Family-based analysis of susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3

Han Zhao¹,²,³, Xinghua Xu¹,²,³, Xiuye Xing¹,²,³, Jianfeng Wang¹,²,³, Lin He⁴,⁵,⁶, Yongyong Shi⁶, Yuhua Shi¹,²,³, Yueran Zhao¹,²,³, and Zi-Jiang Chen¹,²,³,*

¹Center for Reproductive Medicine, Provincial Hospital Affiliated to Shandong University, 324 Jingwu Road, 250021 Jinan, People’s Republic of China ²National Research Center for Assisted Reproductive Technology and Reproductive Genetics, The Key Laboratory for Reproductive Endocrinology of Ministry of Education, 324 Jingwu Road, 250021 Jinan, People’s Republic of China ³Shandong Provincial Key Laboratory of Reproductive Medicine, 324 Jingwu Road, 250021 Jinan, People’s Republic of China ⁴Institutes of Biomedical Sciences, Fudan University, Shanghai, People’s Republic of China ⁵Institute for Nutritional Sciences, Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, 200030 Shanghai, People’s Republic of China ⁶Bio-X Center, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Shanghai Jiao Tong University, 200030 Shanghai, People’s Republic of China

*Correspondence address. Tel: +86-531-85187856; Fax: +86-531-87068226; E-mail: chenzijiang@hotmail.com

Submitted on August 13, 2011; resubmitted on October 6, 2011; accepted on October 13, 2011

BACKGROUND: Polycystic ovary syndrome (PCOS) is a complex endocrine-metabolic disorder. A previous genome-wide association study (GWAS) identified five single nucleotide polymorphisms (SNPs) which were independently associated with PCOS in Han Chinese. To overcome population stratification, a family-based analysis was conducted to validate whether these five SNPs are associated with PCOS.

METHODS: A total of 276 family trios (828 participants) having a proband with PCOS were included in the family-based study. The transmission disequilibrium test (TDT) was used to analyze the association between PCOS and five SNPs rs13429458, rs12478601, rs13405728, rs10818854 and rs2479106 in three susceptible loci 2p16.3, 2p21 and 9q33.3.

RESULTS: A positive association was observed for the SNP rs13429458 (P = 3.74 × 10⁻⁵).

CONCLUSIONS: TDT confirms that SNP rs13429458, in the THADA gene, is significantly associated with risk of PCOS. This family-based analysis enhances our previous case–control GWAS and provides further support for the role of susceptibility loci in PCOS.

Key words: polycystic ovary syndrome / transmission disequilibrium test / single-nucleotide polymorphism / genome wide association study

Introduction

Polycystic ovary syndrome (PCOS) is a common and complex endocrine disorder in women of reproductive age, with a prevalence of 6–8% (Ehrmann et al., 1999; Goodarzi and Azziz, 2006). The syndrome is characterized by the presence of two or more of the following features: chronic oligo-ovulation or anovulation, androgen excess and polycystic ovaries [Rotterdam European Society for Human Reproduction and Embryology/American Society for Reproductive Medicine (ESHRE/ASRM) sponsored by PCOS Consensus Workshop Group, 2004]. Additionally, PCOS is associated with important endocrine-metabolic derangements including dyslipidemia, atherosclerosis, insulin resistance and type 2 diabetes (T2D) (Wild et al., 2000; Carmina, 2009; Kandarakis et al., 2009).

Familial aggregation (Legro et al., 1998; Vink et al., 2006; Menke and Strauss, 2007; Urbanek, 2007) and heritable tendencies of PCOS have long been recognized for this complex and heterogeneous disorder. Several candidate genes have been identified for association studies. Previously, our genome-wide association study (GWAS) also identified three susceptibility loci, 2p16.3, 2p21 and 9q33.3, which were associated with PCOS in Han Chinese (Chen et al., 2011). However, the mode of inheritance for PCOS and the molecular mechanisms underlying PCOS have not been fully clarified. Several factors would complicate the molecular genetics of PCOS, such as population stratification, environmental factors and genetic heterogeneity. To avoid those impacting factors, it is crucial to conduct family-based analysis using the transmission disequilibrium test (TDT) to explore the putative contribution of candidate genes.
The TDT is proposed as a family-based association test for the presence of genetic linkage between a genetic marker and a trait. A specificity of the TDT is that it will detect genetic linkage only in the presence of genetic association. Genetic association can be caused by population structure, while genetic linkage will not be affected, which makes the TDT robust to the presence of population structure.

The three PCOS genome-wide associated susceptibility loci 2p16.3, 2p21 and 9q33.3 mapped to the genomic areas of three genes LHCGR, THADA and DENND1A, respectively. LHCGR encodes a G-protein-coupled receptor for LH and HCG. In the ovary, LHCGR is expressed in granulosa cells at the later stages of pre-ovulatory follicle development. The mid-cycle LH surge triggers ovulation. Single nucleotide polymorphism (SNP) rs13405728 was the most significant at region 2p16.3 and is located in an intron of LHCGR. THADA (thyroid adenoma associated) was identified from chromosomal aberrations containing this genomic region in benign thyroid adenomas (Drieschner et al., 2006). Several SNPs in this region reached genome-wide association significance. After conditional logistic regression analysis, two SNPs (rs13429458 and rs12478601), both located in the gene THADA, were shown to be independently associated with PCOS. DENND1A encodes a domain differentially expressed in normal and neoplastic cells (DENN) that can bind to type-I tumor necrosis factor receptor type 1 (TNFRI) as a negative regulator through a death-domain-death domain interaction (Del Villar and Miller, 2004). Conditional logistic regression analysis showed two independent associations, rs10818854 and rs2479106, both within DENND1A, were associated with PCOS.

In the current study, to further investigate the relationship between the five SNPs mentioned earlier (rs13429458, rs12478601, rs13405728, rs10818854 and rs2479106, independently associated with PCOS) and the pathogenesis of PCOS, a family-based analysis was performed in 276 family trios with PCOS to assess linkage and association between PCOS and the three loci that might provide a better understanding of the contribution of GWAS results to PCOS.

### Materials and Methods

#### PCOS families

A set of 276 Han Chinese family trios consisting of mothers, fathers and offspring affected with PCOS, 828 participants in total, were recruited from the Center for Reproductive Medicine, Provincial Hospital Affiliated to Shandong University, during the period from July 2007 to February 2011. All the PCOS probands were of Han Chinese origin and independent from our previous cases of PCOS in the GWAS (Chen et al., 2011). PCOS was diagnosed according to the 2003 Rotterdam criteria (Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004) with at least two of the following three features: oligoamennorhea or amenorrhea, clinical or biochemical hyperandrogenism and polycystic ovaries on ultrasound. Transvaginal ultrasound was used to detect polycystic ovaries, defined as the presence of at least 1 ovary 10 cm³ or containing at least 12 follicles 2–9 mm in diameter. Ultrasonic examination was performed rectally if subjects were virginal. Androgen excess was defined on the basis of hirsutism (Ferriman–Gallwey score ≥6) or elevated circulating total testosterone ≥60 ng/dl (Ferriman and Gallwey, 1961; Shi et al., 2008). Other related diseases, such as adrenal congenital hyperplasia, Cushing syndrome and androgen-secreting tumors were excluded. This study was approved by the Institutional Review Board of Shandong University. Written informed consent was obtained from all participants.

#### SNP genotyping

Genomic DNA was extracted from peripheral blood using a QIAamp DNA mini kit (QIAGEN, Hilden, Germany) according to the manufacturer’s protocol. All SNPs were amplified using PCR and the respective primers were provided in Table Ⅰ. The PCR was carried out under the following conditions: initial denaturation at 95 °C for 5 min followed by 35 cycles of denaturation at 94 °C for 30 s, annealing for 30 s at 60 °C, extension at 72 °C for 1 min and finally 72 °C for 7 min. The PCR products were first analyzed by agarose gel electrophoresis and then sequenced on an automated sequencer (ABI PRISM 310; Applied Biosystems, Foster City, CA, USA). All typed SNPs were checked for departure from Hardy–Weinberg equilibrium (HWE) in samples of affected probands and their parents.

#### The TDT

Details of TDT were described in our previous study (Xu et al., 2011). Descriptive statistics for individual SNPs, including minor allele frequency (MAF) and HWE were obtained using Haplovlew 4.2 (Barrett et al., 2005). Then association between the five SNPs and PCOS was tested by the TDT analysis which was also performed using Haplovlew 4.2 (Barrett et al., 2005). Statistical significance was considered at the two-tailed P level of 0.05. The TDT is a valid χ² test statistic for the linkage hypothesis, regardless of population history. In the TDT test, by collecting unrelated family trios with PCOS, we analyzed the difference between the probability of parents-to-offspring transmission and the hypothesis of no association (probability of transmission 50%). If a discrepancy exists, the reason would be an association between the polymorphisms and PCOS. For the positive markers, we performed multiple testing using the Haplovlew permutation function for further correction. To attenuate the impact of a mother who possibly had PCOS, the TDT was implemented again after excluding those families which included a mother with irregular menstruation.

#### Results

#### Clinical and metabolic features

As shown in Table II, the mean (±SD) age of 276 women affected by PCOS was 27.45 ± 4.02 years and mean BMI was 25.03 ± 4.39 kg/m².

### Table I PCR primers used for amplification of the five PCOS susceptibility SNPs.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs13429458</td>
<td>5'-CAGCGGTATGATTCTCGTGTG-3' (forward)</td>
</tr>
<tr>
<td></td>
<td>5'-GCTAAAAATCTCACCTGGGACG-3' (reverse)</td>
</tr>
<tr>
<td>rs12478601</td>
<td>5'-AGACTCAGTAGATGCACATCAT-3' (forward)</td>
</tr>
<tr>
<td></td>
<td>5'-TTACCTGCAAACCTTCGAATG-3' (reverse)</td>
</tr>
<tr>
<td>rs13405728</td>
<td>5'-GTGGTCTTTACTCCTAGACACATG-3' (forward)</td>
</tr>
<tr>
<td></td>
<td>5'-CCATCCACATCTCAGTTCAATATC-3' (reverse)</td>
</tr>
<tr>
<td>rs1818854</td>
<td>5'-CAAACCCAGCTGATCGAAT-3' (forward)</td>
</tr>
<tr>
<td></td>
<td>5'-GTGTGAAATCATAGACCGACAC-3' (reverse)</td>
</tr>
<tr>
<td>rs2479106</td>
<td>5'-GACAGCCACTCAAGAACAG-3' (forward)</td>
</tr>
<tr>
<td></td>
<td>5'-AAGCCACCATCCAGTCAC-3' (reverse)</td>
</tr>
</tbody>
</table>
The mean level of total testosterone was $60.97 \pm 33.75$ ng/dl. Among the PCOS probands ($n = 266$), the mean levels of fasting glucose and fasting insulin were 5.47 mmol/l (5.47 ± 1.29 in mean ± SD) and 11.01 mmol/l (11.01 ± 6.91 in mean ± SD), respectively.

**TDT analysis in trios with PCOS**

For all markers typed, no deviations from HWE were observed ($P > 0.05$). Table III lists the results of TDT for each SNP. A significant difference in transmission was found for rs13429458 (transmitted/non-transmitted $= 102:51$; $\chi^2 = 17.000; P = 3.74 \times 10^{-5}$). SNP rs12478601 also had a mildly associated trend with PCOS (transmitted/non-transmitted $= 127:100$; $\chi^2 = 3.211; P = 0.073$). With regard to the other markers, none of them showed evidence for association with PCOS but risk alleles were all over transmitted. The genotyping results of the five SNPs indicated that the MAFs among family-based analysis were consistent with our previous GWAS study.

Moreover, to reduce the possible maternal PCOS effect, 77 families whose mother had a history of irregular menstruation were removed. We performed TDT again on the remaining 199 families and found consistent results (Supplementary data, Table SI). SNP rs13429458 was still significant (transmitted/non-transmitted $= 74:37; \chi^2 = 12.330; P = 4.0 \times 10^{-4}$).

**Correction for multiple testing**

Haploview permutation function was used for multiple testing on positive marker rs13429458. We performed 10,000 permutations, and the rs13429458 association remained statistically significant after correction ($P$ corrected $= 0.0003$).

**Discussion**

In our previous GWAS, we found significant differences between cases with PCOS and controls in the allele-frequency distributions of SNP rs13429458 (2p21), rs12478601 (2p21), rs13405728 (2p16.3), rs10818854 (9q33.3) and rs2479106 (9q33.3) (Chen et al., 2011). Although the cases and controls were matched carefully by geography, it was difficult to classify Chinese samples because the Han Chinese have an extremely long and complex demographic history. One way to overcome population stratification is to use family-based analysis.

In this family-based study, we conducted the TDT and tested the above five markers. SNP rs13429458 showed significant transmission difference, which validated our PCOS GWAS, and implies that
rs13429458 is a robust risk marker for PCOS. SNP rs13429458 is located in the intron of THADA on chromosome 2p21. THADA was initially identified in thyroid adenomas by chromosomal rearrangements, where THADA was disrupted and fused to an intron of peroxisome proliferator-activated receptor gamma (PPARG) (Drieschner et al., 2006). The over-transmission of SNP rs13429458 indicates that gene THADA should be a new potential candidate for PCOS. THADA (a different SNP rs7578597) was detected as a susceptibility gene for T2D in a large European GWAS (Zeggini et al., 2008); however, later T2D studies in Chinese (Hu et al., 2009), Finnish (Stančáková et al., 2009) and Indian (Sanghera et al., 2009) patients failed to replicate this finding. In a Chinese population, THADA had no correlation with T2D but it was associated with abnormal 2h insulin levels in glucose tolerance tests (Hu et al., 2009). These results provide useful information for follow-up functional studies, which may reveal that THADA participates in the pathogenesis of both PCOS and T2D.

With respect to the other four SNPs, they are not statistically associated with PCOS by TDT. One limitation of this study is the relatively small sample size (276 families). We estimated the statistical power of the TDT sample size for the four SNPs (Supplementary data, Table SII). The statistical power of this sample set is only 50–63%, with α level 0.05. To reach 90% power, for example, 552 families are needed for rs12478601 and 759 families for rs2479106. Moreover, there were not so many heterozygous parents (possessing two different alleles for a single trait) of each marker, which could affect the results of the TDT.

Although earlier studies (Legro et al., 1998; Kahsar-Miller et al., 2001; Ehrmann, 2005; Urbanek, 2007) have established that there is familial clustering of PCOS and a genetic predisposition, evidence for the involvement of specific genes or regions has been difficult to confirm (Franks et al., 2001; Urbanek and Spielman, 2002; Ehrmann, 2005). The family-based TDT following our case–control study (Chen et al., 2011) indicates well that the positive association is not likely attributable to population stratification. As our findings implicate PPARG partially identified in thyroid adenomas by chromosomal rearrangements, we especially thank all the families for participating in this study.

Acknowledgements
We thank Li You, Qingmei Zheng, Yuehong Bian, Di Wu and Changming Zhang of Provincial Hospital Affiliated to Shandong University for their sample collecting and technical guidance. We especially thank all the families for participating in this study.

Funding
This research was supported by the National Natural Science Foundation of China (81000238, 81070461, 8100236, 30973170) and National Basic Research Program of China (973 program) (2010CB94500, 2007CB947403).

References
Kandarakis E, Christakou C, Diamanti-Kandarakis E. Metabolic syndrome and polycystic ovary syndrome...and vice versa. Arq Bras Endocrinol Metabol 2009;53:227–237.

Supplementary data
Supplementary data are available at http://humrep.oxfordjournals.org/.

Authors’ roles
H.Z. partially designed, executed and drafted the manuscript; X.X., X.X., Y.S. and Y.Z. collected all clinical data and blood samples; J.W. performed genotyping; L.H. and Y.S. performed TDT analysis; Z.-J.C. designed, supported the study and revised the manuscript.


Urbanek M. The genetics of the polycystic ovary syndrome. *Nat Clin Pract Endocrinol Metab* 2007; **3:**103–111.

Urbanek M, Spielman RS. Genetic analysis of candidate genes for the polycystic ovary syndrome. *Curr Opin Endocrinol Diabetes* 2002; **9:**492–501.

Vink J, Sadrzadeh SM, Lambalk CB, Boomsma DI. Heritability of polycystic ovary syndrome (PCOS) in a Dutch twin-family study. *J Clin Endocrinol Metab* 2006; **91:**2100–2104.

