Urine metabolomic profiles of autism and autistic traits – a twin study

Abishek Arora[1,2], Francesca Mastropasqua[1,2], Sven Bölte[1,3,4] and Kristiina Tammimies[1,2,*]

[1]Center of Neurodevelopmental Disorders (KIND), Centre for Psychiatry Research, Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden

[2]Astrid Lindgren Children's Hospital, Karolinska University Hospital, Region Stockholm, Stockholm, Sweden

[3]Child and Adolescent Psychiatry, Stockholm Health Care Services, Region Stockholm, Stockholm, Sweden

[4]Curtin Autism Research Group, Curtin School of Allied Health, Curtin University, Perth, Western Australia

[*]email: kristiina.tammimies@ki.se

Corresponding author

Kristiina Tammimies, PhD

Postal address: BioClinicum J9:30, Visionsgatan 4, 171 56 Solna, Sweden

E-mail: kristiina.tammimies@ki.se

Abstract

Currently, there are no reliable biomarkers for autism diagnosis. The heterogeneity of autism and several co-occurring conditions are key challenges to establishing these. Here, we used untargeted mass spectrometry-based urine metabolomics to investigate metabolic differences for autism diagnosis and autistic traits in a well-characterized twin cohort (N=105). We identified 208 metabolites in the urine samples of the twins. No clear, significant metabolic drivers for autism diagnosis were detected when controlling for other neurodevelopmental conditions. However, we identified nominally significant changes for several metabolites. For instance, phenylpyruvate (p=0.019) and taurine (p=0.032) were elevated in the autism group, while carnitine (p=0.047) was reduced. We furthermore accounted for the shared factors, such as genetics within the twin pairs, and report additional metabolite differences. Based on the nominally significant metabolites for autism diagnosis, the arginine and proline metabolism pathway (p=0.024) was enriched. We also investigated the association between quantitative autistic traits, as measured by the Social Responsiveness Scale 2nd Edition, and metabolite differences, identifying a greater number of nominally significant metabolites and pathways. A significant positive association between indole-3-acetate and autistic traits was observed within the twin pairs (adjusted p=0.031). The utility of urine biomarkers in autism, therefore, remains unclear, with mixed findings from different study populations.

Lay summary

Earlier literature has suggested that there are chemical molecules (metabolites) in the urine samples of autistic individuals that could be linked with their diagnosis of autism. However, there are still mixed results and uncertainty if any of these metabolites could be used as biomarkers (chemical molecules that indicate a biological condition). In this study, we analysed urine samples of 105 twins from Sweden, of which 48 were identical or fraternal twins, and the analysed metabolite levels were compared between those twins diagnosed with autism and those without a diagnosis. Additionally, we tested if the metabolite levels were different in relation to the level of autistic behaviour. The metabolite levels were measured using mass spectrometry, allowing us to identify hundreds of chemical molecules present in a urine sample followed with testing their significance for autism diagnosis or autistic traits. No single metabolite was found to be highly associated with a diagnosis of autism. However, indole-3-acetate, a metabolite produced by the breakdown of the amino acid tryptophan, was associated with the level of autistic behaviour within the twins. Nevertheless, several metabolites showed

some association with autism or the level of autistic behaviour, while revealing interesting pathways for studies in the future.

Keywords

Autism, Twin cohort, Urine metabolomics, Mass spectrometry, Metabolic pathways.

Introduction

Autism is a neurodevelopmental condition with a heterogeneous presentation that is diagnosed in 1 to 2% of the world population (Lord et al., 2020). Although several genes and genetic variants have been associated with autism phenotypes (Lord et al., 2020; Satterstrom et al., 2020; Zhou et al., 2022), there is no evidence of a single biomarker that can be of benefit to the diagnostic procedure of the condition (Cortese et al., 2023). Finding biomarkers that could aid either in the diagnosis or specifying a subgroup of individuals could help, for instance, in the early diagnosis of autism (Amaral et al., 2019).

Prior research has suggested an association of metabolic anomalies with autism that could potentially aid in the search for biomarkers (Jensen et al., 2022). Metabolomic analyses using urine samples provide a non-invasive method with greater ease of sample collection and handling in comparison to other biospecimens. This is particularly advantageous in clinical settings where the process of diagnosing autism can begin at a younger age (Lord, 2019). Moreso, the ability to detect metabolites in urine samples is well-established using mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR), and the typical human urine metabolome has been mapped (Bouatra et al., 2013). Early efforts to understand the urine metabolome in autism began several decades ago, with initial biochemical studies searching for associations (Himwich et al., 1972; Young et al., 1978). The investigations have shifted to more untargeted and high-precision analytical approaches using NMR (Ma et al., 2021; Mavel et al., 2013) and MS (Mussap et al., 2020; Noto et al., 2014; Timperio et al., 2022).

The interest in autism metabolomics has continued to expand; however, more accessible data are required across different populations to identify robust associations. Such an approach will also make it possible to characterise how both, situational and long-term environmental factors interplay with the genetic backgrounds of autism (Bai et al., 2019). Twin studies are an appropriate tool to investigate the contribution of environmental and genetic factors (Colvert et al., 2015; Frazier et al., 2014) and can be extended to metabolic profiles. It

is especially valuable to analyse differences between monozygotic (MZ) twins who are discordant for autism, to tease out genetic contributions.

In this study, we utilised a rare twin sample, the Roots of Autism and ADHD Twin Study in Sweden (RATSS) (Bölte et al., 2014; Myers et al., 2021), enriched for MZ twins discordant for autism, and their urine sample-based metabolomics to search and validate potential metabolite biomarkers. Furthermore, on the basis of these findings we have mapped the affected metabolic pathways. Our study adds on to the literature exploring the metabolites as possible biomarkers for autism and expands the search for autistic traits.

Methods

Study participants

In this study, individuals (N=105) and 48 complete twin pairs were selected from the RATSS cohort (Bölte et al., 2014; Myers et al., 2021), which is a neurodevelopmental condition (NDC) enriched twin sample recruited from the general Swedish population, for untargeted mass spectrometry-based urine metabolomics. Detailed inclusion and exclusion criteria for RATSS have been previously published (Bölte et al., 2014). Among the recruited twins, we selected participants for this study based on the autism diagnosis status, whether the twin pair was concordant (both with autism diagnosis) or discordant (only one with autism diagnosis) and if they had available urine samples. Furthermore, we age- and sex-matched the non-autistic twin pairs. The study was approved by the Swedish Ethical Review Authority (2016/1452-31). Informed consent was obtained from all participants or their caregivers, based on their age.

During a 2.5-day study visit, a team of clinical professionals conducted a diagnostic evaluation of the participants in line with the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) guidelines (American Psychiatric Association, 2013). The evaluation utilised a combination of diagnostic interviews, a review of medical history documents, and the use of established diagnostic measures (Bölte et al., 2014). This included behavioural assessment tools such as the Autism Diagnostic Interview-Revised (ADI-R) (Rutter et al., 2003), the Autism Diagnostic Observation Schedule 2nd Edition (ADOS-2) (Lord et al., 2012). Furthermore, additional tools were used to establish the diagnosis of other NDCs, if any, such the Diagnostic Interview for ADHD in Adults (DIVA-2) (Kooij & Francken, 2010), the Structured Clinical Interview for DSM-IV (SCID-IV) (First & Gibbon, 2004), and the Adaptive Behavior Assessment System (ABAS) (Harrison & Oakland, 2000) – more detailed information about the same is available in the publication by Bölte et al (2014). Autistic traits were evaluated with the parent-report version of the Social Responsiveness Scale 2nd Edition

(SRS-2) (Constantino, 2005, 2011; Constantino & Gruber, 2012), consisting of 65 items. Intelligence quotient (IQ) was measured by using the Wechsler Intelligence Scale for Children or Adults - IV General Ability Index (GAI) (Tulsky et al., 2001).

Additionally, the participants were asked for a list of their current, regularly used medications during the study visit. As there was a collection of different medications including antidepressants and ADHD medication, all were grouped together and adjusted for in our analyses. No subgrouping was possible for the drugs due to lack of power to detect metabolomic effects of specific drugs.

Urine collection and metabolite extraction

The urine samples were collected at the visit from the participants, aliquoted and stored in the Karolinska Institutet Biobank at -80 °C, until these were suitably transported for analysis to the Proteomics and Metabolomics Facility, University of Tuscia. All samples were handled as per the same stated protocol. Before the metabolomic analysis, the urinary specific gravity was measured following centrifugation at 13,000 g for 10 minutes. Urine aliquots (200 µl) were mixed with 200 µl of methanol:acetonitrile:water (50:30:20), vortexed for 30 minutes, maximum speed at 4 °C and then centrifuged at 16,000 g for 15 minutes at 4 °C. Supernatants were collected for metabolomic analysis.

Ultra-High Performance Liquid Chromatography (UHPLC)

For the experiments, 20 µL of samples were injected into a UPLC system (Ultimate 3000, Thermo Scientific) and were analysed on positive mode: samples were loaded onto a Reprosil C18 column (2.0 mm × 150 mm, 2.5 µm - Dr Maisch, Germany) for metabolite separation. Chromatographic separations were achieved at a column temperature of 30 °C and flow rate of 0.2 mL/min. A linear gradient (0–100%) of solvent A (ddH_2O, 0.1% formic acid) to B (acetonitrile, 0.1% formic acid) was employed over 20 minutes, returning to 100% solvent A in 2 minutes and a 6-minute post-time solvent A hold. Acetonitrile, formic acid, and HPLC-grade water were purchased from Sigma Aldrich.

High Resolution Mass Spectrometry (HRMS)

The UPLC system was coupled online with a mass spectrometer, Q Exactive (Thermo Scientific), scanning in full MS mode (2 µ scans) at a resolution of 70,000 in the 67 to 1000 *m/z* range, target of 1 × 106 ions and a maximum ion injection time (IT) of 35 ms, 3.8 kV spray voltage, 40 sheath gas, and 25 auxiliary gas, operated in negative and then positive ion mode.

Source ionization parameters were: spray voltage, 3.8 kV; capillary temperature, 300 °C; and S-Lens level, 45. Calibration was performed before each analysis against positive or negative ion mode calibration mixes (Piercenet, Thermo Fisher, Rockford, IL) to ensure sub-ppm error of the intact mass.

Metabolite quantification

Data were normalized by urinary specific gravity because creatinine excretion may be abnormally reduced in autistic children (Whiteley et al., 2006). Replicates were exported as mzXML files and processed through MAVEN (Melamud et al., 2010). Mass spectrometry chromatograms were elaborated for peak alignment, matching and comparison of parent and fragment ions, and tentative metabolite identification (within a 10-ppm mass deviation range between observed and expected results against the imported Kyoto Encyclopaedia of Genes and Genomes (KEGG) database (Kanehisa et al., 2023).

Differential metabolomics

Principal component analysis (PCA) was performed in R (R Core Team, 2020) followed by visualisation with the *factoextra* package (https://cran.r-project.org/package=factoextra) to identify any outliers in the dataset. Prior to the PCA, missing values (0.018%) were imputed using the *missMDA* package (Josse & Husson, 2016) in R (R Core Team, 2020). Here, the number of dimensions of the PCA were estimated by cross-validation using the k-fold method followed by imputation of missing values using the PCA model.

Differential metabolomic analysis for 208 metabolites was performed using the *drgee* package (Zetterqvist & Sjölander, 2015) in R (R Core Team, 2020), for overall effects and differences between twin pairs using a generalised estimating equations (GEE) model, with suitable covariates. For this analysis, missing values (0.018%) were not removed from the dataset and were treated as *NA* values. In the GEE model, the metabolites were individually considered as the response variable, while the autism diagnosis status was the predictor variable. Age, sex, body mass index (BMI), medication status and diagnosis status of other NDCs were added as covariates to the model: *Metabolite ~ Autism Diagnosis + Other NDC Diagnosis + age + sex + BMI + medication status*. The model was applied for the whole cohort (Model A) and for differences within twin pairs (Model B) through the *drgee* package (Zetterqvist & Sjölander, 2015) in R (R Core Team, 2020). The same model was then applied to a subset of the cohort based on zygosity. Here, age and sex were removed as covariates when testing for differences between twin pairs as these were intrinsically controlled.

In addition to testing autism diagnosis status as a predictor variable, we used autistic traits measured by the SRS-2 (Constantino, 2005, 2011; Constantino & Gruber, 2012) total raw scores. As IQ levels are known to have an impact on the SRS-2 scores (Hirosawa et al., 2020), along with the previously stated covariates (other than diagnostic status of other NDCs), IQ scores from GAI were added to the models: *Metabolite ~ SRS total raw score + age + sex + BMI + medication status + IQ GAI*. Also, for SRS-2 total raw scores, the models were applied to subsets of the cohort based on zygosity. As previously stated, age and sex were removed as covariates when testing for differences between twin pairs as these were intrinsically controlled. For all instances, p values were extracted from the results of each model and adjusted for multiple comparisons using the false discovery rate (FDR) method (Benjamini & Hochberg, 1995) in R (R Core Team, 2020).

Pathway enrichment analysis

Significant metabolites from the GEE models were tested for enrichment in the Small Molecule Pathway Database (SMPDB) (Jewison et al., 2014) using over-representation analysis (ORA) from the enrichment analysis module of MetaboAnalystR (Pang et al., 2022). The cut-off threshold for entries from metabolite sets to be included was set to 2. ORA was implemented using the hypergeometric test to evaluate whether a particular metabolite set is represented more than expected by chance within the given list of metabolites. One-tailed p values were provided after adjusting for multiple testing. Based on the statistical testing, enrichment dot-plots and pathway networks were generated. In the pathway networks, 2 or more nodes were only connected by an edge when the number of shared metabolites was >25% of the combined metabolites contributing to each node.

Statistical analysis and plots

All statistical analyses were performed in R (v4.1.2) (R Core Team, 2020). The statistical models and tests used for the analyses are described in the methodology relevant to the experimental technique, in the sections above. All plots, unless otherwise stated, were created using the *ggplot2* package (v3.3.5) (Wickham, 2009) or the *pheatmap* package (v1.0.12, https://CRAN.R-project.org/package=pheatmap) in R (v4.1.2) (R Core Team, 2020).

Results

Metabolites do not clearly explain sample differences

UHPLC was coupled with HRMS for the untargeted analysis of metabolites. In the study cohort (N=105 twins), of which 48 form complete twin pairs, we identified 208 metabolites with high reliability following mass spectrometry of urine samples. Principal component analysis (PCA) of the metabolomics data demonstrated no clear drivers of effects across PC1 (23.2%) and PC2 (10.8%) (Figure 1A). The first 10 principal components were able to explain 68.2% of the variation (Figure 1B). The loadings for PC1 (Figure 1C) and PC2 (Figure 1D) did not identify a metabolite that was a major driver of the observed effects, with the top 10 loadings having a contribution of ~1.25% to ~1.5 % each in PC1 and ~1.75 to ~2.5% each in PC2. There were 9 individuals outside the 95% CI ellipses (Figure 1A), who were designated as outliers and excluded from further analysis, leaving a study cohort with 96 individuals, of which 42 form complete twin pairs. PCA was repeated on this subset of 96 samples to check if the prior removal of outliers better explained the variance of the dataset, however, this was to no avail. More outliers were therefore not removed based on further iterations of the PCA. Moreover, a lack of clear clusters following the PCA (Figure 1A) also ruled out the possibility of using the contemporarily favoured orthogonal partial least squares discriminant analysis (OPLS-DA) model for multivariate analysis in the full cohort (Worley & Powers, 2016).

Of these 96 individuals selected for further analysis, the autistic participants had a mean age of 14 years, 11 being female and 24 having a clinical diagnosis of an NDC other than autism. The demographic and clinical features of the study cohort used for further analyses are summarised in Table 1. Specific information on race/ethnicity, socioeconomic status and educational attainment were not recorded.

Table 1. Demographic characteristics of study cohort.

Characteristics	Autism (n=38)	Other (n=58)
Age, mean years (SD) [range]	14.21 (3.26) [8-21]	15.76 (3.14) [9-23]
Sex, n females (%)	11 (29.95%)	22 (37.93%)
MZ:DZ concordant pairs*	2:1	11:3
Autism discordant pairs (n)*	14	
MZ	5	
DZ	9	
BMI, mean (SD) [range]	21 (5.41) [13-38]	21 (3.05) [15-30]
Other NDC diagnosis n (%)	24 (63.16%)	8 (13.79%)
Medication n (%)	24 (63.16%)	15 (23.86%)
SRS-2, mean (SD) [range]	74 (18.66) [0-105]	48 (9.60) [35-78]
IQ, mean (SD) [range]	86 (25.13) [0-138]	102 (13.81) [65-130]

SD: Standard Deviation, MZ: Monozygotic, DZ: Dizygotic, BMI: Body Mass Index, NDC: Neurodevelopmental Condition, SRS-2: Social Responsiveness Scale 2nd Edition, IQ: Intelligence Quotient, * Complete twin pairs.

Subtle changes in metabolite status across study cohort

The peak areas of each identified metabolite from participants of the study cohort (N=96) were analysed for association with autism diagnosis using a GEE model, first for the whole cohort (Model A) and then within twin pairs (Model B). We also adjusted the models for the diagnosis status of other NDCs. As expected, from the prior PCA, no metabolite had a significant association with autism after FDR correction for 208 tests. Based on nominal significance, a handful of metabolites were changed (Table S1A). For instance, deoxycholic acid (p=0.048), orotate (p=0.042), phenylpyruvate (p=0.019) and taurine (p=0.032) were elevated, while carnitine (p= 0.047) was reduced in autism (Figure 2A and Table S1A).

When analysing for differences within twin pairs using the GEE model (Model B), i.e. which corrects for all the shared variation between the twins such as genetics and shared environment, eight metabolites were significantly modulated in the autism group based on the nominal p value (Figure 2A and Table S1A). These were increased hypoxanthine (p=0.019), diiodothyronine (p=0.049) and kynurenic acid (p=0.003), and decreased guanidoacetic acid (p=0.001), indole (p=0.001) and L-arginosuccinate (p=0.031).

The study cohort was then subsetted as per zygosity status and the same GEE models were applied for analysis. In the subset of MZ twins (n=52), 17 metabolites were identified to be significantly different between autistic and other twins (Model A, Figure 2A and Table S1B), including an increase in unique metabolites such as alanine (p=0.029), asparagine (p=0.025) and 2-oxo-4-methylthiobutanoate (p=0.007), and a decrease in creatine (p=0.001) and deoxyinosine (p=0.020). When testing for differences within the MZ twin pairs (Model B), 17 metabolites were differentially modulated (Figure 2A and Table S1B), including additional metabolites that were elevated like glutamine (p=0.008), proline (p=0.022) and homocysteine (p=0.037). A decrease was detected in several metabolites, such as glucose-6-phosphate (p=0.032), L-kynurenine (p=0.043) and D-histidine (p=0.014), among others.

In the subset for dizygotic (DZ) twins (n=44), peak areas of only diidothyronine (p=0.039) were elevated (Model A, Figure 2A and Table S1C). After accounting for differences within twin pairs in the GEE model (Model B), ten metabolites were significantly modulated (Figure 2A and Table S1C), such as an increase in valine (p=0.018), quinoline (p=0.034) and N-acetyl-glutamate (p=0.046). Reductions were detected in four metabolites including isocitrate (p=0.018) and indole (p=0.009).

Association between autistic traits and metabolite differences

We also investigated associations between the metabolites and quantitative autistic traits measured by the SRS-2 total raw scores (Constantino, 2005, 2011; Constantino & Gruber, 2012). We analysed the association between the peak areas of each metabolite identified in the study cohort (N=96) with the SRS-2 using the GEE models and adjusted for IQ (Wechsler GAI) (Tulsky et al., 2001). Again, no metabolite reached statistical significance after FDR multiple test correction. From Model A, 30 nominally significant metabolites were detected (Figure 2B and Table S2A). Amongst those elevated were acetylcarnitine (p=0.023), aspartate (p=0.002) and hydroxyphenylpyruvate (p=0.016); and the only metabolite to be reduced was creatine (p=0.027). When testing for differences within the twin pairs (Model B), five metabolites crossed the nominal significance threshold (Figure 2B and Table S2A), including a decrease in adenosine (p=0.025) and 6-phospho-d-gluconate (p=0.033), and an increase in indole-3-acetate (p=0.0001, adjusted p=0.031).

We further explored the association in the zygosity groups for the autistic traits and the metabolites. In the MZ subset (n=52), 12 metabolites demonstrated nominally significant changes when applying the GEE model (Model A, Figure 2B and Table S2B). An elevation was detected in coenzyme A (p=0.044), AMP (p=0.041) and a decline in levels of DL-pipecolic acid (p=0.024). Following testing for differences within the twin pairs (Model B), six metabolites were identified (Figure 2B and Table S2B), which included an increase in prephenate (p=0.027) and D-glycerylaldehyde 3-phosphate (p=0.003) and a decrease in 5'-S-Methyl-5'-thioadenosine (p=0.001) and thiamine (p=0.032).

In the DZ subset (n=44), 23 metabolites were significantly elevated ($p<0.05$) in association with the autistic traits (Model A, Figure 2B and Table S2C). These were, to name a few, D-methionine (p=0.047), D-gluconate (p=0.038), myo-inositol (p=0.047) and lysine (p=0.043). When accounting for differences between the twin pairs (Model B), 16 metabolites demonstrated significant changes (Figure 2B and Table S2C) where none were reduced, including glycolate (p=0.009), 4-aminobutyrate (p=0.035) and IMP (p=0.012).

Enrichment of identified metabolites in biochemical pathways

The nominally significant metabolites identified from the GEE models were used for ORA as described in the methods. Ten biochemical pathways were identified to be enriched in the autistic twins based on the full cohort analysis (N=96, Model A) (Figure 3A and 3B), out of which there was significant enrichment only in the bile acid biosynthesis pathway (p=0.035,

Table S3A). When ORA analysis was done for the metabolites found to be nominally significant for autism diagnosis between the twin pairs (Model B), ten pathways were identified (Figure 3C and 3D), where only the arginine and proline metabolism was significant (p=0.024, Table S3B).

For the models based on autistic traits in the full cohort (N=96, Model A), 43 biochemical pathways were enriched (Figure 4B), out of which two crossed the significance threshold (Table S3C): glutathione metabolism (p=0.0503) and glutamate metabolism (p=0.0506). The enrichment dot-plot of the top ten pathways is depicted in Figure 4A. On the other hand, eight biochemical pathways were enriched from significant metabolites identified after testing for autistic traits between the twin pairs (Model B, Table S3D), of which methionine metabolism (p=0.001) was significantly detected (Figure 4C and 4D).

Discussion

In this study, we explored the urine metabolome of a subset of individuals from the RATSS cohort (Bölte et al., 2014; Myers et al., 2021) using untargeted mass-spectrometry. While only a single metabolite reached significance after FDR correction, we identified several metabolites as nominally significant for autism diagnosis and autistic traits. Furthermore, the identified metabolites were enriched in a few relevant biochemical pathways. Increased detection of significant metabolites were found for the autistic traits, rather than for the autism diagnoses.

Our study is the first to evaluate differences in the urine metabolome of an autism twin cohort and utilise information on co-occurring NDCs in our analyses. This is an important aspect to account for, when investigating suitable biomarkers in lieu of the high prevalence of co-occurring NDCs in autism (Khachadourian et al., 2023). The size of our study cohort (N=96, with n=38 diagnosed with autism) is larger when compared to several recently published studies using urine-based mass spectrometry, such as 60 available samples from an otherwise larger cohort (Liu et al., 2022), 57 participants (Mussap et al., 2020) and 14 autism discordant sibling pairs (Timperio et al., 2022). The twin cohort also gives us the possibility to analyse the effects of genetic and environmental factors on these potential biomarkers. Despite this, we were not able to find any metabolites that reached the multiple test correction significance. However, a few were nominally associated similarly in the whole cohort and within twin pair analyses (Figure 2A). Our results show that the contribution of the shared factors, including genetics, in the metabolite profiles is strong and therefore it is difficult to find robust biomarkers of autism.

Our study is also the first to use autistic traits as measured by the SRS-2 as a predictor of metabolite changes. Predominantly, we found a large number of metabolites being associated with autistic traits rather than autism diagnosis. This shows potential that in the future, metabolite profiles could be used for detecting differences in continuous autism operationalisations. We observed a significant positive association between indole-3-acetate and autistic traits within the twin pairs. Elevated levels have been previously detected in autistic individuals (Gevi et al., 2016; Mussap et al., 2020). However, a decline has also been reported (Emond et al., 2013). Indole-3-acetate is a known by-product of the tryptophan metabolism modulated by the gut microbiome (Krishnan et al., 2018; Pavlova et al., 2017), with growing evidence indicating a dysbiosis (Kang et al., 2013; Wan et al., 2022).

Despite the fact that we only found nominally associated metabolites, our findings align with previously reported urine-based metabolomic studies for autism diagnosis. For instance, we observed elevated levels of phenylpyruvate (Noto et al., 2014; Timperio et al., 2022) and taurine (Ma et al., 2021; Mavel et al., 2013; Nadal-Desbarats et al., 2014; Yap et al., 2010), as noted earlier in several studies. We also show decreased levels of carnitine, similarly to what was reported in other biological samples (Fahmy et al., 2013; Filipek et al., 2004). However, direct comparisons with metabolomic studies using different biological samples should be critically evaluated due to lack of information about any correlations between such reports, as observed in other multifactorial pathologies (Erben et al., 2021). Our observations are unique when it comes to the association with autistic traits, as these have not been used as a predictor in metabolomic studies of autism before. While individual metabolites provide a snapshot into specific altered metabolic processes, interpretation of pathway level changes may be more relevant in multifactorial conditions like autism (Arora et al., 2023).

Several of the enriched pathways that were detected in this study are related to energy metabolism and mitochondrial function, which are crucial in the developing brain (Bülow et al., 2022). There are earlier studies that have linked such metabolic changes with autism physiology. These include the genetics (Varga et al., 2018) and functional differences (Giulivi et al., 2010) that contribute to atypical mitochondrial function in autism (Frye, 2020; Rossignol & Frye, 2012). We have also identified pathways related to amino acid metabolism, although with relatively lower confidence. This may nevertheless still prove to be an interesting aspect to investigate further, particularly with the recently growing evidence in this direction (Chen et al., 2022; Tărlungeanu et al., 2016; Zhu et al., 2022).

Even with a relatively larger cohort, our study like several others is limited by the size of the clinical cohort to identify robust and reproducible urine-based biomarkers of autism, if

there are any. Furthermore, since urine is an excretory by-product, dietary choices can influence the urine metabolome (Stratakis et al., 2022), both in neurotypical and autistic individuals. Future studies should aim to include this information in their analyses, accompanied with knowledge about the gut-microbiome status of study participants, as this can also be potentially associated with autism status (Taniya et al., 2022). It would also be essential to account for puberty related changes in the metabolism of adolescents while characterising the metabolome (Cominetti et al., 2020; Perng et al., 2019). Lastly, there is a need to replicate such studies in several international cohorts to test the value of identified biomarkers in different populations.

In conclusion, we have identified several urinary metabolites that show nominal association with autism diagnosis status and autistic traits. We have also determined the enrichment of the significant features in metabolic pathways. While our findings highlight no robust urine-based biomarkers for autism, this ever-increasing knowledgebase can serve as signposts for investigations in the future.

Author contributions

Conceptualisation: A.A., S.B. and K.T.; Methodology: A.A. and K.T.; Software development: A.A. and K.T.; Validation: A.A. and K.T.; Formal Analysis: A.A.; Investigation: A.A.; Resources: S.B. and K.T.; Data Curation: A.A.; Writing – Original Draft: A.A., F.M. and K.T.; Writing – Review & Editing: A.A., F.M., S.B. and K.T.; Visualisation: A.A. and K.T.; Supervision: S.B. and K.T.; Project Administration: A.A. and K.T.; Funding Acquisition: S.B. and K.T.

Acknowledgements

We would like to thank all twins and parents who have participated in this research. We would also like to thank the RATSS team, especially Anna Pilfalk, Dr. Karl Lundin Remnelius, Dr. Johan Isaksson and Dr. Janina Neufeld for their valuable contribution to the work presented in this study. Additionally, we would like thank Prof. Antonio Persico for early discussion on designing the project. We acknowledge the Swedish Twin Registry for data access and urine analysis by the Proteomics and Metabolomics Facility, Department of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy and Prof. Lello Zolla.

Funding information

The project was supported by the Swedish Research Council (S.B., and K.T.), Swedish Foundation for Strategic Research (K.T.), the Swedish Brain Foundation – Hjärnfonden (K.T.), the Harald and Greta Jeanssons Foundations (K.T.), Åke Wiberg Foundation (K.T.), Strategic Research Area Neuroscience Stratneuro (K.T.), The Swedish Foundation for International Cooperation in Research and Higher Education STINT (K.T.), and Board of Research at Karolinska Institutet (K.T.). Open access funding provided by Karolinska Institutet.

Conflict of interest statement

A.A., F.M., and K.T. declare no competing interests. S.B. declares no direct conflict of interest related to this article. He discloses that he has in the last three years acted as an author, consultant, or lecturer for Medice and Roche. He receives royalties for textbooks and diagnostic tools from Hogrefe, Kohlhammer, UTB and Liber. S.B. is a partner in NeuroSupportSolutions International AB.

Ethics statement

The study was approved by the Swedish Ethical Review Authority (2016/1452-31).

Data availability statement

The mass spectrometry data is accessible from the corresponding author after necessary clearances. The utilised code is available on GitHub (https://github.com/Tammimies-Lab/RATSS-Metabolomics) or available upon request from the corresponding author.

Supporting information

Additional supporting information may be found in the online version of the article at the publisher's website.

References

Amaral, D. G., Anderson, G. M., Bailey, A., Bernier, R., Bishop, S., Blatt, G., Canal-Bedia, R., Charman, T., Dawson, G., de Vries, P. J., Dicicco-Bloom, E., Dissanayake, C., Kamio, Y., Kana, R., Khan, N. Z., Knoll, A., Kooy, F., Lainhart, J., Levitt, P., … Whitehouse, A. (2019). Gaps in current autism research: The thoughts of the Autism

Research Editorial Board and Associate Editors. *Autism Research*. https://doi.org/10.1002/aur.2101

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Arora, A., Becker, M., Marques, C., Oksanen, M., Li, D., Mastropasqua, F., Watts, M. E., Arora, M., Falk, A., Daub, C. O., Lanekoff, I., & Tammimies, K. (2023). Screening autism-associated environmental factors in differentiating human neural progenitors with fractional factorial design-based transcriptomics. *Scientific Reports*, *13*(1), 10519. https://doi.org/10.1038/s41598-023-37488-0

Bai, D., Yip, B. H. K., Windham, G. C., Sourander, A., Francis, R., Yoffe, R., Glasson, E., Mahjani, B., Suominen, A., Leonard, H., Gissler, M., Buxbaum, J. D., Wong, K., Schendel, D., Kodesh, A., Breshnahan, M., Levine, S. Z., Parner, E. T., Hansen, S. N., … Sandin, S. (2019). Association of Genetic and Environmental Factors With Autism in a 5-Country Cohort. *JAMA Psychiatry*, *76*(10), 1035. https://doi.org/10.1001/jamapsychiatry.2019.1411

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bölte, S., Willfors, C., Berggren, S., Norberg, J., Poltrago, L., Mevel, K., Coco, C., Fransson, P., Borg, J., Sitnikov, R., Toro, R., Tammimies, K., Anderlid, B.-M., Nordgren, A., Falk, A., Meyer, U., Kere, J., Landén, M., Dalman, C., … Lichtenstein, P. (2014). The Roots of Autism and ADHD Twin Study in Sweden (RATSS). *Twin Research and Human Genetics*, *17*(03), 164–176. https://doi.org/10.1017/thg.2014.12

Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., Knox, C., Bjorndahl, T. C., Krishnamurthy, R., Saleem, F., Liu, P., Dame, Z. T., Poelzer, J., Huynh, J., Yallou, F. S., Psychogios, N., Dong, E., Bogumil, R., Roehring, C., & Wishart, D. S. (2013). The Human Urine Metabolome. *PLoS ONE*, *8*(9), e73076. https://doi.org/10.1371/journal.pone.0073076

Bülow, P., Patgiri, A., & Faundez, V. (2022). Mitochondrial protein synthesis and the bioenergetic cost of neurodevelopment. *IScience*, *25*(9), 104920. https://doi.org/10.1016/j.isci.2022.104920

Chen, W.-X., Liu, B., Zhou, L., Xiong, X., Fu, J., Huang, Z.-F., Tan, T., Tang, M., Wang, J., & Tang, Y.-P. (2022). De novo mutations within metabolism networks of amino acid/protein/energy in Chinese autistic children with intellectual disability. *Human Genomics*, *16*(1), 52. https://doi.org/10.1186/s40246-022-00427-7

Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S. R., Woodhouse, E., Gillan, N., Hallett, V., Lietz, S., Garnett, T., Ronald, A., Plomin, R., Rijsdijk, F., Happé, F., & Bolton, P. (2015). Heritability of Autism Spectrum Disorder in a UK Population-Based Twin Sample. *JAMA Psychiatry*, *72*(5), 415. https://doi.org/10.1001/jamapsychiatry.2014.3028

Cominetti, O., Hosking, J., Jeffery, A., Pinkney, J., & Martin, F.-P. (2020). Contributions of Fat and Carbohydrate Metabolism to Glucose Homeostasis in Childhood Change With Age and Puberty: A 12-Years Cohort Study (EARLYBIRD 77). *Frontiers in Nutrition*, *7*, 139. https://doi.org/10.3389/fnut.2020.00139

Constantino, J. N. (2005). *Social Responsivness Scale (SRS)*. Western Psychological Services.

Constantino, J. N. (2011). The Quantitative Nature of Autistic Social Impairment: *Pediatric Research*, *69*(5 Part 2), 55R-62R. https://doi.org/10.1203/PDR.0b013e318212ec6e

Constantino, J. N., & Gruber, C. P. (2012). *Skattningsskala för social responsivitet, andra upplagan (SRS-2) [Social Responsiveness Scale, Second Edition (SRS-2)] [Swedish version by Zander, E. & Bölte, S.]*. Hogrefe.

Cortese, S., Solmi, M., Michelini, G., Bellato, A., Blanner, C., Canozzi, A., Eudave, L., Farhat, L. C., Højlund, M., Köhler-Forsberg, O., Leffa, D. T., Rohde, C., de Pablo, G. S., Vita, G., Wesselhoeft, R., Martin, J., Baumeister, S., Bozhilova, N. S., Carlisi, C. O., … Correll, C. U. (2023). Candidate diagnostic biomarkers for neurodevelopmental disorders in children and adolescents: A systematic review. *World Psychiatry*, *22*(1), 129–149. https://doi.org/10.1002/wps.21037

Emond, P., Mavel, S., Aïdoud, N., Nadal-Desbarats, L., Montigny, F., Bonnet-Brilhault, F., Barthélémy, C., Merten, M., Sarda, P., Laumonnier, F., Vourc'h, P., Blasco, H., & Andres, C. R. (2013). GC-MS-based urine metabolic profiling of autism spectrum disorders. *Analytical and Bioanalytical Chemistry*, *405*(15), 5291–5300. https://doi.org/10.1007/s00216-013-6934-x

Erben, V., Poschet, G., Schrotz-King, P., & Brenner, H. (2021). Comparing Metabolomics Profiles in Various Types of Liquid Biopsies among Screening Participants with and without Advanced Colorectal Neoplasms. *Diagnostics*, *11*(3), 561. https://doi.org/10.3390/diagnostics11030561

Fahmy, S. F., El-hamamsy, M. H., Zaki, O. K., & Badary, O. A. (2013). L-Carnitine supplementation improves the behavioral symptoms in autistic children. *Research in Autism Spectrum Disorders*, *7*(1), 159–166. https://doi.org/10.1016/j.rasd.2012.07.006

Filipek, P. A., Juranek, J., Nguyen, M. T., Cummings, C., & Gargus, J. J. (2004). Relative Carnitine Deficiency in Autism. *Journal of Autism and Developmental Disorders*, *34*(6), 615–623. https://doi.org/10.1007/s10803-004-5283-1

First, M. B., & Gibbon, M. (2004). *The structured clinical interview for DSM-IV axis I disorders (SCID-I) and the structured clinical interview for DSM-IV axis II disorders (SCID-II).*

Frazier, T. W., Thompson, L., Youngstrom, E. A., Law, P., Hardan, A. Y., Eng, C., & Morris, N. (2014). A Twin Study of Heritable and Shared Environmental Contributions to Autism. *Journal of Autism and Developmental Disorders*, *44*(8), 2013–2025. https://doi.org/10.1007/s10803-014-2081-2

Frye, R. E. (2020). Mitochondrial Dysfunction in Autism Spectrum Disorder: Unique Abnormalities and Targeted Treatments. *Seminars in Pediatric Neurology*, *35*, 100829. https://doi.org/10.1016/j.spen.2020.100829

Gevi, F., Zolla, L., Gabriele, S., & Persico, A. M. (2016). Urinary metabolomics of young Italian autistic children supports abnormal tryptophan and purine metabolism. *Molecular Autism*, *7*(1), 47. https://doi.org/10.1186/s13229-016-0109-5

Giulivi, C., Zhang, Y.-F., Omanska-Klusek, A., Ross-Inta, C., Wong, S., Hertz-Picciotto, I., Tassone, F., & Pessah, I. N. (2010). Mitochondrial Dysfunction in Autism. *JAMA*. https://doi.org/10.1001/jama.2010.1706

Harrison, P. L., & Oakland, T. (2000). *Adaptive behavior assessment system*. Psychological Corporation San Antonio, TX.

Himwich, H. E., Jenkins, R. L., Fujimori, M., Narasimhachari, N., & Ebersole, M. (1972). A biochemical study of early infantile autism. *Journal of Autism and Childhood Schizophrenia*, *2*(2), 114–126. https://doi.org/10.1007/BF01537565

Hirosawa, T., Kontani, K., Fukai, M., Kameya, M., Soma, D., Hino, S., Kitamura, T., Hasegawa, C., An, K., Takahashi, T., Yoshimura, Y., & Kikuchi, M. (2020). Different associations between intelligence and social cognition in children with and without

autism spectrum disorders. *PLOS ONE*, *15*(8), e0235380. https://doi.org/10.1371/journal.pone.0235380

Jensen, A. R., Lane, A. L., Werner, B. A., McLees, S. E., Fletcher, T. S., & Frye, R. E. (2022). Modern Biomarkers for Autism Spectrum Disorder: Future Directions. *Molecular Diagnosis & Therapy*, *26*(5), 483–495. https://doi.org/10.1007/s40291-022-00600-7

Jewison, T., Su, Y., Disfany, F. M., Liang, Y., Knox, C., Maciejewski, A., Poelzer, J., Huynh, J., Zhou, Y., Arndt, D., Djoumbou, Y., Liu, Y., Deng, L., Guo, A. C., Han, B., Pon, A., Wilson, M., Rafatnia, S., Liu, P., & Wishart, D. S. (2014). SMPDB 2.0: Big Improvements to the Small Molecule Pathway Database. *Nucleic Acids Research*, *42*(D1), D478–D484. https://doi.org/10.1093/nar/gkt1067

Josse, J., & Husson, F. (2016). MissMDA: A Package for Handling Missing Values in Multivariate Data Analysis. *Journal of Statistical Software*, *70*(1). https://doi.org/10.18637/jss.v070.i01

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., & Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research*, *51*(D1), D587–D592. https://doi.org/10.1093/nar/gkac963

Kang, D.-W., Park, J. G., Ilhan, Z. E., Wallstrom, G., LaBaer, J., Adams, J. B., & Krajmalnik-Brown, R. (2013). Reduced Incidence of Prevotella and Other Fermenters in Intestinal Microflora of Autistic Children. *PLoS ONE*, *8*(7), e68322. https://doi.org/10.1371/journal.pone.0068322

Khachadourian, V., Mahjani, B., Sandin, S., Kolevzon, A., Buxbaum, J. D., Reichenberg, A., & Janecka, M. (2023). Comorbidities in autism spectrum disorder and their etiologies. *Translational Psychiatry*, *13*(1), 71. https://doi.org/10.1038/s41398-023-02374-w

Kooij, J., & Francken, M. (2010). Diagnostic interview for ADHD in adults 2.0 (DIVA 2.0). *Adult ADHD: Diagnostic Assessment and Treatment*, 33–99.

Krishnan, S., Ding, Y., Saedi, N., Choi, M., Sridharan, G. V., Sherr, D. H., Yarmush, M. L., Alaniz, R. C., Jayaraman, A., & Lee, K. (2018). Gut Microbiota-Derived Tryptophan Metabolites Modulate Inflammatory Response in Hepatocytes and Macrophages. *Cell Reports*, *23*(4), 1099–1111. https://doi.org/10.1016/j.celrep.2018.03.109

Liu, W., Li, L., Xia, X., Zhou, X., Du, Y., Yin, Z., & Wang, J. (2022). Integration of Urine Proteomic and Metabolomic Profiling Reveals Novel Insights Into Neuroinflammation in Autism Spectrum Disorder. *Frontiers in Psychiatry*, *13*, 780747. https://doi.org/10.3389/fpsyt.2022.780747

Lord, C. (2019). Recognising the heterogeneity of autism. *The Lancet Psychiatry*, *6*(7), 551–552. https://doi.org/10.1016/S2215-0366(19)30220-2

Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones, E. J. H., Jones, R. M., Pickles, A., State, M. W., Taylor, J. L., & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature Reviews Disease Primers*, *6*(1), 5. https://doi.org/10.1038/s41572-019-0138-4

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). Autism diagnostic observation schedule–2nd edition (ADOS-2). *Los Angeles, CA: Western Psychological Corporation*, *284*.

Ma, Y., Zhou, H., Li, C., Zou, X., Luo, X., Wu, L., Li, T., Chen, X., Mao, M., Huang, Y., Li, E., An, Y., Zhang, L., Wang, T., Xu, X., Yan, W., Jiang, Y., & Wang, Y. (2021). Differential Metabolites in Chinese Autistic Children: A Multi-Center Study Based on Urinary 1H-NMR Metabolomics Analysis. *Frontiers in Psychiatry*, *12*, 624767. https://doi.org/10.3389/fpsyt.2021.624767

Mavel, S., Nadal-Desbarats, L., Blasco, H., Bonnet-Brilhault, F., Barthélémy, C., Montigny, F., Sarda, P., Laumonnier, F., Vourc′h, P., Andres, C. R., & Emond, P. (2013). 1H–13C NMR-based urine metabolic profiling in autism spectrum disorders. *Talanta*, *114*, 95–102. https://doi.org/10.1016/j.talanta.2013.03.064

Melamud, E., Vastag, L., & Rabinowitz, J. D. (2010). Metabolomic Analysis and Visualization Engine for LC−MS Data. *Analytical Chemistry*, *82*(23), 9818–9826. https://doi.org/10.1021/ac1021166

Mussap, M., Siracusano, M., Noto, A., Fattuoni, C., Riccioni, A., Rajula, H. S. R., Fanos, V., Curatolo, P., Barberini, L., & Mazzone, L. (2020). The Urine Metabolome of Young Autistic Children Correlates with Their Clinical Profile Severity. *Metabolites*, *10*(11), 476. https://doi.org/10.3390/metabo10110476

Myers, L., Pan, P., Remnélius, K. L., Neufeld, J., Marschik, P. B., Jonsson, U., & Bölte, S. (2021). Behavioral and biological divergence in monozygotic twin pairs discordant for autism phenotypes: A systematic review. *JCPP Advances*, *1*(2). https://doi.org/10.1111/jcv2.12017

Nadal-Desbarats, L., Aïdoud, N., Emond, P., Blasco, H., Filipiak, I., Sarda, P., Bonnet-Brilhault, F., Mavel, S., & Andres, C. R. (2014). Combined ^{1}H-NMR and ^{1}H– ^{13}C HSQC-NMR to improve urinary screening in autism spectrum disorders. *The Analyst*, *139*(13), 3460–3468. https://doi.org/10.1039/C4AN00552J

Noto, A., Fanos, V., Barberini, L., Grapov, D., Fattuoni, C., Zaffanello, M., Casanova, A., Fenu, G., De Giacomo, A., De Angelis, M., Moretti, C., Papoff, P., Ditonno, R., & Francavilla, R. (2014). The urinary metabolomics profile of an Italian autistic children population and their unaffected siblings. *The Journal of Maternal-Fetal & Neonatal Medicine*, *27*(sup2), 46–52. https://doi.org/10.3109/14767058.2014.954784

Pang, Z., Zhou, G., Ewald, J., Chang, L., Hacariz, O., Basu, N., & Xia, J. (2022). Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. *Nature Protocols*, *17*(8), 1735–1761. https://doi.org/10.1038/s41596-022-00710-w

Pavlova, T., Vidova, V., Bienertova-Vasku, J., Janku, P., Almasi, M., Klanova, J., & Spacil, Z. (2017). Urinary intermediates of tryptophan as indicators of the gut microbial metabolism. *Analytica Chimica Acta*, *987*, 72–80. https://doi.org/10.1016/j.aca.2017.08.022

Perng, W., Rifas-Shiman, S. L., Hivert, M.-F., Chavarro, J. E., Sordillo, J., & Oken, E. (2019). Metabolic trajectories across early adolescence: Differences by sex, weight, pubertal status and race/ethnicity. *Annals of Human Biology*, *46*(3), 205–214. https://doi.org/10.1080/03014460.2019.1638967

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rossignol, D. A., & Frye, R. E. (2012). Mitochondrial dysfunction in autism spectrum disorders: A systematic review and meta-analysis. *Molecular Psychiatry*, *17*(3), 290–314. https://doi.org/10.1038/mp.2010.136

Rutter, M., Le Couteur, A., & Lord, C. (2003). Autism diagnostic interview-revised. *Los Angeles, CA: Western Psychological Services*, *29*(2003), 30.

Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J.-Y., Peng, M., Collins, R., Grove, J., Klei, L., Stevens, C., Reichert, J., Mulhern, M. S., Artomov, M., Gerges, S., Sheppard, B., Xu, X., Bhaduri, A., Norman, U., … Walters, R. K. (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*, S0092867419313984. https://doi.org/10.1016/j.cell.2019.12.036

Stratakis, N., Siskos, A. P., Papadopoulou, E., Nguyen, A. N., Zhao, Y., Margetaki, K., Lau, C.-H. E., Coen, M., Maitre, L., Fernández-Barrés, S., Agier, L., Andrusaityte, S., Basagaña, X., Brantsaeter, A. L., Casas, M., Fossati, S., Grazuleviciene, R., Heude, B., McEachan, R. R., … Chatzi, L. (2022). Urinary metabolic biomarkers of diet quality in European children are associated with metabolic health. *ELife*, *11*, e71332. https://doi.org/10.7554/eLife.71332

Taniya, M. A., Chung, H.-J., Al Mamun, A., Alam, S., Aziz, Md. A., Emon, N. U., Islam, Md. M., Hong, S.-T. shool, Podder, B. R., Ara Mimi, A., Aktar Suchi, S., & Xiao, J. (2022). Role of Gut Microbiome in Autism Spectrum Disorder and Its Therapeutic Regulation. *Frontiers in Cellular and Infection Microbiology*, *12*, 915701. https://doi.org/10.3389/fcimb.2022.915701

Tărlungeanu, D. C., Deliu, E., Dotter, C. P., Kara, M., Janiesch, P. C., Scalise, M., Galluccio, M., Tesulov, M., Morelli, E., Sonmez, F. M., Bilguvar, K., Ohgaki, R., Kanai, Y., Johansen, A., Esharif, S., Ben-Omran, T., Topcu, M., Schlessinger, A., Indiveri, C., … Novarino, G. (2016). Impaired Amino Acid Transport at the Blood Brain Barrier Is a Cause of Autism Spectrum Disorder. *Cell*, *167*(6), 1481-1494.e18. https://doi.org/10.1016/j.cell.2016.11.013

Timperio, A. M., Gevi, F., Cucinotta, F., Ricciardello, A., Turriziani, L., Scattoni, M. L., & Persico, A. M. (2022). Urinary Untargeted Metabolic Profile Differentiates Children with Autism from Their Unaffected Siblings. *Metabolites*, *12*(9), 797. https://doi.org/10.3390/metabo12090797

Tulsky, D. S., Saklofske, D. H., Wilkins, C., & Weiss, L. G. (2001). Development of a General Ability Index for the Wechsler Adult Intelligence Scale—Third Edition. *Psychological Assessment*, *13*, 566–571. https://doi.org/10.1037/1040-3590.13.4.566

Varga, N. Á., Pentelényi, K., Balicza, P., Gézsi, A., Reményi, V., Hársfalvi, V., Bencsik, R., Illés, A., Prekop, C., & Molnár, M. J. (2018). Mitochondrial dysfunction and autism: Comprehensive genetic analyses of children with autism and mtDNA deletion. *Behavioral and Brain Functions*, *14*(1), 4. https://doi.org/10.1186/s12993-018-0135-x

Wan, Y., Zuo, T., Xu, Z., Zhang, F., Zhan, H., Chan, D., Leung, T.-F., Yeoh, Y. K., Chan, F. K. L., Chan, R., & Ng, S. C. (2022). Underdevelopment of the gut microbiota and bacteria species as non-invasive markers of prediction in children with autism spectrum disorder. *Gut*, *71*(5), 910–918. https://doi.org/10.1136/gutjnl-2020-324015

Whiteley, P., Waring, R., Williams, L., Klovrza, L., Nolan, F., Smith, S., Farrow, M., Dodou, K., Lough, W. J., & Shattock, P. (2006). Spot urinary creatinine excretion in pervasive developmental disorders. *Pediatrics International*, *48*(3), 292–297. https://doi.org/10.1111/j.1442-200X.2006.02207.x

Wickham, H. (2009). *Ggplot2*. Springer New York. https://doi.org/10.1007/978-0-387-98141-3

Worley, B., & Powers, R. (2016). PCA as a Practical Indicator of OPLS-DA Model Reliability. *Current Metabolomics*, *4*(2), 97–103. https://doi.org/10.2174/2213235X04666160613122429

Yap, I. K. S., Angley, M., Veselkov, K. A., Holmes, E., Lindon, J. C., & Nicholson, J. K. (2010). Urinary Metabolic Phenotyping Differentiates Children with Autism from Their Unaffected Siblings and Age-Matched Controls. *Journal of Proteome Research*, *9*(6), 2996–3004. https://doi.org/10.1021/pr901188e

Young, J. G., Cohen, D. J., Brown, S. L., & Caparulo, B. K. (1978). Decreased urinary free catecholamines in childhood autism. *Journal of the American Academy of Child Psychiatry*, *17*(4), 671–678. https://doi.org/10.1016/s0002-7138(09)61019-0

Zetterqvist, J., & Sjölander, A. (2015). Doubly Robust Estimation with the R Package drgee. *Epidemiologic Methods*, *4*(1). https://doi.org/10.1515/em-2014-0021

Zhou, X., Feliciano, P., Shu, C., Wang, T., Astrovskaya, I., Hall, J. B., Obiajulu, J. U., Wright, J. R., Murali, S. C., Xu, S. X., Brueggeman, L., Thomas, T. R., Marchenko, O., Fleisch, C., Barns, S. D., Snyder, L. G., Han, B., Chang, T. S., Turner, T. N., … Chung, W. K. (2022). Integrating de novo and inherited variants in 42,607 autism cases identifies mutations in new moderate-risk genes. *Nature Genetics*, *54*(9), 1305–1319. https://doi.org/10.1038/s41588-022-01148-2

Zhu, J., Hua, X., Yang, T., Guo, M., Li, Q., Xiao, L., Li, L., Chen, J., & Li, T. (2022). Alterations in Gut Vitamin and Amino Acid Metabolism are Associated with Symptoms and Neurodevelopment in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *52*(7), 3116–3128. https://doi.org/10.1007/s10803-021-05066-w

Figure legends

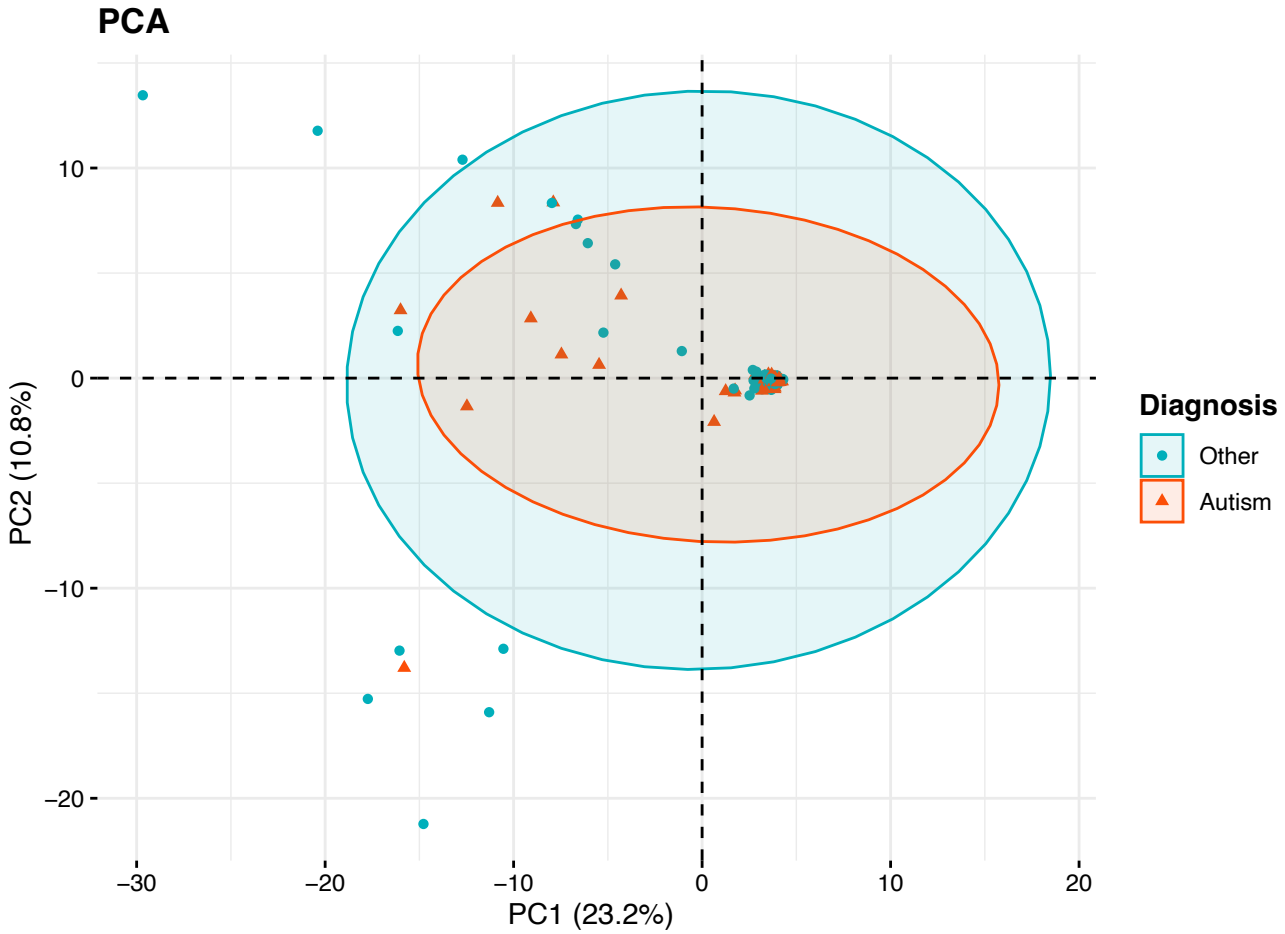
Figure 1. Principal Component Analysis (PCA) of study cohort. **(A)** Scatter plot for individuals based on metabolite contributions and grouped as per autism diagnosis status. **(B)** Scree plot of eigenvalues of principal components from A. **(C)** Loading plot of metabolites contributing to principal component 1 (PC1). **(D)** Loading plot of metabolites contributing to principal component 2 (PC2).

Figure 2. Heatmap of generalised estimating equations (GEE) model estimates. **(A)** Autism diagnosis across full cohort (N=96), monozygotic twins (n=52) and dizygotic twins (n=44) tested using GEE Models A and B. **(B)** Autistic traits across full cohort, monozygotic twins and dizygotic twins tested using GEE Models A and B. (* $p<0.05$, ** $p<0.01$, *** $p<0.001$)

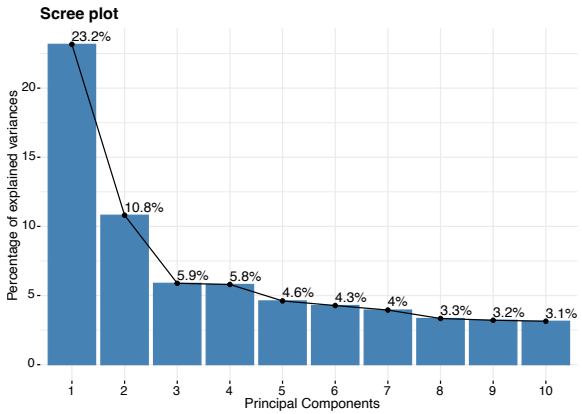
Figure 3. Pathway enrichment of significant metabolites based on autism diagnosis. **(A)** Enrichment dot-plot for top 10 enriched pathways from GEE Model A. **(B)** Interaction network for enriched pathways from GEE Model A. **(C)** Enrichment dot-plot for top 10 enriched pathways from GEE Model B. **(D)** Interaction network for enriched pathways from GEE Model B.

Figure 4. Pathway enrichment of significant metabolites based on autistic traits. **(A)** Enrichment dot-plot for top 10 enriched pathways from GEE Model A. **(B)** Interaction network for enriched pathways from GEE Model A. **(C)** Enrichment dot-plot for top 10 enriched pathways from GEE Model B. **(D)** Interaction network for enriched pathways from GEE Model B.

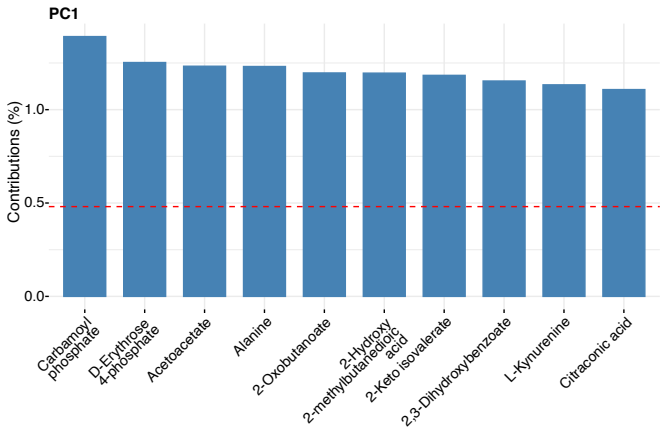
A Principal component analysis – Scatter plot



B Principal component analysis – Scree plot



C Loadings – Principal Component 1



D Loadings – Principal Component 2

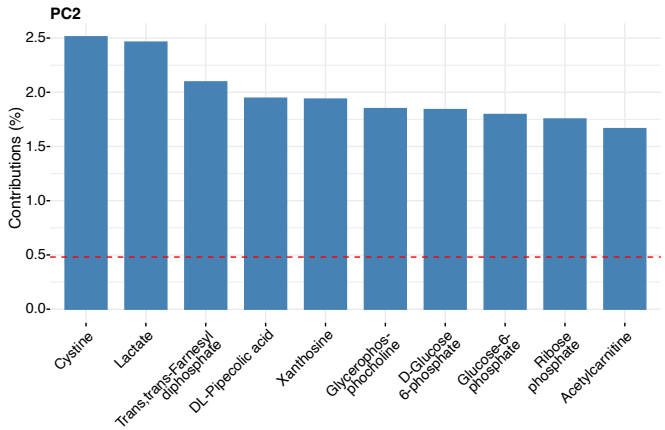


Figure 1

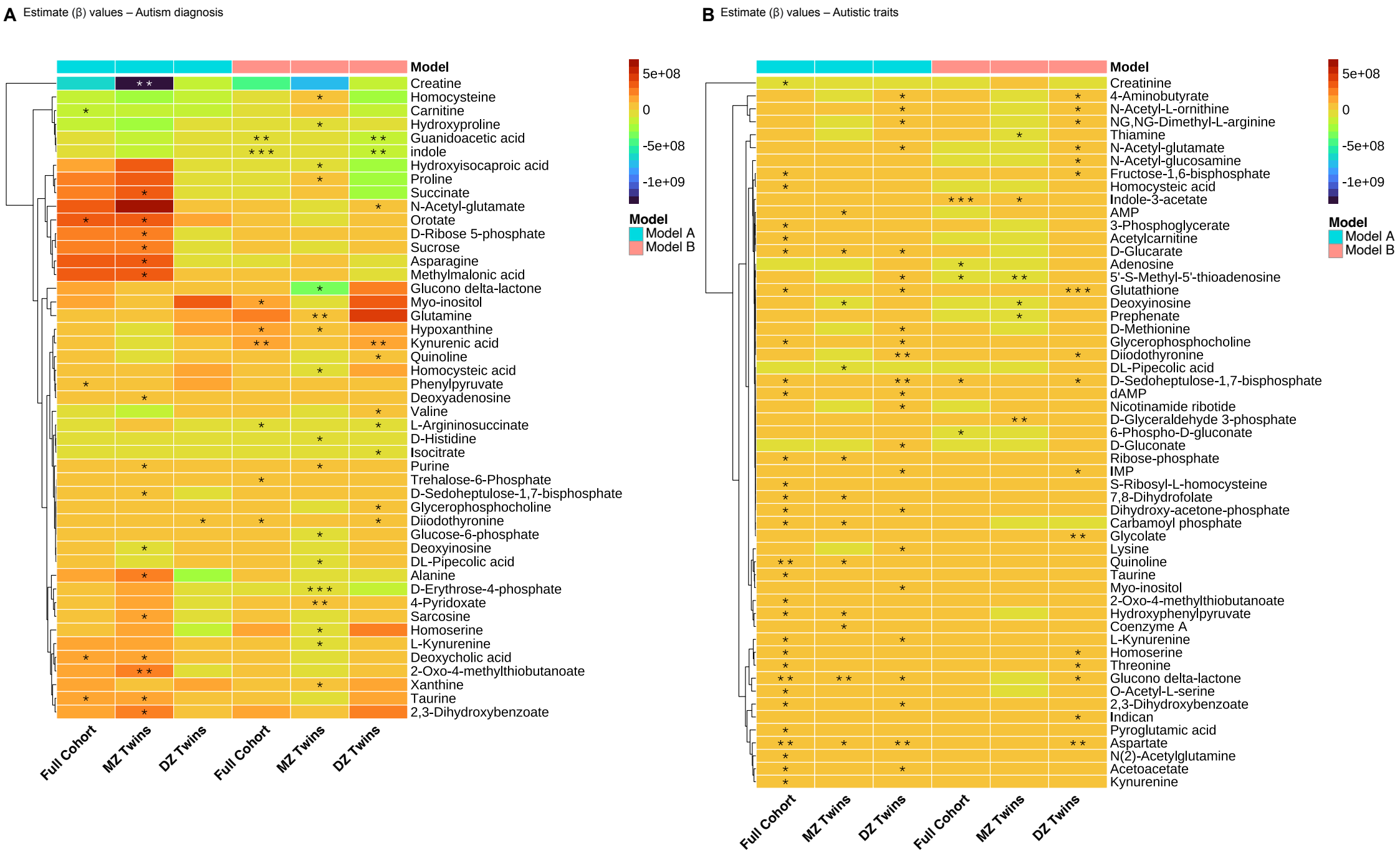

A
Estimate (β) values – Autism diagnosis
Model
Creatine
Homocysteine
Carnitine
Hydroxyproline
Guanidoacetic acid
indole
Hydroxyisocaproic acid
Proline
Succinate
N-Acetyl-glutamate
Orotate
D-Ribose 5-phosphate
Sucrose
Asparagine
Methylmalonic acid
Glucono delta-lactone
Myo-inositol
Glutamine
Hypoxanthine
Kynurenic acid
Quinoline
Homocysteic acid
Phenylpyruvate
Deoxyadenosine
Valine
L-Argininosuccinate
D-Histidine
Isocitrate
Purine
Trehalose-6-Phosphate
D-Sedoheptulose-1,7-bisphosphate
Glycerophosphocholine
Diiodothyronine
Glucose-6-phosphate
Deoxyinosine
DL-Pipecolic acid
Alanine
D-Erythrose-4-phosphate
4-Pyridoxate
Sarcosine
Homoserine
L-Kynurenine
Deoxycholic acid
2-Oxo-4-methylthiobutanoate
Xanthine
Taurine
2,3-Dihydroxybenzoate
Full Cohort
MZ Twins
DZ Twins
Full Cohort
MZ Twins
DZ Twins
5e+08
0
-5e+08
-1e+09
Model
Model A
Model B
B
Estimate (β) values – Autistic traits
Model
Creatinine
4-Aminobutyrate
N-Acetyl-L-ornithine
NG,NG-Dimethyl-L-arginine
Thiamine
N-Acetyl-glutamate
N-Acetyl-glucosamine
Fructose-1,6-bisphosphate
Homocysteic acid
Indole-3-acetate
AMP
3-Phosphoglycerate
Acetylcarnitine
D-Glucarate
Adenosine
5'-S-Methyl-5'-thioadenosine
Glutathione
Deoxyinosine
Prephenate
D-Methionine
Glycerophosphocholine
Diiodothyronine
DL-Pipecolic acid
D-Sedoheptulose-1,7-bisphosphate
dAMP
Nicotinamide ribotide
D-Glyceraldehyde 3-phosphate
6-Phospho-D-gluconate
D-Gluconate
Ribose-phosphate
IMP
S-Ribosyl-L-homocysteine
7,8-Dihydrofolate
Dihydroxy-acetone-phosphate
Carbamoyl phosphate
Glycolate
Lysine
Quinoline
Taurine
Myo-inositol
2-Oxo-4-methylthiobutanoate
Hydroxyphenylpyruvate
Coenzyme A
L-Kynurenine
Homoserine
Threonine
Glucono delta-lactone
O-Acetyl-L-serine
2,3-Dihydroxybenzoate
Indican
Pyroglutamic acid
Aspartate
N(2)-Acetylglutamine
Acetoacetate
Kynurenine
Full Cohort
MZ Twins
DZ Twins
Full Cohort
MZ Twins
DZ Twins
5e+08
0
-5e+08
-1e+09
Model
Model A
Model B

Figure 2

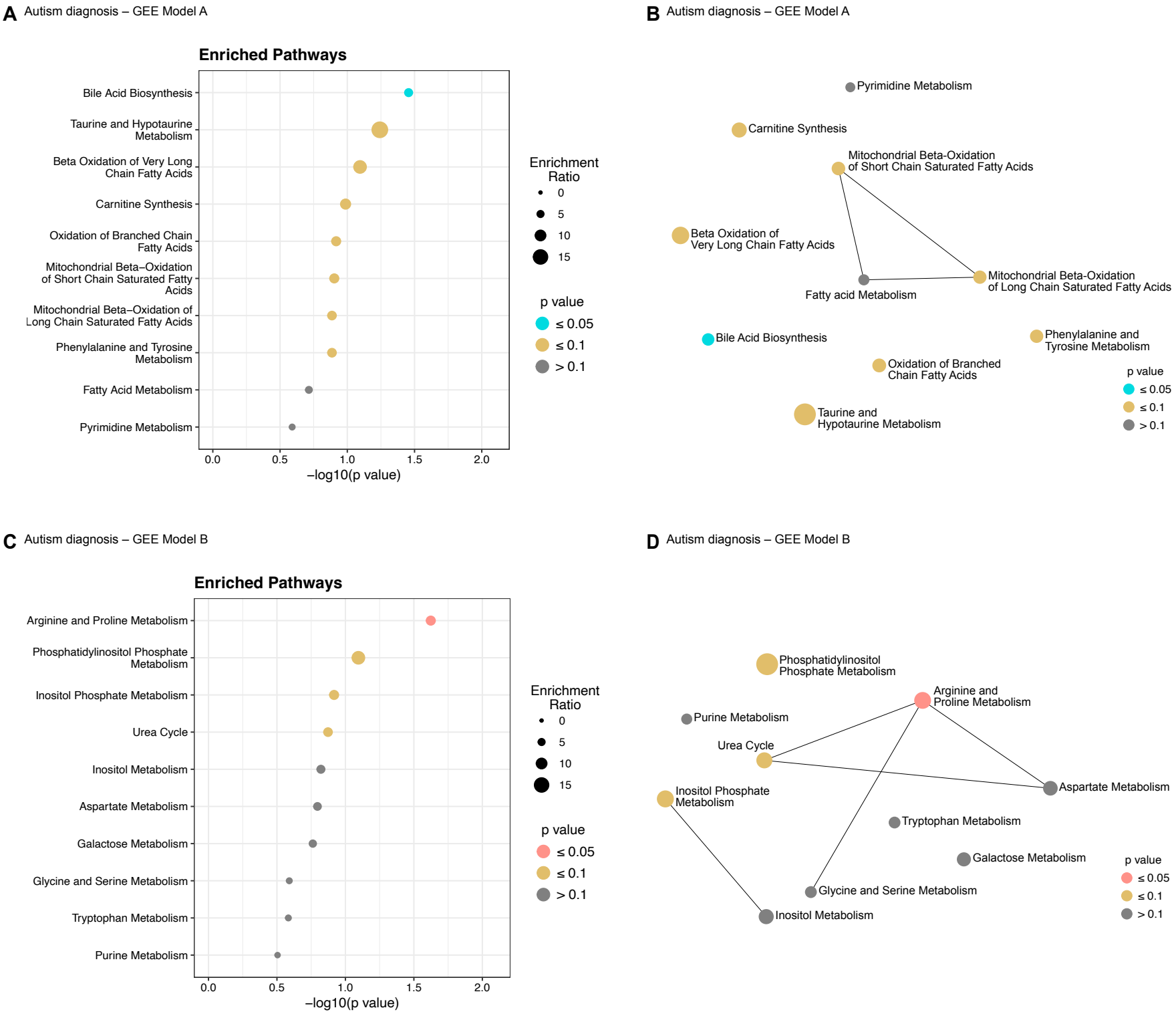

A
Autism diagnosis – GEE Model A
Enriched Pathways
Bile Acid Biosynthesis
Taurine and Hypotaurine Metabolism
Beta Oxidation of Very Long Chain Fatty Acids
Carnitine Synthesis
Oxidation of Branched Chain Fatty Acids
Mitochondrial Beta–Oxidation of Short Chain Saturated Fatty Acids
Mitochondrial Beta–Oxidation of Long Chain Saturated Fatty Acids
Phenylalanine and Tyrosine Metabolism
Fatty Acid Metabolism
Pyrimidine Metabolism
0.0
0.5
1.0
1.5
2.0
–log10(p value)
Enrichment Ratio
0
5
10
15
p value
≤ 0.05
≤ 0.1
> 0.1
B
Autism diagnosis – GEE Model A
Pyrimidine Metabolism
Carnitine Synthesis
Mitochondrial Beta-Oxidation of Short Chain Saturated Fatty Acids
Beta Oxidation of Very Long Chain Fatty Acids
Mitochondrial Beta-Oxidation of Long Chain Saturated Fatty Acids
Fatty acid Metabolism
Bile Acid Biosynthesis
Phenylalanine and Tyrosine Metabolism
Oxidation of Branched Chain Fatty Acids
Taurine and Hypotaurine Metabolism
p value
≤ 0.05
≤ 0.1
> 0.1
C
Autism diagnosis – GEE Model B
Enriched Pathways
Arginine and Proline Metabolism
Phosphatidylinositol Phosphate Metabolism
Inositol Phosphate Metabolism
Urea Cycle
Inositol Metabolism
Aspartate Metabolism
Galactose Metabolism
Glycine and Serine Metabolism
Tryptophan Metabolism
Purine Metabolism
0.0
0.5
1.0
1.5
2.0
–log10(p value)
Enrichment Ratio
0
5
10
15
p value
≤ 0.05
≤ 0.1
> 0.1
D
Autism diagnosis – GEE Model B
Phosphatidylinositol Phosphate Metabolism
Arginine and Proline Metabolism
Purine Metabolism
Urea Cycle
Aspartate Metabolism
Inositol Phosphate Metabolism
Tryptophan Metabolism
Galactose Metabolism
Glycine and Serine Metabolism
Inositol Metabolism
p value
≤ 0.05
≤ 0.1
> 0.1

Figure 3

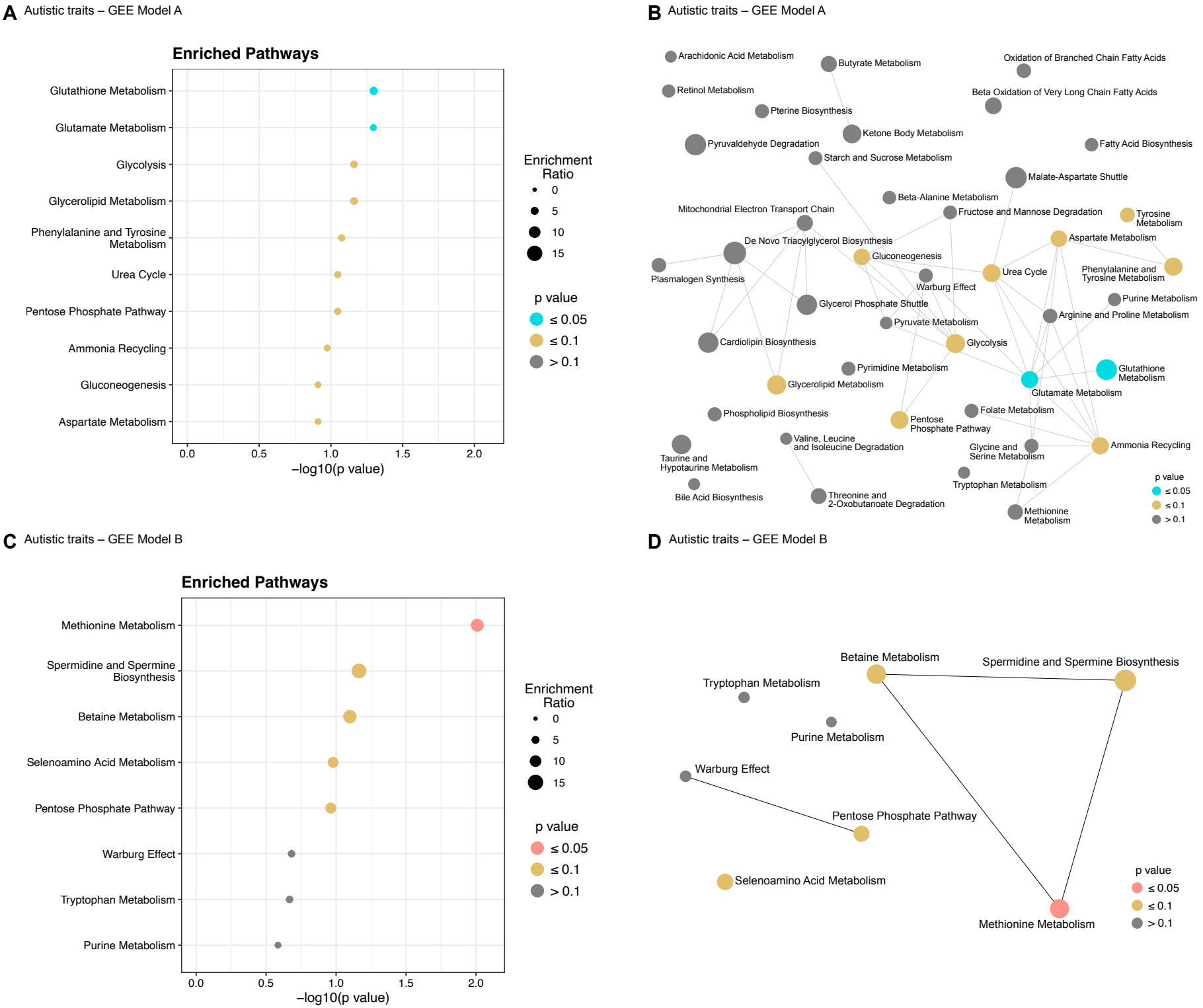

A Autistic traits – GEE Model A
Enriched Pathways
Glutathione Metabolism
Glutamate Metabolism
Glycolysis
Glycerolipid Metabolism
Phenylalanine and Tyrosine Metabolism
Urea Cycle
Pentose Phosphate Pathway
Ammonia Recycling
Gluconeogenesis
Aspartate Metabolism
0.0
0.5
1.0
1.5
2.0
−log10(p value)
Enrichment Ratio
0
5
10
15
p value
≤ 0.05
≤ 0.1
> 0.1
B Autistic traits – GEE Model A
Arachidonic Acid Metabolism
Butyrate Metabolism
Oxidation of Branched Chain Fatty Acids
Retinol Metabolism
Beta Oxidation of Very Long Chain Fatty Acids
Pterine Biosynthesis
Ketone Body Metabolism
Pyruvaldehyde Degradation
Fatty Acid Biosynthesis
Starch and Sucrose Metabolism
Malate-Aspartate Shuttle
Beta-Alanine Metabolism
Mitochondrial Electron Transport Chain
Fructose and Mannose Degradation
Tyrosine Metabolism
De Novo Triacylglycerol Biosynthesis
Aspartate Metabolism
Gluconeogenesis
Plasmalogen Synthesis
Urea Cycle
Phenylalanine and Tyrosine Metabolism
Warburg Effect
Purine Metabolism
Glycerol Phosphate Shuttle
Arginine and Proline Metabolism
Pyruvate Metabolism
Cardiolipin Biosynthesis
Glycolysis
Pyrimidine Metabolism
Glutathione Metabolism
Glycerolipid Metabolism
Glutamate Metabolism
Phospholipid Biosynthesis
Folate Metabolism
Pentose Phosphate Pathway
Valine, Leucine and Isoleucine Degradation
Glycine and Serine Metabolism
Ammonia Recycling
Taurine and Hypotaurine Metabolism
Tryptophan Metabolism
Bile Acid Biosynthesis
Threonine and 2-Oxobutanoate Degradation
Methionine Metabolism
p value
≤ 0.05
≤ 0.1
> 0.1
C Autistic traits – GEE Model B
Enriched Pathways
Methionine Metabolism
Spermidine and Spermine Biosynthesis
Betaine Metabolism
Selenoamino Acid Metabolism
Pentose Phosphate Pathway
Warburg Effect
Tryptophan Metabolism
Purine Metabolism
0.0
0.5
1.0
1.5
2.0
−log10(p value)
Enrichment Ratio
0
5
10
15
p value
≤ 0.05
≤ 0.1
> 0.1
D Autistic traits – GEE Model B
Betaine Metabolism
Spermidine and Spermine Biosynthesis
Tryptophan Metabolism
Purine Metabolism
Warburg Effect
Pentose Phosphate Pathway
Selenoamino Acid Metabolism
Methionine Metabolism
p value
≤ 0.05
≤ 0.1
> 0.1

Figure 4