Progress in The Molecular Genetic Research in Autism Spectrum Disorders

Xiao Han¹, Rui Peng³, Linna Zhang²*# and Zhiwen Shi¹, 4*

*Both the authors contributed equally

¹State Key Laboratory of Genetic Engineering, MOE Key Laboratory of Contemporary Anthropology, and Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai 200438, China
²Huangpu District Mental Health Center, 1162 Qu Xi Road, Shanghai 200023, China
³Obstetrics and Gynecology Hospital, Institute of Reproduction and Development, Fudan University, Shanghai 200011, China
⁴B. Braun Precision Medical Technology (Shanghai) Co., Ltd. 398 Jiang Su Road, Shanghai 200050, China

*Corresponding Authors: Dr. Zhiwen Shi, State Key Laboratory of Genetic Engineering, MOE Key Laboratory of Contemporary Anthropology, and Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University, Shanghai 200438, China, E-mail: 14110700089@fudan.edu.cn

Dr. Linna Zhang, Huangpu District Mental Health Center, 1162 Qu Xi Road, Shanghai 200023, China, E-mail: Zhangln_1988@126.com

Received: 24 May 2019; Accepted: 31 May 2019; Published: 05 June 2019

Abstract

Autism spectrum disorders (ASDs) bring heavy economic and spiritual burden to families and society. Early diagnosis and intervention are helpful to improve the behavioral abnormalities of ASDs patients and their ability of self-care and social interaction. Prenatal diagnosis and genetic testing provide conditions for early screening, but a clear understanding of the genetic causes of ASDs is still needed. The AutDB database integrates the pathogenic mutations published in recent years. But unfortunately, among hundreds of ASDs susceptibility genes, most of the evidence is weak, and only a small part of the pathogenicity is clear. Significant progress has been made in the use of different animal models to verify the causal relationship of ASDs and to study the pathogenesis of ASDs at the molecular and cellular level. Functional clustering analysis reveals that highly heterogeneous genetic factors
ultimately lead to clinical phenotypes through common pathways, suggesting that therapeutic approaches should be tailored to these pathways. Establishing rapid genetic screening methods and developing targeted therapeutic drugs will bring good news to ASDs patients and families.

**Keywords:** Autism spectrum disorders; Prenatal diagnosis; Clinical phenotypes; Therapeutic drugs

1. **Introduction**

ASDs are a set of early-onset neurodevelopmental conditions, characterized by difficulties in social communication and reciprocal interactions, and unusually circumscribed interests and repetitive behaviors. It is estimated that the prevalence of ASDs in 8-year-olds is 13-29‰, and 60-72‰ in the general population [1, 2]. The male-to-female ratio of ASDs is 3:1 [3].

Although the causes of ASDs are still not fully understood, many twin studies over the past 40 years have highlighted the contribution of genes and environment to ASDs. Twin studies provide strong evidence for the heritability of ASDs to range from 64% to 91% [4], suggesting a gene-environment interaction [5]. From an evolutionary perspective, traits of ASDs may be subject to positive selection pressure [6]. Such individuals may have successfully exchanged products or their fixing skills, thereby acquiring resources and improving their reproductive ability, which may help maintain the autism allele in the gene pool. However, there is no consensus on the specific genetic risk factors for ASDs.

Studies on cytogenetics, linkage, association, sequencing have shown that the genetic architecture of ASDs is complex and heterogeneous [7, 8]. Many genetic variants associated with ASDs are highly pleiotropic. High locus heterogeneity has also been reported: about 1000 genes presumed to be involved. Both rare mutations with large effects and common variants with small effects were found to play a role in the epidemic of ASDs. Rare mutations are often found in ASDs, possibly in the form of Mendelian syndrome, chromosomal abnormalities, rare copy number variants and single-nucleotide variants. De novo mutations (microdelete or microduplication of copy number variations or nonsense, splice, and frameshift mutations) occurred in reproductive system (especially parents) may have a significant impact on this outcome [9]. Similarly, copy number variations(CNVs) with moderate effect sizes and variable penetrance may also play a role [8]. In terms of common variants, genome-wide association studies have identified a number of SNPs, but the methods were largely unsuccessful. Several studies identified signals did not reached the genome-wide level of significance [10, 11].

2. **CNVs associated with ASDs**

CNVs, the most common type of structural variations in the human genome, are considered to be a vital contributor to the pathogenesis of ASDs. Large-scale studies comparing the non-ASDs population and population with ASDs showed that the incidence of CNVs in patients was higher than in controls, with an overall frequency of 8-21%, and
10% of detected CNVs were considered clinically relevant [12, 13]. Some rare, de novo CNVs overlap genes have been associated with ASDs, including NLGN3, NLGN4X, SHANK3, ASTN2, PTCHD1, DDX53, NRX1N, CNTN4, CNTNAP2, OXTR, PTCHD1, NLGN1 and SHANK1 [[8, 14]]. In addition, recurrent duplications and deletions have been uncovered as being risk factors for ASDs. These genomic variants include 16p11.2 deletions and duplications (approximately 0.8%), 15q11-13 duplications (approximately 0.5%) and 22q11 deletions and duplications (approximately 0.5%). 7q11.23 duplication (~0.2%), 1q21.1 deletion and duplication (~0.2%) [2, 8]. It also has been reported that the frequency of genome rearrangement of these risk sites accounts for > 1% of ASDs cases [15].

In 2007, the Autism Gene Database (AutDB) authorized by Simon foundation was established to collect pathogenic genes and candidate genes of ASDs. AutDB includes 99 ASD-related syndromic genes and 3022 non-syndromic genes, 4544 CNVs and 158 ASD-related linkage regions. The genetic basis of ASDs is complex, and two main models are proposed to explain the genetic components of ASDs.

### 2.1 Common variant of CNVs

The contribution of common variation to the risk of developing ASD is less clear. To produce a more comprehensive picture, a family study of 466 Swedish patients with ASDs has been conducted [16]. Their unaffected twins, siblings, cousins, and 2580 typically neurotypical individuals were also included. Consistent with previous reports on the effects of common ASDs variants, the data provided by the authors indicated that families with at least two children with ASDs had more frequent CNVs (about 60%) than families with only one affected member. Data from the same study suggest that rare genetic CNVs played a role in only 2.6% of ASDs. Recently, Grove et al. [17] reported a genome-wide meta-analysis, including 18381 ASDs patients and 27969 controls, identifying a set of common risk CNVs closely associated with ASDs for the first time.

### 2.2 Rare CNVs

Numerous studies have focused on the frequency and specificity of rare CNVs. Pinto et al. [14] studied the contribution of rare CNVs in 996 patients with ASDs and their relatives (876 trios) and 1287 neurotypical individuals. Rare CNVs were significantly more common in patients with ASDs than in the neurotypical population (7.6% vs 4.5%), especially given the known loci associated with ASDs and/or intellectual disability. Furthermore, this association was even more important when only deletions were considered. The authors pointed out that many sites (such as SHANK2, SYNGAPI1 and DLGAP) were new candidate sites for ASDs. Recently, the same team [14] expanded the previous study by adding 1,604 families to the original group, recording that 16p11.2 and NRXN1 (2p16.3) were the most common deletions among the affected individuals (0.3%) and 15q11-q13 were the most common duplications (0.25%).

As already reported, rare CNVs often occurred as de novo events [8]. Some authors hypothesized there was an association between ASDs and de novo CNVs because the de novo CNVs mutations were observed more frequently in patients than in controls (3-10% vs 1%). In fact, in 2001, two studies were conducted using simplex families in
Simons Simplex Collection, which was a database, to collect the genetic data of more than 2000 families and to explore the degree of genetic risk of ASDs. In the second study, the hypothesis that ASDs were an unfortunate consequence of a common-risk CNVs was denied, and the research focused on rare events. In fact, CNVs were more common in females than males, and males showed more de novo deletions than duplications. Sanders et al. [18] confirmed previous research results by exome sequencing of 2591 Simons Simplex families. Combined with CNVs data published by Autism Genome Project, they found that there was a strong correlation between ASDs and 6 risk loci (1q21.1, 3q29, 7q11.23, 16p11.2, 15q11.2-13 and 22q11.2).

Yuen et al. sequenced the genomes of 5,205 ASDs patients (N=2620) and healthy controls and detected an average of 400 CNVs (> 2kb in size) in each genome. 7% of subjects showed at least one disease-causing chromosomal variant (N=21) or CNVs of different subtypes (N=152). In particular, the authors also found that 22 CNVs overlapped with some genes were associated with ASDs susceptibility [19].

3. Single-nucleotide polymorphisms associated with ASDs

Several large-scale sequencing projects have been completed [20, 21]. In simplex families, the incidence of rare de novo and inherited single-nucleotide polymorphisms (SNPs) in ASDs cases was higher than their unaffected siblings. In contrast, there was no difference in the incidence of synonymous mutations. It is estimated that SNPs associated with ASDs occur in 1% of the population. Unlike rare variants, SNPs in individuals are thought to have less impact on disease and may even be benign. The most common SNPs are found in genome-wide association studies. These SNPs are thought to exist in large areas of the intervening genome and are expected to be co-inherited. Common SNPs are thought to explain 40% of genetic risk in ASDs [22]. There are some common SNPs identified in several large GWAS researches: rs4307059 5p14.1 region, rs10513025, rs10513026, rs16883317 5p15.2, rs4141463 20p12.1 region, and rs936938 rs6537835. rs1877455 1p13.2 area. These SNPs near CDH9, CDH10, CSDE1, MACROD2, NRAS, SEMA5A, TAS2R1, and TRIM33 genes. Despite these genome-wide discoveries, none of these results have been consistently replicated across studies because of the relatively small sample size.

In order to overcome the phenotypic heterogeneity of ASDs, other studies have focused on a series of related phenotypes of ASDs and reported the susceptibility loci associated with each subphenotype. For example, delayed language was found to be associated with 3p12.3 [23]. This region contains the ROBO2 gene, which encodes a conserved axon-binding receptor, and it has been reported that the expression of ROBO2 in the anterior cingulate cortex is decreased in ASDs cases [24]. Another common genetic association was found in the 7q35 region containing CNTNAP2 gene [25]. These studies indicate that a site-specific sub-phenotype may be widespread in the population [26]. In addition to these examples, several other studies examined common genetic variants associated with learning and cognition. In a recent study of high-functioning ASDs cases, researchers discovered a variant in the ARNT2 gene (rs17225178) [27] that encoded an important transcription factor. Interestingly, ARNT2 is widely known to regulate cellular responses to hypoxia through its dimerization with HIF1A in the central nervous system.
Common variants in individuals with learning difficulties, such as rs789859 in region 3q29 [28], are also associated with ASDs, and the region is missing from scratch in patients with learning difficulties in the neuropsychiatric population [29].

4. Candidate Genes Associated with ASDs

Significant advancement has been made in the ASDs field in recent decades, with efforts toward understanding the molecular basis of synapse. The present study examines many ASDs-related genes which regulate neuronal activity and modulate synaptic strength or number. The development of genetically modified mouse models also have fueled research and increased understanding of potential genetic causes of ASDs. These results raise the possibility that de novo, rare and highly penetrant mutations in clinical cases may lead to the distortion of typical neuronal connectivity, increasing the risk of ASDs [30].

4.1 Transmembrane protein associated with ASDs

Neuroligins (NLGNs) located at the postsynaptic side of the synapse that have been shown to contribute to synaptic neurotransmission through their influence on synaptic formation and distributed at excitatory and inhibitory synapses. Importantly, mutations in NLGN genes have been linked to the abovementioned neuropsychiatric disorders [31], such as missense mutations of NLGN3 and NLGN4 in ASDs [32]. In the NLG R451C (NLGN-3 mutated) model mice, which display autism-like phenotypes, increased dendritic spine density was found [33].

Neurexins (NRXNs) are conserved synaptic cell adhesion molecules. They include three different genes (Nrxn1, Nrxn2, and Nrxn3) and two protein isoforms for each gene in mammals. They participate in synapse formation, including synapse specialization, establishment, maturation, and plasticity. Rare NRXN1 mutations were shown to be associated with ASDs in recent studies[34]. More direct evidence comes from model animals. Mice with Nrxn1a deletion resulted in electrophysiological changes, impaired spatial memory and increased repetitive behaviors, traits that were all consistent with cognitive impairments [35]. Interestingly, Nrxn1a heterozygous knock-out mice showed differences in responsiveness to novelty and accelerated habituation to novel environments compared to wild type (+/+ ) litter-mates, indicating sex-specific differences of the behavioral phenotype.

CNTNs are a six-member subgroup of the immunoglobulin cell adhesion molecule superfamily (IgCAMs) with pronounced brain expression and function. They play an essential role in the formation, maintenance and plasticity of neuronal networks. Recent genetic studies of neuropsychiatric disorders have pinpointed contactin-4 (CNTN4), contactin-5 (CNTN5) and contactin-6 (CNTN6) as candidate genes in neurodevelopmental disorders, particularly in autism spectrum disorders (ASDs) [36].

Human genetic studies have identified multiple ASDs-associated genes that are implicated in the structure and function of neuronal synapses including the cadherin and protocadherin superfamily of calcium dependent neural cell adhesion molecules [37, 38]. PCDH genes are renowned for their roles in dendritogenesis, dendrite arborization
and dendritic spine regulation making them perfect candidate genes for ASDs [39]. Deletions near PCDH10 have also been reported in families with autism[40]. Further supporting evidence for protocadherins in ASDs was the finding that ASDs brains with PCDH10 variants have increased dendritic spine densities compared to controls [41].

4.2 Scaffolding proteins associated with ASDs
SHANK proteins are indispensable scaffolding proteins that function as the “backbone” of the postsynaptic density in forming protein–protein interactions that are integral to the recruitment and to the clustering of target proteins at the postsynaptic density. SHANK proteins are the product of 3 genes—SHANK1, SHANK2, and SHANK3. SHANK proteins have gained attention over the past decade for their roles in ASDs. Over 100 disease causing mutations in this set of proteins are reported in the HGMD database. Genetically modified mouse models also provide strong evidence. Shank3 knockout mice showed the autistic-like social deficits and this phenotype could be durably rescued via social training coupled with optogenetic activation of Pet1 neurons in the dorsal raphe nucleus (Optogenetic activation of dorsal raphe neurons rescues the autistic-like social deficits in Shank3 knockout mice) [42]. Schmeisser et al reported increased self-grooming behavior in adult female Shank2 knockout mice, while digging behavior was strongly decreased [43]. In adult male Shank2 knockout mice, self-grooming was unchanged, but digging behavior was strongly decreased [44]. Marble burying was strongly reduced in adult Shank1+/− heterozygous and Shank1 knockout mice across social contexts, as compared to adult Shank1 wildtype littermate controls [45].

4.3 Ion channels associated with ASDs
Ion channels regulate synaptic activities and play important roles in nervous system disease. The SCN1A gene encodes a subunit of sodium channel, previously associated with epilepsy and indicated as the candidate gene for ASDs and the de novo and rare missense variant (p. P1894L) at a highly conserved position was found in sporadic ASDs [46]. The calcium channel, voltage-dependent, L type, alpha 1C subunit gene (CACNA1C) encodes an alpha-1 subunit of a voltage-dependent calcium channel. Mutations in CACNA1C and other L type calcium channels participate in the pathogenesis of Timothy and Fragile X syndromes, monogenic disorders with ASD-like symptoms [47]. Genome-wide association studies (GWAS) have shown an association between the single nucleotide polymorphism (SNP) rs1006737 in this gene with autism in a Chinese Han population [48]. The KCNMA1 gene, which encodes the alpha-subunit of the large conductance Ca2+-activated K+ (BKCa) channel, a synaptic regulator of neuronal excitability, has found to be physically disrupted in autism [49].

4.4 Proteins at non-synaptic sites associated with ASDs
Some proteins encoded by ASDs candidate genes are at non-synaptic sites. Iuliana.et.al identified variants in one gene, Fanconi-associated nuclease 1 (FANI) has being associated with schizophrenia and ASDs. They suggested that FANI is a key driver in the 15q13.3 locus for the associated psychiatric and neurodevelopmental phenotypes [50]. MECP2, a nuclear protein that is expressed postnatally during mammalian brain development, can bind both methylated and unmethylated CpG sequences and can either repress or activate transcription. Mutations in MECP2 can lead to ASDs repetitive behavior, hypotonia and anxiety [51, 52]. To sum up: about 35% of the ASDs patients
are caused by genetic factors, but almost any known pathogenic mutation in the known genes are not the major cause of ASDs. The genetic etiology of ASDs is caused by a number of relatively rare genetic mutations [53, 54].

5. Drug Development, Genetic Testing and Genetic Counseling

5.1 Drug development for ASDs

Various ASDs mouse models are currently being used for drug screening in the hope of finding the ideal drugs that would improve the core symptoms of ASDs. Risperidone is a dopamine antagonist approved by the Food and Drug Administration to reduce irritability in patients with ASDs, but no drug has been approved for the core symptoms of ASDs.

Inhibiting the mTOR pathway in adult Tsc2+/- mice can restore synaptic plasticity and rescue cognitive and social deficit phenotypes [55]. mTOR binds to different proteins to form two complexes, mTORC1 and mTORC2. Rapamycin and its derivatives can inhibit the serine threonine kinase activity of mTORC1. New mTOR inhibitors, such as Torin 1 and KU-0063794, can inhibit both mTORC1 and mTORC2 at the same time. Now they are in the early stage of clinical research.

The maternal copy of the UBE3A gene is abnormal in Angelman syndrome patients, while the paternal copy of the UBE3A gene is normal but not expressed. So, activating the expression of paternal UBE3A in neurons is a potential therapeutic measure. Drug screening showed that topotecan and irinotecan could activate the expression of paternal UBE3A gene [56]. UBE3A-ATS is a long-chain antisense RNA silencing the expression of paternal UBE3A. Antisense oligonucleotides (ASO) targeting UBE3A-ATS may activate the expression of paternal UBE3A gene in the brain of patients. In Angelman syndrome mice, intraventricular injection of antisense RNA against UBE3A-ATS partially restored the protein level of UBE3A and improved some cognitive deficits [57].

In fragile X syndrome, the absence of FMRP enhances the glutamate-mediated signaling pathway by metabolic glutamate receptor 5 (mGluR5), resulting in synaptic plasticity defects. Pre-clinical studies have shown that mGluR5 antagonists can improve behavioral deficits in mice with fragile X syndrome [58]. Insulin-like growth factor 1 (IGF1) can rescue the abnormal formation of excitatory synapses and synaptic conduction defects [59] in neurons induced by SHANK3 mutations in vitro. IGF1 also partially saved the synaptic structure and function of MeCP2 mutant mice, significantly improving their behavior and prolonging their life span [60].

Oxytocin is widely used in obstetric settings to stimulate uterine contraction. In addition, many studies have shown that oxytocin plays an important role in social interaction. Oxytocin deficiency or insensitivity may be related to social disorders [61]. Exogenous or endogenous oxytocin can save the social deficit of Cntnap2 mutant mice [62]. Oxytocin nasal administration can promote social interaction in normal people and improve social disorders in autistic patients, suggesting the application prospects of oxytocin and its receptor agonists in the clinical treatment of ASDs.
5.2 Genetic testing for ASDs

Finding the pathogenic genetic mutations of ASDs will be helpful for treatment and prognosis. In 2013, the American Society of Medical Genetics issued a series of recommendations for genetic testing [63]. All ASDs children were subjected to Chromosomal Microarray Analysis (CMA) as a preliminary test, including comparative genomic hybridization and single nucleotide polymorphism analysis. Secondary tests included: fragile X mutation detection in all male patients; MECP2 sequencing in all female patients and male patients with clinical symptoms suggesting MECP2 involvement; PTEN sequencing in children with malformed giant symptoms. In general, all ASDs children are required to undergo CMA testing and then further testing based on gender, family history and clinical symptoms. With the technology development, exome sequencing is likely to replace CMA testing. Because exome sequencing costs are similar to CMA testing costs, but exome sequencing can give more information. Parents of ASDs patients often ask about the clinical value of genetic testing. Using the above genetic testing methods, it is estimated that 21% of ASDs children can get diagnostic letters [64].

5.3 Genetic counseling for ASDs

Genetic counseling for patients’ families is also important. The phenotypic diversity and high genetic heterogeneity of ASDs make genetic counseling of ASDs challenging. For parents who have children with ASDs, a very real problem is the risk of ASDs in next births. Unfortunately, this risk varies greatly and is more complex to analyze. Women with ASDs patients in their offspring will increase the risk of ASDs in next births. For example, if a mother has two boys sick, the risk for her next boy is 32%. If only one boy is sick, the risk for the next boy is about 10% [65]. Although these population statistics help families understand the risk of ASDs in their offspring, this wide range of risk limits the value of this information to individual families. High-throughput sequencing and genetic testing of patients and their parents, as well as an accurate understanding of genetic factors, will substantially improve the prediction of disease risk.

6. Prospect

It is believed that with the deepening of research, there will be a clearer clinical classification and understanding of ASDs, which is conducive to the early diagnosis and screening of ASDs. For individuals, the causes of ASDs are complex and diverse. Individualized treatment should be introduced in the treatment of ASDs. As more and more mechanisms of ASDs are elucidated, new therapies are constantly proposed, which also provide a good basis for individualized treatment of ASDs patients.

References


49. Pop-Jordanova N, Plasevska-Karanfilska D. Autism - genetics, electrophysiology and clinical syndromes, Prilozi/Makedonska akademija na naukite i umetnostite, Oddelenje za biolo?ki i medicinski nauki=Contributions/Macedonian Academy of Sciences and Arts, Section of Biological and Medical Sciences 35 (2014): 133-146.


This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license 4.0