



Urinary proteome profiling for children with autism using data-independent acquisition proteomics

Wenshu Meng, Yuhang Huan, Youhe Gao

Gene Engineering Drug and Biotechnology Beijing Key Laboratory, College of Life Sciences, Beijing Normal University, Beijing, China

Contributions: (I) Conception and design: Y Gao; W Meng; (II) Administrative support: Y Gao; (III) Provision of study materials or patients: Y Gao; W Meng; (IV) Collection and assembly of data: W Meng; Y Huan; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Youhe Gao. Beijing Key Laboratory of Gene Engineering Drug and Biotechnology, College of Life Sciences, Beijing Normal University, No. 19 XinJieKouWai Street, HaiDian District, Beijing, China. Email: gaoyouhe@bnu.edu.cn.

Background: Autism is a complex neurodevelopmental disorder. Objective and reliable biomarkers are crucial for the clinical diagnosis of autism. Urine can accumulate early changes of the whole body and is a sensitive source for disease biomarkers.

Methods: The data-independent acquisition (DIA) strategy was used to identify differential proteins in the urinary proteome between autistic and non-autistic children aged 3–7 years. Receiver operating characteristic (ROC) curves were developed to evaluate the diagnostic performance of differential proteins.

Results: A total of 118 differential proteins were identified in the urine between autistic and non-autistic children, of which 18 proteins were reported to be related to autism. Randomized grouping statistical analysis indicated that 91.5% of the differential proteins were reliable. Functional analysis revealed that some differential proteins were associated with axonal guidance signaling, endocannabinoid developing neuron pathway, synaptic long-term depression, agrin interactions at neuromuscular junction, phosphatase and tensin homolog deleted on chromosome 10 (PTEN) signaling and synaptogenesis signaling pathway. The combination of cadherin-related family member 5 (CDHR5) and vacuolar protein sorting-associated protein 4B (VPS4B) showed the best discriminative performance between autistic and non-autistic children with an area under the curve (AUC) value of 0.987.

Conclusions: The urinary proteome could distinguish between autistic children and non-autistic children. This study will provide a promising approach for future biomarker research of neuropsychiatric disorders.

Keywords: Autism; urine; proteome; biomarker; diagnosis

Submitted Apr 28, 2021. Accepted for publication Jun 21, 2021.

doi: 10.21037/tp-21-193

View this article at: <https://dx.doi.org/10.21037/tp-21-193>

Introduction

Autism spectrum disorder (ASD) is a complex neurodevelopmental disorder characterized by difficulties in social interaction and limited repetitive behaviors, interests, or activities (1). The incidence of autism has continued to increase over the past two decades, with the number of patients with autism as high as 1% to 2.5% of the total

population and a male/female ratio of 4:1 (2). Autism usually occurs at an early stage and is a lifelong developmental disorder that places a heavy burden on families and public health.

The etiology and pathological mechanism of autism are uncertain, which brings challenges to its diagnosis and intervention. There is currently no effective treatment for ASD, but some studies have found that behavioral

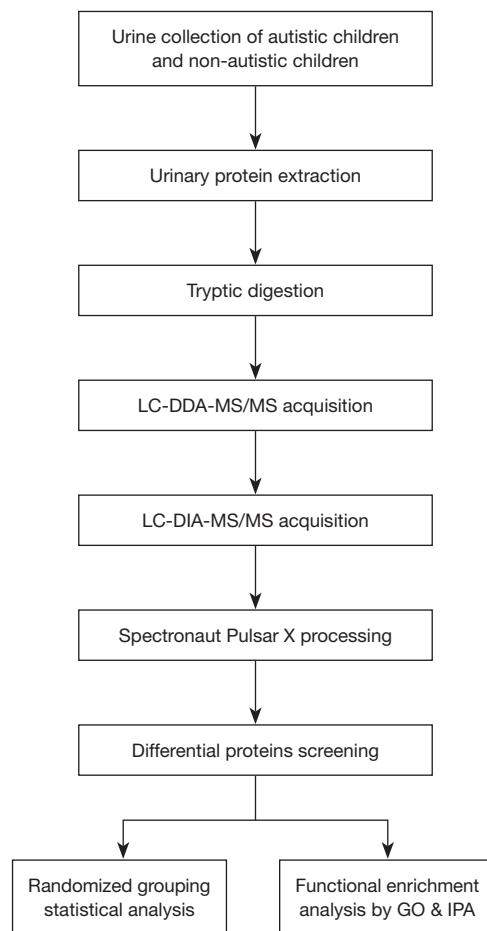


Figure 1 The workflow of urine proteome analysis in children with autism. DDA, data-dependent acquisition; DIA, data-independent acquisition; GO, gene ontology; IPA, ingenuity pathway analysis; LC-MS/MS, liquid chromatography couple with tandem mass spectrometry.

interventions for autistic children can effectively alleviate their symptoms at the early stage (3). Therefore, the early diagnosis is crucial for autism. The clinical diagnosis of autism mainly relies on behavioral and cognitive assessment according to the criteria in the diagnostic and statistical manual of mental disorders, which is certain subjective. Hence, the objective and reliable biomarkers are needed for the diagnosis of autism.

Previous proteomic studies on biomarkers and pathogenetic mechanisms of ASD have focused on blood, saliva and brain tissues (4–8). However, only a few studies have used urine. Urine is a sensitive source for diseases biomarkers. Without the control of homeostatic mechanisms, urine can accumulate early changes of the

whole body (9). In addition, urine collection is simple and non-invasive. There are several clinical studies showed that urine could reflect pathological changes of various diseases involving brain and nervous system, such as Alzheimer's disease (10), familial Parkinson's disease (11), pediatric medulloblastoma (12), and gliomas (13). However, for neuropsychiatric disorders with abnormal social behaviors such as ASD, it is unknown whether urine can show differences.

In this study, the data-independent acquisition (DIA) strategy was used to identify differential proteins in the urinary proteome between autistic and non-autistic children aged 3–7 years. This study aims to investigate whether the urinary proteome can distinguish between autistic children and non-autistic children. The workflow of this study is presented in *Figure 1*.

We present the following article in accordance with the MDAR reporting checklist (available at <https://dx.doi.org/10.21037/tp-21-193>).

Methods

Urine sample collection

In this study, urine samples from 18 autistic children aged 3–7 years from the Fengtai District Sunshine Angel Special Training Center in Beijing and 6 non-autistic children aged 3–6 years from Beijing Normal University were collected (*Table S1*). All ASD patients were diagnosed by child neuropsychiatrists according to criteria defined in the Diagnostic and Statistical Manual of Mental Disorders-Fifth Edition (DSM-V). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) for research on human participants, and the study protocols were approved by the Institutional Review Board at Beijing Normal University (ICBIR_A_0098_006). Written informed consent was obtained from the parents of all participants.

Urinary protein extraction and tryptic digestion

Urine samples were centrifuged at 12,000 ×g for 40 min at 4 °C to remove impurities and large cell debris. The supernatants were precipitated with three volumes of ethanol at -20 °C overnight and then centrifuged at 12,000 ×g for 30 min at 4 °C. The precipitate was resuspended in lysis buffer [8 mol/L urea, 2 mol/L thiourea, 50 mmol/L Tris, and 25 mmol/L dithiothreitol (DTT)]. The Bradford assay was

used to measure the protein concentration of each sample.

The urinary proteins were digested using the filter-aided sample preparation (FASP) method (14). A total of 100 µg protein of each sample was loaded onto a 10 kDa filter device (Pall, Port Washington, NY, USA) and washed twice with UA (8 mol/L urea, 0.1 mol/L Tris-HCl, pH 8.5) and 25 mmol/L NH₄HCO₃. The samples were reduced with 20 mmol/L DTT (Sigma, St. Louis, USA) at 37 °C for 1 h and then alkylated with 50 mmol/L iodoacetamide (IAA, Sigma, St. Louis, USA) in the dark for 40 min. After washing once with UA and twice with 25 mmol/L NH₄HCO₃, the proteins were digested with trypsin (enzyme-to-protein ratio of 1:50) at 37 °C overnight. The peptide mixtures were desalting using Oasis HLB cartridges (Waters, Milford, MA, USA) and then dried by vacuum evaporation.

High-pH reversed-phased peptide fractionation

The peptide samples were dissolved in 0.1% formic acid and diluted to 0.5 µg/µL. For the generation of spectral library, 96 µg of pooled peptides from 4 µg of each sample was fractionated using a high-pH reversed-phased peptide fractionation kit (catalog number: 84868, Thermo, USA). According to the manufacturer's instructions, 10 fractionated samples were obtained and were dried by vacuum evaporation. Then, 10 fractionated samples were dissolved in 20 µL of 0.1% formic acid. One microgram of each fraction was loaded for liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) analysis in data-dependent acquisition (DDA) mode.

LC-MS/MS analysis

An EASY-nLC 1200 chromatography system (Thermo Fisher Scientific, Waltham, MA, USA) and an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) were used for mass spectrometry acquisition and analysis. The iRT reagent (Biognosys, Switzerland) was spiked at a concentration of 1:10 v/v into all urinary samples for calibration of the retention time of the extracted peptide peaks. All peptide samples were loaded on a trap column (75 µm × 2 cm, 3 µm, C18, 100 Å) and a reverse-phase analysis column (75 µm × 25 cm, 2 µm, C18, 100 Å). The eluted gradient was 4–35% buffer B (0.1% formic acid in 80% acetonitrile) at a flow rate of 300 nL/min for 90 min.

In DDA mode, the parameters were set as follows: the full scan from 350 to 1,500 m/z with resolution at 120,000

and MS/MS scan with resolution at 30,000 in Orbitrap; the 30% higher-energy collisional dissociation (HCD) energy; the maximum injection time of 45 ms.

In DIA mode, 1 µg of each sample was analyzed with twice replicates. The DIA method with 36 variable windows was set for DIA acquisition (Table S2). The parameters were set as follows: the full scan from 350 to 1,500 m/z with resolution at 60,000; the DIA scan from 200 to 2,000 m/z with resolution of 30,000; the 32% HCD energy; and the maximum injection time of 100 ms. A quality control (QC) sample of the mixture from each sample was analyzed in DIA acquisition after every four samples.

Spectral library generation and data analysis

The DDA data of 10 fractions were processed using Proteome Discoverer software (version 2.1, Thermo Scientific) and searched against the Swiss-Prot Human database (released in 2018, including 20,346 sequences) appended with the iRT peptide sequence. The search parameters were set as follows: two missed trypsin cleavage sites were allowed; the parent ion mass tolerances were set to 10 ppm; the fragment ion mass tolerances were set to 0.02 Da; the carbamidomethyl of cysteine was set as a fixed modification; and the oxidation of methionine was set as a variable modification. The false discovery rate (FDR) of proteins was less than 1%. A total of 2,184 protein groups, 11,518 peptide groups and 59,341 peptide spectrum matches were identified. The search result was used to set the variable windows for DIA mode. For the generation of spectral library, the DDA raw files were imported to Spectronaut Pulsar X software (Biognosys, Switzerland). All DIA raw files were processed using Spectronaut Plusar X software with default setting. All results were filtered by a Q value cutoff of 0.01. The protein identification was based on two unique peptides.

Statistical analysis

A comparison of proteins between autistic and non-autistic group was conducted using independent samples *t*-test. Group differences resulting in *P*<0.05 were considered statistically significant. Differential proteins were screened with the following criteria: fold change in increasing group ≥1.5 and in decreasing group ≤0.67, *P*<0.01. Receiver operating characteristic (ROC) analysis were performed for individual proteins and protein combinations using Metaboanalyst software (<https://www.metaboanalyst.ca/>)

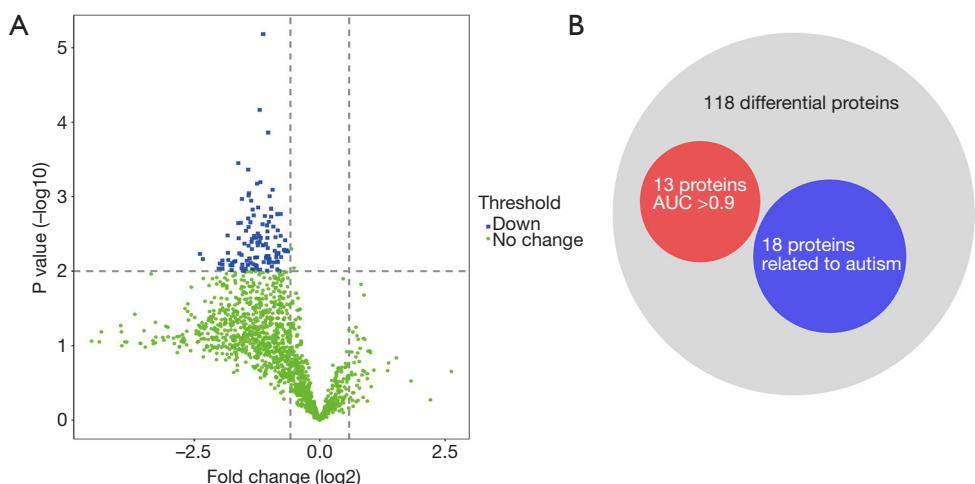


Figure 2 Differential proteins analysis by comparing autistic urine samples with non-autistic urine samples. (A) Volcano plots of differential proteins; (B) Venn diagram of 118 differential proteins, 18 proteins related to autism, and the top 13 differential proteins with AUC values (AUC >0.9). AUC, area under the curve.

www.metaboanalyst.ca).

Functional enrichment analysis

Functional annotation of differential proteins was performed using DAVID 6.8 (<https://david.ncifcrf.gov>) (15) and ingenuity pathway analysis (IPA) software (Ingenuity Systems, Mountain View, CA, USA), including biological process, cellular component, molecular function and pathways. The threshold of significance was set at a $P<0.05$.

Results

Identification of differential proteins in ASD urinary proteome

To investigate differences between autistic and non-autistic children, 24 urinary samples from 18 autistic children and 6 non-autistic children were analyzed after proteolysis by LC-DIA-MS/MS. A total of 1,631 protein groups were identified in this study. A QC sample of the mixture from each sample was analyzed after every four samples. The 95% of the quantile of coefficient of variation (CV) value of the QC was 0.51, and proteins with CV >0.51 were considered outliers. A total of 1,511 proteins with the CV below 0.51 were for subsequent analysis, and the identification and quantification details are listed in online table (available at <https://cdn.amegroups.cn/static/public/tp-21-193-1.pdf>). Among them, 118 differential proteins

were identified between the autistic and non-autistic groups (fold change ≥ 1.5 or ≤ 0.67 , $P<0.01$), the volcano plot of differential proteins is shown in *Figure 2A*. The details of the differential proteins are listed in *Table 1*.

Randomized grouping statistical analysis

Given that the number of proteomic features identified in the samples was higher than the number of samples, the differences between two groups might be randomly generated. A randomized grouping statistical analysis strategy was developed to confirm whether these differential proteins were caused by disease. Twenty-four samples from the autism ($n=18$) and control groups ($n=6$) were randomly divided into two groups and the same criteria were used to screen differential proteins. In a total of 134,596 (C_{24}^2) combinations, the average number of differential proteins was 10. These results showed that only 10 differential proteins could be generated randomly, further indicating that 91.5% of the differential proteins were reliable.

ROC curve analysis

To evaluate the diagnostic performance of differential proteins between autistic and non-autistic children, ROC curves were performed for individual proteins and protein combinations. Among 118 differential proteins, 13 proteins (CDHR5, VPS4B, NICA, LEG1, ARL3, MANBA, VATG1, CO5A2, CHM1B, CDC42, NRP1, F13B, INARI)

Table 1 The differential proteins identified in urine samples between autistic and healthy children

Accession	Description	Ratio of ASD/HC	P value	Reference
Q9NZV1	Cysteine-rich motor neuron 1 protein	0.6476	0.0055	–
P54652	Heat shock-related 70 kDa protein 2	0.6356	0.0052	–
O00461	Golgi integral membrane protein 4	0.6234	0.0054	–
P07204	Thrombomodulin	0.6227	0.0038	–
Q9Y6W3	Calpain-7	0.6125	0.0052	–
Q99816	Tumor susceptibility gene 101 protein	0.5869	0.0065	–
Q9H0E2	Toll-interacting protein	0.5856	0.0034	–
Q9UN37	Vacuolar protein sorting-associated protein 4A	0.5852	0.0017	–
P61204	ADP-ribosylation factor 3	0.5725	0.0046	–
P39060	Collagen alpha-1 (XVIII) chain	0.5723	0.0099	–
P19961	Alpha-amylase 2B	0.5676	0.0057	–
Q8WV92	MIT domain-containing protein 1	0.5622	0.0028	–
Q8NHP8	Putative phospholipase B-like 2	0.5602	0.0075	–
Q04756	Hepatocyte growth factor activator	0.5586	0.0017	–
Q8TAB3	Protocadherin-19	0.5583	0.0095	–
Q9HD42	Charged multivesicular body protein 1a	0.5578	0.0074	–
P27105	Stomatin	0.5542	0.0059	–
P08758	Annexin A5	0.5494	0.0065	–
O00560	Syntenin-1	0.5453	0.0035	–
Q9NZN3	EH domain-containing protein 3	0.5394	0.0053	–
P35247	Pulmonary surfactant-associated protein D	0.5305	0.0076	–
O43633	Charged multivesicular body protein 2a	0.5277	0.0025	–
P62070	Ras-related protein R-Ras2	0.5216	0.0008	–
Q96AP7	Endothelial cell-selective adhesion molecule	0.5173	0.0086	–
O95971	CD160 antigen	0.5132	0.0019	–
Q7LBR1	Charged multivesicular body protein 1b	0.5131	0.0014	–
P51148	Ras-related protein Rab-5C	0.5026	0.0010	–
Q16348	Solute carrier family 15 member 2	0.5023	0.0075	–
P54707	Potassium-transporting ATPase alpha chain 2	0.5007	0.0069	–
P14314	Glucosidase 2 subunit beta	0.4976	0.0022	–
P01033	Metalloproteinase inhibitor 1	0.4935	0.0077	–
P18510	Interleukin-1 receptor antagonist protein	0.4928	0.0063	(16)
Q99538	Legumain	0.4921	0.0068	–
O75348	V-type proton ATPase subunit G 1	0.4911	0.0001	–
Q14108	Lysosome membrane protein 2	0.4893	0.0067	–

Table 1 (continued)

Table 1 (continued)

Accession	Description	Ratio of ASD/HC	P value	Reference
P62879	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	0.4888	0.0049	–
Q92542	Nicastrin	0.4875	0.0036	(17)
P0DJD8	Pepsin A-3	0.4873	0.0029	–
P62834	Ras-related protein Rap-1A	0.4844	0.0099	–
Q9BY43	Charged multivesicular body protein 4a	0.4836	0.0024	–
Q8WUM4	Programmed cell death 6-interacting protein	0.4810	0.0063	–
Q5VW32	BRO1 domain-containing protein BROX	0.4764	0.0066	–
P46109	Crk-like protein	0.4763	0.0043	–
P07711	Procathepsin L	0.4691	0.0018	(18)
Q8NEU8	DCC-interacting protein 13-beta	0.4624	0.0075	–
Q9HBB8	Cadherin-related family member 5	0.4589	7E-06	–
Q96SM3	Probable carboxypeptidase X1	0.4575	0.0045	–
Q9Y3E7	Charged multivesicular body protein 3	0.4569	0.0045	–
Q8IX04	Ubiquitin-conjugating enzyme E2 variant 3	0.4545	0.0023	–
O60939	Sodium channel subunit beta-2	0.4432	0.0036	–
Q6P1N0	Coiled-coil and C2 domain-containing protein 1A	0.4425	0.0006	(19)
P15311	Ezrin	0.4384	0.0095	–
Q9Y6N7	Roundabout homolog 1	0.4366	0.0041	(20)
O75351	Vacuolar protein sorting-associated protein 4B	0.4364	0.0001	–
Q9H0X4	Protein FAM234A	0.4344	0.0034	–
P36543	V-type proton ATPase subunit E 1	0.4281	0.0014	–
O00161	Synaptosomal-associated protein 23	0.4259	0.0031	(21)
P21926	CD9 antigen	0.4255	0.0045	–
Q8IWA5	Choline transporter-like protein 2	0.4246	0.0041	–
Q9H3G5	Probable serine carboxypeptidase CPVL	0.4225	0.0007	–
Q9BYH1	Seizure 6-like protein	0.4207	0.0065	(22)
Q5JXA9	Signal-regulatory protein beta-2	0.4194	0.0093	–
Q9UBD6	Ammonium transporter Rh type C	0.4194	0.0085	–
P10912	Growth hormone receptor	0.4168	0.0036	–
P09936	Ubiquitin carboxyl-terminal hydrolase isozyme L1	0.4139	0.0083	(23)
P02749	Beta-2-glycoprotein 1	0.4112	0.0066	(24)
P09543	2',3'-cyclic-nucleotide 3'-phosphodiesterase	0.4051	0.0036	–
P36405	ADP-ribosylation factor-like protein 3	0.4047	0.0034	–
P08754	Guanine nucleotide-binding protein G(i) subunit alpha	0.4034	0.0018	–
Q9BRG1	Vacuolar protein-sorting-associated protein 25	0.4010	0.0080	–

Table 1 (continued)

Table 1 (continued)

Accession	Description	Ratio of ASD/HC	P value	Reference
Q71RC9	Small integral membrane protein 5	0.3969	0.0015	–
Q9UEF7	Klotho	0.3958	0.0094	–
P02649	Apolipoprotein E	0.3952	0.0094	(25)
P05413	Fatty acid-binding protein, heart	0.3935	0.0037	(26,27)
P63092	Guanine nucleotide-binding protein G(s) subunit alpha isoforms short	0.3930	0.0050	–
P01111	GTPase NRas	0.3903	0.0041	(28)
P03950	Angiogenin	0.3895	0.0011	–
P09488	Glutathione S-transferase Mu 1	0.3872	0.0077	(29)
O00462	Beta-mannosidase	0.3841	0.0022	–
Q16378	Proline-rich protein 4	0.3800	0.0068	–
P41181	Aquaporin-2	0.3767	0.0009	(30)
P60953	Cell division control protein 42 homolog	0.3758	0.0026	–
Q9H223	EH domain-containing protein 4	0.3752	0.0010	–
P10586	Receptor-type tyrosine-protein phosphatase F	0.3749	0.0046	–
P62873	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	0.3740	0.0020	–
P63000	Ras-related C3 botulinum toxin substrate 1	0.3732	0.0004	(31)
P36969	Phospholipid hydroperoxide glutathione peroxidase	0.3730	0.0082	–
Q8NBS9	Thioredoxin domain-containing protein 5	0.3713	0.0082	–
O75954	Tetraspanin-9	0.3650	0.0072	–
O00526	Uroplakin-2	0.3620	0.0029	–
Q9BZM4	UL16-binding protein 3	0.3593	0.0093	–
Q96MM7	Heparan-sulfate 6-O-sulfotransferase 2	0.3540	0.0043	–
O14786	Neuropilin-1	0.3501	0.0099	–
P61088	Ubiquitin-conjugating enzyme E2 N	0.3446	0.0081	–
P05160	Coagulation factor XIII B chain	0.3423	0.0011	–
P61225	Ras-related protein Rap-2b	0.341	0.0091	–
P80303	Nucleobindin-2	0.3409	0.006	–
Q96EY5	Multivesicular body subunit 12A	0.3362	0.0023	–
Q9H1C7	Cysteine-rich and transmembrane domain-containing protein 1	0.3296	0.0058	–
Q96CS7	Pleckstrin homology domain-containing family B member 2	0.3266	0.0036	–
Q13621	Solute carrier family 12 member 1	0.3259	0.0004	(32)
Q13103	Secreted phosphoprotein 24	0.3255	0.0023	–
Q53TN4	Cytochrome b reductase 1	0.3117	0.0074	–
Q8IWV2	Contactin-4	0.3076	0.0075	(33)

Table 1 (continued)

Table 1 (continued)

Accession	Description	Ratio of ASD/HC	P value	Reference
Q14254	Flotillin-2	0.3001	0.0078	–
P17181	Interferon alpha/beta receptor 1	0.2997	0.0084	–
P50897	Palmitoyl-protein thioesterase 1	0.2964	0.0097	–
O60613	Selenoprotein F	0.2927	0.0083	–
Q10471	Polypeptide N-acetylgalactosaminyltransferase 2	0.2832	0.0071	–
P63096	Guanine nucleotide-binding protein G(i) subunit alpha-1	0.2813	0.0057	–
P05997	Collagen alpha-2(V) chain	0.2801	0.0033	–
Q9ULZ9	Matrix metalloproteinase-17	0.2616	0.0072	–
P24821	Tenascin	0.2610	0.0093	(34)
O00159	Unconventional myosin-Ic	0.2604	0.0081	–
Q9GZM7	Tubulointerstitial nephritis antigen-like	0.2516	0.0079	–
P21796	Voltage-dependent anion-selective channel protein 1	0.2480	0.0092	–
P09382	Galectin-1	0.1988	0.0069	–
P02511	Alpha-crystallin B chain	0.1910	0.0059	–

ASD, autism spectrum disorder; HC, healthy control. Ratio of ASD/HC represents each protein abundance in the ASD group divide by protein abundance in the HC group.

showed the good discriminative performance between autistic and non-autistic children (AUC >0.9) (*Figure 2B*, *Table 2*). As shown in *Figure 3*, the combination of CDHR5 and VPS4B showed an AUC of 0.987, which was higher than that of the individual protein. Thus, these differential proteins and protein panels could be potential diagnostic biomarkers for autism.

Function analysis of the differential proteins

Functional annotation of 118 differential proteins was performed by DAVID. The differential proteins were classified into biological process, cellular component, and molecular function. In the biological process category, 62 items were significantly enriched (*Table S3*), of which representative biological processes are presented in *Figure 4A*. These differential proteins were involved in viral budding via host endosomal sorting complex required for transport (ESCRT) complex, multivesicular body assembly, autophagy, small GTPase mediated signal transduction, Ras protein signal transduction, axon guidance, chemical synaptic transmission, and negative regulation of neuron death. In the cellular component category, the majority of differential proteins came from extracellular exosomes (*Figure 4B*). In the

molecular function category, GTPase activity, GTP binding, signal transducer activity and protein homodimerization activity were overrepresented (*Figure 4C*).

To identify the major biological pathways of differential proteins, IPA software was performed for canonical pathways and network analysis. A total of 206 items were significantly enriched (online table available at <https://cdn.amegroups.cn/static/public/tp-21-193-2.pdf>), of which representative pathways are presented in *Figure 4D*. Axonal guidance signaling, endocannabinoid developing neuron pathway, STAT3 pathway, phosphatase and tensin homolog deleted on chromosome 10 (PTEN) signaling, synaptogenesis signaling pathway, synaptic long-term depression, and PI3K/AKT signaling were overrepresented. In addition, IPA network analysis revealed that a total of 25 differential proteins were involved in the top regulator effect network “cell-to-cell signaling and interaction, cellular movement, hematological system development and function” with score 47 (*Figure 5*).

Discussion

In this study, urine proteome in children with autism was analyzed by DIA proteomics, and 118 differential proteins

Table 2 The top 13 differential proteins with AUC values

Accession	Description	AUC
Q9HBB8	Cadherin-related family member 5	0.98148
O75351	Vacuolar protein sorting-associated protein 4B	0.96296
Q92542	Nicastrin	0.93519
P09382	Galectin-1	0.93519
P36405	ADP-ribosylation factor-like protein 3	0.92593
O00462	Beta-mannosidase	0.92593
O75348	V-type proton ATPase subunit G 1	0.91667
P05997	Collagen alpha-2(V) chain	0.91667
Q7LBR1	Charged multivesicular body protein 1b	0.90741
P60953	Cell division control protein 42 homolog	0.90741
O14786	Neuropilin-1	0.90741
P05160	Coagulation factor XIII B chain	0.90741
P17181	Interferon alpha/beta receptor 1	0.90741

AUC, area under the curve.

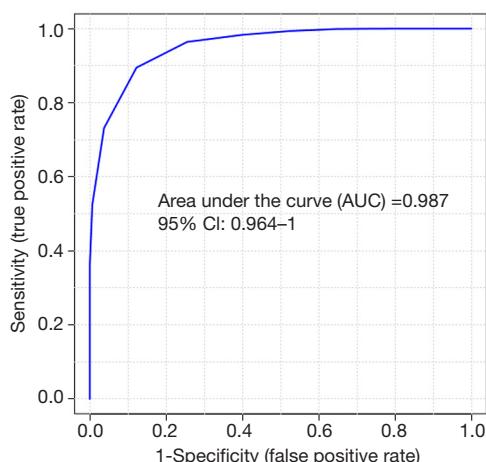


Figure 3 ROC curve analysis of the combination of CDHR5 and VPS4B. ROC, receiver operating characteristic; CDHR5, cadherin-related family member 5; VPS4B, vacuolar protein sorting-associated protein 4B; CI, confidence interval.

were identified between autistic and non-autistic children. Among them, 18 proteins have been reported to be related to autism (*Figure 2B*). For example, interleukin 1 receptor antagonist protein (IL1RA) is an anti-inflammatory cytokine that was downregulated in the serum of autistic patients (16). Nicastrin (NCSTN) plays an important role

in the regulation of short-term and long-term synaptic plasticity (17). Cathepsin L1 (CATL1) stimulates neuronal axon growth (18). *CC2D1A* (19) has been reported to be as candidate genes for autism. The abnormality of ROBO may cause autism by interfering with the serotonergic system or interfering with neurodevelopment (20). SNAP25 was reported to be involved in autism, seizures, and intellectual disability (21), and SNAP23 was downregulated in this study. *SEZ6L* (22) is a candidate gene for autism. Low levels of ubiquitin carboxy-terminal hydrolase isoenzyme L1 (UCHL1) is associated with ubiquitination interference in autism (23). Beta-2-glycoprotein 1 (APOH) was reported to be elevated in the plasma of patients with autism compared with that of control subjects (24). APOE hypermethylation is associated with ASD in the Chinese population (25). The abnormal expression of *FABP7* and *FABP5* genes in individuals with autism was found, and *FABP3* was downregulated in the urine of ASD patients, which plays a key role in cognition and emotional behavior (26,27). *NRAS* (28) is a candidate gene of ASD. *GSTM1* genotype may serve as a moderator of the effect of some prenatal factors on the risk of ASD (29). The expression of AQP4 in the brains of autistic patients was reported to be decreased (30). We found that AQP2 were downregulated in the urine of ASD patients. RAC1 stimulates the initiation and elongation of dendrites, Rac1/PAK/LIMK signaling

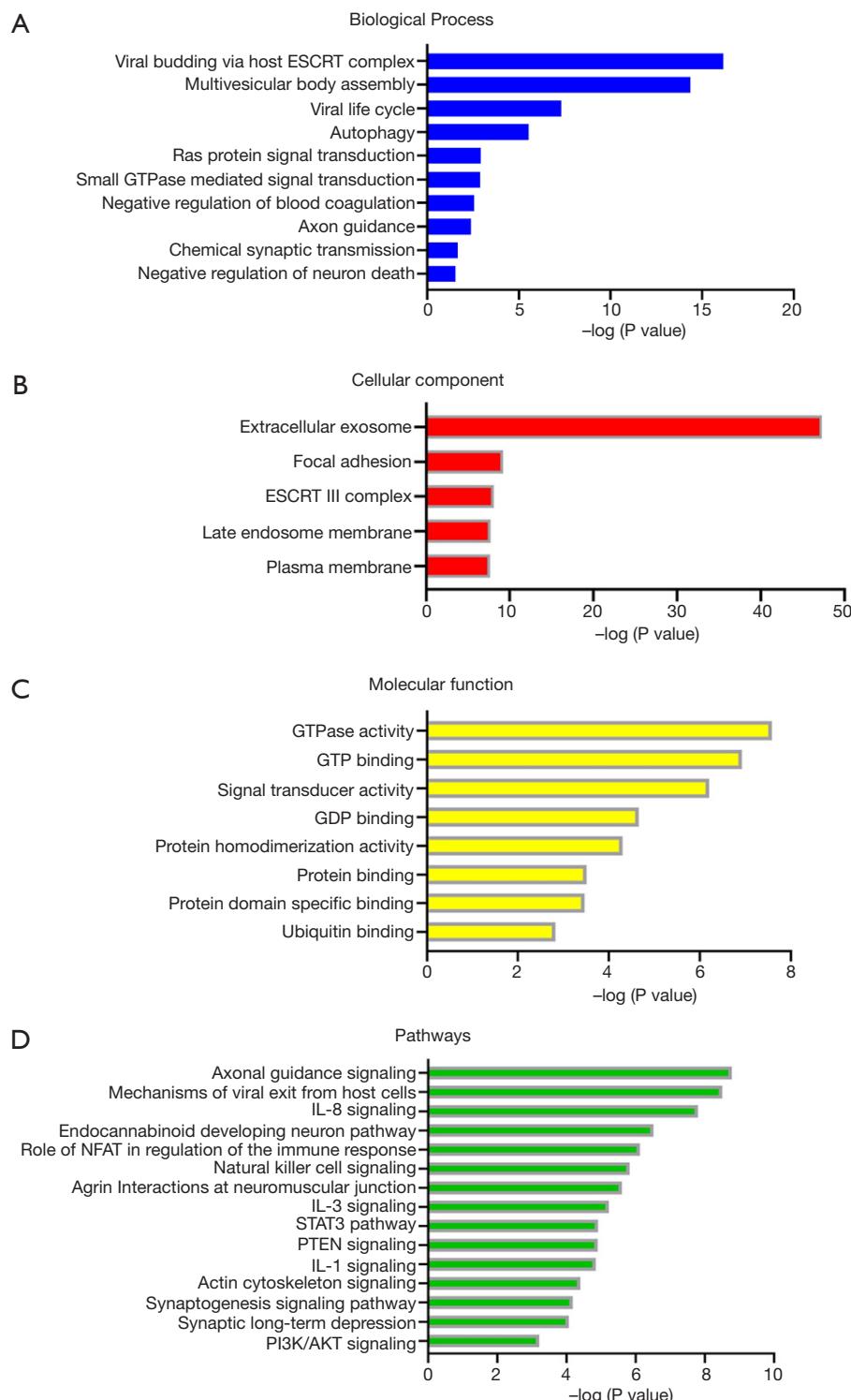


Figure 4 Functional analysis of differential proteins. (A) Biological process; (B) cellular component; (C) molecular function; (D) pathways. ESCRT, endosomal sorting complex required for transport; GTPase, guanosine triphosphate hydrolase; GDP, guanosine diphosphate; GTP, guanosine triphosphate; NFAT, nuclear factor of activated T cells; STAT3, signal transducer and activator of transcription 3; PTEN, phosphatase and tensin homolog deleted on chromosome 10; IL-8, interleukin 8; IL-3, interleukin 3; IL-1, interleukin 1; P13K/AKT, phosphatidylinositol-3 kinase/cellular homolog of the viral oncogene v-Akt.

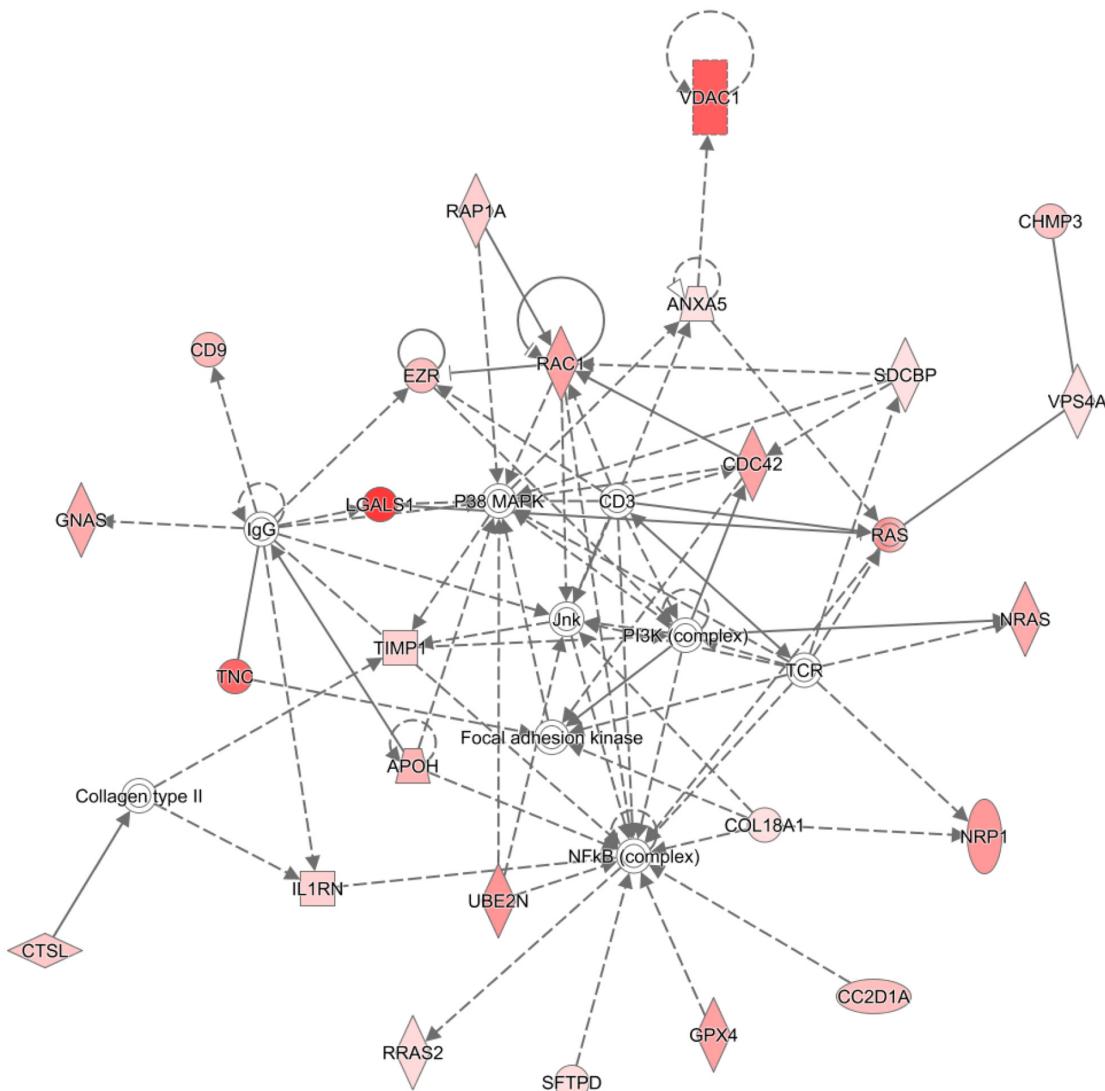


Figure 5 IPA revealed that the top regulator effect network. Red indicates down-regulated proteins in this study. IPA, ingenuity pathway analysis.

promotes actin filament assembly, and actin dysregulation is a pathophysiological mechanism of autism (31). Bumetanide administration can improve the symptoms of autism (32). We found that bumetanide-sensitive sodium-(potassium)-chloride cotransporter 2 (SLC12A1) was downregulated in urine. CNTN4 plays an important role in the formation, maintenance, and plasticity of neuronal networks and disruption of contactin 4 has been reported in ASD patients (33). The mutations in the tenascin C (*TNC*) gene could cause sensory impairment in ASD (34). Although some differential proteins have not been reported to be related to autism,

they also might serve as candidate urinary biomarkers for autism.

In addition, some important pathways were associated with autism. For example, changes in axonal microstructure are considered to be the basis of the cognitive performance of people with autism (35), several differential proteins were involved in axonal guidance signaling. Moreover, the endogenous cannabinoid system is involved in regulating many cellular functions and molecular pathways in autism, such as unbalanced glutamate and gamma-aminobutyric acid (GABA) and glutamate energy transmission, and

disorders of the endogenous cannabinoid system may play an important role in the pathophysiology of autism (36,37). Dysfunction of PTEN signaling may also be combined with changes in other autism-related genes or pathways to influence social behavior (38). Multiple susceptibility genes of autism encode synaptic-related proteins and affect the formation, elimination, transmission and plasticity of synapses (39), 9 proteins (APOE, CDC42, CRKL, NRAS, RAB5C, RAC1, RAP1A, RAP2B, RRAS2) were involved in synaptogenesis signaling pathway and 7 proteins (GNAI1, GNAI3, GNAS, NRAS, RAP1A, RAP2B, RRAS2) were involved in synaptic long-term depression. A large amount of evidence suggests that inflammation may be involved in the pathophysiological process of autism, manifested as a change in proinflammatory cytokine signals (40) and several inflammation-related signals were enriched in this study, such as IL-8 and IL-3 signaling. Thus, urinary proteins might reflect the pathophysiological process of autism and provide new targets for the intervention for autism.

Although autism is a heterogeneous neurological developmental disorder with multiple etiologies, subtypes and developmental trajectories, the urinary proteome between autistic group and non-autistic group showed clear differences, suggesting that autism might have a limited number of common biological pathways (41) or the ASD patients who contributed urine samples in this study might happen to be of similar subtypes.

This preliminary study has some limitations worth noting. First, the number of participants enrolled was limited. Secondly, the subtypes of children with autism in this study was not clear, and different subtypes may have different biomarkers, so whether our findings may be extended to other subtypes of autism is uncertain. Furthermore, whether these candidate urinary biomarkers can be applicable to earlier-age autistic children is unknown. Therefore, a large number of ASD patients with earlier ages and multiple subtypes from multicenter should be considered in future studies. Despite limitations of the study, our results demonstrate that ASD can be reflected in the urine, suggesting that urine proteome is a promising approach for diagnosis of ASD.

Conclusions

The urinary proteome could distinguish between autistic children and non-autistic children. This study will provide a promising approach for future biomarker research of neuropsychiatric disorders.

Acknowledgments

Funding: This work was supported by the National Key Research and Development Program of China (2018YFC0910202, 2016YFC1306300); the Fundamental Research Funds for the Central Universities (2020KJZX002); the Beijing Natural Science Foundation (7172076); the Beijing Cooperative Construction Project (110651103); the Beijing Normal University (11100704); and the Peking Union Medical College Hospital (2016-2.27). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at <https://dx.doi.org/10.21037/tp-21-193>

Data Sharing Statement: Available at <https://dx.doi.org/10.21037/tp-21-193>

Peer Review File: Available at <https://dx.doi.org/10.21037/tp-21-193>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/tp-21-193>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) for research on human participants, and the study protocols were approved by the Institutional Review Board at Beijing Normal University (ICBIR_A_0098_006). Written informed consent was obtained from the parents of all participants.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the

original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Lai MC, Lombardo MV, Baron-Cohen S. Autism. *Lancet* 2014;383:896-910.
- Lyall K, Croen L, Daniels J, et al. The Changing Epidemiology of Autism Spectrum Disorders. *Annu Rev Public Health* 2017;38:81-102.
- Fernell E, Eriksson MA, Gillberg C. Early diagnosis of autism and impact on prognosis: a narrative review. *Clin Epidemiol* 2013;5:33-43.
- Shen L, Zhang K, Feng C, et al. iTRAQ-Based Proteomic Analysis Reveals Protein Profile in Plasma from Children with Autism. *Proteomics Clin Appl* 2018;12:e1700085.
- Corbett BA, Kantor AB, Schulman H, et al. A proteomic study of serum from children with autism showing differential expression of apolipoproteins and complement proteins. *Mol Psychiatry* 2007;12:292-306.
- Ngounou Wetie AG, Wormwood KL, Russell S, et al. A Pilot Proteomic Analysis of Salivary Biomarkers in Autism Spectrum Disorder. *Autism Res* 2015;8:338-50.
- Amal H, Barak B, Bhat V, et al. Shank3 mutation in a mouse model of autism leads to changes in the S-nitroso-proteome and affects key proteins involved in vesicle release and synaptic function. *Mol Psychiatry* 2020;25:1835-48.
- Abraham J, Szoko N, Natowicz MR. Proteomic Investigations of Autism Spectrum Disorder: Past Findings, Current Challenges, and Future Prospects. *Adv Exp Med Biol* 2019;1118:235-52.
- Gao Y. Urine—an untapped goldmine for biomarker discovery? *Sci China Life Sci* 2013;56:1145-6.
- Watanabe Y, Hirao Y, Kasuga K, et al. Molecular Network Analysis of the Urinary Proteome of Alzheimer's Disease Patients. *Dement Geriatr Cogn Dis Extra* 2019;9:53-65.
- Virreira Winter S, Karayel O, Strauss MT, et al. Urinary proteome profiling for stratifying patients with familial Parkinson's disease. *EMBO Mol Med* 2021;13:e13257.
- Hao X, Guo Z, Sun H, et al. Urinary protein biomarkers for pediatric medulloblastoma. *J Proteomics* 2020;225:103832.
- Wu J, Zhang J, Wei J, et al. Urinary biomarker discovery in gliomas using mass spectrometry-based clinical proteomics. *Chin Neurosurg J* 2020;6:11.
- Wiśniewski JR, Zougman A, Nagaraj N, et al. Universal sample preparation method for proteome analysis. *Nat Methods* 2009;6:359-62.
- Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;4:44-57.
- Saghazadeh A, Ataeinia B, Keynejad K, et al. Anti-inflammatory cytokines in autism spectrum disorders: A systematic review and meta-analysis. *Cytokine* 2019;123:154740.
- Lee SH, Sharma M, Südhof TC, et al. Synaptic function of nicastrin in hippocampal neurons. *Proc Natl Acad Sci U S A* 2014;111:8973-8.
- Tohda C, Tohda M. Extracellular cathepsin L stimulates axonal growth in neurons. *BMC Res Notes* 2017;10:613.
- Sener EF, Cikili Uytun M, Korkmaz Bayramov K, et al. The roles of CC2D1A and HTR1A gene expressions in autism spectrum disorders. *Metab Brain Dis* 2016;31:613-9.
- Anitha A, Nakamura K, Yamada K, et al. Genetic analyses of roundabout (ROBO) axon guidance receptors in autism. *Am J Med Genet B Neuropsychiatr Genet* 2008;147B:1019-27.
- Najera K, Fagan BM, Thompson PM. SNAP-25 in Major Psychiatric Disorders: A Review. *Neuroscience* 2019;420:79-85.
- Chapman NH, Nato AQ Jr, Bernier R, et al. Whole exome sequencing in extended families with autism spectrum disorder implicates four candidate genes. *Hum Genet* 2015;134:1055-68.
- Çetin İ, Tezdiç İ, Taraklıoğlu MC, et al. Do Low Serum UCH-L1 and TDP-43 Levels Indicate Disturbed Ubiquitin-Proteosome System in Autism Spectrum Disorder? *Noro Psikiyat Arş* 2017;54:267-71.
- Careaga M, Hansen RL, Hertz-Pannier I, et al. Increased anti-phospholipid antibodies in autism spectrum disorders. *Mediators Inflamm* 2013;2013:935608.
- Hu Z, Yang Y, Zhao Y, et al. APOE hypermethylation is associated with autism spectrum disorder in a Chinese population. *Exp Ther Med* 2018;15:4749-54.
- Shimamoto C, Ohnishi T, Maekawa M, et al. Functional characterization of FABP3, 5 and 7 gene variants identified in schizophrenia and autism spectrum disorder and mouse behavioral studies. *Hum Mol Genet* 2015;24:2409.
- Yamamoto Y, Kida H, Kagawa Y, et al. FABP3 in the Anterior Cingulate Cortex Modulates the Methylation Status of the Glutamic Acid Decarboxylase67 Promoter Region. *J Neurosci* 2018;38:10411-23.
- Herbert MR, Russo JP, Yang S, et al. Autism and

- environmental genomics. *Neurotoxicology* 2006;27:671-84.
- 29. Mandic-Maravic V, Mitkovic-Voncina M, Pljesa-Ercegovac M, et al. Autism Spectrum Disorders and Perinatal Complications-Is Oxidative Stress the Connection? *Front Psychiatry* 2019;10:675.
 - 30. Benga O, Huber VJ. Brain water channel proteins in health and disease. *Mol Aspects Med* 2012;33:562-78.
 - 31. Duffney LJ, Zhong P, Wei J, et al. Autism-like Deficits in Shank3-Deficient Mice Are Rescued by Targeting Actin Regulators. *Cell Rep* 2015;11:1400-13.
 - 32. Zhang L, Huang CC, Dai Y, et al. Symptom improvement in children with autism spectrum disorder following bumetanide administration is associated with decreased GABA/glutamate ratios. *Transl Psychiatry* 2020;10:9.
 - 33. Roohi J, Montagna C, Tegay DH, et al. Disruption of contactin 4 in three subjects with autism spectrum disorder. *J Med Genet* 2009;46:176-82.
 - 34. Yoo HJ, Kim K, Kim IH, et al. Whole exome sequencing for a patient with Rubinstein-Taybi syndrome reveals de novo variants besides an overt CREBBP mutation. *Int J Mol Sci* 2015;16:5697-713.
 - 35. McFadden K, Minshew NJ. Evidence for dysregulation of axonal growth and guidance in the etiology of ASD. *Front Hum Neurosci* 2013;7:671.
 - 36. Krueger DD, Brose N. Evidence for a common endocannabinoid-related pathomechanism in autism spectrum disorders. *Neuron* 2013;78:408-10.
 - 37. Zamberletti E, Gabaglio M, Parolaro D. The Endocannabinoid System and Autism Spectrum Disorders: Insights from Animal Models. *Int J Mol Sci* 2017;18:1916.
 - 38. Zhou J, Parada LF. PTEN signaling in autism spectrum disorders. *Curr Opin Neurobiol* 2012;22:873-9.
 - 39. Guang S, Pang N, Deng X, et al. Synaptopathology Involved in Autism Spectrum Disorder. *Front Cell Neurosci* 2018;12:470.
 - 40. Ashwood P. Differential T Cell Levels of Tumor Necrosis Factor Receptor-II in Children with Autism. *Front Psychiatry* 2018;9:543.
 - 41. Parikshak NN, Swarup V, Belgard TG, et al. Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature* 2016;540:423-7.

Cite this article as: Meng W, Huan Y, Gao Y. Urinary proteome profiling for children with autism using data-independent acquisition proteomics. *Transl Pediatr* 2021;10(7):1765-1778. doi: 10.21037/tp-21-193

Supplementary

Table S1 The composition of age and gender in the ASD and HC group

ID	Gender	Age (years)
ASD (n=18)		
ASD1	Male	5
ASD2	Male	5
ASD3	Male	7
ASD4	Female	4
ASD5	Male	3
ASD6	Male	4
ASD7	Male	3
ASD8	Male	7
ASD9	Female	4
ASD10	Male	4
ASD11	Male	6
ASD12	Male	7
ASD13	Male	6
ASD14	Male	4
ASD15	Male	8
ASD16	Female	4
ASD17	Male	5
ASD18	Male	4
HC (n=6)		
HC1	Male	6
HC2	Male	5
HC3	Male	5
HC4	Male	3
HC5	Male	5
HC6	Male	5

ASD, autism spectrum disorder; HC, healthy control.

Table S2 The variable isolation window of the DIA method

m/z	z	t start (min)	t stop (min)	Isolation window (m/z)
368.5	2	0	90	37
398	2	0	90	22
418.5	2	0	90	19
439	2	0	90	22
458.5	2	0	90	17
474	2	0	90	14
488.5	2	0	90	15
503	2	0	90	14
518	2	0	90	16
534.5	2	0	90	17
549.5	2	0	90	13
562.5	2	0	90	13
576	2	0	90	14
591	2	0	90	16
607.5	2	0	90	17
623	2	0	90	14
637	2	0	90	14
651.5	2	0	90	15
667.5	2	0	90	17
685.5	2	0	90	19
704	2	0	90	18
721.5	2	0	90	17
737.5	2	0	90	15
753.5	2	0	90	17
771	2	0	90	18
789.5	2	0	90	19
808	2	0	90	18
828.5	2	0	90	23
851.5	2	0	90	23
874.5	2	0	90	23
900.5	2	0	90	29
932.5	2	0	90	35
969	2	0	90	38
1013.5	2	0	90	51
1073.5	2	0	90	69
1253.5	2	0	90	291

m/z, mass-to-charge ratio; DIA, data-independent acquisition.

Table S3 All biological processes were enriched in this study

Term	–log (P value)	Proteins
Viral budding via host ESCRT complex	16.161	Q9Y3E7, Q9BY43, Q96EY5, Q9UN37, Q7LBR1, Q8WV92, Q99816, Q8WUM4, O43633, O75351, Q9HD42
Multivesicular body assembly	14.347	Q9Y3E7, Q9BY43, Q96EY5, Q9UN37, Q7LBR1, Q9BRG1, Q99816, Q8WUM4, O43633, O75351, Q9HD42
Cell separation after cytokinesis	12.886	Q9Y3E7, Q9BY43, Q9UN37, Q7LBR1, Q8WV92, Q8WUM4, O43633, O75351, Q9HD42
Nucleus organization	9.538	Q9Y3E7, Q9BY43, Q9UN37, Q7LBR1, Q8WUM4, O43633, O75351, Q9HD42
Mitotic metaphase plate congression	8.377	Q9Y3E7, Q9BY43, Q9UN37, Q7LBR1, Q8WUM4, O43633, O75351, Q9HD42
regulation of mitotic spindle assembly	7.638	Q9Y3E7 Q7LBR1 Q8WUM4 O43633 O75351 Q9HD42
positive regulation of exosomal secretion	7.469	Q9UN37 O00560 Q99816 Q8WUM4 O43633 O75351
viral life cycle	7.310	Q9Y3E7 Q9BY43 Q9UN37 Q99816 Q8WUM4 O43633 O75351
protein transport	7.252	Q16348 Q9UN37 Q96EY5 Q8IX04 O43633 O75351 P51148 P50897 Q9Y3E7 O00161 Q7LBR1 P61204 P62834 Q99816 Q8WUM4 Q9HD42
regulation of centrosome duplication	6.886	Q9Y3E7 Q7LBR1 Q8WUM4 O43633 O75351 Q9HD42
endosomal transport	6.569	Q9Y3E7 Q9BY43 Q9UN37 Q9GZM7 Q9BRG1 Q99816 O43633 O75351
ESCRT III complex disassembly	6.420	Q9UN37 Q7LBR1 O43633 O75351 Q9HD42
positive regulation of viral release from host cell	6.222	Q9Y3E7 Q9UN37 Q99816 O43633 O75351
autophagy	5.523	Q9H0E2 Q9Y3E7 Q9BY43 Q96EY5 Q9UN37 Q9BRG1 Q99816 O43633 O75351
vacuolar transport	3.824	Q9BY43 Q7LBR1 O43633 Q9HD42
ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway	3.824	Q96EY5 Q9UN37 Q99816 O75351
ubiquitin-independent protein catabolic process via the multivesicular body sorting pathway	3.357	Q9UN37 Q8WUM4 O75351
regulation of viral process	3.187	Q9Y3E7 O43633 O75351
positive regulation of viral process	3.187	Q96EY5 Q99816 O75351
Ras protein signal transduction	2.921	P46109 P62070 O00560 P62873 P01111 113
protein folding	2.886	Q8NBS9 P63096 P62879 P02511 P08754 P62873 P14314
small GTPase mediated signal transduction	2.886	P63000 P61204 P62070 P62834 P51148 P36405 P60953 P01111
negative regulation of blood coagulation	2.553	P02749 P02649 P07204
protein polymerization	2.481	Q9Y3E7 Q9BY43 O43633
cell adhesion	2.444	Q8IWV2 P21926 Q9HBB8 P63000 P10586 Q9GZM7 P24821 P39060 Q9Y6N7 Q14254
proteolysis involved in cellular protein catabolic process	2.398	P07711 Q9GZM7 Q9H3G5 Q99538
axon guidance	2.367	P15311 Q8IWV2 P63000 O14786 Q9Y6N7 P01111
microvillus assembly	2.252	P15311 P62834 P61225
protein homooligomerization	2.174	Q9BY43 Q9NZN3 O43633 P02511 Q9H223 P27105
blood coagulation	2.108	P08758 P05160 P63000 Q9NZN3 P07204 P60953
epithelial cell differentiation	1.959	Q9H0E2 O00526 Q9UBD6 P21796
positive regulation of viral life cycle	1.886	Q9UN37 O75351
positive regulation of substrate adhesion-dependent cell spreading	1.721	P46109 P63000 P60953
ephrin receptor signaling pathway	1.699	P63000 O00560 Q92542 P60953
regulation of extracellular exosome assembly	1.699	Q9UN37 O75351
regulation of viral budding via host ESCRT complex	1.699	Q99816 Q8WUM4
positive regulation of viral budding via host ESCRT complex	1.699	Q9UN37 Q99816
late endosomal microautophagy	1.699	Q96EY5 Q99816
chemical synaptic transmission	1.658	P10586 O60939 O00560 P09543 P50897 Q14254
pinocytosis	1.585	P50897 Q9H223
positive regulation of extracellular exosome assembly	1.585	O00560 Q8WUM4
cellular response to glucagon stimulus	1.538	Q9BY43 P02649 P10912
negative regulation of neuron death	1.538	P63092 P62879 P62873
cell division	1.523	P63096 Q9UN37 Q7LBR1 Q99816 P08754 P60953 Q9HD42
platelet degranulation	1.495	P01033 Q13103 P21926 P02749
establishment or maintenance of apical/basal cell polarity	1.481	P15311 P60953
cellular protein modification process	1.481	P61088 O00462 Q99816 Q8IX04
receptor internalization	1.481	P15311 P21926 P10912
regulation of macroautophagy	1.456	P09936 O75348 P36543
vacuole organization	1.409	Q9UN37 O75351
hair follicle placode formation	1.409	Q9UN37 O75351
viral release from host cell	1.409	P63092 P60953
membrane invagination	1.409	Q9BY43 O43633
platelet activation	1.377	P21926 P63000 P61225 P62873
substantia nigra development	1.377	P09936 P09543 P60953
membrane budding	1.337	P15311 Q9HBB8
regulation of microvillus length	1.337	Q9BY43 Q9UN37
cellular response to catecholamine stimulus	1.337	P63092 P62873

ESCRT, endosomal sorting complex required for transport; GTPase, guanosine triphosphate hydrolase.