


RESEARCH

Open Access



# An integrated multi-omic approach demonstrates distinct molecular signatures between human obesity with and without metabolic complications: a case–control study

Fayaz Ahmad Mir<sup>1\*†</sup>, Raghvendra Mall<sup>3,7\*†</sup>, Ehsan Ullah<sup>2\*†</sup>, Ahmad Iskandarani<sup>1</sup>, Farhan Cyprian<sup>4</sup>, Tareq A. Samra<sup>1</sup>, Meis Alkasem<sup>1</sup>, Ibrahim Abdalhakam<sup>1</sup>, Faisal Farooq<sup>2</sup>, Shahrada Taheri<sup>1,5,6</sup> and Abdul-Badi Abou-Samra<sup>1,5,6</sup>

## Abstract

**Objectives** To examine the hypothesis that obesity complicated by the metabolic syndrome, compared to uncomplicated obesity, has distinct molecular signatures and metabolic pathways.

**Methods** We analyzed a cohort of 39 participants with obesity that included 21 with metabolic syndrome, age-matched to 18 without metabolic complications. We measured in whole blood samples 754 human microRNAs (miRNAs), 704 metabolites using unbiased mass spectrometry metabolomics, and 25,682 transcripts, which include both protein coding genes (PCGs) as well as non-coding transcripts. We then identified differentially expressed miRNAs, PCGs, and metabolites and integrated them using databases such as mirDIP (mapping between miRNA-PCG network), Human Metabolome Database (mapping between metabolite-PCG network) and tools like MetaboAnalyst (mapping between metabolite-metabolic pathway network) to determine dysregulated metabolic pathways in obesity with metabolic complications.

**Results** We identified 8 significantly enriched metabolic pathways comprising 8 metabolites, 25 protein coding genes and 9 microRNAs which are each differentially expressed between the subjects with obesity and those with obesity and metabolic syndrome. By performing unsupervised hierarchical clustering on the enrichment matrix of the 8 metabolic pathways, we could approximately segregate the uncomplicated obesity strata from that of obesity with metabolic syndrome.

<sup>†</sup>Fayaz Ahmad Mir, Raghvendra Mall and Ehsan Ullah equally contributed, co-first authors

\*Correspondence:

Fayaz Ahmad Mir  
fmir1@hamad.qa  
Raghvendra Mall  
raghvendra.mall@stjude.org  
Ehsan Ullah  
eullah@hbku.edu.qa

Full list of author information is available at the end of the article



**Conclusions** The data suggest that at least 8 metabolic pathways, along with their various dysregulated elements, identified via our integrative bioinformatics pipeline, can potentially differentiate those with obesity from those with obesity and metabolic complications.

## Introduction

Obesity is a chronic progressive relapsing disease occurring when excessive fat accumulation exerts a deleterious effect on health [1]. Obesity can have multiple adverse health consequences including the development of the metabolic syndrome, characterized by insulin resistance, dyslipidemia, hypertension, and/or type 2 diabetes mellitus. The metabolic syndrome, in turn, increases the risk of atherosclerotic cardiovascular disease such as coronary artery disease, stroke and peripheral artery disease [2]. There is growing interest in defining obesity phenotypes to understand the relationship between excess adiposity and metabolic health. A subgroup, termed “Metabolically Unhealthy Obesity” phenotype (obesity with metabolic syndrome or OBM) develop metabolic complications while another subgroup, termed as “Metabolically Healthy Obesity” (obesity only or OBO), do not demonstrate or progress to the metabolic syndrome, even in later life [3–9].

The various definitions of obesity phenotypes and various disease outcomes studies have resulted in diverse findings. In particular, there is controversy regarding cardiovascular outcomes in OBO individuals; either a similar risk profile or a greater risk profile than healthy weight individuals has been reported [5, 10–16]. A lower risk profile for OBO compared to OBM has been observed. There have been suggestions that OBO status is transient, progressing to OBM; nevertheless, perhaps, the OBO phenotype is associated with longer protection from metabolic deterioration. One potential contributing factor to the development of metabolic consequences in OBM is body fat distribution [17]. Central abdominal adiposity and accompanied high liver fat content have been linked to the OBM phenotype, whereas subcutaneous adiposity has been associated with the OBO phenotype [18, 19]. Apart from fat distribution, the potential pathways mediating adiposity and metabolic phenotypes include alterations in adipogenesis, adipose tissue function and inflammation, angiogenesis, immune dysregulation, and mitochondrial dysfunction [20].

The study of OBO and OBM phenotypes provides insight into the biological pathways differentially affected by excess body fat in each phenotype allowing prognostication and development of future treatments to target specific pathways to protect against metabolic dysfunction. In this work, we used an integrative bioinformatics pipeline to identify enriched metabolic pathways by

integrating data from transcriptional signatures, microRNAs and circulating metabolites, to obtain a more holistic picture of the differential enrichment of metabolic pathways between OBO and OBM.

## Materials and methods

### Study design

The study was a cross-sectional study approved by the institutional review board (IRB) of Hamad Medical Corporation (HMC, IRB protocol #16245/16) and all participants provided written informed consent. Participants were recruited at the Qatar Metabolic Institute, Hamad Medical Corporation, Doha, Qatar. After signing the informed consent, they were invited to the study site for anthropometric measurements and collection of fasting blood samples between 7 and 9 A.M. after 12–14 h of fasting.

Participants included both men and women aged 18 to 65 years with body mass index (BMI)  $\geq 35$  kg/m<sup>2</sup> with or without the metabolic syndrome but without other chronic disease or terminal illness. Participants were classified into two groups, based on presence of the components of the metabolic syndrome (obesity only [OBO] versus obesity with metabolic syndrome [OBM]) using the International Diabetes Federation (IDF) metabolic syndrome criteria: BMI  $> 30$  kg/m<sup>2</sup> or waist circumference above the ethnic threshold (102 cm for men and 88 cm for women) PLUS any 2 of: triglycerides  $\geq 150$  mg/dL (1.7 mmol/L), HDL-cholesterol  $< 40$  mg/dL (1.03 mmol/L) in men or  $< 50$  mg/dL (1.29 mmol/L) in women, blood pressure  $\geq 130 / 85$  mmHg, and fasting blood glucose  $\geq 110$  mg/dL (5.6 mmol/L).

### Blood chemistry assays

For serum collection, whole blood was collected via BD Vacutainer Serum Separation Tubes (BD Biosciences, Franklin Lakes, NJ, USA). Blood samples were kept at room temperature for 30–60 min, and then centrifuged at 3000g for 10 min. Following centrifugation, serum was separated, aliquoted and immediately stored at  $-80$  °C for further use. Blood biochemistry measurements were performed at the HMC clinical laboratory. Measurements included HbA1c via Turbidimetric Inhibition Immunoassay (TINIA Roche Diagnostics, Mannheim, Germany), glucose by enzymatic reference method with hexokinase (Cobas 6000, Roche Diagnostics

International, Switzerland), and triglycerides, total and HDL cholesterol by calorimetric assays.

#### 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  $-80^{\circ}\text{C}$  for storage. The samples were thawed overnight, then total RNA was isolated with a PAXgene Blood RNA Kit including the DNase Set (Qia- gen). 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).

#### MicroRNA (miRNA) profiling

The expression levels of 754 miRNAs were determined using the TaqMan OpenArray Human MicroRNA panels (PN: 4,470,189; Life Technologies Foster City, CA, USA) on a QuantStudio 12 K Flex instrument. For all experimental groups, 3  $\mu\text{L}$  (10 ng) of total RNA was used for reverse transcription (RT) reactions using MegaPlex RT Primers Human Pool Set v3.0 (PN: 4,444,745; Pool A v2.1 and Pool B v3.0) according to the manufacturer's optimized protocol for low sample input for profiling human microRNA, using the OpenArray platform on BioRad c1000 Touch thermal cycler. No-template controls were included. Pre-amplification of RT products was performed using 5  $\mu\text{L}$  RT reaction combined with the matching Megaplex PreAmp Primer Pool A v2.1 or B v3.0 and amplified using the thermal cycler (Applied biosystems). The pre-amplified products were diluted 1:40 in  $0.1 \times \text{TE}$  pH 8. For each experimental set, 10  $\mu\text{L}$  of the diluted products were combined to give a total of 40  $\mu\text{L}$  pooled sample. For both Pool A and Pool B groups, 22.5  $\mu\text{L}$  of the pooled products were combined with an equivalent volume of TaqMan OpenArray Real-Time Master Mix and aliquoted into a 96 well plate. Then, 5  $\mu\text{L}$  from each well were then transferred into a 384 well plate for loading onto OpenArray plates using an AccuFill robotic system. The OpenArray plates were run on a QuantStudio 12 K Flex instrument (Life Technologies) and the raw data files were imported and analyzed using the DataAssist software (Life Technologies). Failed reactions were excluded from analysis and undetermined CT values for sample sets determined to have good amplifications were assigned a threshold value of 40, defining low abundance or absence of miRNA expression. Global mean normalization was used to calculate relative fold change for the miRNA expression.

#### Transcriptomic profiling

A Clariom D Assay-Human system (Applied Biosystems) containing 47,231 expressions and 770 control probes was used for transcriptomics. In brief, 100 ng of total RNA was reverse-transcribed into cDNA using GeneChip™ WT PLUS Reagent Kit (Applied biosystems); 5.5  $\mu\text{g}$  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 [21], oligo v1.60.0 [22], clariomdhumantranscriptcluster.db v8.8.0 [23], pd.clariom.d.human v3.14.1 [24], affycoretools 1.68.1 [25] 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 clariomhumantranscriptcluster.db 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 25,682 genes (including PCGs as well as non-coding genes) with their corresponding expression values for the 39 participant profiles.

#### Metabolomic profiling

Metabolomic profiling was performed using established protocols at Metabolon, Durham, NC, USA. Sample accession, sample preparation, sample processing and QC generation was done at the Anti-Doping Laboratory—Qatar (ADLQ) with final review of QC data and curation by Metabolon. Samples were prepared using the automated MicroLab STAR® system from Hamilton Company. Several recovery standards were added prior to the first step in the extraction process for QC purposes. Proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into five fractions, the sample extracts were stored overnight under nitrogen before preparation for analysis. The study utilized Waters ACQUITY UPLC, and a Thermo Scientific Q-Exactive high resolution/accurate

mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. The extracted five fractions were reconstituted in solvents compatible to each of the four methods. Two fractions were analyzed by two separate reverse phases (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one was analyzed by RP/UPLC-MS/MS with negative ion mode ESI, one analyzed by hydrophilic interaction chromatography (HILIC)/UPLC-MS/MS with negative ion mode ESI, and one sample was reserved for backup. The analysis was done on two types of columns, C18 column (Waters UPLC BEH C18-2.1 × 100 mm, 1.7 μm) and HILIC column (Waters UPLC BEH Amide 2.1 × 150 mm, 1.7 μm). Raw data were extracted, peak-identified and quality control processed using Metabolon's hardware and software. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. A total of 1069 metabolites were profiled; 174 unnamed metabolites were removed leaving 895 metabolites for the rest of the analysis.

Several quality control steps were used to ensure data quality. Assuming values to be missing at random, 191 metabolites having less than 80% non-missing values (more than 20% missing values per metabolite due to biological and/or technical reasons) were removed [26] resulting in 704 metabolites for downstream analysis. Missing values in the remaining data were imputed by replacing the missing values for each metabolite with the minimum value detected for the metabolite. Outliers in samples were identified using principal component analysis (PCA) [27, 28] and only one sample was removed.

#### Data analysis

Statistical characteristics of clinical measurements were calculated by comparing the OBO and OBM samples using R [29]. Student's t-test was used to calculate p-value of the normally distributed measurement; group variance in this case was tested using F-test. For remaining measurements, the Mann–Whitney test was used. The normality of the measurements was tested using the Anderson–Darling test. P-values were not corrected for false discovery rate (FDR) due to small sample size.

#### MiRNA – PCG interaction network

We utilized the microRNA Data Integration Portal, mirDIP v4.1 (<http://ophid.utoronto.ca/mirDIP/>), which consists of over 152 million human microRNA–target predictions collected from 30 different resources [30–32]. The mirDIP integrative score was constructed by taking a statistical consensus from the predictions available through myriad resources and was assigned to each unique miRNA–target interaction to provide a unified

measure of confidence. The integrated scores range, 0 to 1, was used; higher scores correspond to stronger evidence of potential interaction between miRNA and target protein coding gene (PCG); the target PCGs were thus identified.

By focusing on interactions between differentially expressed miRNAs (OBM vs OBO) and PCGs, we initially have a reduced set of 53,625 miRNA–target gene interactions (MTIs). Further, focusing on those MTIs, where the number of available sources for interaction was  $>=9$  (including reporter assay, western blot, microarray, NGS etc.) and the integrated score was  $>=0.25$ , resulted in 17,797 unique MTIs. Additionally, for our use case, we removed those MTIs where the target PCG was not differentially expressed. This resulted in a miRNA–PCG network consisting of 457 nodes (55 DE–miRNAs and 402 differentially expressed protein coding genes (DE–PCGs)) and 991 miRNA–PCG interaction edges.

#### Metabolite set enrichment analysis

Metabolite set enrichment analysis (MSEA) or pathway enrichment analysis was performed using the online version of Metaboanalyst 5.0 (<https://www.metaboanalyst.ca/>). Data file (comma separated values) was prepared using R for Metaboanalyst containing samples, disease status, compounds IDs, and values of metabolites for each sample. The compounds were matched using HMDB IDs provided by Metabolon. For pathway enrichment analysis, autoscaling was used to normalize each variable to be mean centered with a unit variance. KEGG human metabolic pathways corresponding to 84 metabolite sets were used as the reference pathways for the enrichment analysis. Metabolite sets having less than two metabolites in our dataset were filtered out.

#### Single sample gene set enrichment

To obtain an estimate of enrichment of the metabolic pathways, identified through our integrative multi-omics analysis framework, for each sample, we used the Gene Set Variation Analysis (GSVA) function in R [33–35]. 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  $i^{\text{th}}$  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 each  $G(p_i)$  is obtained, an unsupervised hierarchical clustering was then performed to determine whether the enrichment scores of the metabolic pathways can distinguish OBO from OBM using the ‘hclust’ function with ‘ward.D’ method in R. This helps to narrow down the targetable metabolic pathways which are differentially enriched with respect to the phenotype.

**Integrative bioinformatics pipeline**

The myriad omics profiles including miRNA through RT-PCR, transcriptional signatures through Affymetrix and metabolic profiles through LC-MS technique are illustrated in Fig. 1. The differences between OBO and OBM subjects for each omics type are deduced using a Bayesian statistical framework, namely LIMMA [36] (limma package v3.52.4 in R), which identifies differentially expressed miRNAs, PCGs and metabolites at a p-value threshold of 0.05 (no FDR correction owing to small sample size).

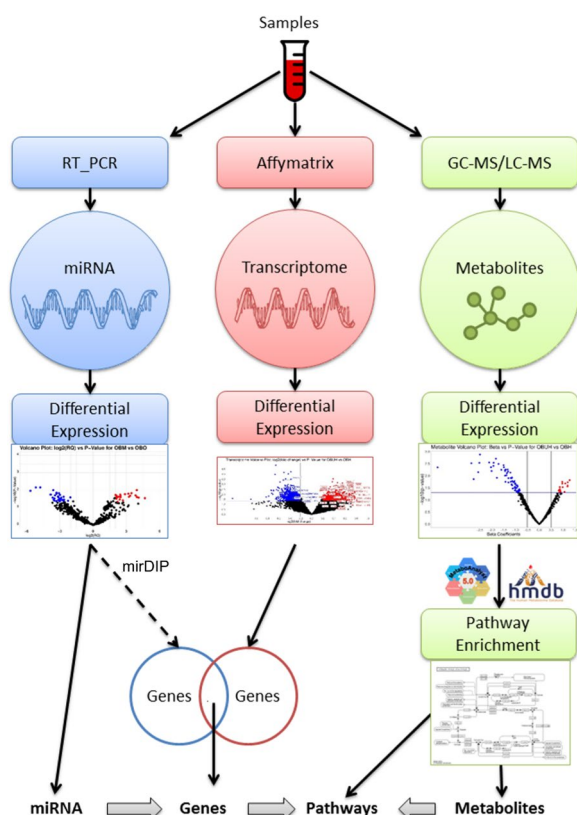
We then perform metabolite set enrichment analysis resulting in enriched metabolic pathways. We used the mirDIP database to obtain information about MTIs. For our use case, we focused on those target PCGs which

are involved in the enriched metabolic pathways and were also differentially expressed with respect to the phenotype ( $p\text{-value} \leq 0.05$ , LIMMA). Thus, we concentrated on the DE-PCGs which act as enzymes/kinases in these enriched metabolic pathways. For each of these DE-PCGs, we obtained the corresponding miRNAs (also differentially expressed) that were known to regulate the expression of these genes with high integrated scores using the mirDIP database. A snapshot of the entire workflow is depicted in Fig. 1 and Additional file 1: Figure S1.

**Results**

**Clinical characteristics**

The clinical characteristics of the sample are summarized in Table 1. As OBO and OBM participants were matched for age and BMI; the mean age and BMI for OBO and OBM groups were not significantly different. The mean ages were 38.1 and 40.5 years and mean BMIs 40.0 and 39.6  $\text{kg/m}^2$ , for OBO and OBM, respectively. There were 11 women and 7 men in the OBO group and 9 women and 12 men in the OBM group. The clinical traits that were significantly different between the two groups include HBA1c ( $p=0.002$ ), fasting triglycerides ( $p=0.001$ ), high-density lipoprotein cholesterol (HDL-cholesterol,  $p=0.008$ ), fasting glucose ( $p=0.009$ ) and insulin levels ( $p=0.05$ ). Other important clinical traits which were not significantly different between the two sets include clinical variables such creatinine, total



**Fig. 1** Comprehensive information about multiple dysregulated elements involved in the metabolic pathways that are differentially enriched between OBO vs OBM

**Table 1** Clinical and biochemical traits of the study subjects. OBO, obesity only and OBM, obesity with metabolic syndrome. Significance was determined by the Student’s t-test

| Feature                          | OBO N = 18   | OBM N = 21   | P Value |
|----------------------------------|--------------|--------------|---------|
| Age (years)                      | 38.1 ± 4.2   | 40.5 ± 7.3   | 0.283   |
| Females (N; %)                   | 11 (61%)     | 9 (42.9%)    | 0.415   |
| Height (cm)                      | 167.4 ± 11.9 | 170.8 ± 9.6  | 0.370   |
| Weight (kg)                      | 113.4 ± 19.6 | 110.9 ± 27.6 | 0.782   |
| BMI ( $\text{kg/m}^2$ )          | 40.0 ± 4.5   | 39.6 ± 3.0   | 0.746   |
| Smoking (%)                      | 6.0          | 33.0         | 0.027   |
| HBA1c (%)                        | 5.5 ± 0.27   | 7.02 ± 1.9   | 0.002   |
| Fasting triglycerides (mmol/L)   | 1.39 ± 0.48  | 2.65 ± 1.52  | 0.001   |
| Total Cholesterol (mmol/L)       | 4.9 ± 1.1    | 4.8 ± 1.1    | 0.855   |
| LDL-Cholesterol (mmol/L)         | 2.8 ± 1.3    | 2.6 ± 1.1    | 0.728   |
| HDL-Cholesterol (mmol/L)         | 1.5 ± 0.7    | 1.0 ± 0.3    | 0.008   |
| Fasting glucose (mmol/L)         | 5.2 ± 0.6    | 7.4 ± 3.4    | 0.009   |
| Creatinine ( $\mu\text{mol/L}$ ) | 67.5 ± 14.1  | 65.3 ± 14.1  | 0.563   |
| Fasting insulin (miU/mL)         | 19.0 ± 13.3  | 27.6 ± 13.2  | 0.053   |
| C-reactive protein (mg/L)        | 12.8 ± 12.5  | 7.1 ± 4.5    | 0.064   |
| Alanine transaminase (U/L)       | 20.7 ± 11.6  | 36.5 ± 35.1  | 0.063   |
| Aspartate aminotransferase (U/L) | 18.8 ± 9.6   | 23.6 ± 15.0  | 0.251   |

cholesterol, and low-density lipoprotein cholesterol (LDL-cholesterol) levels. A greater percentage of the OBM group reported cigarette smoking.

**Enrichment analysis of metabolic pathways**

The results of metabolic pathway enrichment analysis using metabolite signatures in the cohort are shown in Table 2 and Fig. 2 and. Out of 61 metabolic pathways, 12 were significantly ( $p \leq 0.1$ ) enriched. Due to the small sample number, the p-values were relatively high and statistical Q was high; however, no FDR correction was used for p-values. The enriched pathways included Lysine degradation ( $Q = 11.62, p = 0.003$ ); Amino sugar and nucleotide sugar metabolism ( $Q = 14.15, p = 0.005$ ); Arginine and proline metabolism ( $Q = 6.15, p = 0.015$ ); Fructose and mannose metabolism ( $Q = 10.85, p = 0.017$ ); Galactose metabolism ( $Q = 8.53, p = 0.02$ ); Starch and sucrose metabolism ( $Q = 8.35, p = 0.05$ ); D-Glutamine and D-glutamate metabolism ( $Q = 7.45, p = 0.05$ ); Alanine, aspartate and glutamate metabolism ( $Q = 4.80, p = 0.078$ );

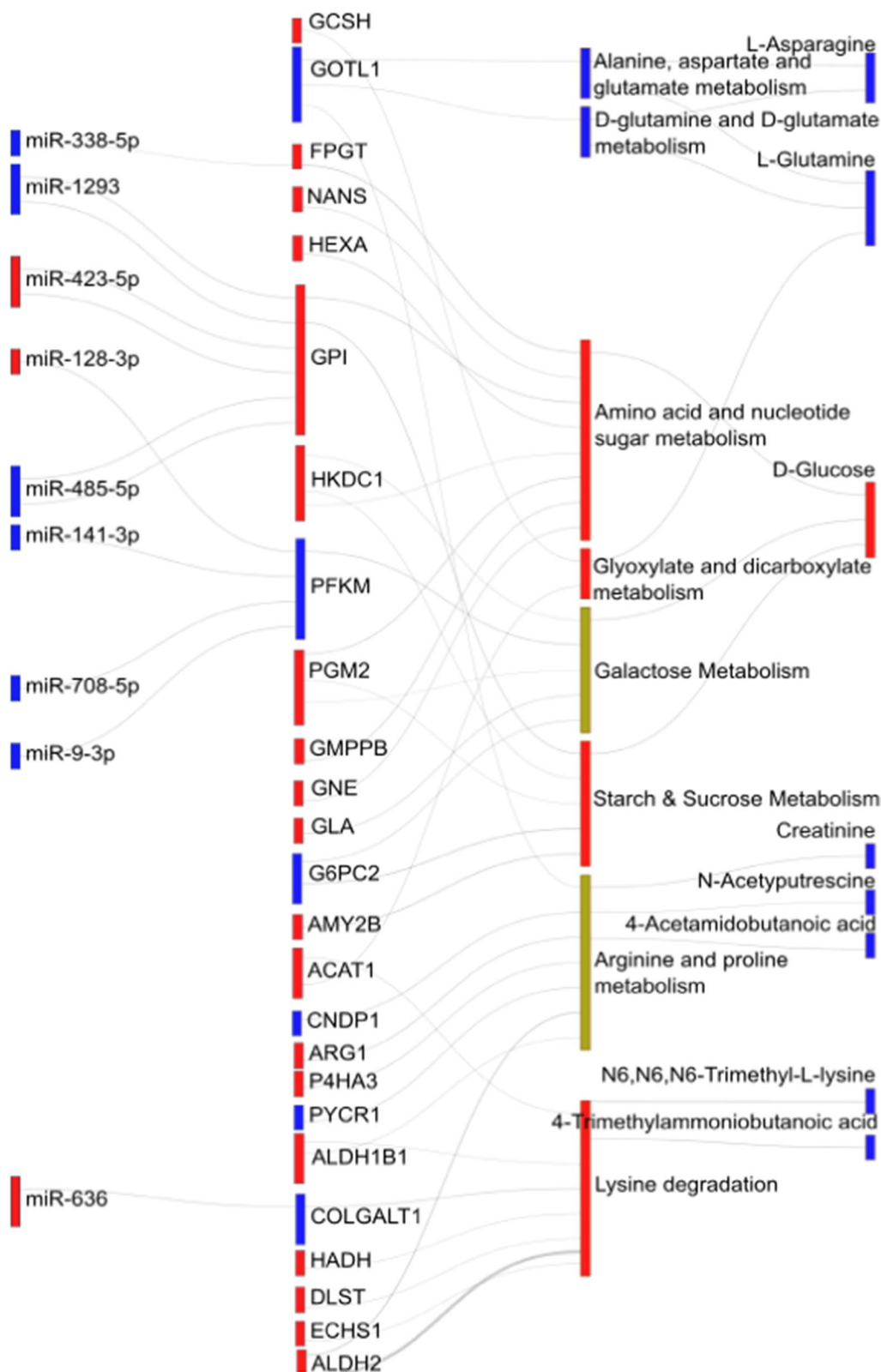
Fatty acid elongation ( $Q = 8.63, p = 0.078$ ); Steroid biosynthesis ( $Q = 7.91, p = 0.09$ ); Nicotinate and nicotinamide metabolism ( $Q = 5.44, p = 0.09$ ); and Glyoxylate and dicarboxylate metabolism ( $Q = 5.09, p = 0.09$ ).

**Regulatory elements in enriched metabolic pathways**

The metabolic pathway enrichment analysis identified 12 enriched pathways. Out of these 12 metabolic pathways, there were only 8 pathways which were associated with differentially expressed metabolites and include differentially expressed PCGs as kinases in these pathways. Hence, we focused on these 8 metabolic pathways for further analysis. The myriad regulatory elements involved in those pathways including the miRNAs, PCGs and metabolites were illustrated in Fig. 2 as a Sankey plot and in Table 2. A total of 8 differentially expressed (DE) metabolites along with 25 DE-PCGs ( $p \leq 0.05$ ) participating as kinases in at least one of these metabolic pathways, are illustrated; together with 9 upstream DE miRNAs ( $p \leq 0.05$ ) showing 12 out of 991 interaction

**Table 2** List of differentially expressed miRNA, genes and metabolites associated with enriched pathways. Here ‘NI’ refers to ‘non-identified’ or no miRNAs identified. The ‘red’ colored regulatory elements are significantly up-regulated, and the ‘blue’ colored regulatory elements are significantly down-regulated based on our framework

| Pathway                                     | miRNA  | Genes   | Metabolites   |
|---|--|---|---|
| Alanine, aspartate and glutamate metabolism | NI   | GOT1L1  | L-Asparagine, L-Glutamine                                   |
| Amino sugar and nucleotide sugar metabolism | miR-338-5p, miR-1293, miR-423-5p, miR-485-5p | FPGT, NANS, HEXA, HKDC1, GMPPB, GNE, GPI                | D-Glucose   |
| Arginine and proline metabolism             | NI   | PGM2, ALDH2, CNDP1, ARG1, ALDH1B1, P4HA3, GOT1L1, PYCR1 | Creatinine, N-Acetylputrescine, 4-Acetamidobutanoic acid    |
| D-Glutamine and D-glutamate metabolism      | NI   | GOT1L1  | L-Asparagine, L-Glutamine                                   |
| Galactose metabolism                        | miR-128-3p, miR-141-3p, miR-708-5p, miR-9-3p | PFKM, GLA, HKDC1, PGM2, G6PC2                           | D-Glucose   |
| Glyoxylate and dicarboxylate metabolism     | NI   | ACAT1, GCSH   | L-Glutamine   |
| Lysine degradation                          | miR-636                                      | HADH, ACAT1, ALDH1B1, DLST, ALDH2, COLGALT1, ECHS1      | N6,N6,N6-Trimethyl-L-lysine, 4-Trimethylammonibutanoic acid |
| Starch and sucrose metabolism               | miR-423-5p, miR-1293, miR-485-5p             | AMY2B, HKDC1, GPI, G6PC2, PGM2                          | D-Glucose   |



**Fig. 2** Sankey Plot highlighting interactions between differentially expressed regulatory elements and their associated metabolic enriched pathways. The 'red' color represents over-expression of a regulatory element while the 'blue' color corresponds to down-regulation of the regulatory element in OBM samples in comparison to the OBO samples, respectively

edges between the DE-miRNA – DE-PCGs and were highlighted in Fig. 2.

The 8 enriched metabolic pathways include amino sugar and nucleotide sugar metabolism, alanine, aspartate and glutamate metabolism, arginine and proline metabolism, D-glutamine and D-glutamate metabolism, galactose metabolism, glyoxylate and dicarboxylate metabolism, lysine degradation and starch and sucrose metabolism (Table 2). A detailed list of DE-MiRNAs, DE-PCGs, metabolites along with their direction of regulation in the enriched pathways were also highlighted (Table 2). For the majority of the metabolic pathways, the DE-miRNAs were down-regulated in OBM samples ('blue' in Fig. 2) while their corresponding target DE-PCGs were primarily up-regulated [37] except PFKM, which was significantly down-regulated. Interestingly, the metabolic pathways D-glutamine and D-glutamate metabolism and alanine, aspartate and glutamate metabolism along with its associated DE-PCG, GOT1L1, and metabolites, L-Asparagine and L-Glutamine were all significantly down-regulated in OBM samples as compared to OBO.

#### GSVA pathway enrichment and clinical implications

We performed single sample gene set enrichment analysis (GSVA) using the gene expression of the DE-PCGs involved in the 8 metabolic pathways to obtain their enrichment values as shown in Fig. 3 (rows corresponding to metabolic pathways and columns represent the participant samples). Using an unsupervised clustering technique, the participant data was segregated into 2 clusters (C1 and C2) corresponding to the OBO and OBM strata respectively. The majority of the 8 metabolic pathways were significantly positively enriched in the cluster C2 while they were negatively enriched in the cluster C1. Thus, the enrichment of these 8 metabolic pathways could be used in a predictive framework to potentially distinguish OBM from OBO.

Moreover, as observed in Fig. 3A, we highlighted several of the clinical traits including age, sex, BMI, diet and exercise status, blood glucose levels and medication information. Interestingly, within the OBM C2 cluster, two predominant sub-clusters could be distinguished based on blood glucose and insulin levels; in one OBM sub-cluster there were more subjects with higher blood glucose and insulin levels (potentially diabetes); and at the same time, they reported greater exercise and smoking with higher enrichment values for the 8 metabolic pathways.

Figure 3B provides a comprehensive comparison of the enrichment values between the OBO and OBM group for each of the 8 metabolic pathways of interest and shows pathways that were most significantly differentially

enriched between the OBO and OBM such as amino sugar and nucleotide sugar metabolism, glyoxylate and dicarboxylate metabolism, lysine degradation and starch and sucrose metabolism.

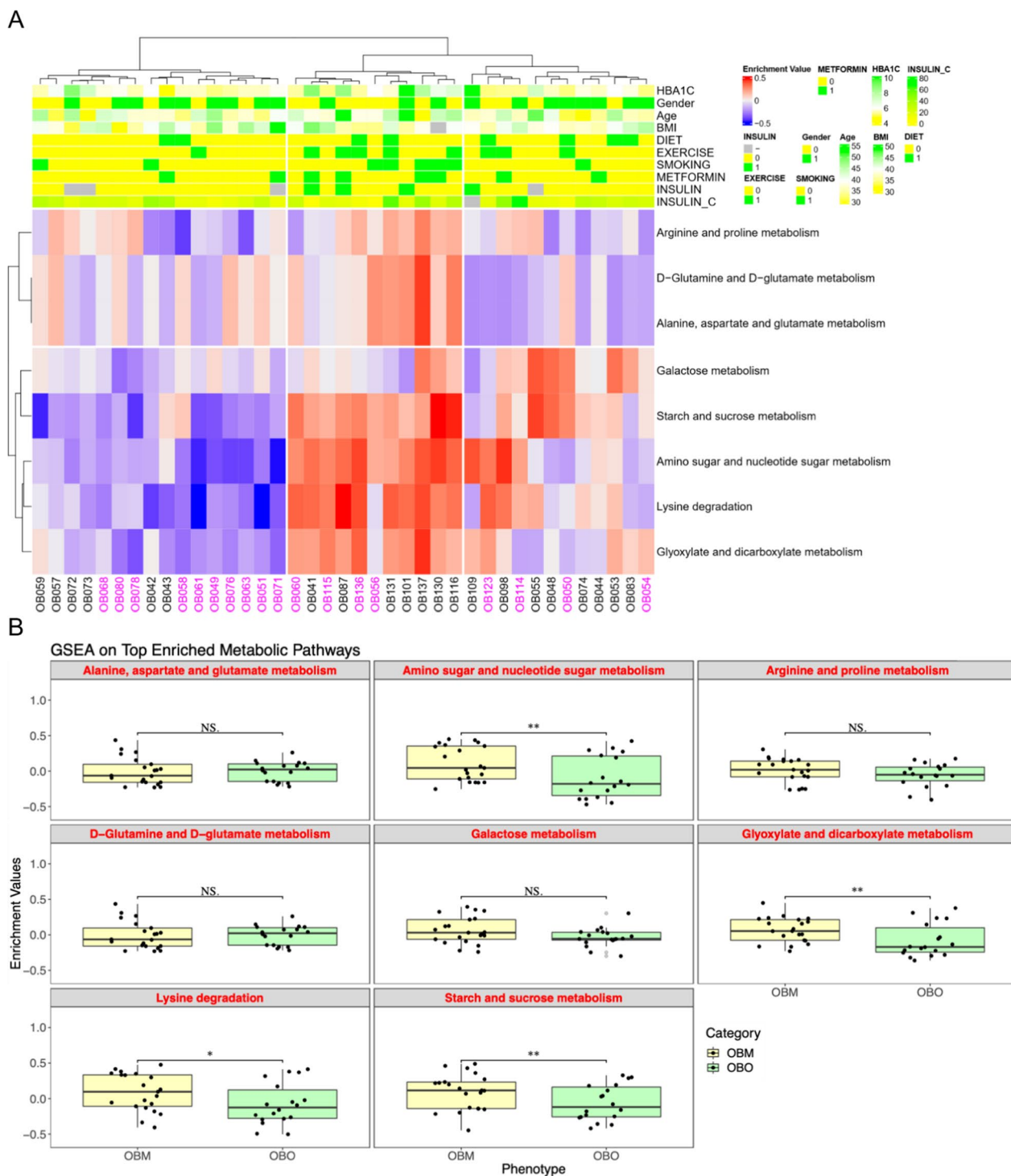
A network-based representation of the myriad regulatory elements involved in the 8 enriched metabolic pathways was illustrated in Fig. 4. We observed two clear clusters of regulatory elements with one cluster comprising alanine, aspartate and glutamate metabolism, D-glutamine and D-glutamate metabolism, glyoxylate and dicarboxylate metabolism, lysine degradation and arginine and proline metabolism pathways along with primarily DE-PCGs and differentially expressed metabolites. There was only one DE-miRNA (miR-636) associated in this cluster. The other cluster comprising galactose metabolism, starch and sugar metabolism and amino sugar and nucleotide sugar metabolism included the remaining 9 DE-miRNAs and DE-PCGs with the inclusion of D-Glucose as the lone differentially expressed metabolite. All metabolic pathways in the second cluster were associated with sugar metabolism and were predominantly enriched in the OBM group (Fig. 3). Interestingly, the PFKM and GPI DE-PCGs associated with the galactose and starch and sucrose metabolism pathways were regulated by maximum number of different DE-miRNAs including 4 DE-miRNAs for PFKM and 3 DE-miRNAs for GPI respectively.

#### Discussion

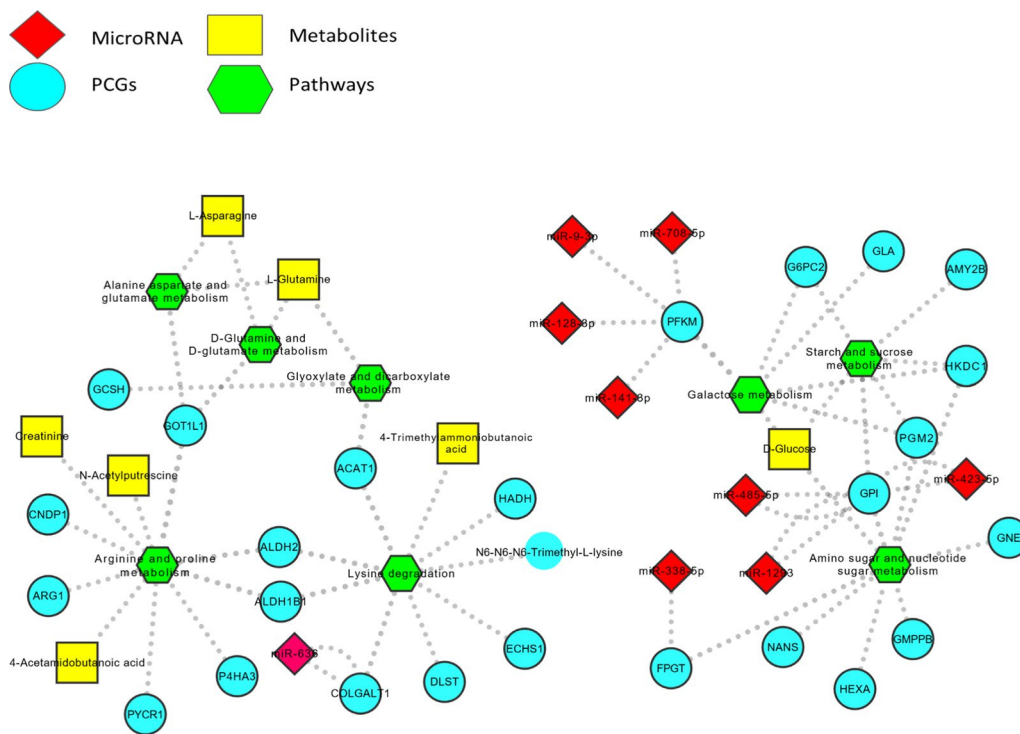
Given that obesity is a complex disorder, a full mechanistic insight requires a coordinated multi-omic approach. Also, deciphering the molecular networks that distinguish obesity with metabolic complications from obesity without metabolic consequences may lead to the identification of critical biomarkers that can be the basis for prognostication and personalized obesity management. In this work, we developed and used an integrative bioinformatics pipeline to identify enriched metabolic pathways by integrating data from transcriptional signatures, microRNAs and circulating metabolites, to obtain a more holistic picture of the differential enrichment of metabolic pathways between OBO and OBM. The identified miRNAs, transcripts, and metabolites are associated with abdominal adiposity, dyslipidemia, insulin resistance, glucose intolerance, and a proinflammatory/pro-thrombotic state.

We identified differentially expressed miRNAs known to influence pathways that regulate metabolism of amino acids (alanine, aspartate, asparagine, glutamate, glutamine, arginine, creatinine and proline), amino acid sugars and nucleotide sugars as well as carbohydrates (glyoxylate, galactose and sucrose). Impaired insulin signaling, a well-known abnormality in the metabolic





**Fig. 3** Clinical implications of enriched metabolic pathways and their correlation with clinical traits. Here the 'purple' colored column labels correspond to OBO samples while the 'black' colored column labels represent the OBM samples in **A**. p-values in **B** were obtained through a t-test procedure and '\*' correspond to p-value  $\leq 0.05$  and p-value  $> 0.01$ , '\*\*' correspond to p-value  $< 0.01$  and p-value  $> 0.001$  and 'NS' stand for 'not-significant' p-value



**Fig. 4** Network representation of the interaction between different regulatory elements associated with the enriched metabolic pathways

syndrome, leads to lower uptake of glucose in skeletal muscles, and an upregulation of hepatic gluconeogenesis resulting in hyperglycemia. Metabolic biomarkers of gluconeogenesis and glycolysis have also been reported as biomarkers of the metabolic syndrome and other cardio-metabolic diseases [37]. Several dysregulated microRNAs in OBM (Additional file 2: Tables S1.1 and S1.2) involved in glucose metabolism have been previously associated with the metabolic syndrome [38–40]. Intriguingly, Mononen et al. found a negative association between hsa-miR-93-3p and HbA1c levels [41]. Elevated insulin secretion reflects increased insulin resistance that is observed in the OBM group [42, 43]. The microRNA, miR-135, specific for the pancreatic islets and suppressing glucose-induced insulin secretion [44] was upregulated in OBM. Wang and colleagues reported that miR-1303 was significantly upregulated in those with type 2 diabetes and was higher in those with complications [45]. On the other hand, the upregulated GNE gene in the OBM group, which is a target of miR-1303, encodes an enzyme that initiates and regulates the biosynthesis of N-acetylneuraminic acid (NeuAc), a precursor of sialic acids. Additionally, the NANS gene, that is higher in OBM, encodes an enzyme called N-Acetylneuraminase Synthase that functions in the biosynthetic pathways of sialic acid. Browning et al. have demonstrated that elevated sialic acid was positively associated with the metabolic syndrome,

independent of BMI in women [46]. Other studies indicate that elevated levels of Neu5Ac during coronary artery disease progression as well as myocardial injury [47, 48]. Furthermore, the GALT gene, which is upregulated in OBM, encodes the enzyme galactose-1-phosphate uridylyltransferase, responsible for converting ingested galactose to glucose. Of note, disturbed galactose metabolism in the elderly and humans with diabetes accelerates cataract formation [49]. Other genes that were found to be higher in OBM encode glycolytic enzymes including Glucose-6-Phosphate Isomerase (GPI), Phosphoglycerate Mutase 1 (PGAM1), Hexosaminidase A, also known as HEXA as well as Phosphoglucomutase 2 (PGM2). Also, COLGALT1, which was down-regulated in OBM, is essential for extracellular matrix reorganization and post-translational collagen glycosylation in the endoplasmic reticulum [50–52]. Although, bi-allelic variants of COLGALT1 have been associated with brain small vessel disease, to date there are no studies that have investigated its role in development of diabetes or cardiovascular complications that result as an outcome of long-standing metabolic syndrome [53]. In addition to insulin resistance, amino acid catabolism is accelerated in diabetes reflected by muscle wasting along with increased uptake of alanine by the liver for gluconeogenesis and the breakdown of branched chain amino acids in muscle [54, 55]. An increasing body of evidence has linked visceral adiposity,

insulin resistance, and diabetes to plasma amino acid levels [56]. Yamaguchi et al. have reported on the usefulness of measuring the plasma free amino acid profiles in evaluating risk for the metabolic syndrome, diabetes, dyslipidemia, and hypertension in a large cohort [57].

Some genes involved in amino acids metabolism that were significantly dysregulated in OBM versus OBO including ARG1 gene encoding arginase I that hydrolyzes arginine to ornithine and urea. Our findings reveal a significant difference in multiple amino acids and their derivatives between OBM and OBO, including L-Glutamine, L-asparagine, D-glutamate and L-Arginine. In this line, sarcosine, an elevated metabolite in the OBM group, is an amino acid that forms as an intermediate in the metabolism of choline in the kidney and liver and is also formed in the hydrolysis of creatine. Notably, sarcosine has not been reported previously in the context of obesity with metabolic complications. Whether sarcosine levels can be used as an early novel marker for developing metabolic syndrome in individuals with obesity but without metabolic complications cannot be determined in our small cohort and thus warrants further investigation. Plasma dimethylglycine, which is an amino acid derivative, has been shown to be independently related to acute myocardial infarction [58]. In addition to other amino acids, lysine degradation is favored to produce Acetyl-CoA during gluconeogenesis in the hepatic mitochondria [59, 60]. Several enzymes are upregulated in hepatic mitochondria of the OBM group to facilitate Acetyl-CoA production (ALDH2, HADH, DSLT, ECHS1, ACAT1, ALDH1B1). Interestingly, miR-636 was identified to promote atherosclerosis development via SP1 and lysine degradation [61]. Altered liver homeostasis has been demonstrated in the metabolic syndrome. We found a higher expression of the ALDH2 gene encoding for aldehyde dehydrogenase 2 which is central to alcohol metabolism. ALDH2 genetic polymorphisms have been also linked with an increased risk of type 2 diabetes in patients with coronary artery disease [62]. We also checked the expression of PCGs in different tissue types to validate our results. Since the samples were taken from the blood serum there exists a high chance that the differentially expressed genes will correspond to higher expression in blood tissue. We searched the PCGs in the GTEx portal (<https://gtexportal.org/>) and observed that these genes were highly expressed across multiple tissue types (Additional file 1: Figure S3).

The identified miRNAs, transcripts, and metabolites that are associated with the 8 enriched metabolic pathways between OBM and OBO groups as listed in Table 2. Lysine degradation happens through formation of saccharopine and the pipercolic acid pathway. The tissue-specific roles of these two pathways are still under investigation [63]. Moreover, lysine degradation is favored to produce

Acetyl-CoA during gluconeogenesis in the hepatic mitochondria [59, 60]. Amino sugar and nucleotide sugar metabolism, galactose metabolism, and starch and sucrose metabolism pathways mainly comprise of sugar metabolism and are associated with metabolic disorders. Arginine and proline metabolism pathway involves the biosynthesis and metabolism of several amino acids including arginine, ornithine, proline, citrulline, and glutamate. Alterations of arginine and proline pathway are associated with many disorders [64].

The strength of our study, in comparison to single omic studies, is that we integrated multi-omic data to deduce pathways differentially altered between OBO and OBM. Our data suggest early changes in key miRNAs and transcripts that have a widespread impact on the metabolism with functional consequences. The clinical significance of the identified network and cause or consequence of the involved elements remain to be validated in larger cohorts, while the biological roles of these elements in the metabolic syndrome can be further investigated for mechanistic insights. The intricate network revealed provides a panel of putative metabolic biomarkers with a comprehensive picture of the underlying biological processes as opposed to single omic studies. A limitation of our study is that the transcriptomic data from whole blood reflects blood cell function whereas the miRNAs and metabolites reflect functioning of all tissues. Despite this limitation, the integration of these data provided meaningful and metabolically relevant pathways that are differentially regulated between OBO and OBM; thus, metabolic pathways in circulating blood cells may reflect the metabolic status of most tissues. Future prospective studies can utilize these biomarkers to identify the predisposition and the sequence of occurrence of hypertension, dysglycemia and/or dyslipidemia in obesity which can accelerate the discovery of new personalized therapeutics.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-023-04074-x>.

**Additional file 1: Figure S1.** Flowchart highlighting different number of miRNAs, PCGs, Metabolites analyzed in our dataset and steps taken to identify key miRNAs, PCGs, Metabolites involved in significantly enriched metabolic pathways between OBM vs OBO. **Figure S2.** Correlation between clinical traits and enriched metabolic pathways. Here each block represents correlation between the enrichment profile of a metabolic pathway and a clinical characteristic based on their values across the 39 participant profiles. Here we only considered those pathways which were significantly enriched and those clinical traits which were highlighted in Fig. 3A. **Figure S3.** Gene expression of 25 identified genes (differentially expressed in the samples) across multiple tissue types in the GTEx portal (<https://gtexportal.org/>). The expression of these genes is higher across multiple tissue types and not biased towards blood tissue from which samples were collected.

**Additional file 2: Table S1:** List of significantly upregulated miRNAs in OBM compared to OBO. Here RQ refers to Relative Quantification measure using the standard formula [24]. An RQ value showcases the fold-change

(FC) of a specific miRNA in two populations. An  $RQ=1$  indicated that a specific miRNA was not differentially expressed in OBM versus OBO samples. Only those miRNAs were considered which were above the limit of quantification (LOQ). **Table S2.** List of significantly downregulated miRNAs in OBM compared to OBO. Here RQ refers to Relative Quantification measure using the standard formula [24]. An RQ value showcases the fold-change (FC) of a specific miRNA in two populations. An  $RQ=1$  indicated that a specific miRNA was not differentially expressed in OBM versus OBO samples. Only those miRNAs were considered which were above the limit of quantification (LOQ). **Table S2.1.** List of metabolites significantly upregulated in OBM compared to OBO. The estimate represents the coefficients of logistic regression used to fit the model. The magnitude of a coefficient indicates the contribution of the metabolite in classifying OBM from OBO. The corresponding p-values of each metabolite are listed in the last column. **Table S2.2.** List of metabolites significantly downregulated in OBM compared to OBO. The estimate represents the coefficients of logistic regression used to fit the model. The magnitude of a coefficient indicates the contribution of the metabolite in classifying OBM from OBO. The corresponding p-values of each metabolite are listed in the last column.

### Funding

Open Access funding provided by the Qatar National Library. Qatar Metabolic Institute (QMI), Hamad Medical Corporation, 16245/16.

### Author details

<sup>1</sup>Qatar Metabolic Institute, Academic Health System, Hamad Medical Corporation, PO Box 3050, Doha, Qatar. <sup>2</sup>Qatar Computational Research Institute (QCRI), Hamad Bin Khalifa University, Doha, Qatar. <sup>3</sup>Department of Immunology, St. Jude Children's Research Hospital, Memphis, USA. <sup>4</sup>College of Medicine, QU Health, Qatar University, Doha, Qatar. <sup>5</sup>National Obesity Treatment Center, Hamad Medical Corporation, Doha, Qatar. <sup>6</sup>Weil Cornell Medicine – Qatar, Doha, Qatar. <sup>7</sup>Biotechnology Research Center, Technology Innovation Institute, P.O. Box 9639, Abu Dhabi, United Arab Emirates.

Received: 17 January 2023 Accepted: 22 March 2023

Published online: 29 March 2023

### References

- Kyle TK, Dhurandhar EJ, Allison DB. Regarding obesity as a disease: evolving policies and their implications. *Endocrinol Metab Clin North Am*. 2016;45:511–20.
- Despres JP, Lemieux I. Abdominal obesity and metabolic syndrome. *Nature*. 2006;444:881–7.
- Heianza Y, Kato K, Kodama S, Ohara N, Suzuki A, Tanaka S, Hanyu O, Sato K, Sone H. Risk of the development of Type 2 diabetes in relation to overall obesity, abdominal obesity and the clustering of metabolic abnormalities in Japanese individuals: does metabolically healthy overweight really exist? The Niigata Wellness Study. *Diabet Med*. 2015;32:665–72.
- McLaughlin T, Abbasi F, Lamendola C, Reaven G. Heterogeneity in the prevalence of risk factors for cardiovascular disease and type 2 diabetes mellitus in obese individuals: effect of differences in insulin sensitivity. *Arch Intern Med*. 2007;167:642–8.
- Meigs JB, Wilson PW, Fox CS, Vasan RS, Nathan DM, Sullivan LM, D'Agostino RB. Body mass index, metabolic syndrome, and risk of type 2 diabetes or cardiovascular disease. *J Clin Endocrinol Metab*. 2006;91:2906–12.
- Rey-Lopez JP, de Rezende LF, Pastor-Valero M, Tess BH. The prevalence of metabolically healthy obesity: a systematic review and critical evaluation of the definitions used. *Obes Rev*. 2014;15:781–90.
- Magkos F. Metabolically healthy obesity: what's in a name? *Am J Clin Nutr*. 2019;110:533–9.
- van Vliet-Ostapchouk JV, Nuotio ML, Slagter SN, Doiron D, Fischer K, Foco L, Gaye A, Gogele M, Heier M, Hiekkalinna T, et al. The prevalence of metabolic syndrome and metabolically healthy obesity in Europe: a collaborative analysis of ten large cohort studies. *BMC Endocr Disord*. 2014;14:9.
- Bluher M. Metabolically Healthy Obesity. *Endocr Rev*. 2020;41:89.
- Caleyachetty R, Thomas GN, Toulis KA, Mohammed N, Gokhale KM, Balachandran K, Nirantharakumar K. Metabolically healthy obese and incident cardiovascular disease events among 35 million men and women. *J Am Coll Cardiol*. 2017;70:1429–37.
- Hinnouho GM, Czernichow S, Dugravot A, Batty GD, Kivimaki M, Singh-Manoux A. Metabolically healthy obesity and risk of mortality: does the definition of metabolic health matter? *Diabetes Care*. 2013;36:2294–300.
- Hinnouho GM, Czernichow S, Dugravot A, Nabi H, Brunner EJ, Kivimaki M, Singh-Manoux A. Metabolically healthy obesity and the risk of cardiovascular disease and type 2 diabetes: the Whitehall II cohort study. *Eur Heart J*. 2015;36:551–9.
- Morkedal B, Vatten LJ, Romundstad PR, Laugsand LE, Janszky I. Risk of myocardial infarction and heart failure among metabolically healthy but obese individuals: HUNT (Nord-Trøndelag Health Study) Norway. *J Am Coll Cardiol*. 2014;63:1071–8.
- Eckel N, Li Y, Kuxhaus O, Stefan N, Hu FB, Schulze MB. Transition from metabolic healthy to unhealthy phenotypes and association with cardiovascular disease risk across BMI categories in 90 257 women (the Nurses' Health Study): 30 year follow-up from a prospective cohort study. *Lancet Diabetes Endocrinol*. 2018;6:714–24.
- Eckel RH, Hokanson JE. The prediction of atherosclerotic cardiovascular disease in type 1 diabetes mellitus: do we just stop Here? *Circulation*. 2016;133:1051–3.
- Hamer M, Bell JA, Sabia S, Batty GD, Kivimaki M. Stability of metabolically healthy obesity over 8 years: the English Longitudinal Study of Ageing. *Eur J Endocrinol*. 2015;173:703–8.
- Vague J. The degree of masculine differentiation of obesities: a factor determining predisposition to diabetes, atherosclerosis, gout, and uric calculous disease. *Am J Clin Nutr*. 1956;4:20–34.
- Stefan N, Kantartzis K, Machann J, Schick F, Thamer C, Rittig K, Balletshofer B, Machicao F, Fritsche A, Haring HU. Identification and characterization of metabolically benign obesity in humans. *Arch Intern Med*. 2008;168:1609–16.
- Stefan N, Schick F, Haring HU. Causes, characteristics, and consequences of metabolically unhealthy normal weight in humans. *Cell Metab*. 2017;26:292–300.
- Richard AJ, White U, Elks CM, Stephens JM. Adipose Tissue: Physiology to Metabolic Dysfunction. In *Endotext*. Feingold KR, Anawalt B, Blackman MR, Boyce A, Chrousos G, Corpas E, de Herder WW, Dhatariya K, Dungan K, Hofland J, eds. South Dartmouth: New York; 2000
- Gautier L, Cope L, Bolstad BM, Irizarry RA. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*. 2004;20:307–15.
- Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. *Bioinformatics*. 2010;26:2363–7.
- JW. M. clariomshumantranscriptcluster.db: Affymetrix clariomshuman annotation data (chip clariomshumantranscriptcluster). . R package version 880 In R package version 880, vol 2017.
- JW M: pd.clariom.d.human: Platform Design Info for Affymetrix Clariom\_D\_Human. R package version 3141; 2016.
- JW M: Affycoretools: Functions useful for those doing repetitive analyses with Affymetrix GeneChips. R package version 1640; 2008.
- Wei R, Wang J, Su M, Jia E, Chen S, Chen T, Ni Y. Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. *Sci Rep*. 2018;8:663.
- Pearson K. On lines and planes of closest fit to systems of points in space. *The London Edinburgh Dublin Philosop Magazine J Sci*. 1901;2:559–72.
- H Hotelling: Analysis of a complex of statistical variables into principal components. *J Educ Psychol*, 24, 417 1933.
- R R: **A Language and Environment for Statistical Computing. R Foundation for Statistical Computing**. 2020.
- Tokar T, Pastrello C, Rossos AEM, Abovsky M, Hauschild AC, Tsay M, Lu R, Jurisica I: **mirDIP 4.1-integrative database of human microRNA target predictions**. *Nucleic Acids Res* 2018, **46**:D360–D370.
- Shirdel EA, Xie W, Mak TW, Jurisica I. NAViGaTing the microneome—using multiple microRNA prediction databases to identify signalling pathway-associated microRNAs. *PLoS ONE*. 2011;6: e17429.
- Mir FA, Mall R, Iskandarani A, Ullah E, Samra TA, Cyprian F, Parray A, Alkaseem M, Abdalhakam I, Farooq F, Abou-Samra AB. Characteristic MicroRNAs Linked to Dysregulated Metabolic Pathways in Qatari Adult



- Subjects With Obesity and Metabolic Syndrome. *Front Endocrinol (Lausanne)*. 2022;13: 937089.
33. Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
  34. Mall R, Saad M, Roelands J, Rinchai D, Kunji K, Almeer H, Hendrickx W, F MM, Ceccarelli M, Bedognetti D: Network-based identification of key master regulators associated with an immune-silent cancer phenotype. *Brief Bioinform* 2021, 22.
  35. Sherif S, Mall R, Almeer H, Naik A, Al Homaid A, Thomas R, Roelands J, Narayanan S, Mohamed MG, Bedri S, et al. Immune-related 3-lncRNA signature with prognostic connotation in a multi-cancer setting. *J Transl Med*. 2022;20:442.
  36. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43: e47.
  37. Brandao-Lima PN, Carvalho GB, Payolla TB, Sarti FM, Rogero MM. Circulating microRNA Related to Cardiometabolic Risk Factors for Metabolic Syndrome: A Systematic Review. *Metabolites*. 2022;12:89.
  38. Mir FA, Ullah E, Mall R, Iskandarani A, Samra TA, Cyprian F, Parray A, Alkasem M, Abdalhakam I, Farooq F, Abou-Samra AB. Dysregulated Metabolic Pathways in Subjects with Obesity and Metabolic Syndrome. *Int J Mol Sci*. 2022;23:9.
  39. Gulyaeva LF, Kushlinskiy NE. Regulatory mechanisms of microRNA expression. *J Transl Med*. 2016;14:143.
  40. McGarrah RW, Crown SB, Zhang GF, Shah SH, Newgard CB. Cardiovascular Metabolomics. *Circ Res*. 2018;122:1238–58.
  41. Mononen N, Lyytikäinen LP, Seppala I, Mishra PP, Juonala M, Waldenberger M, Klopp N, Illig T, Leiviska J, Loo BM, et al. Whole blood microRNA levels associate with glycemic status and correlate with target mRNAs in pathways important to type 2 diabetes. *Sci Rep*. 2019;9:8887.
  42. Cavaghan MK, Ehrmann DA, Polonsky KS. Interactions between insulin resistance and insulin secretion in the development of glucose intolerance. *J Clin Invest*. 2000;106:329–33.
  43. Reaven GM. Insulin resistance/compensatory hyperinsulinemia, essential hypertension, and cardiovascular disease. *J Clin Endocrinol Metab*. 2003;88:2399–403.
  44. Poy MN, Eliasson L, Krutzfeldt J, Kuwajima S, Ma X, Macdonald PE, Pfeffer S, Tuschl T, Rajewsky N, Rorsman P, Stoffel M. A pancreatic islet-specific microRNA regulates insulin secretion. *Nature*. 2004;432:226–30.
  45. Wang C, Wan S, Yang T, Niu D, Zhang A, Yang C, Cai J, Wu J, Song J, Zhang CY, et al. Increased serum microRNAs are closely associated with the presence of microvascular complications in type 2 diabetes mellitus. *Sci Rep*. 2016;6:20032.
  46. Browning LM, Jebb SA, Mishra GD, Cooke JH, O'Connell MA, Crook MA, Krebs JD. Elevated sialic acid, but not CRP, predicts features of the metabolic syndrome independently of BMI in women. *Int J Obes Relat Metab Disord*. 2004;28:1004–10.
  47. Zhang L, Wei TT, Li Y, Li J, Fan Y, Huang FQ, Cai YY, Ma G, Liu JF, Chen QQ, et al. Functional metabolomics characterizes a key role for n-acetylneuraminic acid in coronary artery diseases. *Circulation*. 2018;137:1374–90.
  48. Li MN, Qian SH, Yao ZY, Ming SP, Shi XJ, Kang PF, Zhang NR, Wang XJ, Gao DS, Gao Q, et al. Correlation of serum N-Acetylneuraminic acid with the risk and prognosis of acute coronary syndrome: a prospective cohort study. *BMC Cardiovasc Disord*. 2020;20:404.
  49. Birlouez-Aragon I, Ravelontseho L, Villate-Cathelineau B, Cathelineau G, Abitbol G. Disturbed galactose metabolism in elderly and diabetic humans is associated with cataract formation. *J Nutr*. 1993;123:1370–6.
  50. Baumann S, Hennet T. Collagen accumulation in osteosarcoma cells lacking GLT25D1 collagen galactosyltransferase. *J Biol Chem*. 2016;291:18514–24.
  51. Schegg B, Hulsmeier AJ, Rutschmann C, Maag C, Hennet T. Core glycosylation of collagen is initiated by two beta(1-O)galactosyltransferases. *Mol Cell Biol*. 2009;29:943–52.
  52. Liefhebber JM, Punt S, Spaan WJ, van Leeuwen HC. The human collagen beta(1-O)galactosyltransferase, GLT25D1, is a soluble endoplasmic reticulum localized protein. *BMC Cell Biol*. 2010;11:33.
  53. Teunissen MWA, Kamsteeg EJ, Sallevelt S, Pennings M, Bauer NJC, Vermeulen RJ, Nicolai J. Biallelic Variants in the COLGALT1 gene causes severe congenital porencephaly: a case report. *Neurol Genet*. 2021;7: e564.
  54. Bellary S, Kyrou I, Brown JE, Bailey CJ. Type 2 diabetes mellitus in older adults: clinical considerations and management. *Nat Rev Endocrinol*. 2021;17:534–48.
  55. Wahren J, Felig P, Cerasi E, Luft R. Splanchnic and peripheral glucose and amino acid metabolism in diabetes mellitus. *J Clin Invest*. 1972;51:1870–8.
  56. Hosseinkhani S, Arjmand B, Dilmaghani-Marand A, Mohammadi Fateh S, Dehghanbanadaki H, Najjar N, Alavi-Moghadam S, Ghodssi-Ghassemabadi R, Nasli-Esfahani E, Farzadfar F, et al. Targeted metabolomics analysis of amino acids and acylcarnitines as risk markers for diabetes by LC-MS/MS technique. *Sci Rep*. 2022;12:8418.
  57. Yamaguchi N, Mahbub MH, Takahashi H, Hase R, Ishimaru Y, Sunagawa H, Amano H, Kobayashi-Miura M, Kanda H, Fujita Y, et al. Plasma free amino acid profiles evaluate risk of metabolic syndrome, diabetes, dyslipidemia, and hypertension in a large Asian population. *Environ Health Prev Med*. 2017;22:35.
  58. Svingen GF, Ueland PM, Pedersen EK, Schartum-Hansen H, Seifert R, Ebbing M, Loland KH, Tell GS, Nygard O. Plasma dimethylglycine and risk of incident acute myocardial infarction in patients with stable angina pectoris. *Arterioscler Thromb Vasc Biol*. 2013;33:2041–8.
  59. Information NCFB: PubChem Pathway Summary for Pathway SMP0000037, Lysine Degradation. PathBank 2022.
  60. Cornell NW, Lund P, Krebs HA. The effect of lysine on gluconeogenesis from lactate in rat hepatocytes. *Biochem J*. 1974;142:327–37.
  61. Zhang LL. CircRNA-PTPRA promoted the progression of atherosclerosis through sponging with miR-636 and upregulating the transcription factor SP1. *Eur Rev Med Pharmacol Sci*. 2020;24:12437–49.
  62. Xu F, Chen Y, Lv R, Zhang H, Tian H, Bian Y, Feng J, Sun Y, Li R, Wang R, Zhang Y. ALDH2 genetic polymorphism and the risk of type II diabetes mellitus in CAD patients. *Hypertens Res*. 2010;33:49–55.
  63. Leandro J, Houten SM. The lysine degradation pathway: Subcellular compartmentalization and enzyme deficiencies. *Mol Genet Metab*. 2020;131:14–22.
  64. Patin F, Corcia P, Vourc'h P, Nadal-Desbarats L, Baranek T, Goossens JF, Marouillat S, Dessein AF, Descat A, Madji Hounoum B, et al. Omics to explore amyotrophic lateral sclerosis evolution: the central role of arginine and proline metabolism. *Mol Neurobiol*. 2017;54:5361–74.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

