Proteomic analysis of serum proteins of children with autism

✍Anna L Kaysheva[1], Artur T Kopylov[1], Tatyana O Pleshakova[1], Ivan Y Iourov[2], Svetlana G Vorsanova[2], Yuri B Yurov[2], Andrey A Shchetkin[3], Alexander I Archakov[1], Yuri D Ivanov[1]

[1] Institute of Biomedical Chemistry, IBMC RAMS, Pogodinskaya St. 10/8, Moscow, 119121 Russia
[2] Mental Health Research Center, Kashirskoe Shosse 34, Moscow, 115522 Russia
[3] Center of Medical Prevention, Navaginskaya St. 9, 354000 Russia
✍kaysheva1@gmail.com

RESEARCH

ABSTRACT

A comparative panoramic mass spectrometric analysis of serum samples of children suffering from autistic disorders was carried out. Three families were examined, with nine samples distributed in four control samples of healthy parents and five samples of autistic children. In the course of comparative analysis of protein composition of the serum samples, a small group of potential marker proteins for diagnosis was selected, which are not functionally or structurally related to one another. The group included 15 proteins; one of them, LIM domain-containing protein 1, was identified in four out of the five autism depleted serum samples, and the other 14 were identified in three samples. This could be the basis to develop highly sensitive quantitative mass spectrometric SRM methods to detect autism markers in serum.

Keywords: autism, ESI-MS/MS, mass spectrometry, diagnosis, protein markers

Biotecnología Aplicada 2017;34:2211-2214

RESUMEN

Análisis proteómico del perfil de proteínas séricas de niños con autismo. Se estudiaron muestras de suero de niños autistas mediante el análisis panorámico comparativo de espectrometría de masas. Tres familias fueron analizadas, con nueve muestras distribuidas en cuatro correspondientes a padres sanos y cinco a niños autistas. Se seleccionó un subconjunto de 15 proteínas, sin relación estructural ni funcional entre ellas, como biomarcadores con potencial para el diagnóstico. Una de ellas, la proteína 1 portadora de dominio LIM, estuvo presente en cuatro de las cinco muestras de pacientes estudiadas. Las otras 14 proteínas fueron identificadas en tres de las cuatro muestras sin suero de los pacientes. Estos biomarcadores pudieran servir de base para desarrollar métodos de espectrometría de masa SRM para la detección del autismo a partir de muestras de suero.

Palabras clave: autismo, ESI-MS/MS, espectrometría de masa, diagnóstico, proteínas marcadoras

*I*ntroduction

Autism is a non-communicable disease of increasing prevalence, with global estimates of approximately 67 million people worldwide suffering from autism, and each year this level increases by 14 % [1]. Owing to the difficulty of diagnosing this disorder, no official data on the number of autistic children in Russia is available. However, the groups of individuals at risk, for the development of autism are large enough. To date, more than 80 % of children with autism are disabled persons. In 2012, the Center for disease control in the USA reported, on average, about one case of autism for every 88 children [2]. According to the international classification of diseases ICD-10, autistic disorders include: childhood autism (F84.0), atypical autism (after 3 years) (F84.1), Rett syndrome (F84.2) and Asperger syndrome (F84.5) [3]. Autism can be caused by genetic modifications, organic lesions of the central nervous system, metabolic disorders, the exposure to viral and bacterial infections, chemical exposure to the mother's body during pregnancy, among others. In fact, autism in children is more common than cancer, diabetes and AIDS combined [2].

In the early 2000's, the genes associated with development of autistic disorders were annotated, and their heritability was established [4-6]. Based on genetic data, it was shown that autistic disorders are heterogeneous [7]. Genes associated with development of autistic disorders may participate in other psychiatric and neurological disturbances [8]. There are more than 60 genetic mutations in chromosomes 11, 12, 13, 14 and 16 that are associated with the risk of autism, predominantly of paternal origin [8]. However, careful analysis of proteins associated with these genes showed that most of the observed *de novo* mutations are not associated with autism spectrum disorders [9]. To date, only the protein products of the CHD8 and KATNAL2 genes [9] are acknowledged by the scientific community to be associated with the development of autistic disorders. Therefore, it appears that genomic information is not unambiguous and provides only an indirect view as to risk of development of autistic disorders. A deeper understanding of the causes of autism can be achieved by identifying functional protein markers along with phenotypic and behavioral reactions [10]. It is essential the availability of protein biomarkers able to provide early diagnosis at the preverbal stage [11]. In addition, protein biomarkers will identify targets for drug therapy and determine biological reference points for monitoring behavioral changes or treatment efficiency, and could reveal the underlying causes for the development of autistic disorders.

To date, the scientific literature provides evidence for the increased levels of the following proteins in the serum samples of autistic patients: the complement system proteins [12, 13]: apolipoproteins ApoA1, ApoA4 and PON1 [14]; GFAP, apoptosis factor Bcl-2, a factor of the metabolism of glutamate GAD-2

1. Kopetz PB, Endowed ED. Autism worldwide: Prevalence, perceptions, acceptance, action. J Soc Sci. 2012;8:196-201.

2. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators; Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders--Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. MMWR Surveill Summ. 2012;61(3):1-19.

3. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders. Clinical descriptions and diagnostic guidelines. Geneva: WHO; 2004.

4. Conciatori M, Stodgell CJ, Hyman SL, O'Bara M, Militerni R, Bravaccio C, *et al*. Association between the HOXA1 A218G polymorphism and increased head circumference in patients with autism. Biol Psychiatry. 2004;55(4):413-9.

5. Freitag CM, Staal W, Klauck SM, Duketis E, Waltes R. Genetics of autistic disorders: review and clinical implications. Eur Child Adolesc Psychiatry. 2010;19(3):169-78.

6. Hu VW, Nguyen A, Kim KS, Steinberg ME, Sarachana T, Scully MA, *et al*. Gene expression profiling of lymphoblasts from autistic and nonaffected sib pairs: altered pathways in neuronal development and steroid biosynthesis. PLoS One. 2009;4(6):e5775.

[15], and autoantibodies against brain proteins [16]. Also the levels of C-reactive protein [17] and chemokines during pregnancy were found increased [18]. All these proteins are regarded as potential markers of autism in children.

We have investigated five serum samples of children suffering from autism spectrum disorders using the panoramic mass spectrometric analysis of high resolution. It appears that, in contrast to control samples, at least 15 functionally and structurally related proteins (attributed by us to the group of potential marker proteins) are detectable in their sera, one of them, the LIM domain-containing protein 1, of potential interest for autism diagnosis.

Materials and methods

Reagents

Acetonitrile and TCA were from Merck (Germany). Formic acid was from ACROS Organics (USA). Ethylenediaminetetraacetic acid (EDTA) was from Sigma-Aldrich (USA). Modified trypsin was from Promega (USA).

Serum samples

Four control samples of serum from healthy volunteers and five samples of blood serum of children with autistic disorders was provided by the Veltischev Clinical Pediatric Research Institute of Pirogov, at the Russian National Research Medical University, in Moscow, Russia. Samples of blood serum were obtained from three families. The first family includes healthy mother (aged 48) – C1 sample, two sons with autistic disorder (aged 10 and 20) samples S2 and S3 samples, respectively. The second family includes: healthy mother (aged 34) and father (aged 38) – C4 and C6 samples, respectively, and a son with autistic disorder (aged 6) – S5. The third family includes: healthy mother (aged 37) – C8, and children with autistic disorders (aged 7 and 8) – S7 and S9 samples, respectively. Written informed consents were obtained from the parents of autistic children studied, authorizing the participation in the study and the use of the biological material for the intended research use.

The material under study consisted of serum samples from fresh venous blood obtained in the morning on fasting. Blood was collected in pre-chilled tubes with EDTA, quickly mixed and centrifuged at 4 ºC and 1500 rpm for 10 min. After centrifugation, the supernatant in the tubes was carefully collected with an automated pipette and placed in nine cryovials with a volume of 2 mL. The serum samples were stored in cryovials at -80 ºC until use [19].

The depletion of serum samples from major protein fractions was performed using affinity HPLC columns MARS Hu-14 (Agilent, USA) according to the manufacturer's protocol. Hydrolytic cleavage of proteins in serum samples was performed according to Kopylov *et al*. [20].

Mass spectrometric analysis of the peptide composition in serum samples was conducted for serum-depleted samples: four control (labeled as C1, C4, C6, C8) and five patients with autistic disorder (labeled S2, S3, S5, S7, S9).

Methods

Mass spectrometric protein registration in the serum samples

Hydrolyzed samples of serum (the original and enrichment) were analyzed by reversed phase liquid chromatography (LC) in conjunction with tandem mass spectrometry (MS/MS). Separation of peptides was performed using Ultimate 3000 Nano-flow HPLC chromatographic system (Thermo Scientific, USA) coupled with Orbitrap Q Exactive high resolution mass spectrometer (Thermo Scientific, USA).

Peptides were separated using PepMap C18 analytical column (Thermo Scientific, USA; column dimensions 3 mm × 500 µm, particle size 5 µm) in a linear gradient from 98 % buffer A (water, 0.1 % formic acid) and 2 % buffer B (water, 0.1 % formic acid, 80 % acetonitrile) to 65 % buffer B for 95 min at a flow rate of 0.4 µL/min.

Peptides were analyzed by LC-MS/MS with the the nanospray ionization (NSI) source in three technical replicates for each sample with high-energy fragmentation induced by collision higher energy collisional dissociation (HCD), and normalized collision energy was 27 % (per 400 m/z and z = 2+). Singly charged ions and ions with undetermined charge state were excluded upon MS/MS scanning.

Analysis of mass spectrometric data

MS/MS spectra in RAW format were processed using Mass Hunter (version B 2.0) software (Agilent, USA). To identify the proteins, the Mascot proteomic search engine (hereinafter –Mascot, www.matrixscience.com) was used with the following search parameters: database (Database) – SwissProt "SwissProt 2012_11"; taxonomic group (Taxonomy filter) – human; fixed modifications (Fixed modifications) – none; non-fixed modifications (variable modifications – oxidation (M); enzyme – none; Peptide Mass Tolerance – 20 ppm; Fragment Mass Tolerance – 50 mmu; Use MUDPIT protein scoring – 1; Include Sub-Set Proteins – 1; Use Homology Threshold – 1; Group Protein Families – 1. Peptides and proteins identified with Mascot with index of reliability (score) more than 30 and 50 consequently, and an emPAI factor of less than 0.02, were considered unreliable [21].

In addition, another proteomic search engine, the X!Tandem (http://human.thegpm.org/), was used for protein identification, with the following search options: Taxon – human (SwissProt), Fragment mass error – 0.05 Da, Potential modifications – oxidation (M), Method: Select device & parent δm – Orbitrap (20 ppm), log(e) < –1.5.

Results and discussion

We have conducted a panoramic mass spectrometric analysis (shotgun LC-MS/MS) of protein composition of nine depleted serum samples, including four controls (C1, C4, C6 and C8) and five samples of autistic children (S2, S3, S5, S7 and S9). A complete list of proteins identified in our experiments using two proteomic search engines – Mascot and X!Tandem – is given in Supplementary 1.

Depleted serum samples differ in the number of identified proteins and in protein composition.

7. O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, *et al*. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature. 2012;485(7397):246-50.

8. Guilmatre A, Dubourg C, Mosca AL, Legallic S, Goldenberg A, Drouin-Garraud V, *et al*. Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. Arch Gen Psychiatry. 2009;66(9):947-56.

9. Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, *et al*. Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature. 2012;485(7397):242-5.

10. Hartley SL, Seltzer MM, Hong J, Greenberg JS, Smith L, Almeida D, *et al*. Cortisol response to behavior problems in FMR1 premutation mothers of adolescents and adults with fragile X syndrome: A diathesis-stress model. Int J Behav Dev. 2012;36(1):53-61.

11. Ngounou Wetie AG, Sokolowska I, Wormwood K, Beglinger K, Michel TM, Thome J, *et al*. Mass spectrometry for the detection of potential psychiatric biomarkers. J Mol Psychiatry. 2013;1(1):8.

12. Corbett BA, Kantor AB, Schulman H, Walker WL, Lit L, Ashwood P, *et al*. A proteomic study of serum from children with autism showing differential expression of apolipoproteins and complement proteins. Mol Psychiatry. 2007;12(3):292-306.

13. Woods AG, Sokolowska I, Taurines R, Gerlach M, Dudley E, Thome J, *et al*. Potential biomarkers in psychiatry: focus on the cholesterol system. J Cell Mol Med. 2012;16(6):1184-95.

14. Ngounou Wetie AG, Wormwood K, Thome J, Dudley E, Taurines R, Gerlach M, *et al*. A pilot proteomic study of protein markers in autism spectrum disorder. Electrophoresis. 2014;35(14):2046-54.

15. Eichler EE, Zimmerman AW. A hot spot of genetic instability in autism. N Engl J Med. 2008;358(7):737-9.

16. Elamin NE, Al-Ayadhi LY. Brain autoantibodies in autism spectrum disorder. Biomark Med. 2014;8(3):345-52.

17. Lehti V, Hinkka-Yli-Salomaki S, Cheslack-Postava K, Gissler M, Brown AS, Sourander A. The risk of childhood autism among second-generation migrants in Finland: a case-control study. BMC Pediatr. 2013;13:171.

18. Zerbo O, Yoshida C, Grether JK, Van de Water J, Ashwood P, Delorenze GN, *et al*. Neonatal cytokines and chemokines and risk of Autism Spectrum Disorder: the Early Markers for Autism (EMA) study: a case-control study. J Neuroinflammation. 2014;11:113.

19. Kaysheva AL, Ivanov YD, Frantsuzov PA, Krohin NV, Pavlova TI, Uchaikin VF, *et al*. Mass spectrometric detection of the amino acid sequence polymorphism of the hepatitis C virus antigen. J Virol Methods. 2016;229:86-90.

20. Kopylov AT, Zgoda VG, Lisitsa AV, Archakov AI. Combined use of irreversible binding and MRM technology for low- and ultralow copy-number protein detection and quantitation. Proteomics. 2013;13(5):727-42.

The number of identified proteins using Mascot in depleted serum samples varies within the 25 % range, while the number of identifications revealed through X!Tandem varies within the 40 % range (Figure 1).

Among all four control samples of depleted serum, the matching of protein identifications revealed with Mascot was 78 proteins, while with X!Tandem it was 82 proteins (Figure 2).

In control samples, the protein composition detected by both search engines was very similar. Large groups of apolipoproteins, proteins of the blood coagulation system and complement, immunoglobulins, as well as keratins and the contractile proteins (actin, tropomyosin) were discovered (Supplementary information 1). In the course of identification, two comparison lists for control proteins have been established: the first (for Mascot) included 130 proteins, and the second (for X!Tandem) included 171 proteins. Comparison lists contain the names and UniProtKB identifiers of proteins that were revealed in at least two control samples.

It is interesting to point out that in control serum samples, X!Tandem proteomic search engine identified virtually all the proteins (126 proteins out of 130) from the list of control proteins for Mascot. In control serum samples, some proteins of the complement system, apolipoproteins, hepatocyte growth factor, G-chain of immunoglobulins G, beta-2-glycoprotein and other high- and medium-abundant serum proteins were identified by X!Tandem and were not found by Mascot.

A subsequent comparative analysis with depleted serum samples of children suffering from autistic disorders was conducted with the use of these lists. The analysis of protein composition of autistic depleted serum samples, that differ from proteins in comparison lists of control serum samples, revealed the heterogeneity of results obtained by the two search engines.

The comparison of protein identifications obtained using Mascot and X!Tandem for the serum samples of children suffering from autism spectrum disorders has shown that, in average, 80 % of the proteins identified by Mascot are also identified by X!Tandem (Supplementary information 2). It is also worth noting that X!Tandem proteomic search engine identifies in the samples a greater number of proteins than Mascot. The latter, in its turn, demonstrates better coverage of the amino acid sequence of the identified proteins in comparison with X!Tandem. Detailed comparison of the identifications of the two searches is presented in Supplementary 2. It was evidenced that the selected most popular search engines supplemented each other and allowed the most complete evaluation of the protein composition of the analyzed biological samples.

Among the identifications obtained with the Mascot, only two proteins differ from the comparison list in three out of the five samples collected from autistic children. The greater number of matches between serum samples of autistic children was detected using X!Tandem. Thus, with the use of this search engine, the LIM domain-containing protein 1 (Q9UGP4) was identified in four out of the five samples. In addition, 12 more proteins were identified in three out of the five samples collected from the autistic children (Table 1).

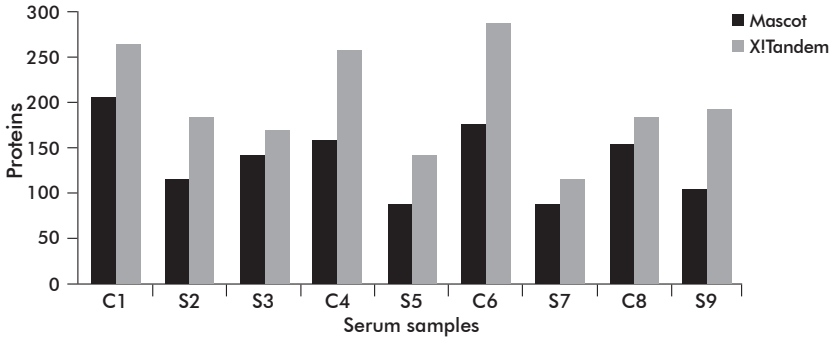

Figure 1. Number of proteins identified with the Mascot and X!Tandem software in depleted control samples (C1, C4, C6 and C8) and in serum samples of children with autistic disorders (S2, S3, S5, S7 and S9). See the text for parameters used for proteins identification by Mascot and X!Tandem software.

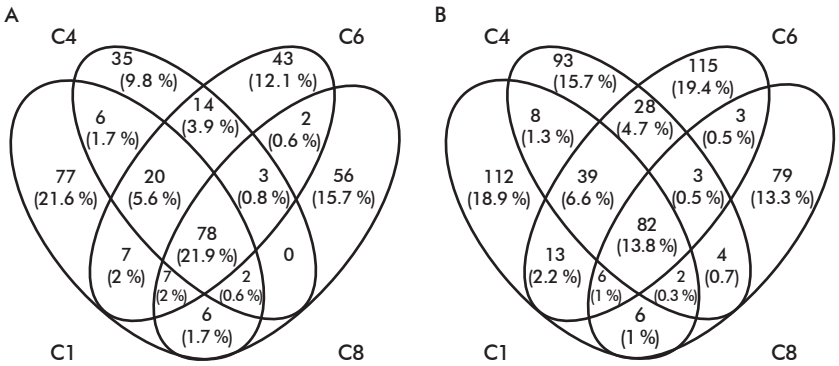

Figure 2. Overlap of protein identifications among depleted control samples (C1, C4, C6 and C8). A) Proteins identified by Mascot. B) Proteins identified by X!Tandem. See the text for parameters used for proteins identification by either software.

Table 1. Group of potential markers of autistic disorders identified by Mascot and X!Tandem

Software	Protein	UniProt AC	Samples
Mascot	Phospholipid transfer protein	P55058	S2, S3, S5
	Attractin	O75882	S2, S3, S5
X!Tandem	LIM domain-containing protein 1	Q9UGP4	S2,S3,S7,S9
	Hepatocyte growth factor-like protein	P26927	S2, S3, S5
	LIM domain-containing protein 1	Q9UGP4	S2,S3,S7
	Lymphoid-restricted membrane protein	Q12912	S2,S3,S7
	Cat eye syndrome critical region protein 2	Q9BXF3	S2,S3,S7
	Proline-, glutamic acid- and leucine-rich protein 1	Q8IZL8	S2,S3,S7
	UNC5C-like protein	Q8IV45	S2,S3,S7
	Adenylyl cyclase-associated protein 1	Q01518	S2,S3,S9
	14-3-3 protein beta/alpha	P31946	S2,S3,S9
	LIM domain-containing protein 1	Q9UGP4	S2,S3,S9
	Cryptochrome-2	Q49AN0	S2,S3,S9
	Protein S100-A6	P06703	S2,S3,S9
	Methionine adenosyltransferase 2 subunit beta	Q9NZL9	S2,S3,S9

Proteins identified only in two serum samples were not analyzed. A group of potential markers was proposed based on these analyses. This group included proteins detected in three and four depleted serum samples of children suffering from autistic disorders (Table 1).

Noteworthy, the most authoritative database STRING 10 (http://string-db.org/), designed for the detection of direct or indirect protein interactions, in

21. Ivanov YD, Kaysheva AL, Frantsuzov PA, Pleshakova TO, Krohin NV, Izotov AA, *et al*. Detection of hepatitis C virus core protein in serum by atomic force microscopy combined with mass spectrometry. Int J Nanomedicine. 2015;10:1597-608.

which 184 millions of interactions for 9.6 millions of proteins are annotated, has not revealed direct or indirect associations (including functional ones) among the proteins from the group of conventional markers. It is interesting that, according to protein knowledgebase UniProt KB (http://www.uniprot.org/), most of the group of potential markers (consisting of 8 proteins) is involved in cell proliferation and differentiation. The rest of the proteins perform transport function (one protein) or participate in the implementation of immune response (three proteins).

In comparison with a few other studies we have revealed significant inconsistencies in the list of potential biomarkers. So, we have revealed the group of apolipoproteins, proteins of the complement system and the C-reactive protein were present in both types of serum samples, control (normal) and autistic [10-14]. Therefore, they were not attributed for us to the group of potential markers. Moreover, in many studies it is shown that the proteins listed above are found in human normal serum [22, 23]. In its turn, we have not identified glutamate metabolism factor GAD-2 and autoantibodies against brain proteins in the control samples or in the autistic serum samples.

Received in October, 2016.
Accepted in April, 2017.

Conclusions

In the course of a exploratory comparative study of five samples of depleted sera of autistic children, it was found that, compared to the control samples, the serum samples of autistic children contain a small group of common proteins that are not functionally or structurally related to one another. We have designated them as potential markers. Among these proteins, only the LIM domain-containing protein 1 was revealed in four out of five samples. The remaining 14 proteins of this group were only detected in three samples of children suffering from autistic disorders. It should be noted that we do not identify a greater overlap when comparing the list of proteins of brothers with autism (S2 and S3) with respect to the other children's.

Nevertheless, in the future we plan to continue the study using a large number of samples. We intend to develop highly sensitive (at the level of 10^{-15} M) quantitative mass spectrometric SRM methods for the evaluation of the selected proteins in serum.

Acknowledgements

This study was supported by the Russian Science Foundation, Grant No.14-25-00132.

22. Anderson L, Anderson NG. High resolution two-dimensional electrophoresis of human plasma proteins. Proc Natl Acad Sci U S A. 1977;74(12):5421-5.

23. Semple SJ, Smith LL, McKune AJ, Hoyos J, Mokgethwa B, San Juan AF, *et al*. Serum concentrations of C reactive protein, alpha1 antitrypsin, and complement (C3, C4, C1 esterase inhibitor) before and during the Vuelta a Espana. Br J Sports Med. 2006;40(2):124-7.