



HAL
open science

Association of estrogen receptor gene variants (ESR1 and ESR2) with polycystic ovary syndrome in Tunisia

Zeineb Douma, Meriem Dallel, Weal Bahia, Assila Ben Salem, Faten Hachani Ben Ali, Wassim Y. Almawi, Corinne Lautier, Sara Haydar, Florin Grigorescu, Touhemi Mahjoub

► To cite this version:

Zeineb Douma, Meriem Dallel, Weal Bahia, Assila Ben Salem, Faten Hachani Ben Ali, et al.. Association of estrogen receptor gene variants (ESR1 and ESR2) with polycystic ovary syndrome in Tunisia. *Gene*, 2020, 741, pp.144560 -. 10.1016/j.gene.2020.144560 . hal-03489803

HAL Id: hal-03489803

<https://hal.science/hal-03489803>

Submitted on 22 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Association of estrogen receptor gene variants (*ESR1* and *ESR2*) with polycystic ovary syndrome in Tunisia

Zeineb Douma^a, Meriem Dallel^a, Weal Bahia^a, Assila Ben Salem^a, Faten Hachani Ben Ali^b, Wassim Y. Almawi^c, Corinne Lautier^d, Sara Haydar^d, Florin Grigorescu^{d,e} and Touhemi Mahjoub^{a*}

a, Laboratory of Human Genome and Multifactorial Diseases (LR12ES07), Faculty of Pharmacy, University of Monastir, Monastir, Tunisia

b, Farhat Hached University Hospital, Unit of Reproductive Medicine, Sousse, Tunisia

c, School of Medicine, Nazarbayev University, Astana, Kazakhstan and Faculty of Sciences, El-Manar University, Tunis, Tunisia

d, University of Montpellier, UMR204 NUTRIPASS (IRD, UM, SupAgro), Montpellier, France

e, Institut Convergences Migrations, Collège de France, Paris, France

Corresponding author information:

Florin Grigorescu
IURC,
Molecular Endocrinology,
641 Avenue du Doyen Gaston Giraud,
34093 Montpellier
France
Email: florin.grigorescu@inserm.fr
Phone: +33 411759895
Fax: +33 67542731
Cellular: +33 626232481

Association of estrogen receptor gene variants (*ESR1* and *ESR2*) with polycystic ovary syndrome in Tunisia

Zeineb Douma^a, Meriem Dallel^a, Weal Bahia^a, Assila Ben Salem^a, Faten Hachani Ben Ali^b, Wassim Y. Almawi^c, Corinne Lautier^d, Sara Haydar^d, Florin Grigorescu^{d,e} and Touhemi Mahjoub^{a*}

a, Laboratory of Human Genome and Multifactorial Diseases (LR12ES07), Faculty of Pharmacy, University of Monastir, Monastir, Tunisia

b, Farhat Hached University Hospital, Unit of Reproductive Medicine, Sousse, Tunisia

c, School of Medicine, Nazarbayev University, Astana, Kazakhstan and Faculty of Sciences, El-Manar University, Tunis, Tunisia

d, University of Montpellier, UMR204 NUTRIPASS (IRD, UM, SupAgro), Montpellier, France

e, Institut Convergences Migrations, Collège de France, Paris, France

Corresponding author information:

Florin Grigorescu
IURC,
Molecular Endocrinology,
641 Avenue du Doyen Gaston Giraud,
34093 Montpellier
France
Email: florin.grigorescu@inserm.fr
Phone: +33 411759895
Fax: +33 67542731
Cellular: +33 626232481

Keywords: estrogen receptors, *ESR1*, *ESR2*, single nucleotide variation, polycystic ovary syndrome, metabolic syndrome, haplotype

Abstract

SNV (*single nucleotide variation*) in estrogen receptor (*ESR1* and *ESR2*) genes are susceptibility markers for complex diseases, such as cancer, metabolic disorders and women infertility. We explored six widely used SNVs in *ESR1* (rs2234693, rs9340799, rs3798577, rs3020314) and *ESR2* (rs1256049, rs4986938) in *polycystic ovary syndrome* (PCOS) in women from Tunisia (n = 254) compared to controls (n = 170). Genotyping was performed by RFLP-PCR or real-time PCR and analyzed in GoldenHelix statistical package. Logistic regression revealed association of rs2234693, rs3798577 and rs3020314 (*ESR1*) and rs1256049 (*ESR2*), the association of rs2234693 (C/T) being the strongest with $P < 4.81 \times 10^{-6}$, 2.88×10^{-5} after Bonferroni correction, OR 0.31, 95%CI (0.18-0.53)). Correlations were found with LH, LH/FSH or hyperandrogenism and even more significant with metabolic syndrome (rs9340799) and hyperglycemia (rs3798577). Among 14 haplotypes reconstructed in *ESR1* gene, four haplotypes (H1 to H4) were associated with PCOS the strongest being that H1 ($P < 0.002$) supported by Bonferroni ($P < 0.033$) and permutation tests ($P < 4 \times 10^{-4}$). In haplotype trend regression, concordant correlations were found with insulin resistance ($P < 0.033$) for H2 and with high blood pressure for H3 ($P < 0.048$). While these data revealed influential role on metabolic rather and hormonal features of PCOS, the association of rs2234693 was the strongest among all ethnic populations studied thus far giving a new insight estrogen receptor gene variation in distant North African populations and their role in metabolic alteration of PCOS.

Abbreviations

AO, Anovulation

COH, Controlled Ovarian Hyperstimulation

Era, Estrogen Receptor alpha

ESR1, *Estrogen, Eeceptor 1*

ESR2, *Estrogen Receptor 2*

FSH, Follicle Stimulating Hormone

HA, Hyperandrogenism

HOMA, Homeostasic Model Assessment, IVF, In Vitro Fertilization

LH, Luteinizing Hormone

MAF, *Minor Allele Frequency*

MetS, Metabolic Syndrome

OR, Odds Ratio

PCOS, Polycystic Ovary Syndrome

SNV, *Single-Nucleotide Variant*

TT, Total Testosterone

W/H, Waist / Hip ratio

1. Introduction

Genetic markers containing SNV (*single nucleotide variation*) of estrogen receptor genes are widely used in detecting genetic susceptibility for complex diseases such as cancer or metabolic disorders (Hegele-Hartung et al. 2004, Paterni et al. 2014). Because estrogen receptors ($ESR\alpha$ and $ESR\beta$) regulate cyclic gonatotropin release via hypothalamic-pituitary-ovarian axis and both are expressed in the human ovary, SNVs in the *ESR1* (Chr 6q25.1) and *ESR2* (Chr 14q23.2) genes were investigated for their contribution in the pathogenesis of *polycystic ovary syndrome* (PCOS) in humans and mouse models (Arao et al. 2019). PCOS is a typical complex disorder, affecting women at reproductive age and characterized by anovulation (AO), hyperandrogenism (HA) and polycystic ovaries at ultrasound scan (Franks et al. 2006, Diamanti-Kandarakis and Dunaif 2012). PCOS is also characterized by obesity, dyslipidaemia and intrinsic insulin resistance, which expose affected women to the risk of developing type 2 diabetes and metabolic syndrome (MetS) (Dunaif et al. 1989). Although extraordinary progress have been made by *Genome Wide Association Studies* (GWAS) in genetics of complex disorders, the pathogenesis of PCOS remains complex with multiple gene-gene and gene-environment interactions (Day et al. 2015, Hiam et al. 2016). Previous studies using SNVs in *ESR1* or *ESR2* genes reported positive associations with PCOS in *Caucasian, Chinese, Iranian, Pakistani, Greek* and *Brazilian* women, albeit with some variation (Valkenburg et al. 2011, Jiao et al. 2018, Khafagi et al. 2014, Liaqat et al. 2015, Nectaria et al. 2012, Silva et al. 2015). Despite this, GWAS in PCOS did not shown positive hits for these two genes in populations studied so far (Day et al. 2015). Although there would be many explanations for this situation, including SNV density and multifaceted PCOS phenotype, ethnic components might considerably contribute to variability of results. This study was aimed to investigate widely used SNVs in *ESR1* (rs2234693, rs9340799, rs3798577 and rs3020314) and *ESR2* (rs1256049 and rs4986938) genes in PCOS women from Tunisia (North Africa). We report here positive association with PCOS of SNVs in *ESR1* and *ESR2* genes, namely of rs2234693 (C/T). Studies of independent SNVs as well as haplotype analysis

in *ESR1* gene indicated correlations with metabolic profile (e.g. insulin resistance) rather than hormonal features of PCOS.

2. Population and methods

2.1. Populations. This study was performed at *Farhat Hached* Hospital, Sousse (Tunisia) where patients (n = 254) were recruited in accordance with the Helsinki Declaration (General Assembly of the World Medical Association 2014). Informed consent was obtained from all subjects and the protocol approved by ethical committee of the *Monastir* University (Tunisia). Diagnosis of PCOS was based Rotterdam criteria in the presence of two out of three criteria: oligo-ovulation or anovulation (AO), presence of polycystic ovaries at ultrasound scan, and clinical and/or biochemical signs of HA (Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004). Women with PCO and AO but without HA were not considered (Azziz et al. 2009). Patients with Cushing's syndrome, non-classical adrenal 21-hydroxylase deficiency, hyperprolactinaemia and androgen-secreting neoplasms were excluded. Patients with *Body Mass Index* (BMI) > 45 kg/m², severe chronic illness and medications affecting the carbohydrate metabolism 3 months before inclusion also were excluded. HA was considered on the basis of clinical hirsutism, androgenic alopecia and plasma levels of total testosterone (TT) with cut off value > 2.2 nmol/L (0.63 ng/ml). Oligovulation or AO was declared when women had < 8 menstrual cycles/year or amenorrhea, respectively. All patients underwent transvaginal ultrasound scan and diagnosis was made in the presence of more than 12 small follicles (2 - 9 mm) in at least one ovary and/or increase of the ovarian volume > 10 ml. Controls (n = 170) were recruited among women with normal menstrual cycle during annual checkup as previously described (Dallel et al. 2019). DNA samples were declared to MESRI (*Ministère de l'Enseignement Supérieur de la Recherche et de l'Innovation*) in France under CODECOH # DC-2014-2226 and analyzed in the frame of European MEDIGENE (FP7-279171) program (https://cordis.europa.eu/project/rcn/101810_en.html). MetS diagnosis was based on the National Cholesterol Education Program (NCEP) and Adult Treatment Panel-III criteria (Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults 2001). Impaired fasting glucose (IFG) was defined by ADA criteria with cut-off value of 5.6 mmol/L glucose (American Diabetes Association, 2018). Serum samples were analyzed for fasting glucose, total cholesterol, low-

density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol, triglycerides (TG) and fasting insulin using commercial kits as described (Dallel et al. 2019, Ben Salem et al. 2015). *Homeostasis Model Assessment* for insulin resistance ($HOMA_{IR}$) was calculated by fasting glycemia (mmol/L) \times fasting insulin (μ U/mL) divided by 22.5. Insulin resistant and non-insulin resistant subjects were classified based on fasting insulin in controls as previously described (Haydar et al. 2019). Serum samples were also assayed for follicular-stimulating hormone (FSH), luteinizing hormone (LH), TT, estradiol (E2), prolactin (PRL) and anti-Müllerian hormone (AMH) by chemiluminescence immunoassays as previously described (Dallel et al. 2018) with intra- and inter-assay coefficients of variation $< 10\%$.

2.2. SNV genotyping. Blood samples were collected on EDTA and genomic DNA was isolated by the salting-out method as described (Ben Salem et al. 2015). Two SNVs (rs3798577 and rs3020314) were genotyped by VIC- and FAM-labeled real-time TaqMan assay (*Applied Biosystems*, Foster City, CA) according to manufacturer's instructions. Replicated blinded quality control samples were included to assess reproducibility, which indicated concordance $> 99\%$. The average successful genotyping rate was 95.9%. Other SNVs (rs2234693, rs9340799, rs1256049 and rs4986938) were detected by RFLP-PCR as previously described (Ghali et al. 2018). Genomic positions of SNVs (GRCh37/hg19, by February 2009) were retrieved from the *Axiom Analysis Suite 1.1* (Affymetrix) as described (Haydar et al. 2019). All SNVs were in Hardy-Weinberg equilibrium and *Minor allele frequency* (MAF) in Tunisians was compared to that of 1000 Genomes Project (www.1000genomes.org/data).

2.3. Statistical methods. Statistical analysis was performed using SNP & Variation Suite (SVS) v8.4.4 program (Golden Helix. Inc., Bozeman, MT) (www.goldenhelix.com) and we applied the MEDISCOPE map file as described (Haydar et al. 2019, Čoklo et al. 2016). Genetic association was performed by logistic regression. Significance was considered when $P < 0.05$ and supported by Bonferroni multiple testing correction. The study was powered (PBAT) > 0.87 for 6 SNVs for 424 samples, which indicated significance when $MAF > 0.2$ and major gene effects ($OR > 2.0$). All basic, additive, recessive and dominant models were tested and data were reported in the “results” section for the basic model and the best values from other models or genotypes. Besides the PCOS, we also tested

the association with MetS *per se*, its components and insulin resistance (nominal variables). For numerical variables, correlations were investigated using the Correlation Trend Test (CTT) and results were expressed as P and R coefficient values. Correlation with HOMA_{IR} index and cumulative criteria for MetS was studied by ANOVA in Statview 5.0 program (SAS, Abacus Concepts, Berkeley, CA), setting 2α level at 5%, as previously described (Lautier et al. 2003) and using *a posteriori* Student-Newman-Keuls test for significance among groups. Haplotypes in *ESR1* gene were analyzed in HAPLOVIEW program (<https://www.broadinstitute.org/haploview>) and results were expressed as P value of χ^2 test after permutation using 10,000 iterations (Haydar et al. 2019, Barrett et al. 2005). Results for haplotype association were reported from the SVS v8.4.4 program while correlations with numerical variables were reported from haplotype regression tests by providing P value for each haplotype and the correlation coefficient. Functional prediction for potential modified transcriptional activity was performed with HaploReg v4.1 (<http://archive.broadinstitute.org/mammals/haploreg>).

3. Results

3.1. *Association of SNVs in ESR1 and ESR2 genes.* This study was conducted in a Tunisian population of 424 subjects composed of 254 PCOS patients and 170 control women (**Table 1**). Clinical and laboratory features indicated that women with PCOS were 1.7 - fold more obese, displaying more severe insulin resistance (1.3-fold increase in $HOMA_{IR}$) and 8-fold increase in the prevalence of MetS. All components of the MetS were more prevalent, particularly central obesity and hyperglycemia ($P < 0.0001$, in χ^2 test) and hypertriglyceridemia ($P < 0.0003$). As expected, PCOS women displayed higher TT levels ($P < 0.0001$, Mann-Whitney) and LH/FSH ratio ($P < 0.0035$), but LH did not reach significance. Hirsutism was present in 28.30% (5.40 % in controls) and acne in 44.40% of PCOS women (10% in controls). No significant differences were observed for waist over hip (W/H) ratio, fasting insulin, PRL, E2, but AMH level was 2.4-fold higher in PCOS ($P < 0.0001$, Mann-Whitney).

When SNVs were tested independently, two SNVs in *ESR1* (rs2234693 and rs3798577) and one in *ESR2* (rs1256049) were associated with PCOS (**Table 2**). In the basic model, association of rs2234693 (C/T) was supported by Bonferroni correction ($P < 0.02$) but the best significance was obtained with the dominant model ($P < 4.81 \times 10^{-6}$, 2.88×10^{-5} after Bonferroni correction, OR 0.31, 95%CI (0.18-0.53)). For rs1256049 of *ESR2* the best association was obtained for genotypes. While rs2234693 and rs3798577 in *ESR1* were frequent (MAF 0.49 for both) rs1256049 in *ESR2* had low MAF (0.03). As shown in **Figure 1**, prevalence of C allele of rs2234693 in *ESR1* was proportional with increasing values of $HOMA_{IR}$ index as function of cumulative criteria for MetS with values ($P < 0.03$, $\alpha = 0.64$, ANOVA).

Next, we studied the influential role of SNVs on hormonal and metabolic features of PCOS. Trend correlations were found for HA, LH, LH/FSH and TT but only the correlation with LH levels was maintained as a trend after Bonferroni correction ($P < 0.051$). By contrast, several SNVs in *ESR1* gene were associated with MetS, central obesity, low HDL and hyperglycemia (as nominal variables), the most significant being that of rs9340799 for MetS and that of rs3798577 for hyperglycemia (**Table 3**).

3.2. *Haplotype association and functional correlation.* To understand influential role of SNVs in the *ESR1* gene we performed the haplotype analysis. In HAPLOVIEW, 15 haplotypes were revealed (**Table 1S**), among which the haplotype TATC remained positively associated after permutation test ($P < 0.0004$). Haplotypes were also tested in GoldenHelix program, which use the Expectation/Maximization (EM) algorithm. A number of 14 haplotypes were identical (haplotype TGCT with frequency 0.01 was not found). In GoldenHelix, logistic regression revealed positive associations with PCOS for haplotypes H1, H2, H3 and H4). Association of H1 (TATC) with frequency of 0.078 was maintained by Bonferroni correction and displayed protective effect. Haplotypes H2, H3 and H4 were pathogenic, all with OR between 2 and 4.6 (**Table 4**).

Based on these associations we explored potential correlations with biochemical and clinical features PCOS. As summarized in **Table 4**, correlations for PCOS features were not striking. Thus, haplotype H8 (TGTC) was associated with HA, but haplotype trend test did not reveal correlations with hormonal levels. By contrast, haplotypes, including those associated with PCOS were well correlated with HOMA_{IR} index or components of MetS such as central obesity, high blood pressure or low HDL. Haplotype H1 and H4 were associated with obesity and correlated with HOMA_{IR}, respectively, indicating that SNVs in the *ESR1* gene better reflect metabolic features rather than PCOS specific hormonal profile. To better understand the significance of these findings, we have search in HAPLOREG database potential annotation for modifications of transcriptional activity (**Table 5**). The rs2234693 (*ESR1*) would interact with AP-4 while rs4986938 in *ESR2* with Pax-6_1, Pax-8 and RAR. Finally, in the hypothesis that ethnic variability could explain variation in PCOS association, we have compared SNV frequency with other ethnic populations. All Tunisian SNVs had comparable MAF compared to Europeans, except rs9340799 (*ESR1*) with MAF of 0.47, less frequent than other populations and rs3020314 (*ESR1*) with MAF 0.39 higher than Europeans (MAF 0.29) but close to Africans (0.75) or Asians (0.80).

4. Discussion

In this paper we provide evidence for genetic association of widely used SNVs in *ESR1* and *ESR2* genes with PCOS in Tunisian population either as independent SNV markers or as complex haplotypes in *ESR1* gene. Association for rs2234693 (C/T) in *ESR1* was the strongest among all ethnic populations studied so far, driving the attention on the contribution of ethnic variability in association studies (Valkenburg et al. 2011, Jiao et al. 2018, Khafagi et al. 2014, Liaqat et al. 2015, Nectaria et al. 2012, Silva et al. 2015). Although the association, was correlated to hormonal profile of PCOS, the significant correlation were found with MetS or its components, including insulin resistance. These data are in agreement to other studies on coronary artery diseases and hyperlipidemia (Lucas et al. 2011, Zhao et al. 2010).

Clinical profile of PCOS patients from Tunisia was similar to other ethnic populations (Diamanti-Kandarakis and Dunaif 2012). It should be noted that MetS and obesity were more prevalent in PCOS women even in the absence of significant differences for W/H ratio, fasting insulin, HOMA_{IR} or HDL levels. These features are in agreement with other studies (GIANT consortium), which indicated more severe obesity in PCOS than the general population (Wojciechowski et al. 2012). Although differences were found for age, it is unlikely that this feature would contribute to genetic findings in PCOS.

The major finding in this study was the association for rs2234693 (C/T) in *ESR1* gene, which was supported by Bonferroni correction, the C allele being pathogenic and T allele protective. The significance ($P < 4.81 \times 10^{-6}$) of this SNV did not reach GWAS significance ($P < 10^{-8}$), in accordance with previous studies ((Day et al. 2015). Among populations studied so far, no association was found for rs2234693 in Caucasian, Han Chinese, Iranian, Greeks and Brazilian (9 Valkenburg et al. 2011, Jiao et al. 2018, Khafagi et al. 2014, Nectaria et al. 2012, Silva et al. 2015) populations, except the Pakistani women (Liaqat et al. 2015) where the C allele was also pathogenic ($P < 0.005$, OR 1.84). In this study, the rs9340799 (G/A) was not significant, but associations were found in Chinese and Pakistani populations (Jiao et al. 2018, Liaqat et al. 2015, 12). For other SNVs in *ESR1* gene, associations were concordant with reports in Chinese population (Jiao et al. 2018). In *ESR2* gene, allele association of rs1256049 (G/A) was not maintained after Bonferroni correction, although

significance was reached with additive model ($P < 0.02$) and for genotypes ($P < 0.009$). Association rs4986938 (G/A) was not significant in contradiction with the study in Pakistani women, although the OR of 0.88 indicated the same protective effect of the A allele (12). Taken together these data suggest that at comparable MAF, there would be some specificity for Tunisians from North Africa in marking the PCOS syndrome. Along this line, it should be noted that same SNVs (e.g. rs3798577 in *ESR1* and rs1256049 in *ESR2*) in Tunisians were also associated with breast cancer ($P < 1.2 \times 10^{-3}$ and 5.8×10^{-4} , respectively) and the rs2234693 was reported with the same pathogenic C allele (Ghali et al. 2018). All together these data drive the attention on the importance of ethnic variation in association studies, including GWAS. Tunisians from North Africa are genetically distant from Europeans and very likely possess different linkage disequilibrium blocking compared to European populations registered in 1000 genome project.

Correlation of SNVs in *ESR1* and *ESR2* genes with hormonal and biochemical parameters of PCOS also deserves discussion. For instance, protective A alleles of rs4986938 and rs9340799, were concordantly negatively correlated with HA (OR 0.49), TT, central obesity and MetS. The C protective allele of rs3798577 was concordantly negatively correlated with LH levels and LH/FSH ratio, but discordantly for hyperglycemia. Concordant was also the protective T allele of rs2234693, negatively correlated with W/H ratio but conversely, correlated with low HDL (nominal variable) indicating rather insulin resistance. The most plausible explanation for these contradictions is the complex effects of haplotypes. Haplotype analysis revealed 3 haplotypes pathogenic for PCOS with $OR > 2.0$. Only two of them contain the pathogenic C allele of rs2234693 (H2 and H3). The most significant associated haplotype (H1) and sustained by Bonferroni correction and contains in the first position T protective allele of rs2234693. Very likely the markers studied are surrogate SNV leaving the possibility that causal variations to be distantly located, which is typical for complex disorders. Moreover, effects of haplotypes should be also considered as function of their frequency. The most significant haplotype associated with PCOS (H1) was at low frequency (0.07) but the most frequent haplotype H3 (CGTT, 0.174) was pathogenic and contained pathogenic alleles for all 4 SNVs (rs2234693, rs9340799, rs3798577 and rs3020314, respectively).

Since haplotypes were most likely associated with metabolic features, including H1 for obesity and H4 for insulin resistance, we cannot exclude the possibility that correlations found in this study would be explained by more prevalent and severe metabolic alterations in PCOS. This is accordance with similar descriptions in Greeks (e.g. for rs2234693) where despite absence of association, C allele of rs2234693 displayed higher glucose/insulin ratio, thus reflecting insulin resistance (Nectaria et al. 2012). This observation is concordant to pathogenic mutations in *ESR1* gene associated to unresponsiveness of receptors, glucose intolerance and severe hyperinsulinemia accompanied by *Acanthosis Nigricans* (Smith et al. 1994). This is also concordant to description of novel binding site for B-*myb* modifying transcription of $ES\alpha$, although we cannot conclude from our data on the potential mechanisms involved (Swartz et al. 2005). Besides the ethnic variation, we cannot exclude the possibility that other features of PCOS not investigated in this study would be correlated with *ESR1* and *ESR2* genes. Indeed, ESRs were considered candidates for PCOS also because of their role in follicle growth and maturation or release of the oocyte (Arao et al. 2019, Silva et al. 2015). Thus *ESR1* gene (rs2234693) was reported as being associated with endometriosis and susceptibility for *controlled ovarian hyperstimulation* (COH) and pregnancy outcomes of IVF (Silva et al. 2015). To continue the comparison with our previous results in North Africa, we should note that haplotypes H2, H3 and H5 associated with PCOS also were synergistically pathogenic for breast cancer, but in Tunisians, haplotype H1 appeared specific for PCOS (Ghali et al. 2018). In conclusion, based on SNV and haplotype data, genes encoding for ESRs are good markers for PCOS in North Africa, particularly for metabolic features and further studies would necessary to completely explain their influential role in PCOS and obtain valuable markers at the clinical scale.

5. References

- American Diabetes Association 2018. Introduction: Standards of Medical Care in Diabetes-2018. *Diabetes Care* 41:S1–S2. <https://doi.org/10.2337/dc18-Sint01>
- Arao Y, Hamilton KJ, Wu S-P, et al (2019) Dysregulation of hypothalamic-pituitary estrogen receptor α -mediated signaling causes episodic LH secretion and cystic ovary. *FASEB J* 33:7375–7386. <https://doi.org/10.1096/fj.201802653RR>
- Azziz R, Carmina E, Dewailly D, et al (2009) The Androgen Excess and PCOS Society criteria for the polycystic ovary syndrome: the complete task force report. *Fertil Steril* 91:456–488. <https://doi.org/10.1016/j.fertnstert.2008.06.035>
- Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265. <https://doi.org/10.1093/bioinformatics/bth457>
- Ben Salem A, Attaoua R, Mtiraoui N, et al (2015) Haplotyping strategy highlights the specificity of FTO gene association with polycystic ovary syndrome in Tunisian women population. *Gene* 565:166–170. <https://doi.org/10.1016/j.gene.2014.12.004>
- Čoklo M, Auguštin DH, Šarac J, et al (2016) Diversity of Y-chromosomal and mtDNA Markers Included in Mediscope Chip within Two Albanian Subpopulations from Croatia and Kosovo: Preliminary Data. *Coll Antropol* 40:195–198
- Dallel M, Sarray S, Douma Z, et al (2018) Differential association of DENND1A genetic variants with polycystic ovary syndrome in Tunisian but not Bahraini Arab women. *Gene* 647:79–84. <https://doi.org/10.1016/j.gene.2018.01.028>
- Dallel M, Sghaier I, Finan RR, et al (2019) Circulating leptin concentration, LEP gene variants and haplotypes, and polycystic ovary syndrome in Bahraini and Tunisian Arab women. *Gene* 694:19–25. <https://doi.org/10.1016/j.gene.2019.01.032>

- Day FR, Hinds DA, Tung JY, et al (2015) Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome. *Nat Commun* 6:8464. <https://doi.org/10.1038/ncomms9464>
- Diamanti-Kandarakis E, Dunaif A (2012) Insulin Resistance and the Polycystic Ovary Syndrome Revisited: An Update on Mechanisms and Implications. *Endocr Rev* 33:981–1030. <https://doi.org/10.1210/er.2011-1034>
- Dunaif A, Segal KR, Futterweit W, Dobrjansky A (1989) Profound peripheral insulin resistance, independent of obesity, in polycystic ovary syndrome. *Diabetes* 38:1165–1174. <https://doi.org/10.2337/diab.38.9.1165>
- Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (2001) Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA* 285:2486–2497. <https://doi.org/10.1001/jama.285.19.2486>
- Franks S, Mccarthy MI, Hardy K (2006) Development of polycystic ovary syndrome: involvement of genetic and environmental factors. *International Journal of Andrology* 29:278–285. <https://doi.org/10.1111/j.1365-2605.2005.00623.x>
- General Assembly of the World Medical Association (2014) World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *J Am Coll Dent* 81:14–18. <https://www.ncbi.nlm.nih.gov/pubmed/25951678>
- Ghali RM, Al-Mutawa MA, Al-Ansari AK, et al (2018) Differential association of ESR1 and ESR2 gene variants with the risk of breast cancer and associated features: A case-control study. *Gene* 651:194–199. <https://doi.org/10.1016/j.gene.2018.02.011>
- Goodarzi MO, Azziz R (2006) Diagnosis, epidemiology, and genetics of the polycystic ovary syndrome. *Best Practice & Research Clinical Endocrinology & Metabolism* 20:193–205. <https://doi.org/10.1016/j.beem.2006.02.005>

- Haydar S, Grigorescu F, Vintilă M, et al (2019) 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* 14:e0214122. <https://doi.org/10.1371/journal.pone.0214122>
- Hegele-Hartung C, Siebel P, Peters O, et al (2004) Impact of isotype-selective estrogen receptor agonists on ovarian function. *Proc Natl Acad Sci USA* 101:5129–5134. <https://doi.org/10.1073/pnas.0306720101>
- Hiam D, Moreno-Asso A, Teede HJ, et al (2019) The Genetics of Polycystic Ovary Syndrome: An Overview of Candidate Gene Systematic Reviews and Genome-Wide Association Studies. *J Clin Med* 8:. <https://doi.org/10.3390/jcm8101606>
- Jiao X, Chen W, Zhang J, et al (2018) Variant Alleles of the ESR1, PPARG, HMGA2, and MTHFR Genes Are Associated With Polycystic Ovary Syndrome Risk in a Chinese Population: A Case-Control Study. *Front Endocrinol (Lausanne)* 9:. <https://doi.org/10.3389/fendo.2018.00504>
- Khafagi Z, Mozdarani H, Behmanesh M, Aleyasin A (2014) Distribution Of FSHR, ESR1 And ESR2 Snps Among Women With Polycystic Ovary Syndrome Undergoing In Vitro Fertilization. *BullEnvPharmacolLife Sci* 3:128–134.<http://www.bepls.com>.
- Lautier C, El Mkadem SA, Renard E, et al (2003) Complex haplotypes of IRS2 gene are associated with severe obesity and reveal heterogeneity in the effect of Gly1057Asp mutation. *Hum Genet* 113:34–43. <https://doi.org/10.1007/s00439-003-0935-3>
- Liaqat I, Jahan N, Krikun G, Taylor HS (2015) Genetic Polymorphisms in Pakistani Women With Polycystic Ovary Syndrome. *Reprod Sci* 22:347–357. <https://doi.org/10.1177/1933719114542015>
- Lucas G, Lluís-Ganella C, Subirana I, et al (2011) Post-genomic update on a classical candidate gene for coronary artery disease: ESR1. *Circ Cardiovasc Genet* 4:647–654. <https://doi.org/10.1161/CIRCGENETICS.111.960583>

- Nectaria X, Leandros L, Ioannis G, Agathocles T (2012) The importance of ER α and ER β gene polymorphisms in PCOS. *Gynecol Endocrinol* 28:505–508. <https://doi.org/10.3109/09513590.2011.649811>
- Paterni I, Granchi C, Katzenellenbogen JA, Minutolo F (2014) Estrogen receptors alpha (ER α) and beta (ER β): subtype-selective ligands and clinical potential. *Steroids* 90:13–29. <https://doi.org/10.1016/j.steroids.2014.06.012>
- 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* 81:19–25. <https://doi.org/10.1016/j.fertnstert.2003.10.004>
- Shi Y, Zhao H, Shi Y, et al (2012) Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome. *Nat Genet* 44:1020–1025. <https://doi.org/10.1038/ng.2384>
- Silva FS, S3oter MO, Sales MF, et al (2015) Estrogen receptor α gene (ESR1) PvuII and XbaI polymorphisms are associated to metabolic and proinflammatory factors in polycystic ovary syndrome. *Gene* 560:44–49. <https://doi.org/10.1016/j.gene.2015.01.037>
- Valkenburg O, Uitterlinden AG, Themmen AP, et al (2011) Genetic polymorphisms of the glucocorticoid receptor may affect the phenotype of women with anovulatory polycystic ovary syndrome. *Hum Reprod* 26:2902–2911. <https://doi.org/10.1093/humrep/der222>
- Wojciechowski P, Lipowska A, Rys P, et al (2012) Impact of FTO genotypes on BMI and weight in polycystic ovary syndrome: a systematic review and meta-analysis. *Diabetologia* 55:2636–2645. <https://doi.org/10.1007/s00125-012-2638-6>
- Zhao T, Zhang D, Liu Y, et al (2010) Association between ESR1 and ESR2 gene polymorphisms and hyperlipidemia in Chinese Han postmenopausal women. *J Hum Genet* 55:50–54. <https://doi.org/10.1038/jhg.2009.122>

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Acknowledgments

This work was performed as collaboration between France and Tunisia in the frame of MEDIGENE project of the European Union (FP7-279171-1). FG was CRHC (*Chargé de Recherche Hors Classe*) at *Institut National de la Santé et de la Recherche Médicale* (INSERM) up to June 2019, currently affiliated to the *Institut Convergences Migrations* (ICM), Collège de France, Paris (France) and *Honorary Professor* in genetics at *Carol Davila* University of Medicine & Pharmacy Bucharest, Romania.

Funding

This research was partially supported by European Union (grant FP7-279171-1).

Author contribution statement

ZD: sample processing, genotyping, Original draft preparation

MD: sample processing, literature search

WB: genotyping

AB: sample processing

FH: patient screening and referral

WA: Analysis of data, biochemical investigation

CL: Statistical analysis

SH: Statistical analysis

FG: Haplotype analysis, project leader, Supervision

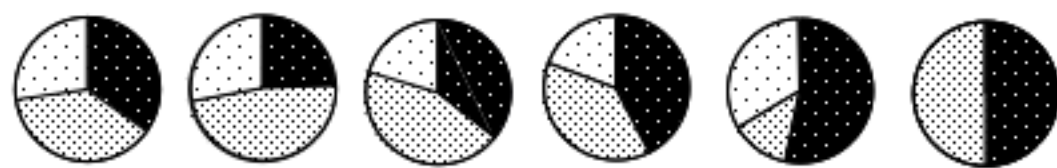
TM: Drafting of manuscript, partial funding, Supervision

Figure legends

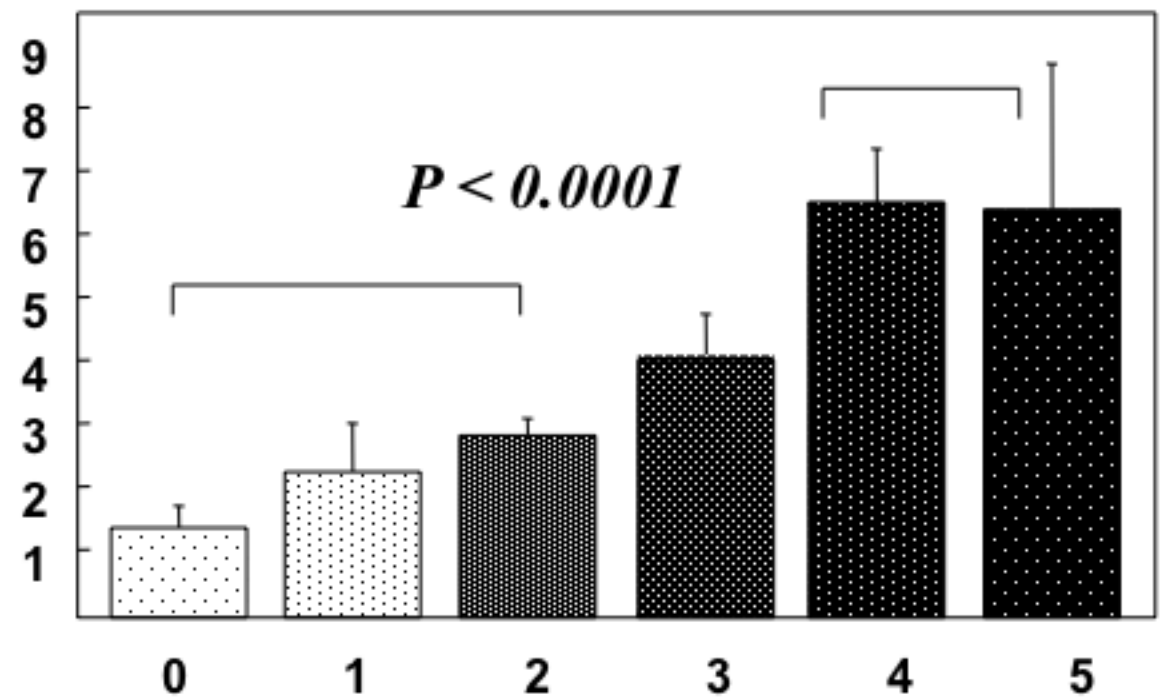
Figure 1. Relationship between cumulative criteria for MetS and HOMA_{IR} index and correlation with genotypes of rs2234693 (TT, CT, and CC) in Tunisian women with PCOS. P value refers to the significance (Student-Newman-Keuls *a posteriori* test) of the correlation between HOMA_{IR} values and cumulative criteria for MetS (ANOVA). Correlation with genotypes is indicated in the “Results” section.

Genotypes of rs2234693

TT  TC  CC 



Insulin resistance
HOMA_{IR}



Cumulative criteria for MetS

Table 1. Clinical and laboratory features of PCOS patients compared to control women from Tunisia

Parameter	Controls	PCOS	P value ^d
<i>n</i>	170	254	NA
Age (years)	35.22 ± 0.43	31.08 ± 0.31	< 0.0001
BMI (kg/m ²)	27.57 ± 0.56	29.14 ± 0.39	< 0.01
Waist (cm)	65.23 ± 2.9	88.67 ± 2.77	< 0.0001
W/H ^a	0.91 ± 0.00	0.93 ± 0.00	NS
BMI-based obesity (%) ^b	27.73	43.38	0.003
Fasting glycemia (mmol/L)	4.13 ± 0.18	6.25 ± 0.22	< 0.0001
Fasting insulin (μU/ml)	12.51 ± 1.70	15.99 ± 1.40	NS
HOMA _{IR}	2.73 ± 0.49	3.63 ± 0.28	NS
Insulin resistance (%) ^c	33.33	75.89	< 0.0001
Cholesterol (mmol/L)	4.24 ± 0.13	5.41 ± 0.13	< 0.0001
HDL (mmol/L)	1.19 ± 0.05	1.28 ± 0.04	NS
TG (mmol/L)	1.21 ± 0.11	1.66 ± 0.09	< 0.0087
Central obesity (%)	36.47	57.87	< 0.0001
Hyper Triglyceridemia	2.94	17.71	0.0003
Hyperglycemia (%) ^d	2.35	29.92	< 0.0001
Low HDL cholesterol (%)	12.94	22.44	0.013
High Blood Pressure (%)	NA	17.32	NA
MetS ^{ATP-III} (%)	2.35	19.29	< 0.0001
Total Testosterone (nmol/L)	0.92 ± 0.14	3.32 ± 0.21	< 0.0001
Hyperandrogeny (%)	7.5	54.47	< 0.0001
LH/FSH	0.70 ± 0.04	0.91 ± 0.05	0.0035
LH (IU/L)	4.79 ± 0.25	5.60 ± 0.33	NS
FSH (IU/L)	7.83 ± 0.35	6.31 ± 0.16	< 0.0001
Prolactin (nmol/L)	183.25 ± 44.77	130.17 ± 22.51	NS
E2 (nmol/L)	0.40 ± 0.05	0.39 ± 0.035	NS
AMH ng/ml	51.32 ± 4.50	124.57 ± 3.84	< 0.0001

a, waist/hip ratio; b, the cut off value of Body Mass Index (BMI) was 30 kg/m²; c, Patients were stratified as insulin resistant or non-insulin resistant using a cut-off value defined as the mean of fasting insulin + 1SD of lean controls, which corresponds to HOMA_{IR} of 2.8, close to the median (2.48) of PCOS patients; Insulin resistance is reported as prevalence; d, statistical significance was obtained by Mann-Whiney test for numerical variables and χ^2 for nominal variable; NS stands for non-significant; NA stands for not applicable.

Table 2. Genetic association with PCOS of SNVs in *ESR1* and *ESR2* genes in Tunisian women

Gene	SNV ID	Position ^a	Ref/Alt allele ^b	MAF	OR [95%CI]	P ^c	Bonferroni correction ^d	P ^e
<i>ESR1</i>	rs2234693	152163335	C/T	0.49	0.65 [0.49-0.87]	0.0037	0.02	4.81 x 10 ⁻⁶ (D)
	rs9340799	152463381	G/A	0.47	0.87 [0.61-1.23]	0.420	1	0.37 (A)
	rs3798577	152421130	T/C	0.49	0.53 [0.29-0.97]	0.039	0.23	0.032 (A)
	rs3020314	152270672	T/C	0.39	1.78 [0.97-3.33]	0.065	0.39	0.036 (A)
<i>ESR2</i>	rs1256049	64724051	G/A	0.03	0.31 [0.12-0.81]	0.014	0.08	0.009 (G)
	rs4986938	64699816	G/A	0.34	0.88 [0.63-1.23]	0.440	1	0.46 (A)

a, position is reporting for GRCh37/hg19; b, Reference and alternative allele are reported from MEDISCOPE map file (*Affymetrix*); c, P value is reported for the “basic model” for all SNVs; d, Bonferroni correction was performed using the following formula: $P_c = 1-(1-P)^n$ where n is the number of comparisons; e, the best values as function of the genetic model: D, dominant; A, Additive; G, genotype association

Table 3. Significant correlations of SNVs in *ESR1* and *ESR2* genes in PCOS Tunisian population

Parameter ^a	Chr	SNV ID	P value	P _{corr} ^b	R coeff	OR ^c	95%CI
LH	6	rs3798577	0.0086	0.051	- 0.342	NA	NA
LH/FSH	6	rs3798577	0.0240	0.140	- 0.294	NA	NA
HA	14	rs4986938	0.0253	0.152	NA	0.49	0.26 - 0.92
Testosterone	6	rs9340799	0.0463	0.270	- 0.172	NA	NA
MetS	6	rs9340799	0.0069	0.041	NA	0.28	0.11 -0.69
Central Obesity	6	rs9340799	0.0570	0.280	NA	0.38	0.14 -1.03
W/H	6	rs2234693	0.0290	0.170	- 0.266	NA	NA
Low HDL	6	rs2234693	0.0237	0.143	NA	0.43	0.24 -0.77
Hyperglycemia	6	rs3798577	0.0023	0.014	NA	9.33	2.04 -42.66

a, Numerical variables were tested by Correlation Trend Test (CTT) while nominal variables were tested by logistic regression; b, P_{corr} refers Bonferroni correction; c, OR and R represent the odds ratio from logistic regression or the R coefficient in the CTT; NA stands for not-applicable.

Table 4. Association of haplotypes in *ESR1* gene with PCOS in Tunisian population and summary effects on phenotype features

ID #	Haplotype	Freq	Association with PCOS			Summary effect of haplotype on phenotype features ^a
			P	P ^{corr}	OR [95%CI]	
H1	TATC	0.078	0.002	0.033	0.23 [0.062 - 0.84]	Associated with obesity (P < 0.015, OR 0.026 [0.07-0.98])
H2	CGCT	0.072	0.008	0.117	4.68 [0.73 - 29.83]	Correlation with HOMA _{IR} (P < 0.033, Coefficient 12.3)*
H3	CGTT	0.174	0.030	0.416	2.03 [0.78 - 5.30]	Association with high blood pressure (P < 0.048, OR 9.38 [0.53-163.74])*
H4	TGCC	0.089	0.031	0.431	4.00 [0.83 - 19.31]	Correlation with HOMA _{IR} (P < 0.035, Coefficient 14.6)
H5	CGTC	0.110	0.074	1.000	0.46 [0.16 - 1.30]	Correlation with HOMA _{IR} (P < 0.050 - Coefficient 10.71)*
H6	TACC	0.048	0.173	1.000	0.46 [0.10 - 2.11]	Correlation with HOMA _{IR} (P < 0.025 - Coeff 14.43)*
H7	CATC	0.045	0.226	1.000	0.55 [0.11 - 2.61]	Association with high blood pressure (P < 0.038, OR 5.99 [0.35-101.94])*
H8	TGTC	0.031	0.377	1.000	0.56 [0.087 - 3.60]	Association with HA (P < 0.013 OR 13.51 [0.24-749.8])* Trend association with obesity (P < 0.09 in PCOS* and P < 0.015)
H9	CATT	0.040	0.381	1.000	1.84 [0.28 12.03]	Correlation with HOMA _{IR} (P < 0.084 - Coefficient 10.18)* Trend association with Low HDL (P < 0.038)*
H10	CACT	0.048	0.496	1.000	1.47 [0.29 - 7.56]	None
H11	TGTT	0.043	0.537	1.000	0.67 [0.14 - 3.30]	Correlation with HOMA _{IR} (P < 0.03 - Coefficient 21.14)* Trend association with obesity (P < 0.043), low HDL (P < 0.011)* Correlation with HOMA _{IR} (P < 0.0009, Coefficient 14.64)
H12	TACT	0.070	0.614	1.000	0.78 [0.22 - 2.81]	Trend association with insulin resistance (P < 0.009)
H13	TATT	0.065	0.678	1.000	0.82 [0.22 - 3.09]	Trend in correlation with dyslipidemia (P < 0.045)*
H14	CACC	0.084	0.686	1.000	1.22 [0.36 - 4.15]	Trend in association with cholesterol levels (P < 0.09)*

a, effects also observed in PCOS population are indicated by *; P^{corr} refers to Bonferroni correction;

Table 5. Minor allele frequency (MAF) of SNVs in *ESR1* and *ESR2* genes in Tunisians compared to other world populations and annotated transcriptional factors

SNV ID	Chr	Allele	AFR ^a	EUR	EAS	TN	Annotated TF ^b
rs2234693	6	C/T	0,43	0,57	0,60	0.49	AP-4_2, E2A_3, HEN1_1, HEN1_2, LBP-1_2, LUN-1, RP58
rs9340799	6	G/A	0,73	0,69	0,80	0.47	Ets_disc5, Hand1_1
rs3798577	6	T/C	0.43	0.47	0.4	0.49	Gene expression of DYX1C1 (probeID ILMN_1812523) in breast tumors ()
rs3020314	6	T/C	0.75	0.29	0.80	0.39	Not annotated
rs1256049	14	G/A	0.10	0.03	0.40	0.03	Not annotated
rs4986938	14	G/A	0.28	0.38	0.12	0.34	CTCF_known1, Nr2f2, Pax-6_1, Pax-8_1, RAR

a, AFR, EUR, EAS and TN stand for African, Europeans, East Asians and Tunisians, respectively; b, TF stands for transcriptional factors; MAF values are reported from 1000 genome project for the minor allele defined in Table 2.