Combining a nontargeted and targeted metabolomics approach to identify metabolic pathways significantly altered in polycystic ovary syndrome

Alice Y. Chang a,⁎, Antigoni Z. Lalia a, Gregory D. Jenkins b, Tumpa Dutta a, Rickey E. Carter b, Ravinder J. Singh c, K. Sreekumaran Naira a

a Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Mayo Clinic, Rochester, MN
b Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN
c Division of Clinical Biochemistry and Immunology, Mayo Clinic, Rochester, MN

ARTICLE INFO
Article history:
Received 26 April 2016
Accepted 1 March 2017

Keywords:
α-Aminoadipic acid
Branched-chain amino acids
Insulin sensitivity
Metabolic syndrome
Vitamin D

ABSTRACT

Objective. Polycystic ovary syndrome (PCOS) is a condition of androgen excess and chronic anovulation frequently associated with insulin resistance. We combined a nontargeted and targeted metabolomics approach to identify pathways and metabolites that distinguished PCOS from metabolic syndrome (MetS).

Methods. Twenty obese women with PCOS were compared with 18 obese women without PCOS. Both groups met criteria for MetS but could not have diabetes mellitus or take medications that treat PCOS or affect lipids or insulin sensitivity. Insulin sensitivity was derived from the frequently sampled intravenous glucose tolerance test. A nontargeted metabolomics approach was performed on fasting plasma samples to identify differentially expressed metabolites, which were further evaluated by principal component and pathway enrichment analysis. Quantitative targeted metabolomics was then applied on candidate metabolites. Measured metabolites were tested for associations with PCOS and clinical variables by logistic and linear regression analyses.

Results. This multiethnic, obese sample was matched by age (PCOS, 37 ± 6; MetS, 40 ± 6 years) and body mass index (BMI) (PCOS, 34.6 ± 5.1; MetS, 33.7 ± 5.2 kg/m²). Principal component analysis of the nontargeted metabolomics data showed distinct group separation of PCOS from MetS controls. From the subset of 385 differentially expressed metabolites, 22% were identified by accurate mass, resulting in 19 canonical pathways significantly altered in PCOS, including amino acid, lipid, steroid, carbohydrate, and vitamin D metabolism. Targeted metabolomics identified many essential amino acids, including branched-chain amino acids (BCAA) that were elevated in PCOS compared with MetS. PCOS

Keywords:
α-Aminoadipic acid
Branched-chain amino acids
Insulin sensitivity
Metabolic syndrome
Vitamin D

Abbreviations: 1,25(OH)2D3, 1,25-dihydroxyvitamin D3; 1,25(OH)2D2, 1,25-dihydroxyvitamin D2; 25(OH)D3, 25-hydroxyvitamin D3; 25(OH)D2, 25-hydroxyvitamin D2; α-AA, α-aminoadipic acid; BCAA, branched-chain amino acid; BMI, body mass index; FDR, false-discovery rate; HDL, high-density lipoprotein; HOMA-IR, homeostasis model of insulin resistance; LDL, low-density lipoprotein; LLOQ, lower limit of quantification; MetS, metabolic syndrome; PCA, principal component analysis; PCOS, polycystic ovary syndrome; QC, quality control; SHBG, steroid hormone-binding globulin; S, insulin sensitivity.

⁎ Corresponding author. Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Mayo Clinic, 200 First St SW, Rochester, MN 55905. Tel.: +1 507 284 2191; fax: +1 507 284 5745.

E-mail address: chang.alice1@mayo.edu (A.Y. Chang).

http://dx.doi.org/10.1016/j.metabol.2017.03.002
0026-0495/© 2017 Elsevier Inc. All rights reserved.
was most associated with BCAA (P = .02), essential amino acids (P = .03), the essential amino acid lysine (P = .02), and the lysine metabolite α-amino adipic acid (P = .02) in models adjusted for surrogate variables representing technical variation in metabolites. No significant differences between groups were observed in concentrations of free fatty acids or vitamin D metabolites. Evaluation of the relationship of metabolites with clinical characteristics showed 1) negative associations of essential and BCAA with insulin sensitivity and sex hormone-binding globulin and 2) positive associations with homeostasis model of insulin resistance and free testosterone; metabolites were not associated with BMI or percent body fat.

Conclusions. PCOS was associated with significant metabolic alterations not attributed exclusively to androgen-related pathways, obesity, or MetS. Concentrations of essential amino acids and BCAA are increased in PCOS, which might result from or contribute to their insulin resistance.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

Polycystic ovary syndrome (PCOS), a condition of androgen excess and chronic anovulation, is the most common endocrine disorder among women of reproductive age [1,2]. Insulin resistance is highly prevalent in PCOS, even among lean women [3], and it is associated with a higher risk of type 2 diabetes mellitus and an increase in cardiovascular risk factors [4]. Limited knowledge about the exact mechanisms underlying the pathophysiology of PCOS has resulted in few available or effective therapies that ameliorate symptoms of PCOS or improve fertility or the metabolic complications of insulin resistance. Genome-wide association studies have identified candidate genes associated with PCOS, but their role in the underlying pathophysiology of PCOS has not been delineated to date. In addition, the lower odds ratios of genetic associations suggest that candidate genes are likely relevant for only a subset of women with PCOS [5–8]. Furthermore, heterogeneity in the diagnosis of PCOS, resulting from diagnostic criteria requiring 2 of 3 features, results in considerable phenotype overlap and variable response to antiandrogen therapies and metformin. This heterogeneity limits the progress of more traditional approaches to identify underlying causes and novel therapies for PCOS.

Genomic [9–11], proteomic [11,12], and metabolomic [13–20] approaches to study the pathogenesis of PCOS have implicated various pathways, including oxidative stress, immune function, and lipid metabolism. However, the activation of these pathways might reflect the high prevalence of obesity in women with PCOS, rather than the underlying pathogenesis of insulin resistance associated with PCOS. Given the great potential for interaction or overlap of metabolic pathways in PCOS vs those of obesity, plasma metabolomics offers detailed profiling of small-molecule breakdown products downstream of genomic and proteomic expression to identify active metabolic pathways associated with specific PCOS phenotypes or response to therapy.

To date, metabolomic studies in PCOS have evaluated small samples of women or evaluated the effect of multidrug therapy on PCOS [14,15,17,21]. In addition, many studies could not examine or control for obesity or differences in insulin sensitivity [12–15,18,19]. Therefore, any detected differences could have been due to obesity and obesity-related insulin resistance, rather than to PCOS. In addition, only 3 nontargeted metabolomics studies included a limited targeted quantitative approach [15,17,22] to further validate candidate pathways and correlations with PCOS characteristics.

In the current study, we sought to determine whether a metabolomics approach could identify a specific plasma metabolic fingerprint that could distinguish between obese women with PCOS and obese controls with metabolic syndrome (MetS). First, we conducted a nontargeted metabolomics analysis to determine candidate metabolic profiles and related pathways that were significantly altered between the two groups. Second, we completed a targeted metabolomics analysis of specific candidate pathways to explore metabolites that were quantitatively different between the two groups and to evaluate their association with clinical characteristics of PCOS.

2. Research Design and Methods

The study protocol was approved by the University of Texas Southwestern Medical Center Institutional Review Board, and all participants provided written informed consent to enroll in the study.

2.1. Study Participants

The plasma samples were previously collected for a study of PCOS comparing coronary function and cardiovascular performance in overweight or obese women with PCOS compared with overweight or obese women with MetS. Women with PCOS were recruited from an academic tertiary care center and county medical center. Controls with MetS were recruited from the same centers and from the Dallas Heart Study [23] to match PCOS participants by age and body mass index (BMI). The following inclusion and exclusion criteria were applied to both groups (women with PCOS and controls with MetS alone). Premenopausal women aged 30 to 50 years were required to have a BMI greater than 25 kg/m² and evidence of prediabetes (elevated fasting glucose, 100–125 mg/dL) or MetS. Participants could not have a previous diagnosis of diabetes mellitus or hypertension or be receiving medication to treat PCOS or hypertension or to
improve insulin sensitivity. Current tobacco users also were excluded. The control group reported regular menstrual cycles since adolescence and had no evidence of hyperandrogenism. Women in the control group without PCOS were studied between days 7 and 14 of the midfollicular phase.

### 2.2. Variable Definitions for the Diagnosis of PCOS and MetS

PCOS was diagnosed by using the Rotterdam criteria [24], requiring specifically the presence of both oligomenorrhea and hyperandrogenism. Oligomenorrhea was defined by fewer than 9 periods per year since menarche or since the age of 20 years when not pregnant, breastfeeding, or taking oral contraceptives. Hyperandrogenism was defined biochemically by an elevated total testosterone or dehydroepiandrosterone or symptoms of severe acne, androgenic alopecia, or Ferriman-Gallwey score ≥ 8.

MetS was defined according to the National Cholesterol Education Program Adult Treatment Panel criteria (at least 3 of 5 criteria): 1) high-density lipoprotein (HDL) cholesterol <50 mg/dL; 2) triglyceride ≥150 mg/dL; 3) impaired fasting glucose ≥100 mg/dL and <126 mg/dL; 4) hypertension or systolic blood pressure ≥130 mmHg or diastolic blood pressure > 85 mmHg; and 5) waist circumference > 88 cm [25].

### 2.3. Insulin Sensitivity

The Bergman minimal model [26] was used to calculate insulin sensitivity (SI), acute insulin response to glucose, disposition index, and glucose effectiveness from serum insulin and plasma glucose concentrations, obtained during the insulin-modified version of the frequently sampled intravenous glucose tolerance test. After obtaining fasting blood samples, dextrose (600 mg/kg; delivered as a 50% solution) was administered through an antecubital vein over a 2-min period. During the next 3 h, blood samples were obtained at 21 time points from an arterialized vein in a hand placed in a hot box warmed to approximately 60 °C.

As a surrogate measure of insulin resistance, we also calculated the index of homeostasis model of insulin resistance (HOMA-IR) from the following formula [27]:

$$\text{HOMA-IR} = \frac{\text{fasting glucose} \text{[mg/dL]}}{\mu\text{IU/mL}} / 405$$

### 2.4. Body Composition

Percent body fat was calculated from measurement of body density from underwater weight, as previously described [28].

### 2.5. Assays

Venous blood was collected in standard blood collection tubes containing ethylene diamine tetra-acetic acid for plasma and in serum separator tubes for serum. Plasma samples were immediately processed and aliquotted into 2-mL screw-top cryovials (Phoenix Research Products) and stored at −80 °C. Samples from baseline fasting time points were thawed only once for the nontargeted and targeted metabolomics analyses.

### 2.6. Nontargeted Metabolomics

Plasma nontargeted metabolomics profiling was performed using liquid chromatography/time-of-flight mass spectrometry (6220 ToF MS; Agilent) operated in both positive and negative electrospray ionization modes, using a scan range of m/z 100–1200 at a resolution of 10,000, as described previously [20,29]. Small-molecule metabolites were extracted from 100 μL of plasma by deproteinization with 80% methanol. Before deproteinization, an internal standard solution (4 μL) of 13C6-phenylalanine (250 ng/μL) was added to each sample and plasma quality control (QC) samples to monitor recovery and reproducibility in metabolite extraction [29]. Dried samples were stored at −20 °C until analysis. Samples were reconstituted in running buffer and analyzed within 48 h of reconstitution. A small fraction of supernatant from each sample was combined into a pool (pooled QC), and the remaining supernatant was split into 2 fractions to be used for polar hydrophilic interaction liquid chromatography and reverse-phase C18 ultra performance liquid chromatography separation. A separate plasma QC sample was analyzed with pooled QC to account for analytic and instrumental variability. Each QC sample was analyzed in duplicate, and chromatographic separation was achieved using hydrophilic interaction liquid chromatography and reverse phase (C18) liquid chromatography separately. The instrument settings were as follows: nebulizer gas temperature, 325 °C; capillary voltage, 3.5 kV; capillary temperature, 300 °C; fragmenter voltage, 150 V; skimmer voltage, 58 V; octapole voltage, 250 V; cycle time, 0.5 s; and run time, 15.0 min [20,29].

#### 2.6.1. Data Analysis

Metabolite peak intensities and differential regulation of metabolites between groups were determined as described previously [20,29]. Each sample was normalized to the median of the baseline and log2 transformed. Data alignment, filtering, and univariate and multivariate statistical and differential analysis were performed using Mass Profiler Professional software (Agilent). Default settings were used, with the exception of the signal-to-noise ratio threshold, mass limit (0.0025 units), and time limit (9 s). Each metabolite was putatively identified on the basis of accurate mass (m/z) against the METLIN database using a detection window of ±7 ppm. The identified metabolites are annotated as Chemical Abstracts Service, Kyoto Encyclopedia of Genes and Genomes, Human Metabolome Project database, and LIPID MAPS identifiers. Identification of selected metabolites was validated using the standards described in Supplemental Table 1.

#### 2.6.2. Statistical Analysis

Metabolites detected in >50% of the samples in any of the study groups were selected for differential expression analyses [20]. After normalization, univariate statistical analysis (unpaired t test analysis) was performed to compare the differentially expressed metabolites between PCOS and MetS (with false-discovery rates [FDRs] of ≤0.05) and was estimated by Q values [20,29]. Unsupervised principal component analysis (PCA) was performed to display variation between PCOS and MetS study groups for data visualization and whether group variation could be explained by these
metabolite variables reducing the dimensionality of the qualitative data.

2.6.3. Pathway Analysis
Identified metabolites that were differentially expressed between groups were used for pathway enrichment using MetaCore (GeneGo) [14,16]. Metabolite identifiers (Chemical Abstracts Service and Kyoto Encyclopedia of Genes and Genomes) were used for each metabolite; identifiers included name, molecular weight, fold change, and differential P value. The P value from the hypergeometric test, generated by MetaCore, represents the enrichment of certain metabolites in a pathway. A P value < .05 indicates significant enrichment. The FDR of 0.15 was applied in the assessment of pathway enrichment to allow a greater number of pathways to be reviewed because of the smaller sample size of metabolites entered into the pathway analysis [20,29]. The issue of multiple testing is accounted for in the programming of the MetaCore tool [30].

2.7. Targeted Metabolomics

2.7.1. Quantitation of Free Fatty Acid and Amino Acids
Quantitative measurements of free fatty acid and 45 amino acid metabolites were performed by tandem mass spectrometry against 12-point calibration curves that underwent the same derivatization with internal standard as described previously [20,29].

2.7.2. Quantitation of Vitamin D Metabolites
25-Hydroxyvitamin D$_2$ [25(OH)D$_2$], 25-hydroxyvitamin D$_3$ [25(OH)D$_3$], 1,25-dihydroxyvitamin D$_2$ [1,25(OH)$_2$D$_2$], and 1,25-dihydroxyvitamin D$_3$ [1,25(OH)$_2$D$_3$] were extracted from serum samples by using deuterated stable isotopes [31]. After derivatization, they were analyzed by liquid chromatography–mass spectrometry/mass spectrometry with multiple-reaction monitoring [31].

2.7.3. Statistical Analysis
To address technical variation in the metabolomics measurements, surrogate variables were created by using the R package sva [32]. Briefly, variation in metabolites was compartmentalized into variation due to interperson variability and technical variation. The surrogate variables were created as principal components of the residual error of a linear model of the metabolites, adjusted for variables that would account for interperson variation (eg, PCOS, age, race, BMI, and 11 other measured characteristics) [33]. These are then used in future analyses to account for any technical variation of the assay, assuming that this residual error is due to technical variation.

Because several metabolites had concentrations below the lower limit of quantification (LLOQ), the relationship of clinical variables and metabolites was modeled as 1) a linear regression of the van der Waerden rank-transformed metabolite on each clinical variable, individually adjusting for 2 surrogate variables for technical variation; and 2) a logistic regression of the metabolite concentration being above the LLOQ on each clinical variable, individually adjusting 2 surrogate variables for technical variation. To test the effect of a clinical variable on a metabolite, the minimum P value is observed between these 2 models, and the null distribution of the minimum P value is created empirically by randomly sorting the metabolite 1000 times. For metabolites with quantifiable concentrations, Spearman partial correlation coefficients were calculated between the metabolites and clinical variables, including S$_i$, HOMA-IR, fasting glucose, free and total testosterone, estrogen, progesterone, steroid hormone-binding globulin (SHBG), low-density lipoprotein (LDL) cholesterol, BMI, percent body fat, and age.

Again, because the concentrations of several metabolites were below the LLOQ, the effect of these metabolites on PCOS was modeled using logistic regression adjusted for 2 surrogate variables, with metabolites being modeled with 2 effects: 1) metabolite concentration above LLOQ and 2) an indicator of metabolite concentration being above LLOQ. To test the effect of the metabolite on PCOS, a likelihood ratio test with 2 degrees of freedom was used to test the 2 metabolite effects while simultaneously adjusting for the surrogate variables.

Groups of metabolites were tested for association with PCOS simultaneously by using random effects models implemented using the globaltest package (eg, essential amino acids, branched-chain amino acids [BCAAs], etc). Because of the small sample size, the effect of the metabolites could be tested in association with either the group term of PCOS vs MetS or individual clinical variables (described below). Additionally, because prior information indicated that phenylalanine and tyrosine were positively correlated, we conducted the more powerful test of summing phenylalanine and tyrosine and testing for the association in the same manner as for individual metabolites above. All metabolites tested in groups had no values below the LLOQ.

For the purposes of exploratory analysis, metabolites were further reviewed that had a P value < .05. Due to sample size, a limited number of covariates could be included in modeling PCOS (at most 2 to enable stable statistical modeling). It appeared that the largest source of variation in the measured metabolites was not explained by the measured variables in this study (these are assumed to contribute to the majority of interindividual variation in the actual metabolites, not the measured metabolites). Thus, these metabolite technical variation variables were included as the only covariates prioritized above the clinical variables. A stratified analysis of the metabolites and clinical variables by PCOS diagnosis would not be possible because the sample size would be cut in half, further exacerbating the sample size issue. Clinical variables were selected on the basis of known and observed differences between women with and without PCOS (age, BMI, insulin sensitivity, acute insulin response, HOMA-IR, fasting glucose, free testosterone, total testosterone, estradiol, progesterone, SHBG, LDL, percent body fat, and race/ethnicity).

Therefore, metabolites were tested in association with clinical variables by pooling together women with PCOS and women with MetS only. In terms of association with clinical variables, using the Bonferroni correction, statistical significance was set at 0.004 for the evaluation of the 14 clinical factors.

Group comparisons for baseline characteristics and vitamin D concentrations were made using Wilcoxon rank-sum test. Statistical analyses for these specific analyses were performed using JMP statistical software (SAS Institute).
3. Results

3.1. Baseline Characteristics

Anthropometric characteristics and baseline biomarkers for the PCOS (n = 20) and MetS (n = 18) cohorts are shown in Table 1. Both groups were matched by age and BMI, and we observed no statistical difference in percent body fat or fasting glucose concentrations. Insulin sensitivity was more impaired in PCOS compared with MetS (P = .02), along with the disposition index, which is the product of insulin sensitivity and the acute insulin response (Table 1). As expected, total testosterone was significantly higher in PCOS (P = .04). Lipid concentrations (total cholesterol, triglycerides, and HDL cholesterol) were not statistically different between groups. The only exception was LDL cholesterol, which was higher in PCOS (P = .05).

3.2. Nontargeted Metabolomics Analysis

As shown in Fig. 1, PCA showed PCOS and obese controls with MetS were distinctly separated on the basis of plasma metabolite differences. The first and second components in PCA explained 19.51% and 9.1% of variations, respectively. From the subset of 385 metabolites that significantly differed between PCOS and MetS, 85 (22%) were identified by accurate mass. Identification of 82 metabolites was validated by using standards (Supplemental Table 1). The identified metabolites characterized 19 significantly altered canonical pathways (P < .05; FDR, <0.15) between PCOS and MetS controls (Fig. 2), and metabolites were involved in networks for lipid, steroid, amino acid, carbohydrate, and vitamin metabolism (Supplemental Table 2). Among the significantly altered pathways were those for vitamin D, N-acyltransferase, prostaglandin biosynthesis and metabolism, regulation of cystic fibrosis transmembrane conductance regulator gating, BCAAs, and saturated fatty acids metabolism. Supplemental Table 3 demonstrates reproducibility of replicate sample values for components 1 and 2.

3.3. Targeted Metabolomics Analysis

Targeted metabolomics analysis was performed to confirm associations of PCOS with leading candidate pathways: 1) amino acids and their metabolites and 2) fatty acids and their metabolites. Vitamin D concentrations were also quantitatively measured with mass spectrometry. Essential amino acids (histidine, methionine, threonine, lysine, valine, isoleucine, leucine, phenylalanine, tryptophan) (Fig. 3A) and the subgroup of BCAAs (valine, isoleucine, leucine) (Fig. 3B) were most strongly associated with PCOS. Within these groups, further analysis of individual amino acids, adjusted for surrogate variables, showed that the metabolites significantly elevated in PCOS were BCAAs, lysine (median PCOS vs control, 167.7 vs 154.1 μmol/L; P = .02) (Supplemental Table 4), and the lysine metabolite α-aminoadipic acid (α-AA) (median PCOS vs control, 0.9 vs 0.7 μmol/L; P = .02). When we analyzed the sum

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PCOS (n = 20)</th>
<th>MetS (n = 18)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td>37 (6)</td>
<td>40 (6)</td>
<td>.06</td>
</tr>
<tr>
<td>Body mass index, mean (SD), kg/m²</td>
<td>34.6 (5.1)</td>
<td>33.7 (5.2)</td>
<td>.57</td>
</tr>
<tr>
<td>Body fat, mean (SD), %</td>
<td>39.0 (6.8)</td>
<td>37.4 (6.2)</td>
<td>.44</td>
</tr>
<tr>
<td>Race/ethnicity, No. (%)</td>
<td></td>
<td></td>
<td>.44</td>
</tr>
<tr>
<td>White</td>
<td>8 (40)</td>
<td>5 (28)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>7 (35)</td>
<td>10 (55)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>5 (30)</td>
<td>3 (17)</td>
<td></td>
</tr>
<tr>
<td>Glucose metabolism, mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fasting glucose, mg/dL</td>
<td>97 (6)</td>
<td>95 (12)</td>
<td>.54</td>
</tr>
<tr>
<td>Fasting insulin, μIU/mL</td>
<td>11.1 (6.9)</td>
<td>8.6 (8.5)</td>
<td>.08</td>
</tr>
<tr>
<td>Sg, min⁻¹ × (μU/mL)⁻¹ × 10⁻⁴</td>
<td>1.9 (1.0)</td>
<td>3.0 (1.2)</td>
<td>.02</td>
</tr>
<tr>
<td>AIRg, min × (μU/mL)</td>
<td>701.6 (539.5)</td>
<td>764.5 (388.6)</td>
<td>.68</td>
</tr>
<tr>
<td>Disposition index*</td>
<td>1135.1 (643.4)</td>
<td>2146.9 (1388)</td>
<td>.01</td>
</tr>
<tr>
<td>Sg, min⁻¹ × 100</td>
<td>1.7 (0.5)</td>
<td>1.9 (0.98)</td>
<td>.36</td>
</tr>
<tr>
<td>HOMA-IR</td>
<td>2.8 (1.5)</td>
<td>2.2 (2.7)</td>
<td>.45</td>
</tr>
<tr>
<td>Sex steroid hormones, mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total testosterone, ng/dL</td>
<td>45.6 (23.3)</td>
<td>32.1 (15.3)</td>
<td>.04</td>
</tr>
<tr>
<td>Free testosterone, ng/dL</td>
<td>8.9 (6.4)</td>
<td>5.6 (4.2)</td>
<td>.10</td>
</tr>
<tr>
<td>Estradiol, pg/mL</td>
<td>107.4 (70)</td>
<td>118.4 (78)</td>
<td>.64</td>
</tr>
<tr>
<td>Sex hormone-binding globulin, mcg/mL</td>
<td>28.4 (15.6)</td>
<td>35.4 (17.7)</td>
<td>.21</td>
</tr>
<tr>
<td>Lipids, mean (SD), mg/dL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cholesterol</td>
<td>191 (26)</td>
<td>175 (39)</td>
<td>.15</td>
</tr>
<tr>
<td>Low-density lipoprotein cholesterol</td>
<td>120 (22)</td>
<td>103 (30)</td>
<td>.05</td>
</tr>
<tr>
<td>High-density lipoprotein cholesterol</td>
<td>50 (8)</td>
<td>51 (12)</td>
<td>.73</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>105 (45)</td>
<td>110 (53)</td>
<td>.78</td>
</tr>
</tbody>
</table>

Abbreviations: AIRg, acute insulin response to glucose; HOMA-IR, homeostasis model of insulin resistance; MetS, metabolic syndrome; PCOS, polycystic ovary syndrome; Sg, glucose effectiveness; Sg, insulin sensitivity index.

* Disposition index = Sg × AIRg.
of the concentrations of phenylalanine and tyrosine (aromatic amino acids), they were higher in PCOS compared with MetS (median PCOS vs control, 109.5 vs 104.2 μmol/L; \( P = .0497 \)).

Individually, phenylalanine and tyrosine were not significantly associated with PCOS (\( P = .16 \) and \( .06 \), respectively).

Free fatty acids and lipid metabolites, when evaluated individually or after grouping by specific characteristics (eg, saturated fatty acids, unsaturated fatty acids, ceramides), were not significantly different between PCOS and MetS. Linoleic acid was the only lipid metabolite associated with PCOS (\( P = .04 \)). Vitamin D concentrations were not significantly different between PCOS and MetS (Table 2).

3.4. Association of Amino Acids and Lipid Metabolites with Clinical Characteristics

Because of the differences in SI and androgens between PCOS and MetS, we questioned whether the observed associations with amino acid metabolites could be associated with measures of SI, body composition, or sex hormones among the entire participant cohort. SI was negatively associated with essential amino acids, including lysine, α-AA, and the BCAAs valine and leucine (Table 3). α-AA was also positively associated with fasting glucose and HOMA-IR. Among the remaining amino acid metabolites measured, asparagine and glycine were negatively associated with SI, and positively associated with HOMA-IR, and tyrosine was positively associated with HOMA-IR (Supplemental Table 5).

BCAAs and essential amino acids were not associated with BMI or percent body fat. Isoleucine was the only amino acid positively associated with LDL cholesterol (\( P = .02 \)). No associations were identified between these groups of amino acids and hormones such as estradiol, progesterone, and total testosterone. However, free testosterone was positively associated with

![Fig. 1](image1.png) Two-dimensional score plot of principal component analysis showing group separation between PCOS and MetS. Each person is represented by 2 replicate data points. PCOS and MetS groups can be seen in clusters separated on 2 components. The first and second components explained 19.51% and 9.17% of variations, respectively. MetS indicates controls with metabolic syndrome (green squares); PCOS, polycystic ovary syndrome (red triangles).

![Fig. 2](image2.png) Metabolic pathways significantly different in PCOS vs MetS. The significance of the pathways was evaluated using \( P \) values and a false-discovery rate < 0.05. CF indicates cystic fibrosis; CFTR, cystic fibrosis transmembrane conductance regulator; HETE, hydroxyl eicosatetraenoic acid; HPETE, hydroperoxy eicosatetraenoic acid; PPAR, peroxisome proliferator-activated receptor.
associated with BCAA (Table 3). SHBG was inversely correlated with lysine, α-AA, valine, leucine, and tryptophan (Table 3). Among the lipid metabolites, linoleic acid was negatively associated with SI and LDL, and stearic acid was positively associated with BMI and percent body fat. Among the ceramides, C16-ceramide was negatively associated with S	extsubscript{1} and SHBG and positively associated with fasting glucose; C18-ceramide was positively associated with LDL, total and free testosterone, and estradiol; and C22-ceramide was positively associated with fasting glucose and total testosterone (Supplemental Table 5).

4. Discussion

In the current study, we identified 385 metabolites and 19 metabolic pathways that distinguished PCOS from age- and BMI-matched controls with MetS. We extended this nontargeted approach by performing a confirmatory, quantitative, targeted metabolomics analysis and showed that PCOS is strongly associated with groups of amino acid metabolites compared with obese women with MetS. Specifically, greater concentrations of BCAA and essential amino acids seemed to distinguish PCOS from MetS.

Comparisons to previous studies of plasma metabolomics profiles in PCOS are limited by small sample sizes, further diluted by subgroups of BMI and insulin resistance status [14,15,21], as well as by different analytic techniques, including less-sensitive nuclear magnetic resonance or nontargeted metabolomics and lack of BMI matching in control groups [14,16,18]. Most prior studies used the Rotterdam criteria for PCOS and did not restrict study participants to those with the more severe phenotype of oligomenorrhea and anovulation. Consequently, those studies might include women with a milder metabolic phenotype and less insulin resistance than this cohort [34]. Together, these factors may account for the conflicting findings of prior studies that reported varying increases and decreases in fatty acid and amino acid metabolites in PCOS vs controls [35]. Our results share the greatest agreement with the largest previously published study that confirmed significantly different metabolites from its nontargeted analysis with standards and reported increased BCAA and lysine in PCOS compared with controls [18]. However, this group also described changes in other amino acids that were not corroborated by our data; these discordant findings could be attributable to their PCOS group not being overweight and to using a control group with a markedly lower BMI than our controls. None of the prior studies performed subsequent quantitative targeted analysis of amino acid profiles, as was done in this study.

![Fig. 3](image)

**Table 2 – Vitamin D concentrations in PCOS vs MetS.**

<table>
<thead>
<tr>
<th>Vitamin</th>
<th>PCOS, median (IQR)</th>
<th>MetS, median (IQR)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>25(OH)D&lt;sub&gt;3&lt;/sub&gt;, ng/mL</td>
<td>19 (14–24)</td>
<td>14 (9–21)</td>
<td>.11</td>
</tr>
<tr>
<td>25(OH)D&lt;sub&gt;2&lt;/sub&gt;, ng/mL</td>
<td>0 (0–4.5)</td>
<td>0 (0–53)</td>
<td>.97</td>
</tr>
<tr>
<td>1,25(OH)&lt;sub&gt;2&lt;/sub&gt;D&lt;sub&gt;3&lt;/sub&gt;, pg/mL</td>
<td>46 (41–61)</td>
<td>47 (37–54)</td>
<td>.42</td>
</tr>
<tr>
<td>1,25(OH)&lt;sub&gt;2&lt;/sub&gt;D&lt;sub&gt;2&lt;/sub&gt;, pg/mL</td>
<td>0 (0–15)</td>
<td>0 (0–36)</td>
<td>.83</td>
</tr>
</tbody>
</table>

Abbreviations: 1,25(OH)<sub>2</sub>D<sub>2</sub>, 1,25-dihydroxyvitamin D<sub>2</sub>; 1,25(OH)<sub>2</sub>D<sub>3</sub>, 1,25-dihydroxyvitamin D<sub>3</sub>; 25(OH)D<sub>3</sub>, 25-hydroxyvitamin D<sub>3</sub>; 25(OH)D<sub>2</sub>, 25-hydroxyvitamin D<sub>2</sub>; MetS, metabolic syndrome; PCOS, polycystic ovary syndrome.

* The entire range reported here.
Extensive evidence from non-PCOS studies suggests that increased BCAAs and essential amino acids are associated with obesity, insulin resistance, and type 2 diabetes mellitus [36,37]. Multiple cohort studies using metabolomics have demonstrated a stronger association of obesity and insulin resistance with BCAA and related metabolites than with lipid metabolites [38-40]. Furthermore, elevated essential and BCAAs predict the development of diabetes mellitus in the Framingham and Malmo cohorts, even after adjusting for insulin resistance [7].

With our detailed characterization of insulin sensitivity and percent body fat, we have demonstrated that obese women with PCOS have greater impairment of insulin sensitivity than obese women with MetS, and this difference could not be explained by percent body fat or androgens. These results suggest that greater insulin resistance, rather than obesity or elevated androgens, is a key factor for the altered metabolomic consequences of PCOS.

To date, the proposed models for increased BCAA concentrations in obesity, insulin resistance, and type 2 diabetes mellitus suggest decreased BCAA catabolism due to decreased expression or action of BCAA catabolic enzymes in adipose and skeletal muscle tissue [36]. Increases in BCAA might also contribute to glucose intolerance by increasing the supply of BCAA metabolites glutamate and alanine for gluconeogenesis. Although we did not observe significant differences in glutamate and alanine concentrations between PCOS and MetS controls, our controls were matched by BMI and had similar percent body fat; possibly, the previously observed differences in obese cohorts might have been attributable to differences in adiposity [38]. Increases in circulating BCAA might interfere with insulin signaling in skeletal muscle and further promote insulin resistance, especially in the context of obesity and a high-fat diet [36,38,41]. Interestingly, improved glucose homeostasis has been reported for animals fed a diet specifically enriched in leucine [42]. In a cohort of healthy adults, leucine ingested with glucose acted synergistically to stimulate insulin secretion [43]. In our cohort of obese individuals, leucine concentrations were not associated with the acute insulin response during the intravenous glucose tolerance test (Supplemental Table 5).

In the present study, we observed greater circulating concentrations of lysine and α-AA in PCOS compared with MetS. The essential amino acid lysine and its metabolite α-AA were negatively associated with HOMA-IR and SHBG, and positively associated with fasting glucose. Consistent with these findings, previous studies have identified α-AA as a novel biomarker for type 2 diabetes mellitus and a potential modulator of β-cell function. In the Framingham cohort, individuals in the highest quartile for α-AA concentrations had greater than a 4-fold increase in risk for type 2 diabetes mellitus during a 12-year period, even in individuals with normal glucose tolerance [44]. A previous interventional study in PCOS identified lysine as a metabolite with marked changes after polytherapy that included insulin sensitizers and an antiandrogen [15]. In adults with glucose intolerance, plasma concentrations of lysine and α-AA were significantly reduced in response to metformin and pioglitazone treatment [45]. Although the mechanisms are not well understood, experimental data have shown that α-AA increased insulin secretion from pancreatic β cells in cell culture and lowered fasting plasma glucose in mice [44]. Hyperglycemia also increased α-AA production in endothelial cell culture [46]. Future human studies are needed to better understand the mechanisms for increased lysine and α-AA and their potential role in glucose homeostasis.

The aromatic amino acids phenylalanine and tyrosine have also been implicated in insulin resistance and metabolic disease [7,36,45], as well as in PCOS [18]. Although we did not observe a strong association with phenylalanine, plasma

<table>
<thead>
<tr>
<th>Table 3 – Association of amino acids with clinical characteristics (N = 38)</th>
<th>BMI</th>
<th>S&lt;sub&gt;1&lt;/sub&gt;</th>
<th>HOMA-IR</th>
<th>Free Testosterone</th>
<th>SHBG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino Acid</td>
<td>ρ</td>
<td>P Value</td>
<td>ρ</td>
<td>P Value</td>
<td>ρ</td>
</tr>
<tr>
<td>Essential</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histidine</td>
<td>0.07</td>
<td>.950</td>
<td>-0.02</td>
<td>.737</td>
<td>0.11</td>
</tr>
<tr>
<td>Methionine</td>
<td>-0.30</td>
<td>.736</td>
<td>-0.31</td>
<td>.108</td>
<td>0.37</td>
</tr>
<tr>
<td>Threonine</td>
<td>0.05</td>
<td>.951</td>
<td>-0.15</td>
<td>.874</td>
<td>0.02</td>
</tr>
<tr>
<td>Lysine</td>
<td>0.20</td>
<td>.212</td>
<td>-0.46</td>
<td>.008</td>
<td>0.43</td>
</tr>
<tr>
<td>α-Aminoadipic acid&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.17</td>
<td>.290</td>
<td>-0.37</td>
<td>.015</td>
<td>0.42</td>
</tr>
<tr>
<td>Branched chain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valine</td>
<td>0.13</td>
<td>.412</td>
<td>-0.25</td>
<td>.029</td>
<td>0.21</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>0.18</td>
<td>.301</td>
<td>-0.44</td>
<td>.010</td>
<td>0.43</td>
</tr>
<tr>
<td>Leucine</td>
<td>0.18</td>
<td>.389</td>
<td>-0.30</td>
<td>.043</td>
<td>0.29</td>
</tr>
<tr>
<td>Aromatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>0.20</td>
<td>.472</td>
<td>-0.07</td>
<td>.422</td>
<td>0.11</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>0.15</td>
<td>.270</td>
<td>0.05</td>
<td>.970</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index; HOMA-IR, homeostasis model of insulin resistance; SHBG, sex hormone-binding globulin; S<sub>1</sub>, insulin sensitivity.

<sup>a</sup> The Spearman partial correlation (ρ) adjusted for surrogate variables representing technical variation in metabolites.

<sup>b</sup> Boldface text shows associations with P < .05.

<sup>c</sup> The insulin sensitivity index (10<sup>-4</sup> × min<sup>-1</sup> × μU<sup>-1</sup> × mL<sup>-1</sup>) was derived from frequently sampled intravenous glucose tolerance tests.

<sup>d</sup> A metabolite of lysine.
concentrations of tyrosine were associated with PCOS, and the sum of phenylalanine and tyrosine concentrations was greater in PCOS compared with MetS.

Our nontargeted metabolomics analysis identified the saturated fatty acid biosynthesis pathway as being significantly altered in the PCOS group. However, when we examined free fatty acid metabolites in the plasma, only linoleic acid was associated with PCOS. Escobar-Morreale et al. [15] demonstrated that obese women with PCOS had a different profile of fatty acid metabolites suggestive of increased lipolysis compared with nonobese women with PCOS who demonstrated suppression of lipolysis. When comparing obese women with PCOS to nonobese women with PCOS or to BMI-matched controls without PCOS, long-chain fatty acids such as linoleic acid were elevated only in the obese women with PCOS [15]; this finding was corroborated by other groups that did not match control groups by BMI [19,21]. The lack of significant differences in the lipid metabolite profile in the current study may be attributable to matching our control group for both obesity and MetS. In the present study, linoleic acid was also positively associated with S1 and negatively associated with LDL cholesterol; these clinical variables were both different in the women with PCOS compared with controls.

4.1. Vitamin D

In the nontargeted metabolomics analysis, the vitamin D2 (ergocalciferol) metabolism network was the most significantly altered pathway in PCOS compared with MetS (Fig. 2). The 25(OH)D2 concentrations measured in PCOS and MetS were indicative of vitamin D deficiency (<20 ng/mL), as defined by the Institute of Medicine [47] and seen in other studies of PCOS [21,48,49]. Vitamin D deficiency has also been associated with metabolic disturbances such as obesity, insulin resistance, and dyslipidemia [50,51]. Although 25(OH)D2 has been associated with insulin resistance in PCOS, as measured by HOMA-IR, these associations were not significant after adjusting for BMI [49,52]. In agreement with other studies [53,54], we did not find any differences in plasma concentrations of vitamin D [25(OH)D3 or 1,25(OH)2D3] between PCOS and BMI-matched controls with MetS. This finding suggests that low vitamin D might be associated with obesity, rather than with PCOS. Although we did not quantify all the intermediate vitamin D metabolites, it is unlikely that we had the power to detect these smaller changes if 25(OH)D2 and 1,25(OH)2D3 concentrations already were similar. Another explanation for the discrepancy between the pathways analysis and the quantitative measurements relates to the analytic technique. Pathway analysis represents the enrichment of certain metabolites in a metabolic pathway. The P value from the hypergeometric test in pathway analysis is generated from the number of metabolites present in a given pathway, instead of from significantly changing metabolites in the pathway. In this manner, the vitamin D pathway might be selected due to enrichment of vitamin D metabolites and its derivatives that are putatively identified by accurate mass but cannot be measured quantitatively because of the lack of availability of standards. The quantitative analysis targeted clinically relevant and routinely tested metabolites of vitamin D, which did not confirm the nontargeted analysis. This illustrates the importance of follow up targeted analysis.

4.2. Strengths and Limitations

Because of the limited number of participants, our results should be considered exploratory until they can be confirmed in a larger cohort. Recognizing the limitations of the sample size, strengths of this study include model development for the targeted analysis, which accounted for technical and interindividual variation. The approach of validating the nontargeted analysis results with quantitative targeted metabolomics is another strength of this study. However, because of the limited sample size, we could not further evaluate the associations of clinical variables with specific metabolites. Here, we reduced the confounding effects of obesity, dyslipidemia, and insulin resistance by choosing obese women with MetS as the control group. However, the obesity and MetS in our cohort might explain the lack of differences between PCOS and controls in free fatty acids, ceramides, and sphingolipids that have been previously reported with insulin resistance [38] and obesity [15,18]. A few ceramide metabolites were associated with S1 but among the entire sample of women, irrespective of PCOS status. Sample size may have limited our ability to discern associations of BCAA metabolites glutamate, glutamine, and alanine that were described previously [36,38]. Participants were studied after a 12-h overnight fast as outpatients. They did not have a standardized diet before samples were obtained for metabolomics analysis, which could increase variability in the measured metabolites and decrease our ability to observe differences. However, these limitations should not have affected the observed associations with PCOS or correlations with clinical variables. Although women with PCOS did not have an induced withdrawal period for timing of blood samples, they were selected because they had a higher likelihood of anovulatory cycles. Because the women without PCOS were studied in the midfollicular phase, differences in estrogen and progesterone concentrations should be minimized. Estradiol and progesterone concentrations were also evaluated and were not associated with the metabolites (Supplemental Table 5).

Although it would be desirable to examine the PGE2 and N-acyltransferase pathways, because of the lack of availability of standards, we could not measure these metabolites. Experiments in the future might explore the functional aspects of these important pathways in PCOS and offer new insights regarding the pathophysiology of PCOS.

In this cross-sectional study, we were unable to determine causality (ie, were metabolic changes caused by PCOS or did they cause the development of PCOS features?). As discussed above, the selection of an obese comparison group with MetS and detailed phenotyping suggests that impaired insulin sensitivity is a key factor in the metabolic consequences or contributors to PCOS.

4.3. Conclusions

In summary, PCOS was associated with significant metabolic alterations when compared with age- and BMI-matched controls with MetS. Although we used a relatively small group of subjects, we observed that greater concentrations of BCAA, essential amino acids, and the lysine metabolite α-AA seemed to distinguish PCOS from MetS. How these specific
amino acid elevations in PCOS might contribute or result from PCOS remains to be determined. The translational implication is that further study of amino acid metabolism in PCOS might determine the underlying pathophysiology of insulin resistance in PCOS and target the development of new therapies. Because BCAA and lysine metabolism is increased in other insulin-resistant individuals without PCOS, future study might also be relevant for the prevention and treatment of type 2 diabetes mellitus. Finally, the current study indicates that a nontargeted metabolomics approach is feasible for identifying candidate pathways characterizing PCOS, and it highlights the importance of a targeted approach for follow-up in the discovery of biomarkers.

**Author Contributions**

A.Y.C. designed the study, collected data, analyzed data, interpreted data, and wrote the manuscript. She is the guarantor of this work.

A.L. wrote the manuscript and contributed to data analysis and interpretation.

G.D.J. analyzed data and contributed to interpretation and manuscript writing.

T.D. analyzed data and contributed to interpretation and manuscript writing.

R.E.C. analyzed data and contributed to interpretation and manuscript writing.

R.J.S. contributed to data analysis and interpretation.

K.S.N. contributed to the design of the study, interpretation of data, and manuscript writing.

**Funding**

Research reported in this publication was supported by the Mayo Clinic Metabolomics Core through grant number U24DK100469 from the National Institute of Diabetes and Digestive and Kidney Diseases, and grant number UL1 TR000135 from the National Center for Advancing Translational Sciences (NCATS), the American Heart Association (Fellow-to-Faculty Transition Award to A.Y.C.), and the Office of Women’s Health Research (Building Interdisciplinary Careers in Women’s Health Award K12HD065987 to A.Y.C.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

The sponsors of the study had no role in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

**Disclosure Statement**

The authors declare no potential conflicts of interests.

**Acknowledgments**

None.

**Conflict of Interest**

None.

**Appendix A. Supplementary data**

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.metabol.2017.03.002.

**REFERENCES**


