Nutrition Journal

RESEARCH

Open Access

Metabolomics and nutrient intake reveal metabolite–nutrient interactions in metabolic syndrome: insights from the Korean Genome and Epidemiology Study

Minyeong Kim[1†], Suyeon Lee[1†], Junguk Hur[2] and Dayeon Shin[1*]

Abstract

Background Despite advances in metabolomics, the complex relationship between metabolites and nutrient intake in metabolic syndrome (MetS) remains poorly understood in the Korean population.

Objective This study aimed to characterize the metabolomic profiles and nutrient intake associated with MetS and to examine their relationships in the Ansan-Ansung cohort of the Korean Genome and Epidemiology Study (KoGES).

Methods Data from 2,306 middle-aged adults (1,109 men and 1,197 women) in the KoGES Ansan-Ansung cohort were analyzed. Plasma metabolites were measured using liquid chromatography-mass spectrometry, identifying 135 metabolites. Nutrient intake was assessed using a validated semi-quantitative food frequency questionnaire covering 23 nutrients. MetS-associated metabolites and nutrients were identified using the Wilcoxon rank-sum test, logistic regression, partial least squares-discriminant analysis, and group least absolute shrinkage and selection operator analysis. Pathway enrichment analysis identified key metabolic pathways, and fixed-effects models were applied to assess metabolite–nutrient relationships based on MetS status.

Results Eleven metabolites, including hexose (FC = 0.95, $P = 7.04 \times 10^{-54}$), alanine, and branched-chain amino acids, and three nutrients including fat, retinol, and cholesterol, were significantly associated with MetS (FC range = 0.87–0.93; all $P < 0.05$). Pathway analysis highlighted disruptions in arginine biosynthesis and arginine–proline metabolism. The MetS group exhibited six unique metabolite–nutrient pairs that were not observed in the non-MetS group, including 'isoleucine–fat,' 'isoleucine–P,' 'proline–fat,' 'leucine–fat,' 'leucine–P,' and 'valerylcarnitine–niacin.' Notably, dysregulated metabolism of branched-chain amino acids, such as isoleucine and leucine, has been implicated in oxidative stress. Importantly, the stochastic gradient descent classifier achieved the best predictive performance among the eight machine learning models (area under the curve, AUC = 0.84), highlighting the robustness of

†Minyeong Kim and Suyeon Lee contributed equally as co-first authors.

*Correspondence:
Dayeon Shin
dyshin@inha.ac.kr

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

classification based on metabolite data. However, the absence of external validation limits the generalizability of these findings.

Conclusions This comprehensive metabolomic analysis of the KoGES Ansan-Ansung cohort revealed distinct metabolic profiles and nutrient intake patterns associated with MetS, highlighting altered metabolite–nutrient relationships and disrupted metabolic pathways. These findings provide new insights into potential associations between metabolic phenotypes and dietary intake, which may help inform individualized dietary approaches related to MetS, such as branched-chain amino acids-restricted diets (valine, isoleucine, leucine), reduced intake of hexose-rich carbohydrates, and modulation of niacin-rich protein sources according to individual metabolic profiles.

Keywords Metabolic syndrome, Metabolic biomarkers, Nutrients, Machine learning algorithms, Personalized nutrition, Precision nutrition

Introduction

The prevention and management of metabolic syndrome (MetS) pose increasing challenges to public health due to its rising global prevalence [1]. MetS is a cluster of metabolic abnormalities—including abdominal obesity, elevated blood pressure, high fasting glucose levels, hypertriglyceridemia, and low high-density lipoprotein (HDL) cholesterol—that collectively increase the risk of chronic conditions such as cardiovascular disease, type 2 diabetes, and chronic kidney disease [2]. In Korea, the prevalence of MetS reached 24.9% in 2021, with a concurrent increase in diabetic kidney disease from 12.4% in 2015 to 25.4% in 2021 [3, 4]. This trend highlights the urgent need for effective strategies to detect and manage MetS, especially in high-risk populations.

Recent studies have highlighted the relevance of metabolomics in uncovering the pathophysiology of MetS. Specific metabolites, such as phosphatidylcholines (PCs) and sphingolipids, are positively associated with lipid abnormalities, while acylcarnitines and amino acids are inversely correlated with obesity markers [5]. Notably, lysoPC a C18:2 has emerged as a potential biomarker associated with all five MetS components and linked to impaired glucose metabolism and increased cardiovascular risk [6–8]. These findings suggest that metabolomic profiling could enhance early identification of individuals at risk of MetS.

Lifestyle modifications, particularly dietary changes, are recommended as the first-line treatment strategy for managing MetS. Despite extensive research on dietary influences, establishing causal links between nutrient intake and MetS remains challenging due to inconsistent findings from clinical trials. For example, a literature review on single nutrients and MetS found that intervention trials showed that taking vitamin D and calcium supplements together was effective in patients with glucose intolerance, whereas randomized controlled trials reported inconsistent clinical improvements with vitamin D supplementation [9]. Although specific dietary interventions, such as improving food quality [10] or altering macronutrient distribution [11], have shown promise in managing MetS and its components, optimal nutritional strategies for MetS prevention are still unclear.

Current healthcare primarily focuses on treating patients after disease onset, which is both costly and less effective for managing chronic and late-stage diseases. Thus, as traditional healthcare approaches often fall short in addressing individual variability, there is growing emphasis on personalized strategies. Metabolomics offers a powerful tool for capturing molecular signatures reflective of diet and metabolic health, providing opportunities for precision nutrition and individualized risk prediction [12, 13]. Precision nutrition involves tailoring dietary interventions based on metabolic responses to specific nutrients or dietary patterns. As noted in previous studies, metabolites associated with strict adherence to Mediterranean and Nordic diets—such as carnitines, sphingomyelins, and phospholipids—are derived from endogenous metabolism and have been linked to cardiometabolic risk markers including blood pressure and insulin resistance [13]. These findings suggest the potential of metabolomics to inform tailored dietary strategies for managing complex conditions such as MetS. Furthermore, beyond informing interventions, metabolomic profiling can enhance the accuracy of dietary assessments themselves. Particularly in nutritional epidemiology, integrating metabolomic biomarkers can improve assessments beyond self-reporting methods [14]. Considering the rising prevalence of MetS and its associated health risks, there is an urgent need for effective diagnostic and prognostic tools. Therefore, developing a prediction model based on serum metabolomic profiles represents a promising avenue for enhancing early detection and personalized risk stratification of MetS, ultimately improving clinical outcomes through tailored interventions [15].

Although previous studies have examined the associations between metabolic profiles and individual components of MetS in Korean populations [16–20], comprehensive analyses directly linking MetS status to metabolite–nutrient interactions remain limited. Moreover, to our knowledge, no study has applied a machine learning approach using metabolite profiles to predict

MetS within a Korean cohort. In addition, although some studies have examined the associations among MetS, metabolic profiles, and dietary intake, most of these investigations have primarily focused on non-Korean populations [21–23].

In view of the complexity of MetS as a multifactorial condition, integrating insights from multiple biological domains through advanced omics technologies is crucial for a deeper understanding of its underlying mechanisms. We hypothesize that MetS alters metabolite-nutrient interactions, with distinct profiles in Korean adults detectable via metabolomics. To address this gap, the present study investigates the associations between serum metabolite levels, nutrient intake, and MetS in middle-aged Korean adults. Additionally, it aims to develop and validate a prediction model for MetS based on metabolomic profiles, providing a robust tool for early identification and intervention strategies tailored to the Korean population.

Methods

Study design and participants

This study utilized data from the Korean Genome and Epidemiology Study (KoGES), conducted by the Korea National Institute of Health, which collects comprehensive information on demographic variables, health status, medical history, biochemical variables, genotypes, and metabolites [24]. The baseline survey was conducted between 2001 and 2002, with follow-up surveys conducted biennially over a span of 20 years. Although KoGES is a prospective cohort study with repeated follow-up surveys, this study utilized data from a single time point (2005–2006) because plasma metabolite data were only available for the specific period. Therefore, this study was designed as a cross-sectional analysis and was limited to the 7,515 individuals who participated in 2005–2006. The response rate at the second follow-up was 76.0%, and no notable differences in baseline characteristics were observed between respondents and non-respondents (i.e., those who participated in the baseline survey only) [24].

A flowchart of the study population is presented in Fig. 1. From the initial cohort of 7,515 participants, exclusions were made based on the following criteria: absence of plasma metabolite data ($n = 5{,}085$), missing dietary information ($n = 14$), incomplete MetS components data ($n = 101$), and missing demographic details, including sex, age, and body mass index (BMI; $n = 9$). Consequently, a total of 2,306 participants, comprising 1,109 men and 1,197 women, were included in the final analysis. This study was approved by the Institutional Review Board of Inha University on March 20, 2024 (protocol number: 240307-1A). Patients and members of the public were not involved in the design, conduct, reporting, or dissemination of this study.

Definition of MetS

The criteria for MetS diagnosis were based on the National Cholesterol Education Program Adult Treatment Panel III (NCEP ATP III) guidelines [25], with waist circumference cutoffs established by the Korean Society for the Study of Obesity [26]. The NCEP ATP III guidelines are still widely used in epidemiological studies and have been applied to diverse populations, including Korean cohorts. Because the international standard cutoffs for waist circumference have been reported to be inappropriate for Asians, including Koreans [27], ethnicity-specific values were applied to accurately reflect the metabolic risk associated with abdominal obesity in the Korean population. MetS was defined as the presence of at least three of the following five risk factors: (1) waist circumference ≥ 90 cm for men and ≥ 85 cm for women, (2) triglyceride levels ≥ 150 mg/dL, (3) HDL cholesterol < 40 mg/dL for men and < 50 mg/dL for women, (4) systolic blood pressure ≥ 130 mmHg, diastolic blood pressure ≥ 85 mmHg, history of hypertension, use of antihypertensive medication, or ongoing hypertension treatment, and (5) fasting blood glucose > 100 mg/dL, a diabetes diagnosis, ongoing diabetes treatment, insulin therapy, or oral diabetes medication.

Metabolite measurement

Targeted metabolomics was performed using electrospray ionization liquid chromatography–mass spectrometry (ESI-LC/MS) and tandem mass spectrometry (MS/MS) with the AbsoluteIDQ™ p180 kit (BIOCRATES Life Sciences AG, Innsbruck, Austria). This kit enables the quantification of 40 acylcarnitines, 21 amino acids, 19 biogenic amines, 1 hexose, 90 glycerophospholipids, and 15 sphingolipids. All measurements were carried out strictly following the manufacturer's protocol (Manual UM-P180). The assay procedures of the AbsoluteIDQ™ p180 kit, as well as the metabolite nomenclature, have been described in detail previously [16]. Briefly, 10 µL of serum was aliquoted onto a 96-well plate with a filter, followed by metabolite extraction. The extracts were analyzed by flow injection analysis–tandem mass spectrometry for acylcarnitines, hexose, glycerophospholipids, and sphingolipids in both positive (acylcarnitines, glycerophospholipids, and sphingolipids) and negative (hexose) ion modes. Amino acids and biogenic amines were quantified using liquid chromatography–tandem mass spectrometry in positive ion mode. Metabolite concentrations were automatically measured using the MetVal™ software package (BIOCRATES Life Sciences AG), with internal standards serving as references for concentration calculations, and were expressed in µM. The quality

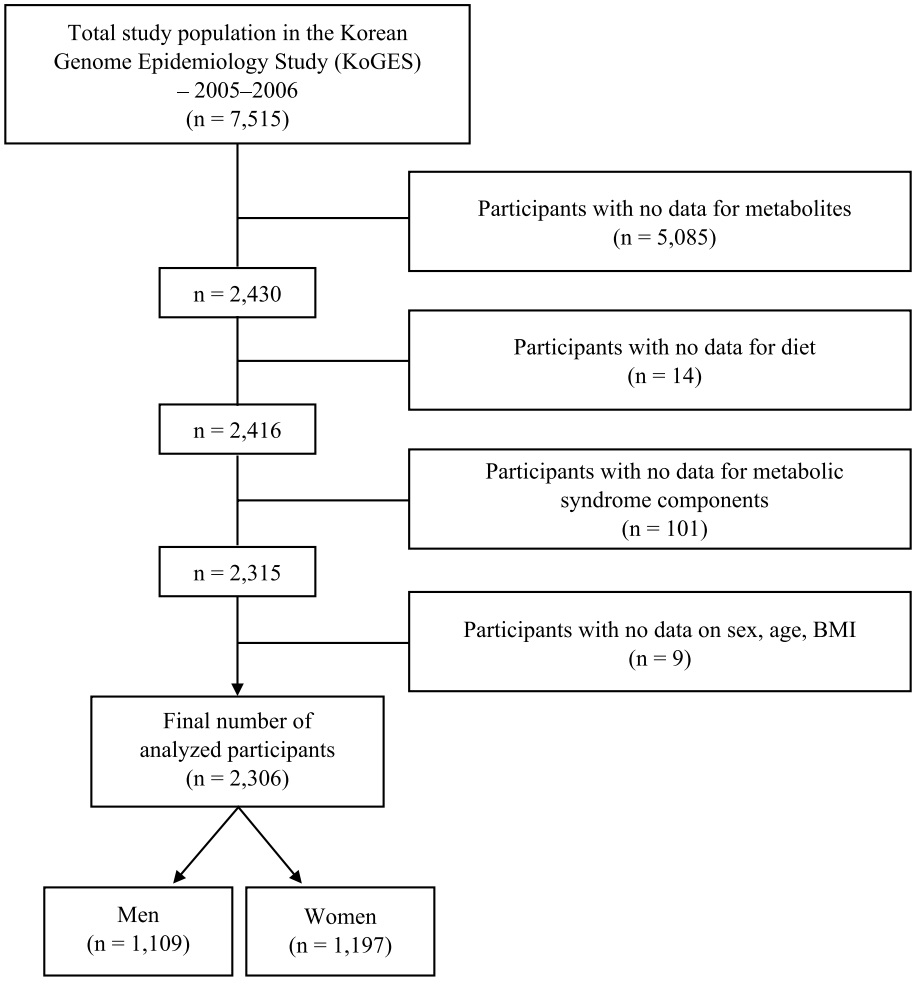

Fig. 1 Flowchart illustrating the participant selection process

of metabolite data was assessed based on the following criteria: (1) coefficient of variation for each metabolite in the reference standards < 25%, (2) at least half of the analyzed metabolite concentrations in the reference standards exceeding the limit of detection, and (3) at least half of the analyzed metabolite concentrations in the experimental samples exceeding the limit of detection [19, 20]. Quality control was performed by the KoGES team prior to data release, and only metabolites meeting predefined analytical standards were provided to researchers. Following the exclusion of 51 metabolites that did not meet quality control standards, a total of 135 metabolites were included in the final analysis. These comprised 13 acylcarnitines, 21 amino acids, 10 biogenic amines, 1 hexose, 78 glycerophospholipids, and 12 sphingolipids.

Dietary assessment

Dietary information was collected at baseline using a semi-quantitative food frequency questionnaire (FFQ), which included 106 food items and reference amounts commonly consumed by Koreans to assess the annual average intake frequency and quantity [28]. The intake frequency of each food item was recorded at nine levels: almost never, once per month, two to three times per month, one to two times per week, three to four times per week, five to six times per week, and once, twice, or three times per day. In addition, intake amounts were classified as large, medium, or small. Daily nutrient intake was estimated using the Korean Food Composition Table [29], which evaluated total energy intake in calories along with 22 nutrients—protein, fat, carbohydrate, calcium, phosphorus, iron (Fe), potassium, vitamin A, sodium, vitamin B1, vitamin B2, niacin, vitamin C, zinc, vitamin B6, folate, retinol, carotene, ash, fiber, vitamin E, and cholesterol. In this study, we included all 23 nutrients systematically assessed in the KoGES dataset, including total energy and the 22 macro- and micronutrients. These were selected not based on a hypothesis-driven approach but to comprehensively capture the full nutritional information available from the validated FFQ and food composition table.

Statistical analyses

Descriptive analysis

Descriptive summaries of the demographic and clinical characteristics were generated for participants with and without MetS. Differences in continuous variables between the MetS and non-MetS groups were assessed using the Wilcoxon rank-sum test, with results expressed as means and standard deviations. Categorical variables were analyzed using the chi-square test and presented as frequencies and percentages.

Differential metabolite and nutrient identification

The intake of each metabolite and nutrient was log-transformed and standardized to a z-score. Four complementary methods—the Wilcoxon rank-sum test, partial least squares-discriminant analysis (PLS-DA), logistic regression, and group least absolute shrinkage and selection operator (LASSO)—were used to identify metabolites associated with MetS. Each method provides a unique perspective on the data and reveals distinct patterns of association. For nutrient intake, only the Wilcoxon rank-sum test and PLS-DA were applied, whereas all four methods were used for metabolite analysis. The Wilcoxon rank-sum test was used to evaluate univariate differences between the MetS and non-MetS groups using the R package *stats*. The PLS-DA, implemented using the R package *mixOmics*, was used to identify metabolites that best distinguished the two groups based on the first latent variable, with importance assessed by the variable importance in projection (VIP) score (VIP > 1). Logistic regression and group LASSO were applied exclusively to metabolites, with logistic regression models adjusted for age, sex, and BMI. Covariates were selected based on their established roles as biological determinants of MetS. Age and sex are known to influence hormonal regulation, fat distribution, and insulin sensitivity, which are key factors in MetS pathophysiology [30]. Additionally, indicators of adiposity such as BMI have been consistently associated with MetS, insulin resistance, and metabolic profiles [5]. *P*-values were corrected for multiple comparisons using the Benjamini–Hochberg method. Group LASSO, implemented using the R package *glmnet*, further explored sub-pathway differences by mapping the metabolites to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways [31] and the Human Metabolome Database (HMDB, http://www.hmdb.ca) [32]. Group LASSO was employed to explore metabolite sub-pathway differences by matching metabolites to KEGG and HMDB pathways, resulting in the inclusion of 41 metabolites in the analysis. The analysis was conducted using matched metabolites as predictors (X) and MetS status as the outcome variable (Y), with metabolites grouped according to KEGG and HMDB pathways. A binary logistic regression model with group penalties was applied using the R package *gglasso*, and significant metabolites associated with MetS were identified based on the model coefficients.

To compare the results across the different methods (Wilcoxon, logistic regression, PLS-DA, and group LASSO), the overlapping metabolites and nutrients identified by each model were visualized using Venn diagrams in the R *VennDetail* package. Each method was chosen for its distinct analytical strength: the Wilcoxon rank-sum test identified unadjusted group differences in metabolite and nutrient levels; logistic regression assessed adjusted associations with MetS while controlling for age, sex, and BMI; PLS-DA was employed to identify variables contributing to group separation; and group LASSO was applied to select entire sub-pathways, enabling pathway-level interpretation by accounting for within-pathway correlation structures. This approach allowed for the identification of variables that were consistently important across multiple methods, as well as those unique to specific models. The use of Venn diagrams facilitated the interpretation of intersections and differences between the methods, providing insights into their respective contributions to identifying key variables.

Pathway enrichment analysis

Pathway enrichment analysis was conducted using the KEGG database and MetaboAnalyst 6.0 (http://www.metaboanalyst.ca/). Metabolites were mapped to KEGG pathways using the R package *KEGGREST* to validate annotations and confirm their biological relevance to metabolic pathways, resulting in the inclusion of 41 annotated metabolites, comprising five acylcarnitines, 21 amino acids, 10 biogenic amines, and one hexose. Enrichment analysis was then performed on the significant metabolites identified by each method—Wilcoxon, logistic regression, PLS-DA, and group LASSO—using MetaboAnalyst 6.0 to determine their pathway associations.

Correlation analysis

Spearman's rank correlation was used to assess the associations between metabolites and nutrients, with the results visualized using the R package *corrplot*. Correlation networks were constructed to explore the relationships among metabolites, nutrients, and MetS components, providing insights into their interconnections and potential influences.

Machine learning-based classification models

To comprehensively evaluate modeling approaches suitable for metabolite-based prediction of MetS, we selected eight machine learning (ML) algorithms based on their demonstrated suitability for metabolomics data, characterized by high dimensionality, multicollinearity,

and complex biochemical interactions. Logistic regression (LR) and stochastic gradient descent (SGD) were included for their interpretability and scalability. These linear models are widely used in metabolomics due to their ability to handle large numbers of predictors while maintaining transparent model structures. Regularization allows LR and SGD to identify a parsimonious set of predictive metabolites, which facilitates biological interpretation and clinical translation [33]. The support vector machine (SVM) was selected for its strong performance in classifying nonlinear relationships between metabolic features and disease states. Kernel-based SVMs have been shown to accurately discriminate metabolic profiles associated with MetS and related phenotypes by projecting complex patterns into higher-dimensional space [34]. Random forest (RF) and AdaBoost (ADB) are ensemble learning algorithms that are particularly effective in metabolomics, where predictor variables are correlated and biological signals are embedded in noisy, heterogeneous data. RF identifies key metabolites through internal feature ranking, aiding biomarker discovery, while ADB enhances classification performance by focusing on samples with borderline metabolic signatures [34, 35]. K-nearest neighbors (KNN) was included as a nonparametric model that classifies individuals based on the overall similarity of their metabolic profiles. KNN has demonstrated competitive performance in metabolomics-based MetS classification tasks, especially when using curated panels of metabolites with strong biological signals [15]. Naïve Bayes (NB) is a probabilistic classification algorithm based on Bayes' Theorem and the assumption of conditional independence among features. It offers fast computation and is known to perform relatively well, particularly when the training dataset is limited. In metabolomics, NB can serve as a lightweight baseline classifier for exploratory analyses or preliminary screening, especially when computational efficiency and interpretability are prioritized [36]. Deep neural network (DNN) is suited to metabolomics data because it can capture complex, nonlinear interactions among metabolites through multilayer architectures. Their ability to learn hierarchical representations allows them to model intricate biological relationships embedded in omics data, which has led to superior classification performance in recent metabolomics and multi-omics disease classification studies [37]. All models were implemented in R, version 4.4.1, utilizing established packages.

Data preprocessing and feature selection

A total of 186 metabolites were initially quantified using the AbsoluteIDQ™ p180 kit. Prior to model development, metabolites were filtered based on predefined quality control criteria: (1) a coefficient of variation (CV) < 25% in reference standards, (2) ≥ 50% of reference standard concentrations exceeding the limit of detection, and (3) ≥ 50% of experimental sample concentrations exceeding the limit of detection. Based on these thresholds, 51 metabolites were excluded, and the remaining 135 metabolites were retained for analysis. All predictor variables were standardized using z-score normalization prior to analysis. The outcome variable was MetS status, defined according to the NCEP ATP III and the Korean Society for the Study of Obesity criteria. It was encoded as a binary variable (1 = MetS, 0 = non-MetS).

Data partitioning and internal validation

The dataset was randomly partitioned into training (70%) and test (30%) subsets using stratified sampling to preserve the class distribution, which was implemented via the *createDataPartition()* function from the *caret* package. Model development and internal validation were conducted exclusively on the training set, while model performance was evaluated on the independent test set. External validation could not be performed due to the unavailability of an independent dataset. This limitation has been acknowledged and addressed in the discussion section.

Model implementation and configuration

All eight models were trained using consistent procedures and default or minimal-tuned hyperparameters as implemented in their respective R packages. SVM was implemented using the *e1071* package with a radial basis function kernel and probability estimates enabled. RF, developed via the *randomForest* package, was trained with 500 trees while retaining default settings for all other parameters. KNN was configured using the *kknn* package with $k = 5$ and a rectangular kernel. LR was fitted using the *glm* function with a binomial logistic regression model that included all available predictors. DNN, implemented via the *neuralnet* package, utilized two hidden layers with 5 and 10 neurons respectively, a logistic activation function, a training threshold of 0.25, and a maximum of 1×10^6 iterations. NB, also implemented using *e1071*, was applied with default settings. ADB was constructed using the *adabag* package with 100 boosting iterations (mfinal = 100) and boos set to TRUE. Finally, SGD was implemented as Lasso-penalized logistic regression using the *glmnet* package through *caret*, with the hyperparameters set to $\alpha = 1$ and $\lambda = 0.01$. No formal hyperparameter optimization procedures were applied beyond manual specification of core model parameters for convergence and interpretability. No automated grid search or cross-validation-based tuning was conducted. Key parameters were set based on established defaults and literature to prioritize interpretability and convergence.

Model evaluation

Model performance was assessed on the independent hold-out test set using the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and overall accuracy (non-error rate). AUC confidence intervals were calculated using DeLong's method, and ROC curves were generated with the *pROC* and *ggplot2* packages. A fixed classification threshold of 0.5 was applied consistently across all models [38].

Reporting quality assessment

To evaluate reporting quality, we utilized the TRIPOD + AI statement, which provides a comprehensive checklist of items that should be reported in studies describing the development and/or validation of predictive AI models [39]. The checklist encompasses all major sections of a manuscript, including the title, abstract, introduction, methods, results, and discussion. In addition to these conventional sections, the TRIPOD + AI statement places particular emphasis on open science practices, patient and public involvement, and equity considerations. Each item's adherence was evaluated as "Yes," "Partial," "No," or "Not Applicable" and was accompanied by the corresponding page number. The completed checklist is available in Supplementary Table 2.

Fixed-effects model for effects of nutrition

A fixed-effects model was applied to examine the interactions between metabolites and nutrients according to MetS status. First, linear regression analyses identified metabolites significantly associated with MetS ($P < 0.05$). Second, relationships between these significant metabolites and nutrient intake were evaluated using the lm() function from the stats package in R, with fixed effects including MetS status and nutrient intake. The fixed-effects models were designed to examine nutrient–metabolite associations stratified by MetS status, without additional covariate adjustment, to isolate interaction patterns across strata. To focus on robust associations, we selected 11 metabolites consistently associated with MetS using four statistical approaches: the Wilcoxon test, logistic regression, PLS-DA, and cluster LASSO. In addition, metabolite-nutrient pairs significantly associated with nutrient intake were included. Finally, the dataset was stratified by MetS status to further evaluate how nutrient intake influences metabolite levels within each group.

Statistical implementation and modeling considerations

All *P*-values were adjusted for the false discovery rate (FDR) using the Benjamini–Hochberg method, with values less than 0.05 considered statistically significant. All statistical analyses, except for pathway enrichment analysis, were conducted using R, version 4.4.1 (R Foundation for Statistical Computing). Visualization was additionally performed using Cytoscape, version 3.10.3. Covariate adjustment was applied only in the logistic regression models to control for age, sex, and BMI, which are well-established confounders in nutritional metabolomics. Other analytic methods—such as the Wilcoxon rank-sum test, PLS-DA, fixed-effects models, and machine learning classifiers—were conducted without covariate adjustment. This approach aligns with standard best practices, where univariate and dimension-reduction methods are employed to explore inherent group differences without introducing potential overadjustment bias [40], while logistic regression is specifically designed to estimate adjusted associations. Machine learning models were similarly built using the full set of features for classification, focusing on predictive performance rather than inference, as is common in biomarker discovery applications.

Results

Clinical characteristics of the study population

A total of 838 participants with MetS and 1,468 non-MetS were included in this study. Table 1 presents the detailed demographic, biochemical, and anthropometric characteristics of the two groups. The mean age of the MetS group was significantly higher than that of the non-MetS group (58.33 ± 8.81 years vs. 55.62 ± 8.91 years, respectively). The proportion of males was higher in the non-MetS group than in the MetS group. BMI, waist circumference, systolic and diastolic blood pressure, and triglyceride and blood glucose levels were significantly higher in the MetS group, whereas HDL cholesterol levels were significantly lower. Furthermore, the prevalence of MetS-related phenotypes, including abdominal obesity, elevated blood pressure, elevated blood glucose, elevated triglycerides, and low HDL cholesterol levels, was significantly higher in the MetS group compared to the non-MetS group.

Identification of metabolic biomarkers

We matched 41 out of 135 metabolites to KEGG and HMDB pathways and identified those significantly associated with MetS using four different methods (Fig. 3). A volcano plot was used to visualize the differences between the two groups based on MetS status using the Wilcoxon rank-sum test (Fig. 2A). Among the 41 metabolites, 27 exhibited significant differences, with hexose and alanine showing the most pronounced differences, both of which were elevated in the MetS group (FC = 0.95 and 1.46, respectively). In addition, among the 23 nutrients analyzed, three—fat, retinol, and cholesterol—showed significant differences and were decreased in the MetS group (FC = 0.93, 0.87, and 0.91, respectively).

Table 1 Characteristics of study participants stratified by the presence of metabolic syndrome

Characteristics	MetS group ($n=838$)	95% CI	non-MetS group ($n=1,468$)	95% CI	*P*-value
Age (years)	58.33 ± 8.81	57.7–58.9	55.62 ± 8.91	55.2–56.1	< 0.0001
Male (%)	371 (44.27)		738 (50.27)		0.0063
BMI (kg/m^2)	26.45 ± 2.93	26.2–26.6	23.38 ± 2.82	23.2–23.5	< 0.0001
Waist circumference (cm)	91.06 ± 7.52	90.6–91.6	81.40 ± 7.66	81.0–81.8	< 0.0001
Systolic blood pressure (mmHg)	125.05 ± 15.94	124.0–126.1	113.47 ± 15.74	112.7–114.3	< 0.0001
Diastolic blood pressure (mmHg)	83.13 ± 9.63	82.5–83.8	76.02 ± 9.80	75.5–76.5	< 0.0001
Triglycerides (mg/dL)	212.25 ± 161.04	201.3–223.2	109.22 ± 54.18	106.4–112.0	< 0.0001
Blood glucose (mg/dL)	104.89 ± 24.44	103.2–106.5	89.88 ± 13.91	89.2–90.6	< 0.0001
HDL cholesterol (mg/dL)	39.00 ± 7.91	38.5–39.5	46.97 ± 10.39	46.4–47.5	< 0.0001
Abdominal obesity (%)	625 (74.58)		253 (17.23)		< 0.0001
Elevated blood pressure (%)	639 (76.25)		409 (27.86)		< 0.0001
Elevated fasting glucose (%)	465 (55.49)		249 (16.96)		< 0.0001
Elevated triglycerides (%)	558 (66.59)		195 (13.28)		< 0.0001
Low HDL cholesterol (%)	684 (81.62)		615 (41.89)		< 0.0001

Categorical variables are presented as numbers and percentages (%), whereas continuous variables are expressed as mean ± standard deviation *P*-values were calculated using the Wilcoxon rank-sum test for continuous variables and the chi-square test for categorical variables

BMI body mass index, *HDL* high-density lipoprotein, *MetS* metabolic syndrome, *CI* confidence interval

Differential metabolites and nutrients between the two groups were identified using PLS-DA with VIP values > 1. Among the 41 metabolites and 23 nutrients, 14 metabolites and 10 nutrients effectively distinguished the MetS group from the non-MetS group. The top 14 metabolites and top 10 nutrients with the highest VIP scores and greatest contributions to the separation between the MetS and non-MetS groups were visualized in the VIP score plot (Fig. 2B). Among the 14 metabolites, only glycine tended to decrease in the MetS group, whereas the other metabolites showed an increase. In contrast, the intake of all 10 nutrients was lower in the MetS group, with hexose having the highest VIP score among metabolites and fat being the most prominent nutrient.

Among the 41 candidate metabolites, 14 exhibited significant differences between the MetS and non-MetS groups based on the Wilcoxon rank-sum test and had VIP values > 1 (Table 2). Similarly, among the 23 candidate nutrients, 10 were identified with VIP values > 1, of which three—fat, retinol, and cholesterol—showed significant differences according to the Wilcoxon rank-sum test (Table 3).

Figure 3 presents Venn diagrams illustrating the overlap of metabolites and nutrients significantly associated with MetS, as identified using Wilcoxon, logistic regression, PLS-DA, and group LASSO. Nutrients were excluded from the logistic regression and group LASSO analyses owing to the lack of significant findings after adjustment, leaving the overlap based solely on Wilcoxon and PLS-DA results. The metabolite analysis identified 11 overlapping metabolites—hexose, alanine, glutamate, valine, isoleucine, proline, leucine, valerylcarnitine, glycine, propionylcarnitine, and phenylalanine—which accounted for 26.8% of the 41 metabolites analyzed and were consistently detected across all four analytical approaches. The nutrient analysis revealed three overlapping nutrients: fat, retinol, and cholesterol. This integrative approach provides a comprehensive understanding of the relationships between metabolites, nutrients, and MetS from multiple analytical perspectives.

Metabolic pathway enrichment analysis

Significantly enriched sub-pathways, derived from 11 metabolites identified by the Wilcoxon, adjusted logistic regression, PLS-DA, and group LASSO models, are illustrated as dot plots in Fig. 4. 'Arginine biosynthesis' and 'arginine and proline metabolism' were particularly significant in the group LASSO analysis, indicating a strong association between these pathways and MetS ($P = 1.37 \times 10^{-65}$; $P = 4.55 \times 10^{-43}$, respectively). Furthermore, 'glutathione metabolism,' 'porphyrin metabolism,' 'glyoxylate and dicarboxylate metabolism,' 'glycine, serine, and threonine metabolism,' 'primary bile acid biosynthesis,' 'valine, leucine, and isoleucine biosynthesis,' and 'valine, leucine, and isoleucine degradation' were identified as significant sub-pathways by all four analytical methods. In contrast, pathways that were not significant are depicted with faded colors and smaller sizes, indicating their relatively lower importance in the interpretation process.

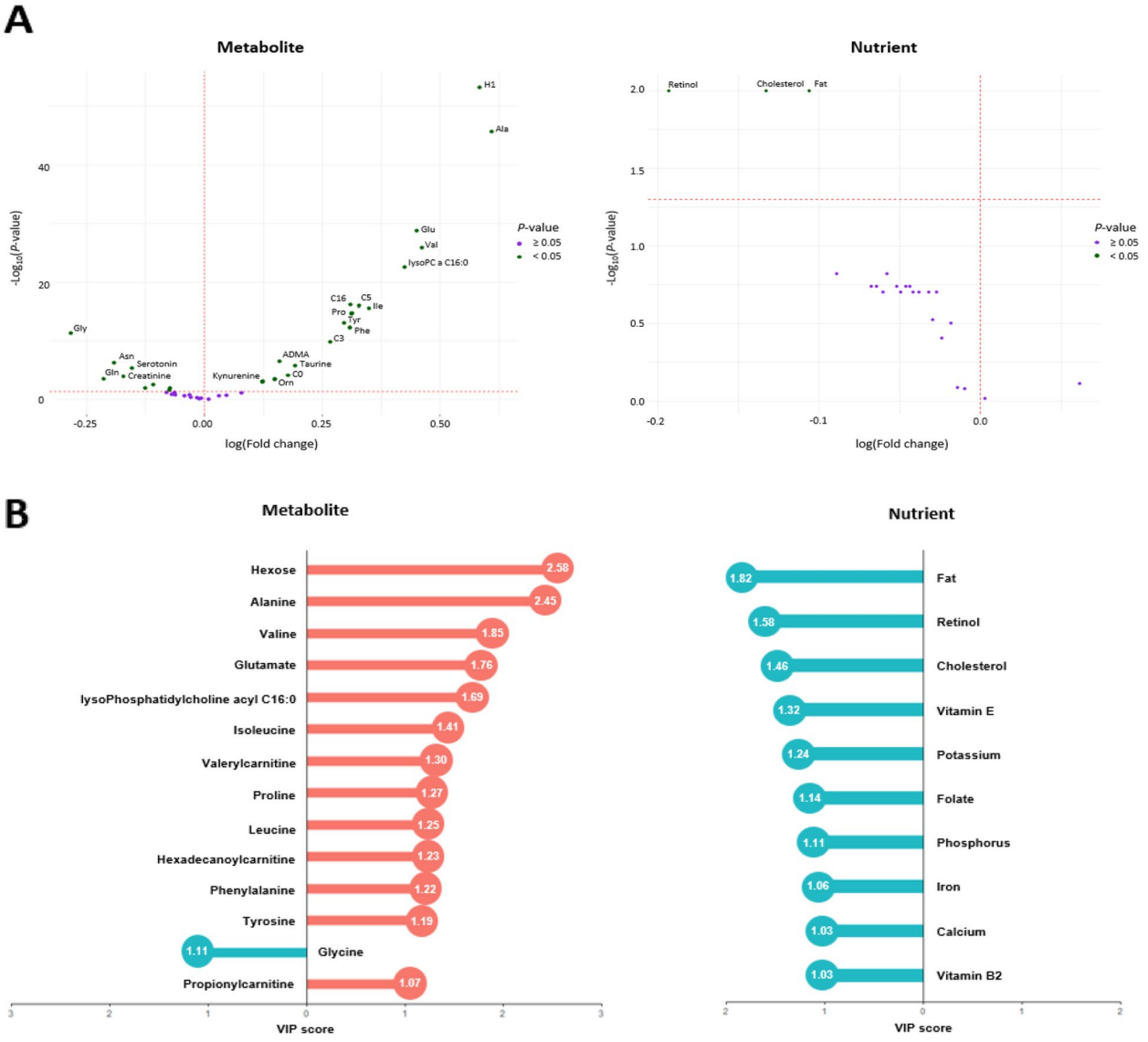

Fig. 2 Identification of metabolic biomarkers and disrupted pathways associated with metabolic syndrome. **A** Volcano plot of metabolites and nutrients identified using the Wilcoxon rank-sum test. The x-axis represents the log(fold-change), where fold-change indicates the relative abundance of each metabolite or nutrient in MetS cases compared with non-MetS. The y-axis represents − log10(adjusted *P*-value), with *P*-values corrected for multiple comparisons using the Benjamini–Hochberg method. Purple dots represent metabolites or nutrients with non-significant *P*-values, whereas green dots represent those with significant *P*-values. The horizontal red line denotes the significance threshold of 0.05. **B** VIP score plot showing the top metabolites and nutrients that differentiate MetS cases from non-MetS, as identified by PLS-DA. All metabolite and nutrient values were standardized (z-scores) before VIP analysis. MetS, metabolic syndrome; PLS-DA, partial least squares-discriminant analysis; VIP, variable importance in projection

Associations between metabolic biomarkers and nutrients

Spearman's correlation coefficients were calculated to explore the potential relationships between the 11 differential metabolites identified by all four methods and the 23 nutrients. The correlation coefficient matrix is shown in Fig. 5, with significant correlation coefficients ranging from – 0.097 to 0.162. Among the 11 metabolites, valine exhibited a significant positive correlation with all nutrients. Valerylcarnitine showed significant correlations with 12 nutrients, with only one, vitamin C, being negatively correlated. Propionylcarnitine was associated with six nutrients and showed a positive correlation with energy and carbohydrates. Proline demonstrated positive correlations with eight nutrients, whereas leucine and isoleucine were positively correlated with 18 and 17 nutrients, respectively. In contrast, glycine exhibited a significant negative correlation with 11 nutrients. Glutamate showed a weak but significant positive correlation with sodium, whereas phenylalanine and alanine did not show any significant correlations with any nutrient. Among all significant correlations, the highest coefficient was observed between valine and niacin (r = 0.16,

Table 2 Differential metabolites identified using the Wilcoxon rank-sum test and PLS-DA

Metabolites	MetS group		non-MetS group		Trend	FC	VIP	*P*-value	BH *P*-value
	Mean	SD	Mean	SD					
Hexose (μM)	5623.39	1498.80	4859.05	953.18	Up	0.95	2.58	7.04E-54	2.89E-52
Alanine (μM)	546.55	111.96	477.00	108.22	Up	1.46	2.45	2.15E-46	4.40E-45
Glutamate (μM)	187.81	81.30	161.31	76.93	Up	1.48	1.76	1.83E-29	2.50E-28
Valine (μM)	230.96	39.87	212.30	39.23	Up	1.48	1.85	1.46E-26	1.50E-25
lysoPhosphatidylcholine acyl C16:0 (μM)	281.35	59.17	256.44	59.58	Up	1.51	1.69	3E-23	2.46E-22
Hexadecanoylcarnitine (μM)	0.16	0.05	0.14	0.07	Up	1.90	1.23	6.91E-17	4.72E-16
Valerylcarnitine (μM)	0.16	0.05	0.14	0.08	Up	1.35	1.30	1.06E-16	6.21E-16
Isoleucine (μM)	87.18	19.62	80.48	18.14	Up	1.56	1.41	3.08E-16	1.58E-15
Proline (μM)	174.27	48.14	160.51	47.92	Up	1.20	1.27	2.13E-15	9.71E-15
Tyrosine (μM)	74.04	14.13	69.89	14.43	Up	1.20	1.19	9.12E-14	3.74E-13
Leucine (μM)	182.04	35.94	171.05	33.47	Up	1.59	1.25	5.55E-13	2.07E-12
Phenylalanine (μM)	104.79	16.71	99.73	16.97	Up	1.24	1.22	6.45E-13	2.21E-12
Glycine (μM)	325.49	82.51	346.74	84.26	Down	1.27	1.11	5.53E-12	1.75E-11
Propionylcarnitine (μM)	0.48	0.15	0.44	0.15	Up	1.22	1.07	1.7E-10	4.99E-10

The mean and SD for the MetS and non-MetS groups are presented, along with the trend (down: lower in the MetS group), FC, VIP score, *P*-value (unadjusted from the Wilcoxon rank-sum test), and BH *P*-value (adjusted for multiple comparisons using the BH method)

Mean and SD values represent raw metabolite concentrations (μM)

BH Benjamini–Hochberg, *FC* fold change, *MetS* metabolic syndrome, *PLS-DA* partial least squares-discriminant analysis, *SD* standard deviation, *VIP* variable importance in projection

Table 3 Differential nutrients identified using the Wilcoxon rank-sum test and PLS-DA

Nutrients	MetS group		non-MetS group		Trend	FC	VIP	*P*-value	BH *P*-value
	Mean	SD	Mean	SD					
Fat (g)	26.15	19.07	28.14	18.61	Down	0.93	1.82	0.0008	0.0100
Retinol (μg)	52.19	48.54	59.66	57.14	Down	0.87	1.58	0.0011	0.0100
Cholesterol (mg)	134.95	113.35	147.98	125.02	Down	0.91	1.46	0.0013	0.0100
Vitamin E (mg)	7.65	4.04	8.13	5.08	Down	0.94	1.32	0.0328	0.1509
Potassium (mg)	2194.67	1029.01	2284.92	1046.23	Down	0.96	1.24	0.0286	0.1509
Folate (μg)	209.13	115.66	219.16	124.00	Down	0.95	1.14	0.0495	0.1821
Phosphorus (mg)	868.51	343.51	896.81	360.18	Down	0.97	1.11	0.0645	0.1821
Fe (mg)	9.44	4.50	9.73	4.70	Down	0.97	1.06	0.0650	0.1821
Calcium (mg)	410.93	241.09	428.57	257.10	Down	0.96	1.03	0.1153	0.1974
Vitamin B2 (mg)	0.86	0.43	0.89	0.44	Down	0.96	1.03	0.0792	0.1821

The mean and SD for the MetS and non-MetS groups are presented, along with the trend (down: lower in the MetS group), FC, VIP score, *P*-value (unadjusted from the Wilcoxon rank-sum test), and BH *P*-value (adjusted for multiple comparisons using the BH method)

BH Benjamini–Hochberg, *FC* fold change, *Fe* iron, *MetS* metabolic syndrome, *PLS-DA* partial least squares-discriminant analysis, *SD* standard deviation, *VIP* variable importance in projection

$P < 0.0001$). All computed Spearman's correlation coefficients are presented in Supplementary Table 1.

We subsequently examined the correlation network to assess the associations between nutrients, metabolites, and MetS components (Supplementary Fig. 1). Overall, the associations between nutrients and metabolites, as well as between nutrients and MetS components, were generally weak. In contrast, a strong association was observed between MetS components and metabolites, with blood glucose and hexose levels showing the strongest positive correlation ($r = 0.67$, $P < 0.0001$).

Associations between metabolic biomarkers, nutrients, and MetS

A fixed-effects model was constructed to investigate potential interactions between nutrients and metabolites that differ by MetS status. Significant differences in metabolite–nutrient pairs were observed between the two groups. In the MetS group, 16 pairs showed significant correlations, with 6 pairs—'isoleucine–fat,' 'isoleucine–P,' 'proline–fat,' 'leucine–fat,' 'leucine–P,' and 'valerylcarnitine–niacin'—being exclusive to this group. In contrast, the non-MetS group exhibited 23 significant pairs, including 13 unique ones, such as 'hexose–vitamin B6,' 'hexose–fiber,' 'valine–folate,' 'valine–ash,' 'valine–fiber,' 'isoleucine–carbohydrate,' 'isoleucine–Fe,' 'leucine–Fe,' 'leucine–fiber,' 'valerylcarnitine–protein,'

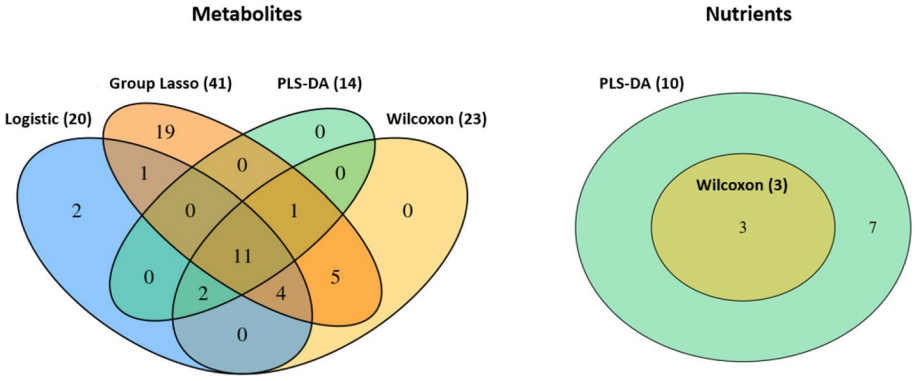

Fig. 3 Overlap of metabolites and nutrients associated with metabolic syndrome across statistical methods. **A** Venn diagram showing the overlap of metabolites identified by Wilcoxon (yellow), adjusted logistic regression (blue), PLS-DA (green), and group LASSO (orange). **B** A separate Venn diagram illustrating the overlap of nutrients identified by Wilcoxon (yellow) and PLS-DA (green). LASSO, least absolute shrinkage and selection operator; PLS-DA, partial least squares-discriminant analysis

'valerylcarnitine–carbohydrate,' 'glycine–vitamin B1,' and 'glycine–vitamin B6' (Fig. 6).

A negative association between glycine and protein levels was observed in both the MetS (estimate = −0.615, $P = 0.002$) and non-MetS group (estimate = −0.486, $P = 0.002$). In the MetS group, positive associations were identified between isoleucine and niacin (estimate = 0.450, $P < 0.001$), as well as between leucine and niacin (estimate = 0.436, $P < 0.001$). In particular, isoleucine (estimate = 0.199, $P = 0.008$), proline (estimate = 0.177, $P = 0.004$), and leucine (estimate = 0.166, $P = 0.015$) showed significant positive associations with fat intake only in the MetS group. Conversely, in the non-MetS group, positive associations were observed between valerylcarnitine and protein (estimate = 0.340, $P = 0.015$), as well as between valerylcarnitine and carbohydrate (estimate = 0.398, $P = 0.019$). Furthermore, in the non-MetS group, niacin showed a positive association with hexose, valine, isoleucine, proline, and leucine $P < 0.005$) (Table 4).

Machine learning-based diagnostic models for MetS using serum metabolomics data: a comparative analysis of eight classification algorithms

Using data from 135 metabolites, classification models were developed using eight machine learning algorithms: SVM, RF, KNN, LR, DNN, NB, ADB, and SGD. The classification performance of these models is presented in Table 5 and illustrated in Fig. 7. Among the eight algorithms, SGD achieved the highest performance, with an AUC of 0.84 (95% confidence interval: 0.81–0.87). The AUC values across models ranged from 0.74 to 0.84, demonstrating their robustness in distinguishing MetS cases from non-MetS.

Discussion

In this study, we identified specific metabolic and nutritional markers associated with MetS in middle-aged Korean adults, highlighting disrupted metabolic pathways that may be relevant for informing future dietary recommendations. A comprehensive analysis using multiple statistical models identified 11 metabolites and three nutrients that were differentially abundant in MetS. Notably, the 'arginine biosynthesis' and 'arginine and proline metabolism' pathways were significantly altered. This study identified novel, MetS-specific nutrient–metabolite relationships through an integrated network analysis, shedding light on the complex interactions between diet and metabolism. Furthermore, incorporating machine learning into metabolomics enabled the characterization of key metabolic markers, offering a new perspective on predictive modeling in a large Korean cohort.

Among the 11 metabolites consistently identified across all four methods, hexose emerged as the most prominent, exhibiting a positive correlation with MetS. A previous study in older cohorts from Japan and the United States also reported a similar association between serum hexose levels and MetS [41]. Among hexoses, glucose functions as the primary energy source in living organisms. During glucose metabolism, a fraction enters the hexosamine biosynthesis pathway (HBP), a relatively minor branch of glycolysis. In this pathway, fructose-6-phosphate is converted to glucosamine-6-phosphate by the rate-limiting enzyme glutamine:fructose-6-phosphate amidotransferase (GFAT), which is subsequently converted to UDP-N-acetylglucosamine (UDP-GlcNAc), the major end product. UDP-GlcNAc is used for O-GlcNAcylation of intracellular proteins at serine/threonine residues, and elevated O-GlcNAcylation has been implicated in insulin resistance and vascular complications of diabetes [42]. In our study, hexose showed a strong correlation with blood glucose levels, a key component of MetS. This finding

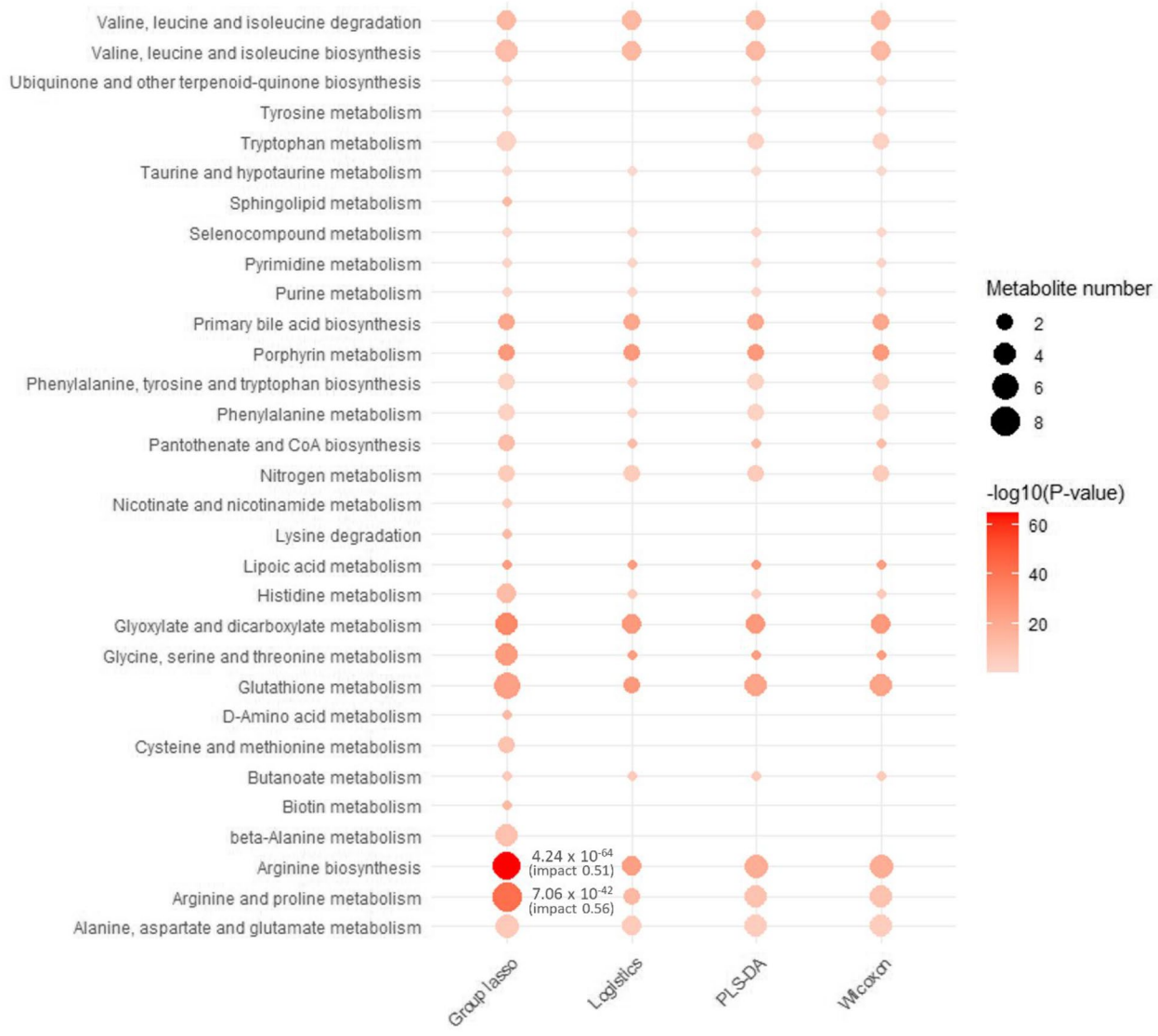

Fig. 4 Pathway enrichment analysis of metabolites associated with metabolic syndrome. Significantly enriched KEGG pathways are shown in a dot plot, where the horizontal axis indicates the analysis methods, and the vertical axis lists the sub-pathway names. Dot color represents statistical significance (− log10(*P*-value)), and dot size indicates the number of matched metabolites. The pathways identified as significantly enriched are labeled with false discovery rate-corrected *p*-values and pathway impact scores. CoA, coenzyme A; KEGG, Kyoto Encyclopedia of Genes and Genomes; LASSO, least absolute shrinkage and selection operator; PLS-DA, partial least squares-discriminant analysis

aligns with a previous KoGES-based study that reported a significant association between serum hexoses and the risk of developing type 2 diabetes [16].

The levels of branched-chain and aromatic amino acids (BCAA and AAA), including leucine, isoleucine, valine, tyrosine, and phenylalanine, were significantly higher in the MetS group. A study conducted in a Mediterranean male population reported significantly higher BCAA and AAA levels in individuals with MetS [43]. Similarly, a Chinese study reported a 2.17-fold higher prevalence of MetS among individuals with the highest BCAA levels [44]. These patterns suggest that elevated levels of BCAA and AAA reflect dysregulated metabolic pathways associated with MetS.

Short-chain acylcarnitines, including valerylcarnitine and propionylcarnitine, were positively associated with MetS, consistent with previous findings [45]. As both are byproducts of BCAA catabolism, their elevated levels in our study may reflect the increased BCAA concentrations observed in the MetS group. Among the 11 candidate metabolites, glycine was the only one significantly lower in individuals with MetS. This finding aligns with consistent reports of reduced plasma glycine levels in MetS among older Chinese adults [46].

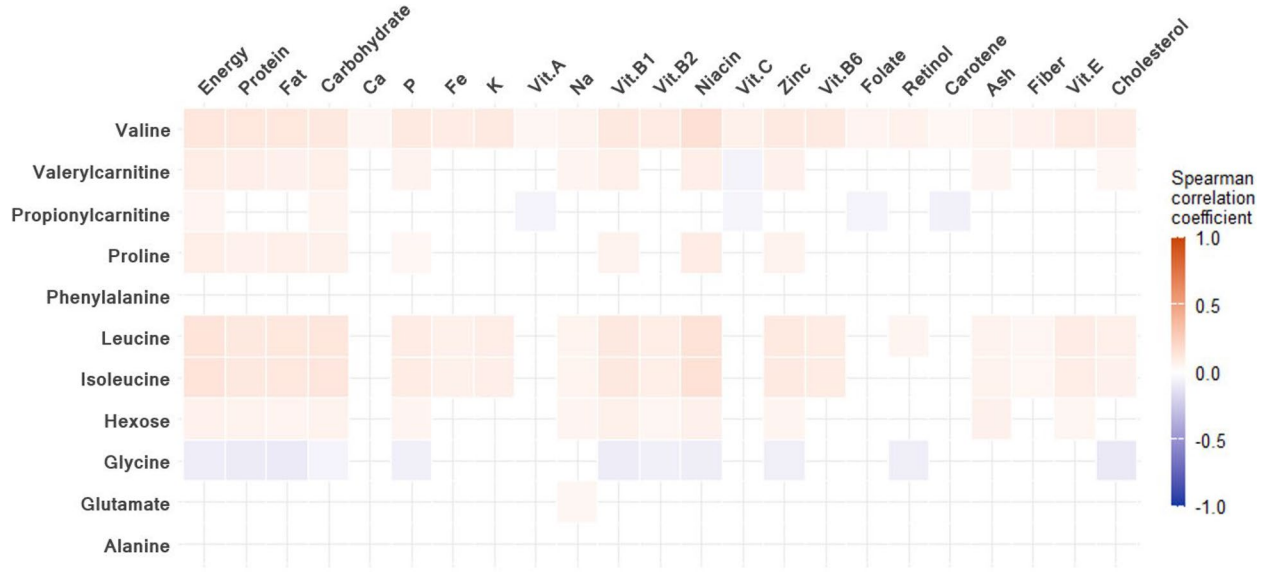

Fig. 5 Correlation coefficient matrix of potential metabolic biomarkers and nutrients. All statistically significant correlations between metabolites (rows) and nutrients (columns) are color-coded, whereas insignificant correlations are left blank. Positive correlations are shown in blue, and negative correlations in red. *P*-values were adjusted for multiple comparisons using the Benjamini–Hochberg method. Ca, calcium; Fe, iron; K, potassium; Na, sodium; P, phosphorus; Vit, vitamin

Regarding nutrient intake, the MetS group had lower levels of fat, retinol, and cholesterol consumption. While some studies have reported higher fat intake among individuals with MetS, these contrasting findings may reflect cultural dietary differences, particularly in the types of fats commonly consumed and the overall dietary patterns, such as whether the primary fat sources are plant-based or animal-based, across populations. This suggests that not only the quantity of fat, but also its quality, may vary across cultures and may be differentially associated with MetS risk across populations. A systematic review highlighted that replacing saturated fats with unsaturated fats reduces the risk of MetS [47], and in Korean adults, very low fat intake—especially below 15% of total energy—has been associated with an increased prevalence of MetS [48]. Population-specific dietary patterns, such as high intake of fermented vegetables and legumes, and low consumption of processed meats and dairy, contrast sharply with Western diets, which are generally richer in saturated fats and refined sugars. These findings underscore the importance of considering culturally specific dietary patterns when examining the associations between nutrient intake and MetS, suggesting that both nutrient quality and overall dietary context may be linked to MetS risk in population-specific ways. Therefore, the observed lower fat intake among Korean adults with MetS may reflect complex interactions between dietary quality, cultural habits, and disease risk, which may guide future research or inform hypothesis generation regarding population-specific recommendations. In addition to macronutrients, bioactive food components such as phytochemicals are emerging as important contributors to metabolic health. Although the KoGES dataset does not explicitly include variables for flavonoids and polyphenols—compounds known to modulate key metabolic pathways—their intake can be indirectly estimated from patterns of phytochemical-rich food consumption [49]. Given their abundance in plant-based diets and well-documented metabolic benefits, these compounds warrant further investigation in future research to profile their associated metabolomics, which may reveal key biochemical pathways involved in the development of MetS. Specifically, elucidating the interactions between phytochemical intake and metabolite concentrations may clarify how these bioactive compounds modulate glucose and lipid metabolism, inflammation, and oxidative stress, thereby identifying novel interactions that influence MetS risk or protection and offering new avenues for precision nutrition.

The 'arginine biosynthesis' and 'arginine and proline metabolism' pathways exhibited the strongest associations with MetS in the group LASSO model. These pathways are closely linked to vascular function through their role in nitric oxide (NO) production. Clinical studies have reported that insufficient L-arginine levels in obese adolescent patients with MetS may impair NO production, which plays a crucial role in vasodilation and blood pressure regulation [50]. Decreased L-arginine has also been associated with vascular dysfunction and key MetS features such as hypertension, impaired glucose tolerance, and obesity [51, 52]. Although antihypertensive medications were included in the diagnostic criteria for

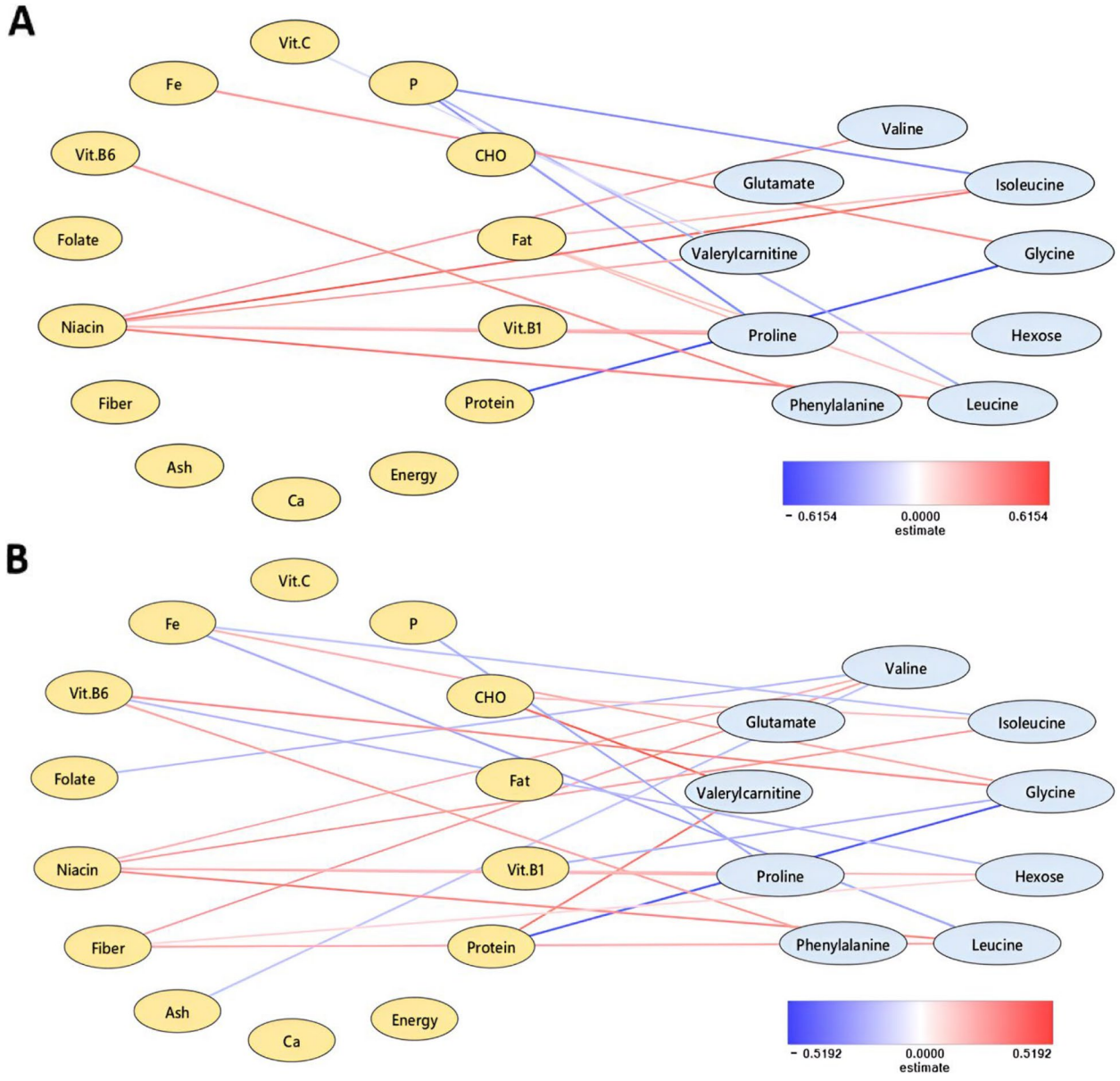

Fig. 6 The effect of nutrients on metabolites based on the presence or absence of metabolic syndrome. **A** MetS group. **B** non-MetS group. Significant nutrient–metabolite associations ($P < 0.05$) are represented by lines, with red indicating positive correlations and blue indicating negative correlations. The intensity of the line color corresponds to the magnitude of the estimated effect, with stronger associations depicted in more prominent colors. Ca, calcium; CHO, carbohydrate; Fe, iron; MetS, metabolic syndrome; P, phosphorus; Vit, vitamin

MetS and therefore could not be treated as confounding variables in this study, future studies aiming to elucidate the role of arginine-related pathways in MetS more precisely should consider including medications related to vascular function and blood pressure as potential confounders. These findings support a potential link between arginine-related metabolism and metabolic alterationsin MetS, with possible implications for cardiovascular health.

In addition to the arginine pathway, several other metabolically enriched pathways identified in this study provide further insight into the pathophysiology of MetS. Glutathione metabolism was enriched across multiple analytical models and may indicate redox imbalance as a contributing factor to insulin resistance and vascular dysfunction [53]. Likewise, the glycine, serine, and threonine metabolism pathways play important roles in antioxidant defense, one-carbon metabolism, and glucose regulation [54, 55]. The lower levels of glycine observed in MetS participants may reflect impaired redox homeostasis and compromised metabolic flexibility. Furthermore, the accumulation of short-chain acylcarnitines

Table 4 Nutrient–metabolite associations stratified by the presence of metabolic syndrome

Fixed effects	Estimate	SE	z	*P*-value
MetS				
Hexose				
Niacin	0.205	0.082	2.504	0.012
Vitamin B6	−0.177	0.121	−1.465	0.143
Fiber	0.005	0.073	0.069	0.945
Glutamate				
Ca	−0.045	0.032	−1.402	0.161
Valine				
Fat	0.030	0.057	0.519	0.604
Carbohydrate	−0.059	0.045	−1.311	0.190
Niacin	0.288	0.075	3.860	< 0.001
Folate	−0.121	0.089	−1.367	0.172
Ash	−0.084	0.055	−1.530	0.126
Fiber	0.064	0.097	0.664	0.507
Isoleucine				
Protein	0.015	0.191	0.080	0.936
Fat	0.199	0.075	2.641	0.008
Carbohydrate	0.033	0.050	0.658	0.511
P	−0.385	0.151	−2.544	0.011
Fe	−0.140	0.092	−1.520	0.129
Niacin	0.450	0.099	4.567	< 0.001
Ash	−0.003	0.0493	−0.069	0.945
Proline				
Fat	0.177	0.061	2.891	0.004
P	−0.388	0.085	−4.557	< 0.001
Niacin	0.300	0.081	3.720	< 0.001
Leucine				
Fat	0.166	0.068	2.430	0.015
P	−0.267	0.108	−2.466	0.014
Fe	−0.105	0.124	−0.842	0.400
Niacin	0.436	0.088	4.937	< 0.001
Folate	−0.152	0.096	−1.589	0.112
Ash	−0.013	0.057	−0.226	0.822
Fiber	0.086	0.109	0.795	0.427
Valerylcarnitine				
Energy	0.086	0.256	0.338	0.736
Protein	−0.112	0.147	−0.762	0.446
Carbohydrate	−0.063	0.164	−0.386	0.700
Niacin	0.260	0.096	2.720	0.007
Vitamin C	−0.117	0.047	−2.517	0.012
Glycine				
Energy	0.512	0.331	1.547	0.122
Protein	−0.615	0.198	−3.110	0.002
Carbohydrate	−0.264	0.212	−1.248	0.212
Fe	0.354	0.116	3.044	0.002
Vitamin B1	−0.180	0.113	−1.591	0.112
Vitamin B6	0.086	0.125	0.688	0.492
Phenylalanine				
Fe	−0.190	0.099	−1.924	0.055
Vitamin B6	0.353	0.104	3.394	0.001
Folate	−0.146	0.076	−1.932	0.054
non-MetS				
Hexose				

Table 4 (continued)

Fixed effects	Estimate	SE	z	*P*-value
Niacin	0.180	0.051	3.543	< 0.001
Vitamin B6	−0.198	0.074	−2.680	0.007
Fiber	0.104	0.045	2.335	0.020
Glutamate				
Ca	−0.009	0.027	−0.346	0.729
Valine				
Fat	0.012	0.050	0.248	0.804
Carbohydrate	0.041	0.039	1.058	0.290
Niacin	0.189	0.064	2.973	0.003
Folate	−0.192	0.070	−2.735	0.006
Ash	−0.172	0.052	−3.309	0.001
Fiber	0.233	0.076	3.085	0.002
Isoleucine				
Protein	0.116	0.142	0.820	0.412
Fat	0.044	0.064	0.682	0.495
Carbohydrate	0.151	0.042	3.634	< 0.001
P	−0.236	0.121	−1.955	0.051
Fe	−0.177	0.071	−2.498	0.013
Niacin	0.239	0.073	3.295	0.001
Ash	−0.006	0.043	−0.135	0.892
Proline				
Fat	0.078	0.054	1.436	0.151
P	−0.240	0.071	−3.376	0.001
Niacin	0.215	0.066	3.273	0.001
Leucine				
Fat	0.093	0.059	1.581	0.114
P	−0.082	0.087	−0.935	0.350
Fe	−0.259	0.097	−2.672	0.008
Niacin	0.311	0.068	4.591	< 0.001
Folate	−0.096	0.072	−1.325	0.185
Ash	−0.097	0.052	−1.888	0.059
Fiber	0.228	0.081	2.812	0.005
Valerylcarnitine				
Energy	−0.519	0.265	−1.960	0.050
Protein	0.340	0.139	2.444	0.015
Carbohydrate	0.398	0.170	2.342	0.019
Niacin	0.027	0.079	0.341	0.733
Vitamin C	−0.238	0.041	−5.856	< 0.001
Glycine				
Energy	0.396	0.274	1.444	0.149
Protein	−0.486	0.158	−3.077	0.002
Carbohydrate	−0.274	0.177	−1.547	0.122
Fe	0.206	0.087	2.367	0.018
Vitamin B1	−0.214	0.084	−2.557	0.011
Vitamin B6	0.304	0.093	3.260	0.001
Phenylalanine				
Fe	−0.130	0.082	−1.586	0.113
Vitamin B6	0.238	0.085	2.782	0.005
Folate	−0.107	0.061	−1.741	0.082

Data were analyzed using a fixed-effects model to account for individual variability, with significant associations defined as $P < 0.05$

Ca calcium, *Fe* iron, *MetS* metabolic syndrome, *P* phosphorus, *SE* standard error

Table 5 Comparative performance metrics of eight machine learning algorithms for metabolic syndrome classification using serum metabolomics

Models	Sensitivity	Specificity	Non-error Rate	AUC (95% CI)
Support vector machine	0.394	0.145	0.236	0.835 (0.804–0.865)
Random forest	0.510	0.877	0.744	0.812 (0.779–0.845)
K-nearest neighbors	0.462	0.841	0.703	0.749 (0.713–0.786)
Logistic regression	0.614	0.809	0.738	0.819 (0.788–0.851)
Deep neural network	0.327	0.282	0.298	0.744 (0.705–0.783)
Naive Bayes	0.705	0.705	0.705	0.749 (0.712–0.786)
AdaBoost	0.606	0.841	0.755	0.822 (0.790–0.853)
Stochastic gradient descent	0.653	0.866	0.789	0.841 (0.811–0.870)

AUC area under the curve, *CI* confidence interval

such as valerylcarnitine and propionylcarnitine points to dysregulated BCAA catabolism, which has been linked to mitochondrial overload and lipid oxidation defects [56, 57]. These metabolite-level disruptions support the hypothesis that mitochondrial dysfunction is a central mechanistic axis in the development of MetS [58, 59]. Impaired synthesis of L-arginine and NO elevates oxidative stress, increasing the cellular demand for glutathione [60]. The metabolism of glycine, serine, and threonine contributes to glutathione production by providing critical precursors like glycine and serine, thus directly linking amino acid availability to redox regulation [61]. At the same time, disruptions in BCAA catabolism lead to mitochondrial overload, which further exacerbates oxidative stress. These interconnected metabolic disturbances collectively create a pro-inflammatory, insulin-resistant environment—hallmarks of MetS, along with endothelial dysfunction [62]. Furthermore, the accumulation of short-chain acylcarnitines reflects impaired β-oxidation capacity within mitochondria. This leads to incomplete lipid breakdown and increased generation of reactive oxygen species (ROS) [63]. The resulting oxidative stress compromises mitochondrial integrity, disrupts the electron transport chain and reduces ATP synthesis [64]. In this context, glutathione metabolism serves as a key antioxidant defense, neutralizing ROS and maintaining mitochondrial membrane stability [65]. Additionally, glycine and serine—through the one-carbon metabolism pathway—support epigenetic regulation and mitochondrial protein maintenance via methylation reactions [66].

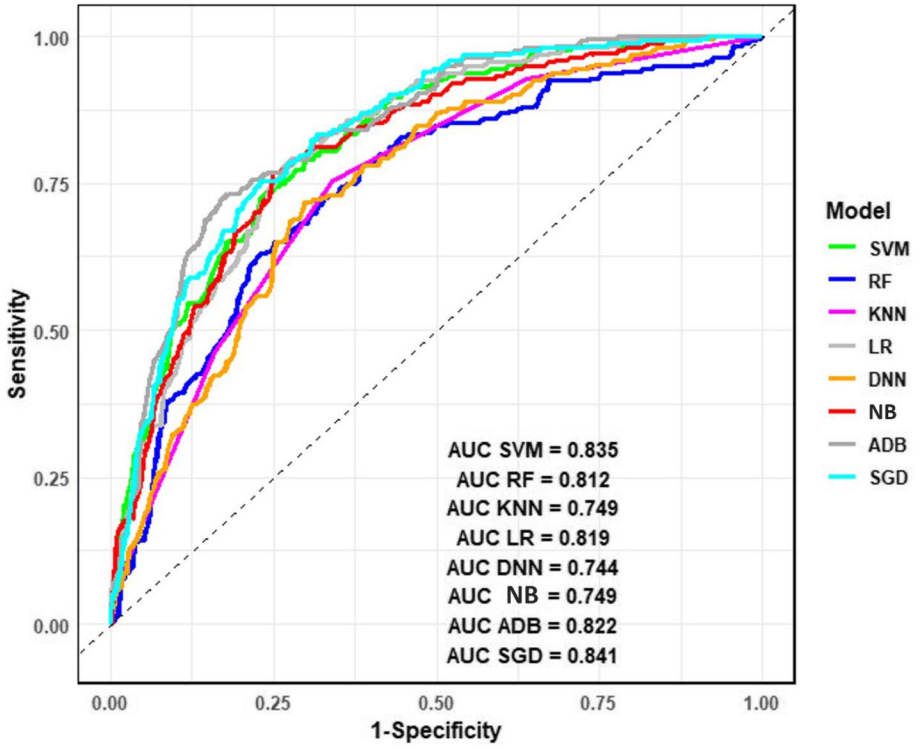

Fig. 7 ROC curves of eight machine learning algorithms for metabolic syndrome classification using serum metabolomics profiles. ADB, AdaBoost; ROC, receiver operating characteristic; AUC, area under the curve; DNN, deep neural network; KNN, k-nearest neighbors; LR, logistic regression; NB, Naive Bayes; RF, random forest; SGD, stochastic gradient descent; SVM, support vector machine

Taken together, these findings underscore the central role of mitochondrial homeostasis in metabolic health. Disruptions in these metabolic pathways can both initiate and perpetuate mitochondrial dysfunction, driving the progression of MetS through mechanisms such as insulin resistance and vascular impairment [62, 67].

Metabolites serve as valuable biomarkers for exploring diet–disease associations. In this study, BCAAs were positively correlated with energy and nutrient intake, consistent with their role in energy metabolism. Excessive nutrient intake or obesity may lead to elevated BCAA turnover, which has been linked to impaired fatty acid oxidation and insulin resistance [68]. Hexoses were positively correlated with all 23 nutrients in this study, highlighting the influence of overall dietary load on metabolic profiles. A cross-sectional analysis from the European Prospective Investigation into Cancer and Nutrition-Potsdam study similarly reported that high consumption of red meat and fish, combined with low intake of whole-wheat bread and tea, was associated with elevated plasma hexose levels [69]. In contrast, glycine showed a significant negative correlation with energy and calorie-rich nutrients, consistent with its protective role against insulin resistance [70]. Valerylcarnitine and propionylcarnitine levels were negatively associated with antioxidant nutrients, consistent with the role of antioxidants in promoting fatty acid oxidation and reducing acylcarnitine accumulation resulting from incomplete oxidation [65]. Although individual metabolite–nutrient correlations were generally weak in this study, such as the correlation between valine and niacin ($r = 0.16$), previous research suggests that combining multiple metabolites can improve the accuracy of dietary predictions [71], underscoring the complexity of underlying metabolic pathways and the need for integrative analytical approaches. Importantly, when statistically significant, correlation coefficients as low as $r = 0.15$ are often considered biologically meaningful, as they may reflect subtle but functionally relevant dietary influences on metabolism, such as protein influencing the biosynthesis of BCAAs, fat modulating mitochondrial β-oxidation, or niacin and riboflavin serving as cofactors in redox-related enzymatic pathways [72]. However, given the large number of metabolite–nutrient correlations examined, the potential for false-positive associations due to multiple testing cannot be completely excluded, even with FDR correction.

We identified metabolite–nutrient associations that differed by MetS status, with 16 pairs significant in the MetS group and 23 in the non-MetS group. Notably, niacin also exhibited significant positive associations with hexose, valine, isoleucine, leucine, and valerylcarnitine in the MetS group. Niacin serves as a precursor to nicotinamide adenine dinucleotide (NAD^+) and nicotinamide adenine dinucleotide phosphate ($NADP^+$), which play key roles in mitochondrial respiration, glycolysis, and lipid β-oxidation [73]. Niacin is mainly obtained from dietary sources such as meat and fish, suggesting that dietary patterns may partly explain the observed associations. The valerylcarnitine-niacin pair showed significantly positive results only in the MetS group. As a metabolite involved in fatty acid β-oxidation and BCAA metabolism, valerylcarnitine reflects metabolic responses to glucose [74]. Its altered levels may indicate mitochondrial dysfunction and impaired energy metabolism associated with insulin resistance [75]. These findings highlight the need for further investigation into valerylcarnitine pathways and their role in metabolic health in MetS.

The unique metabolite–nutrient relationships observed in the non-MetS and MetS groups underscore the complex interplay of metabolic pathways in maintaining health. For instance, vitamin B6 was positively associated with glycine but negatively associated with hexose in the non-MetS group, suggesting its involvement in amino acid and carbohydrate metabolism. Pyridoxal phosphate (PLP), the active form of vitamin B6, functions as a coenzyme in glycine synthesis via serine hydroxymethyltransferase and in glycogenolysis through glycogen phosphorylase. In the non-MetS group, PLP likely contributed to lower hexose levels by promoting glycine production and supporting efficient glucose metabolism [76]. Conversely, in the MetS group, reduced PLP utilization and insulin resistance may have impaired glycine production, leading to hexose accumulation. Insulin resistance further contributes to elevated blood hexose levels by impairing glucose transport to tissues [77]. These findings highlight the critical role of vitamin B6 in metabolic health, particularly in the regulation of glycine and glucose metabolism.

Similarly, the relationship between Fe intake and BCAA levels underscores how nutrient metabolism differs between the MetS and non-MetS groups. In the non-MetS group, Fe intake was negatively associated with BCAA, including leucine and isoleucine, which may reflect a higher intake of non-heme Fe, primarily found in plant-based foods, compared with heme Fe from animal products. Non-heme Fe has been linked to reduced inflammation and improved metabolic function [78]. Conversely, in the MetS group, impaired Fe metabolism, likely owing to chronic inflammation and elevated hepcidin levels, may have inhibited Fe absorption and utilization [79]. These disruptions could explain why Fe intake did not significantly influence BCAA concentrations in the MetS group. Although this study did not differentiate between heme and non-heme Fe, future research should investigate their distinct effects on BCAA metabolism and their potential role in MetS.

The SGD model showed the best performance (AUC = 0.84) in our study, likely because it is well suited for small datasets with a limited number of variables, like our metabolite-only data. Although the number of metabolites was not large, some of them were correlated. SGD handles this kind of structure efficiently by applying regularization, which helps avoid overfitting. It also produces results that are easy to interpret, making it a practical and reliable choice for predicting MetS using metabolite data. These features make SGD particularly suitable for clinical deployment in Korea, where a unified national health insurance system and high electronic medical record (EMR) adoption provide a strong foundation for digital health tools. In practical terms, such a model could be embedded within EMR platforms at primary care clinics or health screening centers to provide automated risk scores for MetS, aiding early intervention. For effective translation into practice, collaboration with the National Health Insurance Service (NHIS), development of user-friendly interfaces for clinicians, and structured training modules will be key to promoting provider engagement and sustained use [80].

This study is the first to examine the association between serum metabolites, nutrient intake, and MetS in middle-aged Korean adults, emphasizing that metabolite–nutrient relationships can vary based on MetS status. The identified metabolite–nutrient pairs associated with MetS provide a foundation for exploring these relationships in other populations, underscoring their broader applicability. Furthermore, by applying machine learning algorithms to metabolomic data, we developed diagnostic models for MetS classification. Among these, the SGD model demonstrated the highest performance (AUC = 0.84), with high specificity (0.87) and moderate sensitivity (0.65). This trade-off suggests its current utility may lie more in confirmatory diagnosis than in early screening. However, the model also shows promise as a tool for early identification, and future refinement—such as incorporating additional clinical or dietary variables—may enhance its sensitivity and broaden its application in preventive strategies. Despite these strengths, the study has some limitations. First, its cross-sectional design limits the ability to infer causal relationships, highlighting the need for future investigations to distinguish between metabolic profile correlations and causality in MetS. Second, the combined analysis of men and women may have masked sex-specific metabolic variations, which should be addressed in future research. Third, the influence of metabolite–nutrient pairs on MetS, as assessed through the fixed-effects model, has not been experimentally validated. Fourth, in this study, dietary intake was assessed using a semi-quantitative FFQ, which is known to be susceptible to recall bias and measurement error. To minimize these biases, we used a validated FFQ instrument specifically developed for the Korean population. The validity and reproducibility of the study were evaluated against 12-day dietary records, which showed that nutrient density-based comparisons had similar levels of agreement. These findings support the reasonable validity of the FFQ in distinguishing differences in nutrient intake levels among Korean adults [28]. Nevertheless, as with all self-reported dietary methods, potential for non-differential misclassification remains, which may attenuate the observed associations between nutrient intake and metabolite levels. Thus, the associations observed in this study may be conservative estimates. Future research employing more objective measures such as biomarkers or multiple 24-h recalls is warranted to validate and extend our findings. Fifth, our analysis employed a minimally adjusted model controlling only for age, sex, and BMI. While these covariates are well-established confounders in nutritional metabolomics, other lifestyle-related variables such as physical activity, medication use, and dietary behaviors were not included. This modeling choice was made to avoid overadjustment bias and preserve the real-world lifestyle differences between MetS and non-MetS groups. Notably, prior studies investigating the associations between dietary exposures and metabolomic profiles have also adopted similar minimally adjusted models, demonstrating that meaningful and biologically relevant findings can be derived even when only basic demographic covariates are controlled for [81, 82]. However, we acknowledge that the absence of these covariates may result in residual confounding, which should be addressed in future studies using more comprehensive adjustment or stratified analyses. Additionally, in this study, a large proportion of participants (67%) from the initial cohort were excluded due to missing plasma metabolite data. Although this exclusion was based on data availability rather than participant characteristics, it may have introduced selection bias and reduced the generalizability of the findings. Finally, external validation of the predictive models could not be performed due to the unavailability of an independent dataset. While internal validation was conducted using a stratified split of the original dataset, this limitation may restrict the generalizability of the findings. To enhance model generalizability, future research should validate these findings in independent populations with diverse demographic and clinical characteristics.

This study identified key nutrients—such as fat, retinol, and cholesterol—as well as metabolites including hexoses, alanine, and valine, that are associated with MetS in a middle-aged Korean cohort. Among the 11 metabolite sub-pathways enriched across all analytical methods, arginine biosynthesis emerged as the most strongly associated with MetS, highlighting its potential role in vascular and metabolic dysfunction. In particular, hexose

was identified as a prominent marker, showing the strongest correlation with blood glucose, a core component of MetS. Fixed-effects modeling revealed distinct metabolite–nutrient relationships stratified by MetS status, with six unique pairs observed in the MetS group and 13 in the non-MetS group. These findings underscore the differential nature of nutrient–metabolite interactions in the context of MetS and support their potential utility as biomarkers for metabolic dysregulation. The distinct metabolic profiles and nutrient–metabolite associations observed in this Korean population should be interpreted within the broader global context. A recent meta-analysis of 28 million adults across six WHO regions (Africa, Americas, South-East Asia, Europe, Eastern Mediterranean, and Western Pacific) reported a global prevalence of elevated fasting glucose of 24.5% [1]. In our Korean cohort, 55.5% of individuals with MetS exhibited elevated fasting glucose levels, substantially exceeding the global average. Even among individuals without MetS, the prevalence was 17.0%, which is slightly lower than the global prevalence based on the general population, but still noteworthy. These findings reinforce the global relevance of dysregulated glucose metabolism as a key hallmark of MetS and underscore its particularly high burden in the Korean population. In this context, our observation of elevated hexose and its strong correlation with blood glucose further supports the central role of glycemic dysregulation across diverse populations, while also reflecting population-specific dietary influences. Future research should incorporate inter-individual variability, including sex-specific differences and genetic polymorphisms, which may modulate these interactions and further inform personalized dietary strategies. Sex hormones have been shown to influence amino acid metabolism [83], insulin sensitivity, and gut microbiota [84], all of which are relevant to MetS pathophysiology. Notably, both one-carbon metabolism and BCAA catabolism may be differentially regulated based on sex or genetic background. For instance, polymorphisms in *MTHFR* can impair folate metabolism and methylation efficiency [85], while variants in the *BCKDH* or *PPM1K* genes affect BCAA breakdown [86]. These inter-individual differences are critical for advancing precision nutrition and personalized medicine. For instance, individuals with genetic variants in BCAA metabolism may benefit from adjusted macronutrient intake, while those with impaired one-carbon metabolism may require tailored folate or B-vitamin supplementation [87, 88]. Integrating metabolomic and genomic data may support the development of more precise individualized nutrition beyond a one-size-fits-all approach. The SGD classifier demonstrated the highest predictive accuracy among the models tested (AUC = 0.84), but external validation was not conducted, limiting the generalizability of the results. Therefore, the generalizability of these predictive models remains limited, and future research involving independent cohorts is warranted to confirm their robustness.

Overall, this study highlights the value of an integrative multi-omics approach for elucidating how dietary factors modulate metabolic phenotypes and molecular pathways underlying MetS. Based on the observed associations, personalized dietary interventions for individuals with MetS may include BCAA-restricted diets to mitigate the metabolic impact of elevated BCAA and their catabolic byproducts, such as acylcarnitines, as well as targeted carbohydrate modulation, such as reducing high-hexose dietary sources, to improve glycemic control. Additionally, regulating the intake of niacin-rich proteins may be beneficial in managing amino acid and hexose metabolism, and targeted micronutrient supplementation, such as vitamin B6 to support glycine metabolism, could further enhance the metabolic environment in MetS.

Abbreviations

AAA	Aromatic amino acid
ADB	AdaBoost
AUC	Area under the curve
BCAA	Branched-chain amino acid
BMI	Body mass index
CI	Confidence interval
DNN	Deep neural network
EMR	Electronic medical record
FDR	False discovery rate
FFQ	Food frequency questionnaire
HDL	High-density lipoprotein
HMDB	Human Metabolome Database
KEGG	Kyoto Encyclopedia of Genes and Genomes
KNN	K-nearest neighbors
LASSO	Least absolute shrinkage and selection operator
LR	Logistic regression
MetS	Metabolic syndrome
NB	Naive Bayes
NO	Nitric oxide
PC	Phosphatidylcholine
PLP	Pyridoxal 5′-phosphate
PLS-DA	Partial least squares-discriminant analysis
RF	Random forest
ROS	Reactive oxygen species
SD	Standard deviation
SGD	Stochastic gradient descent
SVM	Support vector machine
VIP	Variable importance in projection

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12937-025-01189-3.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3. Supplementary Figure 1. Correlation network of metabolic syndrome components, metabolites, and nutrients. Statistically significant correlations between two metabolites are represented as edges, with positive correlations shown in blue and negative correlations in red. The color gradient and thickness of the edges are proportional to the absolute value of the correlation coefficient. *P*-values were adjusted for the false discovery rate using the Benjamini–Hochberg method. ADMA, asymmetric dimethylarginine; Ac-Orn, acetylornithine; Ala, alanine; Arg,

arginine; Asn, asparagine; Asp, aspartate; C0, free carnitine; C2, propionylcarnitine; C3, butyrylcarnitine; C5, isovalerylcarnitine; C8, octanoylcarnitine; Ca, calcium; Cit, citrulline; Fe, iron; Gln, glutamine; Glu, glutamate; Gly, glycine; H1, hexose; HDL, high-density lipoprotein; His, histidine; Ile, isoleucine; K, potassium; Leu, leucine; Lys, lysine; Met, methionine; Na, sodium; Orn, ornithine; P, phosphorus; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; Val, valine; Vit, vitamin.

Acknowledgements
This study was conducted using bioresources from the National Biobank of Korea and the Korea Disease Control and Prevention Agency (NBK-2024-019). The authors gratefully acknowledge Dr. Sunjin Moon for his valuable statistical advice.

Clinical trial number
Not applicable.

Authors' contributions
Authors'contributions Conceptualization: Shin D; Methodology: Kim M, Lee S, Shin D, Hur J; Formal analysis: Kim M, Lee S, Shin D, Hur J; Investigation: Kim M, Lee S, Shin D, Hur J; Writing–original draft: Kim M, Lee S; Writing–review & editing: Kim M, Lee S, Shin D, Hur J; Funding Acquisition: Shin D; Supervision: Shin D.

Funding
This work was supported by a grant from the National Research Foundation of Korea, funded by the Korean Government (MSIT) (RS-2024–00340086).

Data availability
Data from the Korean Genome and Epidemiology Study are available through the procedure described at https://biobank.nih.go.kr/eng/ (accessed on August 12, 2024). In compliance with institutional policies and data privacy regulations, the underlying code and datasets cannot be made publicly available. Nonetheless, all methodological steps and analytical workflows have been transparently described to ensure reproducibility.

Declarations

Ethics approval and consent to participate
The study protocol was reviewed and approved by the Institutional Review Board of Inha University, Korea, on March 20, 2024 (protocol number: 240307-1A).
All participants provided written informed consent prior to participation in the KoGES.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Food and Nutrition, Inha University, Incheon, Republic of Korea
[2]Department of Biomedical Sciences, School of Medicine and Health Sciences, University of North Dakota, Grand Forks, ND, USA

Received: 6 March 2025 / Accepted: 24 July 2025
Published online: 19 August 2025

References

1. Noubiap JJ, Nansseu JR, Lontchi-Yimagou E, Nkeck JR, Nyaga UF, Ngouo AT, et al. Geographic distribution of metabolic syndrome and its components in the general adult population: a meta-analysis of global data from 28 million individuals. Diabetes Res Clin Pract. 2022;188:109924.
2. Neeland IJ, Lim S, Tchernof A, Gastaldelli A, Rangaswami J, Ndumele CE, et al. Metabolic syndrome. Nat Rev Dis Primers. 2024;10(1):77.
3. Kim NH, Seo M-H, Jung JH, Do Han K, Kim MK, Kim NH. 2023 Diabetic kidney disease fact sheet in Korea. Diabetes Metab J. 2024;48(3):463–72.
4. Kim HJ, Kang DR, Kim JY, Kim W, Jeong YW, Chun K-H, et al. Metabolic syndrome fact sheet 2024: executive report. CardioMetabolic Syndrome J. 2024;4(2):70–80.
5. Allam-Ndoul B, Guénard F, Garneau V, Cormier H, Barbier O, Pérusse L, et al. Association between metabolite profiles, metabolic syndrome and obesity status. Nutrients. 2016;8(6):324.
6. Shi M, Han S, Klier K, Fobo G, Montrone C, Yu S, et al. Identification of candidate metabolite biomarkers for metabolic syndrome and its five components in population-based human cohorts. Cardiovasc Diabetol. 2023;22(1):141.
7. Kim M, Kim M, Han JY, Lee S-H, Jee SH, Lee JH. The metabolites in peripheral blood mononuclear cells showed greater differences between patients with impaired fasting glucose or type 2 diabetes and healthy controls than those in plasma. Diabetes Vasc Dis Res. 2017;14(2):130–8.
8. Zhu M, Han Y, Zhang Y, Zhang S, Wei C, Cong Z, et al. Metabolomics study of the biochemical changes in the plasma of myocardial infarction patients. Front Physiol. 2018;9:1017.
9. Kern HJ, Mitmesser SH. Role of nutrients in metabolic syndrome: a 2017 update. Nutr Diet Suppl. 2018. https://doi.org/10.2147/NDS.S148987.
10. Harrison S, Couture P, Lamarche B. Diet quality, saturated fat and metabolic syndrome. Nutrients. 2020;12(11):3232.
11. Abete I, Astrup A, Martínez JA, Thorsdottir I, Zulet MA. Obesity and the metabolic syndrome: role of different dietary macronutrient distribution patterns and specific nutritional components on weight loss and maintenance. Nutr Rev. 2010;68(4):214–31.
12. Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics data integration, interpretation, and its application. Bioinform Biol Insights. 2020. https://doi.org/10.1177/1177932219899051.
13. Noerman S, Landberg R. Blood metabolite profiles linking dietary patterns with health—toward precision nutrition. J Intern Med. 2023;293(4):408–32.
14. Brennan L, Hu FB, Sun Q. Metabolomics meets nutritional epidemiology: harnessing the potential in metabolomics data. Metabolites. 2021;11(10):709.
15. Wang H, Wang Y, Li X, Deng X, Kong Y, Wang W, et al. Machine learning of plasma metabolome identifies biomarker panels for metabolic syndrome: findings from the China suboptimal health cohort. Cardiovasc Diabetol. 2022;21(1):288.
16. Yang SJ, Kwak SY, Jo G, Song TJ, Shin MJ. Serum metabolite profile associated with incident type 2 diabetes in Koreans: findings from the Korean genome and epidemiology study. Sci Rep. 2018;8(1):8207.
17. Lee H, Jang HB, Yoo MG, Chung KS, Lee HJ. Protective effects of dietary MUFAs mediating metabolites against hypertension risk in the Korean genome and epidemiology study. Nutrients. 2019;11(8):1928.
18. Park S. Associations between APOE gene variants and metabolite levels in hypercholesterolemia: a metabolite GWAS study in a Korean cohort. Biomed Sci Lett. 2024;30(3):173–80.
19. Kim YJ, Lee HS, Kim YK, Park S, Kim JM, Yun JH, et al. Association of metabolites with obesity and type 2 diabetes based on FTO genotype. PLoS ONE. 2016;11(6):e0156612.
20. Hahn SJ, Kim S, Choi YS, Lee J, Kang J. Prediction of type 2 diabetes using genome-wide polygenic risk score and metabolic profiles: a machine learning analysis of population-based 10-year prospective cohort study. EBioMedicine. 2022. https://doi.org/10.1016/j.ebiom.2022.104383.
21. Bondia-Pons I, Martinez JA, de la Iglesia R, Lopez-Legarrea P, Poutanen K, Hanhineva K, et al. Effects of short-and long-term Mediterranean-based dietary treatment on plasma LC-QTOF/MS metabolic profiling of subjects with metabolic syndrome features: the Metabolic Syndrome Reduction in Navarra (RESMENA) randomized controlled trial. Mol Nutr Food Res. 2015;59(4):711–28.
22. Pallister T, Jackson MA, Martin TC, Zierer J, Jennings A, Mohney RP, et al. Hippurate as a metabolomic marker of gut microbiome diversity: modulation by diet and relationship to metabolic syndrome. Sci Rep. 2017;7(1):13670.
23. Wan M, Li Q, Lei Q, Zhou D, Wang S. Polyphenols and polysaccharides from *morus alba* L. fruit attenuate high-fat diet-induced metabolic syndrome modifying the gut microbiota and metabolite profile. Foods. 2022;11(12):1818.
24. Kim Y, Han BG, Group K. Cohort profile: the Korean genome and epidemiology study (KoGES) consortium. Int J Epidemiol. 2017;46(2):e20-e.
25. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation. 2002;106(25):3143–421.

26. Seo MH, Lee W-Y, Kim SS, Kang J-H, Kang J-H, Kim KK, et al. 2018 Korean society for the study of obesity guideline for the management of obesity in Korea. J Obes Metab Syndrome. 2019;28(1):40.
27. Hong AR, Lim S. Clinical characteristics of metabolic syndrome in Korea, and its comparison with other A sian countries. J Diabetes Investig. 2015;6(5):508–15.
28. Ahn Y, Kwon E, Shim J, Park M, Joo Y, Kimm K, et al. Validation and reproducibility of food frequency questionnaire for Korean genome epidemiologic study. Eur J Clin Nutr. 2007;61(12):1435–41.
29. Doo M, Kim Y. The consumption of dietary antioxidant vitamins modifies the risk of obesity among Korean men with short sleep duration. Nutrients. 2017;9(7):780.
30. Li FE, Zhang FL, Zhang P, Liu D, Liu HY, Guo ZN, et al. Sex-based differences in and risk factors for metabolic syndrome in adults aged 40 years and above in Northeast China: results from the cross-sectional China national stroke screening survey. BMJ Open. 2021;11(3):e038671.
31. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30.
32. Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. HMDB 3.0—the human metabolome database in 2013. Nucleic Acids Res. 2012;41(D1):D801–7.
33. Evans ED, Duvallet C, Chu ND, Oberst MK, Murphy MA, Rockafellow I, et al. Predicting human health from biofluid-based metabolomics using machine learning. Sci Rep. 2020;10(1):17635.
34. Galal A, Talal M, Moustafa A. Applications of machine learning in metabolomics: disease modeling and classification. Front Genet. 2022;13:1017340.
35. Yagin FH, Colak C, Algarni A, Algarni A, Al-Hashem F, Ardigò LP. Explainable boosting machines identify key metabolomic biomarkers in rheumatoid arthritis. Medicina. 2025;61(5):833.
36. Gong Y, Ding W, Wang P, Wu Q, Yao X, Yang Q. Evaluating machine learning methods of analyzing multiclass metabolomics. J Chem Inf Model. 2023;63(24):7628–41.
37. Kim H, Heo JH, Lim DH, Kim Y. Development of a metabolic syndrome classification and prediction model for Koreans using deep learning technology: the Korea National Health and Nutrition Examination Survey (KNHANES) (2013–2018). Clin Nutr Res. 2023;12(2):138.
38. Du Y, Rafferty AR, McAuliffe FM, Wei L, Mooney C. An explainable machine learning-based clinical decision support system for prediction of gestational diabetes mellitus. Sci Rep. 2022;12(1):1170.
39. Collins GS, Moons KG, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+ AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385.
40. Das SK, Ainsworth HC, Dimitrov L, Okut H, Comeau ME, Sharma N, et al. Metabolomic architecture of obesity implicates metabolonic lactone sulfate in cardiometabolic disease. Mol Metab. 2021;54:101342.
41. Mahajan UV, Varma VR, Huang C-W, An Y, Tanaka T, Ferrucci L, et al. Blood metabolite signatures of metabolic syndrome in two cross-cultural older adult cohorts. Int J Mol Sci. 2020;21(4):1324.
42. Paneque A, Fortus H, Zheng J, Werlen G, Jacinto E. The hexosamine biosynthesis pathway: regulation and function. Genes. 2023;14(4):933.
43. Ntzouvani A, Nomikos T, Panagiotakos D, Fragopoulou E, Pitsavos C, McCann A, et al. Amino acid profile and metabolic syndrome in a male Mediterranean population: a cross-sectional study. Nutr Metab Cardiovasc Dis. 2017;27(11):1021–30.
44. Hu W, Sun L, Gong Y, Zhou Y, Yang P, Ye Z, et al. Relationship between branched-chain amino acids, metabolic syndrome, and cardiovascular risk profile in a chinese population: a cross-sectional study. Int J Endocrinol. 2016;2016(1):8173905.
45. Taghizadeh H, Emamgholipour S, Hosseinkhani S, Arjmand B, Rezaei N, Dilmaghani-Marand A, et al. The association between acylcarnitine and amino acids profile and metabolic syndrome and its components in Iranian adults: data from STEPs 2016. Front Endocrinol. 2023;14: 1058952.
46. Li X, Sun L, Zhang W, Li H, Wang S, Mu H, et al. Association of serum glycine levels with metabolic syndrome in an elderly Chinese population. Nutr Metab. 2018;15:1–9.
47. Julibert A, del Mar Bibiloni M, Tur JA. Dietary fat intake and metabolic syndrome in adults: a systematic review. Nutr Metab Cardiovasc Dis. 2019;29(9):887–905.
48. Park S, Ahn J, Lee B-K. Very-low-fat diets may be associated with increased risk of metabolic syndrome in the adult population. Clin Nutr. 2016;35(5):1159–67.
49. Yoo H, Park K. Association between phytochemical index and osteoporosis in women: a prospective cohort study in Korea. Nutrients. 2023;15(7): 1605.
50. Assumpção CR, Brunini TM, Pereira NR, Godoy-Matos AF, Siqueira MA, Mann GE, et al. Insulin resistance in obesity and metabolic syndrome: is there a connection with platelet l-arginine transport? Blood Cells Mol Dis. 2010;45(4):338–42.
51. Mone P, Pansini A, Jankauskas SS, Varzideh F, Kansakar U, Lombardi A, et al. L-arginine improves cognitive impairment in hypertensive frail older adults. Front Cardiovasc Med. 2022;9: 868521.
52. Hu S, Han M, Rezaei A, Li D, Wu G, Ma X. L-arginine modulates glucose and lipid metabolism in obesity and diabetes. Curr Protein Pept Sci. 2017;18(6):599–608.
53. Lutchmansingh FK, Hsu JW, Bennett FI, Badaloo AV, McFarlane-Anderson N, Gordon-Strachan GM, et al. Glutathione metabolism in type 2 diabetes and its relationship with microvascular complications and glycemia. PLoS ONE. 2018;13(6):e0198626.
54. Díaz-Flores M, Cruz M, Duran-Reyes G, Munguia-Miranda C, Loza-Rodríguez H, Pulido-Casas E, et al. Oral supplementation with glycine reduces oxidative stress in patients with metabolic syndrome, improving their systolic blood pressure. Can J Physiol Pharmacol. 2013;91(10):855–60.
55. Xie W, Wood AR, Lyssenko V, Weedon MN, Knowles JW, Alkayyali S, et al. Genetic variants associated with glycine metabolism and their role in insulin sensitivity and type 2 diabetes. Diabetes. 2013;62(6):2141–50.
56. Mihalik SJ, Goodpaster BH, Kelley DE, Chace DH, Vockley J, Toledo FG, et al. Increased levels of plasma acylcarnitines in obesity and type 2 diabetes and identification of a marker of glucolipotoxicity. Obesity. 2010;18(9):1695–700.
57. Strand E, Pedersen ER, Svingen GF, Olsen T, Bjørndal B, Karlsson T, et al. Serum acylcarnitines and risk of cardiovascular death and acute myocardial infarction in patients with stable angina pectoris. J Am Heart Assoc. 2017;6(2):e003620.
58. Lemos GdO, Torrinhas RS, Waitzberg DL. Nutrients, physical activity, and mitochondrial dysfunction in the setting of metabolic syndrome. Nutrients. 2023;15(5): 1217.
59. Romani M, Berger MM, D'Amelio P. From the bench to the bedside: branched amino acid and micronutrient strategies to improve mitochondrial dysfunction leading to sarcopenia. Nutrients. 2022;14(3):483.
60. Liang M, Wang Z, Li H, Cai L, Pan J, He H, et al. L-Arginine induces antioxidant response to prevent oxidative stress via stimulation of glutathione synthesis and activation of Nrf2 pathway. Food Chem Toxicol. 2018;115:315–28.
61. Wu X, Liu C, Yang S, Shen N, Wang Y, Zhu Y, et al. Glycine-serine-threonine metabolic axis delays intervertebral disc degeneration through antioxidant effects: an imaging and metabonomics study. Oxid Med Cell Longev. 2021;2021(1):5579736.
62. Liu S, Li L, Lou P, Zhao M, Wang Y, Tang M, et al. Elevated branched-chain α-keto acids exacerbate macrophage oxidative stress and chronic inflammatory damage in type 2 diabetes mellitus. Free Radic Biol Med. 2021;175:141–54.
63. Rutkowsky JM, Knotts TA, Ono-Moore KD, McCoin CS, Huang S, Schneider D, et al. Acylcarnitines activate proinflammatory signaling pathways. Am J Physiol Endocrinol Metab. 2014;306(12):E1378–87.
64. Schieber M, Chandel NS. ROS function in redox signaling and oxidative stress. Curr Biol. 2014;24(10):R453–62.
65. Amorim R, Simões IC, Teixeira J, Cagide F, Potes Y, Soares P, et al. Mitochondria-targeted anti-oxidant AntiOxCIN4 improved liver steatosis in western diet-fed mice by preventing lipid accumulation due to upregulation of fatty acid oxidation, quality control mechanism and antioxidant defense systems. Redox Biol. 2022;55:102400.
66. Choi S-W, Friso S. Modulation of DNA methylation by one-carbon metabolism: a milestone for healthy aging. Nurs Res Pract. 2023;17(4):597–615.
67. Wang XP, Yan D, Jin XP, Zhang WY, Shi T, Wang X, et al. The role of amino acid metabolism alterations in acute ischemic stroke: From mechanism to application. Pharmacol Res. 2024;207:107313.
68. Gojda J, Cahova M. Gut microbiota as the link between elevated BCAA serum levels and insulin resistance. Biomolecules. 2021;11(10):1414.
69. Floegel A, Von Ruesten A, Drogan D, Schulze M, Prehn C, Adamski J, et al. Variation of serum metabolites related to habitual diet: a targeted metabolomic approach in EPIC-Potsdam. Eur J Clin Nutr. 2013;67(10):1100–8.
70. Gar C, Rottenkolber M, Prehn C, Adamski J, Seissler J, Lechner A. Serum and plasma amino acids as markers of prediabetes, insulin resistance, and incident diabetes. Crit Rev Clin Lab Sci. 2018;55(1):21–32.

71. Karlsson T, Winkvist A, Rådjursöga M, Ellegård L, Pedersen A, Lindqvist HM. Identification of single and combined serum metabolites associated with food intake. Metabolites. 2022;12(10): 908.
72. Low DY, Mina TH, Sadhu N, Wong KE, Jain PR, Dalan R, et al. Metabolic variation reflects dietary intake in a multi-ethnic Asian population. medRxiv. 2023:2023.12. 04.23299350.
73. Meyer-Ficca M, Kirkland JB. Niacin. Adv Nutr. 2016;7(3):556–8.
74. Henninger J, Eliasson B, Smith U, Rawshani A. Identification of markers that distinguish adipose tissue and glucose and insulin metabolism using a multimodal machine learning approach. Sci Rep. 2021;11(1):17050.
75. Aichler M, Borgmann D, Krumsiek J, Buck A, MacDonald PE, Fox JEM, et al. N-acyl taurines and acylcarnitines cause an imbalance in insulin synthesis and secretion provoking β cell dysfunction in type 2 diabetes. Cell Metabolism. 2017;25(6):1334-47. e4.
76. Caldow MK, Ham DJ, Godeassi DP, Chee A, Lynch GS, Koopman R. Glycine supplementation during calorie restriction accelerates fat loss and protects against further muscle loss in obese mice. Clin Nutr. 2016;35(5):1118–26.
77. Samuel VT, Shulman GI. The pathogenesis of insulin resistance: integrating signaling pathways and substrate flux. J Clin Invest. 2016;126(1):12–22.
78. Kosacka J, Woidt K, Toyka KV, Paeschke S, Klöting N, Bechmann I, et al. The role of dietary non-heme iron load and peripheral nerve inflammation in the development of peripheral neuropathy (PN) in obese non-diabetic leptin-deficient ob/ob mice. Neurol Res. 2019;41(4):341–53.
79. Abu AlSel BT, Mahmoud AA, Hamed EO, Hakim NA, Sindi AA, Jawad NM, et al. Iron homeostasis-related parameters and hepcidin/ferritin ratio: emerging sex-specific predictive markers for metabolic syndrome. Metabolites. 2024;14(9):473.
80. Yang H. Implementation of artificial intelligence in primary care. Korean J Fam Pract. 2023;13(2):73–9.
81. Gibson R, Lau C-HE, Loo RL, Ebbels TM, Chekmeneva E, Dyer AR, et al. The association of fish consumption and its urinary metabolites with cardiovascular risk factors: the International Study of Macro-/Micronutrients and Blood Pressure (INTERMAP). Am J Clin Nutr. 2020;111(2):280–90.
82. Miles FL, Orlich MJ, Mashchak A, Chandler PD, Lampe JW, Duerksen-Hughes P, et al. The biology of veganism: plasma metabolomics analysis reveals distinct profiles of vegans and non-vegetarians in the Adventist Health Study-2 cohort. Nutrients. 2022;14(3):709.
83. Palacios-González B, León-Reyes G, Rivera-Paredez B, Ibarra-González I, Vela-Amieva M, Flores YN, et al. Targeted metabolomics revealed a sex-dependent signature for metabolic syndrome in the Mexican population. Nutrients. 2022;14(18):3678.
84. Santos-Marcos JA, Mora-Ortiz M, Tena-Sempere M, Lopez-Miranda J, Camargo A. Interaction between gut microbiota and sex hormones and their relation to sexual dimorphism in metabolic diseases. Biol Sex Differ. 2023;14(1): 4.
85. Liew S-C, Gupta ED. Methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism: epidemiology, metabolism and the associated diseases. Eur J Med Genet. 2015;58(1):1–10.
86. White PJ, McGarrah RW, Grimsrud PA, Tso SC, Yang WH, Haldeman JM, et al. The BCKDH kinase and phosphatase integrate BCAA and lipid metabolism via regulation of ATP-citrate lyase. Cell Metabolism. 2018;27(6):1281-93. e7.
87. Goni L, Qi L, Cuervo M, Milagro FI, Saris WH, MacDonald IA, et al. Effect of the interaction between diet composition and the PPM1K genetic variant on insulin resistance and β cell function markers during weight loss: results from the nutrient gene interactions in human obesity: implications for dietary guidelines (NUGENOB) randomized trial. Am J Clin Nutr. 2017;106(3):902–8.
88. He Q, Li J. The evolution of folate supplementation–from one size for all to personalized, precision, poly-paths. J Transl Intern Med. 2023;11(2):128–37.

Publisher's Note