International Journal of Diabetes in Developing Countries (April–June 2025) 45(2):495–506
https://doi.org/10.1007/s13410-024-01360-2

ORIGINAL ARTICLE

Novel pleiotropic variants associated with type 2 diabetes and polycystic ovary syndrome detected using a pleiotropic *cFDR* method

Yuan-Yuan Hu[1] · Xiao Wang[2] · Rui-Ke Liu[3] · Zhi-Mei Feng[4] · Zhi Chen[4] · Li-Bo Chen[1] · Jie Shen[2]

Received: 4 May 2023 / Accepted: 29 May 2024 / Published online: 13 June 2024

Abstract

Background Genome-wide association studies (GWASs) on polycystic ovary syndrome (PCOS) and type 2 diabetes (T2D) are usually conducted as single trait, rather than a simultaneous analysis of the related traits. Therefore, the overlapping genetic mechanisms underlying those traits were largely unknown.

Objective This study aims to investigate the overlapping genetic mechanisms between type 2 diabetes and polycystic ovary syndrome and discover the novel pleiotropic variants between those two traits.

Methods We used an established genetic pleiotropic conditional false discovery rate (*cFDR*) approach to discover novel loci associated with PCOS and T2D by incorporating the summary statistics from existing GWASs of these two traits. Lab experiment was also conducted to verify the identified pleiotropic loci. Mendelian randomization approach was also performed to clarify the causal relationship between those two.

Results Both conditional Q-Q and fold enrichment plots were present to demonstrate the pleiotropic enrichment between PCOS and T2D. Using the *cFDR* level of 0.05, we identified 5 loci for PCOS, 1441 loci for T2D, and five of them were associated with both and were validated in the lab experiment. Significant pleiotropic enrichment was observed between PCOS and T2D. However, we did not observe any causal association between PCOS and T2D.

Conclusions These findings may provide novel insights into the etiology of PCOS and T2D and may further influence the disease development both individually and jointly.

Keywords Polycystic ovary syndrome · Type 2 diabetes · Pleiotropic · Conditional FDR

Introduction

Polycystic ovary syndrome (PCOS) is a common endocrine and metabolic disease in women of reproductive age, and it is characterized by chronic anovulation and hyperandrogenemia [1, 2]. As the most commonly urinary disease in female, it was frequently comorbid with complex traits or diseases like insulin resistance, dyslipidemia, and obesity [3–5]. Studies with establishing evidence suggested that [6] women with PCOS are at higher susceptibility of developing glucose intolerance, type 2 diabetes (T2D) and metabolic syndrome [5]. T2D is a long-term chronic metabolic disorder mainly characterized by high blood sugar, insulin resistance, and relative lack of insulin. Epidemiological studies estimate that 7.5% of the PCOS women are T2D compared with healthy women [7, 8], and the prevalence/incidence of T2D in those of PCOS women was about 2.45/3 times higher than that in healthy women [9, 10]. Strong evidence suggested that PCOS and T2D share primary risk factors such as obesity, insulin resistance, dyslipidemia, and metabolic syndrome [4, 6, 10].

Heritability studies have shown that genetic factors have significant contribution to PCOS risk (h^2~38–71%) [11] and T2D risk (h^2~40–70%) [12]. Previous LD score regression analysis

✉ Jie Shen
drshen2020@163.com

1 Department of Endocrinology and Metabolism, Shenzhen Nanshan People's Hospital, Shenzhen 518052, China

2 Department of Endocrinology and Metabolism, Shunde Hospital, Southern Medical University (The First People's Hospital of Shunde Foshan), Foshan 528308, Guangdong, China

3 Department of Endocrinology and Metabolism, SSL Central Hospital of Dongguan City, Dongguan 523326, China

4 Department of Endocrinology and Metabolism, Affiliated Qingyuan Hospital, Guangzhou Medical University, Qingyuan People's Hospital, Guangzhou 510630, China

[13] discovered significant genetic correlations between T2D, fasting insulin, and PCOS, indicating the shared genetic architecture and biological mechanisms between T2D and PCOS. Although genome-wide association studies (GWASs) have identified many genetic loci associated with PCOS or T2D, these loci can only explain 10% of the genetic variance for either PCOS or T2D [14]. Given the high degree of heritability, close relationship, and potential pleiotropy between these two traits, we assume those two are ideal for the further analyses to improve the detection of loci associated with PCOS or T2D or both and explore their common etiology.

GWASs have been extensively applied in diverse ethics to discover SNPs associated with complex diseases or traits. However, the SNPs identified to date only explain limited variance of the disease risk, which is the so-called "missing heritability" [15, 16]. To increase the statistical power, methods of recruiting additional participants have been proposed, which turned out not feasible and too costly. Therefore, several more efficient approaches have been established [17–19] and successfully performed [20, 21] to identify novel loci for various complex diseases.

Pleiotropy refers to the phenomenon that a single gene affects more than two traits [22]. Studies have shown that genetic pleiotropy exists in many correlated diseases and traits, such as bipolar disorder and schizophrenia [23], indicating the shared genetic mechanisms between them. Andreassen et al. [17] established a genetic-pleiotropy-informed conditional false discovery rate (*cFDR*) method by leveraging two independent GWASs from associated traits into a conditional framework. This approach has been widely used in multiple genetic-related diseases and traits by other groups [24–26] and our groups [27–32]. The findings of these studies have clearly demonstrated the utility of this approach in improving statistical power and gene discovery.

In the current study, by using the *cFDR* method [17] on two large and independent GWAS summary statistics of PCOS and T2D [13, 33], we intended to identify additional pleiotropic novel loci between them. The aim of this study is to improve gene detection for PCOS and T2D and explore the potential causal relationship between those two traits with these two current GWASs and to gain some novel insights into shared underlying mechanisms and overlapping genetic variance between them.

Materials and methods

GWAS datasets for PCOS and T2D

The PCOS GWAS summary statistics include data from 10,074 PCOS cases and 103,164 controls of European descent [13]. The dataset was publicly available from https://www.repository.cam.ac.uk/handle/1810/283491. The T2D GWAS summary statistics contain data from 74,124 T2D cases and 824,006 controls of European descent [33]. All the datasets contain the summary statistics for each SNP, providing the *p* values that have undergone genomic control at the individual study level and again after meta-analysis. For the detailed information about the samples and methods for each dataset, please refer to the corresponding consortium papers [13, 33].

Conditional false discovery rate

The *cFDR* approach is very mature and well-established now and has been successfully applied in multiple correlated diseases or traits by many other groups [24–26] and our group [27–32]. Therefore, we briefly outline this approach as follows: First, data was pre-processed according to previous protocols for next *cFDR* analysis. Then conditional QQ plot was present to demonstrate the pleiotropic enrichment between PCOS and T2D, where conditional empirical cumulative distribution functions (cdfs) for PCOS SNP *p* values were conditioned on T2D nominal *p* values, and vice versa. For each nominal *p* value, an estimate of the *cFDR* was obtained from the conditional empirical cdfs. Finally, two *cFDR* tables were computed; *cFDR* result for PCOS was conditioned on T2D and vice versa. Using the *cFDR* level of 0.05, we identified loci associated with PCOS and T2D, respectively. Then, conjunction *FDR* value was defined as the larger one of those two *cFDR* values.

Validation cFDR analysis using F) and FG

To represent as sensitivity and validation analysis, same *cFDR* analysis were performed for female-specific fasting insulin (FI) and PCOS and female-specific fasting plasma glucose (FG) and PCOS. The dataset was publicly available from http://www.diagram-consortium.org/downloads.html. The female-specific FG and FI GWAS summary statistics include data from 73,089 women (FG meta-analysis) and 50,404 women (FI meta-analysis) of European descent [34]. The data was downloaded from https://magicinvestigators.org/downloads/.

Conditional QQ and true discovery rate plots for assessing pleiotropic enrichment

To evaluate the pleiotropic enrichment between two correlated traits, we proposed conditional QQ plots based on different significance levels of conditional phenotypes. The QQ plots were presented as the observed distribution of *p* values against the expected distribution of *p* values under the null hypothesis. We generated the QQ curve for the nominal $-\log_{10}(p)$-value quantiles associated with SNPs that are below each significance level of the conditional trait. The empirical cdfs of the nominal *p* values on the *x*-axis are

plotted against the nominal $-\log_{10}(p)$-values on the *y*-axis. The degree of pleiotropic enrichment is determined by the degree of separations of the lines from the expected null line, and with the decreased *p* values, the separation from the null line will continue. Furthermore, conditional true discovery rate (TDR) plots were also present to demonstrate the pleiotropic enrichment between PCOS and T2D, where TDR equals 1-FDR.

Conditional Manhattan plots for localizing genetic variants

To provide a better visualization of the localization of the SNPs associated with PCOS or/and T2D, we present conditional Manhattan plots. Any SNP with a $-\log_{10}(\text{FDR})$ value greater than 1.3 ($cFDR < 0.05$ and $ccFDR < 0.05$) was considered as significantly associated with the principal phenotype or both.

Functional annotation and gene enrichment analysis

To assess the functional classifications of the identified loci, we performed functional annotation and gene enrichment analysis using the PANTHER classification system (http://www.pantherdb.org/) [35]. First, enter the name of the genes to be analyzed, one on each line or separated by commas. Then, select the species and GO terms (molecular function, biological process, and cell composition). Your results will then be redirected to PANTHER website. These results are based on the relative enrichment of all protein coding genes in the genome you selected. All significant genes identified in this study were classified according to the following categories: family and subfamily, molecular function, biological processes, and pathway. This analysis provided comprehensive biological information, which allows us to partially validate our findings by determining specific genes that are enriched in PCOS- and T2D-related GO terms.

Protein-protein interaction network

To reveal the physical and functional associations of the identified genes, protein-protein interaction (PPI) analyses were conducted by using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (http://string-db.org/).

Laboratory experiments to valid the pleiotropic gene loci

To valid the molecular/biological mechanisms of the identified pleiotropic loci for PCOS (PROX1), we constructed PCOS mice models, performed hematoxylin and eosin (H&E) staining, and western blot analysis.

Animals and PCOS model treatment

A total of 20 3-week-old female C57/BL6 mice were obtained from the animal laboratory of the Southern Medical University (Guangzhou, China). All animals were bred well, housed in a clean laminar room at a controlled temperature of 20–25 °C with a 12-h light/dark cycle. Mice were randomly divided into two groups (PCOS group ($n = 10$) and vehicle group ($n = 10$)). PCOS group mice were injected with dehydroepiandrosterone (DHEA, Sigma, 6 mg/100 g·d, dissolved in 0.1 ml sesame oil [sigma, USA])) as previously described [36] for 21 consecutive days, while the mice of the control group received sesame oil for an equivalent length of time. All mice were maintained under specific relative-comfortable conditions and were approved by the Animal Ethics Committee of Southern Medical University (No. L2015069).

H&E staining

Ovarian tissues were dissected from mice, fixed for 24 h in 4% paraformaldehyde at 4 °C and embedded in paraffin. They were sliced into 4-μm-thick sections and stained with hematoxylin for 5 min at room temperature, alcohol hydrochloric acid for 10 s, and eosin for 1 s. The sections were then examined by light microscopy.

Western blot analysis

Liver tissues were collected, weighed, and ground into powder under liquid nitrogen. The samples were mixed with RIPA lysis buffer (50 mM Tris (pH 7.4), 150 mM nacl, 1% NP-40, 0.5% sodium deoxycholate). The protein concentration was determined using the bicinchoninic acid method. Protein samples were electrophoresed on 10% SDS-PAGE gels and transferred into nitrocellulose membranes. The membranes were then blocked in 5% skimmed milk at room temperature, incubated overnight at 4 °C with the primary antibodies as follows: Rabbit anti-PROX1 antibody (1:1000; cat. no. ab199359; Abcam) and Rabbit anti-beta-actin antibody (1:8,000; #AB0035; Abways). After washing primary antibodies by TBS with Tween for four times, the membranes were then incubated with a horseradish peroxidase (HrP)-conjugated secondary antibody for 1 h at room temperature. The specific reaction band was visualized using the enhanced chemiluminescence system and exposed. Finally, the ImageJ (MD, USA) image analysis software was used to measure the electrophoretic band gray value. The experiments were repeated three times for each sample.

MR analysis

To explore whether there is causal association between PCOS and T2D, we performed bi-directional Mendelian randomization (MR) analysis using the GWASs datasets mentioned above. First, we selected LD independent SNPs that reached $p < 1E{-}5$ for PCOS as our instrumental variables (IVs) (there was only one SNP reached genome-wide significance level). Those SNPs were then extracted from T2D dataset, and proxy SNPs that were in high LD ($r^2 > 0.8$) with the interest SNPs were used when target SNPs were not available. Data harmonization was next performed to ensure the effect SNP on exposure and outcome is the same allele. Finally, inverse variance weighted (IVW)-fixed effects regression method was used to assess the causal effect of PCOS on T2D. Maximum likelihood method (MLM) and weighted median approach were also applied to complement the IVW. Furthermore, MR-Egger regression [37] was performed to estimate the pleiotropy effect among selected IVs. Finally, MR Steiger directionality test [38] was performed to orient this causal relationship. Then, to exclude the reversal causal relationship, bi-directional MR analysis was repeated using T2D as exposure, PCOS as outcome. All the analyses were performed using "TwoSampleMR" package in R program.

Results

Assessment of pleiotropic enrichment

The conditional Q-Q plot for PCOS (A in Figure 1) shows some enrichment conditioning on various strengths of associations with the T2D. The degree of deflection between

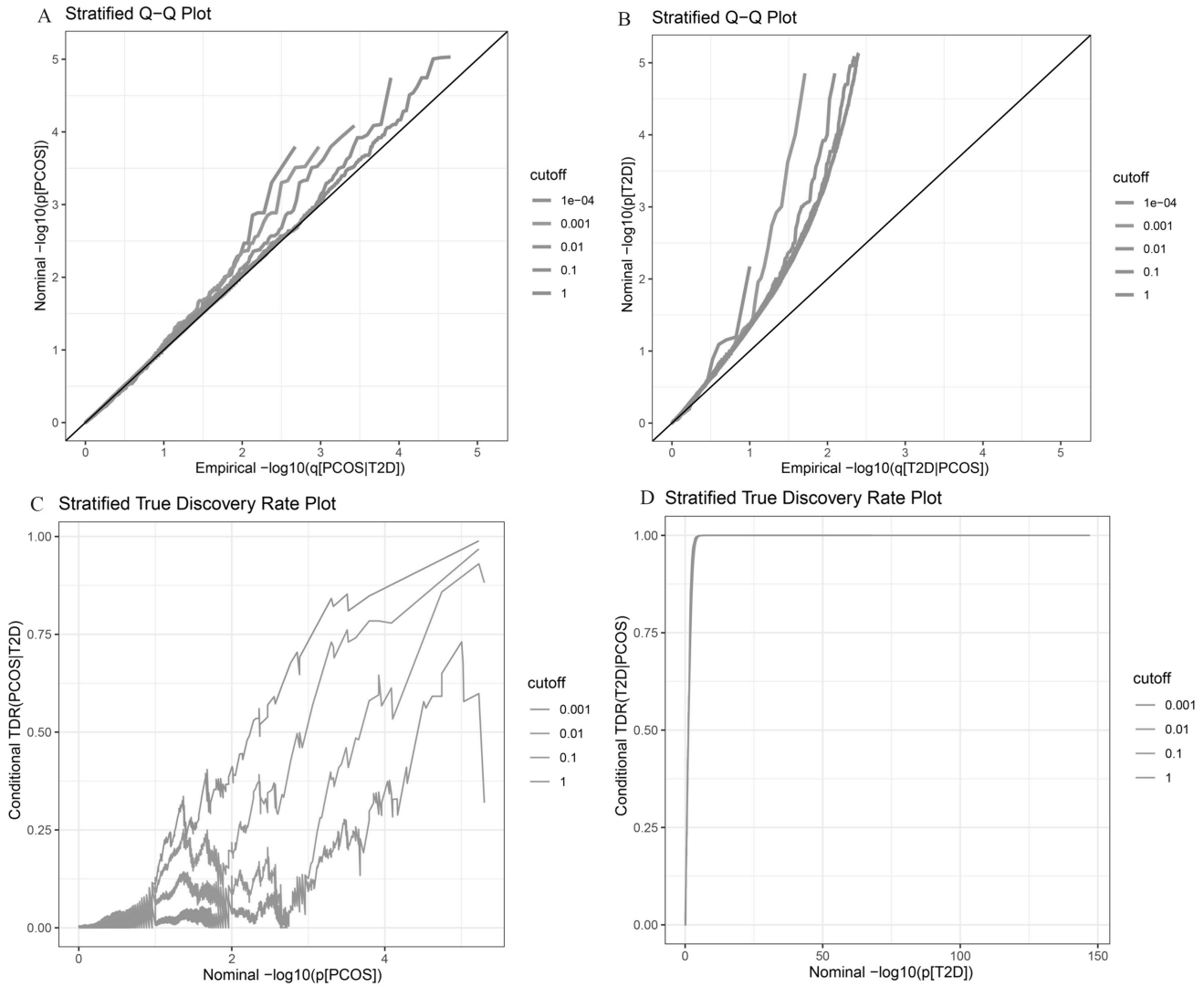

Fig. 1 Assessment of pleiotropic enrichment

curves when including the SNPs that have more significant associations with PCOS indicates an increase in the number of true associations for a given T2D *p* value. Similar enrichment is observed for T2D given PCOS (B in Figure 1), as there appears to be a similar left shift pattern between the different curves. These earlier deflections from the null line indicate a great proportion of true associations for any given T2D nominal *p* value. Furthermore, the upward shift of the curves in the TDR plots also demonstrates significant pleiotropic enrichment between PCOS and T2D (Figure 1C and D).

PCOS loci identified with cFDR

A total of 5 significant SNPs ($cFDR < 0.05$) were identified for PCOS variation (A in Figure 2 and Table S1), and they were mapped to 4 different chromosomes (1, 3, 11, and 16) and annotated to 4 genes. Original GWAS for PCOS [13] and other previous GWASs have identified 19 genetic susceptibility loci associated with PCOS [13, 39–42]. Those 5 SNPs we identified were not previously reported, and only one SNP had *p* values smaller than 1×10^{-5}, while four of them had *p* values larger than 1×10^{-5}. Although those SNPs were not previously reported to be associated with PCOS or PCOS-related traits, four of them were in high LD ($r^2 = 0.8$) with the other traits associated loci like body mass index (BMI), T2D, body fat percentage, waist-hip ratio, and wait circumference (Table S2). SNP rs5030244 was not previously reported in the original PCOS GWAS, and the previous studies did not show its association with PCOS and related traits. For the 4 genes, these 5 SNPs annotated to all of them were newly detected compared to the original PCOS and previous PCOS-related studies.

T2D gene loci identified with cFDR

A total of 1441 significant SNPs ($cFDR < 0.05$) were identified for T2D variation (B in Figure 2 and Table S3).

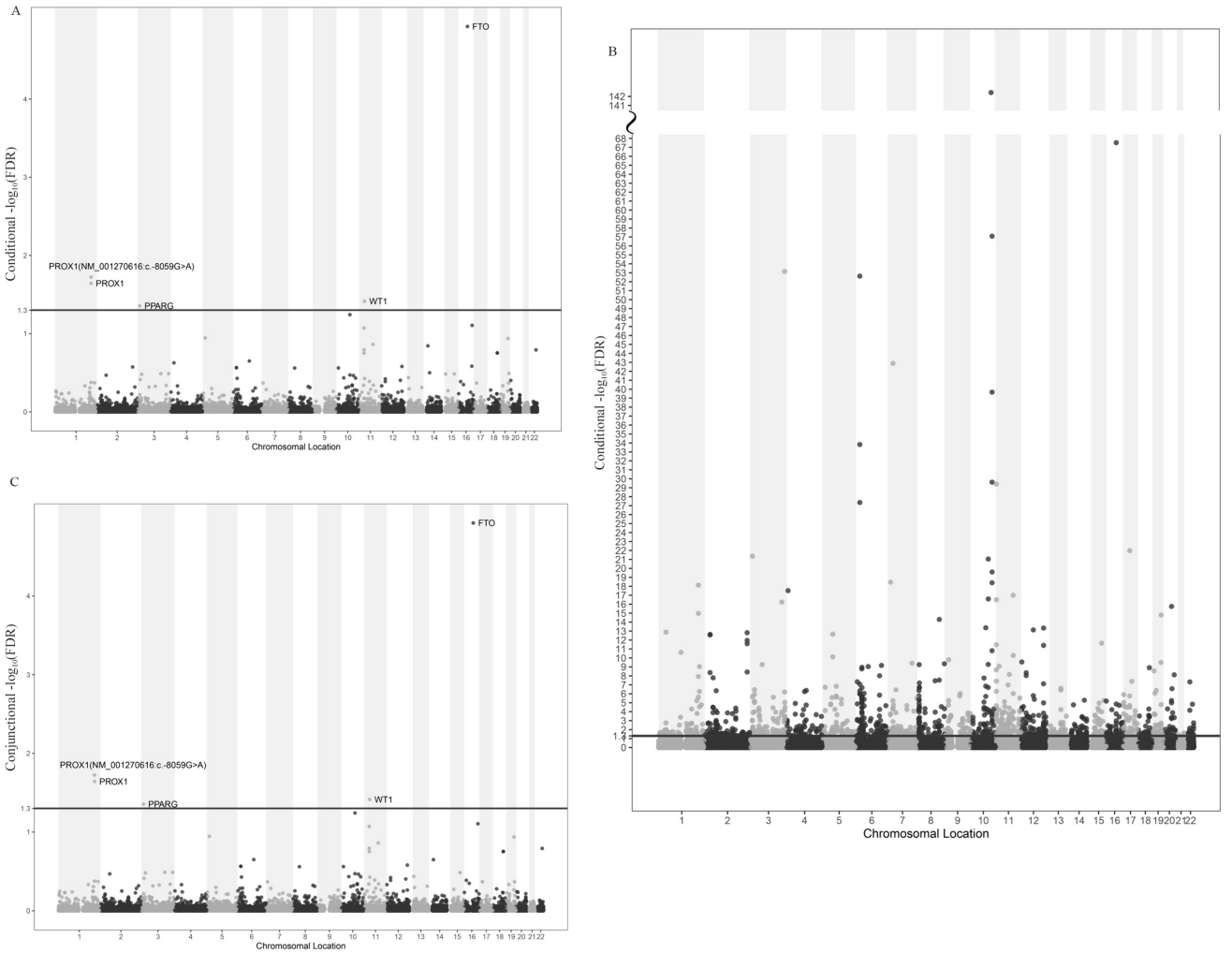

Fig. 2 T2D gene loci identified with *cFDR*; pleiotropic gene loci for both PCOS and T2D

Compared with the original T2D GWAS [33], 580 SNPs had *p* values smaller than 1E−5, while 269 of them reached the standard genome-wide significance of 5E−8. When we restricted the *cFDR* to 0.01, 770 SNPs were left. Of the detected loci that were not reported in the original T2D GWAS, we separated those loci into two categories: genes with *p* values between genome-wide significance and 1E−5, and genes with *p* values between 1E−5 and 0.05 were separately subjected to the PPI analysis from the STRING database; multiple of the genes were previously reported to be associated with T2D or T2D-related studies (Table S4, Figure A and B).

Pleiotropic gene loci for both PCOS and T2D

In this study, by leveraging two GWASs, we identified 5 independent pleiotropic loci (conjunction FDR < 0.05) for both traits (C in Figure 2 and Table 1). Of the 5 identified pleiotropic variants, four SNPs rs9930506 (FTO), rs340839 (PROX1) and rs3767844 (PROX1), and rs2881654 (PPARG) were reported to be significant for T2D in the original T2D GWAS. However, those four SNPs were not previously reported in original GWAS of PCOS or related studies. Furthermore. the remaining SNP rs5030244 was not previously reported in the original PCOS/T2D GWASs and PCOS/T2D-related studies. For the 4 genes, these 5 SNPs annotated to genes FTO, PPARG, and PROX1 were previously reported to be associated with T2D (Table S4), but WT1 was newly detected for T2D. However, for their association with PCOS, the study reported that genes FTO and WT1 were previously reported to be associated with PCOS [43, 44]; genes PROX1 and PPARG were newly detected compared to the original PCOS and previous PCOS-related studies. Detailed information was shown in Table S5. Of the detected 5 pleiotropic loci, four genes were enriched in terms "positive regulation of heart growth," "regulation of lipid storage tissue development," and "regulation of fat cell differentiation." Detailed information of GO term analysis is shown in Tables 2 and 3.

Validation cFDR results

For *cFDR* analysis between PCOS and FI, we observed similar significant separation between the different curves, which indicates a strong pleiotropy between those two traits (A and B in Figure S1). Conditional on association with FI, we identified 3 significant SNPs (3 genes, $cFDR \leq 0.05$) for PCOS variation, and one gene was replicated compared with the main *cFDR* analysis. We identified a total of 4 significant SNPs (5 genes) for FI variation on their association with PCOS, and one SNP and three genes of them were replicated, and we replicated one pleiotropic gene for PCOS and FI (Table S6).

For *cFDR* analysis between PCOS and FG, we identified significant separation between the different curves, which indicates a strong pleiotropy between those two traits (C and D in Figure S1). Conditional on association with FG, we identified 3 significant SNPs (four genes) for PCOS variation, and one gene of them was replicated compared with the main *cFDR* analysis. We identified a total of 37 significant SNPs (38 genes) for FG variation on their association with PCOS, and 13 SNPs, and 23 genes of them were replicated, and we replicated one pleiotropic gene for PCOS and FG (Table S7).

Table 1 Conjunction FDR: pleiotropic loci in PCOS and T2D (cFDR < 0.05)

RSID	Role	Gene	CHR	*p* value A	*p* value B	*cFDR.AcB*	*cFDR.BcA*	Conjunction *FDR*
rs9930506	Intronic	FTO	chr16	5.90E−06	1.50E−68	1.18E−05	3.00E−68	1.18E−05
rs340839	UTR5	PROX1	chr1	0.0014	1.00E−17	0.0189	1.06E−15	0.0189
rs3767844	Intronic	PROX1	chr1	5.00E−04	2.80E−10	0.02275	1.18E−08	0.02275
rs5030244	Intronic	WT1	chr11	0.00016	1.40E−05	0.03848	0.00021	0.03848
rs2881654	Intronic	PPARG	chr3	0.0063	1.00E−24	0.0441	4.39E−22	0.0441

p value A is the *p* value of PCOS; *p* value B is the *p* value of T2D

Table 2 Functional term enrichment analysis

#Term ID	Term description	Gene count	Fold enrichment	Raw *p* value	FDR	Matching proteins in your network (labels)
GO:0040008	Regulation of growth	4	31.11	1.08E−06	1.70E−02	FTO, PPARG, PROX1, and WT1
GO:0050692	DNA binding domain binding	2	> 100	4.24E−07	2.08E−03	PPARG and PROX1
GO:0050693	LBD domain binding	2	> 100	1.02E−06	2.49E−03	PPARG and PROX1

Table 3 Functional ontology terms list of the novel genes

#Term ID	Term description	Parent	Child
GO:0032879	Regulation of localization	Regulation of biological process	Regulation of sequestering of calcium ion Regulation of transport Regulation of cellular localization
GO:0000981	DNA-binding transcription factor activity RNA polymerase II-specific	DNA-binding transcription factor activity	DNA-binding transcription activator activity, RNA polymerase II-specific DNA-binding transcription repressor activity, RNA polymerase II-specific
GO:0006631	Fatty acid metabolic process	Cellular lipid metabolic process	Unsaturated fatty acid metabolic process Very long-chain fatty acid metabolic process Fatty acid catabolic process Fatty acid biosynthetic process Long-chain fatty acid metabolic process
GO:0019216	Regulation of lipid metabolic process	Regulation of primary metabolic process	Regulation of lipid catabolic process Regulation of lipid biosynthetic process

Protein-protein interaction network

The identified T2D-associated genes with *p* values between genome-wide significance and 1E−5 and genes with *p* values between 1E−5 and 0.05 were separately retrieved from the STRING database. For 313 genes with *p* values between genome-wide significance and 1E−5, 238 of them were annotated in this database and clearly enriched into seven clusters (Figure S2 and Table S4), and for the 949 genes with *p* values between 1E−5 and 0.05, 362 of them were annotated in this database and clearly enriched into six clustered, as shown in Figure S3 and Table S4.

Laboratory validation results

Initially, we did not observe any obvious weight difference between vehicle and PCOS group. However, after 21 consecutive days of DHEA injections, the body weight of mice in the PCOS group were significantly increased compared with the control group ($p < 0.01$, Fig 3A). We further observed the morphological changes of ovarian tissue in the two groups under light microscope. In the vehicle group, the cell membrane of the granular cells was intact and continuous, arranged neatly, and the layers were more compact (Fig 3C), while PCOS group showed polycystic ovary changes, increased follicular volume, and disappearance of some oocytes in follicles. Moreover, the granular cell layer was reduced and arranged sparsely (Fig 3D), which demonstrates the successful establishment of the PCOS mice model. Our western blot showed that compared with the control group, the expression of PROX1 protein was significantly decreased in PCOS group ($p < 0.05$) (Figs. 3B and 4).

MR analysis results

After LD clumping, there were 31 LD independent SNPs selected for PCOS as our IVs, then those SNPs were extracted from T2D dataset. After data harmonization, 28 SNPs were left for MR analysis. Standard IVW found that PCOS was associated with increasing risk of T2D (beta = 0.031, se = 0.008, $p = 5.615E{-}05$, Table S8). Consistent with the IVW results, MLM also identify causal association between PCOS and T2D (Table S8), and our egger regression demonstrates no pleiotropy among the selected IVs (beta = 0.014, se = 0.014, p = 0.349). The Steiger test showed that variance explained in the T2D ($r^2 = 0.002$) was less than that in the PCOS ($r^2 = 0.029$) by the instrumenting SNPs ($p = 7.47\text{-}E{-}79$). However, when we performed a sensitivity analysis which only include SNPs that reached significance of $p < 1E{-}6$, no methods reported any association between PCOS and T2D (beta = −0.024, se = 0.017, $p = 0.154$). Our bi-direction MR did not identify any causal association between T2D and PCOS (beta = 0.261, se = 0.140, $P_{IVW} = 0.062$).

Discussion

In this study, by integrating two independent GWASs of PCOS and T2D, we identified significant pleiotropic enrichment between those two traits. Compared to the traditional single trait GWAS, integration of correlated phenotypes allows for the increased detection of significant loci without the need of larger recruitment of individuals in GWASs. By incorporating two independent GWAS summary statistics of PCOS and T2D, we identified 5 loci for PCOS and 1441

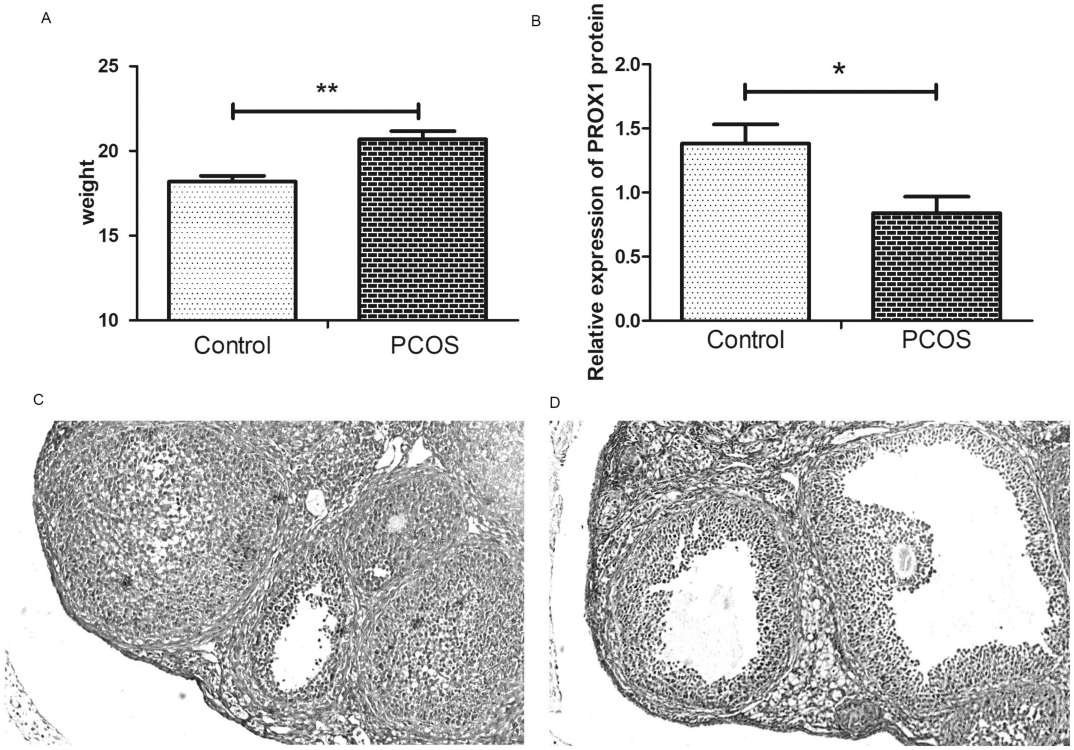

Fig. 3 Laboratory validation results

loci for T2D. Using the genome-wide significance level in the datasets, only 269 SNPs for T2D were significant and no SNPs for PCOS. Previous studies did not report any suggestive association with PCOS and T2D respectively, as shown in Tables S1 and S2. Utilizing the genetic pleiotropic-informed *cFDR* method, we discovered 5 novel genes associated with both PCOS and T2D. The novel findings may help us to further discover the shared genetic mechanisms underlying these two related diseases. The improved identification of novel susceptibility pleiotropic loci may enable us a better understanding of common etiology between diseases, which is of great significance for the clinical prevention and treatment of chronic complex diseases.

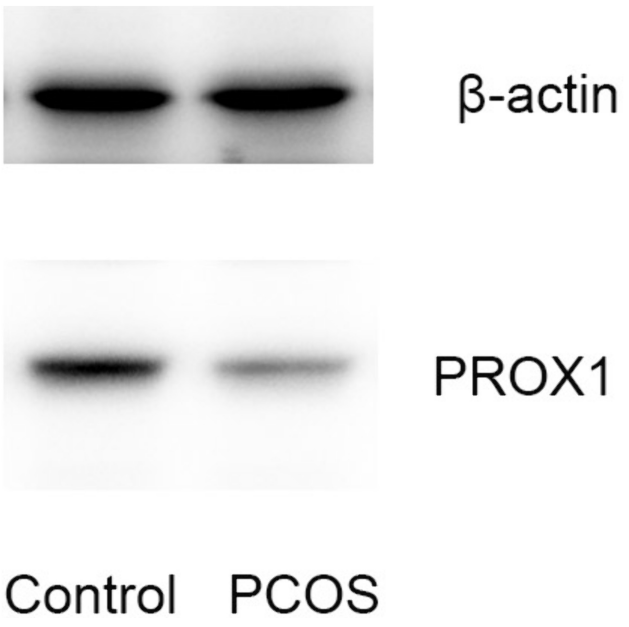

Fig. 4 Western blot showing comparison with the control group; the expression of PROX1 protein was significantly decreased in PCOS group

The 5 pleiotropic signals identified by *ccFDR* approach in this study demonstrated that there was a close relationship and shared genetic mechanisms between these two traits. These 5 pleiotropic SNPs were annotated to 4 genes: FTO, PROX1, WT1, and PPARG. The implementation of *cFDR* method in our study not only enable us successfully to detect known disease-associated genetic variants but also shows the practicability of improved discovery of novel susceptibility loci using existing GWASs summary results. Among those genes, three genes FTO, PROX1, and PPARG were previously reported to be associated with T2D, genes FTO and WT1 were associated with PCOS. Therefore, for T2D, gene WT1 was newly identified, and genes PROX1 and PPARG were novel for PCOS. However, with *ccFDR* framework, those four genes were identified as pleiotropic loci in the current study. Thus, we will discuss genes WT1, PROX1,

and PPARG in the following for their potential functional relevance and significance.

The SNP rs5030244 is located at the intronic region of gene WT1. Gene WT1 was first initiated as a strong candidate predisposition gene for pediatric kidney cancer in 1990s. Moreover, this gene was widely expressed in multiple adult tumor types, including tissues of epithelial, mesenchymal, hematopoietic, and neuronal [45]. Studies reported that WT1 mutation was associated with increasing risk of renal glomerulosclerosis and gonadal dysgenesis; furthermore, germline gene mutation of WT1 may lead to congenital diaphragmatic hernia and heart disease in rare cases. Established evidence support the role of WT1 in kidney development, homeostasis, and disease, which can further explain its function with respect to Wilms' tumor and glomerulosclerosis [45]. The study with RT-PCR quantified the expression of WT1 in granulosa cells of the PCOS patients and healthy controls, and the results showed increased expression of WT1 in PCOS patients; furthermore, the WT1 expression was reported to be moderately correlated with luteinizing hormone levels and the antral follicle counts, which demonstrates the possible mechanisms between hyperandrogenism and polycystic ovaries of PCOS and WT1 [43].

Studies have confirmed that gene PROX1 is related to T2D, but no previous studies reported whether PROX1 is a susceptibility gene of PCOS. In this study, by using *ccFDR* approach, we found that PROX1 is a pleiotropic gene for T2D and PCOS. Previous studies demonstrated the underlying mechanisms of PROX1 in T2D; gene PROX1 can encode a transcription factor that is involved in the proliferation and maturation of postnatal β-cells, thus leading to the reduced secretion of glucose-stimulated insulin and then developed into T2D [46]. Furthermore, PROX1 expression was also reported associated with lipid regulation, which may maintain the circadian rhythm and metabolic homeostasis of cholesterol metabolism by inhibiting the reverse transport of cholesterol in the circulation [47]. Goto et al. [48] reported liver-specific PROX1 knockout mice developed insulin-resistant diabetes. Given this, we assumed that PROX1 may also be involved in the occurrence of insulin resistance in POCS. To partially validate this, we conducted laboratory experiment. As the liver is one of the important organs that regulate glucose metabolism and maintain blood glucose stability, and it is also an important organ involved in insulin resistance [48]; therefore, the liver was selected as the research organ in this study, and we observed the expression of PROX1 protein in the liver of PCOS mice. The decreased expression of PROX1 protein in PCOS mice demonstrates its association with PCOS.

The studies with established evidence have reported that the association between PPRAG and T2D. PPAR-γ has the potential to regulate the adipocyte differentiation, glucose and lipid metabolism, insulin sensitivity, and inflammatory pathways, which demonstrates its involvement in the physiological regulation of systemic metabolism and the etiology of metabolic disorder [49, 50]. Therefore, genetic mutation or variations of PPARG were related to adipose tissue disease, insulin resistance, and T2D. PPAR-γ has many mutations, the most common of which are Pro12A1a (C/G) (rs1801282) in exon 2 and His447His (C/T) (rs3856806) in exon 6. By detecting the polymorphisms of Pro12A1a and His447His, Shaikh et al. [51] found that the mutant genotypes Pro12A1a G and His447His T in PCOS group were significantly lower than those in control group. Dasgupta et al. [52] drew the same conclusion and found that patients with Pro12A1a G and His447His T were less likely to have hyperandrogenemia and hyperinsulinemia. Given the evidence provided by previous research, genetic polymorphisms of PPAR-γ (rs1801282 and rs3856806) may play a role in protecting the host body in the development of PCOS.

We therefore assume that those genes might be involved in certain processes that are significant in the development of PCOS and T2D; however, more future studies are expected to explore the exact mechanisms of the novel gene we identified.

This study has several strengths. First, the statistical power and sample size are increased through the *cFDR* method by combining two large GWAS datasets. Compared to a meta-analysis of the same data [53], this *cFDR* framework allows the detection of disease-related loci regardless of their effect directions. Secondly, MR analysis and sensitivity analysis were applied to assess the potential causal relationship between PCOS and T2D. Compared to the traditional observational studies, MR method could overcome the influence of the confounding factors and reverse causation.

There are also some limitations in this study. First, we could not provide the effect size estimates of pleiotropic loci on the phenotypes due to a lack of individual level data. However, this information can be inferred from the summary beta values in the original GWAS. Second, we could not provide an independent validation of the current results due to a lack of another publicly available PCOS GWAS dataset. However, alternative approaches like GO enrichment and PPI analysis were already applied to check the biological significance of the identified loci. Furthermore, more future experimental studies are needed to further confirm novel findings in our study.

Conclusion

In conclusion, by incorporating two independent GWAS datasets of PCOS and T2D into a *cFDR* framework, we found strong pleiotropic enrichment between those two traits, demonstrating its utility in improving statistical

power. In this study, we discovered 5 novel pleiotropic loci of potential functional significance for PCOS and T2D, and the results may provide us novel insights into the shared genetic mechanisms between these two disorders.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13410-024-01360-2.

Acknowledgments We would like to express our great attitude to DIAGRAM Consortium and the authors of the original PCOS GWAS paper for making the GWAS summary statistics publicly available. We appreciate the support from the Third Affiliated Hospital of Southern Medical University for this project.

Author contribution YYH as the first author performed data analysis and wrote the manuscript. XW performed the lab experiments of the results. RKL and ZMF provided advice and suggestions, while we met some problems during the data analysis process. ZC and LBC gave constructive suggestions during the whole process. JS conceived and initiated this project, provided advice on experimental design, oversaw the implementation of the statistical method, and revised/finalized the manuscript. This study was partially supported by Guangzhou Science and Technology Project (NO: 202102080533), the funder has no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Declarations

Ethics approval The research of human subjects was exempted from ethical approval and consent to participate because the GWAS summary datasets used in this study were publicly available, and the original studies already obtained approval. The research of animal subjects was approved by the Animal Ethics Committee of Southern Medical University.

Competing interests The authors declare no competing interests.

References

1. Polycystic ovary syndrome (PCOS): condition information. January 31, 2017. [cited 2020 10 September]; Available from: https://www.nichd.nih.gov/health/topics/pcos/conditioninfo.
2. Polycystic ovary syndrome (PCOS) fact sheet. December 23, 2014 [cited 2020 10 September]; Available from: https://www.womenshealth.gov/a-z-topics/polycystic-ovary-syndrome.
3. Meier RK. Polycystic ovary syndrome [J]. Nurs Clin North Am. 2018;53(3):407–20.
4. Moghetti P. Insulin resistance and polycystic ovary syndrome [J]. Curr Pharm Des. 2016;22(36):5526–34.
5. Azziz R. Polycystic ovary syndrome [J]. Obstet Gynecol. 2018;132(2):321–36.
6. Ali AT. Polycystic ovary syndrome and metabolic syndrome [J]. Ceska Gynekol. 2015;80(4):279–89.
7. Legro RS, Kunselman AR, Dodson WC, Dunaif A. Prevalence and predictors of risk for type 2 diabetes mellitus and impaired glucose tolerance in polycystic ovary syndrome: a prospective, controlled study in 254 affected women [J]. J Clin Endocrinol Metab. 1999;84(1):165–9.
8. Norman RJ, Masters L, Milner CR, Wang JX, Davies MJ. Relative risk of conversion from normoglycaemia to impaired glucose tolerance or non-insulin dependent diabetes mellitus in polycystic ovarian syndrome [J]. Hum Reprod. 2001;16(9):1995–8.
9. Lo JC, Feigenbaum SL, Yang J, Pressman AR, Selby JV, Go AS. Epidemiology and adverse cardiovascular risk profile of diagnosed polycystic ovary syndrome [J]. J Clin Endocrinol Metab. 2006;91(4):1357–63.
10. Morgan CL, Jenkins-Jones S, Currie CJ, Rees DA. Evaluation of adverse outcome in young women with polycystic ovary syndrome versus matched, reference controls: a retrospective, observational study [J]. J Clin Endocrinol Metab. 2012;97(9):3251–60.
11. Vink JM, Sadrzadeh S, Lambalk CB, Boomsma DI. Heritability of polycystic ovary syndrome in a Dutch twin-family study [J]. The Journal of Clinical Endocrinology & Metabolism. 2006;91(6):2100–4.
12. Almgren P, Lehtovirta M, Isomaa B, Sarelin L, Taskinen MR, Lyssenko V, Tuomi T, Groop L, Botnia G. study, heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia study [J]. Diabetologia. 2011;54(11):2811–9.
13. Day F, Karaderi T, Jones MR, Meun C, He C, Drong A, Kraft P, Lin N, Huang H, Broer L, Magi R, Saxena R, Laisk T, Urbanek M, Hayes MG, Thorleifsson G, Fernandez-Tajes J, Mahajan A, Mullin BH, et al. Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria [J]. PLoS Genet. 2018;14(12):e1007813.
14. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, Strawbridge RJ, Khan H, Grallert H, Mahajan A, Prokopenko I, Kang HM, Dina C, Esko T, Fraser RM, Kanoni S, Kumar A, Lagou V, Langenberg C, et al. Replication, large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes [J]. Nat Genet. 2012;44(9):981.
15. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM. Common SNPs explain a large proportion of the heritability for human height [J]. Nat Genet. 2010;42(7):565–9.
16. Yoo YJ, Pinnaduwage D, Waggott D, Bull SB, Sun L. Genome-wide association analyses of North American Rheumatoid Arthritis Consortium and Framingham Heart Study data utilizing genome-wide linkage results [J]. BMC Proc. 2009;3(Suppl 7):S103.
17. Andreassen OA, Djurovic S, Thompson WK, Schork AJ, Kendler KS, O'Donovan MC, Rujescu D, Werge T, van de Bunt M, Morris AP, McCarthy MI, Roddey JC, McEvoy LK, Desikan RS, Dale AM, G. International Consortium for Blood Pressure, R. Diabetes Genetics, C. Meta-analysis, G. Psychiatric Genomics Consortium Schizophrenia Working. Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors [J]. Am J Hum Genet. 2013;92(2):197–209.
18. Cichonska, A., J. Rousu, P. Marttinen, A. J. Kangas, P. Soininen, T. Lehtimäki, O. T. Raitakari, M.-R. Järvelin, V. Salomaa, M. Ala-Korpela, S. Ripatti, M. Pirinen, metaCCA: summary statistics-based multivariate meta-analysis of genome-wide association studies using canonical correlation analysis [J]. bioRxiv, 2015.
19. Chung D, Yang C, Li C, Gelernter J, Zhao H. GPA: a statistical approach to prioritizing GWAS results by integrating pleiotropy and annotation [J]. PLoS Genet. 2014;10(11):e1004787.
20. Andreassen OA, Thompson WK, Schork AJ, Ripke S, Mattingsdal M, Kelsoe JR, Kendler KS, O'Donovan MC, Rujescu D, Werge T, Sklar P, Psychiatric Genomics C, Bipolar D, Schizophrenia Working G, Roddey JC, Chen CH, McEvoy L, Desikan RS, Djurovic S, Dale AM. Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate [J]. PLoS Genet. 2013;9(4):e1003455.
21. Andreassen OA, McEvoy LK, Thompson WK, Wang Y, Reppe S, Schork AJ, Zuber V, Barrett-Connor E, Gautvik K, Aukrust

P, Karlsen TH, Djurovic S, Desikan RS, Dale AM. G. F. f. O. C. International consortium for blood pressure genome-wide association studies, identifying common genetic variants in blood pressure due to polygenic pleiotropy with associated phenotypes [J]. Hypertension. 2014;63(4):819–26.
22. Stearns FW. One hundred years of pleiotropy: a retrospective [J]. Genetics. 2010;186(3):767–73.
23. Consortium C-DG. o. t. P. G., Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis [J]. Lancet. 2013;381(9875):1371–9.
24. Liley J, Wallace C. A pleiotropy-informed Bayesian false discovery rate adapted to a shared control design finds new disease associations from GWAS summary statistics [J]. PLoS Genet. 2015;11(2):e1004926.
25. Johnston KJA, Adams MJ, Nicholl BI, Ward J, Strawbridge RJ, McIntosh AM, Smith DJ, Bailey MES. Identification of novel common variants associated with chronic pain using conditional false discovery rate analysis with major depressive disorder and assessment of pleiotropic effects of LRFN5 [J]. Transl Psychiatry. 2019;9(1):310.
26. Shadrin AA, Smeland OB, Zayats T, Schork AJ, Frei O, Bettella F, Witoelar A, Li W, Eriksen JA, Krull F, Djurovic S, Faraone SV, Reichborn-Kjennerud T, Thompson WK, Johansson S, Haavik J, Dale AM, Wang Y, Andreassen OA. Novel Loci associated with attention-deficit/hyperactivity disorder are revealed by leveraging polygenic overlap with educational attainment [J]. J Am Acad Child Adolesc Psychiatry. 2018;57(2):86–95.
27. Hu Y, Tan LJ, Chen XD, Liu Z, Min SS, Zeng Q, Shen H, Deng HW. Identification of novel potentially pleiotropic variants associated with osteoporosis and obesity using the cFDR Method [J]. J Clin Endocrinol Metab. 2018;103(1):125–38.
28. Hu Y, Tan LJ, Chen XD, Greenbaum J, Deng HW. Identification of novel variants associated with osteoporosis, type 2 diabetes and potentially pleiotropic loci using pleiotropic cFDR method [J]. Bone. 2018;117:6–14.
29. Zhang Q, Wu KH, He JY, Zeng Y, Greenbaum J, Xia X, Liu HM, Lv WQ, Lin X, Zhang WD, Xi YL, Shi XZ, Sun CQ, Deng HW. Novel common variants associated with obesity and type 2 diabetes detected using a cFDR method [J]. Sci Rep. 2017;7(1):16397.
30. Greenbaum J, Wu K, Zhang L, Shen H, Zhang J, Deng HW. Increased detection of genetic loci associated with risk predictors of osteoporotic fracture using a pleiotropic cFDR method [J]. Bone. 2017;99:62–8.
31. Liu HM, He JY, Zhang Q, Lv WQ, Xia X, Sun CQ, Zhang WD, Deng HW. Improved detection of genetic loci in estimated glomerular filtration rate and type 2 diabetes using a pleiotropic cFDR method [J]. Mol Gen Genomics. 2018;293(1):225–35.
32. Zhang Q, Liu HM, Lv WQ, He JY, Xia X, Zhang WD, Deng HW, Sun CQ. Additional common variants associated with type 2 diabetes and coronary artery disease detected using a pleiotropic cFDR method [J]. J Diabetes Complicat. 2018;32(12):1105–12.
33. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, Payne AJ, Steinthorsdottir V, Scott RA, Grarup N, Cook JP, Schmidt EM, Wuttke M, Sarnowski C, Mägi R, Nano J, Gieger C, Trompet S, Lecoeur C, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps [J]. Nat Genet. 2018;50(11):1505–13.
34. Lagou V, Mägi R, Hottenga J-J, Grallert H, Perry JRB, Bouatia-Naji N, Marullo L, Rybin D, Jansen R, Min JL, Dimas AS, Ulrich A, Zudina L, Gådin JR, Jiang L, Faggian A, Bonnefond A, Fadista J, Stathopoulou MG, et al. Sex-dimorphic genetic effects and novel loci for fasting glucose and insulin variability [J]. Nat Commun. 2021;12(1):24.
35. Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system [J]. Nat Protoc. 2013;8(8):1551–66.
36. Walters KA, Bertoldo MJ, Handelsman DJ. Evidence from animal models on the pathogenesis of PCOS [J]. Best Pract Res Clin Endocrinol Metab. 2018;32(3):271–81.
37. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator [J]. Genet Epidemiol. 2016;40(4):304–14.
38. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data [J]. PLoS Genet. 2017;13(11):e1007081.
39. Chen ZJ, Zhao H, He L, Shi Y, Qin Y, Shi Y, Li Z, You L, Zhao J, Liu J, Liang X, Zhao X, Zhao J, Sun Y, Zhang B, Jiang H, Zhao D, Bian Y, Gao X, et al. Genome-wide association study identifies susceptibility loci for polycystic ovary syndrome on chromosome 2p16.3, 2p21 and 9q33.3 [J]. Nat Genet. 2011;43(1):55–9.
40. Shi Y, Zhao H, Shi Y, Cao Y, Yang D, Li Z, Zhang B, Liang X, Li T, Chen J, Shen J, Zhao J, You L, Gao X, Zhu D, Zhao X, Yan Y, Qin Y, Li W, et al. Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome [J]. Nat Genet. 2012;44(9):1020–5.
41. Hayes MG, Urbanek M, Ehrmann DA, Armstrong LL, Lee JY, Sisk R, Karaderi T, Barber TM, McCarthy MI, Franks S, Lindgren CM, Welt CK, Diamanti-Kandarakis E, Panidis D, Goodarzi MO, Azziz R, Zhang Y, James RG, Olivier M, et al. Genome-wide association of polycystic ovary syndrome implicates alterations in gonadotropin secretion in European ancestry populations [J]. Nat Commun. 2015;6:7502.
42. Day FR, Hinds DA, Tung JY, Stolk L, Styrkarsdottir U, Saxena R, Bjonnes A, Broer L, Dunger DB, Halldorsson BV, Lawlor DA, Laval G, Mathieson I, McCardle WL, Louwers Y, Meun C, Ring S, Scott RA, Sulem P, et al. Causal mechanisms and balancing selection inferred from genetic associations with polycystic ovary syndrome [J]. Nat Commun. 2015;6:8464.
43. Wang Q, Huang T, Shu X, Zhao SG, Liang Y, Muhammad T, Gao F, Zhao H, Liu HB. Wilms' tumor 1 overexpression in granulosa cells is associated with polycystic ovaries in polycystic ovary syndrome patients [J]. Gynecol Obstet Investig. 2018;83(3):241–6.
44. Gao X. The role of FTO and nicotinamide in steroid hormone abnormality of polycystic ovary syndrome. Shandong University; 2020. p. 165.
45. Hastie ND. Wilms' tumour 1 (WT1) in development, homeostasis and disease [J]. Development. 2017;144(16):2862–72.
46. Paul L, Walker EM, Drosos Y, Cyphert HA, Neale G, Stein R, South J, Grosveld G, Herrera PL, Sosa-Pineda B. Lack of Prox1 downregulation disrupts the expansion and maturation of postnatal murine β-cells [J]. Diabetes. 2016;65(3):687–98.
47. Lakbakbi El Yaagoubi, F., H. Charoute, I. Morjane, H. Sefri, H. Rouba, A. Ainahi, M. Kandil, H. Benrahma, A. Barakat, Association analysis of genetic variants with metabolic syndrome components in the Moroccan population [J]. Current Research in Translational Medicine, 2017. 65(3): 121-125.
48. Goto T, Elbahrawy A, Furuyama K, Horiguchi M, Hosokawa S, Aoyama Y, Tsuboi K, Sakikubo M, Hirata K, Masui T, Kubo H, Sakai Y, Uemoto S, Kawaguchi Y. Liver-specific Prox1 inactivation causes hepatic injury and glucose intolerance in mice [J]. FEBS Lett. 2017;591(4):624–35.
49. Cataldi S, Costa V, Ciccodicola A, Aprile M. PPARγ and diabetes: beyond the genome and towards personalized medicine [J]. Current Diabetes Reports. 2021;21(6)
50. Costa V, Gallo MA, Letizia F, Aprile M, Casamassimi A, Ciccodicola A. PPARG: gene expression regulation and next-generation sequencing for unsolved issues [J]. PPAR Res. 2010;2010:409168.

51. Shaikh N, Mukherjee A, Shah N, Meherji P, Mukherjee S. Peroxisome proliferator activated receptor gamma gene variants influence susceptibility and insulin related traits in Indian women with polycystic ovary syndrome [J]. J Assist Reprod Genet. 2013;30(7):913–21.
52. Dasgupta S, Sirisha P, Neelaveni K, Anuradha K, Sudhakar G, Reddy BM. Polymorphisms in the IRS-1 and PPAR-γ genes and their association with polycystic ovary syndrome among South Indian women [J]. Gene. 2012;503(1):140–6.
53. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes [J]. Nat Genet. 2008;40(5):638–45.