

Identification of a robust non-coding RNA signature in diagnosing autism spectrum disorder by cross-validation of microarray data from peripheral blood samples

Wei Cheng, MD, Shanhu Zhou, MD, Jinxia Zhou, MD, Xijia Wang, MD*

Abstract

Novel molecular signatures are needed to improve the early and accurate diagnosis of autism spectrum disorder (ASD), and indicate physicians to provide timely intervention. This study aimed to identify a robust blood non-coding RNA (ncRNA) signature in diagnosing ASD. One hundred eighty six blood samples in the microarray dataset were randomly divided into the training set ($n = 112$) and validation set ($n = 72$). Then, the microarray probe expression profile was re-annotated into the expression profile of 4143 ncRNAs through probe sequence mapping. In the training set, least absolute shrinkage and selection operator (LASSO) penalized generalized linear model was adopted to identify the 20-ncRNA signature, and a diagnostic score was calculated for each sample according to the ncRNA expression levels and the model coefficients. The score demonstrated an excellent diagnostic ability for ASD in the training set (area under receiver operating characteristic curve [AUC] = 0.96), validation set (AUC = 0.97) and the overall (AUC = 0.96). Moreover, the blood samples of 23 ASD patients and 23 age- and gender-matched controls were collected as the external validation set, in which the signature also showed a good diagnostic ability for ASD (AUC = 0.96). In subgroup analysis, the signature was also robust when considering the potential confounders of sex, age, and disease subtypes. In comparison with a 55-gene signature deriving from the same dataset, the ncRNA signature showed an obviously better diagnostic ability (AUC: 0.96 vs 0.68, $P < .001$). In conclusion, this study identified a robust blood ncRNA signature in diagnosing ASD, which might help improve the diagnostic accuracy for ASD in clinical practice.

Abbreviations: ASD = autism spectrum disorder, AUC = area under ROC curve, GEO = Gene Expression Omnibus, HC = healthy control, ncRNAs = non-coding RNAs, PDD-NOS = pervasive developmental disorder-not otherwise specified, ROC = receiver operating characteristic, SE = standard error.

Keywords: autism spectrum disorder, diagnosis, non-coding RNA, signature

1. Introduction

Autism spectrum disorder (ASD) is a heterogeneous set of neurodevelopmental diseases, characterized by deficits in social communication and verbal/nonverbal interaction, as well as restricted and repetitive patterns of interests and behaviors. It has a high prevalence of approximately 0.3% to 1.2%, with 3 main

subtypes of autistic disorder, Asperger disorder and pervasive developmental disorder-not otherwise specified (PDD-NOS).^[1] Despite of onsite before 3 years old, most children are diagnosed with ASD after 4 years old.^[2] Early intensive behavioral interventions could improve the outcomes (e.g., language skills, cognitive performance, and adaptive behavior skills) in some young children with ASD.^[3] Thus, it has a critical need in clinical practice to increase the diagnostic accuracy for ASD.

As functional RNA molecules, non-coding RNAs (ncRNAs) are transcribed from DNA but not translated into proteins, including the types of long non-coding RNA (lncRNA), pseudogene, small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), ribosomal RNA (rRNA), miscellaneous RNA (miscRNA), and so on. ncRNAs have been reported in the pathogenesis of ASD, and aberrant expression of ncRNAs was detected in peripheral blood of ASD patients.^[4,5] With great advances in genetic detection, gene expression profiles are available to identify novel and robust biomarkers. In this study, we obtain the ncRNA expression profiles through microarray probe re-annotation, and identify and validate a blood ncRNA signature for the diagnosis of ASD.

2. Methods

2.1. Data preparation

The database of Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>) was a public functional genomics data

Editor: Lanjing Zhang.

WC and SZ have contributed equally to this study.

Funding: None.

The authors have no conflicts of interest to disclose.

Department of Neurology, Puren Hospital Affiliated to Wuhan University of Science and Technology, Wuhan, China.

* Correspondence: Xijia Wang, Department of Neurology, Puren Hospital Affiliated to Wuhan University of Science and Technology, Wuhan 430081, China (e-mail: wxjy2007@163.com).

Copyright © 2020 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Cheng W, Zhou S, Zhou J, Wang X. Identification of a robust non-coding RNA signature in diagnosing autism spectrum disorder by cross-validation of microarray data from peripheral blood samples. *Medicine* 2020;99:11(e19484).

Received: 31 August 2019 / Received in final form: 4 January 2020 / Accepted: 7 February 2020

<http://dx.doi.org/10.1097/MD.0000000000019484>

repository of array- and sequence-based data, and users could query and download experiments and curated gene expression profiles. In this database, we searched ASD-related datasets of gene expression profiles from inception to June 2019, using the key words including: (“autism spectrum disorder” OR “Asperger” OR “autis*” OR “pervasive developmental disorder” OR “childhood disintegrative disorder”). Datasets were included if meeting the following criteria: detected blood gene expression profiles (probe-tabulated) of both ASD cases and healthy controls (HC); availability of clinical data and corresponding sequences of the microarray probes; the sample size was large enough. Then, the tab-delimited expression value-matrix table was downloaded and \log_2 -transformed. This study was approved by the ethnic committee of Puren Hospital Affiliated to Wuhan University of Science and Technology.

2.2. Probe re-annotation

First, we obtained the microarray probe sequences from the Affymetrix product website (<http://www.affymetrix.com>), as well as the human genome sequences (GRCh38.p12) and comprehensive gene annotation from the GENCODE database (<https://www.gencodegenes.org>).^[6] Then, the probe sequences were aligned to the human genome sequences, using the software of hierarchical indexing for spliced alignment of transcripts (HISAT). No mismatches between probe sequence and reference sequence were allowed. Alignments were stored in SAM format where all alternative best hits were included. Among the successful alignments to the reference genome, we only included the transcripts with at least 1 alignment, each of which was also matched to only 1 transcript. When multiple probes matched to an identical gene, the average expression value across these probes was calculated to represent the corresponded gene. ncRNA expression profiles were extracted after excluding protein-coding RNAs (protein-coding and nonsense mediated decay) according to the RNA types.

2.3. ncRNA signature construction, evaluation, and validation

The samples were randomly divided into the training set and validation set according to the ratio of 6:4. In the training set, a least absolute shrinkage and selection operator (LASSO) penalized generalized linear model was adopted to identify significant ncRNAs.^[7,8] The penalty parameter was estimated by 10-fold cross-validation at 1 standard error (SE) beyond the minimum partial likelihood deviance. Then, the coefficients of significant ncRNAs in the model were extracted to calculate a diagnostic score for each sample in the training set, validation set, and the overall. In receiver operating characteristic (ROC) curve analysis, area under ROC curve (AUC) was calculated to evaluate the diagnostic ability of the signature. Moreover, subgroup analysis was conducted on sex, age, and disease subtypes to assess the diagnostic stability of the signature, and DeLong test for 2 ROC curves was performed to investigate the difference between subgroups. Finally, we also compared the ncRNA signature with a 55-gene signature which derived from the same dataset.^[9]

2.4. External validation

The blood samples of 23 ASD and 23 age- and sex-matched controls were obtained in Puren Hospital from September 2015 to July 2017. Written informed consent was provided before sample collection, and the present study protocol was approved by the ethnic committee of Puren Hospital. Then, total RNA was

extracted using TRIzol reagent (Invitrogen, Waltham, MA, USA), and stored at -80°C . The RNA concentration and purity were measured by the NanoDrop spectrophotometer (Thermo Fisher, Waltham, MA, USA). Total RNA was synthesized into first-strand cDNA using fluorescent-labeled dNTPs (Thermo Fisher, Waltham, MA, USA), before hybridization with a customized microarray which tailed and fixed 20 ncRNA probes (CapitalBio, China). The probe sequences were as follows:

RNU1-16P: GGGACTATGTTTCGTGTTCTCTCCTG
 RNU6-258P: AAGTCGTGAAATAGTCCATATGTTA
 RNU6-485P: CCCTGTGCAAGGATGATATGCAAAT
 RNU6-549P: GCTCACTTCAGTGGTACATATACTA
 RNVU1-15: GAACTCGACTGCATAAATTGTGATA
 RNASSP160: CTATGGCCATAAAACCCTGAACTTG
 IGHV3-47: TCTCCAGAGACAACGCCAAGAAGTC
 TUBB2BP1: CATGCCCTCACCCAAGGTGTCAGAC
 CHST9-AS1: GGAAGTCTTCAGTATCTGTACAAC
 AC074183.2: TGCCCAGGAGAAAGGCTGAAGGACA
 AC136940.3: TTAGATTACGCTTGGCTTCTTTTGA
 RP4-646B12.2: GAAAAAAGTGTGAATCAGTCAGTAC
 RP11-90M2.3: ATTGGTGGGTTTTGATGCCCGAGTA
 RN7SL132P: AGCCCAGGAGCTCATAGTGCCTAT
 AC074117.12: CTGGCGTGGCAGAATACCTCTTTGA
 AC069513.3: GCAGCGCTCTTCCGACAGCCCCA
 RNY1P11: AAAGGGAGTGAGTTATCTCATTTGA
 RP11-162A23.5: CCCACCTCCCCTACCAAAGCCATA
 MYCL2: GACCGCGACTCGTACCATCACTATT
 RNU105B: AGGTCACCTCTCTCCCCAGGCTGTGT

Finally, the ncRNA expression levels were detected by the GenePix microarray scanner (Axon Instrument, Union City, CA, USA), and a diagnostic score was calculated for each sample according to the signature formula.

2.5. Statistical analysis

All statistical analyses were conducted using R 3.6.0 software (The R Foundation, MA, USA). The generalized linear model was constructed with glmnet 2.0–18 package, and ROC curve analysis was performed with ROCR 1.0–7 package. A 2-sided P value $< .05$ was considered statistically significant.

3. Results

3.1. Characteristic of included dataset

The included microarray dataset of GSE18123 was based on the platform of GPL6244 (Affymetrix Human Gene 1.0 ST Array [HuGene-1_0-st]), with a total of 104 ASD (80 men [76.9%] and average age 8.1 years [2–21]) and 82 controls (48 men [58.3%] and average age 8.0 years [2–22]). The subtypes of autistic disorder, Asperger disorder, and PDD-NOS accounted for 39.4% ($n=41$), 14.4% ($n=15$), and 46.2% ($n=48$) respectively. Then, 186 samples were randomly divided into the training set ($n=112$) and validation set ($n=74$).

3.2. Data preprocessing and sample clustering

The GPL6244 platform contained 861,493 sequences (25 bases) aligned to 33,297 probes. After probe re-annotation, a total of 8089 RNAs (32 types) were identified with 18,329 specific probes, among which there were 4143 ncRNAs mapped to 10,258 probes.

Table 1
Twenty non-coding RNAs in the blood diagnostic signature of autism spectrum disorder.

Gene symbol	Gene type	Gene name	Genomic location	Size (bases)	Probe ID	Ensembl ID
AQP4-AS1	Antisense	AQP4 Antisense RNA 1	chr18:26,655,742–27,190,698	534,957	8020717	ENSG00000260372
FTH1P2	Processed pseudogene	Ferritin Heavy Chain 1 Pseudogene 2	chr1:228,687,415–228,687,879	465	7910385	ENSG00000234975
FTH1P3	Processed pseudogene	Ferritin Heavy Chain 1 Pseudogene 3	chr2:27,392,623–27,393,576	954	8051133	ENSG00000213453
GMCL2	Processed pseudogene	Germ Cell-Less 2, Spermatogenesis Associated	chr5:178,184,503–178,187,432	2,930	8116174	ENSG00000244234
HMG2P11	Processed pseudogene	HMG2 Pseudogene 11	chr7:84,876,484–84,877,824	1,341	8133900	ENSG00000232605
IGHV3-47	IG-V pseudogene	Immunoglobulin Heavy Variable 3-47 Pseudogene	chr14:106,518,579–106,519,034	456	7981726	ENSG00000229092
MUC20-OT1	lncRNA	MUC20 Overlapping Transcript	chr3:195,658,062–195,739,964	81,903	8084917	ENSG00000242086
MYCLP1	Processed pseudogene	MYCL Pseudogene 1	chrX:107,272,464–107,276,559	4,096	8169231	ENSG00000204053
POLR2KP2	Processed pseudogene	RNA Polymerase II Subunit K Pseudogene 2	chr13:48,133,956–48,134,553	598	7971561	ENSG00000224510
RN7SL132P	miscRNA	RNA, 7SL, Cytoplasmic 132, Pseudogene	chr7:34,753,424–34,753,719	296	8132286	ENSG00000242653
RNA5SP160	rRNA	RNA, 5S Ribosomal Pseudogene 160	chr4:40,990,154–40,990,273	120	8094772	ENSG00000201736
RNU105B	snoRNA	RNA, U105B Small Nucleolar	chr20:8,831,185–8,831,393	209	8060895	ENSG00000201348
RNU1-16P	snRNA	U1 Small Nuclear 16	chr13:113,478,915–113,479,078	164	7972921	ENSG00000202347
RNU6-258P	snRNA	U6 Small Nuclear 258	chr17:20,126,388–20,126,488	101	8013448	ENSG00000212186
RNU6-485P	snRNA	U6 Small Nuclear 485	chr12:7,118,785–7,118,891	107	7953620	ENSG00000200345
RNU6-549P	snRNA	U6 Small Nuclear 549	chr15:64,671,263–64,671,369	107	7984215	ENSG00000207162
RNVU1-15	snRNA	Variant U1 Small Nuclear 15	chr1:144,412,576–144,412,740	165	7919556; 7919560	ENSG00000207205
RNY1P11	miscRNA	RNY1 Pseudogene 11	chr7:129,164,849–129,164,961	113	8136078	ENSG00000200629
RPS26P39	Processed pseudogene	Ribosomal Protein S26 Pseudogene 39	chr10:123,171,458–123,171,898	441	7936833	ENSG00000227586
TUBB2BP1	Unprocessed pseudogene	Tubulin Beta 2B Class IIb Pseudogene 1	chr6:3,177,039–3,179,764	2,726	8116651	ENSG00000216819

Then, samples were clustered according to the distance in Pearson correlation matrices. When adopted the expression profiles of probes or genes, no outliers were detected (height < 0.2).

3.3. Signature construction, evaluation, and validation

One hundred eighty six blood samples in the microarray dataset were randomly divided into the training set (n=112) and validation set (n=72) according to the ratio of 6:4. In the training set, the LASSO penalized generalized linear model identified 20 significant ncRNAs, which demonstrated an obvious discrepancy between ASD and control samples (Table 1). A diagnostic score was calculated for each sample according to the ncRNA expression levels weighted by their coefficients in the LASSO model.

$$\text{Diagnostic score} = \text{AQP4-AS1} * (-0.169) + \text{FTH1P2} * (-0.070) + \text{FTH1P3} * 0.261 + \text{GMCL2} * (-0.097) + \text{HMG2P11} * (-0.212) + \text{IGHV3-47} * (-0.024) + \text{MUC20-OT1} * (-0.143) + \text{MYCLP1} * (-0.105) + \text{POLR2KP2} * (-0.272) + \text{RN7SL132P} * 0.044 + \text{RNA5SP160} * (-0.184) + \text{RNU105B} * 0.200 + \text{RNU1-16P} * 0.040 + \text{RNU6-258P} * 0.080 + \text{RNU6-485P} * (-0.078) + \text{RNU6-549P} * 0.119 + \text{RNVU1-15} * 0.114 + \text{RNY1P11} * (-0.160) + \text{RPS26P39} * (-0.083) + \text{TUBB2BP1} * (-0.075)$$

The score displayed an excellent diagnostic ability for ASD in the training set (AUC=0.96), validation set (AUC=0.97), and the overall (AUC=0.96) (Fig. 1). No significant difference was detected between the women and men in the training set (AUC: 0.96 vs 0.97, P=.890), validation set (AUC: 0.96 vs 1.00, P=.205), and the overall (AUC: 0.95 vs 0.98, P=.231) (Fig. 2). It was also insignificant between the older and younger cases in the training set (AUC: 0.93 vs 0.99, P=.088), validation set (AUC:

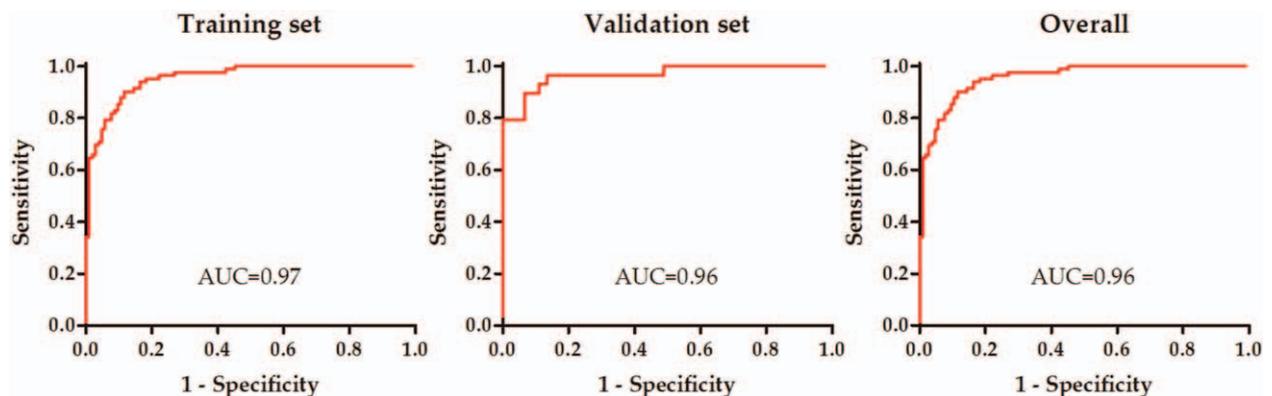


Figure 1. Receiver operating characteristic (ROC) curve analysis of the diagnostic signature. AUC=area under ROC curve.

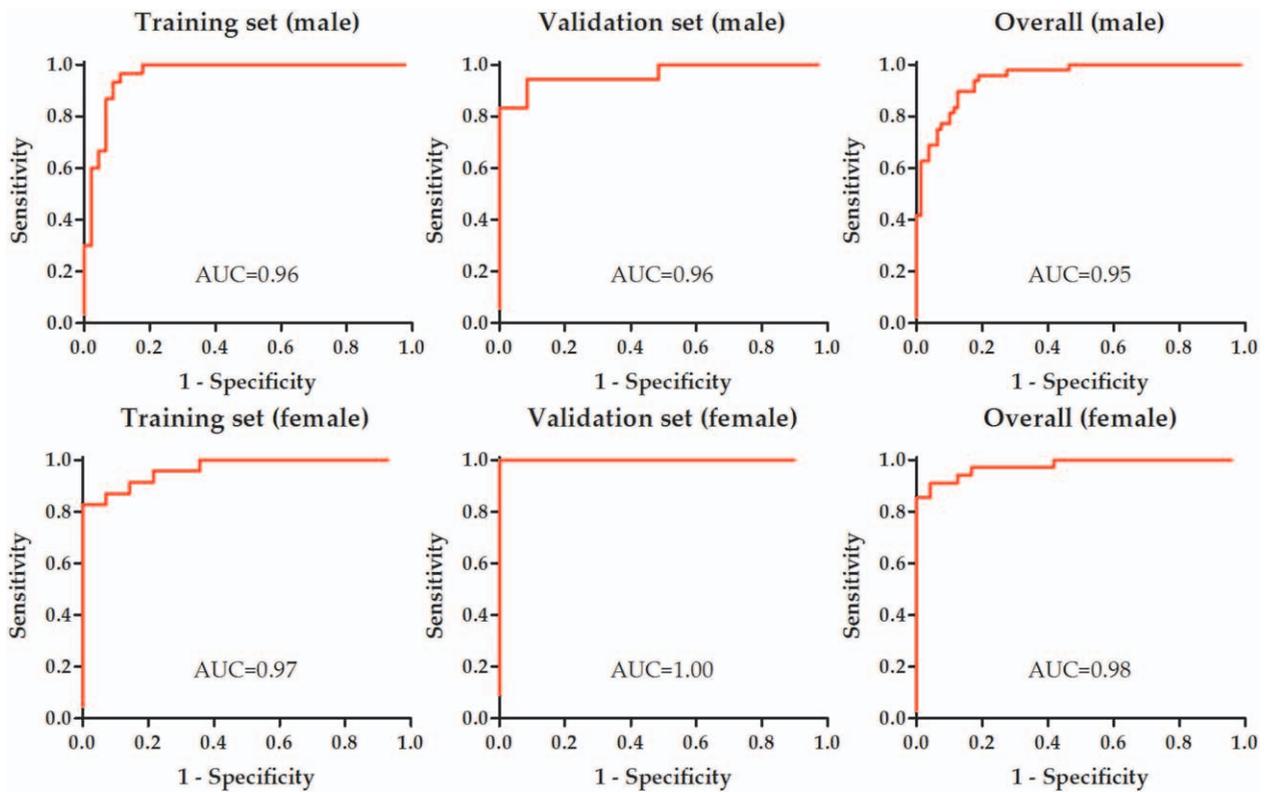


Figure 2. Receiver operating characteristic (ROC) curve analysis of the diagnostic signature in the subgroups of different sexes. AUC=area under ROC curve.

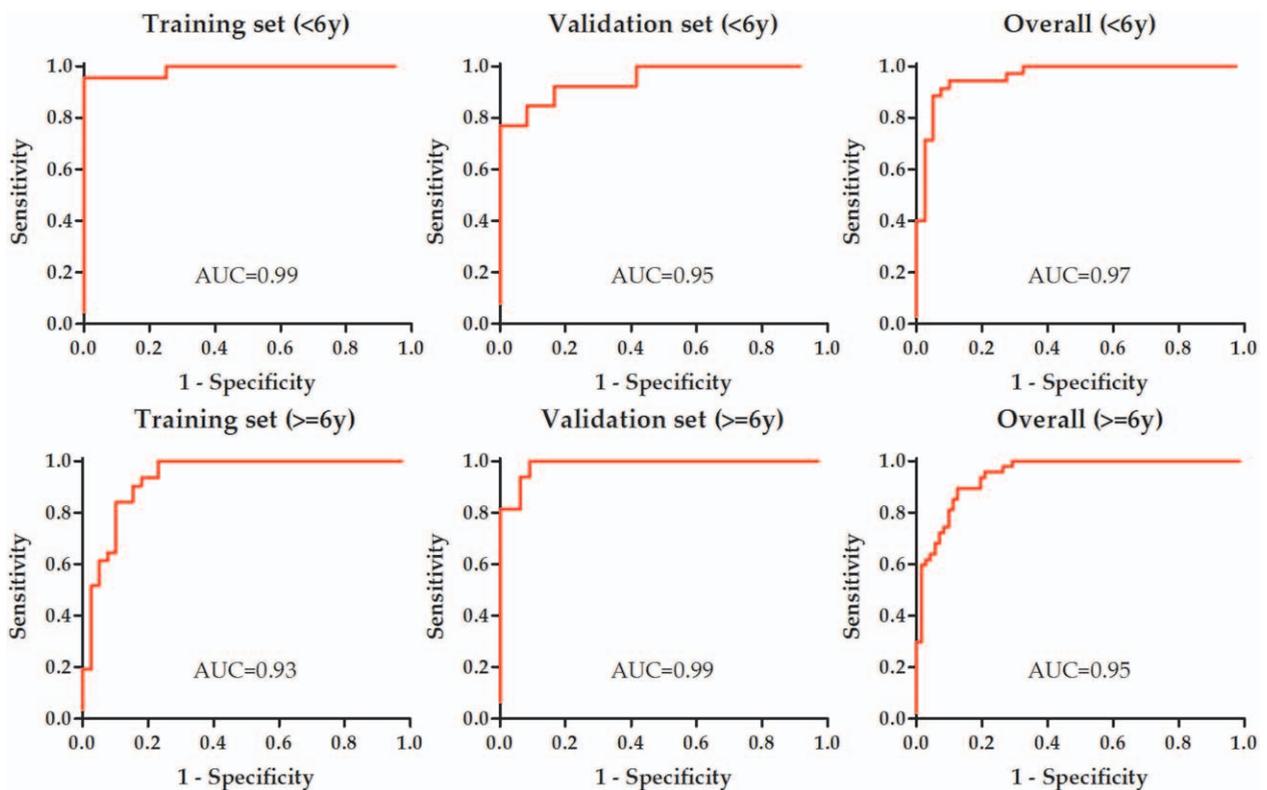


Figure 3. Receiver operating characteristic (ROC) curve analysis of the diagnostic signature in the subgroups of different ages. AUC=area under ROC curve.

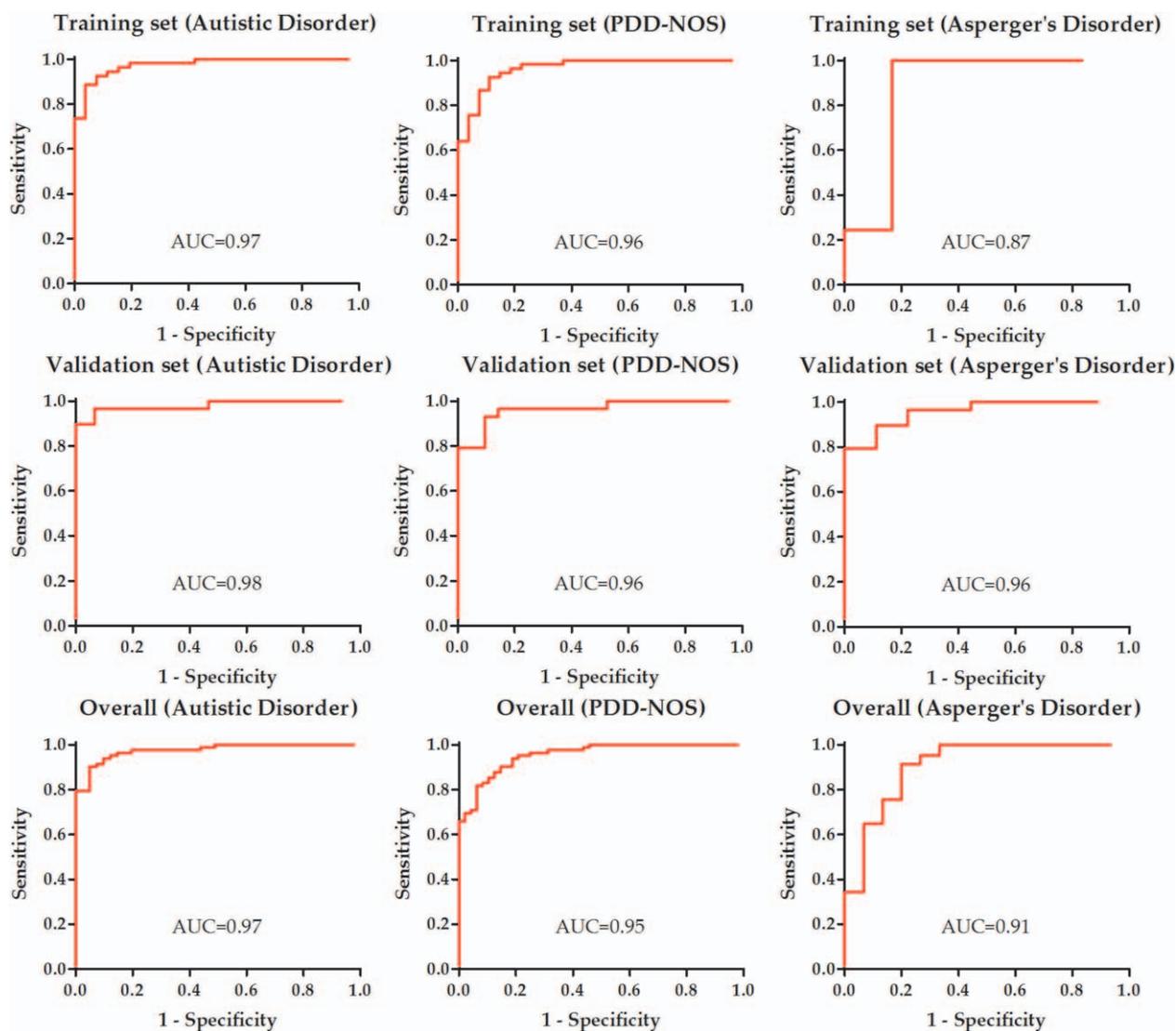


Figure 4. Receiver operating characteristic (ROC) curve analysis of the diagnostic signature in the subgroups of different disease subtypes. AUC = area under ROC curve.

0.99 vs 0.95, $P = .366$, and the overall (AUC: 0.95 vs 0.97, $P = .337$) (Fig. 3).

As for disease subtypes, the score had a good power in diagnosing autistic disorder (AUC = 0.97 in training set, 0.98 in validation set, and 0.97 in the overall), PDD-NOS (AUC = 0.96 in training set, 0.96 in validation set, and 0.95 in the overall), and Asperger disorder (AUC = 0.87 in training set, 0.96 in validation set, and 0.91 in the overall) (Fig. 4). No significant difference was detected between the subtypes ($P > .05$).

3.4. External validation

The blood samples were collected from 23 ASD patients (14 men [60.9%] and average age 8.0 years [2–16]), and 23 age- and sex-matched controls. According to the signature formula, the score displayed a good diagnostic ability for ASD (AUC = 0.96) (Fig. 5).

3.5. Comparison with a 55-gene signature

The previously published 55-gene signature derived from the same dataset.^[9] In the training set, the 55-gene signature showed

no obvious difference than the ncRNA signature (AUC: 0.96 vs 0.96, $P = 1.000$) (Fig. 6). In the validation set, the 55-gene signature displayed a poorer performance than the ncRNA signature (AUC: 0.71 vs 0.97, $P < .001$). In general, the ncRNA signature showed a better diagnostic ability than the 55-gene signature (AUC: 0.96 vs 0.68, $P < .001$).

4. Discussion

In this study, we adopted the methods of probe re-annotation and penalized generalized linear model to identify a novel and robust blood ncRNA signature in diagnosing ASD. The signature showed an excellent stability in the subgroup analyses of age, sex, and disease subtypes. It also displayed a higher diagnostic efficiency than the 55-gene signature.

The ncRNA signature consisted of 20 genes, covering 5 kinds of ncRNAs (lncRNA, pseudogene, miscRNA, rRNA, snoRNA, and snRNA). lncRNAs are untranslated RNA molecules with >200 nucleotides in length, which perform a wide variety of functions and play an important role in the development of

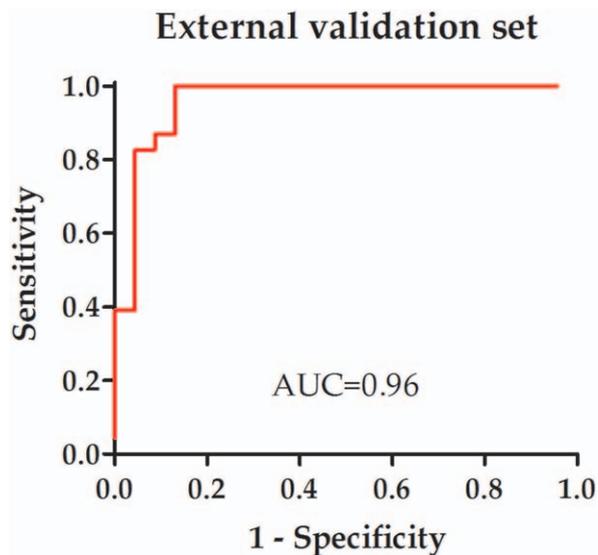


Figure 5. Receiver operating characteristic (ROC) curve analysis of the diagnostic signature in the external validation set. AUC=area under ROC curve.

mental discords.^[10] AQP4-AS1 had a moderate to high expression in nervous system (brain, cortex, and cerebellum) (based on the GTEx database). In GWAS Catalog, AQP4-AS1 was associated with the human phenotypes of blood protein measurement, lobe attachment, breast carcinoma, hair color, and glomerular filtration rate. MUC20-OT1 was involved into the phenotypes of mean corpuscular hemoglobin, mean corpuscular volume, mean corpuscular hemoglobin concentration, eosinophil count, iron biomarker measurement, and transferrin measurement.

Pseudogenes are gene copies that have lost the original coding ability, and they have been reported in the pathogenesis of ASD.^[11] IGHV3-47 was associated with the phenotype of blood protein measurement. FTH1P3 was involved into the progression and prognosis of multiple cancers.^[12,13] GMCL2 was related with protein binding (GO:0005515). MYCLP1 was associated

with the phenotype of alcoholic pancreatitis. POLR2K2 had a moderate to high expression in nervous system (brain, cortex, and cerebellum). RPS26P39 was associated with the phenotypes of sleep duration, age at menopause, and neutropenia, response to gemcitabine, pancreatic carcinoma. TUBB2BP1 had a moderate to high expression in nervous system (brain, cortex, and cerebellum), which was related with the phenotypes of red blood cell distribution width, mean corpuscular hemoglobin, and Crohn disease.

Small ncRNAs are regarded as new class of biomarkers and potential therapeutic targets in neurodegenerative diseases.^[14] In the signature, snRNAs are most frequent. As the components of spliceosome, snRNAs are fairly conserved with a uridine-rich non-coding sequence of <200nt, and involved into the splicing of precursor mRNA. RNU1-16P had a moderate to high expression in nervous system (brain, cortex, and cerebellum) and whole blood. RNU6-258P was related with intelligence.^[15] Compared with normally developed children, the intelligence development in ASD children was significantly delayed.^[16,17] There were positive correlations between age and mean corpuscular volume, and red cell distribution width in ASD children, and 24.1% cases had iron deficiency and 15.5% had anemia.^[18] RNU6-549P was associated with the GWAS phenotypes of mathematical ability, self reported educational attainment, coronary artery disease, and factor VII activating protease measurement. Previous studies suggested impaired metacognitive monitoring, mathematics under-achievement, and educational needs in ASD.^[19–21] RNVU1-15 had a moderate to high expression in nervous system and whole blood. It was associated with U1 snRNP (GO:0005685), mRNA 5'-splice site recognition (GO:0000395), pre-mRNA 5'-splice site binding (GO:0030627). A growing number of alternative splicing regulators have been reported in relation with ASD.^[22,23]

The limitations in this study should be also acknowledged. First, the sample size is not as large as we expected. Second, the method of probe re-annotation could not cover all ncRNAs. In the future, a large-scale prospective designed study was needed to validate this ncRNA signature.

In conclusion, through probe re-annotation and penalized generalized linear model, we identified a novel and robust blood ncRNA signature in diagnosing ASD, which might help improve the diagnostic accuracy for ASD in clinical practice.

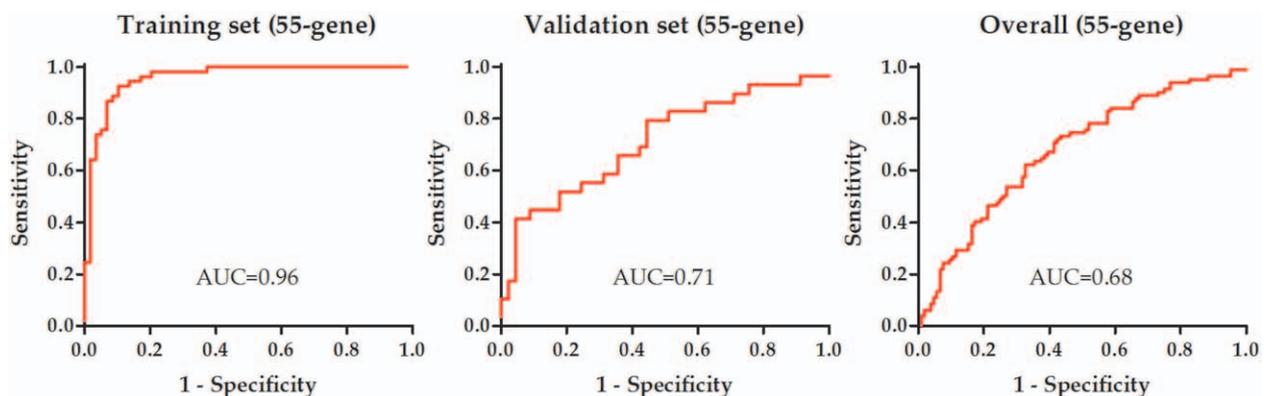


Figure 6. Receiver operating characteristic (ROC) curve analysis of the 55-gene diagnostic signature. AUC=area under ROC curve.

Author contributions

Conceptualization: Wei Cheng.

Data curation: Wei Cheng, Jinxia Zhou.

Formal analysis: Wei Cheng.

Methodology: Jinxia Zhou.

Software: Shanhu Zhou.

Supervision: Shanhu Zhou, Jinxia Zhou.

Visualization: Shanhu Zhou.

Writing – original draft: Wei Cheng.

References

- [1] Simashkova NV, Boksha IS, Klyushnik TP, et al. Diagnosis and management of autism spectrum disorders in Russia: clinical-biological approaches. *J Autism Dev Disord* 2019;49:3906–14.
- [2] Christensen DL, Baio J, Van Naarden Braun K, et al. Centers for Disease C, Prevention Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. *MMWR Surveill Summ* 2016;65:1–23.
- [3] Warren Z, McPheeters ML, Sathe N, et al. A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics* 2011; 127:e1303–11.
- [4] Sayad A, Omrani MD, Fallah H, et al. Aberrant expression of long non-coding RNAs in peripheral blood of autistic patients. *J Mol Neurosci* 2019;67:276–81.
- [5] Yu D, Jiao X, Cao T, et al. Serum miRNA expression profiling reveals miR-486-3p may play a significant role in the development of autism by targeting ARID1B. *Neuroreport* 2018;29:1431–6.
- [6] Han LO, Li XY, Cao MM, et al. Development and validation of an individualized diagnostic signature in thyroid cancer. *Cancer Med* 2018;7:1135–40.
- [7] Zhou J, Hu Q, Wang X, et al. Development and validation of a novel and robust blood small nuclear RNA signature in diagnosing autism spectrum disorder. *Medicine (Baltimore)* 2019;98:e17858.
- [8] Mhawech-Fauceglia P, Izevbaye I, Spindler T, et al. Genomic heterogeneity in peritoneal implants: a differential analysis of gene expression using nanostring Human Cancer Reference panel identifies a malignant signature. *Gynecol Oncol* 2020;156:6–12.
- [9] Kong SW, Collins CD, Shimizu-Motohashi Y, et al. Characteristics and predictive value of blood transcriptome signature in males with autism spectrum disorders. *PLoS One* 2012;7:e49475.
- [10] Tang J, Yu Y, Yang W. Long noncoding RNA and its contribution to autism spectrum disorders. *CNS Neurosci Ther* 2017;23:645–56.
- [11] Cappuccio G, Attanasio S, Alagia M, et al. Microdeletion of pseudogene chr14.232.a affects LRFN5 expression in cells of a patient with autism spectrum disorder. *Eur J Hum Genet* 2019;27:1475–80.
- [12] Yuan H, Jiang H, Wang Y, et al. Increased expression of lncRNA FTH1P3 predicts a poor prognosis and promotes aggressive phenotypes of laryngeal squamous cell carcinoma. *Biosci Rep* 2019;39: BSR20181644.
- [13] Liu M, Gao X, Liu CL. Increased expression of lncRNA FTH1P3 promotes oral squamous cell carcinoma cells migration and invasion by enhancing PI3K/Akt/GSK3b/Wnt/beta-catenin signaling. *Eur Rev Med Pharmacol Sci* 2018;22:8306–14.
- [14] Watson CN, Belli A, Di Pietro V. Small non-coding RNAs: new class of biomarkers and potential therapeutic targets in neurodegenerative disease. *Front Genet* 2019;10:364.
- [15] MacArthur J, Bowler E, Cerezo M, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* 2017;45:D896–901.
- [16] Bedford SA, Park MTM, Devenyi GA, et al. Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and cortical morphometry in autism spectrum disorder. *Mol Psychiatry* 2019.
- [17] Dryburgh E, McKenna S, Reikik I. Predicting full-scale and verbal intelligence scores from functional Connectomic data in individuals with autism Spectrum disorder. *Brain Imaging Behav* 2019.
- [18] Hergüner S, Keleşoğlu FM, Tanıdır C, et al. Ferritin and iron levels in children with autistic disorder. *Eur J Pediatr* 2012;171:143–6.
- [19] Maras K, Gamble T, Brosnan M. Supporting metacognitive monitoring in mathematics learning for young people with autism spectrum disorder: a classroom-based study. *Autism* 2019;23:60–70.
- [20] Hetzroni O, Agada H, Leikin M. Creativity in autism: an examination of general and mathematical creative thinking among children with autism spectrum disorder and children with typical development. *J Autism Dev Disord* 2019;49:3833–44.
- [21] Sagers B, Tones M, Dunne J, et al. Promoting a collective voice from parents, educators and allied health professionals on the educational needs of students on the autism spectrum. *J Autism Dev Disord* 2019;49:3845–65.
- [22] Singh RK, Cooper TA. Pre-mRNA splicing in disease and therapeutics. *Trends Mol Med* 2012;18:472–82.
- [23] Gonatopoulos-Pournatzis T, Wu M, Braunschweig U, et al. Genome-wide CRISPR-Cas9 interrogation of splicing networks reveals a mechanism for recognition of autism-misregulated neuronal microexons. *510.e12-L 524.e12 Mol Cell* 2018;72.