

Comprehensive Analysis of Serum Metabolites and Whole Blood Cell Transcriptome reveals the Dysregulated Metabolic Pathways in Metabolically Healthy Obesity.

Fayaz Mir

FMir1@hamad.qa

Hamad Medical Corporation <https://orcid.org/0000-0003-1104-0852>

Ehsan Ullah

Hamad Bin Khalifa University

Ahmad Iskandarani

Raghvendra Mall

Technology Innovation Institute <https://orcid.org/0000-0003-1779-3150>

Farhan Cyprian

Qatar University

Mutasem Shraim

Meis Alkasem

Hamza Al Halabi

<https://orcid.org/0009-0003-1422-4507>

Shahrad Taheri

Weill Cornell Medical College in Qatar

Abdul-Badi Abou-Samra

Hamad Medical Corporation

Article

Keywords:

Posted Date: October 14th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-5216460/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: There is **NO** conflict of interest to disclose

Abstract

Background

Obesity stands as a formidable public health challenge, contributing to a spectrum of diseases, including cardiovascular disorders, and type 2 diabetes mellitus. Individuals with obesity classified as “metabolically healthy” have susceptibility to various diseases later in life. These diseases often linked to dysregulated metabolic pathways. Our objective was to explore potential associations between serum metabolites and features of metabolic diseases in metabolically healthy subjects with obesity.

Methods

We analyzed a dataset of 40 subjects with obesity only (OBO, n = 20) versus age-matched lean healthy controls (LHC, n = 20). We measured 711 serum metabolites and whole blood transcriptomes. Pathway enrichment analysis was employed to uncover meaningful insights into the association between metabolite concentrations and the observed phenotypic changes. Finally, Transcriptome profiling and subsequent gene set enrichment was done to identify the differentially enriched pathways between the LHC to OBO subjects.

Results

A total of 116 metabolites, mostly lipids, were significantly different ($p < 0.05$) between the 2 groups. Notably the metabolites demonstrated a distinct metabolic signature differentiating OBO from LHC group. The differentially expressed metabolites include lipids, amino acids, nucleotides, peptides, partially characterized molecules, cofactors/vitamins, carbohydrates, xenobiotics, and energy-related metabolites. Pathway enrichment scores indicated that out of 26 metabolic pathways, 14 pathways were differentially activated between the 2 groups. Among these, 5 major metabolic pathways significantly enriched and had maximum difference in mean activity between the two groups were aminoacyl-tRNA biosynthesis, phosphonate and phosphinate metabolism, pyrimidine metabolism, glutathione metabolism and lysine degradation.

Conclusions

Our results indicate that obesity is characterized by a distinctive metabolomic signature emphasizing the perturbed pathways involving amino acids and lipid metabolism.

Introduction

Obesity is one of the greatest public health challenges and is complicated by several chronic diseases including cardiovascular disease, type 2 diabetes mellitus (T2D) and several common cancers ^{1–4}. Obesity is associated with an increased the risk of infection ^{5,6}. Overall, obesity considerably reduces the quality of life and life expectancy ⁷.

Excess adiposity is associated with dysregulation multiple metabolic pathways. Various epidemiologic studies have shown that individuals with obesity are at increased risk for type 2 diabetes mellitus (T2DM) and that obesity is linked to insulin resistance, beta-cell dysfunction, and imbalanced fat tissue metabolism ⁸. However, a subgroup of individuals with obesity but no metabolic disorders has been described as “metabolically healthy” or having “obesity only” (OBO) ⁹. The OBO phenotype has a favorable lipid profile and normal or only slightly altered insulin sensitivity, despite a similar

amount of body fat ⁹. People with OBO have a stable metabolic status and possess a lower risk for metabolic diseases and mortality than people with obesity and metabolic syndrome (OBM), however compared to lean healthy control (LHC) subjects they are at increased risk for diabetes ¹⁰. The OBO phenotype may represent a temporary state with subsequent progression to metabolic derangements ¹¹. Subclinical metabolic derangements are likely to be present in OBO group, which may become through ageing, a major risk factor for a variety of chronic diseases including cardiovascular and metabolic disorders.

Understanding obesity phenotypes through study of alterations in metabolic pathways may enhance the screening, prevention, and treatment of obesity associated metabolic dysfunction. The objective of this study, using untargeted metabolomics, was to identify the underlying metabolic pathways and metabolic mechanisms that are differentially regulated in OBO versus LHC.

Materials and Methods

Study Population

The study was approved by the institutional review board (IRB) of Hamad Medical Corporation under study protocol (IRB protocol #16245/16). Participants with normal body mass index (lean healthy control, LHC) or with obesity only (OBO) were recruited at the Qatar Metabolic Institute and provided written informed consent. For this study, we included men and women with BMI ≤ 25 kg/m² for the LHC group and BMI ≥ 35 kg/m² for the OBO group, and who are generally healthy and free of any chronic illness. Fasting blood samples were collected, centrifuged at 1200 × g at 4°C for 10 min to isolate serum which was aliquoted and stored at -80°C until measurements.

Statistical Analysis

Means, standard deviations, and p-values were computed for study participants in the LHC and OBO groups. Missing values (less than 0.5% of all parameters) were substituted with the median for the corresponding variables. Subsequent to this imputation, various statistical tests were employed based on the type and distribution of the values. For categorical variables, the chi-square test was utilized, while the Wilcoxon test was performed for continuous variables that did not adhere to a normal distribution. To assess the normality of variables, the Shapiro–Wilk test was employed. For those variables demonstrating normal distribution, Student's t-test was applied, providing a comprehensive and robust data analysis.

Metabolomics Profiling and Quality Control

Serum samples from the LHC and OBO groups were subjected to untargeted, ultrahigh-performance liquid chromatography–tandem mass spectroscopy (UPLC–MS/MS) and were curated by Metabolon on serum metabolites. After normalizing samples across batches to correct for minor instrument variations by scaling each batch's medians to one and scaling the rest of the data points proportionally, batch-normalized data were generated. Within ten super classes, 1069 metabolites were profiled. These super classes included lipids (n = 360, 33.68%), amino acids (n = 203, 18.99%), xenobiotics (n = 169, 15.81%), peptides (n = 47, 4.40%), nucleotides (n = 38, 3.55%), cofactors and vitamins (n = 31, 2.90%), carbohydrates (n = 20, 1.87%), partially characterized molecules (n = 17, 1.59%), energy related metabolites (n = 10, 0.94%), and unknown (n = 174, 16.28%). After removing the unknown metabolites, 895 metabolites were used for further analysis. To ensure data quality, data was pre-processed using the criteria suggested by Wei et al. ¹²; metabolites and samples with more than 20% missing data were removed. After removing 184 metabolites, 711 metabolites were available for further analysis. Data imputation was performed by replacing values of metabolite below detection limit with the minimum detectable value for that metabolite.

Identification of outliers was conducted through the application of principal component analysis (PCA). Specifically, samples with first five principal component values falling outside the range of ($\mu \pm 2.5$ SD) were deemed outliers. Subsequently, one sample from the LHC group and three samples from the OBO group were identified as outliers and removed from the dataset (refer to Supplementary Figure S1).

To mitigate the potential impact of outliers on the data, a Winsorization procedure was employed. This involved setting metabolite values above the 90th percentile to the 90th percentile and values below the 10th percentile were replaced with the 10th percentile for each metabolite. This approach helps to normalize extreme values and ensure a more robust dataset for downstream analyses.

Following the aforementioned quality control steps, a total of 711 metabolites and 36 samples were retained for subsequent analyses. Among these, 19 samples belonged to the LHC group, while 17 samples belonged to the OBO. These refined datasets were utilized for further exploration and interpretation of metabolic patterns in the subsequent stages of our analysis.

Univariate Statistical Analysis

Logistic regression was used for assessing discrepancies among individual metabolites, with adjustments made for age, sex, and BMI to minimize their influence on outcomes. Owing to the limited sample size, the analysis did not incorporate false discovery rate (FDR) control. The 'glm' function within the R stats package (R Core Team, version 4.2.1, 2022, Vienna, Austria) facilitated the computation of p-values and effect sizes in logistic regression. The effect size was instrumental in determining the direction of change in metabolite concentration between LHC and OBO: a positive effect size indicated a higher concentration in LHC compared to OBO, and conversely.

Pathway Enrichment Analysis

The pathway enrichment analysis was conducted using MetaboAnalystR v3.0¹³. Metabolite sets were obtained by providing HMDB (Human Metabolome Database) IDs, which were supplied by Metabolon for matching the compounds in the metabolite sets. The auto-normalization option was employed to preprocess the data, transforming it to zero mean and unit variance.

Quantitative enrichment analysis (QEA)¹⁴ was then applied, utilizing metabolite concentrations to construct a generalized linear model for the estimation of the Q-stat (a statistical measure) for each metabolite set. The Q-stat serves as an indicator of the correlation between compound concentrations and the specified phenotype. This approach enables the identification of metabolite sets that exhibit a significant change in concentration in only a few compounds or those with numerous compounds displaying small, correlated changes.

The enrichment ratio (ER) was calculated as the ratio of the Q-stat for the given data to its expected value by chance. An ER greater than one for a particular metabolite set indicates differential metabolite concentrations, signifying potential biological relevance. This comprehensive approach was employed to uncover meaningful insights into the association between metabolite concentrations and the observed phenotypic changes.

The visualization of KEGG pathway maps was accomplished through the annotation process using the online tool Pathview (<https://pathview.uncc.edu/>)^{15,16}. Pathview facilitated the annotation of KEGG pathway maps by filling the metabolite nodes with colors that depict the ratio of normalized metabolite concentrations between LHC and OBO samples. Specifically, the color scheme employed in the visualization was designed to convey information about the relative metabolite concentrations in OBO compared to LHC. The color assignment was as follows: red represented

higher concentrations of metabolites, yellow indicated similar concentrations, and blue signified lower concentrations in OBO as compared to LHC.

This approach not only provided a visually intuitive representation of the metabolite data but also allowed for a quick and comprehensive assessment of the relative abundance of metabolites across the LHC and OBO groups, contributing to a more insightful interpretation of the pathway maps.

RNA isolation and quality control

Whole blood (2.5 mL) was collected into PaXgene Blood RNA Tubes (PreAnalytix). The tubes were inverted 8–10 times then placed at room temperature for 2–3 h and then frozen at – 80°C for storage. The samples were thawed overnight, then total RNA was isolated with a PAXgene Blood RNA Kit including the DNase Set (Qiagen). The concentrations and purity of the RNA samples were evaluated spectrophotometrically (Nanodrop ND-1000, Thermo, Wilmington, DE USA). The RNA isolation process was validated by analyzing the integrity of several RNAs with the RNA 6000 Nano Chip Kit (Agilent). The presence of the small RNA fraction was confirmed by the Agilent Small RNA Kit (Agilent).

Transcriptomic Profiling

A Clariom D Assay-Human system (Applied Biosystems) containing 47,231 expressions and 770 control probes was used for transcriptomic analysis from the whole blood samples. In brief, 100 ng of total RNA was reverse-transcribed using GeneChip™ WT PLUS Reagent Kit (Applied biosystems); 5.5 µg of ss-cDNA was fragmented and labeled then hybridized for 16 h to the GeneChip™ cartridge array in a GeneChip Hybridization Oven 640. Arrays were washed and stained in a GeneChip Fluidics Station 450 and scanned in a GeneChip Scanner 3000 7G. The signal values were evaluated using the Affymetrix® GeneChip™ Command Console software.

The CEL files containing the transcriptional signatures were further analyzed using the affy v1.74.0 ¹⁷, oligo v1.60.0 ¹⁸, pd.clariom.d.human v3.14.1 ¹⁹, affycoretools 1.68.1 ²⁰ packages in R. The raw transcriptional data was read through the 'read.celfiles' function in the oligo package. The quality control steps were performed using the 'rma' function in the oligo package; these include background correction, quantile normalization, and a log transformation of the transcript counts. The background transcript was filtered out using the 'getMainProbes' function in the affy coretools package. The 'annotateEset' function in the affy package was used by passing the transcriptional matrix along with the pd.clariom.d.human for mapping the probe IDs to their corresponding ENTREZ IDs and Gene Symbols. By selecting only those probe IDs which can map to a corresponding ENTREZ ID, we obtained the final transcriptional matrix comprising 32,304 genes (including PCGs as well as non-coding genes) with their corresponding expression values for 21 profiles (including 11 LHC and 10 OBO subjects). Next, the differences between LHC (control) and OBO (case) subjects was deduced using a Bayesian statistical framework, namely LIMMA (limma package v3.52.4 in R) ²¹, which identifies differentially expressed PCGs (DE-PCGs) at a p-value threshold of 0.05 (with FDR correction).

Single sample gene set enrichment.

To obtain an estimate of enrichment of the metabolic pathways, identified through our metabolomic profiling framework, for each sample, we used the Gene Set Variation Analysis (GSVA) function in ^{22–24} GSVA is a non-parametric, unsupervised technique, used to estimate gene set enrichment scores as a function of genes inside and outside the gene set analogously to a competitive gene set test. In our scenario, the gene set comprises the list of DE-PCGs which are participating as kinases/enzymes in the metabolic pathways. Mathematically, this can be represented as $G(p_i) = (g_1, \dots, g_k)$, where $G(p_i)$ represents the gene set corresponding to i^{th} metabolic pathway and consists of k DE-

PCGs with respect to the phenotype of interest. Each of these $G(p_i)$ is passed to GSVA along with the transcriptomic expression matrix keeping all other parameters to default settings to obtain the single sample enrichment scores for each $G(p_i)$. Once the enrichment score for $G(p_i)$ for each subject, we perform the Wilcoxon-ranksum test to identify the differentially enriched pathways w.r.t. whole blood transcriptomics when comparing the LHC to OBO patients.

Results

Baseline Characteristics of the Study Population

The study included a total of 40 participants, who were categorized into two groups: the LHC group and the OBO group. To ensure comparability, participants in both groups were carefully matched based on age, resulting with each group consisting of 20 subjects.

The baseline characteristics of the participants are summarized in Table 1. The mean age of LHC and OBO groups was 35.40 ± 6.61 and 35.95 ± 7.20 years, respectively, and the BMI was 23.31 ± 2.69 and $40.43 \pm 4.77 \text{ kg/m}^2$, respectively. Furthermore, the OBO group exhibited a higher number of females compared to the LHC group, with a statistically significant difference observed ($p \leq 0.012$).

Table 1

Cohort characteristics with reported mean (standard deviation) for all variables, along with corresponding p -values indicating differences between the means. Significantly different values ($p < 0.05$) are underlined. M: males; F: females; All: males and females in the group; p_g : p -value corresponding to differences between the means for males and females within the group; p_m : p -value corresponding to differences between the means for males across the groups; p_f : p -value corresponding to differences between the means for females across the groups; p_a : p -value corresponding to differences between the means for all, across the groups.

Parameters	LHC				OBO			LHC vs OBO			
	M	F	All	p_g	M	F	All	p_g	p_m	p_f	p_a
	<i>n</i> = 14	<i>n</i> = 6	<i>n</i> = 20		<i>n</i> = 5	<i>n</i> = 15	<i>n</i> = 20				
ALT (uL)	19.29 (7.16)	24.35 (12.15)	20.81 (8.92)	0.48	29.18 (13.60)	20.97 (18.03)	23.02 (17.08)	0.37	0.09	0.68	0.61
AST (uL)	20.29 (5.12)	22.50 (10.01)	20.95 (6.74)	0.93	29.18 (13.48)	18.00 (9.83)	20.80 (11.58)	0.06	0.08	0.36	0.96
Age (years)	34.86 (6.88)	36.67 (6.35)	35.40 (6.61)	0.59	39.60 (3.13)	34.73 (7.82)	35.95 (7.20)	0.20	0.16	0.33	0.76
Albumin (gm/L)	42.07 (9.38)	40.50 (4.59)	41.60 (8.14)	0.70	41.32 (4.09)	36.50 (3.37)	37.71 (4.06)	0.07	0.87	0.10	0.07
BMI	23.74 (2.38)	22.30 (3.34)	23.31 (2.69)	0.29	39.82 (4.11)	40.63 (5.08)	40.43 (4.77)	0.97	< 0.01	< 0.01	< 0.01
C-Peptide (ng/ml)	1.82 (0.71)	2.06 (0.53)	1.89 (0.66)	0.47	2.91 (0.70)	3.26 (1.40)	3.17 (1.25)	0.60	0.01	0.01	< 0.01
CRP (mg/L)	0.79 (0.63)	0.60 (0.37)	0.74 (0.56)	0.50	4.58 (4.50)	18.90 (21.46)	15.32 (19.59)	0.03	0.13	0.01	< 0.01
Total Cholesterol (mmol/L)	4.92 (0.82)	4.80 (0.97)	4.89 (0.84)	0.84	5.76 (1.45)	4.10 (0.54)	4.52 (1.10)	< 0.01	0.16	0.12	0.24
Creatinine (Umol/L)	86.07 (10.47)	61.17 (10.28)	78.60 (15.49)	< 0.01	80.00 (9.06)	55.87 (8.55)	61.90 (13.64)	< 0.01	0.18	0.31	< 0.01
Fasting Glucose (mmol/L)	4.70 (0.38)	4.62 (0.78)	4.68 (0.51)	0.59	4.94 (0.50)	4.99 (0.32)	4.98 (0.36)	0.72	0.21	0.12	0.02
HDL-Cholesterol (mmol/L)	1.29 (0.18)	1.52 (0.15)	1.36 (0.20)	0.02	1.20 (0.19)	1.40 (0.29)	1.35 (0.28)	0.14	0.32	0.53	0.97
HbA1C	5.30 (0.27)	5.13 (0.31)	5.25 (0.29)	0.21	5.48 (0.23)	5.53 (0.33)	5.52 (0.30)	0.78	0.19	0.02	0.01
Insulin	8.64 (4.56)	8.85 (4.05)	8.70 (4.30)	0.92	18.48 (6.31)	20.87 (15.39)	20.28 (13.57)	0.74	< 0.01	0.01	< 0.01
LDL-Cholesterol (mmol/L)	3.14 (0.76)	2.92 (0.96)	3.08 (0.81)	0.74	3.94 (1.67)	2.15 (0.46)	2.60 (1.17)	0.01	0.38	0.07	0.14
Triglycerides (mmol/L)	1.07 (0.31)	0.80 (0.20)	0.99 (0.30)	0.07	1.34 (0.39)	1.11 (0.40)	1.17 (0.40)	0.23	0.37	0.10	0.23

There were statistically significant differences between the OBO and LHC groups in various clinical parameters. The mean values for BMI, C-Peptide, CRP, glucose, HbA1C, and insulin were notably higher in the OBO group compared to the LHC group (all $p < 0.05$). Conversely, the mean value of creatinine was lower in the OBO group compared to the LHC group. Similar trends in mean values for gender were observed between the two groups.

Univariate Analysis

Univariate analysis was conducted on the metabolomics profile using logistic regression. Following quality control measures, a total of 711 metabolites were selected after quality control for analysis. Among these, 116 metabolites exhibited statistically significantly different levels ($p < 0.05$) in the OBO group compared to the LHC group. The corresponding fold change in metabolite concentrations was visually represented against the p -values in Fig. 1A.

Figure 1B summarizes the superpathways of the metabolites that displayed significant differences between the LHC and OBO groups. Notably, 91 (78%) of the significantly different metabolites in OBO, relative to LHC, demonstrated an increase. These included lipids (N = 36), amino acids (N = 27), nucleotides (N = 11), peptides (N = 7), partially characterized molecules (N = 4), cofactors/vitamins (N = 3), carbohydrates (N = 2), and xenobiotics (N = 1).

Conversely, the remaining 25 (22%) metabolites exhibited a decrease in OBO compared to LHC. This subgroup included lipids (N = 14), xenobiotics (N = 5), amino acids (N = 2), cofactors/vitamins (N = 2), nucleotides (N = 1), and energy-related metabolites (N = 1).

Below, specific metabolites of interest will be discussed in detail, shedding light on their potential implications and contributions to the observed metabolic differences between the OBO and LHC groups. This detailed analysis aims to provide a deeper understanding of the metabolic alterations associated with the experimental conditions under investigation.

Sphingomyelins

The most prominent disparities among metabolites in the OBO compared to LHC were observed in the lipid category, with a significance threshold set at p -value < 0.05 . Notably, several sphingomyelins demonstrated a significant increase in the OBO group relative to the LHC group (Fig. 2).

The significantly different sphingomyelins include sphingomyelin (d18:1/18:1, d18:2/18:0) (p -value = 0.005), sphingomyelin (d18:0/18:0, d19:0/17:0)* (p -value = 0.005), sphingomyelin (d18:1/20:2, d18:2/20:1, d16:1/22:2)* (p -value = 0.008), sphingomyelin (d18:1/22:2, d18:2/22:1, d16:1/24:2)* (p -value = 0.008), stearyl sphingomyelin (d18:1/18:0) (p -value = 0.009), sphingomyelin (d18:2/21:0, d16:2/23:0)* (p -value = 0.01), sphingomyelin (d18:2/24:2)* (p -value = 0.01), sphingomyelin (d18:2/16:0, d18:1/16:1)* (p -value = 0.012), sphingomyelin (d18:2/18:1)* (p -value = 0.012), sphingomyelin (d18:0/20:0, d16:0/22:0)* (p -value = 0.017), sphingomyelin (d17:2/16:0, d18:2/15:0)* (p -value = 0.017), sphingomyelin (d18:1/20:1, d18:2/20:0)* (p -value = 0.017), sphingomyelin (d18:1/19:0, d19:1/18:0)* (p -value = 0.027), sphingomyelin (d18:2/23:0, d18:1/23:1, d17:1/24:1)* (p -value = 0.032), sphingomyelin (d18:2/24:1, d18:1/24:2)* (p -value = 0.039).

Amino Acids

A comprehensive analysis of amino acid-related molecules revealed significant differences between the LHC and OBO groups, as depicted in Fig. 3. Among the 27 identified amino acid-related molecules, taurine ($p < 0.013$) and guanidinoacetate ($p < 0.008$) were exceptions, exhibiting lower concentrations in the OBO group compared to the LHC group. Intriguingly, all other amino acid-related substances demonstrated higher concentrations in the OBO group.

These findings underscore a distinctive metabolic profile associated with amino acid-related molecules in the OBO group relative to the LHC group. The observed variations in the concentrations of these substances suggest a potential role in delineating the metabolic distinctions between the two cohorts. The specific identification and detailed examination of these amino acid-related molecules may contribute valuable insights into the metabolic pathways and processes that underlie the observed differences, further enhancing our understanding of the physiological implications associated with the experimental conditions.

Pathway Enrichment Analysis

The outcomes of the metabolic pathway enrichment analysis are summarized in Table 2. Specifically, the presented pathways have achieved a significance threshold of p -value < 0.05 . Notably, the enriched pathways predominantly encompass metabolic processes related to amino acids and sugars. This observation suggests a pronounced involvement of these pathways in the metabolic alterations observed between the experimental groups, emphasizing their potential significance in elucidating the underlying mechanisms and biological implications associated with the studied conditions. The detailed information provided in Table 2 serves to offer a comprehensive overview of the enriched pathways, facilitating a focused examination of the specific metabolic processes contributing to the observed differences between the LHC and OBO groups. A comprehensive and detailed graphical representation of the pathways is provided in the supplementary material.

Table 2

Pathway enrichment analysis of metabolites, reporting only pathways with a p -value < 0.05. "Total" represents the number of metabolites identified by KEGG in each pathway, while "Hits" denotes the number of metabolites associated in the present study.

Pathway	Total	Hits	Statistic Q	Expected Q	FDR p-value
Arginine and proline metabolism	38	12	13.828	2.8571	0.00034
Arginine biosynthesis	14	12	12.934	2.8571	0.00034
D-Glutamine and D-glutamate metabolism	6	3	23.949	2.8571	0.00042
Cysteine and methionine metabolism	33	9	11.493	2.8571	0.00052
Nitrogen metabolism	6	2	30.633	2.8571	0.00052
Butanoate metabolism	15	4	17.629	2.8571	0.00066
Nicotinate and nicotinamide metabolism	15	6	19.419	2.8571	0.00084
Glycine, serine and threonine metabolism	33	12	10.259	2.8571	0.00156
Glyoxylate and dicarboxylate metabolism	32	7	12.459	2.8571	0.00285
Thiamine metabolism	7	1	30.428	2.8571	0.00298
Porphyrin and chlorophyll metabolism	30	4	13.419	2.8571	0.00360
Glycerolipid metabolism	16	3	14.555	2.8571	0.00419
Glutathione metabolism	28	6	11.569	2.8571	0.00437
Taurine and hypotaurine metabolism	8	4	13.946	2.8571	0.00580
Histidine metabolism	16	6	10.91	2.8571	0.00585
Alanine, aspartate and glutamate metabolism	28	13	9.3267	2.8571	0.00585
Glycerophospholipid metabolism	36	4	14.007	2.8571	0.00972
Phosphonate and phosphinate metabolism	6	1	22.814	2.8571	0.01091
Ether lipid metabolism	20	1	19.651	2.8571	0.02174
Steroid hormone biosynthesis	85	7	9.0929	2.8571	0.02266
Lysine degradation	25	5	9.7453	2.8571	0.02531
Steroid biosynthesis	42	2	13.844	2.8571	0.03058
Sphingolipid metabolism	21	5	10.437	2.8571	0.03135
Pyrimidine metabolism	39	9	6.9363	2.8571	0.04337
Aminoacyl-tRNA biosynthesis	48	19	7.856	2.8571	0.04708
Pantothenate and CoA biosynthesis	19	9	6.53	2.8571	0.04708

We next performed the GSVA analysis to obtain pathway enrichment scores also referred to as activity for each sample. After obtaining the pathway activities, we performed the Wilcoxon-ranksum test to identify differentially enriched metabolic pathways between OBO vs LHC subjects as depicted in Fig. 4. As shown in Fig. 4, we observed that majority of the metabolic pathways have negative activity in OBO when compared to LHC (as indicated by mean activity). This suggests these metabolic pathways lose their activity profile in OBO patients when compared to LHC

subjects as per their blood transcriptomics. The top 5 metabolic pathways which were significantly enriched and had maximum difference in mean activity between the two groups were: a) Aminoacyl-tRNA biosynthesis (mean Δ Activity = -0.761, p-value = 1.1e-05); b) Phosphonate and phosphinate metabolism (mean Δ Activity = -0.628, p-value = 6.2e-03); c) Pyrimidine metabolism (mean Δ Activity = -0.480, p-value = 1.08e-04); d) Glutathione metabolism (mean Δ Activity = -0.467, p-value = 5.67e-06); e) Lysine degradation (mean Δ Activity = -0.465, p-value = 5.67e-06) as shown in Fig. 4. A total of 14 metabolic pathways were differentially activated (enriched) out of the total of 26 metabolic pathways.

Discussion

Utilizing untargeted metabolomics, our investigation sought to characterize the metabolic profiles of individuals with obesity only (OBO) “without metabolic syndrome” in comparison to lean healthy control (LHC) individuals. Through comprehensive analysis, we successfully identified enriched metabolic pathways and identified specific metabolites exhibiting notable concentration variations between the two groups.

Our findings highlight the influence of excess adiposity on diverse metabolic pathways, with particular emphasis on those associated with amino acids and lipid metabolism. The intricate interplay within these pathways sheds light on the molecular alterations that accompany obesity, offering valuable insights into the metabolic shifts occurring in individuals with this condition. The alteration in the arginine biosynthesis pathway in obesity is associated with potential impairments in nitric oxide production, influencing vascular function ²⁵. Also, obesity induces perturbations in D-glutamine and D-glutamate metabolism, potentially affecting neurotransmission and glutamate signaling ²⁶. The dysregulation of cysteine and methionine metabolism, closely linked to oxidative stress heightened in obesity, leads to an increased production of reactive oxygen species ²⁷. Nitrogen metabolism, crucial for protein synthesis and cellular function, is also impacted by obesity, potentially influencing the delicate balance between anabolism and catabolism ²⁸. The dysregulation in butanoate metabolism suggests challenges in energy homeostasis for individuals with obesity ²⁹. Moreover, alterations in glyoxylate and dicarboxylate metabolism in obesity may contribute to disruptions in energy metabolism and intermediary pathways ³⁰. Disturbances in histidine metabolism in obesity may affect histamine levels, potentially contributing to inflammation and metabolic dysfunction. Lower histidine levels are associated with increased inflammation and oxidative stress, which positively correlate with insulin resistance (IR) and metabolic disorders ³¹. The exploration of amino acid metabolism revealed significant perturbations, providing evidence of altered biochemical processes in individuals with obesity. Moreover, the impact on sugar metabolism pathways elucidates the intricate connection between metabolic dysregulation and the development of obesity.

This study elucidated noteworthy disparities in metabolite profiles especially related to lipids for lean and obese healthy people. Among the identified metabolites, several sphingomyelins exhibited a substantial increase in the obese healthy group compared to the lean healthy group, suggesting an elevated synthesis of lipids in the former. These findings align with previous research associating sphingolipids with key metabolic processes ^{32,33}. In a prior investigation ³², sphingolipids were implicated in loss of insulin sensitivity, pro-inflammatory state characteristic of diabetes, and triggering cell death and dysfunction in critical organs like the pancreas and heart. However, despite these connections, the specific roles of distinct sphingolipid species and the metabolic pathways they participate in remain elusive. Our study not only validates and supplements the findings of previous research but also emphasizes the need for further exploration. The observed increase in sphingomyelins can help researchers to delve deeper into the roles of these specific sphingolipids in the development of metabolic disorders.

Reduced levels of guanidinoacetate have been identified as a factor contributing to alterations in creatine metabolism ³⁴. This phenomenon bears significant implications for both muscular development and the onset of cardiovascular diseases ³⁵. The intricate relationship between guanidinoacetate and creatine metabolism underscores their pivotal

roles in physiological processes. Furthermore, the synthesis of creatine, a vital compound for energy transfer within cells, especially in muscle tissue, relies significantly on guanidinoacetate, a precursor to creatine ³⁶.

The findings of our current study contribute to a comprehensive understanding of the metabolic transitions associated with obesity by comparing lean healthy individuals with obese metabolically healthy subjects. In our previous research ^{37–40}, we focused on comparing metabolically healthy and unhealthy obese individuals. By extending our investigation to include lean and obese metabolically healthy groups, we gain valuable insights into the sequential progression from a lean healthy state to obesity, and subsequently, to metabolically unhealthy conditions. Notably, our analysis reveals distinct metabolic signatures in these transitional phases. During the initial transition from lean to obese metabolically healthy, significant alterations in the activity of lipid and amino acid pathways are observed. In contrast, the subsequent shift to metabolically unhealthy states is characterized by pronounced changes in sugar-related pathways. These findings shed light on the nuanced metabolic dynamics accompanying the progression from lean health to obesity and provide crucial information for developing targeted interventions to prevent or manage metabolic disorders associated with obesity.

The limitations of the present study include the small number of participants imposed by the strict age-matching criteria; potentially limiting the generalizability of results to broader populations. The small participant pool underscores the importance of replicating the study with larger and more diverse cohorts to enhance the robustness and external validity of the findings. The other limitation is the cross-sectional nature of the study design which only provides a snapshot of associations among variables; this precludes the establishment of causal relationships between different factors. Furthermore, the reliance on blood serum as the primary source of metabolites and transcripts poses other limitations. Serum metabolomics offers valuable insights into systemic changes but lacks the specificity to elucidate tissue-specific dysregulations; and whole blood transcriptomic limits analysis to mostly white blood cells. Understanding the intricacies of metabolic dysregulations at the tissue level is essential for a comprehensive understanding of the underlying mechanisms.

In summary, our untargeted metabolomics approach has successfully unraveled the distinctive metabolic signatures associated with obesity, emphasizing the perturbed pathways involving amino acids and lipid metabolism. This research contributes to the broader comprehension of obesity-related metabolic alterations, fostering opportunities for further investigation and the development of targeted interventions in the pursuit of effective therapeutic strategies.

Declarations

Author Contributions: F.A.M, E.U, R.M., A.I and A.B.A.S conceived the study. F.A.M., A.I., H.A.H., F.C., M.A. and M.S. collected, purified, and harmonized the biological samples. E.U. built the computational pipeline and performed bioinformatics analysis. R.M. provided their inputs to the analysis. A.B.A.S and S T supervised the analysis. F.A.M. and E.U. wrote the manuscript. All authors proofread and approved the manuscript.

Disclosure: The authors declared no conflict of interest.

Funding: Research conducted on Qatar cohort was funded by Qatar Metabolic Institute (QMI), Hamad Medical Corporation, Doha Qatar.

Statement of Ethics:

The study was conducted according to the Ministry of Public Health (MOPH) guidelines and approved by the Institutional Review Board Research Ethics Committee of the Hamad Medical Corporation (HMC, IRB protocol #16245/16).

Data Availability Statement: All data generated or analyzed during this study are included in this article and its supplementary material files. Further enquiries can be directed to the corresponding author.

References

1. Kivimaki M, Strandberg T, Pentti J, Nyberg ST, Frank P, Jokela M *et al.* Body-mass index and risk of obesity-related complex multimorbidity: an observational multicohort study. *Lancet Diabetes Endocrinol* 2022; 10(4): 253–263.
2. Isomaa B, Almgren P, Tuomi T, Forsen B, Lahti K, Nissen M *et al.* Cardiovascular morbidity and mortality associated with the metabolic syndrome. *Diabetes Care* 2001; 24(4): 683–9.
3. Alexander CM, Landsman PB, Teutsch SM, Haffner SM, Third National H, Nutrition Examination S *et al.* NCEP-defined metabolic syndrome, diabetes, and prevalence of coronary heart disease among NHANES III participants age 50 years and older. *Diabetes* 2003; 52(5): 1210–4.
4. Lakka HM, Laaksonen DE, Lakka TA, Niskanen LK, Kumpusalo E, Tuomilehto J *et al.* The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. *JAMA* 2002; 288(21): 2709–16.
5. Pugliese G, Liccardi A, Graziadio C, Barrea L, Muscogiuri G, Colao A. Obesity and infectious diseases: pathophysiology and epidemiology of a double pandemic condition. *Int J Obes (Lond)* 2022; 46(3): 449–465.
6. Yates T, Summerfield A, Razieh C, Banerjee A, Chudasama Y, Davies MJ *et al.* A population-based cohort study of obesity, ethnicity and COVID-19 mortality in 12.6 million adults in England. *Nat Commun* 2022; 13(1): 624.
7. Collaboration NCDRF. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *Lancet* 2016; 387(10026): 1377–1396.
8. Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS *et al.* Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA* 2003; 289(1): 76–9.
9. Karelis AD, Brochu M, Rabasa-Lhoret R. Can we identify metabolically healthy but obese individuals (MHO)? *Diabetes Metab* 2004; 30(6): 569–72.
10. Guo F, Garvey WT. Cardiometabolic disease risk in metabolically healthy and unhealthy obesity: Stability of metabolic health status in adults. *Obesity (Silver Spring)* 2016; 24(2): 516–25.
11. Eckel N, Meidtner K, Kalle-Uhlmann T, Stefan N, Schulze MB. Metabolically healthy obesity and cardiovascular events: A systematic review and meta-analysis. *Eur J Prev Cardiol* 2016; 23(9): 956–66.
12. Wei R, Wang J, Su M, Jia E, Chen S, Chen T *et al.* Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. *Sci Rep* 2018; 8(1): 663.
13. Pang Z, Chong J, Li S, Xia J. MetaboAnalystR 3.0: Toward an Optimized Workflow for Global Metabolomics. *Metabolites* 2020; 10(5).
14. Xia J, Wishart DS. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res* 2010; 38(Web Server issue): W71–7.
15. Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* 2009; 10: 161.
16. Luo W, Pant G, Bhavnasi YK, Blanchard SG, Jr., Brouwer C. Pathview Web: user friendly pathway visualization and data integration. *Nucleic Acids Res* 2017; 45(W1): W501–W508.
17. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004; 20(3): 307–15.
18. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. *Bioinformatics* 2010; 26(19): 2363–7.

19. MacDonald JW. Platform Design Info for Affymetrix Clariom_D_Human. 2016; (R package version 3141).
20. MacDonald JW. affycoretools: Functions useful for those doing repetitive analyses with Affymetrix GeneChips. 2008; (R package version 1640;).
21. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43(7): e47.
22. Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013; 14: 7.
23. Mall R, Saad M, Roelands J, Rinchai D, Kunji K, Almeer H *et al.* Network-based identification of key master regulators associated with an immune-silent cancer phenotype. *Brief Bioinform* 2021; 22(6).
24. Sherif S, Mall R, Almeer H, Naik A, Al Homaid A, Thomas R *et al.* Immune-related 3-lncRNA signature with prognostic connotation in a multi-cancer setting. *J Transl Med* 2022; 20(1): 442.
25. Atawia RT, Bunch KL, Toque HA, Caldwell RB, Caldwell RW. Mechanisms of obesity-induced metabolic and vascular dysfunctions. *Front Biosci (Landmark Ed)* 2019; 24(5): 890–934.
26. He R, Zheng R, Zheng J, Li M, Wang T, Zhao Z *et al.* Causal Association Between Obesity, Circulating Glutamine Levels, and Depression: A Mendelian Randomization Study. *J Clin Endocrinol Metab* 2023; 108(6): 1432–1441.
27. Furukawa S, Fujita T, Shimabukuro M, Iwaki M, Yamada Y, Nakajima Y *et al.* Increased oxidative stress in obesity and its impact on metabolic syndrome. *J Clin Invest* 2004; 114(12): 1752–61.
28. Simonson M, Boirie Y, Guillet C. Protein, amino acids and obesity treatment. *Rev Endocr Metab Disord* 2020; 21(3): 341–353.
29. van Deuren T, Blaak EE, Canfora EE. Butyrate to combat obesity and obesity-associated metabolic disorders: Current status and future implications for therapeutic use. *Obes Rev* 2022; 23(10): e13498.
30. Tong L, Tian M, Ma X, Bai L, Zhou J, Ding W. Metabolome Profiling and Pathway Analysis in Metabolically Healthy and Unhealthy Obesity among Chinese Adolescents Aged 11–18 Years. *Metabolites* 2023; 13(5).
31. Feng RN, Niu YC, Sun XW, Li Q, Zhao C, Wang C *et al.* Histidine supplementation improves insulin resistance through suppressed inflammation in obese women with the metabolic syndrome: a randomised controlled trial. *Diabetologia* 2013; 56(5): 985–94.
32. Russo SB, Ross JS, Cowart LA. Sphingolipids in obesity, type 2 diabetes, and metabolic disease. *Handb Exp Pharmacol* 2013; (216): 373–401.
33. Green CD, Maceyka M, Cowart LA, Spiegel S. Sphingolipids in metabolic disease: The good, the bad, and the unknown. *Cell Metab* 2021; 33(7): 1293–1306.
34. Ingoglia F, Chong JL, Pasquali M, Longo N. Creatine metabolism in patients with urea cycle disorders. *Mol Genet Metab Rep* 2021; 29: 100791.
35. Aksentijevic D, Zervou S, Eykyn TR, McAndrew DJ, Wallis J, Schneider JE *et al.* Age-Dependent Decline in Cardiac Function in Guanidinoacetate-N-Methyltransferase Knockout Mice. *Front Physiol* 2019; 10: 1535.
36. da Silva RP, Nissim I, Brosnan ME, Brosnan JT. Creatine synthesis: hepatic metabolism of guanidinoacetate and creatine in the rat in vitro and in vivo. *Am J Physiol Endocrinol Metab* 2009; 296(2): E256-61.
37. Mir FA, Ullah E, Mall R, Iskandarani A, Samra TA, Cyprian F *et al.* Dysregulated Metabolic Pathways in Subjects with Obesity and Metabolic Syndrome. *Int J Mol Sci* 2022; 23(17).
38. Mir FA, Mall R, Iskandarani A, Ullah E, Samra TA, Cyprian F *et al.* Characteristic MicroRNAs Linked to Dysregulated Metabolic Pathways in Qatari Adult Subjects With Obesity and Metabolic Syndrome. *Front Endocrinol (Lausanne)* 2022; 13: 937089.

39. Mir FA, Mall R, Ullah E, Iskandarani A, Cyprian F, Samra TA *et al.* An integrated multi-omic approach demonstrates distinct molecular signatures between human obesity with and without metabolic complications: a case-control study. *J Transl Med* 2023; 21(1): 229.

40. Mir FA, Abdesselem HB, Cyprian F, Iskandarani A, Doudin A, Samra TA *et al.* Inflammatory protein signatures in individuals with obesity and metabolic syndrome. *Sci Rep* 2023; 13(1): 22185.

Figures

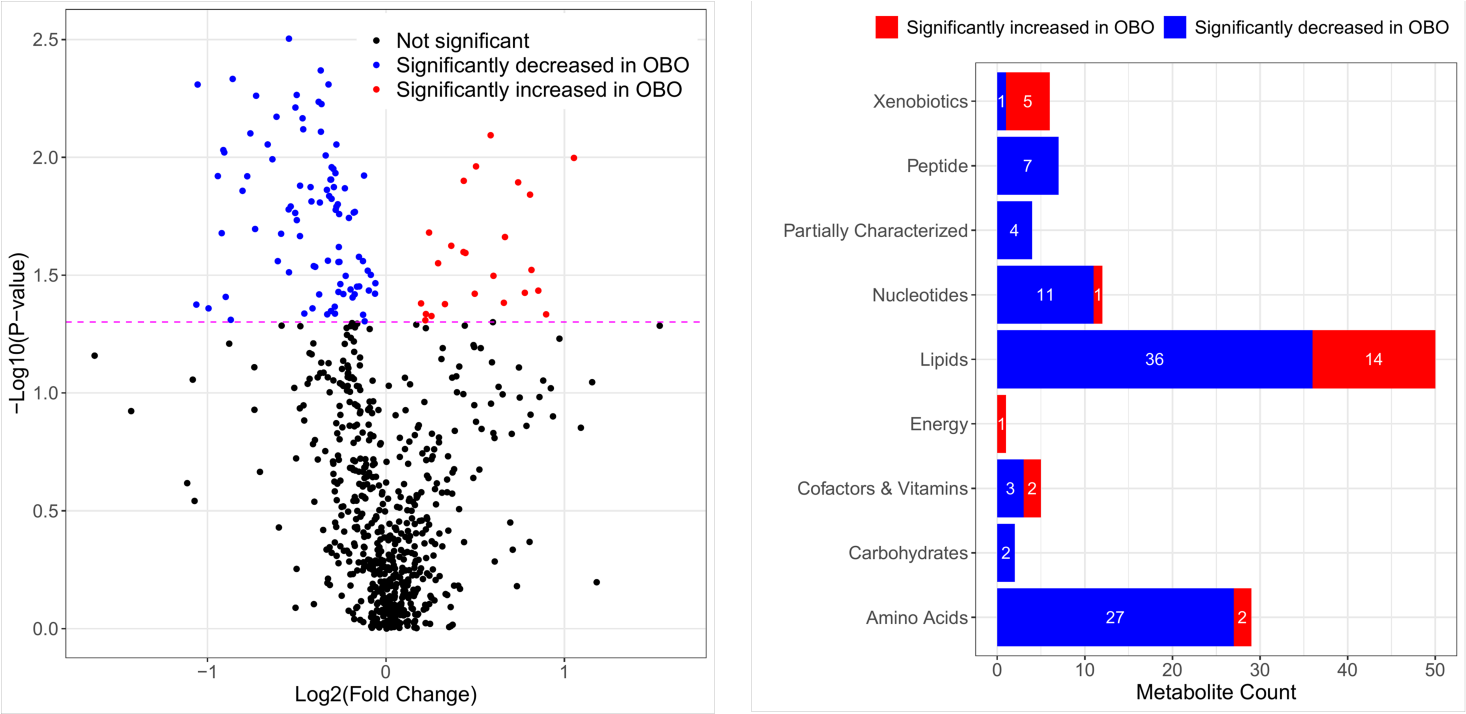


Figure 1

Changes in metabolite concentrations in OBO vs. LHC groups.

(A) Volcano plot illustrating the log₂(fold-change) versus -log₁₀(p-value) of metabolites. Metabolites positioned above the dashed line indicate statistical significance with $p < 0.05$. **(B)** Bar plot presenting the quantification of metabolites significantly altered in concentration within specific super pathways, highlighting the number of metabolites that exhibit increased or decreased levels in the OBO compared to LHC groups.

■ LHC ■ OBO

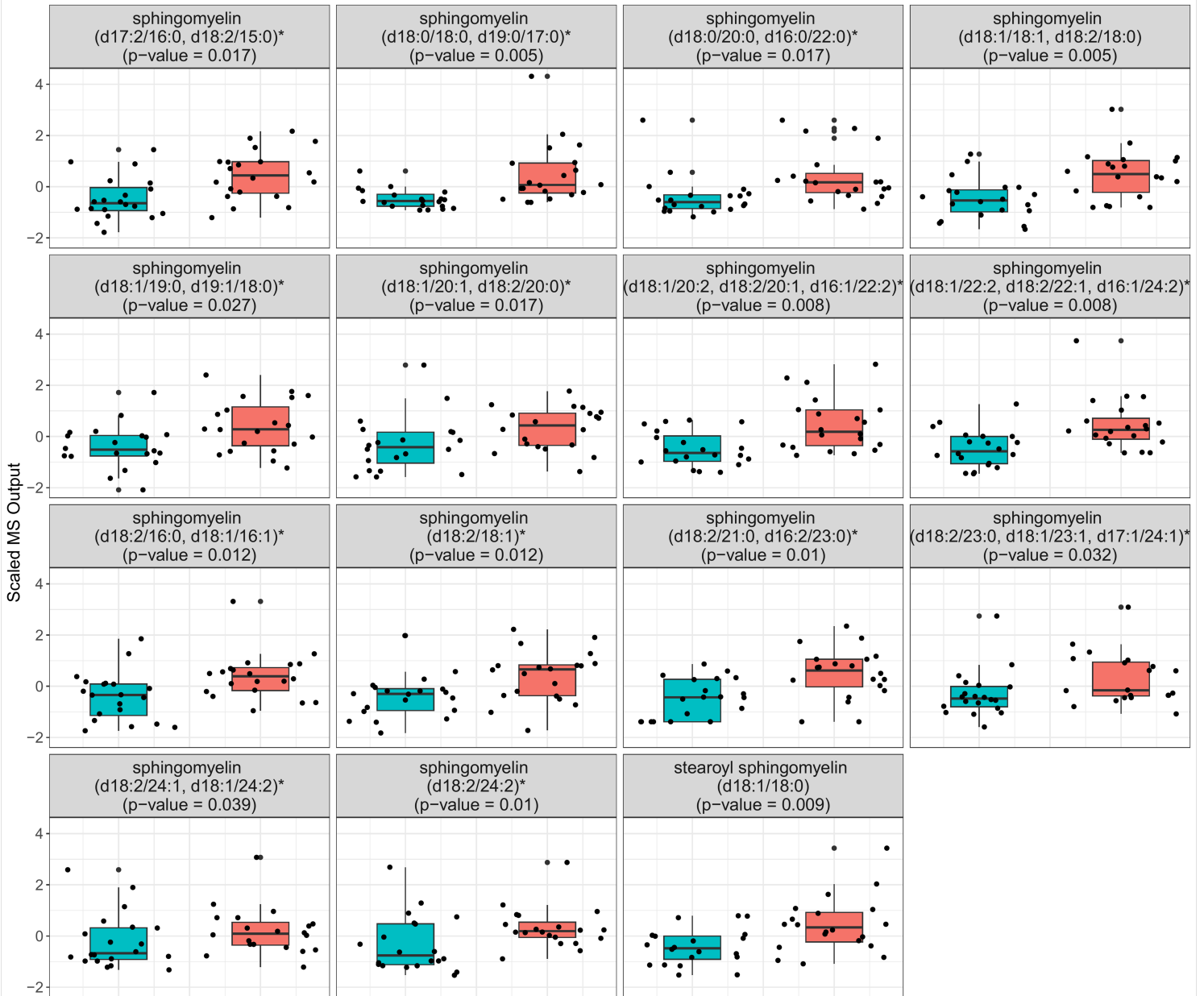


Figure 2

Distribution of statistically significant sphingomyelins observed in OBO versus LHC groups.

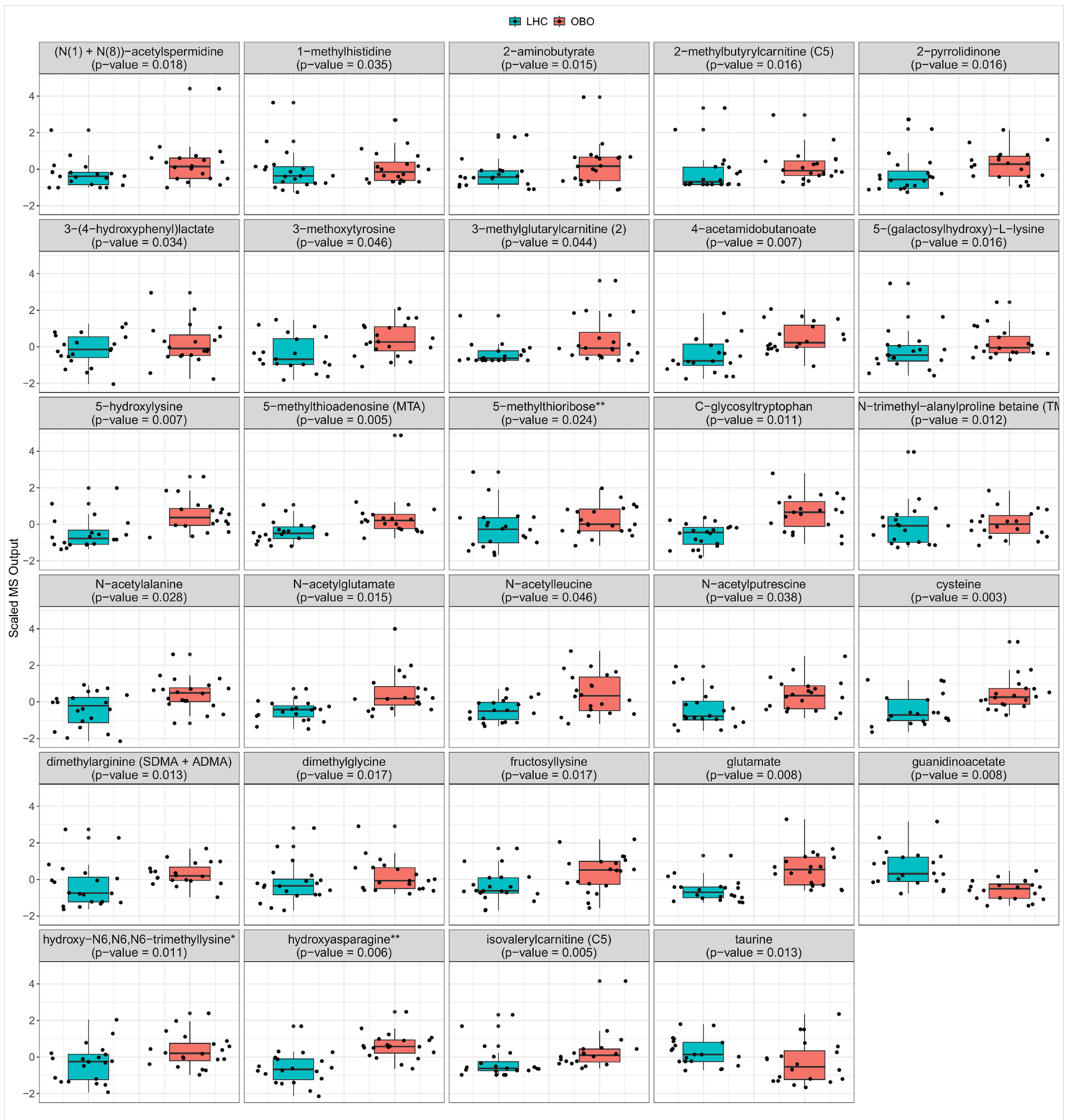


Figure 3

Distribution of statistically significant amino acids in lean healthy (LHC) and metabolically healthy individuals with obesity (OBO) groups.

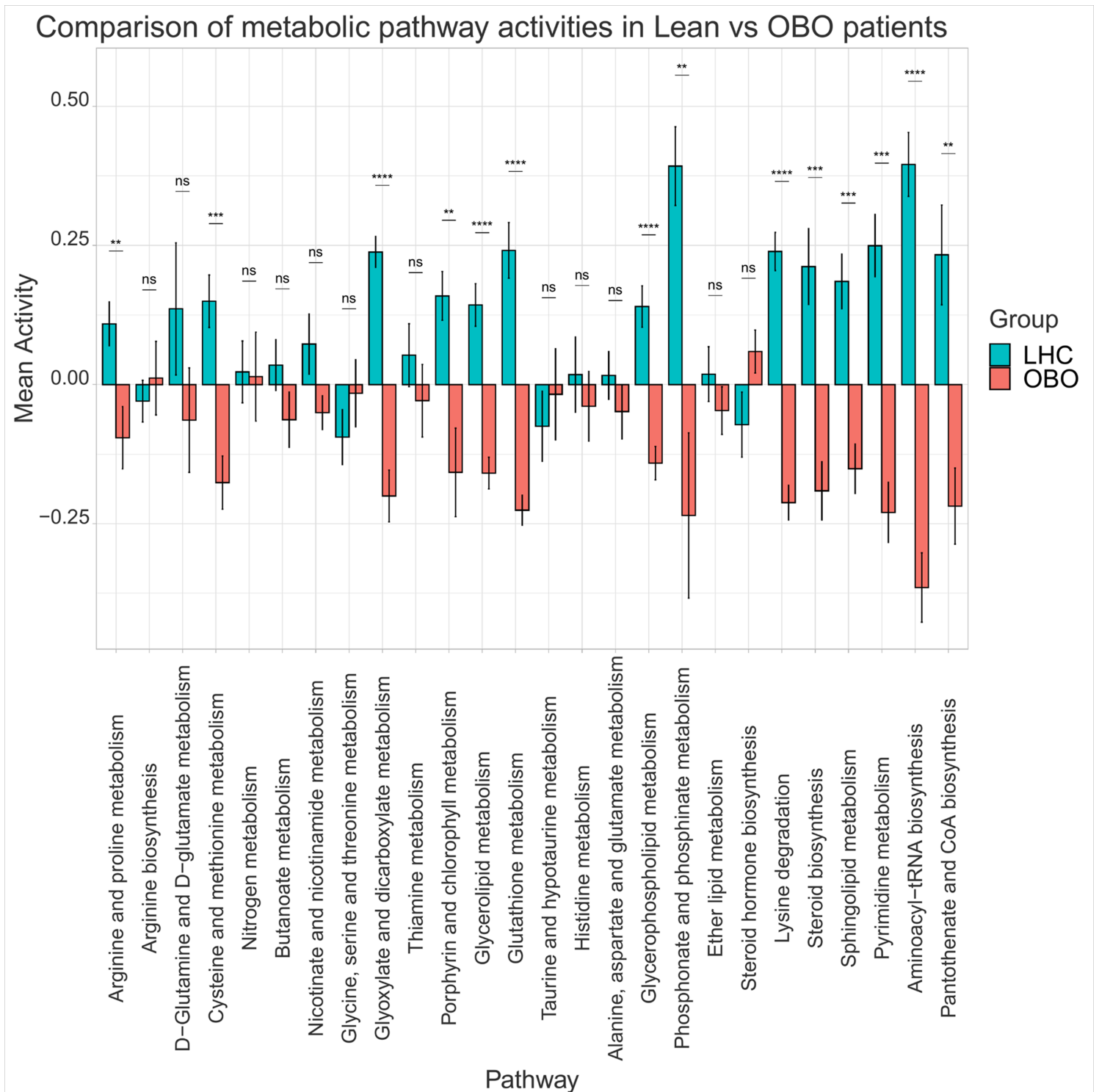


Figure 4

Comparative analysis of metabolic pathway activities between OBO and LHC groups.