Human Reproduction, Vol.28, No.2 pp. 538–544, 2013
Advanced Access publication on December 2, 2012 doi:10.1093/humrep/des424

human reproduction

ORIGINAL ARTICLE *Reproductive genetics*

Genotype–phenotype correlations of PCOS susceptibility SNPs identified by GWAS in a large cohort of Han Chinese women

Linlin Cui[1,2,3,4,†], Han Zhao[1,2,3,4,†], Bo Zhang[5,†], Zhongyu Qu[6], Jiayin Liu[7], Xiaoyan Liang[8], Xiaoming Zhao[9], Junli Zhao[10], Yingpu Sun[11], Peng Wang[1,2,3,4], Tao Li[1,2,3,4], Yuhua Shi[1,2,3,4], and Zi-Jiang Chen[1,2,3,4,*]

[1]Center for Reproductive Medicine, Provincial Hospital Affiliated to Shandong University, 324 Jingwu Road, Jinan 250021, China [2]National Research Center for Assisted Reproductive Technology and Reproductive Genetics, Jinan 250021, China [3]The Key laboratory for Reproductive Endocrinology of Ministry of Education, Jinan 250021, China [4]Shandong Provincial Key Laboratory of Reproductive Medicine, 324 Jingwu Road, Jinan 250021, China [5]Reproductive Medicine Center, The Maternal and Child Health Hospital of Guangxi Zhuang Autonomous Region, Nanning 530003, China [6]Department of Ultrasonography, Provincial Hospital Affiliated to Shandong University, Jinan 250100, China [7]Department of Reproductive Medicine, The First Affiliated Hospital with Nanjing Medical University, Nanjing 210129, China [8]Department of Reproductive Medicine, The Sixth Affiliated Hospital of Sun Yat-sen University, GuangZhou 510080, China [9]Center for Reproductive Medicine, Renji Hospital Affiliated to Shanghai Jiaotong University, Shanghai 200127, China [10]Center for Reproductive Medicine, Affiliated Hospital of Ningxia Medical University, Yinchuan 750003, China [11]Center for Reproductive Medicine, The First Affiliated Hospital of Zhengzhou University, Zhengzhou 450052, China

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

Submitted on August 21, 2012; resubmitted on November 6, 2012; accepted on November 9, 2012

STUDY QUESTION: Are there any correlations between the phenotypes of polycystic ovary syndrome (PCOS) and the genotypes of the PCOS susceptibility single nucleotide polymorphisms (SNPs) in THADA, DENND1A and LHCGR?

SUMMARY ANSWER: The PCOS susceptibility genes, THADA and DENND1A, carry risk alleles that are associated with endocrine and metabolic disturbances in patients with PCOS.

WHAT IS KNOWN ALREADY: PCOS is a heterogeneous endocrinopathy characterized by oligo-anovulation, hyperandrogenism and polycystic ovaries. In a previous genome-wide association study, the SNP variants rs13429458, rs12478601, rs2479106, rs10818854 and rs13405728 in the THADA, DENND1A and LHCGR genes were identified as being independently associated with PCOS. The aim of this study was to identify any additional correlations between the phenotypes of PCOS and genotypes of the five SNPs described in the previous study.

STUDY DESIGN, SIZE, DURATION: In the present cross-sectional study, a total of 1731 PCOS patients and 4964 controls were enrolled.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Patients were diagnosed according to Rotterdam criteria. Clinical information was collected from the patients and controls. Endocrine and metabolic parameters were evaluated for phenotype–genotype correlation analyses.

MAIN RESULTS AND THE ROLE OF CHANCE: Using a recessive model, the AA group for rs13429458 in THADA was associated with increased luteinizing hormone (LH) ($P < 0.01$) and testosterone (T) ($P = 0.02$) levels in subjects with PCOS; the LH/follicle-stimulating hormone ratio was also higher in the AA group ($P < 0.01$). Also using a recessive model, the CC genotype of rs12478601, also in THADA, was associated with increased levels of low-density lipoprotein ($P = 0.02$). Using a dominant model, the GG + AG group for rs2479106 in DENND1A was associated with elevated serum insulin levels 2 h after a glucose load in the patients with PCOS ($P = 0.02$). All of the comparisons were adjusted for age and BMI.

[†] These authors contributed equally to this work.

LIMITATIONS, REASONS FOR CAUTION: The relatively younger age of the participants may represent a considerable bias when evaluating metabolic alterations as a function of different genotypes, as significant metabolic disturbances may emerge later in life. Furthermore, the sample sizes of several sub-genotype groups were relatively small; to some extent this limited the statistical power of the analysis.

WIDER IMPLICATIONS OF THE FINDINGS: The PCOS susceptibility genes, *THADA* and *DENND1A*, carry risk alleles that are associated with endocrine and metabolic disturbances in PCOS patients of Han Chinese descent. The findings have shown genuine heterogeneity, stratified on the basis of both clinical findings and genotypes. Replication of these results is expected in other ethnic groups.

STUDY FUNDING/COMPETING INTEREST(S): This research was supported by the National Basic Research Program of China (973 program) (2010CB945002, 2012CB944700), the National Natural Science Foundation of China (81000238, 81070461, 81000236, 30973170), the Graduate Independent Innovation Foundation of Shandong University (GIIFSDU) (21300070613242, 21300070613246), the Science Research Foundation item of no-earnings health vocation (201002013) and the National Key Technology Research and Development Program (2011BAI17B00). There are no competing interests.

Key words: PCOS / phenotype / THADA / DENND1A

Introduction

Polycystic ovary syndrome (PCOS) is a common endocrinopathy with a prevalence ranging from 6 to 8% in women of reproductive age (Azziz *et al*., 2004). It is characterized by hyperandrogenism (HA), menstrual irregularity and polycystic ovarian morphology, and commonly leads to anovulatory infertility and metabolic complications, including metabolic syndrome, type 2 diabetes (T2D) and cardiovascular disease (CVD; Legro *et al*., 2005; de Groot *et al*., 2011). PCOS is a multifactorial disease caused by complex interactions between environmental factors and predisposing polygenic backgrounds (Vink *et al*., 2006). However, the heterogeneity and uncertain etiology of PCOS have made it difficult to identify candidate genes.

Genome-wide association studies (GWASs) offer a potentially powerful approach to identify the associations between millions of single nucleotide polymorphisms (SNPs) and specific traits or disorders (Hirschhorn and Daly, 2005). Our first GWAS in a Han Chinese cohort revealed three novel susceptibility loci for PCOS, based on the Rotterdam criteria, on chromosomes 2p16.3 (rs13405728), 2p21 (rs13429458, rs12478601) and 9q33.3 (rs10818854, rs2479106). The associated linkage disequilibrium (LD) blocks contained the susceptibility genes *LHCGR*, *THADA* and *DENND1A*, respectively (Chen *et al*., 2011), which were partially confirmed in some recent studies in European cohorts (Lerchbaum *et al*., 2011; Eriksen *et al*., 2012; Goodarzi *et al*., 2012; Welt *et al*., 2012). The *LHCGR* gene, which is located on chromosome 2p16.3, encodes the luteinizing hormone/choriogonadotropin receptor (LHCGR) belonging to the family of G-protein coupled receptors (Atger *et al*., 1995). It is expressed in a variety of tissues, including the ovary, testis and several other non-gonadal tissues (Rahman and Rao, 2009). The thyroid adenoma-associated gene (*THADA*), mapping to chromosome 2p21, was first identified as a target gene in thyroid benign tumors (Rippe *et al*., 2003). A recent GWAS reported that *THADA* is a novel T2D-associated gene that affects pancreatic beta-cell function (Zeggini *et al*., 2008). *DENND1A* is located on chromosome 9q33.3 and encodes a domain of DENN, which is differentially expressed in normal and neoplastic cells and functions as a guanine nucleotide exchange factor for the early endosomal small GTPase Rab35 (Marat and McPherson, 2010).

According to the Rotterdam criteria, four PCOS subtypes have been defined, including Subtype A [the presence of oligo-anovulation (OA) and polycystic ovaries (PCO) without HA], Subtype B (the presence of HA and PCO without OA), Subtype C (the presence of HA and OA with normal sonographic appearance of the ovaries) and Subtype D (the presence of OA, HA and PCO). Different pathophysiological features have been demonstrated between the subtype groups (Yilmaz *et al*., 2011; Panidis *et al*., 2012). However, the influence of the susceptibility SNPs identified in the previous GWAS on the specific subtypes and their associations with endocrine and metabolic characteristics remain unknown.

In this study, clinical data of the patients tested in the first GWAS were collected to identify correlations between their genotypes and phenotypes. Key features, including endocrine and metabolic parameters, were analyzed to determine their potential genetic associations.

Materials and Methods

Subjects

The study cohort consisted of 1731 PCOS patients with complete clinical data. Participants originated from our first GWAS and were recruited at the reproductive medical center of Shandong Provincial Hospital affiliated to Shandong University. All of them were unrelated individuals of reproductive age with no hormonal therapy for at least three months prior to the test.

The diagnosis of PCOS was based on the Rotterdam Consensus proposed in 2003 (Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group, 2004). OA was determined by a menstrual cycle more than 35 days in length or a history of $\leq$8 menstrual cycles in a year (Azziz *et al*., 2004). Polycystic ovarian (PCO) morphology was determined when $\geq$12 follicles measuring 2–9 mm in diameter were scanned in one or both ovaries or the ovarian volume was above 10 ml. HA was confirmed if there was evidence of hyperandrogenemia and/or hirsutism. Patients were grouped according to the four subtypes introduced in the revised consensus of Rotterdam (described in Table I). The diagnosis of PCOS was made only when the other etiologies for hyperandrogenemia and ovulatory dysfunction were excluded, i.e. congenital adrenal

Table I Subtypes of the PCOS subjects.

	Number of cases (%)
OA	1460 (84.3)
HA	1100 (63.5)
PCO	1519 (87.8)
Subtypes	
A (OA + PCO)	631 (36.5)
B (HA + PCO)	271 (15.7)
C (HA + OA)	212 (12.2)
D (OA + HA + PCO)	617 (35.6)

OA, oligo-/anovulation; HA, hyperandrogenism; PCO, polycystic ovary.

hyperplasias, 21-hydroxylase deficiency, androgen-secreting tumors, Cushing's syndrome, thyroid disease and hyperprolactinemia (Chen *et al.*, 2011).

A total of 4964 age-matched healthy women who were referred for routine physical examination or tubal factor infertility were enrolled. All subjects in the control group had regular menstrual cycles (26–35 days) and normal ovarian morphology (antral follicle counts <12 in each ovary). Total testosterone levels and hirsutism scores were also evaluated for exclusion of HA.

The study was approved by the Institutional Review Board of Reproductive Medicine of Shandong University and written informed consent was obtained from all participants.

Clinical and biochemical measurements

Medical-history and anthropometric data (height, weight) were obtained from PCOS patients during a visit to the clinic. All patients underwent transvaginal ultrasound scanning for PCO and modified Ferriman–Gallwey scoring for the identification of hirsutism (Hatch *et al.*, 1981). Subjects were deemed hirsute if the score was six or higher (Ferriman and Gallwey, 1961). The body mass index (BMI) was calculated using the following formula: weight (kg)/height $(m)^2$.

After an overnight fast, blood sampling was performed during the early follicular phase (between 2 and 5 days of the menstrual cycle) for the detection of hormones including follicular-stimulating hormone (FSH), luteinizing hormone (LH), total testosterone (T) and prolactin (PRL) by chemiluminescence immunoassays, all with intra- and inter-assay coefficients of variation <10%. Hyperandrogenemia was defined as a serum total testosterone above 60 ng/dl (Shi *et al.*, 2007).

Each subject underwent a 75-g oral glucose tolerance test (OGTT). Plasma glucose levels at 0 min and 2 h after OGTT were measured using the oxidase method and insulin levels were measured by chemiluminescence immunoassays. The homeostasis model assessment (HOMA) of insulin resistance was derived by the calculation: fasting plasma glucose (mmol/l) * fasting insulin (μIU/ml)/22.5 (Matthews *et al.*, 1985).

Serum cholesterol (CHOL), triglycerides (TG), low-density lipoprotein (LDL) and high-density lipoprotein (HDL) were evaluated in fasting blood samples using the precipitation and enzymatic methods.

SNP determination

Determination of gene variants in *THADA* (rs13429458, rs12478601), *DENND1A* (rs2479106, rs10818854) and *LHCGR* (rs13405728) was performed according to the previous literature regarding PCOS susceptibility loci (Chen *et al.*, 2011).

Genomic DNA extraction was performed by a standard process using Flexi Gene DNA kits (Qiagen). The five SNPs selected were analyzed in genome-wide genotyping using the Affymetrix Genome-Wide Human SNP Array 6.0 and replicated using the ligation detection reaction in our previous GWAS. The success rate of genotyping for all SNPs was above 96%, and the genotype distributions obeyed Hardy–Weinberg equilibrium.

Statistical analysis

The allele frequency and genotype differences were calculated by PLINK (v.1.05, http://pngu.mgh.harvard.edu/purcell/plink). The case–control genetic power was calculated by Genetic Power Calculator (http://pngu.mgh.harvard.edu/~purcell/gpc/). A Bonferroni corrected *P*-value of 0.003 (15 tests in total) was set as the threshold in association with tests of genetic models. Clinical data analysis was performed using Statistical Package for the Social Sciences for Windows (version 13.0; SPSS Inc., Chicago, IL, USA). Considering the small numbers in the homozygous minor allele groups, an appropriate genetic model was adopted for each SNP and the Mann–Whitney test was used to determine any significant differences for between-group comparisons, with data presented as the median (and the interquartile range) because the distribution remained non-normal even after log or square root transformation. Analysis of covariance was used for age and BMI adjustment. A *P*-value <0.05 was considered statistically significant in genotype–phenotype analyses.

Results

The clinical features of the PCOS cohort are shown in Supplementary data, Table SI. A total of 1731 PCOS subjects, a subgroup of patients in our first GWAS, were enrolled due to the availability of detailed clinical data. The average age of the subjects was 28 years, with a mean BMI of 24.46 kg/m^2. A total of 631 patients presented with Subtype A (36.5%), 271 patients presented with Subtype B (15.7%), 212 patients presented with Subtype C (12.2%) and 617 patients presented with Subtype D (35.6%; Table I).

Associations between the rs13429458, rs12478601, rs2479106, rs10818854 and rs13405728 variants and PCOS identified in the previous GWAS were verified in the present cohort (Supplementary data, Table SII; Lerchbaum *et al.*, 2011). The allele frequencies of the five SNPs in the different sub-phenotypic groups are presented in Table II. The rs13429458 SNP was associated with all four subtypes ($P = 1.00E-03$, $P = 1.30E-03$, $P = 5.00E-03$ and $P = 1.75E-05$, respectively). For rs12478601, rs2479106 and rs10818854, statistically significant differences were found for Subtypes A ($P = 5.78E-06$, $P = 9.08E-05$ and $P = 1.68E-06$), B ($P = 9.00E-04$, $P = 2.67E-02$ and $P = 1.29E-02$) and D ($P = 1.80E-07$, $P = 3.60E-03$ and $P = 6.67E-05$). For rs13405728, there was a significant association with Subtype A ($P = 5.00E-04$), C ($P = 2.00E-04$) and D ($P = 2.00E-04$).

The genetic models of the five SNPs were analyzed (Supplementary data, Table SIII) and significant differences were observed for the additive and dominant models. Associations were also identified for the recessive models of rs13429458, rs12478601 and rs13405728. Given that statistical power could be limited by a small number of homozygous minor alleles and thus result in false-positive relationships, the additive model was not taken into account. Further genotype–phenotype correlation analyses were performed using the recessive models of rs13429458, rs12478601 and rs13405728, and the dominant

Table II Allele frequency distribution among different sub-phenotypes.

SNP	MAF (case/control)	P	OR
rs13429458(C/**A**)			
Subtype A ($n = 631$)[a]	0.146/0.185	1.00E−03	1.33 (1.12–1.57)
Subtype B ($n = 271$)[a]	0.128/0.185	1.30E−03	1.55 (1.18–2.03)
Subtype C ($n = 212$)[a]	0.130/0.185	5.00E−03	1.52 (1.13–2.05)
Subtype D ($n = 617$)[a]	0.134/0.185	1.75E−05	1.47 (1.23–1.76)
rs12478601(T/**C**)			
Subtype A ($n = 631$)[a]	0.227/0.291	5.78E−06	1.40 (1.21–1.62)
Subtype B ($n = 271$)[a]	0.222/0.291	9.00E − 04	1.45 (1.16–1.80)
Subtype C ($n = 212$)	0.265/0.291	2.45E − 01	1.14 (0.91–1.43)
Subtype D ($n = 617$)[a]	0.217/0.291	1.80E − 07	1.48 (1.28–1.72)
rs2479106(**G**/A)			
Subtype A ($n = 631$)[a]	0.273/0.223	9.08E−05	1.31 (1.15–1.51)
Subtype B ($n = 271$)[a]	0.265/0.223	2.67E−02	1.26 (1.03–1.54)
Subtype C ($n = 212$)	0.252/0.223	1.56E−01	1.18 (0.94–1.48)
Subtype D ($n = 617$)[a]	0.261/0.223	3.60E−03	1.24 (1.07–1.42)
rs10818854(A/G)			
Subtype A ($n = 631$)[a]	0.130/0.087	1.68E−06	1.57 (1.30–1.89)
Subtype B ($n = 271$)[a]	0.121/0.087	1.29E−02	1.44(1.08–1.93)
Subtype C ($n = 212$)	0.111/0.087	1.07E−01	1.30 (0.94–1.80)
Subtype D ($n = 617$)[a]	0.123/0.087	6.67E−05	1.47 (1.22–1.78)
rs13405728(G/A)			
Subtype A ($n = 631$)[a]	0.191/0.238	5.00E−04	1.32 (1.13–1.54)
Subtype B ($n = 271$)	0.201/0.238	6.51E−02	1.24 (0.99–1.56)
Subtype C ($n = 212$)[a]	0.158/0.238	2.00E−04	1.67 (1.27–2.19)
Subtype D ($n = 617$)[a]	0.187/0.238	2.00E−04	1.36 (1.16–1.59)

MAF, minor allele frequency.
Subtype A: OA (oligo-/anovulation) + PCO(polycystic ovary); Subtype B: HA (hyperandrogenism) + PCO; Subtype C: HA + OA; Subtype D: OA + HA + PCO.
[a]Statistical significance was found between patients and control subjects ($P < 0.05$).

Table III The association of phenotype and genotype of rs13429458.[a]

	AA ($n = 1206$)	CA + CC ($n = 411$)	P
Age (year)[b]	28 (5)	28 (4)	0.02
BMI (kg/m^2)	24.62 (5.85)	24.46 (6.06)	0.50
mFG	1 (2)	1 (2)	0.89
FSH (IU/l)	6.25 (2.08)	6.14 (2.05)	0.61
LH (IU/l)[b]	9.25 (7.45)	8.37 (6.2)	0.02
PRL (ng/dl)	15.58 (11.03)	15.56 (12.2)	0.30
T (ng/dl)[b]	59.42 (33.07)	56.32 (30.04)	0.03
LH/FSH[b]	1.47 (1.29)	1.33 (0.99)	0.02
Fasting glucose (mmol/l)	5.12 (0.89)	5.16 (0.83)	0.15
2 h-Glucose (mmol/l)	6.37 (2.26)	6.15 (2.01)	0.06
Fasting insulin (μIU/ml)	10.03 (8.05)	9.99 (7.67)	0.88
2 h-Insulin (μIU/ml)	54.28 (56.23)	44.79 (56.47)	0.14
HOMA	2.21 (2.00)	2.24 (1.99)	0.84
CHOL (mmol/l)	4.41 (1.31)	4.48 (1.35)	0.31
TG (mmol/l)	1.07 (0.9)	1.09 (0.92)	0.74
LDL (mmol/l)	2.33 (0.94)	2.23 (1.11)	0.13
HDL (mmol/l)	1.17 (0.51)	1.22 (0.58)	0.28

Data are presented as the median (interquartile range).
BMI (kg/m^2), body mass index; mFG, modified Ferriman–Gallwey hirsutism score; FSH (IU/l), follicle-stimulating hormone; LH (IU/l), luteinizing hormone; PRL (ng/dl), prolactin; T (ng/dl), testosterone; Glucose (mmol/l); Insulin (μIU/ml); HOMA, homeostasis model assessment; CHOL (mmol/l), cholesterol; TG (mmol/l), triglyceride; LDL (mmol/l), low-density lipoprotein; HDL (mmol/l), high-density lipoprotein.
[a]The Mann–Whitney test was selected.
[b]Statistical significance was found when compared with the other group ($P < 0.05$).

models of rs2479106 and rs10818854, which were proved to be more significant than the other two models (dominant and recessive; Lerchbaum *et al.*, 2011; Simonis-Bik *et al.*, 2010). All tested powers of selected models were calculated by the Genetic Power Calculator and reached 0.8 at $P = 0.001$.

rs13429458 (*THADA*)

Using a recessive model, patients with the AA genotype of rs13429458 were older and had higher serum LH and T levels than the patients with the AC and CC genotypes ($P = 0.02$ and $P = 0.03$, respectively; Table III). The LH/FSH ratio was also higher in the patients with the AA genotype ($P = 0.02$; Table III). The statistical significance was independent of age and BMI ($P < 0.01$, $P = 0.02$ and $P < 0.01$, respectively, after adjusting for age and BMI).

rs12478601 (*THADA*)

Using a recessive model for rs12478601, the average age of the subjects with PCOS with the CC genotype were significantly higher ($P = 0.01$; Table IV), and these patients had elevated serum LDL levels ($P = 0.02$; Table IV). The LDL levels remained increased in

Table IV The association of the phenotype and genotype of rs12478601.[a]

	CC ($n = 959$)	TC + TT ($n = 631$)	P
Age (year)[b]	28 (5)	28 (4)	0.01
BMI (kg/m^2)	24.94 ± 4.09	24.65 ± 4.05	0.17
mFG	1 (2)	0 (3)	0.38
FSH (IU/l)	6.26 (2.11)	6.28 (2.04)	0.86
LH (IU/l)	9.12 (7.39)	9.04 (6.73)	0.47
PRL (ng/dl)	15.83 (10.75)	15.16 (12.18)	0.81
T (ng/dl)	58.74 (32.02)	58.60 (32.45)	0.61
LH/FSH	1.47 (1.25)	1.43 (1.23)	0.19
Fasting glucose (mmol/l)	5.14 (0.88)	5.11 (0.85)	0.48
2 h-Glucose (mmol/l)	6.37 (2.24)	6.22 (2.21)	0.27
Fasting insulin (μIU/ml)	9.80 (8.14)	10.02 (7.85)	0.33
2 h-Insulin (μIU/ml)	54.05 (54.08)	48.37 (57.12)	0.40
HOMA	2.17 (1.99)	2.28 (2.03)	0.30
CHOL (mmol/l)	4.45 (1.32)	4.40 (1.32)	0.41
TG (mmol/l)	1.03 (0.89)	1.13 (0.87)	0.16
LDL (mmol/l)[b]	2.33 (0.99)	2.23 (1.04)	0.02
HDL (mmol/l)	1.20 (0.48)	1.19 (0.57)	0.77

Data are presented as the median (interquartile range).
BMI (kg/m^2), body mass index; WHR, waist to hip ratio; mFG, modified Ferriman–Gallwey hirsutism score; FSH (IU/l), follicle-stimulating hormone; LH (IU/l), luteinizing hormone; PRL (ng/dl), prolactin; T (ng/dl), testosterone; Glucose (mmol/l); Insulin(μIU/ml); HOMA, homeostasis model assessment; CHOL (mmol/l), cholesterol; TG(mmol/l), triglyceride; LDL (mmol/l), low-density lipoprotein; HDL (mmol/l), high-density lipoprotein.
[a]The Mann–Whitney test was selected.
[b]Statistical significance was found when compared with the other group ($P < 0.05$).

Table V The association of the phenotype and genotype of rs2479106.[a]

	GG + AG ($n = 744$)	AA ($n = 870$)	P
Age (year)	28 (5)	28 (4)	0.19
BMI (kg/m^2)	24.76 (5.47)	24.44 (6.2)	0.66
mFG	0 (3)	1 (2)	0.13
FSH (IU/l)	6.21 (2.05)	6.28 (2.06)	0.42
LH (IU/l)	9.10 (7.43)	15.80 (6.92)	1.00
PRL (ng/dl)	15.46 (11.72)	15.80 (10.86)	0.78
T (ng/dl)	58.46 (31.90)	58.23 (31.67)	0.47
LH/FSH	1.45 (1.28)	1.43 (1.23)	0.87
Fasting glucose (mmol/l)	5.10 (0.90)	5.12 (0.88)	0.97
2 h-Glucose (mmol/l)	6.28 (2.27)	6.34 (2.19)	0.84
Fasting insulin (μIU/ml)	10.01 (7.87)	9.81 (0.88)	0.30
2 h-Insulin (μIU/ml)[b]	55.67 (57.16)	51.15 (54.55)	0.02
HOMA	2.22 (2.03)	2.19 (1.95)	0.53
CHOL (mmol/l)	4.38 (1.29)	4.49 (1.35)	0.54
TG (mmol/l)	1.04 (0.88)	1.08 (0.91)	0.35
LDL (mmol/l)	2.31 (0.98)	2.34 (1.00)	0.75
HDL (mmol/l)	1.19 (0.55)	1.19 (0.50)	0.78

Data are presented as the median (interquartile range).
BMI(kg/m^2), body mass index; mFG, modified Ferriman–Gallwey hirsutism score; FSH (IU/l), follicle-stimulating hormone; LH (IU/l), luteinizing hormone; PRL (ng/dl), prolactin; T (ng/dl), testosterone; Glucose(mmol/l); Insulin(μIU/ml); HOMA, homeostasis model assessment; CHOL (mmol/l), cholesterol; TG (mmol/l), triglyceride; LDL (mmol/l), low-density lipoprotein; HDL (mmol/l), high-density lipoprotein.
[a]The Mann–Whitney test was selected.
[b]Statistical significance was found when compared with the other group ($P < 0.05$).

the subjects with the CC genotype after adjusting for age and BMI ($P = 0.02$).

rs2479106 (*DENND1A*)

Compared with the non-risk group, the GG + AG group showed increased levels of insulin 2 h after a 75-g glucose load, when using the dominant model for rs2479106 ($P = 0.02$; Table V). The difference remained statistically significant after adjusting for age and BMI ($P = 0.02$).

rs10818854 (*DENND1A*)

Using a dominant model, the different genotype groups for rs10818854 did not show any significant differences in phenotype (Supplementary data, Table SIV).

rs13405728 (*LHCGR*)

Although a lower fasting insulin level and HOMA were observed in patients with the AA genotype at rs13405728 using a recessive model (Supplementary data, Table SV), no significant differences were found after adjusting for age and BMI ($P = 0.23$ and $P = 0.19$, respectively).

Discussion

Five susceptibility SNPs of PCOS had been identified in the previous GWAS of a Han Chinese cohort, suggesting the clinical relevance of several novel candidate genes, including *THADA*, *DENND1A* and *LHCGR*. In the present study, correlations between the risk genotypes and the clinical characteristics were analyzed in a subgroup of the PCOS patients tested in the GWAS.

Using genotype–phenotype correlation analysis, the five susceptibility SNPs appear to have different contributions to each subtype. The SNPs in *THADA* and *DENND1A* were associated with specific subtypes including PCO, whereas LHCGR (rs13405728) may be more likely to contribute to the subtype in patients with oligo- or anovulation. The correlation analysis implicates different genetic backgrounds that have the potential to influence specific subtypes, and provides genetic support to the diagnosis of PCOS. However, additional studies on a larger scale are needed to confirm this hypothesis because the number of patients evaluated in several of the sub-phenotype groups was low, which may weaken the statistical power.

Subsequent quantitative trait analysis revealed an association between the rs13429458 AA genotype in the *THADA* gene and increased levels of T. As a key pathophysiological feature of PCOS, hyperandrogenemia significantly affects the progression of the

syndrome. It has been implied that excess androgen accelerates early follicle development and reduces the atresia rate of early antral follicles, which could induce polycystic changes in the ovary (Gleicher *et al*., 2011) Moreover, the polymorphism at rs13429458 was also identified to be associated with subtypes involving HA, which further indicated that the *THADA* gene possibly supports androgen excess in PCOS. Considering the low prevalence of HA in East Asian women, it is preferable to replicate these findings in a Caucasian population, which is more hirsute in general.

Additionally, individuals with the rs13429458 AA genotype had higher LH concentrations, which could persistently stimulate the ovarian theca cell and result in an overproduction of androgen in the gonadotrophin-dependent stages of folliculogenesis (Dumesic and Abbott, 2008). Hypersecretion of LH in patients with PCOS has been shown to be related to anovulation due to the aberrant expression of the receptor, increased LH/FSH values and the absence of the LH peak (Zhu *et al*., 2010). Furthermore, the significant association between the SNP at rs13429458 and PCOS subtypes with OA also suggested that there might be a contribution of *THADA* to anovulation in PCOS.

The rs12478601 SNP is independently located in the *THADA* gene and was identified in the previous GWAS. In the current study, elevated LDL concentrations in the subgroup with the CC genotype demonstrate the association between the SNP in *THADA* and dyslipidemia for the first time. Because increased LDL levels are an important component of metabolic syndrome and an independent risk factor of CVD, these results suggest that patients with the CC genotype have the potential for developing hyperlipemia and additional severe complications.

Using a dominant model, in the GG + GA group for rs2479106 in the *DENNDIA* gene, the 2 h insulin levels after OGTT were significantly higher, suggesting that patients with these genotypes have an increased risk of insulin resistance. Insulin resistance and compensatory hyperinsulinemia are recognized as causes of hyperandrogenemia and, more importantly, can lead to T2D, hypertension and CVD, which are long-term complications of PCOS (Saad *et al*., 1991; Mather *et al*., 2000).

No associations were observed between the rs10818854 genotype in *DENND1A* and the key features of PCOS. These results may be due to the younger age of the patients in the study, as potential associations may become more significant over time. Additional research is necessary to test this hypothesis.

A previous study of 545 Caucasian women with PCOS reported that the rs13405728 genotype was associated with glucose and insulin metabolism, which was not observed in our study (Lerchbaum *et al*., 2011). The conflicting results may be due to differences in genetic backgrounds and environmental factors in the Chinese and Caucasian populations. The relatively young age of the subjects in this study may be an additional influential factor.

Notably, the increased type I error caused by multiple tests is often rectified by the Bonferroni correction. In the present study, the corrected P value should be 0.003(0.05/17), considering the 17 parameters involved. Compared with the corrected *P*-value, the significance that emerged cannot be regarded as statistically relevant. However, since the Bonferroni correction is considered too conservative when the tests are done more than 10 times (Moran, 2003), it is not reasonable to deny the associations of the three SNPs in the *THADA* and *DENND1A* genes and the pathophysiological characteristics such as increased levels of LH, T, LDL or insulin. The results, to some extent, indicate the increased risk in patients with a risk allele, and should be replicated in a large population from different ethnic groups.

In addition, there are several limitations to the present study. First, the low average age of the participants may represent a considerable bias when evaluating endocrine alterations as a function of different genotypes, and significant endocrine and metabolic disturbances may emerge later in life. Secondly, the sample size in several sub-genotype groups was low (e.g. rs13405728 GG genotype, rs13429458 CC genotype and rs10818854 AA genotype), which limited the statistical power of the analysis. Thus, as a compromised method, the genetic models were adopted under the current circumstance and provided a preliminary indication for genotype–phenotype associations. Validation on a larger scale is expected in the future. Thirdly, although the phenotypic variations identified between the sub-genotype groups were moderate, they are consistent with the probable small contribution of individual risk alleles in a multigenic complex disorder.

The associations observed in the present study suggest a higher risk of exhibiting several of the key pathophysiological characteristics of PCOS, including hypersecretion of LH and T, which is crucial for anovulation, and dyslipdemia and hyperinsulinemia, which can be a sign of metabolic disturbances and an underlying increased risk of long-term complications. The results provide inspiration for further functional studies and may help identify individuals who are more likely to respond to targeted therapies. The differences between the sub-genotype groups indicate an increased risk in the patients with specific risk alleles. To our knowledge, this is the first study in a Han Chinese population with PCOS to report genotype–phenotype correlations in the *THADA* and *DENND1A* genes. Replication of the study in larger cohorts of patients with various genetic backgrounds should be pursued. Additional studies regarding gene functions will be required to gain substantial insights into the etiological mechanisms of the effect of these genes on PCOS phenotypes.

In conclusion, the various phenotypes of PCOS may be due to differences in genetic backgrounds. Based on the data collected from a large cohort of PCOS patients, the *THADA* gene contributes to lipid metabolic disorders and hypersecretion of T and LH, while the *DENND1A* gene is partially responsible for insulin resistance in PCOS patients.

Supplementary data

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

Acknowledgements

We are grateful to Yuehong Bian, Di Wu and Changming Zhang of the Provincial Hospital Affiliated to Shandong University for sample collecting and technical support. We especially thank all the subjects for participating in this study.

Authors' roles

L.C. and H.Z. contributed to the conception and design of the study, to the analysis of data and to drafting the article. B.Z., Z.Q., J.L., X.L., X.Z., J.Z. and Y.S. contributed to the acquisition of data and revised

the article. T.L. contributed to the analysis of data. P.W. and Y.S. revised the article. Z.-J.C. oversaw the entire project and revised the manuscript. All authors gave their final approval of the version to be published.

Funding

This research was supported by the National Basic Research Program of China (973 program) (2010CB945002, 2012CB944700), the National Natural Science Foundation of China (81000238, 81070461, 81000236, 30973170), the Graduate Independent Innovation Foundation of Shandong University (GIIFSDU) (21300070613242, 21300070613246), the Science Research Foundation item of no-earnings health vocation (201002013) and the National Key Technology Research and Development Program (2011BAI17B00).

Conflict of interest

There are no conflicts of interests to declare.

References

Atger M, Misrahi M, Sar S, Le Flem L, Dessen P, Milgrom E. Structure of the human luteinizing hormone-choriogonadotropin receptor gene: unusual promoter and 5′ non-coding regions. *Mol Cell Endocrinol* 1995; **111**:113–123.

Azziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, Yildiz BO. The prevalence and features of the polycystic ovary syndrome in an unselected population. *J Clin Endocrinol Metab* 2004;**89**:2745–2749.

Chen ZJ, Zhao H, He L, Shi Y, Qin Y, Li Z, You L, Zhao J, Liu J, Liang X *et al.* Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3. *Nat Genet* 2011;**43**:55–59.

de Groot PC, Dekkers OM, Romijn JA, Dieben SW, Helmerhorst FM. PCOS, coronary heart disease, stroke and the influence of obesity: a systematic review and meta-analysis. *Hum Reprod Update* 2011; **17**:495–500.

Dumesic DA, Abbott DH. Implications of polycystic ovary syndrome on oocyte development. *Semin Reprod Med* 2008;**26**:53–61.

Eriksen MB, Brusgaard K, Andersen M, Tan Q, Altinok ML, Gaster M, Glintborg D. Association of polycystic ovary syndrome susceptibility single nucleotide polymorphism rs2479106 and PCOS in Caucasian patients with PCOS or hirsutism as referral diagnosis. *Eur J Obstet Gynecol Reprod Biol* 2012;**163**:39–42.

Ferriman D, Gallwey JD. Clinical assessment of body hair growth in women. *J Clin Endocrinol Metab* 1961;**21**:1440–1447.

Gleicher N, Weghofer A, Barad DH. The role of androgens in follicle maturation and ovulation induction: friend or foe of infertility treatment? *Reprod Biol Endocrinol* 2011;**9**:116.

Goodarzi MO, Jones MR, Li X, Chua AK, Garcia OA, YD IC, Krauss RM, Rotter JI, Ankener W, Legro RS *et al.* Replication of association of DENND1A and THADA variants with polycystic ovary syndrome in European cohorts. *J Med Genet* 2012;**49**:90–95.

Hatch R, Rosenfield RL, Kim MH, Tredway D. Hirsutism: implications, etiology, and management. *Am J Obstet Gynecol* 1981;**140**:815–830.

Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;**6**:95–108.

Legro RS, Gnatuk CL, Kunselman AR, Dunaif A. Changes in glucose tolerance over time in women with polycystic ovary syndrome: a controlled study. *J Clin Endocrinol Metab* 2005;**90**:3236–3242.

Lerchbaum E, Trummer O, Giuliani A, Gruber HJ, Pieber TR, Obermayer-Pietsch B. Susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21, and 9q33.3 in a cohort of Caucasian women. *Horm Metab Res* 2011;**43**:743–747.

Marat AL, McPherson PS. The connecdenn family, Rab35 guanine nucleotide exchange factors interfacing with the clathrin machinery. *J Biol Chem* 2010;**285**:10627–10637.

Mather KJ, Kwan F, Corenblum B. Hyperinsulinemia in polycystic ovary syndrome correlates with increased cardiovascular risk independent of obesity. *Fertil Steril* 2000;**73**:150–156.

Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 1985;**28**:412–419.

Moran MD. Arguments for rejecting the sequential Bonferroni in ecological studies. *Oikos* 2003;**100**:403–405.

Panidis D, Tziomalos K, Misichronis G, Papadakis E, Betsas G, Katsikis I, Macut D. Insulin resistance and endocrine characteristics of the different phenotypes of polycystic ovary syndrome: a prospective study. *Hum Reprod* 2012;**27**:541–549.

Rahman NA, Rao CV. Recent progress in luteinizing hormone/human chorionic gonadotrophin hormone research. *Mol Hum Reprod* 2009; **15**:703–711.

Rippe V, Drieschner N, Meiboom M, Murua Escobar H, Bonk U, Belge G, Bullerdiek J. Identification of a gene rearranged by 2p21 aberrations in thyroid adenomas. *Oncogene* 2003;**22**:6111–6114.

Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod* 2004; **19**:41–47.

Saad MF, Knowler WC, Pettitt DJ, Nelson RG, Charles MA, Bennett PH. A two-step model for development of non-insulin-dependent diabetes. *Am J Med* 1991;**90**:229–235.

Shi Y, Guo M, Yan J, Sun W, Zhang X, Geng L, Xu L, Chen Z. Analysis of clinical characteristics in large-scale Chinese women with polycystic ovary syndrome. *Neuro Endocrinol Lett* 2007;**28**:807–810.

Simonis-Bik AM, Nijpels G, van Haeften TW, Houwing-Duistermaat JJ, Boomsma DI, Reiling E, van Hove EC, Diamant M, Kramer MH, Heine RJ *et al.* Gene variants in the novel type 2 diabetes loci CDC123/CAMK1D, THADA, ADAMTS9, BCL11A, and MTNR1B affect different aspects of pancreatic beta-cell function. *Diabetes* 2010; **59**:293–301.

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

Welt CK, Styrkarsdottir U, Ehrmann DA, Thorleifsson G, Arason G, Gudmundsson JA, Ober C, Rosenfield RL, Saxena R, Thorsteinsdottir U *et al.* Variants in DENND1A are associated with polycystic ovary syndrome in women of European ancestry. *J Clin Endocrinol Metab* 2012;**97**:E1342–1347.

Yilmaz M, Isaoglu U, Delibas IB, Kadanali S. Anthropometric, clinical and laboratory comparison of four phenotypes of polycystic ovary syndrome based on Rotterdam criteria. *J Obstet Gynaecol Res* 2011; **37**:1020–1026.

Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 2008; **40**:638–645.

Zhu JQ, Zhu L, Liang XW, Xing FQ, Schatten H, Sun QY. Demethylation of LHR in dehydroepiandrosterone-induced mouse model of polycystic ovary syndrome. *Mol Hum Reprod* 2010;**16**:260–266.