

# Potential serum biomarkers from a metabolomics study of autism

Han Wang, PhD; Shuang Liang, PhD; Maoqing Wang, PhD; Jingquan Gao, PhD; Caihong Sun, PhD; Jia Wang, PhD; Wei Xia, MD; Shiyong Wu, PhD; Susan J. Sumner, PhD; Fengyu Zhang, PhD; Changhao Sun, PhD; Lijie Wu, PhD

**Background:** Early detection and diagnosis are very important for autism. Current diagnosis of autism relies mainly on some observational questionnaires and interview tools that may involve a great variability. We performed a metabolomics analysis of serum to identify potential biomarkers for the early diagnosis and clinical evaluation of autism. **Methods:** We analyzed a discovery cohort of patients with autism and participants without autism in the Chinese Han population using ultra-performance liquid chromatography quadrupole time-of-flight tandem mass spectrometry (UPLC/Q-TOF MS/MS) to detect metabolic changes in serum associated with autism. The potential metabolite candidates for biomarkers were individually validated in an additional independent cohort of cases and controls. We built a multiple logistic regression model to evaluate the validated biomarkers. **Results:** We included 73 patients and 63 controls in the discovery cohort and 100 cases and 100 controls in the validation cohort. Metabolomic analysis of serum in the discovery stage identified 17 metabolites, 11 of which were validated in an independent cohort. A multiple logistic regression model built on the 11 validated metabolites fit well in both cohorts. The model consistently showed that autism was associated with 2 particular metabolites: sphingosine 1-phosphate and docosahexaenoic acid. **Limitations:** While autism is diagnosed predominantly in boys, we were unable to perform the analysis by sex owing to difficulty recruiting enough female patients. Other limitations include the need to perform test–retest assessment within the same individual and the relatively small sample size. **Conclusion:** Two metabolites have potential as biomarkers for the clinical diagnosis and evaluation of autism.

## Introduction

Autism (OMIM:209850) is a neurodevelopmental disorder that presents in childhood with clinical manifestations of impairment in language and communication, social interaction and responsiveness, and restricted and repetitive patterns of interest or behaviour.<sup>1</sup> Recent statistics estimate that the prevalence of autism has increased to 1 per 68 children younger than 8 years.<sup>2</sup>

Autism has a strong genetic component<sup>3,4</sup> and has been associated with several genetic variants, such as copy number variants disrupting the genes *NRXN1*, *CNTN4*, *NLGNs* and *SHANK3*<sup>5–10</sup> and de novo mutations (e.g. *CHD8*, *KATNAL2*, *SCN2A*, *NTNG1*).<sup>11–13</sup> Although these genes are functionally of interest and contribute to the understanding of the biology of autism,<sup>14</sup> they account for only a very small proportion of patients. Some of these rare single point genetic mutations, or even common genetic variants, have also been associated in-

consistently with autism.<sup>4,9,15,16</sup> At present, the diagnosis of autism relies mainly on some observational tools that may involve a great variability. Few biomarkers are established for autism. Identifying metabolomic biomarkers may contribute to the improvement of the clinical diagnosis of autism and provide some objective tools that could be used to evaluate the outcome of interventions or treatment.

Metabolomics is the study of a biologic process involving all metabolites that are end products of the cellular process in a whole organism. Identifying metabolite fingerprints between patients and healthy individuals that specific cellular processes leave behind may provide insight into the pathophysiology of human diseases.<sup>17</sup> Metabolomics has been applied to identify serum and urine metabolic markers in several complex human diseases, including obesity,<sup>18</sup> diabetes<sup>19</sup> and coronary artery disease.<sup>20</sup>

Evidence suggests that central nervous system disease may present with metabolic pathway disturbance.<sup>21</sup> Several

**Correspondence to:** L. Wu, Department of Child and Adolescent Health, or C. Sun, Department of Nutrition and Food Hygiene, School of Public Health, Harbin Medical University, 157 Baojian Road, Nangang District, Harbin 150081, China; wljhyd@ems.hrbmu.edu.cn; Sun2002changhao@yahoo.com

Submitted Jan. 9, 2014; Revised May 16, 2014; Revised Aug. 13, 2014; Revised Oct. 30, 2014; Revised Dec. 15, 2014; Accepted Mar. 20, 2015; Early-released Sept. 22, 2015

DOI: 10.1503/jpn.140009

©2016 8872147 Canada Inc.

metabolomic studies have been conducted to map antipsychotic effects and to identify potential metabolites as biomarkers for the diagnosis of schizophrenia.<sup>22,23</sup> A few metabolomic studies of autism have recently been carried out using different techniques in a variety of biospecimens. A urine study using a <sup>1</sup>H nuclear magnetic resonance (NMR) spectroscopy method has shown an elevated level of taurine and decreased glutamate in urine samples of patients with autism.<sup>24</sup> Another study conducted in Saudi Arabia has suggested plasma fatty acids as diagnostic markers for autism,<sup>25</sup> specifically showing an increase in most saturated fatty acids and a decrease in polyunsaturated fatty acids. Ming and colleagues,<sup>26</sup> using a combination of liquid and gas chromatography-based mass spectrometry, detected abnormal amino acid metabolism, increased oxidative stress and altered gut microbiomes in urinary specimens of individuals with autism-spectrum disorder (ASD). Emond and colleagues,<sup>27</sup> using a gas chromatography–mass spectrometry (GC-MS)–based approach in an analysis of urine samples, built a multivariate statistical model that captured global biochemical signatures of autistic individuals, which enabled them to be distinguished very well from healthy children. Kuwabara and colleagues,<sup>28</sup> using capillary electrophoresis time-of-flight mass spectrometry (CE-TOF MS) for high-throughput profiling of metabolite levels in plasma, identified deviated levels of plasma metabolites associated with oxidative stress and mitochondrial dysfunction in individuals with ASD.

Compared with other analytical technologies, ultra-performance liquid chromatography (UPLC) coupled with mass spectrometry (MS) has higher resolution, high sensitivity and rapid separation. This method has been widely used to investigate subtle metabolite alterations in complex mixtures.<sup>29–33</sup> Here, we report a serum metabolomic analysis of a cohort of patients with autism and healthy individuals in the Han Chinese population using ultra-performance liquid chromatography quadrupole time-of-flight tandem mass spectrometry (UPLC/Q-TOF MS/MS) to identify metabolites and potential serum biomarkers associated with autism. We validated the potential biomarkers using an independent autism case–control cohort.

## Methods

### *Study design and participants*

Our study involved a 2-stage design with a discovery cohort of autism cases and controls and an independent cohort of cases and controls for validation. The first cohort was used for a metabolomic discovery analysis of serum to identify metabolites that could contribute to the discrimination of cases and controls. Suggested metabolites were then validated in an independent cohort of cases and controls. Cases and controls in both cohorts were matched for sex and age. No participants used any dietary supplements, such as vitamins and fatty acids, or other medications. Before initiating the study, we obtained approval from the Ethics Committee of Harbin Medical University. Parents or legal guardians of all participants provided written informed consent for chil-

dren who participated in the study after all study procedures had been explained. The study was conducted in accordance with the Declaration of Helsinki.

We recruited patients with autism aged 3–6 years through the Child Development and Behavioural Research Center (CDBRC) of the Harbin Medical University in Harbin, the capital metropolitan city of Heilongjiang Province in north-eastern China, between May 2010 and August 2011. All participants were from the Chinese Han population. The diagnosis of each patient was made independently by at least 2 experienced psychiatrists after evaluating potential participants according to DSM-IV criteria. We excluded individuals with Asperger syndrome, Rett syndrome, pervasive developmental disorder not otherwise specified (PPD-NOS) and fragile X syndrome. We used the Autism Behaviour Checklist (ABC),<sup>34</sup> Childhood Autism Rating Scales (CARS)<sup>35</sup> and the developmental quotient (DQ)<sup>36</sup> for development delay<sup>37</sup> diagnosis to evaluate all patients with autism. We recruited healthy controls from 2 kindergarten classes in Harbin, and they were clinically examined to ensure they had no features of developmental delay or autistic traits.

### *Biospecimen collection and processing*

We collected whole blood (5 mL) samples from each participant in the morning (7:30–8:30 am) following at least 10 hours of fasting. Each participant had been given a standardized dietary recipe and had engaged in moderate physical activity for 1 week before blood was drawn. Samples were transferred into evacuated tubes that contained anticoagulant sodium dihydrogen phosphate (EDTA-K3), centrifuged at 2000g for 10 minutes at 4°C. Serum was extracted and stored at –80°C for sample analysis.

Prior to analysis, serum samples were thawed at 4°C. In total 200 µL of serum was added into a vial and extracted with methanol (600 µL), followed by vigorous vortex for 2 min and then by centrifugation at 14 000g for 10 min at 4°C. Supernatant was taken up and dried under nitrogen at 37°C. Residues were redissolved in 400 µL of acetonitrile/water (1:2), vortexed for 2 min and centrifuged at 14 000g for 10 min at 4°C. The supernatant was placed into an autosampler vial for UPLC-QTOF-MS/MS analysis in both positive and negative electrospray ionization (ESI) modes.

### *Ultra-performance liquid chromatography and mass spectrometry analysis*

Both the discovery and validation cohort samples were processed and analyzed using the same instrument and methods. We transferred a pretreated sample (2 µL) into a 100 µm × 2.1 µm × 1.7 µm BEH C18 column (Waters Corporation) using an Acquity ultra-performance liquid chromatography system. The column temperature was maintained at 35°C, and the flow rate was 0.35 mL/min. Samples from 5 children with autism and 5 controls were run alternately. The gradient consisted of 2 solutions: 1) water with 0.1% formic acid and 2) acetonitrile. The proportioning of mobile phases in gradient elution were as follows. The positive mode involved

2%–20% acetonitrile for 0–1.5 min, 20%–70% for 1.5–6 min and 70%–98% for 6–10 min; the concentration was held at 98% for 2 min, then returned to 2% for 12–14 min and finally held at 2% for 14–16 min. The negative mode involved 2%–30% acetonitrile for 0–2.0 min, 30%–70% for 2–3 min, 70%–75% for 3–7.5 min and 75%–98% for 7.5–10.5 min; the concentration was held at 98% for 1.5 min, returned to 2% for 12–14 min and finally held at 2% for 14–16 min.

We performed MS analysis using a Waters Micromass Q-TOF spectrometer. The MS parameters were set up as follows: the instrument operated with a positive or negative ion mode, capillary voltage of 3000 V (positive) or 2800 V (negative), sample cone voltage of 35 V, desolvation temperature of 320°C, desolvation gas flow of 600 L/h, extraction cone voltage of 3.0 V, collision energy of 6 eV, source temperature of 110°C and cone gas flow of 15 L/h. Centroid data were collected from 50 to 1000 m/z, with a scan time of 0.4 s and an interscan delay of 0.1 s. To avoid possible contamination and keep the signal stable, the Q-TOF mass spectrometer system was tuned for optimum accuracy and reproducibility using leucine enkephalin for the positive ESI mode ( $[M+H]^+$  = 556.2771) and negative ESI mode ( $[M-H]^-$  = 554.2615).

#### *Analytical method validation and quality control*

Precision and repeatability, measured by relative standard deviation (RSD), are important in analytical chemistry. To assure precision and repeatability of our sample analysis, we used pooled samples in both the discovery cohort (total 40 samples: 6 for precision and 14 for repeatability for each the positive and negative mode) and the validation cohort (total 52 samples: 6 for precision and 20 for repeatability for each the positive and negative mode) analyses for quality control (QC).<sup>38–40</sup> The pooled samples were prepared by mixing equal volumes of serum (100  $\mu$ L) from 15 control and 15 autism samples in each analysis, respectively. We analyzed these QC samples together with study samples. Whereas precision was evaluated using 6 continuous injections of the QC sample before the study sample sequence, repeatability was evaluated using 1 QC sample injected for every 10 study samples (5 autism and 5 control). Six single ions with different m/z were randomly selected according to retention time. We examined precision and reproducibility of the QC sample using the RSD of retention time and the peak intensity of 6 ions (RSD < 15% by U.S. Food and Drug Administration guidance<sup>41</sup>). We also performed principal components analysis (PCA) on all QC samples.

At the beginning of sample analysis, a new UPLC column was used and equilibrated after running 20 blanks under the chromatographic conditions described previously. We ran the UPLC/Q-TOF MS/MS first in negative ion mode and then in positive ion mode in both the discovery and validation samples. To avoid sequence effect, we used a randomized crossover approach for running the samples, along with a blank sample to clean chromatographic carryover effect. The running sequence was 1 blank, 5 autism, 1 blank, 5 control, 1 QC and 1 blank, with roughly 70 samples per day. When an examination of the blank injections showed any

presence of carryover effect or long running peaks, we inserted a blank sample run into the sequence. All samples before and after the blank samples were reanalyzed to ensure the carryover effect was washed out. After negative and positive ion mode detection, the chromatographic system and MS interface was also flushed.

#### *Data preprocessing and analysis*

The UPLC/Q-TOF MS/MS data were imported to MarkerLynx Application Manager version 4.1 Masslynx SCN 714 (Waters Corporation). We used the MarkerLynx ApexTrack peak integration for peak detection and alignment. The ApexTrack peak parameters were set as follows: peak width at 5% height, 1 s, and peak to-peak baseline noise (calculated automatically). Collection parameters were set as follows: mass window 0.05 Da, retention time window 0.2 min, minimum intensity 80, noise elimination level 6.0, and "Yes" for deisotope data. After being recognized and aligned, we normalized the intensity of each ion to the summed total ion intensity of each chromatogram. The data-reduction process was handled in accordance with the 80% rule.<sup>38</sup> Only data in the range of 0.40–10.5 min were used. The data were in a 3-dimensional matrix containing retention time, mass-to-charge ratio (m/z) pairs and ion intensity information (variables). We performed a distribution-based data imputation<sup>42,43</sup> for some nondetected values. The data were then exported to SIMCA-P 11.5 software (Umetrics AB, UMEA) for further analysis. We used the Pareto scaling method for data transformation before multivariate statistical analysis.

In metabolomic analysis, the confirmation of metabolite molecular weight and further structure elucidation are considered the most challenging steps. We determined a molecular formula according to the exact mass and isotope pattern. Additional tandem mass spectrometry (MS/MS) experiments were carried out to identify metabolites. We conducted MS/MS experiments to obtain fragmentation patterns of selected metabolites for producing their structure information, and we calculated accurate molecular weight. The mass tolerance between measured m/z values and exact mass of the component of interest was set to within 30 ppm.<sup>44–46</sup> The UPLC-MS/MS product ion spectrum of a metabolite was matched with its structure information using MassFragment application manager software (MassLynx version 4.1, Waters Corporation). A metabolite was detected and identified based on accurate mass, MS/MS information and metabolite structure information from related databases: HMDB ([www.hmdb.ca](http://www.hmdb.ca)) and METLIN (<http://metlin.scripps.edu>). Finally, metabolites were confirmed by comparison of retention time and fragmentation pattern with authentic standards. In addition, resolution and mass accuracy were 5000 ppm and 10 ppm, respectively, which can provide high confidence level in compound identification.

As autism is more prevalent in boys than girls, we also analyzed male participants separately; we did not analyze female participants separately owing to insufficient samples, especially in the validation cohort. In the validation stage, we focused on analyzing those metabolites identified in the discovery stage;

metabolites were considered to be candidates if they could be replicated consistently in an independent sample.

We built a partial least squares discriminant analysis (PLS-DA) model to assess the classification of autism and control participants while optimizing the identification of changes in serum metabolites. It is a type of PLS regression where the dependent variable is a binary outcome (i.e., disease v. healthy), and the independent variables are multiple components derived from metabolites detected and selected based on criteria of variable importance in projection (VIP). We evaluated goodness of fit using R2Y, which is the proportion of the variation in Y explained by the model.<sup>47</sup> The predictability was measured using Q2, which is defined as 1 minus the ratio of the prediction error sum of squares (PRESS) to the (mean corrected) total sum of squares (TSS) of the response Y.<sup>47-50</sup> Values of R2Y and Q2 close to 1.0 indicate a better model.<sup>50,51</sup> We validated the PLS-DA model through permutation analysis (800 times) to reduce possible false-positive findings using SIMCA-P 11.5 (Umetrics AB, UMEA). This software displays the plots of correlation coefficients between the original Y and the permuted Y versus the cumulative R2Y and Q2 as well as fitted regression lines. The criteria for model validity are as follows: 1) all Q2 values on the permuted data set to the left are lower than the Q2 value on the actual data set to the right and 2) the regression line (line joining the point of observed Q2 to the centroid of a cluster of permuted Q2 values) has a negative value of intercept on the Y axis.<sup>47</sup>

Individual discriminating metabolites associated with autism and controls were selected primarily according to the VIP (> 1.0) values and a statistical test for difference ( $p < 0.05$ ) between patients with autism and controls. While some suggest using a VIP greater than 2<sup>52</sup> as a criterion for selecting an individual metabolite, we chose a VIP greater than 1, similar to other studies,<sup>53,54</sup> in order to analyze more metabolites. Our approach took into account both predicting capability and statistical significance of individual metabolites in association with autism. As some metabolites were not normally distributed, we used a *t* test and a nonparametric Mann-Whitney *U* test, which is robust to deviation from the assumption of data

distribution that parametric methods require, to test for the difference in metabolites between cases and controls for metabolomics discovery.

Finally, we performed multiple logistic regression analysis of top potential metabolites from discovery and validation analysis using stepwise selection (criteria: entry = 0.15, removal = 0.15). Receiver operating characteristics (ROC) analysis was used to evaluate predictive ability of potential metabolic biomarkers. Area under the curve (AUC), best cut-off point, sensitivity and specificity were determined using the maximum value of the Youden index. The analysis was performed using SPSS version 13.01. We performed stepwise logistic regression analysis of the validation sample to build a separate model to validate the multiple logistic regression model built on the discovery sample.

## Results

The discovery cohort, comprising 73 patients with autism and 63 healthy controls, underwent the metabolomic discovery analysis of serum to identify metabolites that could contribute to the discrimination of cases and controls. Suggested metabolites were then validated in the independent cohort of 100 cases and 100 controls. The discovery sample comprised 59 male and 14 female patients with autism (mean age  $4.6 \pm 0.8$  [range 3–6] yr) and 51 male and 12 female controls (mean age  $4.1 \pm 0.7$  [range 3–6] yr) matched for age and sex. The validation sample comprised 86 male and 14 female patients with autism (mean age  $4.4 \pm 0.8$  [range 3–6] yr) and 81 male and 19 female controls (mean age  $4.5 \pm 0.8$  [range 3–6] yr) matched for age and sex. The proportion of patients with developmental delay was 80.8% (59 of 73) in the discovery cohort and 74.0% (74 of 100) in the validation cohort. The characteristics of study participants are shown in Table 1.

### Serum metabolite profiling

The precision and repeatability of the sample analysis were excellent for both the discovery and validation samples (Appendix 1, Table S1, available at [jpn.ca](http://jpn.ca)). For the precision, the RSD for retention time ranged from 0.02% to 0.23% and for peak intensity from 2.26% to 7.37% in the discovery sample; for the repeatability, the RSD for retention time ranged from 0.04% to 0.22% and for peak intensity from 1.46% to 6.06%. These measurements were far below 15%, indicating the analytical methods were good according to the FDA guidance. The PCA of the QC sample also provided confidence that our sample running was stable (Appendix 1, Fig. S1A and B). Representative baseline peak intensity chromatograms of serum indicated that the sample metabolites attained suitable separation on UPLC C18 column by gradient elution. Typical single UPLC-QTOF/MS base peak intensity chromatograms of a patient with autism and a healthy control are presented in Figure 1. At a retention time of 2.88 min in the patient with autism and 2.87 min in the control the relative abundance was much higher in the patient (almost > 95%) than the control (50%). A total of 1516 and 1499 potential metabolites were determined in positive and negative modes, respectively, in the discovery analysis.

**Table 1: Demographic and clinical characteristics of participants in the discovery and validation studies**

Category	Discovery cohort		Validation cohort	
	Autism	Control	Autism	Control
No. participants	73	63	100	100
No. male:female	59:14	51:12	86:14	81:19
Age, mean $\pm$ SD, yr	$4.6 \pm 0.8$	$4.1 \pm 0.7$	$4.4 \pm 0.8$	$4.5 \pm 0.8$
DQ, no.				
Healthy	14	63	26	100
Delayed	59	0	74	0
ABC total score, mean $\pm$ SD	$89.1 \pm 28.6$	—	$95.3 \pm 28.7$	—
CARS total score, mean $\pm$ SD	$36.5 \pm 4.6$	—	$35.7 \pm 4.9$	—

ABC = Autism Behaviour Checklist; CARS = Childhood Autism Rating Scales; DQ = developmental quotient; SD = standard deviation.

The PLS-DA analysis of the discovery sample revealed a clear clustering trend between patients with autism and controls for metabolites in both the positive and negative mode (Fig. 2). Three principal components were extracted from metabolites in positive mode ( $R^2Y = 0.995$ ;  $Q^2 = 0.989$ ) and 3 in negative mode ( $R^2Y = 0.984$ ;  $Q^2 = 0.965$ ), and the models fit well. The PLS-DA models built on permuted data sets showed that the observed model was unlikely obtained by chance. Except for the 1 value from the actual data, all permuted  $R^2Y$ s were below or around 0.6 for positive and negative modes. Similarly, all permuted  $Q^2$ s were below or around 0 for positive and negative modes. All  $R^2Y$ s and  $Q^2$ s are much lower than the original points to the right (Fig. 3), suggesting the model fit was valid. This indicated that patients with autism had distinct metabolite alterations in serum.

#### *Difference in metabolic profiling between patients with autism and healthy controls*

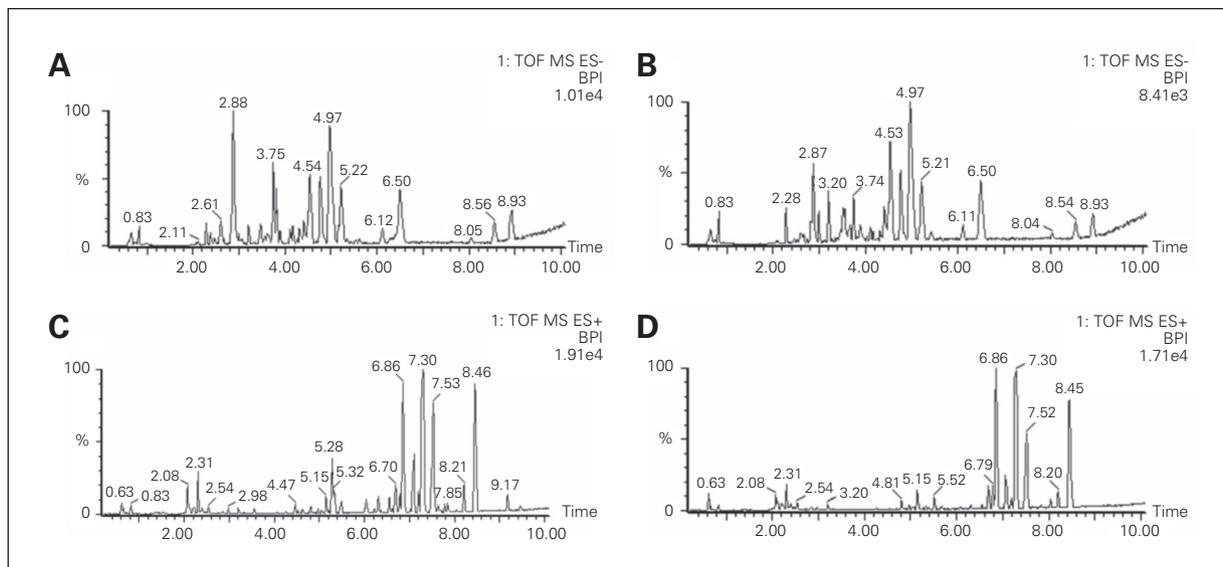
A total of 63 of 228 potential metabolites (or UPLC/Q-TOF MS peaks) in the positive mode and 72 of 218 in the negative mode that met the criterion of  $VIP > 1$  also showed significant differences ( $p < 0.05$ ) between patients with autism and controls. Only 17 metabolites (6 in the positive and 11 in the negative mode) were confirmed by available databases: phytosphingosine, arachidonic acid (ARA), docosapentaenoic acid (DPA), sphingosine 1-phosphate (S1P), uric acid (UA), adrenic acid, docosahexaenoic acid (DHA), LysoPE(0:0/20:2(11Z,14Z)), LysoPE (18:0/0:0), LPA (18:2(9Z,12Z)/0:0), LysoPE(0:0/16:0), L-acetylcarnitine, decanoylcarnitine, pregnanetriol, LysoPC (18:3(6Z,9Z,12Z)), LysoPC (20:3 (5Z,8Z,11Z)) and 9,10-epoxyoctadecenoic acid (9,10-EOA) (Table 2). More detailed descriptions of the relative contents for 17 differential metabolites are provided in Appendix 1, Table S2. Seven of 17 metabolites— phytosphingosine,

ARA, DPA, S1P, UA, adrenic acid and DHA — were identified using reference compounds; chemical structures and mass fragment information on principal metabolites are shown in Appendix 1, Figure S2. In particular, mass errors of adrenic acid and UA were 21.13 and 23.95, respectively, and were identified using both databases and reference compounds, suggesting that the analysis was appropriate for compound identification. Although authentic standards of the other 10 metabolites were not commercially available, they were supported by a library database (Appendix 1, Fig. S3). Unfortunately, the rest of the metabolites were neither available in a related database nor verified using a standard substance.

Analysis of the male-only sample also revealed a similar set of metabolites (data not shown). Sixteen of the 17 metabolites differed significantly between patients with autism and healthy controls in the discovery sample (all  $p < 0.05$ ; Appendix 1, Table S3); LysoPC (20:3(5Z, 8Z, 11Z)) did not differ between the groups ( $p = 0.07$ ). This was expected, because our samples were predominantly male and largely proportional to overall diagnosis of the disease.

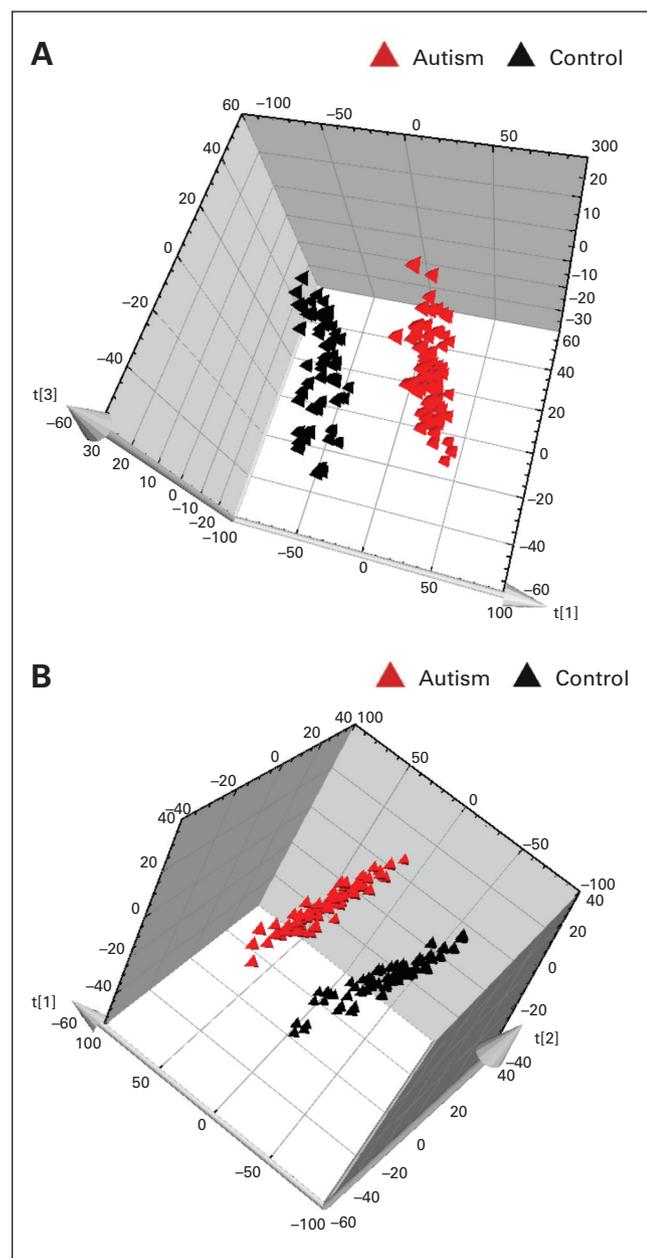
#### *Validation study of top metabolites associated with autism*

A validation study of 17 metabolites was performed in the independent cohort of cases and controls. As in the discovery analysis, the experiments on the validation sample were carried out with good precision and reproducibility (Appendix 1, Table S1 and Fig. S1C and D). We observed significant differences between patients with autism and controls in overall samples for 12 of 17 metabolites: 3 in the positive mode (decanoylcarnitine, phytosphingosine, pregnanetriol) and 9 in the negative mode (DHA, DPA, adrenic acid, S1P, LPA(18:2(9Z,12Z)/0:0), LysoPE(0:0/16:0), LysoPE(18:0/0:0), UA and 9,10-EOA). Unfortunately, 5 serum metabolites were



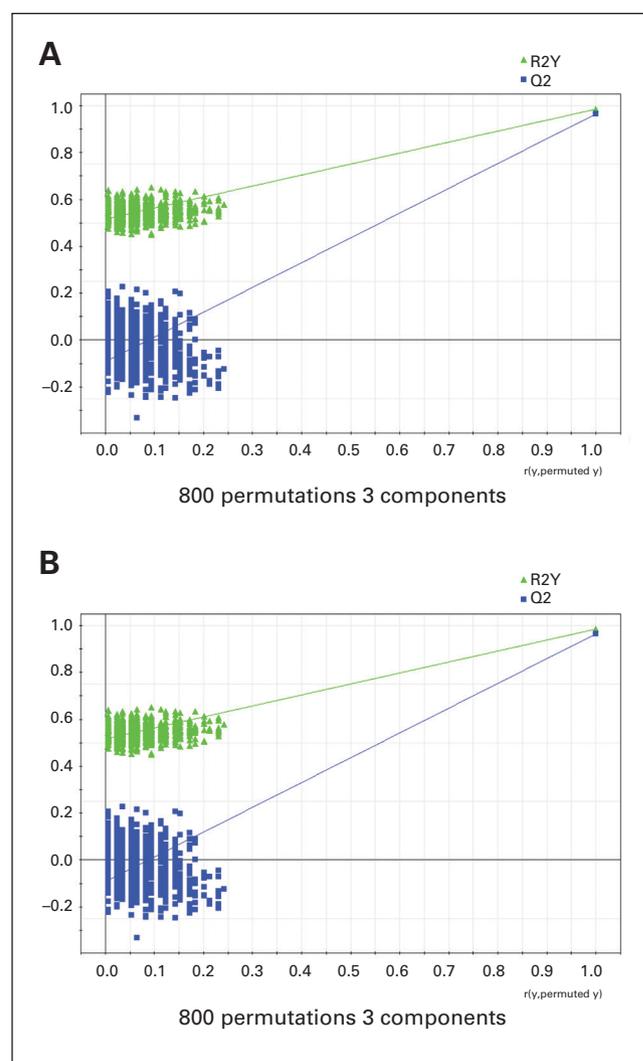
**Fig. 1:** Typical ultra-performance liquid chromatography quadrupole time-of-flight tandem mass spectrometry (UPLC-QTOF/MS) base peak intensity (BPI) chromatograms for 1 autism and 1 control serum sample in both the positive and negative modes. Positive mode: (A) autism versus (B) control sample; negative mode: (C) autism versus (D) control sample. The Y axis shows relative abundance (%), and the X axis shows retention time in minutes, which was used to identify individual metabolites.

not consistently validated (Appendix 1, Table S2). We noted that 1 metabolite, phytosphingosine, was validated in overall samples ( $p = 0.023$ ), but was not replicated in the male-only sample ( $p = 0.32$ ). Another metabolite, L-acetylcarnitine, was not validated in the overall sample ( $p = 0.11$ ), but appeared significant in the male-only validation sample ( $p = 0.026$ ; Appendix 1, Table S2 and S3). In summary, 11 of 17 metabolites were consistently validated in both overall and male-only samples,



**Fig. 2:** Partial least squares discriminant analysis 3-dimensional plot of principal component in children with autism and healthy controls, by positive and negative mode. **(A)** Positive mode:  $R^2Y = 0.995$ ,  $Q^2 = 0.989$ , 3 principal components. **(B)** Negative mode:  $R^2Y = 0.984$ ,  $Q^2 = 0.965$ , 3 principal components. Controls are shown in red, and patients with autism are shown in black.  $t[1]$  = first principal component;  $t[2]$  = second principal component;  $t[3]$  = third principal component.

and all of them had a good predictability. Specifically, S1P and LPA(18:2(9Z,12Z)/0:0) each had a good predictability of disease status ( $AUC > 0.80$ ; Table 3). We also observed an association between the CARS score and decanoylcarnitine ( $r = -0.2564$ ,  $p = 0.07$ ) and pregnanetriol ( $r = -0.3240$ ,  $p = 0.020$ ) and between the ABC score and pregnanetriol ( $r = 0.2697$ ,  $p = 0.06$ ), DHA ( $r = -0.2611$ ,  $p = 0.06$ ), DPA ( $r = -0.3175$ ,  $p = 0.023$ ) and S1P ( $r = 0.2746$ ,  $p = 0.05$ ; Table 4).



**Fig. 3:** Plot of  $R^2Y$  and  $Q^2$  from 800 permutation tests in partial least squares discriminant analysis models for the metabolomics analysis in the discovery sample in the **(A)** positive and **(B)** negative modes. The Y axis shows  $R^2Y$  and  $Q^2$ , and the X axis shows the correlation of observed and permuted data. The 2 points on the right side correspond to  $R^2Y$  and  $Q^2$  of the observed data set. Other points on the left side correspond to  $R^2Y$ s and  $Q^2$ s of permuted data sets. The 2 plots indicate that the 2 models were well-guarded against overfitting. The criteria for model validity are as follows: 1) all  $Q^2$  values on the permuted data set to the left are lower than the  $Q^2$  value on the actual data set to the right and 2) the regression line (line joining the point of observed  $Q^2$  to the centroid of a cluster of permuted  $Q^2$  values) has a negative value of intercept on the Y axis.<sup>47</sup>

Multiple logistic regression analysis

To assess how multiple metabolites collectively classify the disease status of autism, we built a logistic regression model using stepwise selection. The regression model was first built on 11 metabolites for the discovery sample, controlling for age and sex. Two metabolites entered the multiple regression model: DHA (standardized [std]  $\beta = -0.4838, p < 0.001$ ) and

S1P (std  $\beta = 0.5100, p < 0.001$ ). These 2 metabolites were then entered in the model built on the independent validation sample: DHA (std  $\beta = -0.445, p < 0.001$ ) and S1P (std  $\beta = 1.124, p < 0.001$ ; Table 5). We calculated the sensitivity and specificity based on estimates of the final model built on the discovery sample for the validation sample, and the model fit very well (AUC = 0.898, sensitivity = 90%, specificity = 74%; Appendix 1, Table S4 and Fig, S4).

**Table 2: Metabolites identified as differentially expressed in the positive and negative modes in the discovery cohort**

RT, min	Measured mass, Da*	Calculated mass, Da†	Mass error, ppm‡	Elemental composition	Postulated identity	VIP	Fold change§
<b>Positive</b>							
0.79	204.1255	204.1236	-9.308	C <sub>9</sub> H <sub>17</sub> NO <sub>4</sub>	L-acetylcarnitine	2.12	0.56
4.95	316.2504	316.2488	-5.0593	C <sub>17</sub> H <sub>33</sub> NO <sub>4</sub>	Decanoylcarnitine	1.64	0.59
5.4	318.3025	318.3008	-5.3408	C <sub>18</sub> H <sub>39</sub> NO <sub>3</sub>	Phytosphingosine	7.92	8.44
6.81	337.2764	337.2743	-6.2263	C <sub>21</sub> H <sub>36</sub> O <sub>3</sub>	Pregnanetriol	1.70	1.37
6.37	518.3271	518.3247	-4.6303	C <sub>26</sub> H <sub>48</sub> NO <sub>7</sub> P	LysoPC(18:3(6Z,9Z,12Z))	1.83	1.33
7.19	546.3535	546.3560	4.5758	C <sub>28</sub> H <sub>52</sub> NO <sub>7</sub> P	LysoPC(20:3(5Z,8Z,11Z))	1.53	1.13
<b>Negative</b>							
0.82	167.0165	167.0205	23.9497	C <sub>5</sub> H <sub>4</sub> N <sub>4</sub> O <sub>3</sub>	Uric acid	2.63	0.74
4.66	295.221	295.2273	21.3399	C <sub>18</sub> H <sub>32</sub> O <sub>3</sub>	9,10-epoxyoctadecenoic acid	1.85	1.36
8.56	303.2283	303.2324	13.5212	C <sub>20</sub> H <sub>32</sub> O <sub>2</sub>	Arachidonic acid	3.09	0.66
7.97	327.2312	327.2324	3.6671	C <sub>22</sub> H <sub>32</sub> O <sub>2</sub>	Docosahexaenoic acid	3.79	0.37
9.44	329.2444	329.2481	11.2379	C <sub>18</sub> H <sub>32</sub> O <sub>3</sub>	Docosapentaenoic acid	1.95	0.41
10.06	331.2567	331.2637	21.1316	C <sub>22</sub> H <sub>36</sub> O <sub>2</sub>	Adrenic acid	1.72	0.38
4.14	378.2365	378.2409	11.6329	C <sub>18</sub> H <sub>38</sub> NO <sub>5</sub> P	Sphingosine 1-phosphate	3.99	2.04
6.13	433.2403	433.2355	-11.0793	C <sub>21</sub> H <sub>39</sub> O <sub>7</sub> P	LPA(18:2(9Z,12Z)/0:0)	2.35	2.36
4.90	452.2784	452.2777	-1.5477	C <sub>21</sub> H <sub>44</sub> NO <sub>7</sub> P	LysoPE(0:0/16:0)	2.47	1.47
6.36	480.3128	480.3090	-7.9115	C <sub>23</sub> H <sub>48</sub> NO <sub>7</sub> P	LysoPE(18:0/0:0)	2.49	1.64
4.54	504.3077	504.3090	2.5778	C <sub>25</sub> H <sub>48</sub> NO <sub>7</sub> P	LysoPE(0:0/20:2(11Z,14Z))	3.19	1.14

RT = retention time; VIP = variable importance in the projection.  
 \*Molecular weight of metabolites from UPLC-MS/MS.  
 †Calculated using the molecular weight calculator in Mass Lynx version 4.1.  
 ‡(Calculated mass - measured mass) ÷ measured mass × 1 000 000.  
 §Ratio of mean relative amount between the autism and control groups.

**Table 3: Receiver operator characteristic curve analysis of 11 metabolites in positive and negative mode in the validation study**

Biomarkers	AUC	Best cutoff (normalized peak intensity)	Sensitivity, %	Specificity, %	Maximum of Youden index*
<b>Positive mode</b>					
Decanoylcarnitine	0.596	13.11	72	52	0.24
Pregnanetriol	0.607	24.59	36	89	0.25
<b>Negative mode</b>					
Uric acid	0.639	120.04	69	58	0.27
9,10-Epoxyoctadecenoic acid	0.682	8.91	60	86	0.46
Docosahexaenoic acid	0.741	21.36	67	76	0.43
Docosapentaenoic acid	0.719	12.87	59	78	0.37
Adrenic acid	0.682	22.37	50	84	0.34
Sphingosine 1-phosphate	0.871	38.89	79	81	0.60
LPA(18:2(9Z,12Z)/0:0)	0.886	3.94	77	88	0.65
LysoPE(0:0/16:0)	0.721	41.66	61	81	0.42
LysoPE(18:0/0:0)	0.728	29.47	65	75	0.40

AUC = area under the curve.  
 \*Sensitivity + specificity - 1.

We also built a multiple regression model in the male-only sample, using metabolites that were predominant in both the discovery and validation samples. In the discovery sample, age, sex, DHA (std  $\beta = -0.3862$ ,  $p = 0.008$ ) and S1P (std  $\beta = 0.7098$ ,  $p < 0.001$ ) were entered in the regression model and explained 37.40% of the variation. In the validation sample analysis, age, DHA (std  $\beta = -0.3473$ ,  $p = 0.06$ ) and S1P (std  $\beta = 1.2561$ ,  $p < 0.001$ ) were entered in the regression model and explained 55.07% of variation (Table 5). Similarly, a logistic regression model built on the discovery sample fit well in the validation sample (AUC = 0.915, specificity = 82.6%, sensitivity = 86.4%; Appendix 1, Table S4 and Fig. S4).

## Discussion

We performed a metabolomic analysis of serum in a cohort of patients with autism and controls and identified 17 metab-

olites that had best classification of the disease status. Eleven of these metabolites were validated in an independent cohort of cases and controls in both overall and male-only samples in a univariate analysis. Multiple logistic regression analysis of these 11 metabolites identified 2 metabolites, DHA and S1P, that were significant predictors of autism. These findings were based on 2 independent case-control cohorts using multiple logistic regression analysis of 11 replicated metabolites, in which we controlled for potential heterogeneity.

We observed that DHA, DPA, UA and 9,10-EOA were associated with autism. The 9,10-EOA is a peroxidation product of linoleic acid, which consumes excessive unsaturated fatty acids, such as DHA, and leads to a lower level of DHA. Previous studies have revealed significant alterations in fatty acid profiles of patients with autism compared with age-matched healthy children.<sup>25,55,56</sup> Children with autism tend to have a lower level of polyunsaturated fatty acid, which contributes to

**Table 4: Spearman correlation between of 11 metabolites relative contents with the total score of ABC, CARS, in validation stage**

Metabolites	ABC		CARS	
	<i>r</i>	<i>p</i> value	<i>r</i>	<i>p</i> value
Positive				
Decanoylcarnitine	-0.1359	0.34	-0.2564	0.07
Pregnanetriol	0.2697	0.06	0.3240	0.020
Negative				
Uric acid	0.0453	0.75	-0.2228	0.12
9,10-Epoxyoctadecenoic acid	0.2066	0.15	-0.0471	0.74
Docosahexaenoic acid	-0.2611	0.06	-0.1280	0.37
Docosapentaenoic acid	-0.3175	0.023	-0.1637	0.25
Adrenic acid	-0.1892	0.18	-0.1320	0.36
Sphingosine 1-phosphate	0.2746	0.05	-0.0657	0.65
LPA(18:2(9Z,12Z)/0:0)	-0.0632	0.66	-0.0042	0.98
LysoPE(0:0/16:0)	0.0220	0.88	0.1833	0.20
LysoPE(18:0/0:0)	-0.1230	0.39	0.1169	0.41

ABC = Autism Behaviour Checklist; CARS = Childhood Autism Rating Scales.

**Table 5: Multiple logistic regression models for the discovery and validation cohorts**

Factor	Discovery cohort*					Validation cohort†				
	Estimate	SE	Wald $\chi^2$	<i>p</i> value	Std $\beta$	Estimate	SE	Wald $\chi^2$	<i>p</i> value	Std $\beta$
Overall sample										
Intercept	0.2082	1.2091	0.0296	0.86		-0.5312	1.8438	0.0830	0.77	
Sex	0.1669	0.4643	0.1293	0.72	0.0322	1.2835	0.5900	4.7329	0.030	0.2633
Age	-0.3435	0.2135	2.5894	0.11	-0.1538	-0.7798	0.3039	6.5855	0.010	-0.3270
Docosahexaenoic acid	-0.0514	0.0141	13.2953	< 0.001	-0.4838	-0.0719	0.0215	11.1970	< 0.001	-0.4450
Sphingosine 1-phosphate	0.0500	0.0121	17.0152	< 0.001	0.5100	0.1171	0.0184	40.6966	< 0.001	1.1242
Male-only										
Intercept	-0.3142	1.3564	0.0537	0.82		-0.2483	1.9556	0.0161	0.90	
Age	-0.4685	0.2489	3.5416	0.06	-0.2030	-0.7711	0.3525	4.7870	0.029	-0.3042
Docosahexaenoic acid	-0.0398	0.0150	6.9810	0.008	-0.3862	-0.0575	0.0238	5.8337	0.016	-0.3473
Sphingosine 1-phosphate	0.0675	0.0146	21.3914	< 0.001	0.7098	0.1348	0.0223	36.6394	< 0.001	1.2561

SE = standard error; Std = standardized.

\*Overall sample:  $R^2 = 0.3081$ ; max-rescaled  $R^2 = 0.4108$ . Male-only sample:  $R^2 = 0.374$ ; max-rescaled  $R^2 = 0.4632$ .

†Overall sample:  $R^2 = 0.4572$ ; max-rescaled  $R^2 = 0.6096$ . Male-only sample:  $R^2 = 0.4741$ ; max-rescaled  $R^2 = 0.6324$ .

maintenance of the structural and functional integrity of the central nervous system.<sup>57–67</sup> A double-blind randomized placebo-controlled study found that DHA or eicosapentaenoic acid added to ARA supplementation could reduce behavioural problems in children with autism.<sup>68–70</sup> This is consistent with our findings that DHA was significantly associated with CARS score, which measures the severity of autistic behaviours (Table 4). These findings may further support that abnormalities of fatty acid are associated with autism.<sup>71</sup>

We also found that S1P was associated with autism. Sphingomyelin metabolism has been associated with abnormality in cerebral white matter,<sup>72,73</sup> and recent studies have provided support for the hypothesis of an abnormal developmental trajectory of white matter in patients with autism.<sup>74,75</sup> In addition, S1P may play some biological roles in immune function.<sup>76</sup> We suspected that alterations of sphingolipid metabolism in children with autism may be associated with abnormality in cerebral white matter. In addition, our study indicates that children with autism tend to have a lower level of UA in serum. Although the mechanism underlying lower serum UA is not clear, it has been recommended for use in autism screening; a previous study reported that less than 5% (1 of 32) of patients with autism have been found to have lower serum UA.<sup>77</sup>

In the discovery cohort, our male-only analysis was largely consistent with that of the overall sample, except that 1 metabolite — LysoPC (20:3(5Z, 8Z, 11Z)) — was less significant. This was expected because both the discovery and the validation samples were predominantly male. However, the slight differences in validation analyses between the overall and male-only samples were noted. In the male-only validation analysis phytosphingosine was not significant, likely owing to the reduction in sample size. It should be noted that while L-acetylcarnitine was not significant in the overall sample, it became significant in the male-only sample, even when the sample size was reduced, suggesting that the difference in L-acetylcarnitine between patients with autism and controls may be sex-sensitive, possibly owing to sex bias. Lower levels of L-acetylcarnitine and decanoylcarnitine have been found in children with autism and may reflect mitochondrial dysfunction.<sup>78, 79, 80</sup> A double-blind, parallel, multicentre comparison of L-acetylcarnitine with placebo in patients with attention-deficit/hyperactivity disorder and in boys with fragile X syndrome indicated that L-actylcarnitine significantly improved social behaviour.<sup>81</sup>

The alteration in these serum biomarkers could partly contribute to the variety of clinical manifestations of autism. The significant changes in serum pregnanetriol between children with autism and controls indicated perturbations in estrogen metabolism in children with autism. Moreover, the estrogen could have significant effects on the metabolism of polyunsaturated fatty acid.<sup>82,83</sup> We speculated that altered metabolism of hormones may be one of the causes of abnormal fatty acid metabolism. We also found abnormal lysophosphatide metabolism, such as the level changes in LPA (18:2(9Z,12Z)/0:0), LysoPE(0:0/16:0), LysoPE(18:0/0:0), in children with autism. We suspected that alterations of lysophosphatide metabolism in children with autism may be associated with abnormality in polyunsaturated fatty acids and sphingomyelin metabolism.

In our study, significant alterations in the fatty acid profiles (DHA, DPA and S1P) were found in the 2 cohorts. We present some evidence related to fatty acid  $\beta$ -oxidation and mitochondrial dysfunction (9,10-EOA and decanoylcarnitine). A multiple logistic regression model built on DHA and S1P in the discovery sample was confirmed in the validation sample with high specificity and sensitivity. If validated in additional larger samples, DHA and S1P could be used for clinical diagnosis of autism or clinical evaluation of treatment in patients with autism.

### Limitations

We should point out some limitations of this study. First, as autism was diagnosed predominantly in boys, we were unable to perform sex-specific analysis in girls owing to insufficient samples of female patients. Second, we did not examine the stability of the 11 significant metabolites over time after the initial test in the discovery sample; instead, we validated the discovery results in another independent sample. While this study was designed with 2 independent sample cohorts, the sample size was still relatively small. Our findings need to be further validated with a larger sample or with a longitudinal study that could control for intraindividual variations. Further assessment of the specificity for disease prediction should be carried out with other neurodevelopmental disorders, such as ADHD or some childhood metabolic diseases.

### Conclusion

We identified 11 metabolites consistently in 2 independent cohorts of patients with autism and controls that may have some clinical applications for autism. Two metabolites (DHA and S1P) were consistently replicated in a multiple logistic regression model. Our study may also provide a novel insight into the genetic etiology of autism.

**Acknowledgments:** This work was supported by the Natural Science Foundation of China (No: 81273094).

**Affiliations:** From the Department of Child and Adolescent Health, School of Public Health, Harbin Medical University, Harbin, Heilongjiang, China (Wang, Liang, Gao, Sun, Wang, Xia, Wu); the Center for Endemic Disease Control, China Center for Disease Control and Prevention and Harbin Medical University, Harbin, Heilongjiang, China (Wang); the Department of Nutrition and Food Hygiene, School of Public Health, Harbin Medical University, Harbin, Heilongjiang, China (Wang, Sun); the Department of Nursing, Daqing Campus, Harbin Medical University, Daqing, Heilongjiang, China (Gao); the Advanced Analytic Division, SAS Institute, Inc, Cary, North Carolina, USA (Wu); Systems and Translational Sciences, Research Triangle Institute, Research Triangle Park, North Carolina, USA (Sumner); and Lieber Institute for Brain Development, Johns Hopkins University Medical Campus, Baltimore, Maryland, USA (Zhang).

**Competing interests:** None declared.

**Contributors:** F. Zhang, Changhao Sun and L. Wu designed the study. H. Wang, S. Liang, M. Wang, J. Gao, Caihong Sun, J. Wang, W. Xia and L. Wu acquired the data, which H. Wang, M. Wang, S. Wu, S. Sumner, F. Zhang, Changhao Sun and L. Wu analyzed. All authors wrote and reviewed the article and approved it for publication.

## References

- Bailey A, Phillips W, Rutter M. Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *J Child Psychol Psychiatry* 1996;37:89-126.
- Developmental Disabilities Monitoring Network Surveillance Year 2010 Principal Investigators; Centers for Disease Control and Prevention (CDC). Prevalence of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *MMWR Surveill Summ*. 2014;63:1-21.
- Hallmayer J, Glasson EJ, Bower C, et al. On the twin risk in autism. *Am J Hum Genet* 2002;71:941-6.
- Wang K, Zhang H, Ma D, et al. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* 2009;459:528-33.
- Durand CM, Betancur C, Boeckers TM, et al. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet* 2007;39:25-7.
- Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. *Science* 2007;316:445-9.
- Weiss LA, Shen Y, Korn JM, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008;358:667-75.
- Glessner JT, Wang K, Cai G, et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 2009;459:569-73.
- Weiss LA, Arking DE, Daly MJ, et al. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 2009;461:802-8.
- Sanders SJ, Ercan-Sencicek AG, Hus V, et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* 2011;70:863-85.
- Sanders SJ, Murtha MT, Gupta AR, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 2012;485:237-41.
- O'Roak BJ, Vives L, Girirajan S, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 2012;485:246-50.
- Neale BM, Kou Y, Liu L, et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 2012;485:242-5.
- State MW, Sestan N. Neuroscience. The emerging biology of autism spectrum disorders. *Science* 2012;337:1301-3.
- Anney R, Klei L, Pinto D, et al. A genome-wide scan for common alleles affecting risk for autism. *Hum Mol Genet* 2010;19:4072-82.
- Xia K, Guo H, Hu Z, et al. Common genetic variants on 1p13.2 associate with risk of autism. *Mol Psychiatry* 2014;19:1212-9.
- Li M, Wang B, Zhang M, et al. Symbiotic gut microbes modulate human metabolic phenotypes. *Proc Natl Acad Sci U S A* 2008;105:2117-22.
- Serkova NJ, Jackman M, Brown JL, et al. Metabolic profiling of livers and blood from obese Zucker rats. *J Hepatol* 2006;44:956-62.
- Gloyn AL, Faber JH, Malmodin D, et al. Metabolic profiling in Maturity-onset diabetes of the young (MODY) and young onset type 2 diabetes fails to detect robust urinary biomarkers. *PLoS ONE* 2012;7:e40962.
- Brindle JT, Antti H, Holmes E, et al. Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabolomics. *Nat Med* 2002;8:1439-44.
- Kaddurah-Daouk R, Krishnan KR. Metabolomics: a global biochemical approach to the study of central nervous system diseases. *Neuropsychopharmacology* 2009;34:173-86.
- Yang J, Chen T, Sun L, et al. Potential metabolite markers of schizophrenia. *Mol Psychiatry* 2013;18:67-78.
- He Y, Yu Z, Giegling I, et al. Schizophrenia shows a unique metabolomics signature in plasma. *Transl Psychiatry* 2012;2:e149.
- Yap IK, 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;9:2996-3004.
- El-Ansary AK, Bacha AG, Al-Ayahdi LY. Plasma fatty acids as diagnostic markers in autistic patients from Saudi Arabia. *Lipids Health Dis* 2011;10:62.
- Ming X, Stein TP, Barnes V, et al. Metabolic perturbation in autism spectrum disorders: a metabolomics study. *J Proteome Res* 2012;11:5856-62.
- Emond P, Mavel S, Aidoud N, et al. GC-MS-based urine metabolic profiling of autism spectrum disorders. *Anal Bioanal Chem* 2013;405:5291-300.
- Kuwabara H, Yamasue H, Koike S, et al. Altered metabolites in the plasma of autism spectrum disorder: a capillary electrophoresis time-of-flight mass spectroscopy study. *PLoS ONE* 2013;8:e73814.
- Wilson ID, Nicholson JK, Castro-Perez J, et al. High resolution "ultra performance" liquid chromatography coupled to oa-TOF mass spectrometry as a tool for differential metabolic pathway profiling in functional genomic studies. *J Proteome Res* 2005;4:591-8.
- O'Connor D, Mortishire-Smith R. High-throughput bioanalysis with simultaneous acquisition of metabolic route data using ultra performance liquid chromatography coupled with time-of-flight mass spectrometry. *Anal Bioanal Chem* 2006;385:114-21.
- Wang M, Yang X, Wang F, et al. Calcium-deficiency assessment and biomarker identification by an integrated urinary metabolomics analysis. *BMC Med* 2013;11:86.
- Fan L, Zhang W, Yin M, et al. Identification of metabolic biomarkers to diagnose epithelial ovarian cancer using a UPLC/QTOF/MS platform. *Acta Oncol (Madr)* 2012;51:473-9.
- Wang C, Feng R, Sun D, et al. Metabolic profiling of urine in young obese men using ultra performance liquid chromatography and Q-TOF mass spectrometry (UPLC/Q-TOF MS). *J Chromatogr B Analyt Technol Biomed Life Sci* 2011;879:2871-6.
- Volkmar FR, Cicchetti DV, Dykens E, et al. An evaluation of the Autism Behavior Checklist. *J Autism Dev Disord* 1988;18:81-97.
- Ozonoff S, Goodlin-Jones BL, Solomon M. Evidence-based assessment of autism spectrum disorders in children and adolescents. *J Clin Child Adolesc Psychol* 2005;34:523-40.
- Frankenburg WK, Dodds JB. The Denver developmental screening test. *J Pediatr* 1967;71:181-91.
- First LR, Palfrey JS. The infant or young child with developmental delay. *N Engl J Med* 1994;330:478-83.
- Bijlsma S, Bobeldijk I, Verheij ER, et al. Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. *Anal Chem* 2006;78:567-74.
- Sangster T, Major H, Plumb R, et al. A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabolomic analysis. *Analyst* 2006;131:1075-8.
- Gika HG, Theodoridis GA, Wingate JE, et al. Within-day reproducibility of an HPLC-MS-based method for metabolomic analysis: application to human urine. *J Proteome Res* 2007;6:3291-303.
- Guidance for Industry Bioanalytical Method Validation. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM) 2013:5.
- Helsel DR. Summing nondetects: incorporating low-level contaminants in risk assessment. *Integr Environ Assess Manag* 2010;6:361-6.
- Baccarelli A, Pfeiffer R, Consonni D, et al. Handling of dioxin measurement data in the presence of non-detectable values: overview of available methods and their application in the Seveso chloracne study. *Chemosphere* 2005;60:898-906.
- Jensen ON, Podtelejnikov AV, Mann M. Identification of the components of simple protein mixtures by high-accuracy peptide mass mapping and database searching. *Anal Chem* 1997;69:4741-50.
- Lim H, Eng J, Yates III Jr, et al. Identification of 2D-gel proteins: a comparison of MALDI/TOF peptide mass mapping to  $\mu$  LC-ESI tandem mass spectrometry. *J Am Soc Mass Spectrom* 2003;14:957-70.
- Pelander A, Ojanperä I, Laks S, et al. Toxicological screening with formula-based metabolite identification by liquid chromatography/time-of-flight mass spectrometry. *Anal Chem* 2003;75:5710-8.
- Mahadevan S, Shah SL, Marrie TJ, et al. Analysis of metabolomic data using support vector machines. *Anal Chem* 2008;80:7562-70.

48. Szymanska E, Saccenti E, Smilde AK, et al. Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics* 2012;8:3-16.
49. Westerhuis JA, van Velzen EJ, Hoefsloot HC, et al. Discriminant Q2 (DQ2) for improved discrimination in PLS-DA models. *Metabolomics* 2008;4:293-6.
50. Westerhuis JA, Hoefsloot HC, Smit S, et al. Assessment of PLS-DA cross validation. *Metabolomics* 2008;4:81-9.
51. Jones OA, Spurgeon DJ, Svendsen C, et al. A metabolomics based approach to assessing the toxicity of the polyaromatic hydrocarbon pyrene to the earthworm *Lumbricus rubellus*. *Chemosphere* 2008;71:601-9.
52. Perez-Enciso M, Tenenhaus M. Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach. *Hum Genet* 2003;112:581-92.
53. Zhang T, Wu X, Yin M, et al. Discrimination between malignant and benign ovarian tumors by plasma metabolomic profiling using ultra performance liquid chromatography/mass spectrometry. *Clin Chim Acta* 2012;413:861-8.
54. Li S, Liu H, Jin Y, et al. Metabolomics study of alcohol-induced liver injury and hepatocellular carcinoma xenografts in mice. *J Chromatogr B Analyt Technol Biomed Life Sci* 2011;879:2369-75.
55. Wiest MM, German JB, Harvey DJ, et al. Plasma fatty acid profiles in autism: a case-control study. *Prostaglandins Leukot Essent Fatty Acids* 2009;80:221-7.
56. Bell JG, Miller D, MacDonald DJ, et al. The fatty acid compositions of erythrocyte and plasma polar lipids in children with autism, developmental delay or typically developing controls and the effect of fish oil intake. *Br J Nutr* 2010;103:1160-7.
57. Stillwell W, Wassall SR. Docosahexaenoic acid: membrane properties of a unique fatty acid. *Chem Phys Lipids* 2003;126:1-27.
58. Jump DB. The biochemistry of n-3 polyunsaturated fatty acids. *J Biol Chem* 2002;277:8755-8.
59. Muskiet FA, Fokkema MR, Schaafsma A, et al. Is docosahexaenoic acid (DHA) essential? Lessons from DHA status regulation, our ancient diet, epidemiology and randomized controlled trials. *J Nutr* 2004;134:183-6.
60. Florent S, Malaplate-Armand C, Youssef I, et al. Docosahexaenoic acid prevents neuronal apoptosis induced by soluble amyloid-beta oligomers. *J Neurochem* 2006;96:385-95.
61. Nabekura J, Noguchi K, Witt MR, et al. Functional modulation of human recombinant gamma-aminobutyric acid type A receptor by docosahexaenoic acid. *J Biol Chem* 1998;273:11056-61.
62. Kim HY, Akbar M, Lau A, et al. Inhibition of neuronal apoptosis by docosahexaenoic acid (22:6n-3). Role of phosphatidylserine in antiapoptotic effect. *J Biol Chem* 2000;275:35215-23.
63. Bazan NG. Lipid signaling in neural plasticity, brain repair, and neuroprotection. *Mol Neurobiol* 2005;32:89-103.
64. Tamiji J, Crawford DA. The neurobiology of lipid metabolism in autism spectrum disorders. *Neurosignals* 2010;18:98-112.
65. Burdge GC, Jones AE, Wootton SA. Eicosapentaenoic and docosapentaenoic acids are the principal products of alpha-linolenic acid metabolism in young men\*. *Br J Nutr* 2002;88:355-63.
66. Haggarty P, Ashton J, Joynson M, et al. Effect of maternal polyunsaturated fatty acid concentration on transport by the human placenta. *Biol Neonate* 1999;75:350-9.
67. Anderson GJ, Connor WE, Corliss JD. Docosahexaenoic acid is the preferred dietary n-3 fatty acid for the development of the brain and retina. *Pediatr Res* 1990;27:89-97.
68. Kurlak LO, Stephenson TJ. Plausible explanations for effects of long chain polyunsaturated fatty acids (LCPUFA) on neonates. *Arch Dis Child Fetal Neonatal Ed* 1999;80:F148-54.
69. Amminger GP, Berger GE, Schafer MR, et al. Omega-3 fatty acids supplementation in children with autism: a double-blind randomized, placebo-controlled pilot study. *Biol Psychiatry* 2007;61:551-3.
70. Yui K, Koshiba M, Nakamura S, et al. Effects of large doses of arachidonic acid added to docosahexaenoic acid on social impairment in individuals with autism spectrum disorders: a double-blind, placebo-controlled, randomized trial. *J Clin Psychopharmacol* 2012;32:200-6.
71. Richardson AJ, Ross MA. Fatty acid metabolism in neurodevelopmental disorder: a new perspective on associations between attention-deficit/hyperactivity disorder, dyslexia, dyspraxia and the autistic spectrum. *Prostaglandins Leukot Essent Fatty Acids* 2000;63:1-9.
72. Heipertz R, Pilz H, Scholz W. The fatty acid composition of sphingomyelin from adult human cerebral white matter and changes in childhood, senium and unspecific brain damage. *J Neurol* 1977;216:57-65.
73. Sugita M, Dulaney JT, Moser HW. Structure and composition of sulfatides isolated from livers of patients with metachromatic leukodystrophy: galactosyl sulfatide and lactosyl sulfatide. *J Lipid Res* 1974;15:227-33.
74. Barnea-Goraly N, Kwon H, Menon V, et al. White matter structure in autism: preliminary evidence from diffusion tensor imaging. *Biol Psychiatry* 2004;55:323-6.
75. Mak-Fan KM, Morris D, Vidal J, et al. White matter and development in children with an autism spectrum disorder. *Autism* 2013;17:541-57.
76. Spiegel S, Milstien S. The outs and the ins of sphingosine-1-phosphate in immunity. *Nat Rev Immunol* 2011;11:403-15.
77. Schaefer GB, Lutz RE. Diagnostic yield in the clinical genetic evaluation of autism spectrum disorders. *Genet Med* 2006;8:549-56.
78. Filipek PA, Juranek J, Nguyen MT, et al. Relative carnitine deficiency in autism. *J Autism Dev Disord* 2004;34:615-23.
79. Celestino-Soper PB, Violante S, Crawford EL, et al. A common X-linked inborn error of carnitine biosynthesis may be a risk factor for nondysmorphic autism. *Proc Natl Acad Sci U S A* 2012;109:7974-81.
80. Chauhan A, Gu F, Essa M, et al. Brain region-specific deficit in mitochondrial electron transport chain complexes in children with autism. *J Neurochem* 2011;117:209-20.
81. Torrioli MG, Vernacotola S, Peruzzi L, et al. A double-blind, parallel, multicenter comparison of L-acetylcarnitine with placebo on the attention deficit hyperactivity disorder in fragile X syndrome boys. *Am J Med Genet A* 2008;146A:803-12.
82. St-Pierre MV, Kullak-Ublick GA, Hagenbuch B, et al. Transport of bile acids in hepatic and non-hepatic tissues. *J Exp Biol* 2001;204:1673-86.
83. Nemoto Y, Toda K, Ono M, et al. Altered expression of fatty acid-metabolizing enzymes in aromatase-deficient mice. *J Clin Invest* 2000;105:1819-25.