Health Science Reports

WILEY

Health Science Reports
Open Access

ORIGINAL RESEARCH OPEN ACCESS

Multi-Omic Profiling of Gut Microbiota and Fecal Metabolites in Patients With Polycystic Ovary Syndrome: A Cross-Sectional Study

Yu-Mei Li[1] | Fang-Fang He[2] | Donge-Liu[1] | Bin-Xu[1] | Shuyi-Li[1]

[1]Department of Assisted Reproduction, Xiangya Hospital, Central South University, Changsha, P.R. China | [2]Reproductive Center of Chengdu Jinjiang District Maternal and Child Health Hospital, Chengdu, P.R. China

Correspondence: Yu-Mei Li (liyumei@csu.edu.cn)

Received: 20 March 2025 | **Revised:** 19 May 2026 | **Accepted:** 26 May 2026

Funding: Natural Science Foundation of Hunan Province, Grant/Award Number: 2022JJ70077

Keywords: fecal metabolomics | intestinal flora | metagenomics | polycystic ovary syndrome

ABSTRACT

Background and Aim: Intestinal flora composition in polycystic ovary syndrome (PCOS) varies, and the relationship between intestinal flora, fecal metabolites, clinical characteristics, and PCOS pathogenesis remains unclear. This study aimed to elucidate the gut microbiota characteristics of patients with PCOS, focusing on changes in normal-weight individuals, to provide new insights into its pathogenesis.

Methods: We combined 16S rRNA gene sequencing re-analysis with metagenomics and metabolomics to investigate gut microbiota and fecal metabolome alterations in PCOS. We re-analyzed our previous data on normal-weight women with PCOS (PCOS, $n = 24$; healthy controls [HC], $n = 12$) and the public databases (PCOS, $n = 98$; HC, $n = 71$) to further investigate the structure and function of the PCOS intestinal flora. Subsequently, from our previous study samples, we selected 10 patients residing in the Kaifu district, and their fecal samples (normal-weight PCOS group, $n = 6$; HC group, $n = 4$) were analyzed using metagenomic sequencing and non-targeted fecal metabolomics. Finally, the correlations among intestinal flora, fecal metabolites, and clinical indicators were evaluated.

Results: Based on the 16S rRNA data reanalysis, there were no significant differences in beta and alpha diversity between PCOS and normal controls. However, the PCOS group displayed a significantly higher relative abundance of Ruminococcus, Lachnospiraceae, and Escherichia-Shigella ($p < 0.05$) but a significantly lower relative abundance of Prevotella ($p < 0.05$) compared with the HC group. Subsequent metagenomics and metabolomics analyses revealed functional alterations, particularly in pathways related to secondary bile acid and lipid metabolism. Furthermore, Ruminococcus and Roseburia were positively correlated with Homeostasis Model Assessment of Insulin Resistance (HOMA-IR) and negatively correlated with high-density lipoprotein (HDL) in patients with normal-weight PCOS.

Conclusions: This study highlights gut microbial dysbiosis as a key feature of PCOS. Reanalysis of 16S rRNA data revealed specific taxonomic shifts without altering overall diversity, notably an enrichment of *Ruminococcus* and a depletion of

Abbreviations: AMH, Anti-Müllerian hormone; BCAA, Branched-chain amino acid; CRP, C-reactive protein; E2, Estradiol; FFA, Free fatty acid; FSH, Follicle-stimulating hormone; GLP-1, Glucagon-like peptide-1.; HDL, High-density lipoprotein; IL-6, Interleukin-6; IR, Insulin resistance; LDA, Linear discriminant analysis; LDL, Low-density lipoprotein; LH, Luteinizing hormone; OUT, Operational taxonomic unit; P, Progesterone; PRL: prolactin; PCOS, Polycystic ovary syndrome; SCFA, Short-chain fatty acid; SHBG, Sex hormone-binding globulin; T, Testosterone; TC, Total cholesterol; TG, Triglyceride; TNF-α, Tumor necrosis factor-alpha.

Yu-Mei Li and Fang-Fang He contributed equally to this work.

Yumei Li and Fangfang He are co-first authors for this study.

Prevotella. Furthermore, our metagenomics study identified functional reprogramming in pathways related to secondary bile acid and lipid metabolism. Crucially, even in normal-weight PCOS patients, these microbial alterations significantly correlated with adverse metabolic profiles (heightened insulin resistance and lower HDL levels), highlighting the microbiome as a potential therapeutic target.
Ethical Review No: CHiECRT1900028223.

1 | Background

Polycystic ovary syndrome (PCOS) is a multifaceted endocrine and metabolic condition [1, 2]. It is a primary cause of anovulatory infertility, accounting for roughly half to over two-thirds of such cases in women [3]. Furthermore, individuals with PCOS face an increased susceptibility to metabolic disturbances such as insulin resistance (IR), abnormal lipid levels, obesity, type 2 diabetes, and cardiovascular complications [4–6]. The precise causes and mechanisms underlying PCOS are still unclear. However, it is thought that a combination of factors, including neuroendocrine factors, genetic predisposition, immune system function, lifestyle choices, and the composition of the gut microbiome, contribute to its development [7].

The gut microbiota, often referred to as the human body's "second genome," influences metabolic processes and immune responses through its interactions with the external environment [8]. Emerging evidence suggests a notable link between the composition of the gut microbiota and the occurrence of PCOS [9]. However, findings from existing studies, which predominantly rely on 16S rRNA gene sequencing, have been inconsistent. While many studies report marked reductions in alpha diversity [10–14] and shifts in beta diversity [14, 15], others have found no significant changes [16, 17], including our own earlier work [18]. Similarly, conflicting results exist regarding the abundance of key genera like *Bacteroides* [14] [16] highlighting the limitations of current approaches and the potential influence of diverse populations and methodologies [19].

Crucially, 16S rRNA gene sequencing provides limited taxonomic resolution (typically at the genus level) and offers no direct insight into the functional capabilities of the microbiome. To overcome this, some studies have begun to integrate metabolomics. For instance, analyses of serum and fecal metabolites have linked shifts in secondary bile acids and short-chain fatty acids (SCFAs) to PCOS [15, 16, 20]. Yet, research combining deep microbial functional profiling with metabolomics in PCOS remains limited. Fecal metabolites, which more directly reflect the interplay between the host and gut flora, are particularly under-investigated in this context [21]. Therefore, a critical gap exists in our understanding of how specific microbial species—and their encoded metabolic functions—are linked to the distinct fecal metabolic signatures in PCOS.

To address this gap, the present study utilized a dual-tiered approach. First, we conducted a comprehensive reanalysis of 16S rRNA sequencing data (from our previous cohort and a public database) to robustly characterize taxonomic shifts in PCOS. Building upon these findings, our second objective was to perform an integrated multi-omics analysis, combining shotgun metagenomics with untargeted fecal metabolomics, in a targeted subset of normal-weight women with PCOS and healthy controls. We hypothesized that specific microbial species and functional pathways are altered in PCOS and are directly correlated with distinct fecal metabolite profiles. By moving beyond genus-level associations to a higher-resolution, function-centric view, this study aims to [1]: identify robust taxonomic shifts through 16S data reanalysis [2]; elucidate species-level and functional microbial signatures via metagenomics; and [3] uncover novel microbe-metabolite interactions that may contribute to its pathophysiology, providing a stronger foundation for potential diagnostic biomarkers and therapeutic targets.

2 | Methods

2.1 | Study Participants

For the initial 16S rRNA data reanalysis phase, participants were recruited between August and December 2019 from the Department of Assisted Reproduction at Xiangya Hospital, Central South University. The study group comprised 24 women diagnosed with PCOS, aged 18 to 35, and with a normal BMI. A control group of 12 women with normal BMI seeking assisted reproduction services at the same institution during the same timeframe was also recruited; these women's infertility was attributed to male factor or fallopian tube issues. Ethical approval was obtained from both the Ethics Committee of the Department of Assisted Reproduction (Xiangya Hospital, Central South University) and the China Registered Clinical Trial Ethics Review Committee (Ethical Review No: CHiECRT1900028223) prior to the commencement of the study, and all participants gave their written informed consent. The inclusion criteria specified a normal weight range, as determined by a BMI of 18.5–23.9 kg/m^2 [22, 23]. In our published article, we specified the inclusion and exclusion criteria, and the collection of specimens and clinical data [18], and the 16S rRNA sequencing data from this cohort was reanalyzed in the present study.

For the subsequent metagenomics and metabolomics phase, from our previously recruited patient cohort, we selected 10 female participants residing in Kaifu District, Changsha City. The fecal samples from the normal-weight PCOS group ($n = 6$) and the HC group ($n = 4$) were analyzed using metagenomic sequencing and non-targeted fecal metabolomics. Finally, the association between intestinal flora, fecal metabolites, and clinical indicators was investigated.

2.2 | Collection of Public Data

To robustly validate our 16S findings, raw sequencing data and supporting information for the 16S rRNA gene V3–V4 regions were obtained from the NCBI Sequence Read Archive database

under the accession numbers SRP085887 [10], PRJNA786067 [24], and SRP262631 [25]. All included patients met the following conditions: 1. All patients used the new PCOS diagnosis criteria from the 2003 Rotterdam Conference in the Netherlands; and 2. All patients adopted the 16S rRNA gene V3–V4 regions; however, the data did not differentiate between patients with normal-weight PCOS and those who were obese. In Supplementary Table S2, we summarize the basic attributes of four datasets. This includes a summary for each dataset regarding its sample size per group, geographical region, study population, baseline participant characteristics, and sequencing technology employed.

2.3 | DNA Extraction and PCR Amplification

2.3.1 | 16S rRNA Gene Sequencing

The DNA extraction and PCR amplification of the fecal samples and the Illumina MiSeq sequencing and preprocessing of data have been described in our previous studies [18]. Our raw data was analyzed along with other downloads of raw data from databases (PCOS = 122, HC = 83). Paired-end reads were demultiplexed and trimmed to remove barcode and primer sequences. The resulting paired reads, targeting the 16S rRNA gene, were then merged using FLASH (v1.2.8). Quality filtering was performed with fqtrim (v0.94) to generate high-quality reads. Vsearch (v2.3.4) was employed to remove chimeric sequences. DADA2 [26] was subsequently used for dereplication, resulting in a feature table and representative feature sequences. Alpha and beta diversity analyses were performed using QIIME2 [27] after rarefying the data to the lowest sequence count observed across all samples. This rarefied data was then used to calculate the relative abundance of bacterial taxa. QIIME2 was used for the alpha and beta diversity calculations, with visualizations generated using R (v3.5.2). Taxonomic assignment was performed via BLAST against the SILVA (Release 138,https://www.arb-silva.de/documentation/release138/) and NT-16S databases (https://www.ncbi.nlm.nih.gov/).

2.4 | Metagenomic Sequencing and Analysis

Metagenomic libraries were prepared and sequenced by LC-Bio Technology Co. Ltd. (Hangzhou, China) using the Illumina NovaSeq. 6000 platform with paired-end 150 bp reads. Initial sequence processing involved adapter removal with cutadapt (v1.9). Low-quality reads were then trimmed using fqtrim (v0.94) with a sliding-window approach, removing reads with more than 5% ambiguous bases ("N"). To eliminate host DNA contamination, quality-filtered reads were aligned against the host genome using bowtie (v2.2). The remaining reads were de novo assembled into contigs for each sample using MEGAHIT (v1.2.9) [28], which were subsequently utilized for taxonomic and functional annotation. Open reading frame (ORF) prediction was performed on the assembled contigs using MetaGeneMark (v3.26), and the predicted coding sequences (CDS) from all samples were clustered into a non-redundant gene catalog using CD-HIT (v4.6.1) [29]. Taxonomic classification of the microbial community was achieved using DIAMOND (v0.9.14) by aligning sequences against the NR database. Functional annotation was performed by mapping sequences to the KEGG database. The raw metagenomic sequencing data are available in the NCBI Sequence Read Archive under accession number PRJNA1177371.

2.5 | Analysis of Untargeted Metabolomics

Metabolites were extracted from the thawed samples using a chilled 80% methanol solution. Specifically, 50 mg of each sample was mixed with 0.5 mL of pre-cooled 80% methanol and incubated at −20°C for 30 min. Following centrifugation at 20,000 g for 15 min, the supernatant was collected and evaporated to dryness under vacuum. The resulting residue was reconstituted in 100 μL of 80% methanol, and the samples were stored at −80°C until LC-MS analysis. Quality control (QC) samples were generated by pooling 10 μL aliquots from each extraction. All samples were analyzed via LC-MS in a randomized order. Chromatographic separation was performed using a Thermo Fisher Scientific UltiMate 3000 UPLC system equipped with an ACQUITY UPLC T3 column (100 mm × 2.1 mm, 1.8 μm particle size, Waters). The column temperature was maintained at 40°C. Mobile phase A consisted of 5 mM ammonium acetate and 5 mM acetic acid in water, and mobile phase B was acetonitrile. The flow rate was set to 0.3 mL/min. The gradient elution program was as follows: 0–0.8 min, 2% B; 0.8–2.8 min, linear gradient from 2% to 70% B; 2.8–5.6 min, linear gradient from 70% to 90% B; 5.6–6.4 min, linear gradient from 90% to 100% B; 6.4–8.0 min, 100% B; 8.0–8.1 min, linear gradient from 100% to 2% B; 8.1–10 min, 2% B. A Q-Exactive high-resolution tandem mass spectrometer (Thermo Scientific) was used to acquire MS and MS/MS data. MS/MS spectra were acquired at a resolution of 17,500, with an automatic gain control (AGC) target of 1e5 and a maximum injection time of 50 ms. Stepped normalized collision energy (NCE) values of 20, 40, and 60 eV were applied for fragmentation. A TripleTOF 6600 high-resolution tandem mass spectrometer (SCIEX, Framingham, MA, USA) was used for the detection of metabolites eluting from the column. This Q-TOF system was operated in both positive and negative ion modes.

Mass spectrometry raw data (.raw files) were processed using Compound Discoverer 3.1.0 (Thermo Fisher Scientific, USA). The workflow included peak picking, retention time alignment (both within and between groups), adduct merging, gap filling, background noise filtering, and metabolite annotation. Metabolites were identified based on their retention time and m/z values. Peak intensities were recorded, and the resulting data, including feature molecular weight, retention time, peak area, and identification assignments, were exported. Metabolite annotation was performed by matching experimental molecular mass, names, and formulas against the online KEGG and HMDB databases. Further data preprocessing was conducted using the metaX package. Features detected in fewer than 50% of quality control (QC) samples or 80% of biological samples were excluded. Missing values in the remaining data were imputed using the k-nearest neighbor (KNN) algorithm. Principal component analysis (PCA) was employed to assess batch effects and identify potential outliers using the preprocessed dataset. Probabilistic Quotient Normalization (PQN) was applied to normalize ion intensities across samples.

2.6 | Statistical Analysis

All statistical analyses were performed using SPSS version 22.0. Quantitative demographic and clinical variables were assessed for normal distribution. Normally distributed data are presented as mean ± standard deviation (SD) and were compared between groups using an independent samples t-test. Data not following a normal distribution are expressed as median and interquartile range (IQR) and were compared using the Wilcoxon rank-sum test. Categorical demographic and clinical data are presented as counts and percentages and were compared using the Chi-square test. For all baseline comparisons, a two-tailed p-value < 0.05 was considered statistically significant.

Differential abundance analysis of microbial taxa was conducted using the Wilcoxon rank-sum test. A taxon was considered significantly different if it had a p-value < 0.05 and an absolute $\log_2$ (fold change) > 1. Spearman's rank correlation was used to assess the monotonic relationships between bacterial abundances and clinical parameters due to the non-normal distribution of the microbiome data. In addition, to explore the linear trends between specific genera of interest (*Ruminococcus* and *Roseburia)* and key metabolic markers (HDL, HOMA-IR), a supplementary Pearson correlation analysis was also performed. Linear discriminant analysis (LDA) effect size (LEfSe) was performed to identify taxa with differential abundances between the groups. An alpha value of 0.05 was used for the Kruskal-Wallis test, and the threshold for the logarithmic LDA score was set to 3.0.

For the analysis of metabolomic data, Student's t-tests were first used to identify differences in metabolite concentrations between the two phenotypes. The resulting p-values were then adjusted for multiple hypothesis testing using the Benjamini-Hochberg false discovery rate (FDR) procedure. A supervised partial least squares discriminant analysis (PLS-DA) was also performed using the metaX package to identify variables that could discriminate between the groups. Features with a variable importance in projection (VIP) score > 1.0 were considered important contributors to the model. Metabolic pathway enrichment analysis was performed on metabolites identified as significantly different, defined as those meeting the combined criteria of a fold change ≥ 2 or ≤ 0.5, an FDR-adjusted p-value ≤ 0.05, and a VIP score > 1.0. This analysis was conducted in MetaboAnalyst 5.0, using the *Homo sapiens* (human) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway library. The hypergeometric test was applied to identify pathways significantly enriched with differential metabolites, using a p-value < 0.05 as the threshold for significance.

TABLE 1 | Clinical, biochemical and hormonal features of participants. Data are presented as mean ± standard deviation. PCOS, polycystic ovary syndrome group ($n = 6$); HC, healthy control group ($n = 4$). *$p < 0.05$ versus HC group. p-values were calculated using independent-samples t-test, Wilcoxon rank-sum test or Chi-square test according to variable types.

Parametes	PCOS ($n = 6$)	HC ($n = 4$)
Age (years)	26.33 ± 1.36	27.25 ± 1.22
Menstrual cycle (days)	85.50 ± 31.32[a]	29.25 ± 1.5
BMI (kg/m^2)	21.46 ± 1.37	21.54 ± 1.32
Waist to hip ratio	0.85 ± 0.03	0.67 ± 0.01
FPG (mmol/L)	5.28 ± 1.67	5.04 ± 0.25
FINs (uU/mL)	6.94 ± 0.90	6.15 ± 1.57
HOMA-IR	3.15 ± 1.33[a]	1.37 ± 0.33
FSH (IU/L)	6.60 ± 1.56	6.81 ± 0.22
LH (IU/L)	16 ± 7.14[a]	5.65 ± 1.95
LH/FSH	2.38 ± 0.80[a]	0.83 ± 0.3
T (nmol/L)	1.60 ± 0.67[a]	0.67 ± 0.17
P (nmol/l)	0.93 ± 0.52B[a]	0.54 ± 0.16B
E2 (pmol/L)	231.16 ± 55.7	180.80 ± 48.57
AMH	7.74 ± 1.70[a]	4.21 ± 1.27
TBA (umol/l)	3.78 ± 1.19	3.27 ± 1.16
ALT (U/L)	23.05 ± 14.39	14 ± 3.39
AST (U/L)	21.53 ± 5.42	22.85 ± 7.35
Lithic (umol/L)	356.68 ± 43.29[a]	266.8 ± 20.95
TG (mmol/L)	1.33 ± 0.47	0.81 ± 0.21
TC (mmol/L)	4.38 ± 1.47	4.21 ± 1.04
HDL (mmol/L)	1.35 ± 0.22[a]	1.69 ± 0.20
LDL (mmol/L)	3.07 ± 0.55	2.98 ± 0.88
FFA (mmol/L)	0.48 ± 0.16	0.52 ± 0.17
hsCRP (mg/L)	0.8 ± 0.72	0.71 ± 0.57
IL-6 (pg/mL)	2.05 ± 0.26	1.19 ± 0.46
TNF-α (pg/mL)	4.95 ± 1.04	3.76 ± 1.43

Abbreviations: ALT, alanine aminotransferase; AMH, Anti-Mullerian hormone; AST, aspartate aminotransferase; BMI, body mass index; E2, Oestradiol; FINS, fasting plasma insulin; FFA, free fatty acid; FPG, fasting plasma glucose; FSH, follicular stimulating hormone; HDL, high density lipoprotein; HOMA-IR, homeostasis model assessment of insulin resistance; hsCRP, high-sensitivity C-reactive Protein; IL-6, Interleukin-6; LDL, low density lipoprotein; LH, luteinizing hormone; P, Progesterone; PCOS, polycystic ovary syndrome; PCOS-IR, PCOS with insulin resistance; PCOS-NIR, PCOS without insulin resistance; T, Testosterone; TBA, total bile acid; TC, total cholesterol; TG, Triglyceride; TNF-α: tumor necrosis factor-α; WHR, the ratio of waist to hip. [a]$p < 0.05$ for the HC versus PCOS group. p-values were calculated using the independent samples t-test, Wilcoxon rank-sum test, or Chi-square test, depending on the data type.

3 | Results

3.1 | Clinical Characteristics of the Participants

Table 1 summarizes the clinical, hormonal, and metabolic data of the targeted subset of 10 participants selected for the metagenomics and metabolomics analyses. Clinical data for the larger cohort used in the 16S reanalysis have been previously described [18]. There were no significant differences ($p > 0.05$) in age, height, or BMI among the two groups. The menstrual cycle length (in days) and HOMA-IR of the PCOS group were significantly higher than those of the healthy control group (HC). Regarding sex hormones, patients with PCOS exhibited higher levels of Testosterone(T), Luteinizing Hormone(LH), and LH/follicle-stimulating hormone (FSH) relative to the control patients ($p < 0.05$).

3.2 | Alterations of Gut Microbiota Composition in PCOS Based on the 16S rRNA Data

Rarefaction curves from the Operational Taxonomic Units (OTUs) indicated high sampling coverage (~ 99%) in all samples

(Figure S1A). The curve seemingly flattened, indicating that the sequencing depth was sufficient to reflect the diversity of the species in samples. Figure 1 displays the number of OTUs collected from the 205 samples. The Venn diagram indicates 2,947 and 2,518 unique OTUs in the PCOS and HC groups, respectively, with 1,697 OTUs common to the PCOS versus the normal control group (Figure 1A). Alpha diversity analysis revealed that variations in the Chao (Figure 1B) and Shannon indices (Figure S1B) between PCOS and HC groups were statistically nonsignificant. Beta diversity was analyzed using principal coordinates analysis (PCoA) analysis (Figure 1C), and Bray-Curtis (Figure S1C). The diversity analysis of the two groups displayed no significant differences.

Figure S1D depicts the phylum abundance distribution in the bacterial kingdom of the 205 samples. Compared with the HC group, the PCOS group exhibited significantly higher levels of Fusobacteria and Verrucomicrobiota (Figure 1D). Sankey plots demonstrate the relative abundance of microbial taxa at the phylum (center) and genus (right) levels for various samples (left) (Figure 1E). Based on genus-level bacterial community characteristics, a hierarchical heatmap illustrates the top 30 most abundant genera across all samples from both the PCOS and HC groups (Figure S1E). Furthermore, Figure S1F highlights 22 genera that were found to be significantly differentially abundant between the PCOS and HC groups, as determined using the Mann-Whitney U test ($p < 0.05$). At the genus level, *Blautia, Escherichia-Shigella, Ruminococcus, Fusobacterium*, and *Akkermansia* were significantly higher in PCOS than in controls, whereas *Paraprevotella, Desulfovibrim, Selenomonadaeae*, and *Streptococcus* spp. were lower. Linear discriminant analysis (LDA) distribution plot analysis (LDA score > 3) revealed that the predominant bacteria in PCOS were Lachnospiraceae, *Blautia*, Enterobacteriaceae, *Escherichia-Shigella*, *Ruminococcus*, and *Fusobacterium*. The main bacteria in HC were Selenomonadaeae, Rikenellaceae, *Streptococcus*, and *Paraprevotella* (Figure 1F).

3.3 | Functional Analysis Based on Metagenomic Sequencing

To further investigate the potential interaction effects between PCOS and the gut microbiome and identify whether gut microbial changes at the species level were associated with genes or functions of gut bacteria in the PCOS group, metagenomic sequencing was applied to this fecal samples of six PCOS and four matched HC. LDA revealed that Lachnospiraceae, *Ruminococcus, Fusobacterium*, and *Roseburia* were the important characteristic genera in the PCOS group compared with the control group (Figure 2A1, Figure 2A2). We annotated the fecal gene catalogs using the KEGG pathway. At KEGG level 3, the PCOS group displayed significant differences in secondary bile acid biosynthesis, proteasome, polycyclic aromatic hydrocarbon degradation, and the biosynthesis of type II polyketide products (Figure 2B). Additionally, we identified several KEGG orthologies that differed significantly in relative abundance between the groups, as illustrated in Figure S2A. It demonstrated significant enrichments in map01100 (metabolic pathways), map01110 (biosynthesis of secondary metabolites), map01120 (microbial metabolism in diverse environments), map01230 (biosynthesis of amino acids), map01200 (carbon metabolism), and map02010 (ABC transporters) (Figure S2B).

3.4 | Differential Metabolites and KEGG Pathways

After quality control, differential metabolites were identified between the PCOS and HC groups using variable importance for the projection (VIP) values of ≥ 1 and $p < 0.05$ as criteria (volcanic map of differential metabolites displayed in Figure 3A). A total of 59 distinct metabolites were detected in both negative (NEG) and positive (POS) modes. Specifically, compared with the HC group, 30 metabolites were up-regulated, while 29 were down-regulated in the PCOS group. We focused on 16 metabolites with VIP > 2, including hesperetin, 7Z, and 10Z-Hexadecadienoic acid. Table S1 displays detailed information on the metabolic products.

To identify the biological pathways most impacted by the observed metabolic changes, we performed a KEGG pathway enrichment analysis using the set of differentially abundant metabolites. The analysis revealed that several key pathways were significantly enriched in the PCOS group compared to the HC group (Figure 3B). Among the most significantly impacted pathways were 'Metabolic pathways' (ko01100), 'Thermogenesis' (ko04714), 'Linoleic acid metabolism' (ko00591), and 'Glycerolipid metabolism' (ko00561). Collectively, these results strongly indicate that metabolic perturbations in PCOS are driven by alterations in lipid and amino acid metabolism. To identify metabolic pathways altered in the PCOS group relative to the Healthy Control (HC) group, a Gene Set Enrichment Analysis (GSEA) was performed, and the top 30 most significantly enriched metabolite sets are visualized in Figure 3C. The PCOS group exhibited higher levels of nucleotide, linoleic acid, and purine metabolism than the control group (Figure 3D–F).

3.5 | Correlation Analysis

To investigate the potential associations of the significantly altered gut microbial taxa with host physiology, we performed a correlation analysis between their relative abundances and key clinical parameters (Figure 4). Spearman's correlation analysis was used to investigate the relationship between certain bacteria and metabolites and PCOS-related indicators (Figure 4A-B). *Ruminococcus* and *Roseburia* displayed a positive correlation with HOMA-IR but a negative correlation with HDL. Similarly, using Pearson correlation analysis, we examined the relationship between *Ruminococcus* (Figure S3A) and *Roseburia* (Figure S3B) with HDL and HOMA-IR. Serum LH levels were positively correlated with *Fusobacterium* abundance. Moreover, we performed Spearman's correlation analysis on the significantly different metabolites obtained through metabolomics and the significantly different species (at the species level) obtained by metagenomic sequencing analysis. The results are displayed in Figure S4.

4 | Discussion

Many studies currently comparing the gut flora and metabolite differences between patients with PCOS and HCs involve gut

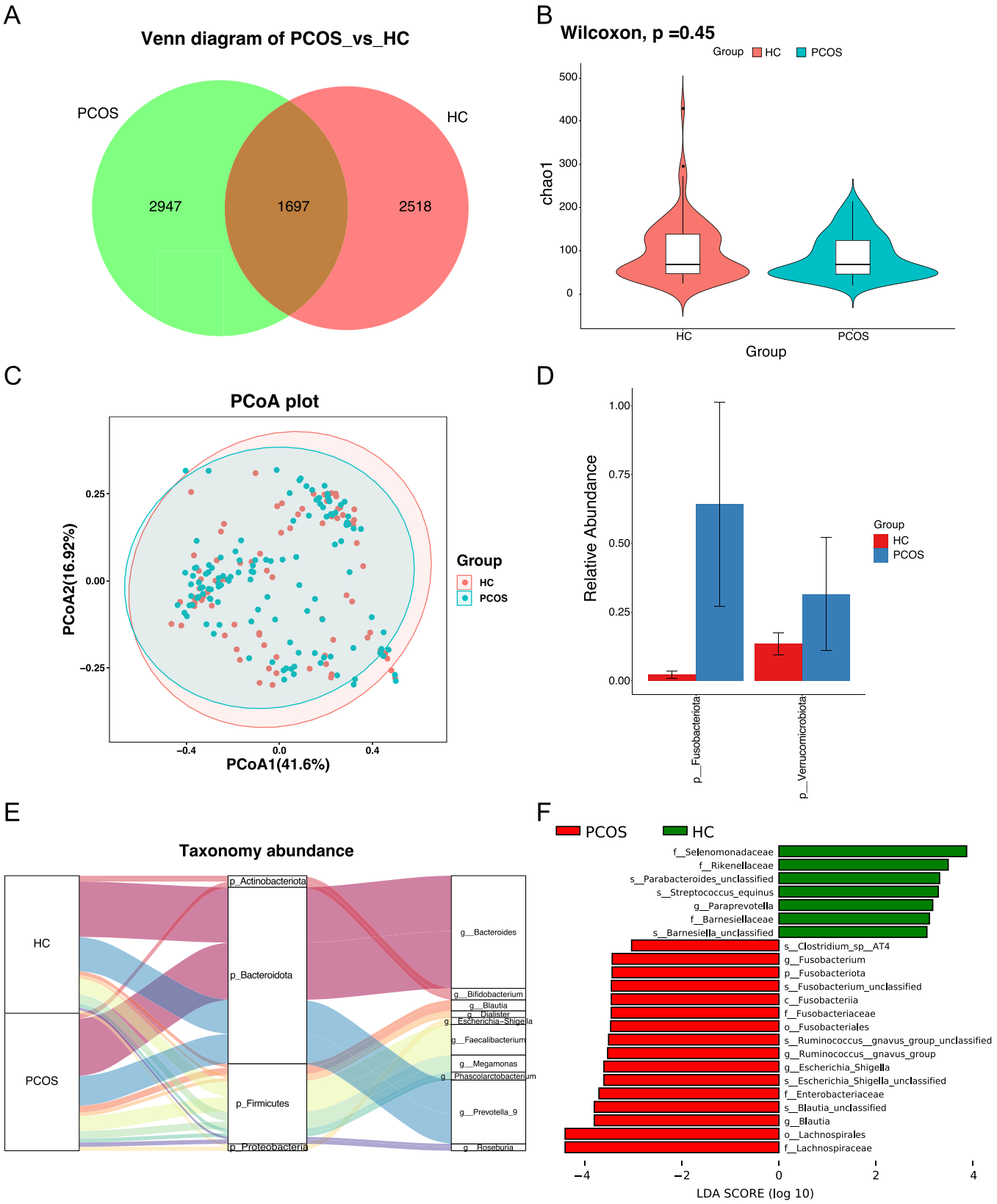

FIGURE 1 | Gut microbial composition in the combined cohort. Analysis was performed on 16S rRNA gene sequencing data from our in-house normal-weight PCOS cohort ($n = 24$) and healthy controls (HC, $n = 12$), combined with public datasets detailed in Table S2. (A) Venn diagram illustrating shared and unique Operational Taxonomic Units (OTUs) between the PCOS and HC groups. (B) Alpha diversity assessed by the Chao1 index. Group differences were evaluated using the Mann-Whitney U test; data are presented as violin plots. (C) Principal Coordinate Analysis (PCoA) plot based on the weighted UniFrac distance metric, visualizing the beta diversity between PCOS (red) and HC (blue) samples. (D) Bar plot showing the average relative abundance of major bacterial phyla. (E) Sankey plot illustrating the taxonomic flow from sample groups (left) to phylum (center) and genus (right) levels. (F) Linear Discriminant Analysis (LDA) Effect Size (LEfSe) histogram showing differentially abundant taxa.

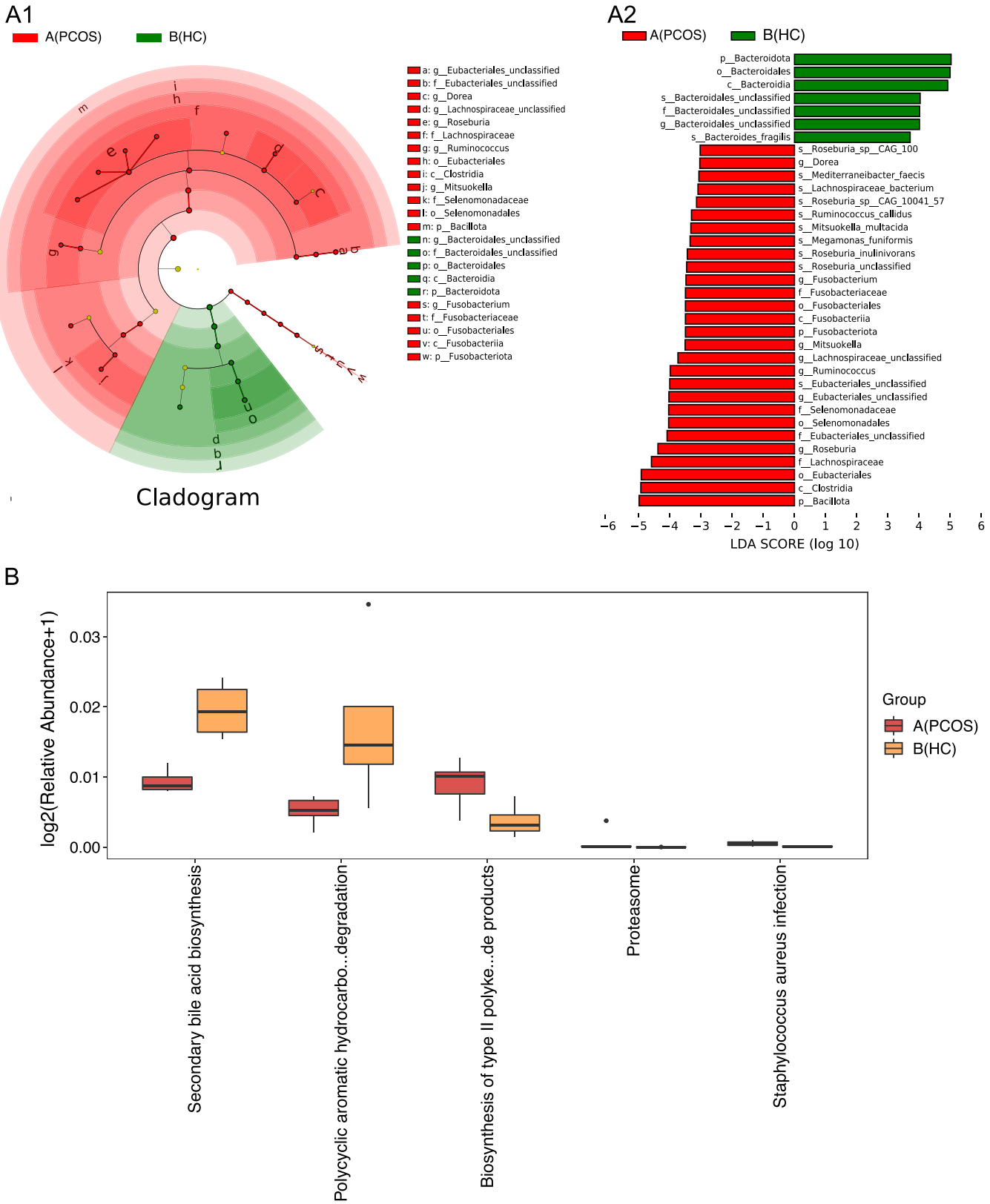

FIGURE 2 | Taxonomic Composition and Functional Pathways of the Microbial Community. Data presented in this figure were generated from shotgun metagenomic sequencing of our in-house normal-weight PCOS cohort (PCOS: $n = 6$, HC: $n = 4$). (A1) Taxonomic cladogram illustrating the phylogenetic distribution of the microbial taxa that were significantly more abundant in the respective groups. The concentric circles represent seven distinct taxonomic ranks, radiating from the innermost circle outwards: Kingdom, Phylum, Class, Order, Family, Genus, and Species. (A2) Histogram of the Linear Discriminant Analysis (LDA) scores for taxa with significant differential abundance. Only taxa with an LDA score > 3.0 are shown. (B) KEGG pathway enrichment column chart. The abscissa is the name of pathway level 3.

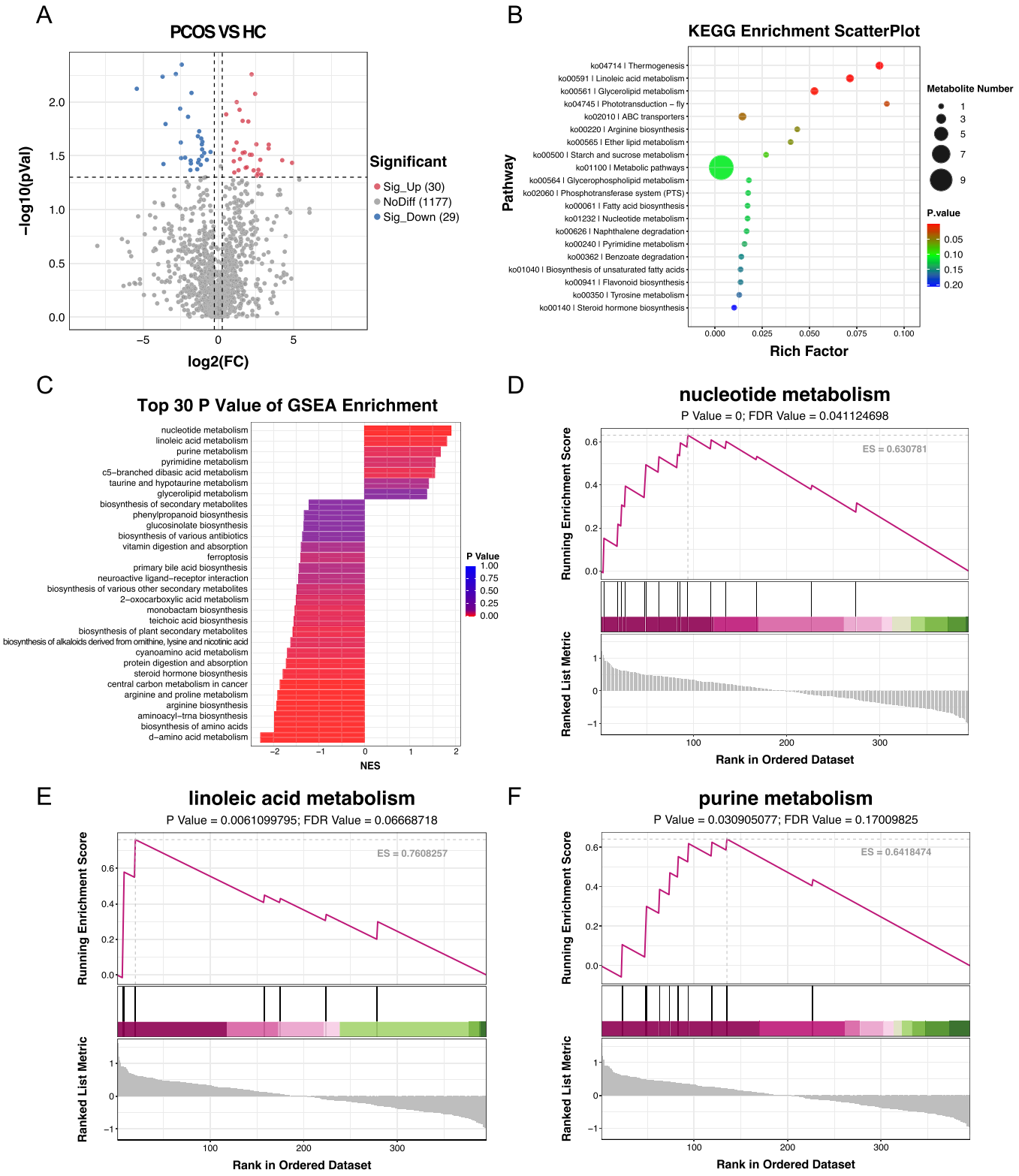

FIGURE 3 | Legend on next page.

microbial dysbiosis, which may trigger PCOS symptoms through many metabolic pathways, including those involving carbohydrates, short-chain fatty acids, lipopolysaccharides, bile acids, and the gut-brain axis. The gut microbiota may be involved in PCOS development with therapeutic potential [30]; however, the mechanisms remain poorly understood. Various observational studies of the gut microbiota at different levels of classification, such as phyla, have yielded conflicting results. Many scholars have used Mendelian randomization studies in recent years to investigate the differences in the gut microbiota between patients with PCOS and HCs, and the results of these studies are still being debated [31, 32]. While Yang et al. also investigated the gut microbiota in women with PCOS using publicly available data, our study differs in several key aspects [33]. First, we focused specifically on studies utilizing 16S

rRNA gene sequencing of the V3-V4 region to minimize methodological bias. Second, we applied stricter inclusion criteria based on PCOS diagnostic criteria to ensure greater homogeneity within the study populations. Although this resulted in a smaller number of included datasets compared to Yang et al., this approach allowed us to reduce potential confounding factors related to sequencing region and diagnostic heterogeneity, providing a more focused analysis. A supplementary table (Table S2) provides detailed characteristics of the datasets included in our re-analysis. In this study, we investigated the intestinal flora of patients with PCOS using 16S sequencing data from public databases and selected study data containing V3–V4 sequencing fragments. Additionally, we used metagenomics and metabolomics to sequence and analyze PCOS intestinal microorganisms and metabolites, revealing the function of intestinal flora in PCOS. Firmicutes, Bacteroidetes, Proteobacteria, and Actinobacteria were the main taxa. *Bacteroides* abundance was significantly higher in patients with PCOS than in HCs. Increased Firmicutes/Bacteroidetes ratios in patients with PCOS are closely associated with gut microbiota disruption and hormonal changes such as testosterone [34]. Studies have reported that patients with PCOS exhibit a lower percentage of relative abundance (RA) of Bacteroidetes [10, 35]. In our study, only *Fuso*bacteria and Verrucomicrobiota were discovered in abundance in patients with PCOS compared to the normal control group, consistent with the results of Chen [36]. Ruminococcaceae, Lachnospiraceae, and *Roseburia* were discovered to be more prevalent in women with obesity and metabolic syndrome [37]. A clinical study in women with gestational diabetes mellitus demonstrated that the relative abundance of Ruminococcaceae increased by twofold when compared with HCs [38]. We conducted a metagenomics investigation of normal-weight patients with PCOS and discovered a positive association between *Ruminococcus, Roseburia*, and HOMA-IR. *Ruminococcus spp*. can cause an increase in the production of inflammatory cytokines, which is positively correlated with disease state [39]. Zeng et al. discovered that Lachnospiraceae and Ruminococcaceae (both Firmicutes) were significantly increased in patients with PCOS-IR, which was consistent with the results of Eyupoglu [40]. At the family level, we discovered no significant difference in Ruminococcaceae; however, at the genus level, *Ruminococcus gnavus* increased in patients with PCOS, consistent with the study of Dong [41]. *Ruminococcus gnavus* derives energy from the epithelioglycans in the intestinal mucus layer. It was significantly higher in patients with inflammatory bowel disease than in the control group [42]. Liu et al. discovered that Ruminococcaceae were reduced in patients with PCOS, whereas Torres et al. discovered that the relative abundance of *Faecalibacterium prausnitzii* was higher in women with PCOS, whereas *Ruminococcus bromii* was lower. Therefore, Ruminococcaceae require further differentiation at the genus and species levels. Levels of *Lachnospiraceae UCG-008* and *Lachnospiraceae NK4A136* were significantly increased in the PCOS-IR rat model [38], and rather than generalizing the effects at the family level, we should evaluate changes at the genus or strain level.

Escherichia/Shigella was identified as a distinctive genus of PCOS in our study using LEfSe analysis, consistent with the results of a recent meta-analysis [43]. Several previous studies have demonstrated that patients with PCOS, especially those with obesity, have a higher percentage of *Escherichia/Shigella* in their gut [10, 13, 24]. This increase is negatively correlated with growth hormone-releasing peptide; growth hormone-releasing peptide and peptide YY (PYY) levels are lower in patients with PCOS compared with HCs [15]. Serotonin, PYY, and growth hormone-releasing peptides are mediators of the brain-gut axis. This could explain why patients with PCOS are more likely to develop depression than HCs. *Escherichia/Shigella* levels in the gut microbiota of patients with depression were significantly higher [44]. Using the SF-36 questionnaire survey [20], Yu et al. discovered that patients with PCOS exhibited a significantly lower quality of life score than the HC group. Further mechanistic studies are needed to demonstrate whether changes in the gut microbiota are associated with depressive tendencies.

As documented in our previous study [18], compared with healthy patients, patients with PCOS exhibited significantly reduced levels of *Prevotella*, an essential short-chain fatty acid-producing microbiota that regulates nutrition absorption and hormone levels in the gut. *Prevotella* produces anti-inflammatory metabolites that inhibit Th17 polarization and promote the differentiation of anti-inflammatory Treg/Tr1 cells in the gut [42]. F*urthermore*, *Prevotella* has been discovered to be involved in the synthesis of branched-chain amino acids [45]. Mice on a high-fat diet demonstrated higher BCAA levels after ingesting *Prevotella* for 2–3 weeks and different levels of IR. Previous studies have used Mendelian randomization to investigate the relationship between PCOS and the intestinal microbiota. Regardless of the level of the intestinal microbiota taxa, various observational studies have yielded contradictory results. This result can be attributed to differences in sample size, grouping criteria, and microbial identification methods. We collected 16S rRNA V3–V4 sequencing data from patients with PCOS and re-analyzed the sequencing data, discovering that there were still discrepancies, if not opposing results, when compared

FIGURE 3 | Differential Metabolite and Pathway Enrichment Analysis Comparing PCOS Patients to Healthy Controls (PCOS: $n = 6$, HC: $n = 4$). (A)The volcano plot illustrates metabolites identified between the PCOS and HC groups. The y-axis represents the -log10 (p-value), and the x-axis represents the log2 (Fold Change). Each dot represents a metabolite. Red dots indicate significantly up-regulated metabolites (log2FC ≥ 1, $p < 0.05$), blue dots indicate significantly down-regulated metabolites (log2FC ≤ -1, $p < 0.05$), and gray dots represent non-significant metabolites. (B) Bubble plot of the KEGG pathway enrichment analysis for differentially abundant metabolites. The y-axis shows the significant pathway names. The x-axis represents the pathway enrichment ratio (Impact Factor). The size of each bubble corresponds to the number of matched metabolites in the pathway, and the color scale reflects the p-value from the enrichment analysis, with darker red indicating higher statistical significance. (C) Gene Set Enrichment Analysis (GSEA) of KEGG pathways. The plot displays the top 30 metabolite sets from the comparison between the PCOS and Healthy Control groups, ranked by p-value and FDR. The x-axis represents the Normalized Enrichment Score (NES) for each set, the y-axis lists the KEGG metabolite set names, and the color scale indicates the nominal p-value. (D, E, F) Gene Set Enrichment Analysis Enrichment Score Plots. Each plot corresponds to a specific metabolite set, showing the Enrichment Score (ES) and FDR values. The plots include a line chart of the ES values, a mark of metabolite set members in the ranked metabolite list, and a heat map of ranked metabolite expression values (red for treatment, green for control).

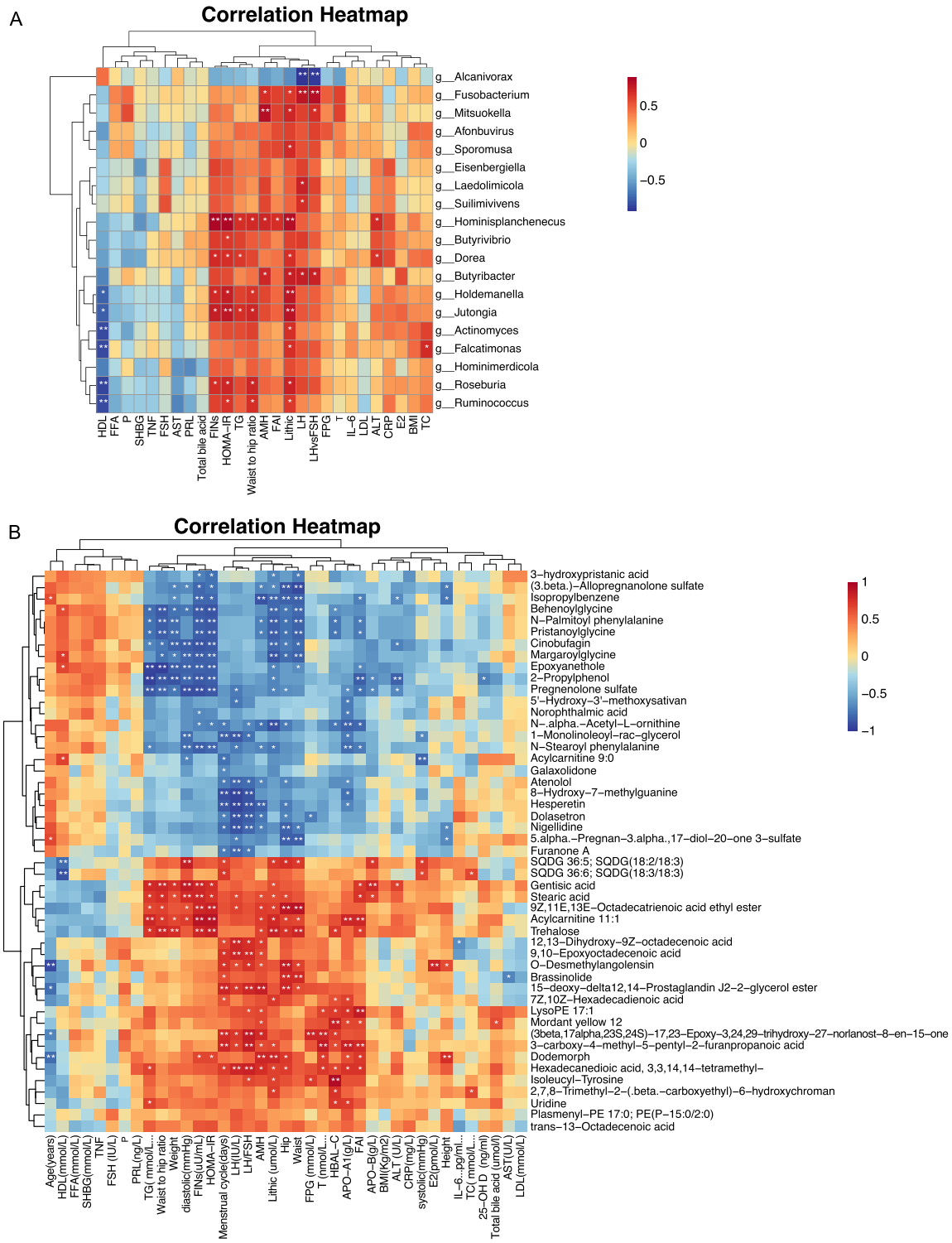

FIGURE 4 | Spearman correlation heatmap between host clinical parameters and the relative abundance of significantly altered fecal microbial taxa. (A) Heatmap of Spearman's rank correlations between host clinical parameters and the relative abundance of the top 20 most abundant microbial genera. (B) Heatmap of Spearman's rank correlations between clinical parameters and the top 50 most abundant fecal metabolites. In both heatmaps, the color scale represents the Spearman's correlation coefficient (R), with red indicating a positive correlation and blue indicating a negative correlation. Asterisks denote the level of statistical significance: *$p < 0.05$, **$p < 0.01$.

with the results of a single study. This suggests that future research should exclude confounding factors such as people in different regions and with different dietary patterns to further explore the changes in intestinal flora in PCOS and provide a theoretical basis for PCOS treatment.

Simultaneously, we have extended our previous studies to further analyze the gut flora and fecal metabolites of patients with PCOS using metagenomics and metabolomics. The PCOS group exhibited significant enrichment compared to the control group in secondary bile acid biosynthesis, proteasome, polycyclic

aromatic hydrocarbon degradation, biosynthesis of type II polyketide products, and the upregulation of nucleotide, linoleic acid, and purine metabolism. Compared with the control group, the PCOS group demonstrated higher levels of nucleotide, linoleic acid, and purine metabolisms. Previous studies have mostly focused on the interaction between intestinal flora and serum metabolites in patients with PCOS. Yang et al. [46] discovered that three metabolites [Ganglioside GM3, Cer (d16:2/22:0), and 3Z,6Z,9Z-Pentacosatriene] were associated with the intestinal flora of PCOS. Yu et al. [20] identified 15 metabolites that differed in serum between PCOS and normal controls. Lysophosphatidylcholine concentration was negatively correlated with *Prevotella*_9 and positively correlated with PCOS characteristic group 0319_6G20. It is proposed that certain gut flora and metabolites may play an important role in insulin resistance and mood change in patients with PCOS. However, in a mouse model of PCOS, intestinal metabolites were more predictive of disease than intestinal bacterial characteristics [47]. Zhou et al. used fecal metabolomics to identify seven metabolites that are characteristic of patients with obesity and PCOS. Correlation analysis revealed a linear relationship between intestinal flora, characteristic fecal metabolites, and serum sex hormones. This suggests that gut bacteria *Veillonella* and *Lachnospira*, through the arachidonic acid metabolic pathway, may influence blood glucose and cholesterol levels in women with obesity and PCOS. However, we discovered several metabolite differences, including decreased Hesperetin and elevated 7Z,10Z-Hexadecadienoic acid in patients with PCOS; Hesperetin was negatively correlated with the LH and AMH levels in patients with PCOS.

5 | Study Limitations

Several limitations of this study should be acknowledged. First, while our re-analysis of public 16S rRNA gene sequencing data increased the overall sample size, the inherent heterogeneity across studies (e.g., in sequencing protocols, patient populations, and diagnostic criteria) precluded a detailed, stratified analysis of different PCOS phenotypes. This challenge highlighted the need for our more focused, in-house investigation. Second, the sample size of our multi-omics cohort was modest **(total $n = 10$; 6 PCOS and 4 HC)**, which, as is common in exploratory multi-omics studies, may limit the statistical power and generalizability of our metabolomic and correlation findings. Consequently, these results should be considered preliminary. Although we controlled for geographic location, other potential confounders such as detailed dietary habits could not be fully accounted for. Finally, the cross-sectional design of our study allows us to identify strong associations, but it cannot establish causality. Therefore, future longitudinal studies with larger, well-phenotyped cohorts are necessary to validate these findings and elucidate the precise causal mechanisms linking the gut microbiome to PCOS pathophysiology.

6 | Conclusion

In conclusion, our reanalysis of 16S rRNA data identified a distinct gut microbial dysbiosis in PCOS, characterized by increased Ruminococcus and Escherichia-Shigella alongside depleted Prevotella. Building on this, our metagenomics study revealed a functional reprogramming of the microbiome, particularly in secondary bile acid and lipid metabolism pathways. Furthermore, in normal-weight women with PCOS, specific taxa like Ruminococcus and Roseburia correlated directly with metabolic dysfunctions, including heightened HOMA-IR and lower HDL cholesterol. These findings demonstrate that the gut microbiome actively contributes to PCOS pathogenesis and presents a promising therapeutic target for novel interventions aimed at improving metabolic health in affected women.

Author Contributions

Yu-Mei Li: conceptualization, investigation, funding acquisition, writing – review and editing, validation, methodology, project administration, resources, supervision, data curation. **Fang-Fang He:** investigation, writing – original draft, writing – review and editing, visualization, validation, methodology, software, project administration, supervision, data curation. **Donge-Liu:** writing – review and editing, investigation, conceptualization, methodology, data curation, supervision. **Bin-Xu:** data curation, supervision, writing – review and editing, project administration, visualization, investigation. **Shuyi-Li:** data curation, supervision, resources, project administration, visualization, software, investigation.

Acknowledgments

We are grateful to the women who participated in the survey. This study was supported by Natural Science Foundation of Hunan Provincial (No. 2022JJ70077). The funding source had no role in the study design, collection, analysis, or interpretation of data, the writing of the report, or the decision to submit the article for publication.

Ethics Statement

This study was approved by the Ethics Committee of the Department of Assisted Reproduction (Xiangya Hospital, Central South University) and carried out in accordance with the ethical principles for medical research involving human subjects addressed in the Helsinki Declaration, Written informed consent was obtained from all individual participants included in the study. The China Registered Clinical Trial Ethics Review Committee (Ethical Review No.: CHiECRT1900028223).

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. The datasets are available from the corresponding author on reasonable request. Raw sequencing reads have been deposited in the NCBI Sequence Read Archive under the accession number PRJNA1177371.

Transparency Statement

They affirm that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

References

1. S. B. Sherman, N. Sarsour, M. Salehi, et al., "Prenatal Androgen Exposure Causes Hypertension and Gut Microbiota Dysbiosis," *Gut Microbes* 9 (2018): 1–22.

2. L. Qin, C. Huang, X. Yan, Y. Wang, Z. Li, and X. Wei, "Long Non-Coding RNA H19 Is Associated With Polycystic Ovary Syndrome in Chinese Women: A Preliminary Study," *Endocrine Journal* 66, no. 7 (2019): 587–595.

3. R. J. Norman, D. Dewailly, R. S. Legro, and T. E. Hickey, "Polycystic Ovary Syndrome," *Lancet* 370, no. 9588 (2007): 685–697.

4. D. Macut, J. Bjekić-Macut, D. Rahelić, and M. Doknić, "Insulin and the Polycystic Ovary Syndrome," *Diabetes Research and Clinical Practice* 130 (2017): 163–170.

5. N. F. Goodman, R. H. Cobin, W. Futterweit, et al., "American Association of Clinical Endocrinologists, American College of Endocrinology, and Androgen Excess and PCOS Society Disease State Clinical Review: Guide to the Best Practices in the Evaluation and Treatment of Polycystic Ovary Syndrome—Part 2," *Endocrine Practice* 21, no. 12 (2015): 1415–1426.

6. Y. Guo, Y. Qi, X. Yang, et al., "Association between Polycystic Ovary Syndrome and Gut Microbiota," *PLoS One* 11, no. 4 (2016): e0153196.

7. R. L. Rosenfield and D. A. Ehrmann, "The Pathogenesis of Polycystic Ovary Syndrome (Pcos): The Hypothesis of PCOS as Functional Ovarian Hyperandrogenism Revisited," *Endocrine Reviews* 37, no. 5 (2016): 467–520.

8. K. Z. Coyte, J. Schluter, and K. R. Foster, "The Ecology of the Microbiome: Networks, Competition, and Stability," *Science* 350, no. 6261 (2015): 663–666.

9. K. Tremellen and K. Pearce, "Dysbiosis of Gut Microbiota (DOGMA)—A Novel Theory for the Development of Polycystic Ovarian Syndrome," *Medical Hypotheses* 79, no. 1 (2012): 104–112.

10. R. Liu, C. Zhang, Y. Shi, et al. Dysbiosis of Gut Microbiota Associated With Clinical Parameters in Polycystic Ovary Syndrome. 2017;8(324).

11. M. Insenser, M. Murri, R. Del Campo, M. Á. Martínez-García, E. Fernández-Durán, and H. F. Escobar-Morreale, "Gut Microbiota and the Polycystic Ovary Syndrome: Influence of Sex, Sex Hormones, and Obesity," *The Journal of Clinical Endocrinology & Metabolism* 103, no. 7 (2018): 2552–2562.

12. P. J. Torres, M. Siakowska, B. Banaszewska, et al., "Gut Microbial Diversity in Women With Polycystic Ovary Syndrome Correlates With Hyperandrogenism," *The Journal of Clinical Endocrinology & Metabolism* 103, no. 4 (2018): 1502–1511.

13. B. Zeng, Z. Lai, L. Sun, et al., "Structural and Functional Profiles of the Gut Microbial Community in Polycystic Ovary Syndrome With Insulin Resistance (IR-PCOS): A Pilot Study," *Research in Microbiology* 170, no. 1 (2019): 43–52.

14. L. Lindheim, M. Bashir, J. Münzker, et al., "Alterations in Gut Microbiome Composition and Barrier Function Are Associated With Reproductive and Metabolic Defects in Women With Polycystic Ovary Syndrome (PCOS): A Pilot Study," *PLoS One* 12, no. 1 (2017): e0168390.

15. J. Zhang, Z. Sun, S. Jiang, et al., "Probiotic Bifidobacterium Lactis V9 Regulates the Secretion of Sex Hormones in Polycystic Ovary Syndrome Patients through the Gut-Brain Axis," *mSystems* 4, no. 2 (2019): e00017-19.

16. L. Zhou, Z. Ni, W. Cheng, et al., "Characteristic Gut Microbiota and Predicted Metabolic Functions in Women With PCOS," *Endocrine Connections* 9, no. 1 (2020): 63–73.

17. X. Qi, C. Yun, L. Sun, et al., "Gut Microbiota-Bile Acid-Interleukin-22 Axis Orchestrates Polycystic Ovary Syndrome," *Nature Medicine* 25, no. 8 (2019): 1225–1233.

18. F. He and Y. Li, "The Gut Microbial Composition in Polycystic Ovary Syndrome With Insulin Resistance: Findings From a Normal-Weight Population," *Journal of Ovarian Research* 14, no. 1 (2021): 50.

19. P. Li, P. Shuai, S. Shen, et al., "Perturbations in Gut Microbiota Composition in Patients With Polycystic Ovary Syndrome: A Systematic Review and Meta-Analysis," *BMC Medicine* 21, no. 1 (2023): 302.

20. Z. Yu, E. Qin, S. Cheng, et al., "Gut Microbiome in PCOS Associates to Serum Metabolomics: A Cross-Sectional Study," *Scientific Reports* 12, no. 1 (2022): 22184.

21. X. Bai, J. Ma, X. Wu, et al., "Impact of Visceral Obesity on Structural and Functional Alterations of Gut Microbiota in Polycystic Ovary Syndrome (PCOS): A Pilot Study Using Metagenomic Analysis," *Diabetes, Metabolic Syndrome and Obesity* 16 (2023): 1–14.

22. W. He, Q. Li, M. Yang, et al., "Lower BMI Cutoffs to Define Overweight and Obesity in China," *Obesity* 23, no. 3 (2015): 684–691.

23. E. E. Oakkar, J. Stevens, K. P. Truesdale, and J. Cai, "BMI and All-Cause Mortality Among Chinese and Caucasians: The People's Republic of China and the Atherosclerosis Risk in Communities Studies," *Asia Pacific Journal of Clinical Nutrition* 24, no. 3 (2015): 472–479.

24. G. Yin, F. Chen, G. Chen, et al., "Alterations of Bacteriome, Mycobiome and Metabolome Characteristics in PCOS Patients With Normal/Overweight Individuals," *Journal of Ovarian Research* 15, no. 1 (2022): 117.

25. L. Zhou, Z. Ni, J. Yu, W. Cheng, Z. Cai, and C. Yu, "Correlation between Fecal Metabolomics and Gut Microbiota in Obesity and Polycystic Ovary Syndrome," *Frontiers in Endocrinology* 11 (2020): 628.

26. B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes, "DADA2: High-Resolution Sample Inference From Illumina Amplicon Data," *Nature Methods* 13, no. 7 (2016): 581–583.

27. E. Bolyen, J. R. Rideout, M. R. Dillon, et al., "Author Correction: Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2," *Nature Biotechnology* 37, no. 9 (2019): 1091.

28. D. Li, C. M. Liu, R. Luo, K. Sadakane, and T. W. Lam, "MEGAHIT: An Ultra-Fast Single-Node Solution for Large and Complex Metagenomics Assembly via Succinct de Bruijn Graph," *Bioinformatics* 31, no. 10 (2015): 1674–1676.

29. L. Fu, B. Niu, Z. Zhu, S. Wu, and W. Li, "CD-HIT: Accelerated for Clustering the Next-Generation Sequencing Data," *Bioinformatics* 28, no. 23 (2012): 3150–3152.

30. X. Zhao, Y. Jiang, H. Xi, L. Chen, and X. Feng, "Exploration of the Relationship Between Gut Microbiota and Polycystic Ovary Syndrome (PCOS): A Review," *Geburtshilfe und Frauenheilkunde* 80, no. 2 (2020): 161–171.

31. J. Wang, P. L. Fiori, G. Capobianco, C. Carru, and Z. Chen, "Gut Microbiota and Polycystic Ovary Syndrome, Focus on Genetic Associations: A Bidirectional Mendelian Randomization Study," *Frontiers in Endocrinology* 15 (2024): 1275419.

32. J. Sun, M. Wang, and Z. Kan, "Causal Relationship between Gut Microbiota and Polycystic Ovary Syndrome: A Literature Review and Mendelian Randomization Study," *Frontiers in Endocrinology* 15 (2024): 1280983.

33. B. Jobira, D. N. Frank, L. Pyle, et al., "Obese Adolescents With PCOS Have Altered Biodiversity and Relative Abundance in Gastrointestinal Microbiota," *Journal of Clinical Endocrinology & Metabolism* 105, no. 6 (2020): e2134–e2144.

34. B. Jobira, D. N. Frank, L. J. Silveira, et al., "Hepatic Steatosis Relates to Gastrointestinal Microbiota Changes in Obese Girls With Polycystic Ovary Syndrome," *PLoS One* 16, no. 1 (2021): e0245219.

35. F. Chen, Z. Chen, M. Chen, et al., "Reduced Stress-Associated FKBP5 DNA Methylation Together With Gut Microbiota Dysbiosis Is Linked With the Progression of Obese PCOS Patients," *NPJ Biofilms and Microbiomes* 7, no. 1 (2021): 60.

36. A. Chávez-Carbajal, K. Nirmalkar, A. Pérez-Lizaur, et al., "Gut Microbiota and Predicted Metabolic Pathways in a Sample of Mexican

Women Affected by Obesity and Obesity Plus Metabolic Syndrome," *International Journal of Molecular Sciences* 20, no. 2 (2019): 438.

37. K. Mokkala, N. Houttu, T. Vahlberg, E. Munukka, T. Rönnemaa, and K. Laitinen, "Gut Microbiota Aberrations Precede Diagnosis of Gestational Diabetes Mellitus," *Acta Diabetologica* 54, no. 12 (2017): 1147–1149.

38. M. Gurung, Z. Li, H. You, et al., "Role of Gut Microbiota in Type 2 Diabetes Pathophysiology," *EBioMedicine* 51 (2020): 102590.

39. N. D. Eyupoglu, K. Ergunay, A. Acikgoz, Y. Akyon, E. Yilmaz, and B. O. Yildiz, "Gut Microbiota and Oral Contraceptive Use in Overweight and Obese Patients With Polycystic Ovary Syndrome," *Journal of Clinical Endocrinology & Metabolism* 105, no. 12 (2020): e4792–e4800.

40. S. Dong, J. Jiao, S. Jia, et al., "16S rDNA Full-Length Assembly Sequencing Technology Analysis of Intestinal Microbiome in Polycystic Ovary Syndrome," *Frontiers in Cellular and Infection Microbiology* 11 (2021): 634981.

41. A. B. Hall, M. Yassour, J. Sauk, et al., "A Novel *Ruminococcus* Gnavus Clade Enriched in Inflammatory Bowel Disease Patients," *Genome Medicine* 9, no. 1 (2017): 103.

42. Q. Zhu and N. Zhang, "Gut Microbiome Composition in Polycystic Ovary Syndrome Adult Women: A Systematic Review and Meta-Analysis of Observational Studies," *Reproductive Sciences* 31 (2024): 1800–1818.

43. H. Rong, X. Xie, J. Zhao, et al., "Similarly in Depression, Nuances of Gut Microbiota: Evidences from a Shotgun Metagenomics Sequencing Study on Major Depressive Disorder Versus Bipolar Disorder With Current Major Depressive Episode Patients," *Journal of Psychiatric Research* 113 (2019): 90–99.

44. H. K. Pedersen, V. Gudmundsdottir, H. B. Nielsen, et al., "Human Gut Microbes Impact Host Serum Metabolome and Insulin Sensitivity," *Nature* 535 (2016): 376–381.

45. Z. Yang, H. Fu, H. Su, et al., "Multi-Omics Analyses Reveal the Specific Changes in Gut Metagenome and Serum Metabolome of Patients With Polycystic Ovary Syndrome," *Frontiers in Microbiology* 13 (2022): 1017147.

46. B. Ho, D. Ryback, B. Benson, et al., "Gut Metabolites Are More Predictive of Disease and Cohoused States Than Gut Bacterial Features in a Polycystic Ovary Syndrome-Like Mouse Model," *mSystems* 6, no. 5 (2021): e0114920.

47. Y. Gao, L. Yang, Q. Yao, J. Wang, and N. Zheng, "Butyrate Improves Recovery from Experimental Necrotizing Enterocolitis by Metabolite Hesperetin through Potential Inhibition the PI3K-Akt Pathway," *Biomedicine & Pharmacotherapy* 176 (2024): 116876.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Figure S1: Analysis of the composition of intestinal microflora ($N = 24$ cases, $N = 12$ controls). (A) Rank abundance curve. (B) Alpha diversity assessed by the Shannon index. Group differences were evaluated using the Mann-Whitney U test. (C) Non-metric Multidimensional Scaling (NMDS) analysis based on the Bray-Curtis distance. (D) Bar chart of relative abundance at the phylum level. (E) Hierarchical clustering heatmap of the 30 most abundant bacterial genera. (F)Differentially abundant bacterial genera between PCOS and HC groups (Mann-Whitney U test, $p < 0.05$).

Figure S2: Functional Analysis of Differential Features Using KEGG Pathways (PCOS: $n = 6$, HC: $n = 4$). (A) Bar chart showing the relative abundance of differentially abundant KEGG Orthologies (KOs). The x-axis lists the KOs, and the y-axis represents the log2-transformed relative abundance. Group comparisons were performed using a t-test. (B) Bubble plot of KEGG pathway enrichment analysis. The Rich Factor (x-axis) is the ratio of the number of differential features to the total number of features in a given pathway. The bubble size corresponds to the number of differential features, and the color indicates the enrichment p-value.

Figure S3: Correlation of *Ruminococcus* and *Roseburia* with metabolic markers. Scatter plots show the Pearson correlation between the relative abundance of (A) *Ruminococcus* and (B) *Roseburia* with HDL cholesterol and HOMA-IR across all subjects in the in-house cohort ($n = 10$). The solid line represents the linear regression fit, and the shaded area indicates the 95% confidence interval. The Pearson's correlation coefficient (r) and *p*-value are displayed on each plot.

Figure S4: Correlation map among gut microbiota, fecal metabolites (PCOS: $n = 6$, HC: $n = 4$). The color of spots represents the Spearman correlation R-value between each bacterial taxa and clinical parameters. The color scale indicates the R-value, with red representing positive correlations and blue representing negative correlations. Statistical significance is denoted by symbols: $^{*}p < 0.05$, $^{**}p < 0.01$.

Table S1: Detailed information on the differential fecal metabolites between PCOS patients and healthy controls.The table lists metabolites with a $|\text{Log2(Fold Change)}| \geq 2$. *p*-values were calculated using a t-test, and VIP (Variable Importance in Projection) scores were derived from PLS-DA analysis.

Table S2: Characteristics of Datasets Included in Re-analysis.This table summarizes the key features of the publicly available 16S rRNA sequencing datasets used in the combined analysis presented in Figure 1.