Differential Metabolites in Chinese Autistic Children: A Multi-Centre Study Based on Urinary ^{1}H-NMR Metabolomics Analysis

Yu Ma
Children's Hospital of Fudan University
Hao Zhou
Guizhou Provincial People's Hospital
Chunpei Li
Children's Hospital of Fudan University
Xiaobing Zou
The Third Affiliated Hospital, Sun Yat-Sen University
Xuerong Luo
The Second Xiangya Hospital of Central South University
Lijie Wu
School of Public Health, Harbin Medical University
Tingyu Li
Children's Hospital of Chongqing Medical University
Xiang Chen
The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University
Meng Mao
Chengdu Women and Children's Hospital
Yi Huang
West China Hospital of Sichuan University
Erzhen Li
Capital Institute of Paediatrics
Yanpeng An
School of Life Sciences, Fudan University
Lili Zhang
Children's Hospital of Fudan University
Tianqi Wang
Children's Hospital of Fudan University
Xiu Xu
Children's Hospital of Fudan University
Weili Yan
Children's Hospital of Fudan University
Yonghui Jiang
Yale School of Medicine
Yi Wang (yiwang@shmu.edu.cn)
Children's Hospital of Fudan University https://orcid.org/0000-0002-7051-3739

Research

Keywords: Autism spectrum disorder, 1H-NMR analysis, metabolomics

Posted Date: August 31st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-64437/v1

Version of Record: A version of this preprint was published at Frontiers in Psychiatry on May 11th, 2021. See the published version at https://doi.org/10.3389/fpsyt.2021.624767.

Abstract

Background: Autism spectrum disorder (ASD) is a group of early-onset neurodevelopmental disorders. However, there is no valuable biomarker for the early diagnosis of ASD. Our large-scale and multi-centre study aims to identify metabolic variations between ASD and healthy children and to investigate differential metabolites and associated pathogenic mechanisms.

Methods: 117 autistic children and 119 healthy children were recruited from research centres of 7 cities. Urine samples were assayed by ^{1}H-NMR metabolomics analysis to detect metabolic variations. Multivariate statistical analysis, including principal component analysis (PCA) and orthogonal projection to latent structure discriminant analysis (OPLS-DA), as well as univariate analysis were used to assess differential metabolites between the ASD and control groups. The differential metabolites were further analysed by receiver operating characteristics (ROC) curve analysis and metabolic pathways analysis.

Results: Compared with the control group, the ASD group showed higher levels of glycine, guanidinoacetic acid, creatine, hydroxyphenylacetylglycine, phenylacetylglycine and formate and lower levels of 3-aminoisobutanoic acid, alanine, taurine, creatinine, hypoxanthine and N-methylnicotinamide. ROC curve showed relatively significant diagnostic values for hypoxanthine (area under the curve (AUC) = 0.657, 95% CI 0.588 to 0.726), creatinine (AUC = 0.639, 95% CI 0.569 to 0.709), creatine (AUC = 0.623, 95% CI 0.552 to 0.694), N-methylnicotinamide (AUC = 0.595, 95% CI 0.523 to 0.668) and guanidinoacetic acid (AUC = 0.574, 95% CI 0.501 to 0.647) in the ASD group. Combining the metabolites creatine, creatinine and hypoxanthine, the AUC of the ROC curve reached 0.720 (95% CI 0.659 to 0.777). Significantly altered metabolite pathways associated with differential metabolites were glycine, serine and threonine metabolism, arginine and proline metabolism, and taurine and hypotaurine metabolism.

Limitations: Due to the restriction of research centres, inconsistent urine collection times may have affected the quality of metabolic analyses. Moreover, it is necessary to compare metabolite levels that vary with ASD severity to better clarify the pathogenesis of ASD.

Conclusions: Urinary amino acid metabolites significantly altered in children with ASD. Amino acid metabolic pathways might play important roles in the pathogenic mechanisms of ASD.

Clinical Trial registration: National Study on Autism Spectrum Disorder in China (NCT02200679). Registered July 24, 2014. https://www.clinicaltrials.gov/ct2/show/NCT02200679

Background

Autism spectrum disorder (ASD) is a group of early-onset neurodevelopmental disorders characterized by social communication difficulties, narrow interests, and repetitive stereotyped behaviours [1]. In Western countries, the prevalence of ASD has increased in recent years, ranging from 1.57% in 2009 to 1.69% in 2018 [2–3]. The prevalence of ASD among children aged 6 to 12 years is approximately 0.70% in China [4]. The large cost associated, mainly consisting of special education services and parental productivity loss, has caused a heavy burden to society and families [5]. Despite the lack of effective drug treatments, several studies highlight the potential benefits of early diagnosis and parent-mediated interventions, which have to some extent improved children's social and communicative abilities [6–7]. Recently, genetic testing has increased the rate of early diagnosis of ASD to 20%-30% [8–10]. However, early diagnosis remains a challenge for nongenetic ASD, which is mainly based on combining clinician observation with caregiver reports [11]. Currently, there is an urgent need to find valuable biomarkers for the early diagnosis of ASD.

Some metabonomic studies have indicated the presence of elevated biomarkers in blood and urine samples from ASD patients, and these biomarkers include pyruvate, lactate, and mitochondrial-related enzymes [12–14]. Major analytical techniques for metabolomics are nuclear magnetic resonance spectroscopy (NMR) and chromatography, including gas chromatography–mass spectrometry (GC-MS) and liquid chromatography–mass spectrometry (LC-MS). The advantages of ^{1}H-NMR are that the sample preprocessing is simple and non-destructive and that the detection of metabolites is comprehensive [15–16]. In 2010, Yap et al. first used urinary ^{1}H-NMR analysis to detect potential biomarkers for ASD [17]. Over the last few years, with urinary ^{1}H-NMR

analysis, some discriminating metabolites have been identified in ASD patients: higher levels of hippurate, glycine, tryptophan and D-threitol and lower levels of glutamate, creatine, valine, betaine and 3-methylhistidine. Further analyses indicated possible pathogenic mechanisms involving gut microbial metabolism, oxidative stress conditions and amino acid metabolism [18–19]. Overall, ^{1}H-NMR analysis shows great potential for the identification of biochemical signatures and investigation of the disease mechanisms of ASD. However, previous ^{1}H-NMR analyses lack large-scale sample sizes to confirm the significance and connection between metabolites and ASD. We aimed to conduct a large-scale and multi-centre study to identify metabolic variations between ASD and control groups through urinary ^{1}H-NMR metabolomics analysis and to investigate potential biological mechanisms related to candidate biomarkers.

Methods

1. Participants

The study was conducted from January 2014 to December 2016. All the participants were recruited from research centres of 7 cities (Shanghai, Guangzhou, Changsha, Chongqing, Chengdu, Wenzhou, and Beijing). Participants were drawn from ASD and control group.

Autistic children from both hospitals and local autism rehabilitation of each research centre were enrolled in the ASD group. The inclusion criteria for ASD group were: (1) Children aged 2 to 18 years; (2) No limitation on the gender; (3) The diagnosis of ASD was based on the Diagnostic and Statistical Manual of Mental Disorders 5 (DSM-5) criteria [1] and confirmed with the Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview-Revised (ADI-R) criteria by trained clinical psychiatrists from each research centre; (4) Urine sample was available. Exclusion criteria for ASD group were: (1) Symptomatic autism (such as Rett syndrome and fragile X syndrome); (2) Other mental illness (such as attention-deficit hyperactivity disorder); (3) Other neurological disorders (such as epilepsy and central nervous system infections); (4) Inherited metabolic diseases; (5) History of brain injury; (6) Taking non-essential drug or dietary supplement before (72 hours) and during sample collection.

Healthy children from the health examination centre of each research centre were enrolled in the control group. The inclusion criteria for control group were: (1) Children with no abnormality in health examination and typical development; (2) Age- and sex-matching with ASD group; (3) Urine sample was available. Exclusion criteria for control group were: (1) Clinical evidence of ASD diagnosis; (2) mental illness (such as attention-deficit hyperactivity disorder); (3) Neurological disorders (such as epilepsy and central nervous system infections); (4) Inherited metabolic diseases; (5) History of brain injury; (6) Taking non-essential drug or dietary supplement before (72 hours) and during sample collection.

After all eligible participants and their parents provided informed consent, they were invited to participate in the study. The study was approved by the institutional ethics committee at the Children's Hospital of Fudan University.

2. Sample collection

First morning urine specimens were collected from all participants during the research period. Samples were collected in 15 mL urine collection tubes without preservative. Each sample was centrifuged and aliquoted into 1.5 mL EP tubes. Afterwards, samples were numbered ("1" represents the ASD group, and "2" represents the control group) and stored at -80°C immediately until ^{1}H-NMR analysis.

3. ^{1}H-NMR spectroscopy experiments and data processing

3.1 Sample preparation

A 500 μL urine sample was added to a 1.5 mL EP tube prefilled with 14 μL KF (5 M) solution. After vortexing, the sample was allowed to rest for 10 minutes, followed by centrifugation (12000 rpm, 4°C) for 10 minutes. A total of 450 μL liquid supernatant was added to an NMR tube preloaded with 10 μL EDTA-d12 (0.1 M). The NMR tube cap was covered and mixed by hand. Finally,

45 μL Na^+/K^+ buffer (1.5 M, pH = 7.40) was added to the NMR tube, which was mixed by hand and then placed in the NMR spectrometer for data collection. The sample preparation process for ^{1}H-NMR analysis is shown in Additional file 2: Fig. S1.

3.2 ^{1}H-NMR spectroscopy experiments

All ^{1}H-NMR spectroscopy experiments were performed at 298 K using a Bruker AVIII 600 MHz NMR spectrometer (Bruker BioSpin, Germany) with a proton resonance frequency of 600.13 MHz.

The NOESYGPPR1D pulse sequence [RD-90°-t1-90°-tm-90°-ACQ] was used to collect the spectra. The parameters were set as follows: spectral width (SW) was 20 ppm, recycle delay (RD) was 2 s, mixing time (t_m) was 80 ms, t_1 was 4 μs, 90° pulse length was 14.8 s, data time was 1.36 s, data points were 32 K, and free induction decay (FID) accumulation was 64 times.

3.3 Data processing

Spectra were processed using MestReNova software (MestReNova 8.1, Spain). The ^{1}H-NMR FID signals were multiplied by an exponential function equivalent to a line broadening of 1 Hz before performing an automatic Fourier transformation. The phase distortion and baseline of each spectrum were manually adjusted. The internal standard trimethylsilyl propanoic acid (TSP, δ=0 ppm) was used as the baseline to calibrate the chemical shifts. The regions of the ^{1}H-NMR spectra (δ 0.3-9.5 ppm) were divided into consecutive integrated spectral regions of equal width (δ 0.004 ppm). The spectral region of the water (δ 4.71-5.055 ppm) and urea (δ 5.6-6.12 ppm) peaks were removed from each spectrum to minimize variations caused by the presaturation of the residual water and urea resonances. Mnova software was used to correct the spectra with obvious chemical shifts after the integration. Metabolites were assigned by referencing the values for chemical shifts in J-resolved (JRES), COSY, TOCSY, HSQC, and HMBC spectra and literature reports. A series of 2D NMR spectra were acquired for selected samples.

To eliminate the instrument differences of sensitivity and stability and to reduce the analysis errors caused by the concentration differences of the samples, two normalization methods were performed. (1) Creatinine normalization: the ^{1}H-NMR spectra were normalized by using the creatinine methylene resonance (δ= 4.05 ppm) as a reference. (2) Total area normalization [20]: the integrated area in each bucket was normalized by the total sum of peak intensities to eliminate the effects of variable concentration among different samples.

4. Data and statistical analysis

4.1 Clinical characteristics of participants

Difference in age between the ASD and control groups was evaluated by Student's t-test when the distribution was normal or the Mann-Whitney U test when it was skewed. Difference in sex was investigated using chi-square test. Statistical analyses were performed by using the SPSS statistical package program (version 20, SPSS Inc., Chicago, IL, USA), and P <0.05 was considered statistically significant.

4.2 Multivariate analysis

The normalized data were imported into the SIMCA-P+ software package (version 13.0, Umetrics, Sweden) for multivariate statistical analysis, including principal component analysis (PCA) and orthogonal projection to latent structure discriminant analysis (OPLS-DA) [21-24]. PCA was first used to observe the overall distribution among samples and the stability of the whole analysis process. Abnormal data were removed according to the overall aggregation trend in all the samples, after which OPLS-DA was used to distinguish the overall difference in the metabolic profile and to find differential metabolites between the groups.

To prevent model overfitting, an internal validation method was used to verify the validity of the model. The OPLS-DA model was validated by a 7-fold cross-validation (CV) [25]. R^2 and Q^2 were two parameters to assess the quality of the model. OPLS-DA was further validated by variance analysis of the cross-validated residuals (CV-ANOVA) [26], and the model was considered valid at P < 0.05.

After multiplying the loading value of each variable with its standard deviation, backtracking conversion of the data was performed. Then, the data were assessed by multivariate analysis and imported into mapping software based on MATLAB (version 7.1, USA) to plot the loading diagram of the correlation coefficient. The Pearson correlation coefficient represented the linear correlation between the variable and the first principal component score value of the OPLS-DA model and was used to determine whether the variation contributed significantly to the intergroup differentiation. The significance was evaluated by the threshold value of the absolute correlation coefficient, which was determined according to the confidence interval of the sample size. If the absolute value of the Pearson correlation coefficient of the variable was higher than the threshold value ($P < 0.05$), the content of the variable was considered significantly different between groups.

4.3Univariate analysis

Differential metabolites between the two groups were selected for univariate analysis. Data from the total area normalized peak area are expressed as the mean ± SD. An independent sample t-test was used for comparing the two groups when the distribution was normal. The non-normal distribution data were evaluated by Mann-Whitney U test. $P < 0.05$ was considered statistically significant.

4.4 Receiver operating characteristics (ROC) curve analysis

The sensitivity and specificity of metabolites with significant differences between the two groups in the diagnosis of ASD were evaluated using a ROC curve analysis. ROC curves were generated by using MedCalc statistical software (version 19.1.7, Belgium). The area under the curve (AUC) was used to measure the overall degree of identification power. An AUC greater than 0.7 was considered acceptable. Optimal cut-off points were determined by maximizing the Youden's J index (J = sensitivity + specificity - 1). The sensitivity, specificity, false-negative rate (FNR), false-positive rate (FPR), positive predictive value (PPV), negative predictive value (NPV), likelihood ratio positive (LR+), and likelihood ratio negative (LR–) were calculated to compare the diagnostic accuracy of the metabolites. Further analysis was performed after each group was stratified by sex and age. Logistic regression was used to analyse the combined metabolites.

4.5 Metabolic pathways and network analysis

MetaboAnalyst (version 4.0, https://www.metaboanalyst.ca/) was used for pathway and network analyses [27]. All differential metabolites were imported into the pathway analysis module to obtain matched pathways according to the *P* values from the pathway enrichment analysis and pathway impact values from the pathway topology analysis. Metabolite set enrichment analysis (MESA) was performed to identify significantly affected metabolic pathways. A metabolite-gene-disease interaction network was established to detect the connections of differential metabolites and associated pathways.

Results

1. Characteristics of the ASD and control groups

A total of 117 children with ASD were enrolled in the ASD group, and 119 healthy children were enrolled in the control group. The mean ages of the ASD and control groups were 10.12 ± 2.60 and 9.90 ± 1.73, respectively, with no significant difference between them ($P = 0.468$). The male to female ratios were 4.82:1 in the ASD group and 3.58:1 in the control group, with no significant difference between them ($P = 0.388$).

2. ^{1}H-NMR spectrum of urine samples

^{1}H-NMR spectra of urine samples from all participants were collected. A typical ^{1}H-NMR spectrum is shown in Fig. 1. The keys for metabolites in Fig.1 are given in Additional file 1: Table S1. A total of 39 metabolites were identified in the ^{1}H NMR spectra of urine samples.

3. Multivariate analysis of ^{1}H-NMR spectra of urine samples

3.1 Creatinine normalization analysis

After creatinine normalization, PCA was performed on ^{1}H-NMR spectra of urine samples. Discriminant variables obtained from the PCA score plot (Fig. 2A) were $R^2X = 0.403$ and $Q^2 = 0.135$. The OPLS-DA score plot (Fig. 2B) showed $R^2X = 0.257$ and $Q^2 = 0.0138$. CV-ANOVA of the OPLS-DA model indicated that the model was not valid ($P = 0.523$).

3.2 Total area normalization analysis

The ^{1}H-NMR spectra of the two groups were analysed by PCA after total area normalization. The parameters of the PCA score plot (Fig. 3A) were $R^2X = 0.366$ and $Q^2 = 0.0656$. The OPLS-DA score plot (Fig. 3B) showed $R^2X = 0.274$, $Q^2 = 0.0565$ and CV-ANOVA $P = 0.009$. These results suggested that the model was valid.

4. Differential metabolites between the ASD and control groups

To distinguish the metabolic differences between the two groups, a polychromatic correlation coefficient loading plot (Fig. 4) was drawn. The colour of the polychromatic loading plot was encoded by the absolute value of the correlation coefficient. The warmer the colour is, the higher the absolute value of the correlation coefficient and the greater the contribution to the intergroup differentiation. The threshold of the absolute value of the Pearson's correlation coefficient was determined to be 0.182. The variables corresponding to the correlation coefficient with an absolute value greater than 0.182 contributed significantly to the intergroup differentiation ($P < 0.05$). Differential metabolites identified by the Pearson's correlation coefficient were showed in Table 1. Compared with the control group, the ASD group showed higher levels of glycine, guanidinoacetic acid, creatine, hydroxyphenylacetylglycine, phenylacetylglycine and formate and lower levels of 3-aminoisobutanoic acid, alanine, taurine, creatinine, hypoxanthine and N-methylnicotinamide.

The normalized peak areas of differential metabolites in each group are summarized in Fig. 5. A total of 6 metabolites, including 3-aminoisobutanoic acid ($P = 0.0425$), creatine ($P = 0.0009$), creatinine ($P < 0.0001$), hypoxanthine ($P < 0.0001$), formate ($P = 0.0267$) and N-methylnicotinamide ($P = 0.0149$), showed significant differences in the normalized peak areas between the two groups (Additional file 1: Table S2).

5. Candidate biomarkers and potential biological mechanisms

5.1 The sensitivity and specificity of metabolites in the diagnosis of ASD

The diagnostic accuracies of differential metabolites in the two groups were evaluated by ROC curve analysis (Table 2, Additional file 3: Fig. S2). The ROC curve showed relatively significant diagnostic values of hypoxanthine (AUC = 0.657, 95% CI 0.588 to 0.726), creatinine (AUC = 0.639, 95% CI 0.569 to 0.709), creatine (AUC = 0.623, 95% CI 0.552 to 0.694), N-methylnicotinamide (AUC = 0.595, 95% CI 0.523 to 0.668) and guanidinoacetic acid (AUC = 0.574, 95% CI 0.501 to 0.647) for ASD (Additional file 4: Fig. S3). The AUC of ROC analysis of the creatine/creatinine ratio was 0.6480 (95% CI 0.579 to 0.718). For each metabolite, there was no significant difference in AUCs between males and females (Additional file 1: Table S3). Compared with age stratification of 7-9 years old, the metabolites guanidinoacetic acid and creatine showed significantly higher diagnostic accuracy for ASD in the age stratification of 13-15 years old (AUC of guanidinoacetic acid = 0.802, 95% CI 0.566 to 0.944, $P = 0.0282$; AUC of creatine = 0.823, 95% CI 0.589 to 0.955, $P = 0.0344$. Results shown in Additional file 1: Table S4). By combining the metabolites of creatine, creatinine and hypoxanthine, the AUC of the ROC curve reached 0.720 (95% CI 0.659 to 0.777), with a sensitivity of 80.34% and specificity of 52.94%.

5.2 Correlated metabolic pathways and networks

The main metabolic pathways associated with differential metabolites are shown in Additional file 5: Fig. S4. According to the bubble plot of the metabolic pathway impact, there were significant metabolite changes in the glycine, serine and threonine metabolism, glyoxylate and dicarboxylate metabolism and taurine and hypotaurine metabolism pathways (Fig. 6A). The plot of metabolite set enrichment analysis (MESA) listed the significant enrichment pathways of differential metabolites (Fig. 6B). Glycine, serine and threonine metabolism, primary bile acid biosynthesis and aminoacyl-tRNA biosynthesis were the pathways

with the most significant enrichment. The metabolite-gene-disease interaction network provides a global view of the connection of the differential metabolites and the potential functional relationships among metabolites, connected genes, and target diseases (Additional file 6: Fig. S5).

Discussion

Complex aetiologies and atypical symptoms pose very large challenges to the early diagnosis of ASD. Urinary ^{1}H-NMR analysis provides a fast and comprehensive assessment to detect potential biomarkers of ASD [28]. Our study was based on the ^{1}H-NMR analysis of urine samples from ASD and control groups to identify candidate metabolites and associated pathogenic mechanisms. Compared with other studies of urinary 1H-NMR metabolomics analysis of ASD, our study sample was large size, multi-centre, and representative. Besides, our study used two normalization methods, creatinine normalization and total area normalization. The analysis using creatinine normalization showed no significant differences between the two groups, though the total area normalization did detect a difference. The variation of creatinine was confirmed by the total area normalization. Whiteley et al. also found that excretion of urinary creatinine in the group of pervasive developmental disorders, which included ASD, was significantly lower than controls [29]. The abnormal of creatinine metabolism might be caused by rigidity in food choice and various exclusion diets associated with ASD [30]. Therefore, total area normalization was more suitable for our study.

The results obtained from the ^{1}H-NMR analysis revealed that the levels of glycine, guanidinoacetic acid, creatine, hydroxyphenylacetylglycine, phenylacetylglycine and formate were higher in the ASD group than those in the control group. Moreover, the levels of 3-aminoisobutanoic acid, alanine, taurine, creatinine, hypoxanthine and N-methylnicotinamide were lower in the ASD group than those in the control group. The normalized peak areas of 3-aminoisobutanoic acid, creatine, creatinine, hypoxanthine, formate and N-methylnicotinamide differed significantly between the two groups. The metabolite levels of glycine, taurine, creatine and creatinine were consistent with those from previous reports [19, 31]. Our study is the first to detect variations in the metabolites of 3-aminoisobutanoic acid, hydroxyphenylacetylglycine, hypoxanthine and N-methylnicotinamide in urine samples from children with ASD. The results of ROC analysis indicated that creatine, creatinine and hypoxanthine have the potential to be biomarkers for the diagnosis of ASD. The creatine/creatinine ratio slightly improved the diagnostic accuracy. Recent research has reported significant female-related alterations of creatine and creatinine, and the creatinine/creatine ratio might be a good predictor of ASD in female subjects [32]. We further found that combining the metabolites creatine, creatinine and hypoxanthine as a diagnostic indicator can largely improve the diagnostic accuracy for ASD.

The extension study of the metabolic pathway analysis demonstrated a possible imbalance of amino acid metabolism in ASD children. Differential metabolites between the ASD and control groups involved glycine, serine and threonine metabolism, arginine and proline metabolism, taurine and hypotaurine metabolism, and glutathione metabolism pathways (Additional file 7: Fig. S6). In accordance with previous observations, amino acid metabolism disorder plays an important role in the pathogenesis mechanism of ASD [33–37]. Creatine and creatinine, which show significant metabolism alterations in ASD, play an essential role in maintaining a high level of energy supply for the brain [38]. Studies have indicated that creatine deficiency occurs in some ASD cases, and creatine may be a therapeutic target for ASD [39–42]. Creatine is biosynthesized from glycine and arginine with an intermediate metabolite of guanidinoacetic acid. Glycine acts as an excitatory neurotransmitter in the early developmental stage. As the nervous system matures, it transforms to the major inhibitory neurotransmitter. If the transformation does not occur, an abnormal level of glycine may result in neural disorders, including ASD [43–45]. The taurine and hypotaurine metabolism pathway also differed significantly between the two groups. Park et al. reported that taurine, as an antioxidant and regulator of inflammation, might be a valid biomarker for ASD [46]. Combined with vitamin D3, taurine showed benefits in the treatment of ASD [47–48]. All metabolic pathways interact with each other and constitute a complex network with related genes and diseases (Additional file 6: Fig. S5).

Many studies have reported that ASD is associated with abnormal gut microbial metabolism [49–51]. Gut microbiota metabolites, including phenylalanine, tyrosine, hippurate and tryptophan, have been reported to be factors in the development of ASD [52–55]. In our study, phenylacetylglycine, a gut microbial co-metabolite, had a slight variation between the two groups.

Phenylacetylglycine is the end product of the phenylalanine metabolism pathway [56]. However, there are no studies on the relationship between phenylacetylglycine and the pathogenesis of ASD.

Limitations

There are some limitations in our study. Due to the restriction of research centres, inconsistent urine collection times may have affected the quality of metabolic analyses. Our study lacks of healthy siblings as controls to remove the effects of confounding factors, such as heredity and environment. Though ^{1}H-NMR analysis is the regular NMR method, some studies have used 2D HSQC-NMR to improve urinary screening in ASD. Compared to ^{1}H-NMR analysis, ^{1}H-^{13}C HSQC-NMR analysis shows the advantage of improving the metabolite detection accuracy and the discrimination ability [18, 31]. Moreover, it is necessary to compare metabolite levels that vary with ASD severity to better clarify the pathogenesis of ASD.

Conclusions

In our study, ^{1}H-NMR metabolomic analysis was used to investigate urinary metabolism patterns in the ASD and control groups. We revealed that urinary amino acid metabolites significantly altered in children with ASD. A series of variations in amino acid metabolism pathways, including glycine, serine and threonine metabolism, arginine and proline metabolism, and taurine and hypotaurine metabolism might play important roles in the pathogenic mechanisms of ASD. Further studies of differential metabolites are needed to improve the understanding of ASD pathogenesis.

Abbreviations

ASD: Autism spectrum disorder

PCA: Principal component analysis

OPLS-DA: Orthogonal projection to latent structure discriminant analysis

NMR: Nuclear magnetic resonance spectroscopy

DSM-5: Statistical manual of mental disorders 5

ADOS: Autism diagnostic observation schedule

ADI-R: Autism diagnostic interview-revised

SW: Spectral width

RD: Recycle delay

FID: Free induction decay

CV: Cross-validation

CV-ANOA: Variance analysis of the cross-validated residuals

ROC: Receiver operating characteristics

AUC: Area under the curve

FNR: False-negative rate

FPR: False-positive rate

PPV: Positive predictive value

NPV: Negative predictive value

LR+: Likelihood ratio positive

LR−: Likelihood ratio negative

MESA: Metabolite sets enrichment analysis

Declarations

Acknowledgments

We sincerely thank all the patients, controls and parents who provided written consent to participate in this study, as well as the clinicians who contributed to the recruitment of the participants.

Funding

This project was supported by the National Health Commission of the People's Republic of China (201302002, Clinical Trials NCT02200679), Shanghai Hospital Development Centre (SHDC12015113) and Guangdong Key Project in "Development of new tools for diagnosis and treatment of Autism" (2018B030335001).

Ethics declarations

Consent for publication

Not applicable.

Ethics approval and consent to participate

The study was approved by the institutional ethics committee at the Children's Hospital of Fudan University. All eligible participants and their parents provided informed consent.

Competing interests

All authors declare that they have no competing interests.

Availability of data and materials

Not applicable

Author information

Yu Ma, Hao Zhou and Chunpei Li contributed equally to this work.

Affiliations

Department of Neurology, Children's Hospital of Fudan University, Shanghai, 201102, China.

Yu Ma, Hao Zhou, Chunpei Li, Lili Zhang, Tianqi Wang & Yi Wang

Department of Pediatrics, Guizhou provincial people's hospital, Guiyang, 550002, China.

Hao Zhou

Child Development Behaviour Centre, The Third Affiliated Hospital, Sun Yat-Sen University, Guangzhou, 510630, China.

Xiaobing Zou

Department of Psychiatry, The Second Xiangya Hospital of Central South University, Changsha, 410011, China.

Xuerong Luo

Department of Children and Adolescent Health, School of Public Health, Harbin Medical University, Harbin, 150081, China.

Lijie Wu

Department of Child Health Care, Children's Hospital of Chongqing Medical University, Chongqing, 400014, China.

Tingyu Li

The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, 325027, China.

Xiang Chen

Department of Child Health Care, Chengdu Women and Children's Hospital, Chengdu, 610091, China.

Meng Mao

Department of Psychiatry, West China Hospital of Sichuan University, Chengdu, 610041, China.

Yi Huang

Capital Institute of Paediatrics, Beijing, 100020, China.

Erzhen Li

State Key Laboratory of Genetic Engineering, Metabonomics and Systems Biology Laboratory, School of Life Sciences, Fudan University, Shanghai, 200438, China.

Yanpeng An

Department of Child Health Care, Children's Hospital of Fudan University, Shanghai, 201102, China.

Xiu Xu

Department of Clinical Epidemiology, Children's Hospital of Fudan University, Shanghai, 201102, China.

Weili Yan

Department of Genetics and Paediatrics, Yale School of Medicine, New Haven, CT 06520, USA.

Yonghui Jiang

State Key Laboratory of Medical Neurobiology, Fudan University, Shanghai, 201132, China.

Yi Wang

Authors' contributions

YW conceived of the study. YM and HZ contributed to the analysis, synthesis and interpretation of the results, and wrote the manuscript. CL, LZ and TW contributed to the sample collection. XZ, XL, LW, TL, XC, MM, YH and ELcontributed to the diagnosis of ASD and sample collection from each research centre. YA, XX, WY and YJ provided guidance for the study. All of the authors contributed to the preparation of the manuscript.

Corresponding author

Correspondence to Yi Wang

References

1. American Psychiatric Association. DSM-5. Autism Spectrum Disorder. Arlington: American Psychiatric Association; 2013.
2. Baron-Cohen S, Scott FJ, Allison C, et al. Prevalence of autism-spectrum conditions: UK school-based population study. Br J Psychiatry. 2009;194(6):500–9.
3. Baio JB, et al. Centre for Disease Control and Prevention. Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2018. MMWR Surveill Summ. 2018;67:1–23.
4. Zhou H, Xu X, Yan WL, et al. Prevalence of Autism Spectrum Disorder in China: A Nationwide Multi-center Population-based Study Among Children Aged 6 to 12 Years. Neurosci Bull. 2020 Jun 30.
5. Buescher AV, Cidav Z, Knapp M, Mandell DS. Costs of autism spectrum disorders in the United Kingdom and the United States. JAMA Pediatr. 2014;168:721–8.
6. Sanchack KE, Thomas CA. Autism spectrum disorder: primary care principles. Am Fam Physician. 2016;94(12):972–9.
7. Tachibana Y, Miyazaki C, Ota E, et al. A systematic review and meta-analysis of comprehensive interventions for pre-school children with autism spectrum disorder (ASD). PLoS ONE. 2017;12(12):e0186502.
8. Tsiplova K, Zur RM, Marshall CR, et al. A microcosting and cost-consequence analysis of clinical genomic testing strategies in autism spectrum disorder. Genet Med. 2017;19(11):1268–75.
9. Drmic IE, Walker SP, Jiang Y, et al. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. Am J Hum Genet. 2013;93(2):249–63.
10. Tammimies K, Marshall CR, Walker S, et al. Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. JAMA. 2015;314(9):895.
11. Lord C, Elsabbagh M, Baird G, et al. Autism spectrum disorder. Lancet. 2018;392(10146):508–20.
12. Khemakhem AM, Frye RE, El-Ansary A, et al. Novel biomarkers of metabolic dysfunction is autism spectrum disorder: potential for biological diagnostic markers. Metab Brain Dis. 2017;32:1983–97.
13. Afaf EA, Geir B, Khemakhem AM, et al. Metabolism-associated markers and childhood autism rating scales (CARS) as a Measure of Autism Severity. J Mol Neurosci. 2018;65(3):265–76.
14. Hollis F, Kanellopoulos AK, Bagni C. Mitochondrial dysfunction in Autism Spectrum Disorder: clinical features and perspectives. Curr Opin Neurobiol. 2017;45:178–87.
15. Ryan D, Robards K, Prenzler PD, et al. Recent and potential developments in the analysis of urine: A review. Anal Chim Acta. 2011;684(1–2):8–20.
16. Kim HK, Choi YH, Verpoorte R. NMR-based plant metabolomics: where do we stand, where do we go? Trends Biotechnol. 2011;29(6):267–75.
17. Yap IKS, Angley M, Veselkov KA, et al. Urinary metabolic phenotyping differentiates children with autism from their unaffected siblings and age-matched controls. J Proteome Res. 2010;9(6):2996–3004.
18. Mavel S, Nadal-Desbarats L, Blasco H, et al. ^{1}H-^{13}C NMR-based urine metabolic profiling in autism spectrum disorders. Talanta. 2013;114:95–102.
19. Lussu M, Noto A, Masili A, et al. The urinary ^{1}H-NMR metabolomics profile of an Italian autistic children population and their unaffected siblings. Autism Res. 2017;10(6):1058–66.
20. Craig A, Cloareo O, Holmes E, Nicholson JK, et al. Scaling and normalization effects in NMR spectroscopic metabonomic data sets. Anal Chem. 2006;78(7):2262–7.
21. Worley B, Powers R. PCA as a practical indicator of OPLS-DA model reliability. Curr Metabolomics. 2016;4(2):97–103.
22. Souihi N, Lindegren A, Eriksson L, Trygg J. OPLS in batch monitoring - Opens up new opportunities. Anal Chim Acta. 2015;857:28–38.

23. Pinto RC. Chemometrics Methods and Strategies in Metabolomics. Adv Exp Med Biol. 2017;965:163–90.
24. Worley B, Powers R. Multivariate Analysis in Metabolomics. Curr Metabolomics. 2013;1(1):92–107.
25. Triba MN, Le Moyec L, Amathieu R, et al. PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. Mol Biosyst. 2015;11(1):13–9.
26. Røraas T, Støve B, Petersen PH, Sandberg S. Biological Variation: The Effect of Different Distributions on Estimated Within-Person Variation and Reference Change Values. Clin Chem. 2016;62(5):725–36.
27. Chong J, Wishart DS, Xia J. Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. Curr Protoc Bioinformatics. 2019;68(1):e86.
28. Silva RA, Pereira TCS, Souza AR, et al. ^{1}H NMR-based metabolite profiling for biomarker identification. Clin Chim Acta. 2020;502:269–79.
29. Whiteley P, Waring R, Williams L, et al. Spot urinary creatinine excretion in pervasive developmental disorders. Pediatr Int. 2006 Jun;48(3):292–7.
30. Bandini LG, Curtin C, Phillips S, et al. Changes in Food Selectivity in Children with Autism Spectrum Disorder. J Autism Dev Disord. 2017 Feb;47(2):439–46.
31. Nadal-Desbarats L, Aïdoud N, Emond P, et al. Combined ^{1}H-NMR and ^{1}H-^{13}C HSQC-NMR to improve urinary screening in autism spectrum disorders. Analyst. 2014;139(13):3460–8.
32. Xiong X, Liu D, He W, et al. Identification of gender-related metabolic disturbances in autism spectrum disorders using urinary metabolomics. Int J Biochem Cell Biol. 2019;115:105594.
33. Smith AM, King JJ, West PR, et al. Amino acid dysregulation metabotypes: potential biomarkers for diagnosis and individualized treatment for subtypes of autism spectrum disorder. Biol Psychiatry. 2019;85(4):345–54.
34. Liu A, Zhou W, Qu L, et al. Altered urinary amino acids in children with autism spectrum disorders. Front Cell Neurosci. 2019;13:7.
35. Li C, Shen K, Chu L, et al. Decreased levels of urinary free amino acids in children with autism spectrum disorder. J Clin Neurosci. 2018;54:45–9.
36. Delaye JB, Patin F, Lagrue E, et al. Post hoc analysis of plasma amino acid profiles: towards a specific pattern in autism spectrum disorder and intellectual disability. Ann Clin Biochem. 2018;55(5):543–52.
37. Ming X, Stein TP, Barnes V, et al. Metabolic perturbance in autism spectrum disorders: a metabolomics study. J Proteome Res. 2012;11(12):5856–62.
38. Fons C, Campistol J. Creatine defects and central nervous system. Semin Pediatr Neurol. 2016;23(4):285–9.
39. Schulze A, Bauman M, Tsai AC, et al. Prevalence of creatine deficiency syndromes in children with nonsyndromic autism. Pediatrics. 2016;137(1):e20152672.
40. Schulze A. Creatine deficiency syndromes. Handb Clin Neurol. 2013;113:1837–43.
41. Fons C, Campistol J. Creatine Defects and Central Nervous System. Semin Pediatr Neurol. 2016;23(4):285–9.
42. Aydin HI. Creatine Transporter Deficiency in Two Brothers with Autism Spectrum Disorder. Indian Pediatr. 2018;55(1):67–8.
43. Kaila K, Price TJ, Payne JA, et al. Cation-chloride cotransporters in neuronal development, plasticity and disease. Nat Rev Neurosci. 2014;15(10):637–54.
44. Ito S. GABA and glycine in the developing brain. J Physiol Sci. 2016;66(5):375–9.
45. Tyzio R, Nardou R, Ferrari DC, et al. Oxytocin-mediated GABA inhibition during delivery attenuates autism pathogenesis in rodent offspring. Science. 2014;343(6171):675–9.
46. Park E, Cohen I, Gonzalez M, et al. Is Taurine a Biomarker in Autistic Spectrum Disorder? Adv Exp Med Biol. 2017;975:3–16.
47. Omura Y, Lu D, Jones MK, Nihrane A, et al. Early Detection of Autism (ASD) by a Non-invasive Quick Measurement of Markedly Reduced Acetylcholine & DHEA and Increased β-Amyloid (1–42), Asbestos (Chrysotile), Titanium Dioxide, Al, Hg & often Coexisting Virus Infections (CMV, HPV 16 and 18), Bacterial Infections etc. in the Brain and Corresponding Safe Individualized Effective Treatment. Acupunct Electrother Res. 2015;40(3):157–87.

48. Zheng HF, Wang WQ, Li XM, et al. Body fluid levels of neuroactive amino acids in autism spectrum disorders: a review of the literature. Amino Acids. 2017;49(1):57–65.
49. Fattorusso A, Di Genova L, Dell'Isola GB, et al. Autism Spectrum Disorders and the Gut Microbiota. Nutrients. 2019;11(3):521.
50. Ding HT, Taur Y, et al. Gut Microbiota and Autism: Key Concepts and Findings. J Autism Dev Disord. 2017;47(2):480–9.
51. Bezawada N, Phang TH, Hold GL, et al. Autism Spectrum Disorder and the Gut Microbiota in Children: A Systematic Review. Ann Nutr Metab. 2020:1–14.
52. Clayton TA. Metabolic differences underlying two distinct rat urinary phenotypes, a suggested role for gut microbial metabolism of phenylalanine and a possible connection to autism. FEBS Lett. 2012;586(7):956–61.
53. Srikantha P, Mohajeri MH. The possible role of the microbiota-gut-brain-axis in autism spectrum disorder. Int J Mol Sci. 2019;20(9):2115.
54. Sitkin SI, Tkachenko EI, et al METABOLIC DYSBIOSIS OF THE GUT MICROBIOTA AND ITS BIOMARKERS. Eksp Klin Gastroenterol. 2016;12(12):6–29.
55. Agus A, Planchais J, et al. Gut Microbiota Regulation of Tryptophan Metabolism in Health and Disease. Cell Host Microbe. 2018;23(6):716–24.
56. Delaney J, Neville WA, et al. Phenylacetylglycine, a putative biomarker of phospholipidosis: its origins and relevance to phospholipid accumulation using amiodarone treated rats as a model. Biomarkers. 2004;9(3):271–90.

Tables

Table 1
Pearson's correlation coefficient of discriminant metabolites

Metabolites	Pearson's correlation coefficient*
3-Aminoisobutanoic acid	0.208
Alanine	0.220
Taurine	0.242
Glycine	-0.237
Guanidinoacetic acid	-0.424
Creatine	-0.278
Creatinine	0.268
Hydroxyphenylacetylglycine	-0.190
Phenylacetylglycine	-0.233
Hypoxanthine	0.370
Formate	-0.248
N-methylnicotinamide	0.348
Unknown	-0.296

* The negative value indicates that the metabolite content in the urine samples of the control group is significantly lower than that of the ASD group. Conversely, the positive value indicates that the metabolite of the control group is significantly higher than that of the ASD group.

Table 2
Diagnostic accuracies of differential metabolites between the ASD and control groups

Metabolites	AUC	Cut-off	Sensitivity (%)	Specificity (%)	FNR (%)	FPR (%)	PPV (%)	NPV (%)	LR+	LR-
3-Aminoisobutanoic acid	0.568	0.5890	67.52	50.42	32.48	49.58	57.2	61.2	1.36	0.64
Alanine	0.508	0.4190	73.50	31.93	26.50	68.07	51.5	55.1	1.08	0.83
Taurine	0.502	0.5683	44.44	61.34	55.56	38.66	53.1	52.9	1.15	0.91
Glycine	0.560	1.2698	41.88	70.59	58.12	29.41	58.3	55.3	1.42	0.82
Guanidinoacetic acid	0.574	1.2573	76.07	40.34	23.93	59.66	55.6	63.2	1.27	0.59
Creatine	0.623	1.8805	44.44	77.31	55.56	22.69	65.8	58.6	1.96	0.72
Creatinine	0.639	4.6505	55.56	69.75	44.44	30.25	64.4	61.5	1.84	0.64
Hydroxyphenylacetylglycine	0.504	0.1710	26.50	78.99	73.50	21.01	55.4	52.2	1.26	0.93
Phenylacetylglycine	0.571	0.2742	87.18	26.89	12.82	73.11	54.0	68.1	1.19	0.48
Hypoxanthine	0.657	0.0359	88.03	36.13	11.97	63.87	57.5	75.4	1.38	0.33
Formate	0.529	0.0768	19.66	91.60	80.34	8.40	69.7	53.7	2.34	0.88
N-methylnicotinamide	0.595	0.0104	74.36	49.58	25.64	50.42	59.2	66.3	1.47	0.52

AUC, area under the curve; FNR, false-negative rate; FPR, false-positive rate; PPV, positive predictive value; NPV, negative predictive value; LR+, likelihood ratio positive; and LR−, likelihood ratio negative.

Figures

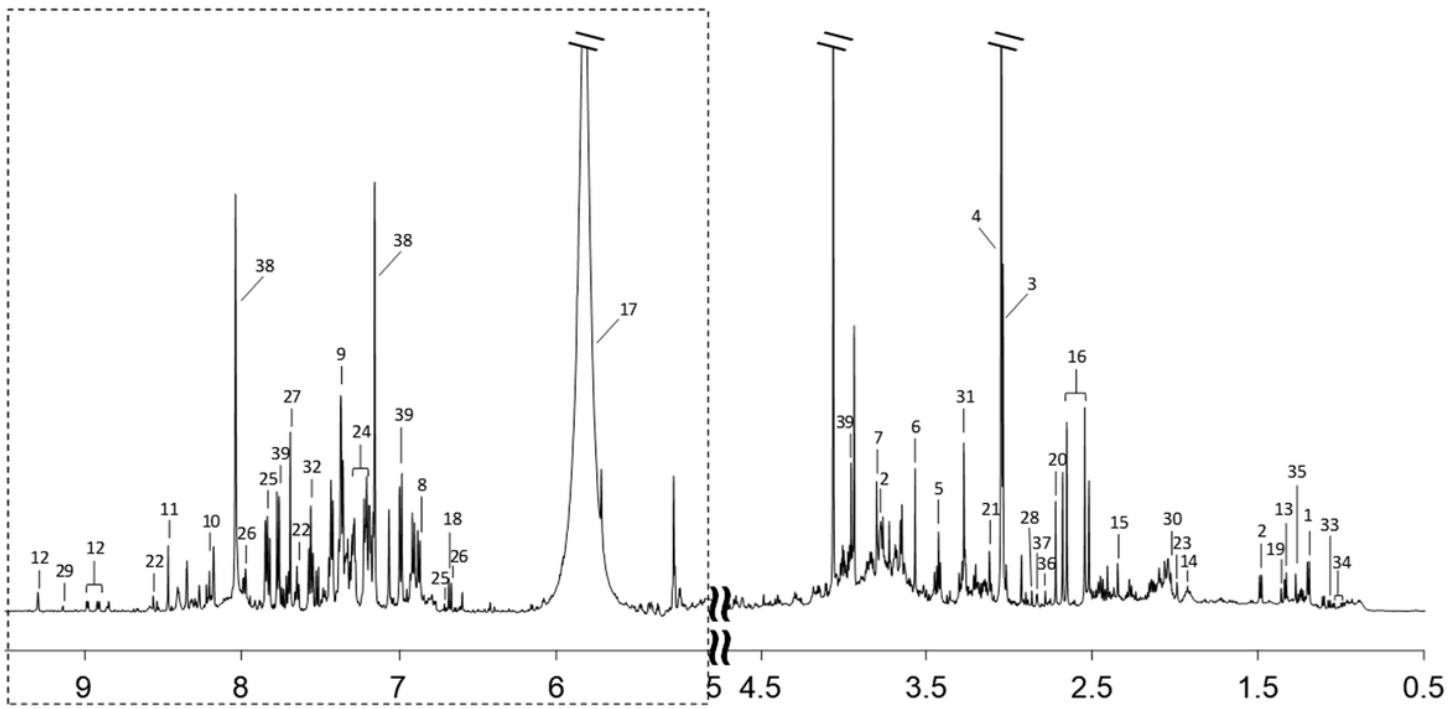

Figure 1

A typical 600 MHz 1H-NMR spectrum of a urine sample. Region of dashed box is vertically magnified 8 times. Metabolites corresponding to each number are listed in Additional file 1: Table S1.

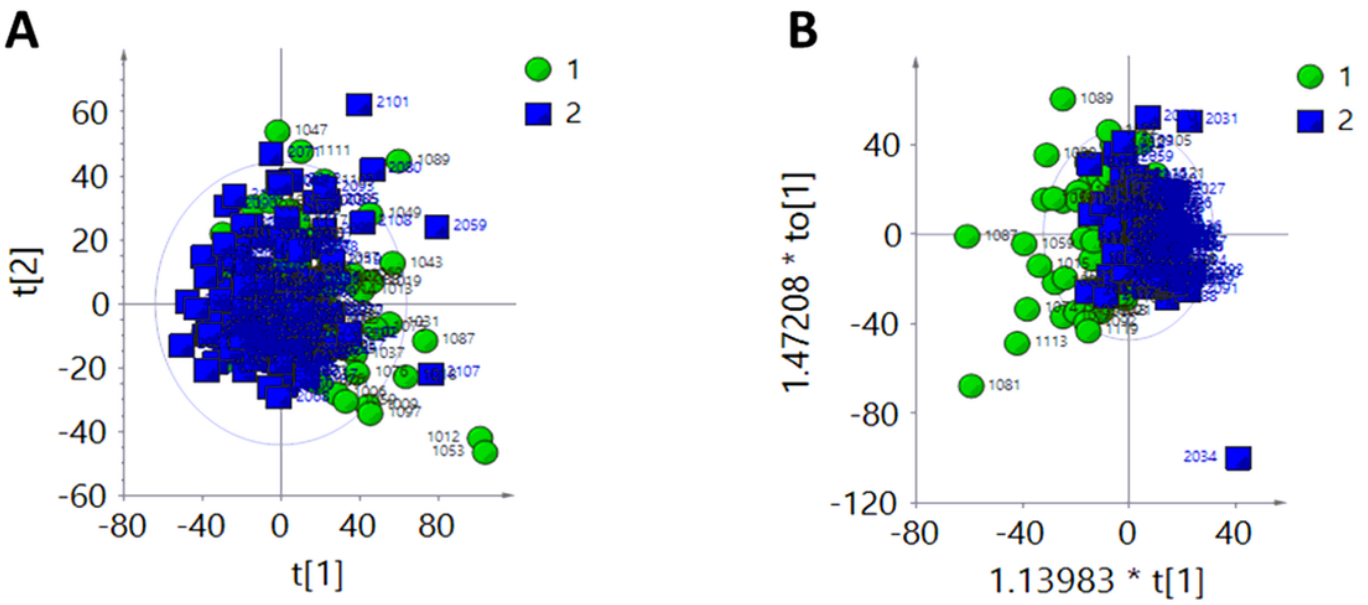

Figure 2

After the creatinine normalization, multivariate analysis results of 1H-NMR spectra of urine samples from the ASD group (green dots, 1) and control group (blue squares, 2). (A) PCA score plot: R2X = 0. 403, Q2 = 0.135. (B) OPLS-DA score plot: R2X = 0.257, Q2 = 0.0138, CV-ANOVA P = 0.523.

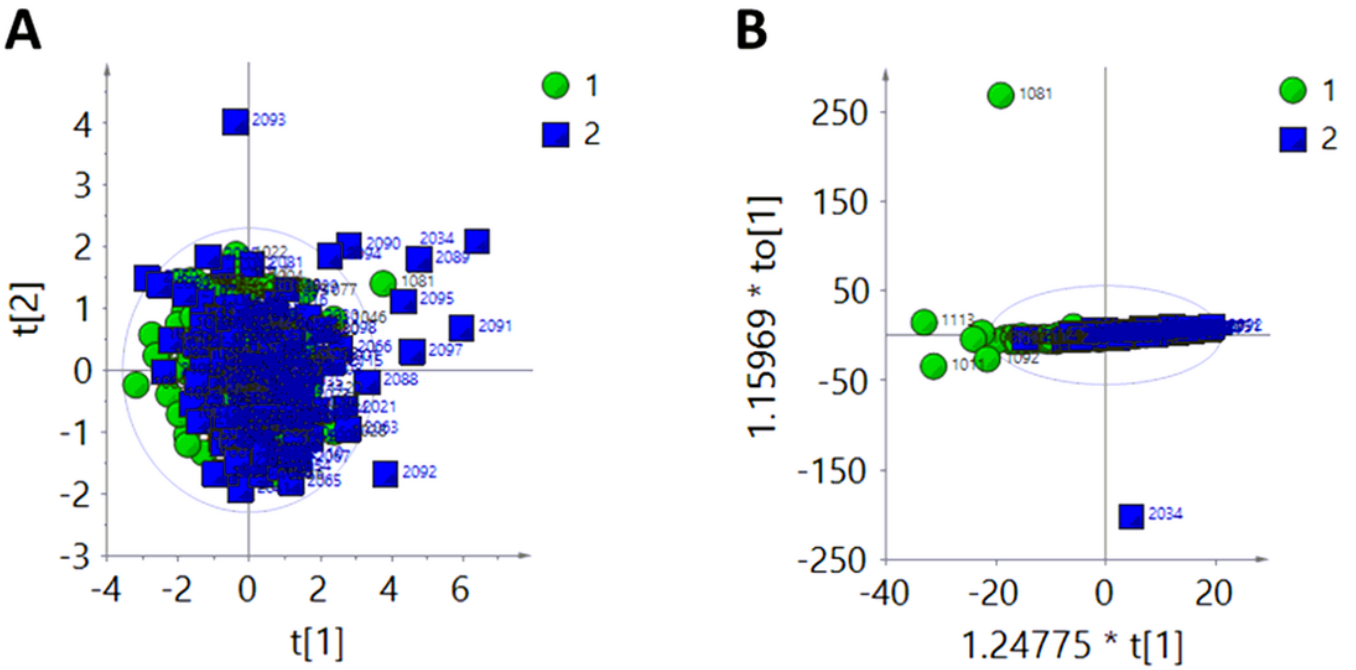

Figure 3

After the total area normalization, multivariate analysis results of 1H-NMR spectra of urine samples from the ASD group (green dots, 1) and control group (blue squares, 2). (A) PCA score plot: R2X = 0.366, Q2 = 0.0656. (B) OPLS-DA score plot: R2X = 0.274,

Q2 = 0.0565, CV-ANOVA P = 0.009.

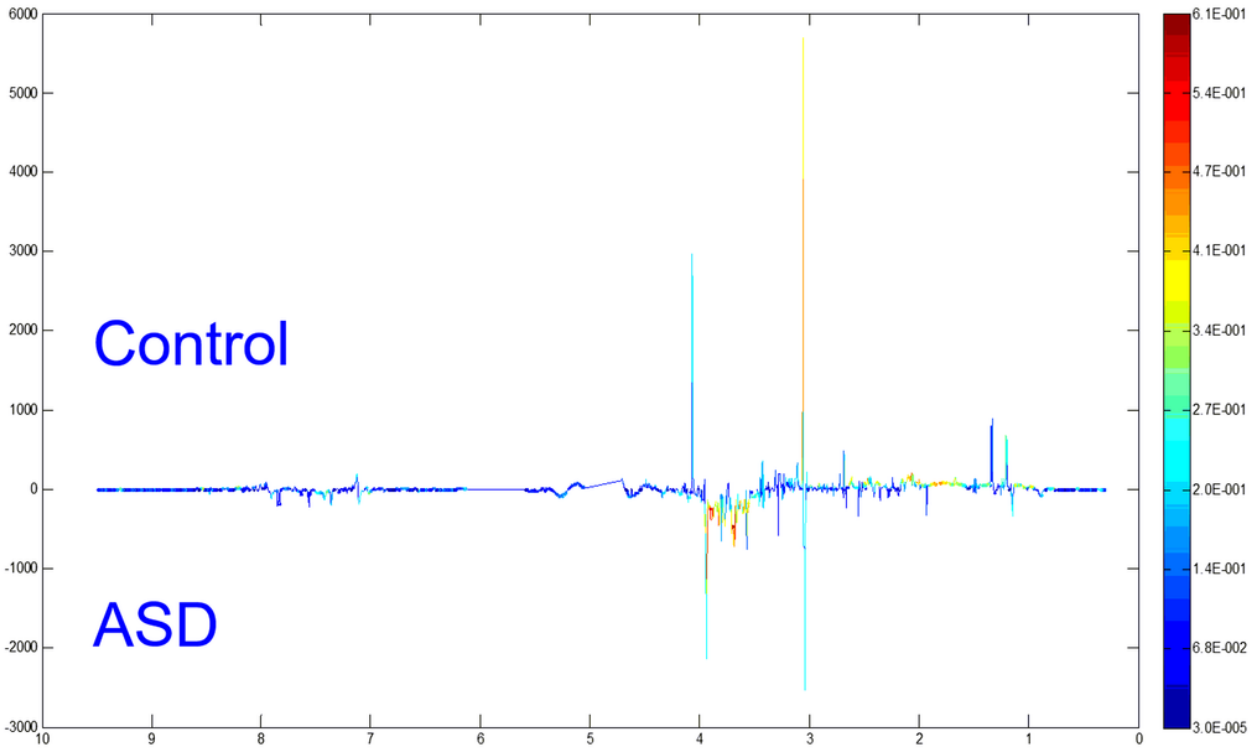

Figure 4

OPLS-DA model: a polychromatic correlation coefficient loading plot.

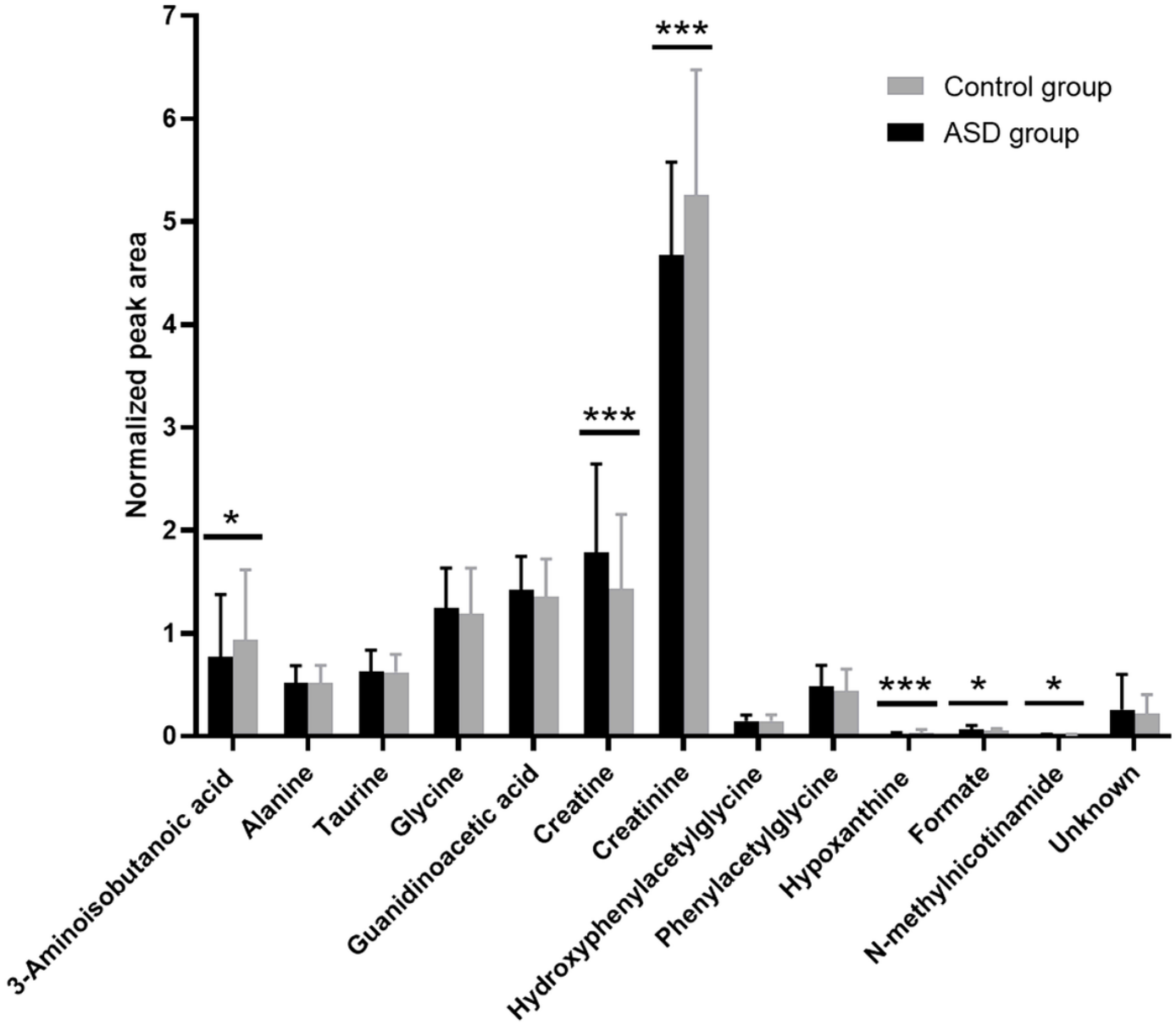

Figure 5

Comparation of the normalized peak area of the differential metabolites between the ASD and control groups. The vertical axis represents the average peak area after the normalization of the corresponding spectral peaks for the metabolites. * $P < 0.05$, *** $P < 0.001$.

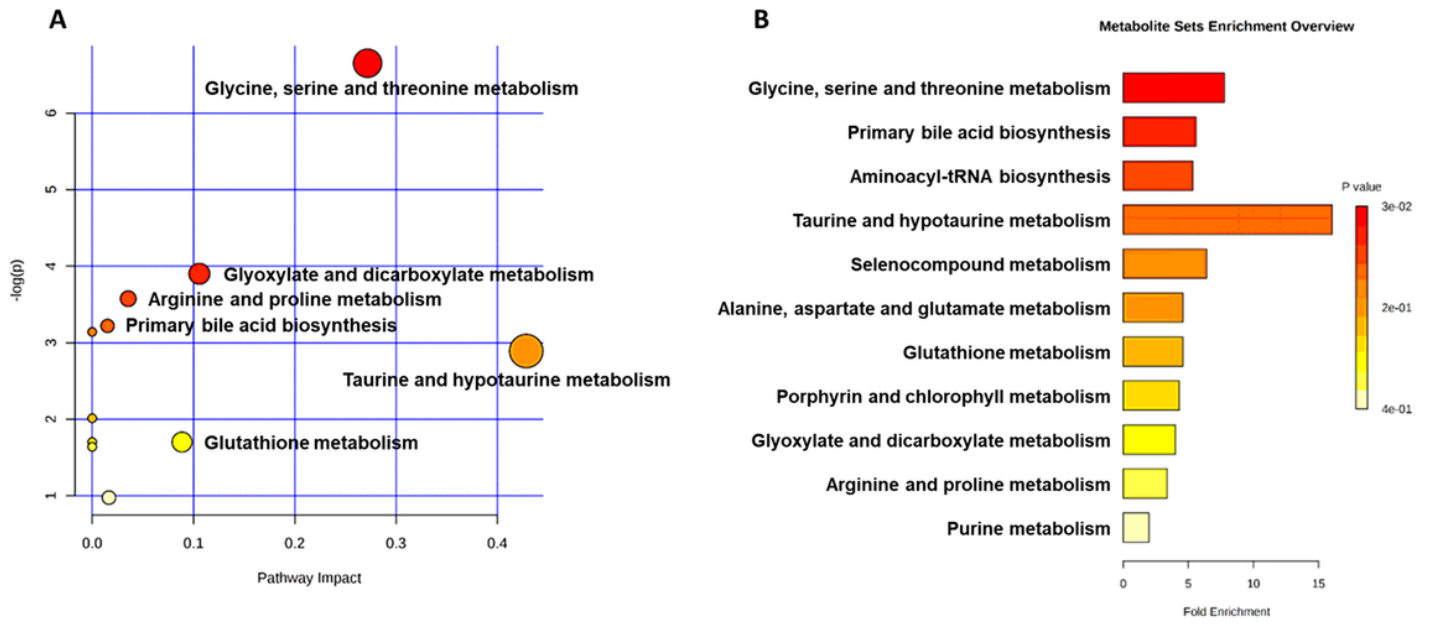

Figure 6

Metabolic pathway analyses utilizing the MetaboAnalyst functional interpretation tools. (A) Bubble plot of metabolic pathway impact. The metabolic pathways are shown as bubbles. The X coordinate and size of the bubble represent the value of pathway impact in the topology analysis. The Y coordinate and colour of the bubble represent the P value of the enrichment analysis. The darker red colour and larger size indicate a more significant metabolite change in the corresponding pathway. (B) Metabolite set enrichment analysis (MESA) plot. Significantly enriched pathways are represented by bars. The colour and length of the bar are based on the P value and fold enrichment, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile7FigS6.tif
- Additionalfile6FigS5.tif
- Additionalfile5FigS4.tif
- Additionalfile4FigS3.tif
- Additionalfile3FigS2.tif
- Additionalfile2FigS1.tif
- Additionalfile1Table.docx