including oestrogen receptor *ESR1* and *ESR2*, *LHCGR*, *FSHR* and *FSHB* genes in relation with hormonal profile. Genotyping was performed Affymetrix gene chip (MEDISCOPE) containing 750,000 SNVs over the entire genome (34-36). Preliminary Principal Component Analysis (PCA) indicated that French population clustered with that of CEU (Utah residents with northern and western European ancestry) population but differently from Tunisians, located somehow between standard CEU and YRI Africans (Yoruba from Ibadan, Nigeria). Moreover, several outliers were observed in population from Southern France, likely representing immigrant population from North Africa and settled in France. As

Table 1. Genes near SNV signals of GWAS in PCOS populations

Site#	Gene name	Locus
1	<i>LHCGR</i> (luteinizing hormone/choriogonadotropin receptor)	Chr 2p16.3
2	<i>THADA</i> (THADA armadillo repeat containing)	Chr 2p21
3	<i>FSHR</i> (follicle stimulating hormone receptor)	Chr 2p16.3
4	<i>ERBB4</i> (erb-b2 receptor tyrosine kinase 4)	Chr 2q34
5	<i>RAD50</i> (RAD50 double strand break repair protein)	Chr 5q31.1
6	<i>GATA4</i> (GATA binding protein 4)	Chr 8p23.1
7	<i>PLGRKT</i> (plasminogen receptor with a C-terminal lysine)	Chr 9p24.1
8	<i>KHDRBS3</i> (KH RNA binding domain containing, signal transduction associated 3)	Chr 8q24.23
9	<i>DENND1A</i> (DENN domain containing 1A)	Chr9q33.3
10	<i>C9orf3</i> or <i>AOPEP</i> (aminopeptidase O - putative)	Chr 9q22.32
11	<i>ZBTB16</i> (zinc finger and BTB domain containing 16)	Chr 11q23.2
12	<i>ARL14EP</i> (ADP ribosylation factor like GTPase 14 effector protein)	Chr 11p14.1
13	<i>FSHB</i> (follicle stimulating hormone subunit beta)	Chr 11p14.1
14	<i>YAP1</i> (YES associated protein 1)	Chr 11q22.1
15	<i>RAB5B</i> (<i>RAB5B</i> , member <i>RAS</i> oncogene family)	Chr 12q13.2
16	<i>HMGA2</i> (high mobility group AT-hook 2)	Chr 12q14.3
17	<i>GYS2</i> (glycogen synthase 2)	Chr 12p12.1
18	<i>KRR1</i> (<i>KRR1</i> small subunit processome component homolog)	Chr 12q21.2
19	<i>TOX3</i> (<i>TOX</i> high mobility group box family member 3)	Chr 16q12.1
20	<i>INSR</i> (insulin receptor)	Chr 19p13.2
21	<i>SUMO1P1</i> (<i>SUMO1</i> pseudogene 1)	Chr 20q13.2
22	<i>MAPRE1</i> (microtubule associated protein RP/EB family member 1)	Chr 20q11.21

Table 2. The most significant SNVs obtained in GWAS and meta-analysis for PCOS

SNV ID	Alleles	Afr	CEU	Asians	Gene	Location
rs2268361	C/T	0.25	0.63	0.47	FSHR	Chr 2p16.3
rs2349415	T/C	0.52	0.68	0.78	FSHR	Chr 2p16.3
rs6166	C/T	0.58	0.55	0.69	FSHR	Chr 2p16.3
rs13405728	A/G	0.32	0.07	0.29	LHCGR	Chr 2p16.3
rs12468394	C/A	0.46	0.53	0.31	THADA	Chr 2p21
rs13429458	A/C	0.20	0.13	0.24	THADA	Chr 2p21
rs12478601	C/T	0.82	0.59	0.32	THADA	Chr 2p21
rs7563201	G/A	0.31	0.51	0.28	THADA	Chr 2p21
rs804279	A/T	0.61	0.71	0.79	GATA4/NEIL2	Chr 8p23.1
rs10818854	G/A	0.07	0.04	0.05	DENND1A	Chr9q33.3
rs2479106	A/G	0.8	0.3	0.19	DENND1A	Chr9q33.3
rs10986105	T/G	0.13	0.04	0.07	DENND1A	Chr9q33.3
rs9696009	G/A	0.46	0.06	0.06	DENND1A	Chr9q33.3
rs11031006	G/A	0.03	0.14	0.03	FSHB	Chr 11p14.1
rs11031005	T/C	0.01	0.14	0.03	FSHB	Chr 11p14.1
rs2059807	A/G	0.82	0.6	0.29	INSR	Chr 19p13.2

Table 3. Minor Allele Frequencies (MAF) of 8 SNVs in LHCGR gene in European (CEU), French, Tunisian and African populations

SNV ID	Alleles	Africans	Tunisians	Europeans	French
rs10865234	C/T	0.56	0.85	0.93	0.87
rs11691408	C/T	0.30	0.23	0.31	0.30
rs1404052	C/T	0.70	0.88	0.93	0.97
rs1404051	A/C	0.25	0.37	0.62	0.59
rs34790224	C/T	0.10	0.13	0.22	0.18
rs17398156	C/G	0.03	0.07	0.09	0.11
rs11892581	A/G	0.48	0.25	0.31	0.31
rs6746530	G/T	0.47	0.25	0.31	0.31
rs12478427	C/A	0.02	0.03	0.04	0.05

Figure 1. Linkage disequilibrium (LD) pattern in French, Tunisian and African populations of the selected region of LHCGR gene. A, French; B, Tunisian; C, African YRI population. Data for YRI African population was obtained from HapMap (<http://hapmap.ncbi.nlm.nih.gov>). Corresponding SNVs with their MAF in various populations are indicated in Table 3.

expected, our studies revealed that French population presents a stronger LD-pattern than Tunisians. A typical example is indicated in Figure 1 on the LHCGR gene between positions 48910113 and 48913019 (GRCh37/hg19, by February 2009). In the region selected for this example, there are 8 SNVs which were depicted in HAPLOVIEW program in French, Tunisian and African populations, while their minor allele frequencies (MAF) are indicated in Table 3.

The LD-pattern suggests that LD in Tunisians was eroded during generations more than in Europeans or there would be some admixture with Sub-Saharan Africans. These aspects complicate the interpretation of replication of results from GWAS and prompt investigators to search more attentively on genetic structure of ethnic groups. A different LD pattern may be explained in several ways either supposing an admixture of Tunisian population with populations from Sub-Saharan Africa or simply by the historical evolution of Tunisian populations. Genetic studies in complex disorders imply a prerequisite knowledge of the anthropological structure of ethnic populations and we believe that anthropological analysis would have consequences on both directions from Europe to North

Africa and *vice versa*. Studies performed in Tunisian population indicated multiple influences and gene flows. Tunisians are inhabitants of Tunisia, a country of Maghreb and Northwest Africa, who speak a common language (Tunisian or Derja) and present a common culture and identity (37). Tunisia is bordered by Libya in the South, Algeria in the west and the Mediterranean Sea in the North. There are 24 governorates in the country and the capital Tunis is the most populated (37). Tunisian population is a mosaic of people of different origins being in majority descendants of ancestral Berbers. Tunisians are speaking Tunisian Arabic language but some minorities are still speaking Berber language. Berber people were admixed with other ethnic groups. According to archeological evidences modern humans established in North Africa 160,000 years ago. Aterian culture dating 40,000 years ago was followed by settlement of people from the Near-East. In historical times, admixture occurred during Phoenician, Roman and Vandal invasions, particularly in the region of Carthage. Arab-Berbers represent around 98% of the population while a fraction of 2% is represented by other ethnic groups like Europeans and Jewish (37). Berbers are autochthonous inhabitants of Tunisia and considered as descendants of Iberomaurusian and Capsian culture originating 5000 y BC. The genetic structure of the population was studied by Alu/STR markers, mtDNA and Chr Y lineages and more recently by autosomal SNVs (38, 39). After Arab invasion, Berber population migrated in the South and took refuge in the mountains, explaining poor admixture with Arabs. Thus, Chr Y uniparental markers in Berbers indicated the presence of E1b1b1b lineage, although some groups were admixed with Arabic populations as shown by haplogroup J1e in high proportion of Berbers living in plains. Other studies on HLA class I and class II genes indicated common features with West Mediterranean populations, concordant with some studies with mtDNA, indicating a close relation with Iberians (38, 39). Genetic studies indicated less genetic distance with European compared to other North African populations. Tunisians also admixed with Sub-Saharan Africans. Interestingly, North-Tunisia contains more Berber and Sub-Saharan African components than Southern Tunisia, explained by the main source of African flow as trans-Saharan trade.

One interesting aspect of genetics in Mediterranean population is the possible gene flow from North Africa into Europeans. As shown by Botigué *et al.*, 2013, higher genetic diversity in Southern Europe compared to the North might be explained by several

hypotheses (40). The first supposed the recolonization of Europe from glacial refugia 20,000 years ago while the second proposed the demic diffusion from the East during introduction of agriculture. Another hypothesis proposes to explain genetic diversity by migration of population from North Africa. Using autosomal SNVs markers, the proportion of sub-Saharan gene flow was estimated around 1-3% in Southern Europeans (40). Some short haplotypes are shared with Yoruba population. The study used SNVs by ADMIXTURE program and showed that North African ancestry in Europe ranged between 5 and 14%. More detailed studies on gene flow suggested that migration from North Africa occurred 230-300 years ago involving 5-7 generations. This raises the question whether migration from North Africa affected the pattern of allele associated with disease risk. Based on GWAS database it was found that for 134 diseases, the majority of alleles reflect an expected pattern of neutral divergence except for multiple sclerosis (MS) which shows a significant divergence from random drift for the pathogenic alleles. Such studies were not performed for PCOS related alleles. For many reasons, PCOS merits to be studied from anthropological point of view and this on both sides of the Mediterranean Sea. In our view (41), these studies may reveal new aspects of the portability of GWAS data on ethnic populations around the Mediterranean area and point to the urgent need to increase the number of genetic data in international programs.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Diamanti-Kandarakis E, Dunaif A. Insulin resistance and the polycystic ovary syndrome revisited: an update on mechanisms and implications. *Endocr Rev.* 2012; 33(6):981-1030.
2. Chen ZJ, Zhao H, He L, Shi Y, Qin Y, Shi Y, Li Z, You L, Zhao J, Liu J, Liang X, Zhao X, Zhao J, Sun Y, Zhang B, Jiang H, Zhao D, Bian Y, Gao X, Geng L, Li Y, Zhu D, Sun X, Xu JE, Hao C, Ren CE, Zhang Y, Chen S, Zhang W, Yang A, Yan J, Li Y, Ma J, Zhao Y. Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. *Nat Genet.* 2011; 43(1):55-59.
3. Shi Y, Zhao H, Shi Y, Cao Y, Yang D, Li Z, Zhang B, Liang X, Li T, Chen J, Shen J, Zhao J, You L, Gao X, Zhu D, Zhao X, Yan Y, Qin Y, Li W, Yan J, Wang Q, Zhao J, Geng L, Ma J, Zhao Y, He G, Zhang A, Zou S, Yang A, Liu J, Li W, Li B, Wan C, Qin Y, Shi J, Yang J, Jiang H, Xu JE, Qi X, Sun Y, Zhang Y, Hao C, Ju X, Zhao D, Ren CE, Li X, Zhang W, Zhang Y, Zhang J, Wu D, Zhang C, He L, Chen ZJ. Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. *Nat Genet.* 2012; 44(9):1020-1025.

4. Lee H, Oh JY, Sung YA, Chung H, Kim HL, Kim GS, Cho YS, Kim JT. Genome-wide association study identified new susceptibility loci for polycystic ovary syndrome. *Hum Reprod*. 2015; 30(3):723-731.
5. Hayes MG, Urbanek M, Ehrmann DA, Armstrong LL, Lee JY, Sisk R, Karaderi T, Barber TM, McCarthy MI, Franks S, Lindgren CM, Welt CK, Diamanti-Kandarakis E, Panidis D, Goodarzi MO, Azziz R, Zhang Y, James RG, Olivier M, Kissebah AH; Reproductive Medicine Network, Stener-Victorin E, Legro RS, Dunaif A. Genome-wide association of polycystic ovary syndrome implicates alterations in gonadotropin secretion in European ancestry populations. *Nat Commun*. 2015; 518: 197-206.
6. Day FR, Hinds DA, Tung JY, Stolk L, Styrkarsdottir U, Saxena R, Bjornnes A, Broer L, Dunger DB, Halldorsson BV, Lawlor DA Laval G, Mathieson I, McCardle WL, Louwers Y, Meun C, Ring S, Scott RA, Sulem P, Uitterlinden AG, Wareham NJ, Thorsteinsdottir U, Welt C, Stefansson K, Laven JSE, Ong KK, Perry JRB. Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. *Nat Commun*. 2015; 6:8464.
7. Day F, Karaderi T, Jones MR, Meun C, He C, Drong A, Kraft P, Lin N, Huang H, Broer L, Magi R, Saxena R, Laik T, Urbanek M, Hayes MG, Thorleifsson G, Fernandez-Tajes J, Mahajan A, Mullin BH, Stuckey BGA, Spector TD, Wilson SG, Goodarzi MO, Davis L, Obermayer-Pietsch B, Uitterlinden AG, Anttila V, Neale BM, Jarvelin MR, Fauser B, Kowalska I, Visser JA, Andersen M, Ong K, Stener-Victorin E, Ehrmann D, Legro RS, Salumets A, McCarthy MI, Morin-Papunen L, Thorsteinsdottir U, Stefansson K; 23andMe Research Team, Styrkarsdottir U, Perry JRB, Dunaif A, Laven J, Franks S, Lindgren CM, Welt CK. Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. *PLoS Genet*. 2018; 14(12):e1007813.
8. Kim S-H, Liu M, Jin HS, Park S. High Genetic Risk Scores of ASIC2, MACROD2, CHRM3, and C2orf83 Genetic Variants Associated with Polycystic Ovary Syndrome Impair Insulin Sensitivity and Interact with Energy Intake in Korean Women. *Gynecol Obstet Invest*. 2019; 84(3):225-236.
9. Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group (2004) Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril*. 2004; 81:19-25.
10. Wolf WM, Wattick RA, Kinkade ON, Olfert MD. Geographical Prevalence of Polycystic Ovary Syndrome as Determined by Region and Race/Ethnicity. *Int J Environ Res Public Health*. 2018; 15(11).
11. Zawadzki JK, Dunaif A. Diagnostic criteria for polycystic ovary syndrome: Towards a rational approach. In *Polycystic Ovary Syndrome*; Dunaif A, Givens JR, Haseltine FP, Merriam GE, Eds; Blackwell Scientific Inc: Boston, MA, USA. 1992; pp. 377-384.
12. Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE, Witchel SF; Task Force on the Phenotype of the Polycystic Ovary Syndrome of The Androgen Excess and PCOS Society. The Androgen Excess and PCOS Society criteria for the polycystic ovary syndrome: the complete task force report. *Fertil Steril*. 2009; 91(2):456-488.
13. Johnson T, Kaplan L, Ouyang P, Rizza R. National Institutes of Health Evidence-Based Methodology Workshop on Polycystic Ovary syndrome; National Institutes of Health: Bethesda, MD, USA. 2012; pp. 1-14.
14. Okoroh EM, Hooper WC, Atrash HK, Yusuf HR, Boulet SL. Prevalence of polycystic ovary syndrome among the privately insured, United States, 2003-2008. *Am J Obstet Gynecol*. 2012; 207(4):299.e1-7.
15. Skiba MA, Islam RM, Bell RJ, Davis SR. Understanding variation in prevalence estimates of polycystic ovary syndrome: a systematic review and meta-analysis. *Hum Reprod Update*. 2018; 24(6):694-709.
16. Goodarzi MO, Quiñones MJ, Azziz R, Rotter JI, Hsueh WA, Yang H. Polycystic ovary syndrome in Mexican-Americans: prevalence and association with the severity of insulin resistance. *Fertil Steril*. 2005; 84(3):766-769.
17. Dunaif A, Sorbara L, Delson R, Green G. Ethnicity and polycystic ovary syndrome are associated with independent and additive decreases in insulin action in Caribbean-Hispanic women. *Diabetes*. 1993; 42(10):1462-1468.
18. 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(7):2434-2438.
19. Ding T, Hardiman PJ, Petersen I, Wang F-F, Qu F, Baio G. The prevalence of polycystic ovary syndrome in reproductive-aged women of different ethnicity: a systematic review and meta-analysis. *Oncotarget*. 2017; 8(56):96351-96358.
20. 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(11): E2412-2420.
21. Wijeyaratne CN, Dilini Udayangani SA, Balen AH. Ethnic-specific polycystic ovary syndrome: epidemiology, significance and implications. *Expert Rev Endocrinol Metab*. 2013; 8(1):71-79.
22. Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. *PLoS Med*. 2012; 9(12):e1001356.
23. Feichtinger M, Göbl C, Weghofer A, Feichtinger W. Reproductive outcome in European and Middle Eastern/North African patients. *Reprod Biomed Online*. 2016; 33(6):684-689.
24. Musmar S, Afaneh A, Mo'alla H. Epidemiology of polycystic ovary syndrome: a cross sectional study of university students at An-Najah national university-Palestine. *Reprod Biol Endocrinol*. 2013; 11:47.
25. Tehrani FR, Simbar M, Tohidi M, Hosseiniapanah F, Azizi F. The prevalence of polycystic ovary syndrome in a community sample of Iranian population: Iranian PCOS prevalence study. *Reprod Biol Endocrinol*. 2011; 9:39.
26. Sanad AS. Prevalence of polycystic ovary syndrome among fertile and infertile women in Minia Governorate, Egypt. *Int J Gynaecol Obstet*. 2014; 125(1):81-82.
27. Zhao H, Lv Y, Li L, Chen Z-J. Genetic Studies on Polycystic Ovary Syndrome. *Best Pract Res Clin Obstet Gynaecol*. 2016; 37:56-65.
28. Hiam D, Moreno-Asso A, Teede HJ, Laven JSE, Stepto NK, Moran LJ, Gibson-Helm M. The Genetics of Polycystic Ovary Syndrome: An Overview of Candidate Gene Systematic Reviews and Genome-Wide Association Studies. *J Clin Med*. 2019; 8(10).
29. Dallel M, Sarray S, Douma Z, Hachani F, Al-Ansari AK, Letaifa DB, Mahjoub T, Almawi WY. Differential association of DENND1A genetic variants with polycystic ovary syndrome in Tunisian but not Bahraini Arab women. *Gene*. 2018; 647:79-84.
30. Dallel M, Sghaier I, Finan RR, Douma Z, Hachani F, Letaifa DB, Mahjoub T, Almawi WY. Circulating leptin concentration, LEP gene variants and haplotypes, and polycystic ovary syndrome in Bahraini and Tunisian Arab women. *Gene*. 2019; 694:19-25.
31. Ben Salem A, Attaoua R, Mtraoui N, Belkahl S, Ezzidi I, Ajina M, Souissi M, Pouchet P, Vintila M, Grigorescu F, Mahjoub T. Common polymorphisms of calpain-10 and the risk of polycystic ovary syndrome in Tunisian population: a case-control study. *Mol Biol Rep*. 2014; 41(10):6569-6574.
32. Ben Salem A, Megdich F, Kacem O, Souayeh M, Hachani Ben Ali F, Hizem S, Janhai F, Ajina M, Abu-Elmagd M, Assidi M, Al Qahtani MH, Mahjoub T. Vascular endothelial growth factor (VEGFA) gene variation in polycystic ovary syndrome in a Tunisian women population. *BMC Genomics*. 2016; 17(Suppl 9).
33. Grigorescu F. New Genetic Approaches in Understanding Susceptibility for Metabolic Syndrome in Immigrant Populations

- Around Mediterranean Area. *Acta Endocrinol (Buchar)*. 2012; 8(1):87-98.
34. Čoklo M, Auguštin DH, Šarac J, Novokmet N, Sindik J, Lewis AP, Petranovic MZ, Mulahasanović LK, Khusnutdinova E, Litvinov S, Haydar S, Lautier C, Normand C, Attaoua R, Vintila M, Bosch-Comas A, Suarez H, Jares P, Gomis R, Missoni S, Marjanović D, Grigorescu F. Diversity of Y-chromosomal and mtDNA Markers Included in Mediscope Chip within Two Albanian Subpopulations from Croatia and Kosovo: Preliminary Data. *Coll Antropol*. 2016; 40(3):195-198. PMID: 29139639
35. Haydar S, Grigorescu F, Vintilă M, Cogne Y, Lautier C, Tutuncu Y, Brun JF, Robine JM, Pugeat M, Normand C, Poucheret P, Gheorghiu ML, Georgescu C, Badiu C, Băculescu N, Renard E, Ylli D, Badiou S, Sutra T, Cristol JP, Mercier J, Gomis R, Macias JM, Litvinov S, Khusnutdinova E, Poiana C, Pasquali R, Lauro D, Sesti G, Prudente S, Trischitta V, Tsatsoulis A, Abdelhak S, Barakat A, Zenati A, Ylli A, Satman I, Kanninen T, Rinato Y, Missoni S. Fine-scale haplotype mapping of MUT, AACS, SLC6A15 and PRKCA genes indicates association with insulin resistance of metabolic syndrome and relationship with branched chain amino acid metabolism or regulation. *PLoS ONE*. 2019; 14(3):e0214122.
36. Haydar S, Lautier C, Grigorescu F. Branched chain amino acids at the edge between mendelian and complex disorders. *Acta Endocrinol (Buchar)*. 2018; 14(2):238-247.
37. Elloumi-Zghal H, Chaabouni Bouhamed H. Genetics and genomic medicine in Tunisia. *Mol Genet Genomic Med*. 2018; 6(2): 134-159.
38. Cherni L, Pereira L, Goios A, Loueslati BY, Khodjet el Khil H, Gomes I, Gusmão L, Alves C, Slama A, Amorim A, Elgaaied AB. Y-chromosomal STR haplotypes in three ethnic groups and one cosmopolitan population from Tunisia. *Forensic Sci Int*. 2005; 152(1):95-99.
39. Kefi R, Hsouna S, Ben Halim N, Lasram K, Romdhane L, Messai H, Abdelhak S. Phylogeny and genetic structure of Tunisians and their position within Mediterranean populations. *Mitochondrial DNA*. 2015; 26(4):593-604.
40. Botigué LR, Henn BM, Gravel S, Maples BK, Gignoux CR, Corona E, Atzmon G, Burns E, Ostrer H, Flores C, Bertranpetit J, Comas D, Bustamante CD. Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc Natl Acad Sci USA*. 2013; 110(29):11791-11796.
41. Grigorescu F. Medigene Program Rewarded with the European Stars Award 2017. *Acta Endocrinol (Buchar)*. 2017; 13(4):512-513.