Identifying Genes Associated With Autism Spectrum Disorders by Random Walk Method With Significance Tests

YAN ZHAO[ID]1, PING ZHAO2, HAIYAN LIANG3, AND XIAOLIN ZHANG3

[1]Key Laboratory of Cardiovascular Proteomics of Shandong Province, Department of Geriatrics, Qilu Hospital of Shandong University, Jinan 250012, China
[2]Department of Rehabilitation, Changqing District Hospital of Traditional Chinese Medicine, Jinan 250399, China
[3]College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China

Corresponding author: Yan Zhao (zyssly@163.com)

ABSTRACT Autism spectrum disorders (ASD) are generally defined as a development disorder typically characterized by social interaction and communication ailments and stereotyped actions due to combined genetic and environmental factors. Different critical aspects contribute to ASD, and consensus has been reached among autism researchers about its predominant genetic factors. However, the pathogenesis of ASD has not been fully revealed, and a systematic method must be developed to identify the genes related to this disease. Here, we predicted new ASD-associated genes by random walk method on the basis of prior-known ASD genes from AutDB. New genes such as RAC3, AC1 (ADCY1), PKC (PRKC) gamma, EPH receptor A5, WNT3A, calretinin, RAS-R, KLF4, and calpain 3 were found to play an irreplaceable role in ASD pathogenesis.

INDEX TERMS Autism spectrum disorders, disease gene, random network, database, protein–protein network.

I. INTRODUCTION

Autism spectrum disorder (ASD) is a widely reported developmental illness with typical symptoms of social interaction and communication disorders and restricted and repetitive behaviors and pathogenesis of combined genetic and environmental factors. According to a global report [1], approximately 24.8 million people suffered from such disease in 2015. Even in developed countries, more than 1.5% of child population is clinically diagnosed with ASD as early as 2017 [1], indicating that this ailment is one of the most mental threatening diseases for people worldwide [1].

ASD pathogenesis has been widely studied, and some critical factors for its initiation and progression have already been identified. This systematic disease involves multiple tissues, organs, and regulatory systems including neural (the brain), immune, and digestive systems [2], [3]. Among these, the neural system is the most important with typical pathological symptoms associated with ASD, and three of its regions are correlated with the pathogenesis of this illness. Cerebellum abnormalities participate in the pathogenesis of ASD.

Further studies on the communications between cerebellum and parietal lobes via brainstem confirmed that pathogenesis on such region may trigger motor dysfunction and dyspraxia, typical symptoms of ASD [3]. In addition to cerebellum, amygdala regions are also correlated with ASD pathogenesis. An abnormal metabolism pattern in such brain region is correlated with the progression of ASD, especially in the early initiation stage [4].

Some typical physical regions in the neural system are correlated with ASD pathogenesis. Genetic background is essential in studying the initiation and progression of ASD. Multiple key pathological factors of ASDs have been identified in multiple tissues at the genetic level. One work focused on the neural system, which was introduced with a regional genetic variation induced syndrome named as fragile X syndrome (FXS), the leading genetic pathogenesis for ASD experienced by more than 5% of patients [5]. In addition, patients with FXS share similar symptoms with ASD such as stereotyped behaviors [6], indicating that fragile X syndrome and its typical genetic variation patterns are correlated with ASD by regulating specific behaviors of this illness. In addition to the whole or regional genome abnormality at genetic levels, specific factors, such as the

The associate editor coordinating the review of this manuscript and approving it for publication was Quan Zou[ID].

functional molecule oxytocin, are pathologically correlated with ASDs. Oxytocin and its specific system regulate sculpting socio-sexual behaviors. Pathogenesis involving such system has been observed in ASD and is correlated with problems in multiple neural systems including language and social communication, sensory, and cognitive [7]. Genes such as *SHANK3* [8] and *PTCHD1* at X-chromosome [9] also participate in ASD pathogenesis via regulation of the neural system and contribute to certain typical symptoms of this disease. Other genes related to ASD, such as *ITGB3* [10], *RORA* [11], oxytocin-encoding genes, and arginine–vasopressin receptors [12], are also functionally correlated with the pathogenesis of such disease. The above genes and genetic abnormalities have been functionally correlated with the neural system, thereby validating the key role of genetic factors affecting neural system during ASD. In addition to the regulatory factors involving the neural system, an endogenous glycosylated vitamin D binding protein was found in macrophages, a key component of the focal immune system, and participates in the correction and repair of dysregulated tissues in pathological regions. Hence, this protein may also be abnormally expressed during ASD.

In summary, the comprehensive pathological mechanisms of ASD have not been fully revealed. Although some genes were found to play a role in the development of ASD, studies on genes that affect ASD are still lacking. Therefore, developing a systematic method to identify the genes related to ASD is necessary. Here, we predicted ASD-associated genes by random walk method based on knowledge on prior-known ASD genes from AutDB [13] and biological network from STRING database [14]. Several newly identified genes can be confirmed, thus validating the efficacy and accuracy of our prediction technique.

II. METHODS
A. DATA
A total of 990 ASD genes were identified from AutDB (http://autism.mindspec.org/autdb/) database with version in 2018 [13] listed in **Supplementary file S1**. AutDB manually curated these genes from literature, and each association has corresponding evidence in references. These genes were considered as seed genes to identify novel ASD-related genes on the STRING network [14], a comprehensive functional association network that integrates various databases, literatures, high throughput analysis, and functional experiments.

In 2020, a new version of AutDB updated the ASD gene number to 1141. As shown in **Figure 1**, 978 ASD genes were also included in the old version of AutDB, whereas 163 genes were newly added. These newly added ASD genes would be used for independent test, which are listed in **Supplementary file S2**.

B. RANDOM WALK WITH RESTART (RWR) ALGORITHM FOR DETERMINING ASD-ASSOCIATED GENES
Recently, network methods are more and more popular to investigate various biological and medical
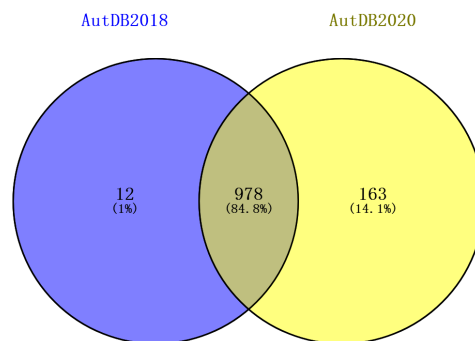
FIGURE 1. Venn diagram to show the common and different ASD genes in two versions of AutDB. Most genes are common in two versions. In the AutDB published in 2020, 163 new ASD genes are added compared with the old version published in 2018.

problems [15]–[22]. RWR algorithm, as a typical network-based feature ranking algorithm [23]–[27], can simulate a random walker starting from one or several seed nodes and walking on a network, thus updating the weight (probability) vector of network nodes in an iterative way.

In the mathematical description, P_i is the probability vector after the i_{th} updating procedure, and the next new probability vector P_{i+1} is updated as follows:

$$P_{i+1} = (1 - \lambda)AP_i + \lambda P_0 \tag{1}$$

where A represents the column-wise normalized adjacency matrix of the given network (e.g., PPI network), and λ represents the probability of returning to seed nodes (e.g., restart action of random walk) and indicates the importance of seed nodes, and P_0 is the initial probability vector. The RWR algorithm stops when the probability vector becomes convergent, and output P_{i+1} as the final result. Each element in this final vector indicates the probability of the corresponding nodes associated with seed nodes (e.g., as ASD-associated genes).

A total of 990 prior-known ASD genes were picked up as seed nodes in the RWR algorithm. An initialization vector P_0 was constructed and consisted of 19,247 elements, which correspond to the 990 ASD-associated genes set to 1/990. The other elements were set to zero, and the probability of returning to seed nodes λ was set as 0.8. The algorithm termination rule is $\|P_{i+1} - P_i\|_1 < 10^{-6}$, and the genes assigned with probabilities higher than a threshold as 10^{-5} were picked up as candidates for further investigations.

C. SIGNIFICANCE TESTS FOR SELECTING ASD-ASSOCIATED GENES
The RWR algorithm can detect candidate ASD-associated genes on the basis of available ASD genes. Owing to the high dependence on the structure of PPI network, RWR algorithm might select some nodes (genes) with special positions in PPI network to introduce some false-positive candidate genes. Increasing the reliability of selected candidate genes on the basis of their important functions is also necessary. Therefore, three significance tests, namely, permutation test, association

test, and function test were further applied to rank and select ASD-associated genes with high confidence.

1) PERMUTATION TEST

The permutation test aims to evaluate the probability values produced by RWR algorithm. m Ensembl ID sets with given size were randomly produced, and RWR algorithm was conducted on the PPI network with this set members as seed nodes to ensure that the tested ASD-associated genes from RWR can receive their probability values in a test. After all m sets were tested, each tested ASD-associated gene should have one actual probability value and random m ones. The P value of permutation test can be calculated as follows:

$$Pv(g) = \frac{\Omega}{m}, \qquad (2)$$

where Ω represented the number of randomly produced sets on which candidate gene g received higher probability values than its actual one. A candidate gene with a small P value is likely to be associated with ASD because it has less association with randomly produced sets.

2) ASSOCIATION TEST

Among the candidate ASD-associated genes passing the permutation test, those with strong associations with prior-known ASD genes were selected. In the PPI network, the interacting proteins always share similar functions; hence, each candidate gene g was linked to ASD genes, and the MIS was measured as follows:

$$MIS(g) = \max\{I(g, g')|g' \text{is a ASD gene}\} \qquad (3)$$

where $I(g, g')$ indicates the confidence score of the interaction between g and g'. Candidate genes with high MISs are likely to be actual ASD-associated genes.

3) FUNCTION TEST

The validated ASD-associated genes must be highly related to biological processes. The gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways are two popular sources of biological processes. The enrichment theory was employed to quantify the relationship between one gene and GO terms or KEGG pathways under the following assumption: the important candidate ASD-associated genes should exhibit similar relationships with at least one ASD genes on the GO terms and KEGG pathways. For one gene g and one GO term or KEGG pathway T, their relationship defined by enrichment theory was calculated as follows:

$$S(g, T) = -\log_{10}(\sum_{k=m}^{n} \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}), \qquad (4)$$

where N is the total number of genes/proteins in humans, M is the number of genes/proteins annotated to T, n is the number of genes/proteins in H that consist of g and its direct neighbors in the PPI network reported in STRING, and m is the number of genes/proteins that are in H and annotated to T. Our in-house program using R function phyper was developed to calculate this S value. Accordingly, each gene can be encoded into a vector by collecting all outcomes of this function. For gene g, the obtained vector was formulated as $E(g)$. The relationship between two genes g and g' on their GO terms and KEGG pathways can be measured by the following:

$$Q(g, g') = \frac{E(g) \cdot E(g')}{\|E(g)\| \cdot \|E(g')\|} \qquad (5)$$

Similar to MIS, another measurement, MFS, can be calculated for each gene g as follows:

$$MFS(g) = \max\{Q(g, g')|g' \text{is a ASD gene}\} \qquad (6)$$

The candidate ASD-associated genes with high MFS, which is important for biological function, must be selected.

III. RESULTS AND DISCUSSION

In this study, based on the RWR algorithm and three significance tests, we designed a computational method to discover novel ASD-associated genes. The entire procedures are illustrated in Figure 2.

A. ASD-ASSOCIATED GENES

Among the 990 ASD genes from AutDB database with 2018 version, RWR analysis assigned weights or probabilities for all genes on the network, and those with probability larger than 10^{-5} remained as initial candidates. Three significance tests helped in extracting important genes and controlling false-positive genes. RWR candidates with P-value less than 0.05, MIS larger than 900 (the cutoff of the highest confidence score in STRING), and MFS larger than 0.97 were further selected and are listed in **Supplementary file S3**. Finally, 147 genes were determined as ASD-associated genes. Some representative genes and their quality measurements are listed in **Table 1,** and the network interactions consisting of all 147 genes are displayed in **Figure 3**. It can be observed that most of these genes had strong associations with each other, indicating that they may share similar functions. Thus, if one gene can be confirmed to be ASD, others may also be related to ASD. Functional enrichment analysis indicated that these predicted ASD-associated genes participate in many disease-relevant functions (**Table 2** and **Supplementary file S4**). As mentioned in Section II.A, 163 newly added ASD genes were picked up for the independent test. A Venn diagram for the predicted ASD genes and newly added genes was plotted, as shown in **Figure 4**. Clearly, three predicted genes (*PRKCA*, *ACTL6B*, and *EGR3*) were included in the recently updated AutDB in 2020, indicating that our computational method can really discover novel ASD-associated genes.

B. FUNCTIONAL RELEVANCE OF ASD-ASSOCIATED GENES

A total of 147 genes are functionally correlated with ASD as confirmed by recent publications, thus validating the efficacy
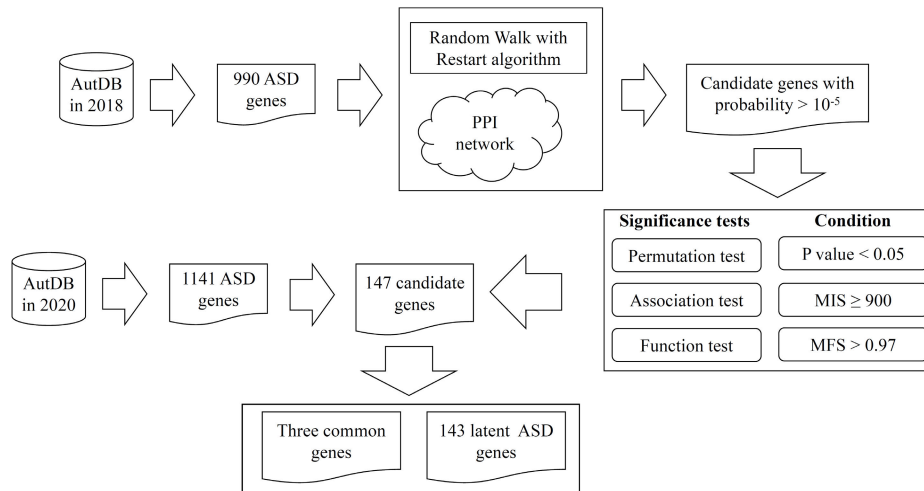
FIGURE 2. Entire procedures to Identify genes associated with autism spectrum disorders.

TABLE 1. Representative predicted candidate ASD genes.

Ensembl ID	Name	Probability	P-value	MIS	MFS
ENSP00000304283	RAC3	1.14E-04	<0.001	998	0.9984
ENSP00000297323	AC1	4.54E-05	0.04	966	0.9968
ENSP00000263431	PKCG	4.78E-05	<0.001	928	0.9968
ENSP00000273854	EPHA5	3.72E-05	<0.001	902	0.9963
ENSP00000284523	WNT3A	4.59E-05	0.002	998	0.9935
ENSP00000307508	CALB2	5.64E-05	<0.001	943	0.9934
ENSP00000246792	RRAS	4.81E-05	0.043	997	0.9933
ENSP00000363804	KLF4	4.22E-05	0.004	997	0.9910
ENSP00000380349	CAPN3	4.37E-05	0.015	998	0.9926

TABLE 2. Enriched pathways from KEGG database.

Pathway name	Pathway id	P-value
Pathways in cancer	KEGG:05200	2.00E-18
Axon guidance	KEGG:04360	1.55E-17
Proteoglycans in cancer	KEGG:05205	2.26E-17
Melanogenesis	KEGG:04916	2.69E-17
Rap1 signaling pathway	KEGG:04015	8.17E-16
mTOR signaling pathway	KEGG:04150	8.50E-16
Signaling pathways regulating pluripotency of stem cells	KEGG:04550	2.05E-15
Breast cancer	KEGG:05224	6.06E-15
Ras signaling pathway	KEGG:04014	9.81E-14
Hepatocellular carcinoma	KEGG:05225	1.43E-10
Gastric cancer	KEGG:05226	2.05E-10
Human cytomegalovirus infection	KEGG:05163	3.97E-10
Wnt signaling pathway	KEGG:04310	7.90E-10
VEGF signaling pathway	KEGG:04370	2.46E-09
Cushing syndrome	KEGG:04934	3.48E-09
Cholinergic synapse	KEGG:04725	3.98E-09
Apelin signaling pathway	KEGG:04371	6.59E-08
EGFR tyrosine kinase inhibitor resistance	KEGG:01521	8.87E-08
Chemokine signaling pathway	KEGG:04062	9.35E-08
Basal cell carcinoma	KEGG:05217	9.82E-08

and accuracy of our prediction results. Owing to the limitation of the manuscript's length, only the representative genes were selected for further detailed discussion.

RAC3 is the first gene in our prediction list. This gene encodes a GTPase of the RAS superfamily and regulates diverse essential biological processes including cell growth, cytoskeletal reorganization, and protein kinase activation [28], especially in the neural system as the major pathological region of ASD. With regard to the correlations between *RAC3* and ASD, this predicted gene contributes to the regulation of neuronal development by interacting with a homologue named *RAC1* [29], [30]. Considering that the
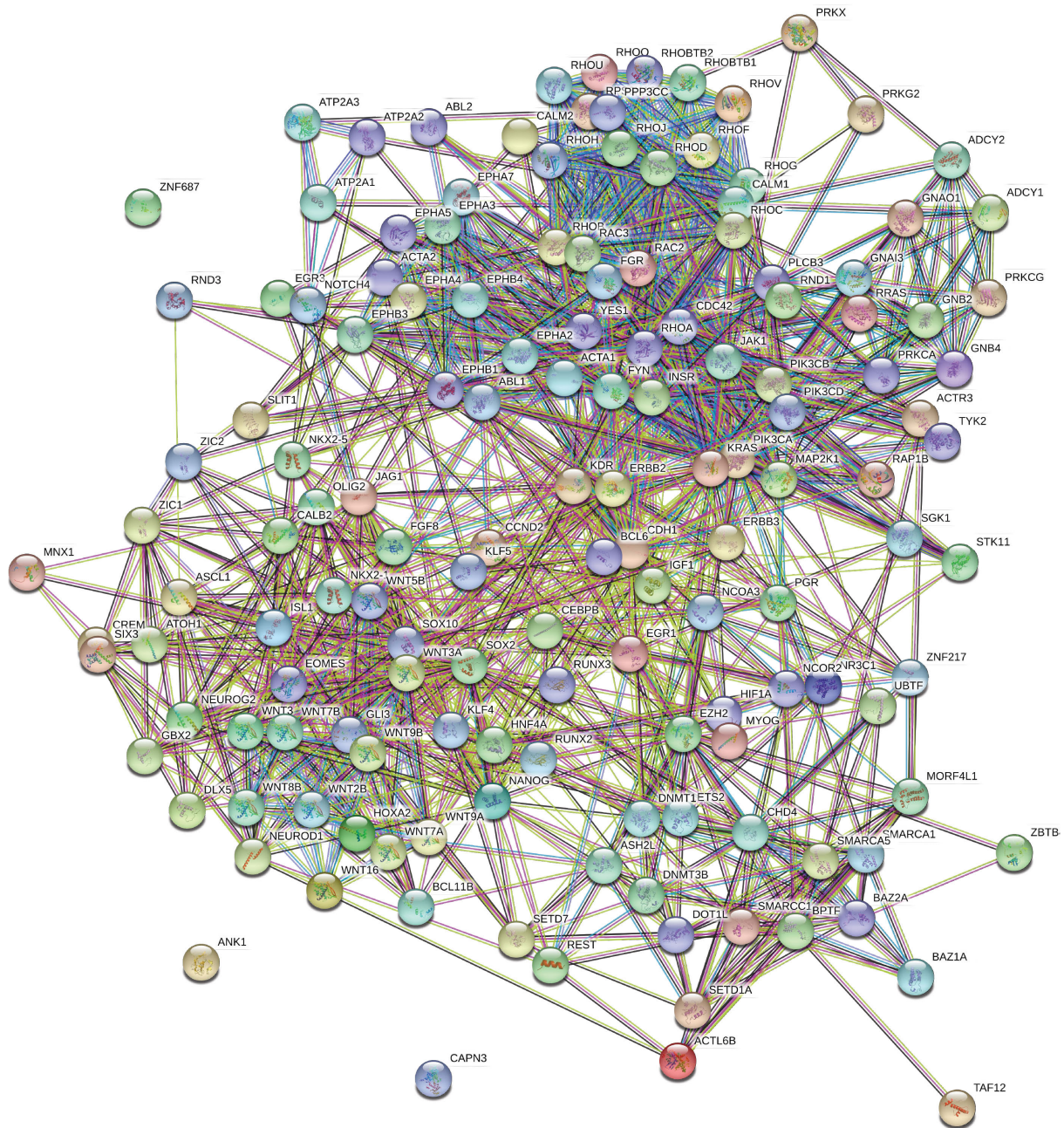
FIGURE 3. STRING network of predicted candidate ASD genes. Most of these genes have strong associations with each other, suggesting that if one gene can be confirmed to be ASD genes, others may also be related to ASD.

abnormal development of neural system is one of the major pathogenesis of ASD [31], [32], predicting such gene as a candidate pathogenic factor for ASD is highly reasonable. This finding validates the efficacy and accuracy of our prediction.

In addition to *RAC3*, the next predicted gene is *AC1*(*ADCY1*) encoding a functional adenylate cyclase a group of enzymes that regulates the essential functions of living cells [33]. With regard to the correlations between adenylate cyclase and ASD, reduced sociability as a typical symptom for ASD is correlated with AC1 at the transcrip-

tomic regulatory level [33]. Further studies on the whole genome level also identified typical genetic variants located inside the gene body of adenylate cyclase 1 (*AC1*) from an ASD family, indicating that such gene may also participate in ASD pathogenesis at the genomic level [34]. In summary, gene *AC1* from our optimal prediction list is definitely an ASD associated gene, thus validating the efficacy and accuracy of our prediction.

As another typical gene, PKC gamma (*PKCG*, also known as *PRKCG*) is involved in multiple cellular signaling pathways in multiple biological systems [35], [36], including
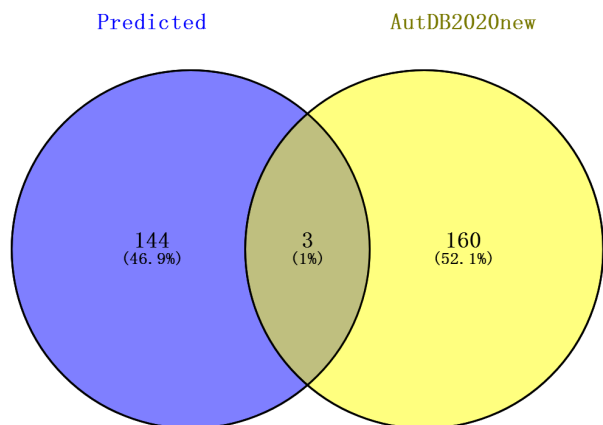
FIGURE 4. Venn diagram to show the relationship between predicted ASD genes and newly added ASD genes in AutDB.

the neural system. ASD regression can be attributed to pathological changes in multiple G-protein-coupled signaling pathways including *PKCG*-associated pathways functioning in the central neural system [37]. Protein kinases such as the candidate gene *PKCG* are the key regulators in such G-protein coupled signaling pathways. Hence, *PKCG* can be identified as a potential regulator for ASD.

The next predicted gene is EPH receptor A5 (*EPHA5*), an ephrin receptor that participates in the development of neural system by interfering with neuron migration and axon targeting [38], [39]. Considering that ASD can also be partially attributed to the abnormal development and dysfunction of the neural system, the candidate gene *EPHA5* may also participate in ASD pathogenesis.

The next identified gene is *WNT3A*, a famous member of the *Wnt* signaling pathway [40] and is correlated with

multiple proliferative biological functions under either physical or pathological conditions. With regard to the internal correlations among the predicted genes *WNT3A* and ASD, the abnormal alterations at the genomic levels, such as the microdeletions and micro duplications of *WNT3A* and related genes from the *Wnt* signaling pathway, have been widely identified in patients with ASD [41]. This finding indicates the potential regulatory role of these genes for ASD. Early brain development [42], [43] and synaptic functioning [44] are also functionally connected to canonical *Wnt* signaling pathway, thus further confirming our prediction of *WNT3A* as a candidate ASD-associated gene.

The next gene is ***CALB2***, which encodes an important protein named as calretinin [45]. Identified in the neural system, *CALB2* regulates the calcium related metabolism in multiple neuronal and endocrine cells, specifically in the pathogenic region of ASD, the cerebellum. A long-standing interest has been directed for its possible role in the hippocampus of patients with ASD and a selective increase in the density of *CALB2*-immunoreactive interneurons in the dentate [46]. This finding suggests that *CALB2* regulates ASD through its increased expression. And an aberrant splicing variant and non-synonymous SNPs on Ca2+ - dependent activator protein have been identified in patients with ASD [47]. These studies implicate calretinin to function in the neuropathology and behavioral characteristics of patients with developmental disorders. In conclusion, *CALB2* can be a critical gene in the ASD development, thereby corresponding with our prediction on its specific role in ASD.

According to our prediction list, ***RRAS*** encoding a specific Rap GTP binding protein named as RAS-related is an optimal gene correlated with ASD. RAS-related regulates neural progenitors [48]. With regard to the correlations
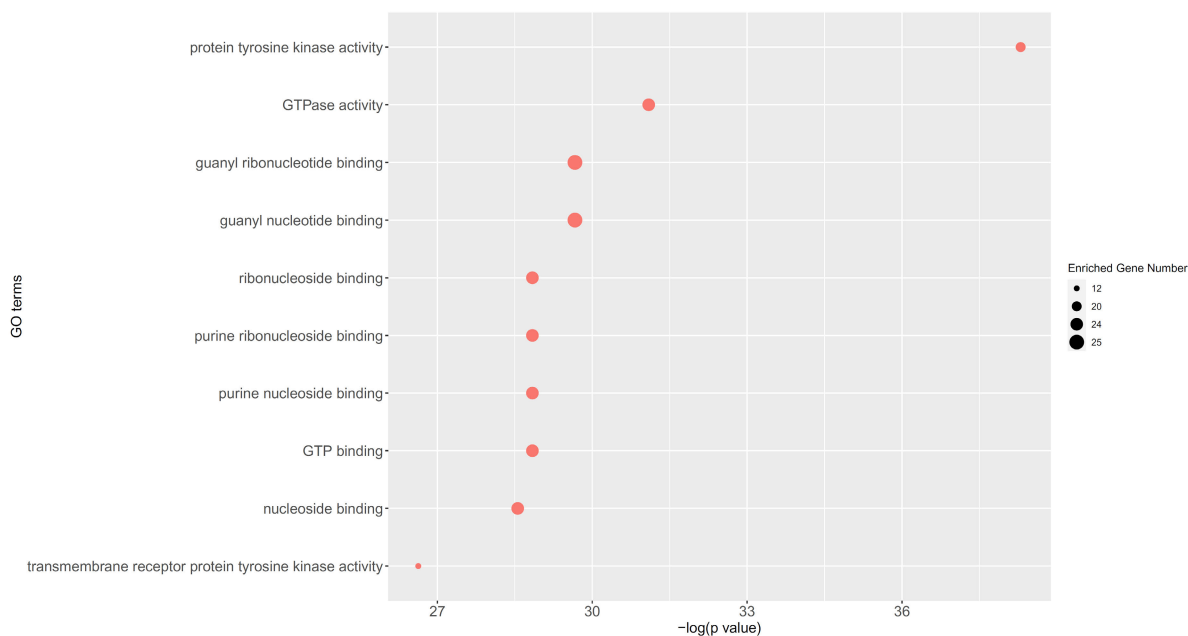
FIGURE 5. General functional distribution of the predicted ASD-associated genes.

between RAS-related and ASD, researchers from NY State Institute for Basic Research in Developmental Disabilities in 2013 reported that this candidate gene and its correlated pathway contribute to the regulation of neuronal cell migration and pathogenesis of ASD [49]. Combined with effective pathways such as RAS signaling pathway, *RRAS* suppresses the development and maturation of central neural system [49], a major system implicated in the pathogenesis for ASD [50]. This finding validated that our predicted gene is definitely an ASD-associated pathological factor.

KLF4 as another effective predicted gene and participates in the regulation of gene expression as a member of zinc finger DNA-binding proteins. With regard to the detailed correlations between this gene and ASD pathogenesis, its inhibition may further suppress the expression of cyclin B1 and D1 and has been functionally correlated with abnormal axon growth in the central neural system [51]. This finding indicated that kruppel-like factors 4 may pathologically regulate ASD by affecting axon regeneration.

Finally, *CAPN3* encodes a member of calpain family, a family of effective calcium-dependent, non-lysosomal cysteine proteases (proteolytic enzymes). Proteins from Calpain family are activated by brain-derived neurotrophic factor (BDNF) and further regulate memory formation [52]. Therefore, this predicted gene may also be critical for BDNF-associated learning and memory procedures. In 2013, researchers from University of Southern California confirmed that calpains including our predicted gene play an irreplaceable role in neural system development via PTEN signaling pathway [52]. Considering the internal correlations between ASD and neural system development, *CAPN3* was reasonably predicted as an ASD-associated gene.

To present the functional distribution pattern of our predicted genes, we also applied gene enrichment analyses using R package *clusterProfiler* to show the general functional distribution of the predicted ASD-associated genes (**Figure 5**).

IV. CONCLUSION

ASD-associated genes were determined based on prior-known ASD genes and biological network information by using the random walk method combined with three significance tests. Many new candidate genes associated with ASD were identified and found to be enriched in many disease-related functions and densely interactive on the biological network. Recent literature also supports the function or pathogen relevance of the predicted genes with ASD. Among which, three genes were confirmed as ASD-related in the updated autDB. Therefore, our newly presented computational approach may identify some potential biomarkers for ASD and provide basic research foundation for further studies on the detailed pathological mechanism of ASD.

SUPPLEMENTARY MATERIAL

Supplementary file S1: List of prior-known ASD genes in 2018; Supplementary file S2: List of prior-known ASD genes in 2020; Supplementary file S3: List of candidate ASD-associated genes predicted by RWR analysis; Supplementary file S4: Functional enrichment analysis for candidate ASD-associated genes; Figure S1 Overlap of different versions of autDB; Figure S2 Independent test with new version of autDB.

REFERENCES

[1] K. Lyall, L. Croen, J. Daniels, M. D. Fallin, C. Ladd-Acosta, B. K. Lee, B. Y. Park, N. W. Snyder, D. Schendel, H. Volk, G. C. Windham, and C. Newschaffer, "The changing epidemiology of autism spectrum disorders," *Annu. Rev. Public Health*, vol. 38, pp. 81–102, Mar. 20 2017.

[2] S. H. Fatemi et al., "Consensus paper: Pathological role of the cerebellum in autism," *Cerebellum*, vol. 11, no. 3, pp. 777–807, Sep. 2012.

[3] J. D. Schmahmann, D. L. Rosene, and D. N. Pandya, "Motor projections to the basis pontis in rhesus monkey," *J. Comparative Neurol.*, vol. 478, pp. 248–268, Oct. 2004.

[4] N. M. Kleinhans, T. Richards, K. E. Weaver, O. Liang, G. Dawson, and E. Aylward, "Brief report: Biochemical correlates of clinical impairment in high functioning autism and Asperger's disorder," *J. Autism Develop. Disorders*, vol. 39, pp. 1079–1086, Jul. 2009.

[5] L. M. McCary and J. E. Roberts, "Early identification of autism in fragile X syndrome: A review," *J. Intellectual Disab. Res.*, vol. 57, pp. 803–814, Sep. 2013.

[6] R. N. Hernandez, R. L. Feinberg, R. Vaurio, N. M. Passanante, R. E. Thompson, and W. E. Kaufmann, "Autism spectrum disorder in fragile X syndrome: A longitudinal evaluation," *Amer. J. Med. Genet. A*, vol. 149A, pp. 1125–1137, Jun. 2009.

[7] E. Quattrocki and K. Friston, "Autism, oxytocin and interoception," *Neurosci. Biobehav. Rev.*, vol. 47, pp. 410–430, Nov. 2014.

[8] L. Soorya, A. Kolevzon, J. Zweifach, T. Lim, Y. Dobry, L. Schwartz, Y. Frank, A. Wang, G. Cai, E. Parkhomenko, D. Halpern, D. Grodberg, B. Angarita, J. P. Willner, A. Yang, R. Canitano, W. Chaplin, C. Betancur, and J. D. Buxbaum, "Prospective investigation of autism and genotype-phenotype correlations in 22q13 deletion syndrome and SHANK3 deficiency," *Mol. Autism*, vol. 4, no. 1, p. 18, Jun. 2013.

[9] A. Noor et al., "Disruption at the PTCHD1 locus on Xp22.11 in autism spectrum disorder and intellectual disability," *Sci. Transl. Med.*, vol. 2, Sep. 2010, Art. no. 49ra68.

[10] V. Napolioni, F. Lombardi, R. Sacco, P. Curatolo, B. Manzi, R. Alessandrelli, R. Militerni, C. Bravaccio, C. Lenti, M. Saccani, C. Schneider, R. Melmed, T. Pascucci, S. Puglisi-Allegra, K.-L. Reichelt, F. Rousseau, P. Lewin, and A. M. Persico, "Family-based association study of ITGB3 in autism spectrum disorder and its endophenotypes," *Eur. J. Hum. Genet.*, vol. 19, pp. 353–359, Mar. 2011.

[11] T. Sarachana, M. Xu, R.-C. Wu, and V. W. Hu, "Sex hormones in autism: Androgens and estrogens differentially and reciprocally regulate RORA, a novel candidate gene for autism," *PLoS ONE*, vol. 6, no. 2, Feb. 2011, Art. no. e17116.

[12] I. Cataldo, A. Azhari, and G. Esposito, "A review of oxytocin and arginine-vasopressin receptors and their modulation of autism spectrum disorder," *Frontiers Mol. Neurosci.*, vol. 11, p. 27, Feb. 2018.

[13] W. Pereanu, E. C. Larsen, I. Das, M. A. Estévez, A. A. Sarkar, S. Spring-Pearson, R. Kollu, S. N. Basu, and S. Banerjee-Basu, "AutDB: A platform to decode the genetic architecture of autism," *Nucleic Acids Res.*, vol. 46, no. D1, pp. D1049–D1054, Jan. 2018.

[14] D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen, and C. V. Mering, "STRING V11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D607–D613, Nov. 2018.

[15] X. Zhao, L. Chen, Z.-H. Guo, and T. Liu, "Predicting drug side effects with compact integration of heterogeneous networks," *Current Bioinf.*, vol. 14, no. 8, pp. 709–720, Dec. 2019.

[16] J. Che, L. Chen, Z.-H. Guo, S. Wang, and Aorigele, "Drug target group prediction with multiple drug networks," *Combinat. Chem. High Throughput Screening*, vol. 23, no. 4, pp. 274–284, May 2020.

[17] J.-P. Zhou, L. Chen, and Z.-H. Guo, "IATC-NRAKEL: An efficient multi-label classifier for recognizing anatomical therapeutic chemical classes of drugs," *Bioinformatics*, vol. 36, pp. 1391–1396, Oct. 2019.

[18] L. Wang, Z. Xuan, S. Zhou, L. Kuang, and T. Pei, "A novel model for predicting LncRNA-disease associations based on the LncRNA-MiRNA-Disease interactive network," *Current Bioinf.*, vol. 14, no. 3, pp. 269–278, Mar. 2019.

[19] N. Srivastava, B. N. Mishra, and P. Srivastava, "In-silico identification of drug lead molecule against pesticide exposed-neurodevelopmental disorders through network-based computational model approach," *Current Bioinf.*, vol. 14, no. 5, pp. 460–467, Jun. 2019.

[20] L. Wang, P. Ping, L. Kuang, S. Ye, F. M. Buland Lqbal, and T. Pei, "A novel approach based on bipartite network to predict human microbe-disease associations," *Current Bioinf.*, vol. 13, no. 2, pp. 141–148, Mar. 2018.

[21] J. Li, L. Chen, S. Wang, Y. Zhang, X. Kong, T. Huang, and Y.-D. Cai, "A computational method using the random walk with restart algorithm for identifying novel epigenetic factors," *Mol. Genet. Genomics*, vol. 293, no. 1, pp. 293–301, Sep. 2018.

[22] Y.-D. Cai, Q. Zhang, Y.-H. Zhang, L. Chen, and T. Huang, "Identification of genes associated with breast cancer metastasis to bone on a protein-protein interaction network with a shortest path algorithm," *J. Proteome Res.*, vol. 16, pp. 1027–1038, Jan. 2017.

[23] S. Köhler, S. Bauer, D. Horn, and P. N. Robinson, "Walking the interactome for prioritization of candidate disease genes," *Amer. J. Hum. Genet.*, vol. 82, no. 4, pp. 949–958, Apr. 2008.

[24] X. Yu, T. Zeng, and G. Li, "Integrative enrichment analysis: A new computational method to detect dysregulated pathways in heterogeneous samples," *BMC Genomics*, vol. 16, no. 1, p. 918, Nov. 2015.

[25] H. Liang, L. Chen, X. Zhao, and X. Zhang, "Prediction of drug side effects with a refined negative sample selection strategy," *Comput. Math. Methods Med.*, vol. 2020, May 2020, Art. no. 1573543.

[26] L. Chen, T. Liu, and X. Zhao, "Inferring anatomical therapeutic chemical (ATC) class of drugs using shortest path and random walk with restart algorithms," *Biochimica Biophysica Acta (BBA)-Mol. Basis Disease*, vol. 1864, no. 6, pp. 2228–2240, Jun. 2018.

[27] L. Chen, Y.-H. Zhang, Z. Zhang, T. Huang, and Y.-D. Cai, "Inferring novel tumor suppressor genes with a protein-protein interaction network and network diffusion algorithms," *Mol. Therapy-Methods Clin. Develop.*, vol. 10, pp. 57–67, Jun. 2018.

[28] A. Hajdo-Milašinović, S. I. J. Ellenbroek, S. van Es, B. van der Vaart, and J. G. Collard, "RAC1 and RAC3 have opposing functions in cell adhesion and differentiation of neuronal cells," *J. Cell Sci.*, vol. 120, no. 4, pp. 555–566, Feb. 2007.

[29] F. Zeidán-Chuliá, J. L. Rybarczyk-Filho, A. B. Salmina, B.-H.-N. de Oliveira, M. Noda, and J. C. F. Moreira, "Exploring the multifactorial nature of autism through computational systems biology: Calcium and the rho GTPase RAC1 under the spotlight," *NeuroMol. Med.*, vol. 15, no. 2, pp. 364–383, Jun. 2013.

[30] S. Corbetta, S. Gualdoni, G. Ciceri, M. Monari, E. Zuccaro, V. L. J. Tybulewicz, and I. De Curtis, "Essential role of RAC1 and RAC3 GTPases in neuronal development," *FASEB J.*, vol. 23, no. 5, pp. 1347–1357, May 2009.

[31] C. Bacon, M. Schneider, C. Le Magueresse, H. Froehlich, C. Sticht, C. Gluch, H. Monyer, and G. A. Rappold, "Brain-specific Foxp1 deletion impairs neuronal development and causes autistic-like behaviour," *Mol. Psychiatry*, vol. 20, no. 5, pp. 632–639, May 2015.

[32] D. H. Ebert and M. E. Greenberg, "Activity-dependent neuronal signalling and autism spectrum disorder," *Nature*, vol. 493, no. 7432, pp. 327–337, Jan. 2013.

[33] X. Chen, H. Cao, A. Saraf, L. S. Zweifel, and D. R. Storm, "Overexpression of the type 1 adenylyl cyclase in the forebrain leads to deficits of behavioral inhibition," *J. Neurosci. Off. J. Soc. Neurosci.*, vol. 35, pp. 339–351, 2015.

[34] Y.-H. Jiang *et al.*, "Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing," *Amer. J. Hum. Genet.*, vol. 93, no. 2, pp. 249–263, Aug. 2013.

[35] L. S. Miraucourt, R. Dallel, and D. L. Voisin, "Glycine inhibitory dysfunction turns touch into pain through PKCgamma interneurons," *PLoS ONE*, vol. 2, no. 11, p. e1116, Nov. 2007.

[36] F.-S. Sheu, R. M. Marais, P. J. Parker, N. G. Bazan, and A. Routtenberg, "Neuron-specific protein F1GAP-43 shows substrate specificity for the beta subtype of protein kinase C," *Biochem. Biophys. Res. Commun.*, vol. 171, no. 3, pp. 1236–1243, Sep. 1990.

[37] L. Ji, A. Chauhan, and V. Chauhan, "Reduced activity of protein kinase C in the frontal cortex of subjects with regressive autism: Relationship with developmental abnormalities," *Int. J. Biol. Sci.*, vol. 8, no. 7, pp. 1075–1084, 2012.

[38] J. G. Flanagan, N. W. Gale, T. Hunter, E. B. Pasquale, and M. Tessier-Lavigne, "Unified nomenclature for Eph family receptors and their ligands, the ephrins," *Cell*, vol. 90, pp. 403–404, Aug. 1997.

[39] K. S. Cramer and I. J. Miko, "Eph-ephrin signaling in nervous system development," *FResearch*, vol. 5, p. 413, Mar. 2016.

[40] A. Aulehla, C. Wehrle, B. Brand-Saberi, R. Kemler, A. Gossler, B. Kanzler, and B. G. Herrmann, "WNT3A plays a major role in the segmentation clock controlling somitogenesis," *Develop. Cell*, vol. 4, no. 3, pp. 395–406, Mar. 2003.

[41] H. Kalkman, "A review of the evidence for the canonical WNT pathway in autism spectrum disorders," *Mol. Autism*, vol. 3, no. 1, p. 10, Oct. 2012.

[42] J. Fu, H.-M. Ivy Yu, T. Maruyama, A. J. Mirando, and W. Hsu, "GPR177/mouse wntless is essential for wnt-mediated craniofacial and brain development," *Develop. Dyn.*, vol. 240, no. 2, pp. 365–371, Feb. 2011.

[43] J. Coyle-Rink, L. D. Valle, T. Sweet, K. Khalili, and S. Amini, "Developmental expression of WNT signaling factors in mouse brain," *Cancer Biol. Therapy*, vol. 1, no. 6, pp. 640–645, Nov. 2002.

[44] P. C. Salinas, "WNT signaling in the vertebrate central nervous system: From axon guidance to synaptic function," *Cold Spring Harbor Perspect. Biol.*, vol. 4, no. 2, Feb. 2012, Art. no. a008003.

[45] J. H. Rogers, "Calretinin: A gene for a novel calcium-binding protein expressed principally in neurons," *J. Cell Biol.*, vol. 105, no. 3, pp. 1343–1353, 1987.

[46] Y. A. Lawrence, T. L. Kemper, M. L. Bauman, and G. J. Blatt, "Parvalbumin-, calbindin-, and calretinin-immunoreactive hippocampal interneuron density in autism," *Acta Neurologica Scandinavica*, vol. 121, no. 2, pp. 99–108, Feb. 2010.

[47] T. Sadakata and T. Furuichi, "Developmentally regulated Ca^{2+}-dependent activator protein for secretion 2 (CAPS2) is involved in BDNF secretion and is associated with autism susceptibility," *Cerebellum*, vol. 8, no. 3, pp. 312–322, Sep. 2009.

[48] P. M. Vassiliev, A. A. Spasov, L. R. Yanaliyeva, A. N. Kochetkov, V. V. Vorfolomeyeva, V. G. Klochkov, and D. T. Appazova, "Neural network modeling of the multitarget rage inhibitory activity," *Biochemistry (Moscow), Suppl. B, Biomed. Chem.*, vol. 13, no. 3, pp. 256–263, Jul. 2019.

[49] K. Yang, F. Cao, A. M. Sheikh, M. Malik, G. Wen, H. Wei, W. T. Brown, and X. Li, "Up-regulation of Ras/Raf/ERK1/2 signaling impairs cultured neuronal cell migration, neurogenesis, synapse formation, and dendritic spine development," *Brain Struct. Function*, vol. 218, no. 3, pp. 669–682, May 2013.

[50] C. A. Walsh, E. M. Morrow, and J. L. Rubenstein, "Autism and brain development," *Cell*, vol. 135, no. 3, pp. 396–400, 2008.

[51] D. L. Moore, M. G. Blackmore, Y. Hu, K. H. Kaestner, J. L. Bixby, V. P. Lemmon, and J. L. Goldberg, "KLF family members regulate intrinsic axon regeneration ability," *Science*, vol. 326, no. 5950, pp. 298–301, Oct. 2009.

[52] V. Briz, Y. T. Hsu, Y. Li, E. Lee, X. Bi, and M. Baudry, "Calpain-2-mediated PTEN degradation contributes to BDNF-induced stimulation of dendritic protein synthesis," *J. Neurosci. Off. J. Soc. Neurosci.*, vol. 33, no. 10, pp. 4317–4328, 2013.

YAN ZHAO was born in Jinan, China. She received the bachelor's degree in nursing from Weifang Medical College, in 2004 and the master's degree in public health from the School of Public Health, Shandong University, in 2011. She has been working with the Qilu Hospital of Shandong University, since 1988. In 2009, she was employed as an Associate Professor of nursing by Shandong University, and she currently works with the Department of Geriatric Neurology.

PING ZHAO was born in Jinan, China. She received the bachelor's degree from Shandong University, in 2008. She worked with Changqing Chengguan Hospital of Jinan from 1991 to 2001. Since 2002, she has been working with the Changqing District Hospital of Traditional Chinese Medicine, Jinan. In 2006, she was employed as a chief nurse, and she currently works with the Department of Rehabilitation.

HAIYAN LIANG received the B.S. degree in information and computing science from the East China Institute of Technology, in 2018. She is currently pursuing the master's degree in computer science with the College of Information Engineering, Shanghai Maritime University.

XIAOLIN ZHANG was born in Xinyang, China. She received the B.S. degree in computer science and technology and the LL.B. degree from Henan Normal University, in 2018.

During her undergraduate course, she obtained the scholarship certificate from Henan Normal University and published two academic articles.

• • •