Abstract. – OBJECTIVE: Transcription factor 7-like 2 (TCF7L2) gene variants confer risk for type 2 diabetes and metabolic traits. We investigated the role of TCF7L2-variants in polycystic ovarian syndrome (PCOS), which is a common endocrine metabolic disorder affecting women of reproductive age. We tested whether TCF7L2 variants are in linkage to and/or in linkage disequilibrium [LD], namely linkage and association) with PCOS.

PATIENTS AND METHODS: Within 212 families from the Italian peninsular population, we analyzed 78 variants using Pseudomarker software for linkage to and LD with PCOS under the dominant model with complete penetrance (D1). In a secondary analysis, we tested the variants under the recessive models with complete penetrance (R1), dominant with incomplete penetrance (D2), and recessive with incomplete penetrance (R2). We tested through in silico analysis the risk variants to detect any potential functional effects.

RESULTS: We identified a total of 14 variants in the TCF7L2 gene significantly linked to and/or in LD with the risk of PCOS (p < 0.05) across different models.

CONCLUSIONS: This study is the first to report TCF7L2 linkage and linkage disequilibrium in Italian families with PCOS.

Key Words: Transcription factor 7-like 2, TCF7L2, Association, Gene, Single nucleotide polymorphism, SNP, Polycystic Ovarian Syndrome, PCOS, Ovary, Hyperandrogenism, Oligomenorrhea, Anovulation, Chronic, Irregular, Menstrual, Infertility, Obesity, Hyperinsulinemia, Insulin Resistance, Hirsutism, Cardiovascular Events, Cerebrovascular, Lipid, Metabolic, Disorders, Type 2 diabetes mellitus, T2D, Impaired Glucose Tolerance, Pregnancy-linked Complications, Gestational Diabetes, Venous Thromboembolism, and endometrial cancer.

Introduction

Polycystic ovarian syndrome (PCOS) is a complex genetic disease characterized by heritable reproductive, metabolic, and endocrine abnormalities. It is considered a polygenic pathology that might result from the interaction of susceptible genomic variants and environmental factors. The classical clinical characteristics of PCOS are hyperandrogenism, oligomenorrhea, chronic anovulation, and hyperinsulinemia with insulin resistance (IR). PCOS has been associated with the potential risk of cardiovascular and cerebrovascular events, lipid metabolic disorders, type 2 diabetes mellitus (T2D), impaired glucose tolerance (IGT), pregnancy-linked complications, gestational diabetes, venous thromboembolism, and endometrial cancer. T2D and PCOS significantly overlap in pathophysiological features with insulin resistance as a central driver, and share genetic features between them. Transcription factor 7-like 2 gene (TCF7L2) is the most potent T2D locus reported to date, and it has been replicated in several European Review for Medical and Pharmacological Sciences 2023; 27: 7346-7351

M. AMIN1, L. DEL BOSQUE-PLATA2, C. GRAGNOLI3,4,5

1INSERM, US-14-Orphanet, Paris, France
2National Institute of Genomic Medicine, Nutrigenetics and Nutrigenomic Laboratory, Mexico City, Mexico
3Department of Medicine, Division of Endocrinology, Creighton University School of Medicine, Omaha, NE, USA
4Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, USA
5Molecular Biology Laboratory, Bios Biotech Multi-Diagnostic Health Center, Rome, Italy

Corresponding Author: Claudia Gragnoli, MD, Ph.D; e-mail: claudia.gragnoli@gmail.com
Novel linkage and association of \textit{TCF7L2} variants with PCOS in Italian families

ethnic groups\textsuperscript{7}. Thus, \textit{TCF7L2} was proposed as a possible PCOS-candidate gene. Several associations of \textit{TCF7L2} variants were previously reported with PCOS-metabolic phenotypes\textsuperscript{8}. \textit{TCF7L2} T2D risk alleles have been associated with markers of pancreatic β-cell dysfunction conferring risk for T2D in the general population\textsuperscript{9} and glucose intolerance in women with PCOS and dysglycemia\textsuperscript{10}. The \textit{TCF7L2} PCOS-risk has been associated with decreased insulin sensitivity in women of European ancestry\textsuperscript{8}. The \textit{TCF7L2} T2D risk allele (rs12255372 T) was associated with a trend towards decreased insulin secretion in Greek women with PCOS\textsuperscript{11}. Nevertheless, negative results have also been reported, such as in Turkish women, in which the \textit{TCF7L2} rs7903146 variant was not associated with PCOS\textsuperscript{12}, and in South Brazilian women, in which the \textit{TCF7L2} rs7903146 and rs11196236 were not associated with PCOS\textsuperscript{13}. Thus, more studies are needed.

We recently reported \textit{TCF7L2} linkage to and linkage disequilibrium with T2D and major depressive disorder\textsuperscript{14}. Of note, PCOS is associated not only with T2D\textsuperscript{15} but also with depression and anxiety\textsuperscript{16}. Thus, in the present study, we aimed to investigate whether \textit{TCF7L2} was linked to or in LD (that is, linkage and association) with the PCOS phenotype in peninsular Italian families.

\textbf{Patients and Methods}

In the current study, we studied 212 Italian families ascertained first for familial T2D and then phenotyped for PCOS, by following the PCOS Rotterdam diagnostic criteria (presence of at least two of these three characteristics: chronic anovulation or oligomenorrhea, clinical or biochemical hyperandrogenism, and/or polycystic ovaries)\textsuperscript{7}. Informed consent was obtained from all subjects before participation, and the study followed the Helsinki Declaration guidelines. The BioEthical Committee approved the study.

We genotyped 78 single nucleotide polymorphisms (SNPs) within the \textit{TCF7L2} gene using microarray in the family members after excluding identical twins and subjects with uncertain paternity. The genotyped markers were tested using Pseudomarker\textsuperscript{18} for parametric linkage to and/or LD with PCOS and the dominant model with complete penetrance (D1). Sequentially, the markers were tested for linkage and/or LD using the recessive model with complete penetrance (R1), the dominant model with incomplete penetrance (D2), and the recessive model with incomplete penetrance (R2).

\textit{In Silico Analysis}

Using different \textit{in silico} functional prediction tools, we analyzed the risk variants for their potential role in transcription-factor binding, splicing, and regulatory potential (SNPnexus\textsuperscript{24}, SNP2TFBS\textsuperscript{25}, RegulomeDB\textsuperscript{26}, and mirSNP\textsuperscript{27}).

\textbf{Statistical Analysis}

Genotyping and Mendelian errors were excluded using PLINK\textsuperscript{19}. We considered \textit{p}-values < 0.05 statistically significant. To infer the presence of LD blocks, we calculated the correlation coefficient between the detected risk variants using the LD matrix of the Tuscany Italian population derived from the 1000 Genomes Project (https://www.internationalgenome.org/data-portal/population/TSI).

Also, by analyzing the LD matrix of the above-mentioned Tuscany population, we tested for LD the identified PCOS-risk variant rs7917983 risk allele (C) with the risk allele T of the \textit{TCF7L2}-variant (rs7903146), which was previously extensively studied in PCOS\textsuperscript{20-25}.

\textbf{Results}

We identified a total of 14 SNPs in the \textit{TCF7L2} gene significantly linked to and/or associated with the risk of PCOS (\textit{p} < 0.05) across different models (Table I, Figure 1). All risk variants are intronic, and two variants were in LD block (Set01). All the variants are novel and were not previously reported with PCOS.

\textit{In Silico Findings}

All risk variants reported in our study intersect with active chromatin transcription in the adipose tissue and ovarian tissue (except rs17130183 and rs34855922, which intersect with repressed chromatin in the ovaries) (RegulomeDB\textsuperscript{26-28}).

\textbf{Discussion}

In this study, we report in Italian families the novel linkage and LD (i.e., linkage and association) of 14 variants in the \textit{TCF7L2} gene.
with the risk of PCOS. The TCF7L2 gene is known to confer a risk for T2D in several populations, including the one under study, as well as confer a risk for PCOS. Two variants in our study (rs6585205 and rs112775103) were previously linked to and/or in LD with the risk of T2D, and two variants (rs10885398 and rs74825300) were previously linked to and/or

### Table I. Polycystic Ovarian Syndrome (PCOS) TCF7L2-Risk Single Nucleotide Polymorphisms (SNPs).

<table>
<thead>
<tr>
<th>Model</th>
<th>SNP</th>
<th>Position</th>
<th>Ref</th>
<th>Alt</th>
<th>Consequence</th>
<th>LD block</th>
<th>Reported in PCOS or in a related phenotype?</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2, R1, R2</td>
<td>rs2094405</td>
<td>112955930</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>D1, D2, R1, R2</td>
<td>rs10885398</td>
<td>112956071</td>
<td>A</td>
<td>G</td>
<td>A</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>D1, D2, R1, R2</td>
<td>rs2296784</td>
<td>112963800</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>Intrinsic</td>
<td>Set01</td>
</tr>
<tr>
<td>D1, D2, R1, R2</td>
<td>rs61875109</td>
<td>112969804</td>
<td>C</td>
<td>A</td>
<td>A</td>
<td>Intrinsic</td>
<td>Set01</td>
</tr>
<tr>
<td>R1, R2</td>
<td>rs12573128</td>
<td>112971038</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>D2</td>
<td>rs7917983</td>
<td>112973123</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>R1, R2</td>
<td>rs74825300</td>
<td>113030349</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>D1</td>
<td>rs35198330</td>
<td>113045025</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>R1, R2</td>
<td>rs6585205</td>
<td>113099405</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>D2</td>
<td>rs17130183</td>
<td>113105435</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>D1</td>
<td>rs34855922</td>
<td>113111835</td>
<td>A</td>
<td>G</td>
<td>G</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>D1</td>
<td>rs362943</td>
<td>113128849</td>
<td>G</td>
<td>A</td>
<td>A</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
<tr>
<td>D1, D2</td>
<td>rs112775103</td>
<td>113153279</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>Intrinsic</td>
<td>Independent</td>
</tr>
</tbody>
</table>

1Models: D1: dominant, complete penetrance, D2: dominant, incomplete penetrance, R1: recessive, complete penetrance, R2: recessive, incomplete penetrance; variants highlighted in red were found in our previous study in T2D and MDD (14).

2Polycystic ovarian syndrome, polycystic ovary, PCOS, irregular menses, anovulation, infertility, oligomenorrhea, obesity, metabolic syndrome, insulin resistance, type 2 diabetes, T2D, hyperandrogenism, hirsutism, alopecia.

![](image.png)

**Figure 1.** Parametric Analysis Results of Polycystic Ovarian Syndrome (PCOS) TCF7L2-Risk Single Nucleotide Polymorphisms (SNPs). Legend. For each TCF7L2-risk single nucleotide polymorphisms in PCOS, we present the \(-\log_{10}(p)\) as a function of the significant \((p < 0.05)\) test statistics (Linkage, Linkage Disequilibrium [LD], Linkage, LD [NoLinkage, Linkage|LD, and LD + Linkage) and per inheritance model. D1: dominant, complete penetrance, D2: dominant, incomplete penetrance, R1: recessive, complete penetrance, R2: recessive, incomplete penetrance. The most significant test statistics are underlined.
in LD with the risk of T2D-MDD comorbidity in Italian families, although with different risk alleles. It is possible that these disorders are contributed to by different alleles within the variants influencing expression or gene regulation or that the risk variants potentially comorbid for the disorders are in LD with other pathogenic variants. This illustrates the complexity of reciprocal intertwining between mental and metabolic disorders in their genetic pathogeneses and indicates that risks could be mediated by allele-specific roles or a network of variants acting through epistasis. However, the variant (rs6585205) has also been studied with schizophrenia in the Chinese population, but no association has been found. The pathogenic role played by the TCF7L2-variants reported in our study in the risk of PCOS could not be fully elucidated. The risk allele (C) of the variant (rs7917983) was previously found to have a role in subcutaneous fat mass, potentially relating to obesity and insulin resistance. Within the Tuscany population tested, the same rs7917983 risk allele (C) also highly correlates— that is, it is in LD— with the risk allele T of the TCF7L2-variant (rs7903146), which has been extensively studied in PCOS albeit with inconsistent results. The remaining risk variants, however, have not been previously implicated in any of PCOS-related phenotypes (e.g., irregular menses, anovulation, infertility, oligomenorrhea, obesity, insulin resistance, T2D, hyperandrogenism, hirsutism). Their functional roles, therefore, remain to be discovered. The risk variants intersect with active chromatin in the adipose tissue and thus with potential positive gene expression. This is inconsistent with two other PCOS-risk variants previously associated with TCF7L2 gene downregulation in the adipose tissue of Finnish patients with T2D. This gene expression discrepancy in the adipose tissue could be due to variant-specific differential roles within the framework of LD between the 2 variants, variable population genetic architecture, which would require further genotyping, sequencing, and testing, or likely epistatic interactions across variants, whose functional network-related outcomes should be analyzed as a whole, albeit this is a complex task.

Limitations

This study has been conducted in a homogenous monoethnic population, and it needs to be replicated in other ethnic groups in order to reach more solid conclusions. Furthermore, functional studies are needed to confirm the implication of the TCF7L2 gene and its reported variants in the pathogenesis of PCOS.

Conclusions

Due to the high morbidity and mortality that PCOS confers, it is fundamental to understand its pathogenesis in order to develop strategies for prevention, treatment, and possibly a cure. In this study, further TCF7L2 pleiotropic effects were detected in PCOS. These effects will need to be replicated in other ethnic groups than the one we studied.

Conflict of Interest

The Authors declare that they have no conflict of interests.

Acknowledgements

We thank the families who participated in the study, and we thank Bios Biotech Multi-Diagnostic Health Center, Rome, Italy, for data access and financial, medical, and laboratory staff support.

Ethics Approval

The Bios Ethical Committee approved this study (Prot. PR/Mg/Cg/31708). The Helsinki Declaration guidelines were followed.

Informed Consent

Subjects provided written informed consent prior to participation.

Funding

This publication was supported in part with the funds received under Nebraska Laws 2021, LB 380, Section 109 awarded to C.G. (PI), Creighton University School of Medicine, through the Nebraska Department of Health and Human Services (DHHS). Its contents represent the views of the authors and do not necessarily represent the official views of the State of Nebraska or DHHS.

Authors’ Contribution

C.G. conceived and supervised the project, including statistical analysis and manuscript drafting. L.B.P. helped with literature search, data assignments, and manuscript drafting. M.A. helped with the bioinformatic analysis and manuscript drafting.
Availability of Data and Materials
The study data are available upon reasonable request; they are not publicly available, due to privacy restrictions.

Authors’ Information
C.G. is Professor of Medicine, Chief of Endocrinology, Endowed Puller Chair at Creighton University School of Medicine, Omaha, NE, and Adjunct Professor of Public Health Sciences, Penn State University College of Medicine, Hershey, PA; L.B.P. is a Ph.D. research scientist at the National Institute of Genomic Medicine, Nutrigenetics, and Nutrigenomic Laboratory, Mexico City, Mexico; M.A. has an MD and a Ph.D in Genetics (University of Paris) and is a geneticist at Orphane | INSERM-US14, Paris, France, and member of the executive committee, European Reference Network for the Intellectual Disabilities (ERN-ITHACA) and the International Consortium of Gene Curation Coalition (GenCC).

ORCID ID
C.G: 0000-0002-3873-6617
L.B.P: 0000-0003-3224-378X
M.A.: 0000-0003-2876-0784

References
8) Yalamanchi SK, Sam S, Cardenas MO, Holaday LW, Urbanek M, Dunaif A. Association of fibril-3 and transcription factor-7-like 2 gene variants with metabolic phenotypes in PCOS. Obesity (Silver Spring) 2012; 20: 1273-1278.
Novel linkage and association of TCF7L2 variants with PCOS in Italian families

19) Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81: 559-575.


