Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Curr Opin Neurol. Author manuscript; available in PMC 2020 August 01.

Published in final edited form as:

Curr Opin Neurol. 2019 August ; 32(4): 627–634. doi:10.1097/WCO.0000000000000718.

Recent Genetic and Functional Insights in Autism Spectrum Disorder

Moe Nakanishi[1], **Matthew P Anderson**[2,3,4], **Toru Takumi**[1]

[1]RIKEN Brain Science Institute, Wako, Saitama 310-0198, Japan

[2]Departments of Neurology and Pathology, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02115, USA.

[3]Boston Children's Hospital Intellectual and Developmental Disabilities Research Center, 300 Longwood Avenue, Boston, MA 02115, USA.

[4]Program in Neuroscience, Harvard Medical School, 300 Longwood Avenue, Boston, MA 02115, USA.

Abstract

Purpose of this review—Recent advances in genetic technologies allowed researchers to identify large numbers of candidate risk genes associated with autism spectrum disorder (ASD). Both strongly penetrant rare variants and the accumulation of common variants with much weaker penetrance contribute to the etiology of ASD. To identify the highly-confident candidate genes, software and resources have been applied, and functional evaluation of the variants has provided further insights for ASD pathophysiology. These studies ultimately identify the molecular and circuit alteration underlying the behavioral abnormalities in ASD. In this review, we introduce the recent genetic and genomic findings and functional approaches for ASD variants providing a deeper understanding of ASD etiology.

Recent findings—Integrated meta-analysis that recruited a larger number of ASD cases has helped to prioritize ASD candidate genes or genetic loci into a highly-confidence candidate genes for further investigation. Not only coding but also non-coding variants have been recently implicated to confer the risk of ASD. Functional approaches of genes or variants revealed the disruption of specific molecular pathways. Further studies combining ASD genetics and genomics with recent techniques in engineered mouse models show molecular and circuit mechanisms underlying the behavioral deficits in ASD.

Summary—Advances in ASD genetics and the following functional studies provide significant insights into ASD pathophysiology at molecular and circuit levels.

Keywords

autism; non-coding variant; common variant; mouse model; CNV

Corresponding to Toru Takumi, MD, PhD, RIKEN Center for Brain Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan, Tel:+81 48 467 5906, Fax:+81 48 467 6079, toru.takumi@riken.jp and Matthew P. Anderson, MD, PhD, Center for Life Science, 330 Brookline Ave, E/CLS-645, Boston, MA 02215, USA, Tel: (617) 735-3202, FAX: (617) 735-3249, Matthew_Anderson@bidmc.harvard.edu.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Introduction

Autism spectrum disorder (ASD) is one of the most common neurodevelopmental disorders (NDDs) with prevalence estimated to be 1-2.47 % of children[1,2]. According to Diagnostic and Statistical Manual (DSM) 5, it is diagnosed by two core symptoms including impaired social communication and interaction, as well as repetitive and restrictive behaviors. Although individuals with ASD share core features, ASD is a clinically heterogeneous group of disorders. They often show a complex combination of medical, neurologic, and behavioral comorbid symptoms including seizure, impaired motor skills, intellectual disability (ID), speech delay, sleep disorder, and gastrointestinal problems. These symptoms generally appear before two years of age and continue throughout life. Although the research on ASD has made great progress during the past decades, no standard treatments are available likely due to the clinical and genetic heterogeneity of ASD[3].

Both genetic and environmental factors contribute to the etiology of ASD, however, the most progress has recently been made in understanding the genetic defects underlying the disorder. Early twin and family studies revealed that ASD is strongly genetically influenced with an estimated heritability of ASD of 40-90%[4-7]. In the past decade, large-scale genomic studies have identified hundreds of genetic defects including single-nucleotide variants (SNVs) and genomic copy number variations (CNVs) that are associated with ASD[8-10]. Moreover, recent studies have highlighted both rare and common variants, and coding and noncoding genetic changes that significantly contribute to the genetic basis of ASD, consistent with a heterogeneous and complex array of etiologies in ASD pathogenesis. Subsequent translational studies based on these genetic findings have uncovered specific molecular, circuit and behavioral deficits caused by these ASD-associated genetic changes.

In this review, we introduce the recent findings of genetic studies and the functional investigation of the consequences of these ASD-associated genetic and genomic defects using cells and genetically engineered animal models.

Identification of genetic variants involved in ASD

Over the past decade, there have been remarkable advances in genetic technologies such as next generation sequencing that has allowed researchers to identify large numbers of candidate risk variants in ASD. So far both rare and common variants have been suggested to confer risk for ASD but only the functional impact of a minor subset of the rare variants have been well studied. Whole-exome sequencing (WES) and microarray-based comparative genomic hybridization studies for CNVs have characterized the substantial impact of rare variants, especially newly arising *de novo* variants in ASD for a number of years[8,11-15]. Although hundreds of “ASD candidate genes” have been identified, the contribution of each individual candidate gene to the ASD population is very low (< 0.5%). To identify “high-confidence candidate genes” in which deleterious mutations affecting the same gene occur in many cases of ASD, a progressively greater numbers of ASD cases and families have been recruited for such analyses. Rubeis *et al.*, analyzed 3,871 ASD cases and ancestry-matched controls and uncovered that ASD-risk genes were enriched in FMRP-targets, synaptic genes, and genes related to the regulation of transcription or chromatin remodeling[8]. Considering

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the great genetic heterogeneity and diversity of functions implicated by these ASD-associated genes, still larger cohorts were needed and were recently integrated with meta-analysis to provide further insights. A recent very large meta-analysis combined the *de novo* SNVs from 10,927 ASD, ID and developmental delay (DD) and analyzed 12,172 *de novo* variants[16]. They identified recurrent mutations in 253 genes and 124 of them reached genome-wide association study (GWAS) threshold suggesting a strong contribution to ASD disease etiology. In a network level analysis, some specific biological function modules were significantly enriched including the regulation of transcription from RNA polymerase II promoters, functions relating to neurotransmitter and synaptic signaling, transmembrane receptor protein serine/threonine kinase signaling, and c-Jun N-terminal kinase. These studies have provided solid evidence to prioritize specific candidate genes and molecular pathways for further investigation of ASD pathogenesis.

So far genetic studies have mostly recruited thousands of ASD cases and controls of European American ancestry. However, more recent studies have recruited substantial numbers of ASD and control individuals of other ethnicities. In the past year, two studies recruited Japanese cohorts: 262 trios with an individual affected by ASD or 1,108 ASD cases[17,18]. These studies basically replicated the previous findings reported for ASD individuals of European descent, such as the excess of deleterious *de novo* variants in ASD when consider all affected genes in aggregate, with an enrichment of mutations in FMRP targets, synaptic genes and genes essential in mice suggesting a shared pathway and etiology between Caucasians and Japanese[19]. Interestingly, among identified genes, the *ATP2B2* gene that was not strongly implicated in previous studies in Caucasians was enriched for *de novo* mutations in the Japanese ASD cohort. Since differences of race and ethnicity in ASD etiology are still poorly understood, genetic analysis of various ethnicities may provide additional unique insights into ASD pathogenic mechanisms.

Estimating the pathogenesis of variants identified in individuals

Each person harbors millions of variants, which typically include more than 10,000 peptide-sequence altering variants, and more than 100 protein-truncating variants[20]. So the identification of the disease-causing gene variant(s) in an individual's exome or genome remains a challenging problem. Interpreting CNVs that delete or duplicate the dosage of a gene is relatively easier but estimating the functional impact of gene sequence variants is particularly important to the interpretation of SNVs. Recent studies have applied several criteria and computational approaches to enrich potentially functional or pathogenic SNVs from among the numerous variants identified in each ASD case. One of the easiest criteria is the classification of the variant type. Among all SNVs, nonsense (a mutation in which a sense codon that corresponds to one of the twenty amino acids specified by the genetic code is changed to a chain-terminating codon), frameshift, and splice site nucleotide mutations are relatively easy to interpret and are expected to result in a loss-of-function (LOF) of the target gene. These are sometimes called "likely gene-disruptive variants" and are considered to be most deleterious. In the case of peptide-sequence altering missense variants, *in silico* prioritization tools have been developed to help predict their functional impact (Table 1). Although many tools have been developed, the SIFT, PolyPhen2 and Combined Annotation-Dependent Depletion (CADD) are the ones most frequently used in ASD studies[21-23].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The characteristics of recent commonly used software were well reviewed by Eilbeck *et al.* [24]. More recently, gene constraint metrics such as Residual Variation Intolerant Score (RVIS) and probability of LOF intolerance (pLI) are used to enrich ASD risk genes[25,26]. They were developed based on the recent population-scale variant databases such as ESP6500 and ExAC, and statistically model the tolerance of a gene to amino-acid change or LOF variations[27,28]. Since recent studies have suggested ASD-risk genes are less intolerant of mutations, this has become a favored approach for screening ASD candidate genes. Although it has been noted that the applications of these tools have missed some ASD candidate genes, these methods remain useful in setting a threshold for significance when identifying those genes to prioritize for further investigations[16].

Functional testing of ASD-associated variants by using cells and animals

In parallel with genetic approaches, functional evaluation of ASD-associated variants by *in vitro* and *in vivo* assays is crucial to facilitating an understanding of ASD pathophysiology and for the future rational design of therapeutic strategies. This functional testing approach provides insights that cannot be uncovered by genetic studies alone. We explain how functional studies are important in two types of ASD risk genes.

i) ASD risk genes associated with LOF

So far over 1000 genes with alterations in ASD have been identified in recent genetic studies and these are suggested to be high-confidence candidate genes based on the recurrent finding of gene-disruptive variants. This strategy successfully identifies genes using LOF in associated with disease. Over the past several years, strong ASD candidate genes with LOF mutations have been characterized in cell and model animal experimental systems. Chromodomain helicase DNA binding protein 8 (CHD8) is one of the genes most frequently mutated in ASD[8,29]. Recent mouse studies have identified synaptic, transcriptional, and behavioral abnormalities caused by *CHD8* mutations and moreover the study of *Chd8* point mutant mice suggests these abnormalities could be sexually dimorphic[30-32]. Haploinsufficiency of other strong candidate genes associated with ASD and ID, *CHD2* and *SETD2* have been recently reported to cause behavioral and synaptic abnormalities in the mutant mice[33-35]. Characterizing these genes especially using model animals provides insights in molecular, circuit, and behavioral abnormalities related with ASD that may lead to the future development of treatments for these specific genetic subtypes of ASD.

ii) ASD risk genes associated with gain-of-function (GOF) or specific pathways

Genetic studies have identified subsets of ASD genes associated with LOF, but this approach would miss genes containing possible gain-of-function (GOF) mutations. Most *in silico* tools are designed to predict nucleotide changes that give rise to a gene LOF; computational approaches that predict GOF variants remains difficult. Therefore, genes with GOF variants or variants causing unpredictable functional alterations tend to be undervalued despite the large number of missense variants that have been identified in individuals with ASD. For example, large numbers of missense but not gene-disrupting variants have been identified in *CACNA1D* encoding the voltage-gated L-type calcium channel in ASD cases. Because of this LOF bias, limited numbers of human genetic studies have supported its potential role in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ASD. However, functional studies of the effects of these ASD-associated missense variants in *CACNA1D* have revealed that they result in a GOF increase in activity of the channel. In such cases, functional studies are critical to identifying these ASD candidate genes[36,37].

Functional assessment of ASD-associated variants can sometimes uncover unexpected molecular pathways involved in ASD pathogenesis. Recently an ASD-associated missense variant in *SHANK3*, a well-established ASD-candidate gene, was reported to disrupt a novel molecular pathway supported by this gene product. Researchers found that an ASD-associated S685I mutation in *SHANK3* specifically diminishes Shank3-ABI1 interactions, which turned out to be critical for dendritic spine development and synaptic transmission. Moreover, behavioral assay of the knock-in mouse carrying this mutation caused ASD-associated behaviors[38]. So far several potential therapeutic strategies targeting the other *SHANK3* pathways have been proposed[39,40]. However, the finding of *Shank3* S685I indicates that the therapy aimed at correcting a specific Shank3-associated pathway may not be equally applicable to all ASD patients carrying pathogenic *SHANK3* mutations. A recent study identifies that ASD-associated *NLGN1* missense variants unexpectedly affected several distinct processes (e.g. protein misfolding and increased cleavage of extracellular domain)[41]. These findings indicate the complex functional outcome caused by different ASD-associated gene variants and the necessity of evaluating them with biological assays.

Functional analyses of ASD-related CNVs

Not only SNVs but also CNVs contribute to the pathogenesis of ASD with high penetrance. There have been substantial functional studies using mouse models of CNVs since the first CNV model of ASD was developed[10,42]. A recent study shows the significance of serotonin (5-HT) during a developmental stage in 15q11-13 duplication (15q dup) mice[43]. The 5-HT level in the brain of 15q dup mice is decreased[44] and 5-HT neural activity in the dorsal raphe nucleus (DRN) of 15q dup mice is also decreased[43]. These phenotypes may impair neocortical excitation/inhibition balance, correct sensory stimulus tuning and social behavior. Furthermore, restoration of normal 5-HT levels in 15q dup mice reveals the reversibility of ASD-related symptoms in the adult. Decreased DR 5-HT activity during social contact and reduced DR 5-HT neuron activity are also observed in 16p11.2 deletion mice[45]. The decrease in sociability in 16p11.2 deletion mice is rescued by activation of activation of DR 5-HT neurons and pharmacological activation of nucleus accumbens 5-HT1b receptors.

ASD genetics combined with genome engineering and AAV viral vectors in model organisms to resolve the circuit basis of behavioral problems in ASD as above. A recent study mapped the neuronal circuit deficits underlying impaired sociability produced by the increased dosages of the *UBE3A* gene found in a strongly penetrant CNV in ASD, maternal 15q11-13 triplication [extranumerary isodicentric chromosome 15q, idic(15)][46]. Glutamatergic synaptic transmission from ventral tegmental area glutamatergic neurons was impaired by the interaction of increased UBE3A and seizures [a frequent comorbidity in idic(15) and in human idiopathic ASD] because these repress expression of a gene encoding the synapse organizing protein CBLN1 that physically binds presynaptic NRXN1 and postsynaptic GluD1. Both *NRXN1* and *GRID1* (gene for GluD1) are frequently deleted

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

genes in the ASD CNVs[47]. Consistent with the findings of the Krishnan *et al.*' study[46], a recent study also finds the role for UBE3A in the nucleus and shows a function of UBE3A as a transcriptional regulator of the innate immune system in the brain[48].

Contributions of common variants

Although estimates vary across studies, common genetic variations, both coding and noncoding, are thought to account for approximately 20-60% of ASD liability[49-52]. In contrast, *de novo,* extremely rare SNV or CNV can have a larger effect but explain only <10% of overall liability[51]. Recent studies assessing both common and rare variants simultaneously suggest that the accumulation of both types of variants in an individual may have an etiologic role. It was previously thought that a single deleterious *de novo* variant in an individual may be sufficiently penetrant to fully explain the disease, and that common variants were significant only in cases without a strong acting variant. However, two recent large-scale studies suggested the significant contribution of common variants even in ASD cases with a known penetrant deleterious variant. Weiner *et al.*, analyzed 6,454 ASD families and uncovered common polygenic variation still contributes to risk in ASD cases carrying a very deleterious *de novo* variant[53]. Furthermore, Niemi *et al.* examined more extreme cases[54]. They examined 6,987 cases of very severe NDD including ASD with morphological and/or physiological abnormalities in the central nervous system (CNS). Even in the extreme cases in whom monogenic causes were strongly suspected, they found part of the disease risk could be attributed to common variations. These studies show the effects of common variation are not negligible in most of ASD and severe NDD cases with or without highly penetrant rare mutations and highlight the complex genetic basis of ASD and NDD.

Recent studies of common variations, especially GWAS studies, report the genetic overlap between ASD and other psychiatric or NDDs. According to the most recent, largest GWAS study, the polygenic risk for ASD had significant overlap with schizophrenia, major depression and ADHD. Interestingly, common polygenic risk for ASD has been repeatedly suggested to have a positive correlation with educational attainment and IQ[55-59]. Considering that rare deleterious variants have the opposite associated, showing a lower IQ and more ID, the way by which rare and common variants confer the risk for ASD likely differs.

The contribution of non-coding variants in ASD

Recently whole-genome sequencing (WGS) is increasingly being used in ASD studies instead of WES mainly because of the falling cost of sequencing. Unlike WES that sequences only the protein-coding regions, WGS reads the entire genome enabling the study of noncoding genome sequences as well. The noncoding genome occupies ~98.5% of the genome, and contain the functional transcriptional regulatory elements that decides when and where or in which cell types a gene is expressed. In addition, human-specific regions (human accelerated regions) contained within the noncoding genome regions might be linked to human-specific traits and their disruption might be linked to neurological or cognitive dysfunction[60]. Researchers have found some evidences that noncoding sequence

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

variations may account for ASD although the evidence is still weak compared to the evidence for the coding variants[61-65]. Brandley *et al.* focused the rare structural variants (SVs; e.g. deletions, duplications, insertions and inversions) affecting highly-intolerant cis-regulatory elements[66]. They found some recurrent rare noncoding SVs in ASD cases, such as *Leo1* promoter disrupting variants. In addition, they show intolerant SVs affecting cis-regulatory elements (e.g. transcription start sites, fetal brain promoters and 3'UTRs) were over-transmitted from father to the ASD but not the control sibling. This observation may indicate that these SVs confer risk for ASD. Furthermore, the recent largest WGS of 1902 ASD quartet families provided significant insights in the contribution of *de novo* noncoding SNVs and small indels[67]. First, they assigned annotation categories to *de novo* noncoding variants and found no noncoding category was significantly associated with ASD after correction for multiple testing. However, further analysis using *de novo* risk score developed by machine learning detected a significant contribution of noncoding to ASD risk. In particular, noncoding variants in evolutionary conserved distal promoter regions showed the most robust signal and similar results were observed in a previous WGS study using a different analytic approach [62]. Overall, these studies show a weak but significant contribution of noncoding variants to ASD risk and future analysis of larger cohorts and improving the resources available to annotate of noncoding variants will provide further insights.

Conclusion

To elucidate the etiology of ASD, comprehensive analyses from genetic to translational studies are essential. Recent advances in genetics have provided increasing insights into the complex genetic basis of ASD. The evidence suggests ASD can be caused by genetic defects that include the following: 1) genomic segment CNVs that include micro-deletions, micro-duplications and even higher increases of genomic segment dosage that can involve multiple genes but often highlights a specific gene when assessing the overlap across many ASD cases; 2) strongly penetrance gene coding sequence SNV mutations that cause either a heterozygous (for steeply dose-sensitive genes) or homozygous LOF and in others cases heterozygous GOF mutations; 3) common variants that modify these penetrant CNV and SNV genetic changes; and 4) the possibility of polygenic mechanisms becoming fully penetrant only when two or more genetic changes occur in molecules in a common molecular pathway. The identification of non-coding variants in ASD is just beginning to emerge and the current evidence suggests these may interfere with conserved gene regulatory elements. Analysis of larger number of ASD cases with integrated meta-analysis has helped to prioritize ASD candidate genes or genetic loci into a highly-confident candidate set for further investigation. The identification of potential ASD candidate genes by genetic and computational approaches is still not perfect, but the development of new methods and online resources will continue to increase success in ASD genetics analytics. An important approach to obtaining further evidence that a gene defect found in ASD has an etiologic role in the disorder is the development of functional methods of validating each variant, for example, the highly efficient techniques of genome editing now possible using CRISPR/Cas9 is sure to accelerate the pace of discovery. Ultimately, these genetic defects can then be used to map the specific neuronal circuit defects that underlie behavioral deficits

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in ASD, providing a deeper understanding of ASD pathophysiology and pointing the way to candidate targets for therapeutic intervention.

Acknowledgments

We acknowledge Invitation Fellowship for Research in Japan from Japan Society of Promotion of Science (JSPS), which allows us this collaboration, and the support of KAKENHI (16H06316, 16H06463, 16K13110) from JSPS and Ministry of Education, Culture, Sports, Science, and Technology, Intramural Research Grant for Neurological and Psychiatric Disorders of NCNP, the Takeda Science Foundation and Smoking Research Foundation. This work was also supported by funding to M.P.A. from The National Institute of Mental Health (R01MH112714, R01MH114858, and 1R21MH100868), The National Institute of Neurological Disorders and Stroke (1R01NS08916), The Eunice Kennedy Shriver National Institute of Child Health and Human Development (1R21HD079249), The Nancy Lurie Marks Family Foundation, Landreth Foundation, Autism Speaks/National Alliance for Autism Research, and the Simons Foundation.

References

1. Investigators DDMNSYP, (CDC) CfDCaP: Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2010. MMWR Surveill Summ 2014, 63:1–21.
2. Xu G, Strathearn L, Liu B, Bao W: Prevalence of Autism Spectrum Disorder Among US Children and Adolescents, 2014-2016. JAMA 2018, 319:81–82. [PubMed: 29297068]
3. Jeste SS, Geschwind DH: Disentangling the heterogeneity of autism spectrum disorder through genetic findings. Nat Rev Neurol 2014, 10:74–81. [PubMed: 24468882]
4. Steffenburg S, Gillberg C, Hellgren L, Andersson L, Gillberg IC, Jakobsson G, Bohman M: A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. J Child Psychol Psychiatry 1989, 30:405–416. [PubMed: 2745591]
5. Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M: Autism as a strongly genetic disorder: evidence from a British twin study. Psychol Med 1995, 25:63–77. [PubMed: 7792363]
6. Hallmayer J, Cleveland S, Torres A, Phillips J, Cohen B, Torigoe T, Miller J, Fedele A, Collins J, Smith K, et al.: Genetic heritability and shared environmental factors among twin pairs with autism. Arch Gen Psychiatry 2011, 68:1095–1102. [PubMed: 21727249]
7. Liu X, Takumi T: Genomic and genetic aspects of autism spectrum disorder. Biochem Biophys Res Commun 2014, 452:244–253. [PubMed: 25173933]
8. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, Kou Y, Liu L, Fromer M, Walker S, et al.: Synaptic, transcriptional and chromatin genes disrupted in autism. Nature 2014, 515:209–215. [PubMed: 25363760]
9. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, Murtha MT, Bal VH, Bishop SL, Dong S, et al.: Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. Neuron 2015, 87:1215–1233. [PubMed: 26402605]
10. Takumi T, Tamada K: CNV biology in neurodevelopmental disorders. Curr Opin Neurobiol 2018, 48:183–192. [PubMed: 29331932]
11. Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, Ercan-Sencicek AG, DiLullo NM, Parikshak NN, Stein JL, et al.: De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 2012, 485:237–241. [PubMed: 22495306]
12. O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD, et al.: Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature 2012, 485:246–250. [PubMed: 22495309]
13. Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V, et al.: Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature 2012, 485:242–245. [PubMed: 22495311]
14. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE, et al.: The contribution of de novo coding mutations to autism spectrum disorder. Nature 2014, 515:216–221. [PubMed: 25363768]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

15. Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, et al.: Strong association of de novo copy number mutations with autism. Science 2007, 316:445–449. [PubMed: 17363630]

16. Coe BP, Stessman HAF, Sulovari A, Geisheker MR, Bakken TE, Lake AM, Dougherty JD, Lein ES, Hormozdiari F, Bernier RA, et al.: Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. Nat Genet 2019, 51:106–116. [PubMed: 30559488] ** This is analysis of *de novo* mutation data from 10,927 individuals with DD and ASD. This study identified the candidate neurodevelopmental disease genes needed for further investigation.

17. Takata A, Miyake N, Tsurusaki Y, Fukai R, Miyatake S, Koshimizu E, Kushima I, Okada T, Morikawa M, Uno Y, et al.: Integrative Analyses of De Novo Mutations Provide Deeper Biological Insights into Autism Spectrum Disorder. Cell Rep 2018, 22:734–747. [PubMed: 29346770]

18. Kushima I, Aleksic B, Nakatochi M, Shimamura T, Okada T, Uno Y, Morikawa M, Ishizuka K, Shiino T, Kimura H, et al.: Comparative Analyses of Copy-Number Variation in Autism Spectrum Disorder and Schizophrenia Reveal Etiological Overlap and Biological Insights. Cell Rep 2018, 24:2838–2856. [PubMed: 30208311]

19. Georgi B, Voight BF, Bućan M: From mouse to human: evolutionary genomics analysis of human orthologs of essential genes. PLoS Genet 2013, 9:e1003484. [PubMed: 23675308]

20. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR, et al.: A global reference for human genetic variation. Nature 2015, 526:68–74. [PubMed: 26432245]

21. Kumar P, Henikoff S, Ng PC: Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009, 4:1073–1081. [PubMed: 19561590]

22. Adzhubei I, Jordan DM, Sunyaev SR: Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013, Chapter 7:Unit7.20.

23. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J: A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014, 46:310–315. [PubMed: 24487276]

24. Eilbeck K, Quinlan A, Yandell M: Settling the score: variant prioritization and Mendelian disease. Nat Rev Genet 2017, 18:599–612. [PubMed: 28804138]

25. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB: Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet 2013, 9:e1003709. [PubMed: 23990802]

26. Samocha KE, Robinson EB, Sanders SJ, Stevens C, Sabo A, McGrath LM, Kosmicki JA, Rehnström K, Mallick S, Kirby A, et al.: A framework for the interpretation of de novo mutation in human disease. Nat Genet 2014, 46:944–950. [PubMed: 25086666]

27. Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, Gabriel S, Rieder MJ, Altshuler D, Shendure J, et al.: Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature 2013, 493:216–220. [PubMed: 23201682]

28. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al.: Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016, 536:285–291. [PubMed: 27535533]

29. Bernier R, Golzio C, Xiong B, Stessman HA, Coe BP, Penn O, Witherspoon K, Gerdts J, Baker C, Vulto-van Silfhout AT, et al.: Disruptive CHD8 mutations define a subtype of autism early in development. Cell 2014, 158:263–276. [PubMed: 24998929]

30. Katayama Y, Nishiyama M, Shoji H, Ohkawa Y, Kawamura A, Sato T, Suyama M, Takumi T, Miyakawa T, Nakayama KI: CHD8 haploinsufficiency results in autistic-like phenotypes in mice. Nature 2016, 537:675–679. [PubMed: 27602517]

31. Platt RJ, Zhou Y, Slaymaker IM, Shetty AS, Weisbach NR, Kim JA, Sharma J, Desai M, Sood S, Kempton HR, et al.: Chd8 Mutation Leads to Autistic-like Behaviors and Impaired Striatal Circuits. Cell Rep 2017, 19:335–350. [PubMed: 28402856]

32. Jung H, Park H, Choi Y, Kang H, Lee E, Kweon H, Roh JD, Ellegood J, Choi W, Kang J, et al.: Sexually dimorphic behavior, neuronal activity, and gene expression in Chd8-mutant mice. Nat Neurosci 2018, 21:1218–1228. [PubMed: 30104731]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

33. Kim YJ, Khoshkhoo S, Frankowski JC, Zhu B, Abbasi S, Lee S, Wu YE, Hunt RF: Chd2 Is Necessary for Neural Circuit Development and Long-Term Memory. Neuron 2018, 100:1180–1193.e1186. [PubMed: 30344048]
34. Deliu E, Arecco N, Morandell J, Dotter CP, Contreras X, Girardot C, Käsper EL, Kozlova A, Kishi K, Chiaradia I, et al.: Haploinsufficiency of the intellectual disability gene SETD5 disturbs developmental gene expression and cognition. Nat Neurosci 2018, 21:1717–1727. [PubMed: 30455454]
35. Moore SM, Seidman JS, Ellegood J, Gao R, Savchenko A, Troutman TD, Abe Y, Stender J, Lee D, Wang S, et al.: Setd5 haploinsufficiency alters neuronal network connectivity and leads to autistic-like behaviors in mice. Transl Psychiatry 2019, 9:24. [PubMed: 30655503]
36. Pinggera A, Lieb A, Benedetti B, Lampert M, Monteleone S, Liedl KR, Tuluc P, Striessnig J: CACNA1D de novo mutations in autism spectrum disorders activate Cav1.3 L-type calcium channels. Biol Psychiatry 2015, 77:816–822. [PubMed: 25620733]
37. Pinggera A, Mackenroth L, Rump A, Schallner J, Beleggia F, Wollnik B, Striessnig J: New gain-of-function mutation shows CACNA1D as recurrently mutated gene in autism spectrum disorders and epilepsy. Hum Mol Genet 2017, 26:2923–2932. [PubMed: 28472301]
38. Wang L, Pang K, Han K, Adamski CJ, Wang W, He L, Lai JK, Bondar VV, Duman JG, Richman R, et al.: An autism-linked missense mutation in SHANK3 reveals the modularity of Shank3 function. Mol Psychiatry 2019.* The functional characterization of the *SHANK3* mutation highlighted the importance of missense variants towards a profound understanding of disease pathogenesis and targeted therapies.
39. Bidinosti M, Botta P, Krüttner S, Proenca CC, Stoehr N, Bernhard M, Fruh I, Mueller M, Bonenfant D, Voshol H, et al.: CLK2 inhibition ameliorates autistic features associated with SHANK3 deficiency. Science 2016, 351:1199–1203. [PubMed: 26847545]
40. Vicidomini C, Ponzoni L, Lim D, Schmeisser MJ, Reim D, Morello N, Orellana D, Tozzi A, Durante V, Scalmani P, et al.: Pharmacological enhancement of mGlu5 receptors rescues behavioral deficits in SHANK3 knock-out mice. Mol Psychiatry 2017, 22:784. [PubMed: 27113996]
41. Nakanishi M, Nomura J, Ji X, Tamada K, Arai T, Takahashi E, Bućan M, Takumi T: Functional significance of rare neuroligin 1 variants found in autism. PLoS Genet 2017, 13:e1006940. [PubMed: 28841651]
42. Nakatani J, Tamada K, Hatanaka F, Ise S, Ohta H, Inoue K, Tomonaga S, Watanabe Y, Chung YJ, Banerjee R, et al.: Abnormal behavior in a chromosome-engineered mouse model for human 15q11-13 duplication seen in autism. Cell 2009, 137:1235–1246. [PubMed: 19563756]
43. Nakai N, Nagano M, Saitow F, Watanabe Y, Kawamura Y, Kawamoto A, Tamada K, Mizuma H, Onoe H, Monai H, et al.: Serotonin rebalances cortical tuning and behavior linked to autism symptoms in 15q11-13 CNV mice. Sci Adv 2017, 3:e1603001. [PubMed: 28691086]
44. Tamada K, Tomonaga S, Hatanaka F, Nakai N, Takao K, Miyakawa T, Nakatani J, Takumi T: Decreased exploratory activity in a mouse model of 15q duplication syndrome; implications for disturbance of serotonin signaling. PLoS One 2010, 5:e15126. [PubMed: 21179543]
45. Walsh JJ, Christoffel DJ, Heifets BD, Ben-Dor GA, Selimbeyoglu A, Hung LW, Deisseroth K, Malenka RC: 5-HT release in nucleus accumbens rescues social deficits in mouse autism model. Nature 2018, 560:589–594. [PubMed: 30089910]
46. Krishnan V, Stoppel DC, Nong Y, Johnson MA, Nadler MJ, Ozkaynak E, Teng BL, Nagakura I, Mohammad F, Silva MA, et al.: Autism gene Ube3a and seizures impair sociability by repressing VTA Cbln1. Nature 2017, 543:507–512. [PubMed: 28297715]
47. Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, Zhang H, Estes A, Brune CW, Bradfield JP, et al.: Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. Nature 2009, 459:569–573. [PubMed: 19404257]
48. Furumai R, Tamada K, Liu X, Takumi T: UBE3A regulates the transcription of IRF, an anti-viral immunity. Hum Mol Genet 2019.
49. Anney R, Klei L, Pinto D, Almeida J, Bacchelli E, Baird G, Bolshakova N, Bölte S, Bolton PF, Bourgeron T, et al.: Individual common variants exert weak effects on the risk for autism spectrum disorders. Hum Mol Genet 2012, 21:4781–4792. [PubMed: 22843504]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

50. Klei L, Sanders SJ, Murtha MT, Hus V, Lowe JK, Willsey AJ, Moreno-De-Luca D, Yu TW, Fombonne E, Geschwind D, et al.: Common genetic variants, acting additively, are a major source of risk for autism. Mol Autism 2012, 3:9. [PubMed: 23067556]
51. Gaugler T, Klei L, Sanders SJ, Bodea CA, Goldberg AP, Lee AB, Mahajan M, Manaa D, Pawitan Y, Reichert J, et al.: Most genetic risk for autism resides with common variation. Nat Genet 2014, 46:881–885. [PubMed: 25038753]
52. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, Daly MJ, Price AL, Neale BM, Consortium SWGotPG: LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015, 47:291–295. [PubMed: 25642630]
53. Weiner DJ, Wigdor EM, Ripke S, Walters RK, Kosmicki JA, Grove J, Samocha KE, Goldstein JI, Okbay A, Bybjerg-Grauholm J, et al.: Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. Nat Genet 2017, 49:978–985. [PubMed: 28504703]
54. Niemi MEK, Martin HC, Rice DL, Gallone G, Gordon S, Kelemen M, McAloney K, McRae J, Radford EJ, Yu S, et al.: Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. Nature 2018, 562:268–271. [PubMed: 30258228]
55. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, Duncan L, Perry JR, Patterson N, Robinson EB, et al.: An atlas of genetic correlations across human diseases and traits. Nat Genet 2015, 47:1236–1241. [PubMed: 26414676]
56. Clarke TK, Lupton MK, Fernandez-Pujals AM, Starr J, Davies G, Cox S, Pattie A, Liewald DC, Hall LS, MacIntyre DJ, et al.: Common polygenic risk for autism spectrum disorder (ASD) is associated with cognitive ability in the general population. Mol Psychiatry 2016, 21:419–425. [PubMed: 25754080]
57. Hagenaars SP, Harris SE, Davies G, Hill WD, Liewald DC, Ritchie SJ, Marioni RE, Fawns-Ritchie C, Cullen B, Malik R, et al.: Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. Mol Psychiatry 2016, 21:1624–1632. [PubMed: 26809841]
58. Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, Turley P, Chen GB, Emilsson V, Meddens SF, et al.: Genome-wide association study identifies 74 loci associated with educational attainment. Nature 2016, 533:539–542. [PubMed: 27225129]
59. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, Pallesen J, Agerbo E, Andreassen OA, Anney R, et al.: Identification of common genetic risk variants for autism spectrum disorder. Nat Genet 2019, 51:431–444. [PubMed: 30804558] ** The most recent and largest GWAS study conducted meta-analysis of 18,381 individuals with ASD. It showed genome-wide significant loci in ASD as well as the genetic correlation with other traits which were novel in this study or confirmed the previous findings.
60. Doan RN, Bae BI, Cubelos B, Chang C, Hossain AA, Al-Saad S, Mukaddes NM, Oner O, Al-Saffar M, Balkhy S, et al.: Mutations in Human Accelerated Regions Disrupt Cognition and Social Behavior. Cell 2016, 167:341–354.e312. [PubMed: 27667684]
61. Turner TN, Hormozdiari F, Duyzend MH, McClymont SA, Hook PW, Iossifov I, Raja A, Baker C, Hoekzema K, Stessman HA, et al.: Genome Sequencing of Autism-Affected Families Reveals Disruption of Putative Noncoding Regulatory DNA. Am J Hum Genet 2016, 98:58–74. [PubMed: 26749308]
62. Turner TN, Coe BP, Dickel DE, Hoekzema K, Nelson BJ, Zody MC, Kronenberg ZN, Hormozdiari F, Raja A, Pennacchio LA, et al.: Genomic Patterns of De Novo Mutation in Simplex Autism. Cell 2017, 171:710–722.e712. [PubMed: 28965761]
63. Werling DM, Brand H, An JY, Stone MR, Zhu L, Glessner JT, Collins RL, Dong S, Layer RM, Markenscoff-Papadimitriou E, et al.: An analytical framework for whole-genome sequence association studies and its implications for autism spectrum disorder. Nat Genet 2018, 50:727–736. [PubMed: 29700473]
64. Takata A: Estimating contribution of rare non-coding variants to neuropsychiatric disorders. Psychiatry Clin Neurosci 2019, 73:2–10. [PubMed: 30293238]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

65. Short PJ, McRae JF, Gallone G, Sifrim A, Won H, Geschwind DH, Wright CF, Firth HV, FitzPatrick DR, Barrett JC, et al.: De novo mutations in regulatory elements in neurodevelopmental disorders. Nature 2018, 555:611–616. [PubMed: 29562236]

66. Brandler WM, Antaki D, Gujral M, Kleiber ML, Whitney J, Maile MS, Hong O, Chapman TR, Tan S, Tandon P, et al.: Paternally inherited cis-regulatory structural variants are associated with autism. Science 2018, 360:327–331. [PubMed: 29674594]

67. An JY, Lin K, Zhu L, Werling DM, Dong S, Brand H, Wang HZ, Zhao X, Schwartz GB, Collins RL, et al.: Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder. Science 2018, 362.* This is the most recent study assessing the SNVs and small indels affecting non-coding region in ASD cases. This study first showed genome-wide significant contribution of non-coding variants in ASD.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Key points

- Recent meta-analysis recruiting the largest number of ASD subjects or non-Caucasians identifies novel or promising candidate genes associated with ASD.
- Recent WGS studies suggest the significant contribution of not only coding but also non-coding variants in ASD.
- Functional characterization of GOF and LOF variants associated with ASD leads to a deeper understanding of the pathogenesis of ASD.
- Studies combining mouse models recapitulating genetic features in patients with ASD and recent genomic techniques show the specific defects of neuronal circuits underlying behavioral deficits in ASD and provide the potential therapeutic targets.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Estimating the likely pathogenic variants associated with ASD

Criteria	Explanation	Interpretation
A. Allele frequency in healthy subjects		
1000 genomes project	Database of 2,504 genomes sequenced from healthy subjects	Rare variants are more likely to have larger impact or pathogenic effects compared to common ones. We should take note that the criteria of healthy is varies among databases
Exome sequencing project 6500	Database of 6,503 exomes sequenced from healthy subjects	
ExAc	Datasets of 60,706 exomes sequenced from unrelated healthy subjects	
GnomAD	Datasets of 125,748 WES and 15,708 WGS from unrelated healthy subjects are available	
B. Inheritance pattern		
De novo	Newly arising mutations in patients.	*De novo* variants are more likely to be penetrant compared to inherited ones. The impact of maternally-inherited variants could be underestimated because of the female protective effect in ASD.
Inherited	Mutations inherited from father or mother to patients	
C. Types of variants		
Indel	Small insertions or deletions of bases	Nonsense, stoploss, splicing site mutations and indels are most likely to impact protein function. On the other hand, only subset of missense variants will impact protein function. Synonymous mutations do not alter amino acid sequence or protein function.
Nonsense	Mutations causing protein-truncation	
Stoploss	Mutations disrupting the stop codon resulting in abnormal extention of proteins	
Missense	Mutations causing a change to the amino acid	
Splicing site	Mutations affecting the splicing sites possibly causing mis-splicing	
Synonymous	Mutation which don't alter the amino acid sequence	
D. Genetic intorelance		
pLi	A gene score of the probability of loss-of-function intolerance determined by the number of observed variants and that of expected variants.	Mutations in intorerant genes are more likely to be deleterious
RVIS	A gene intolerance score determined by the number of observed nonsynonymous variants and that of synonymous variants	
E. *in silico* tools to predict the impact of SNVs		
SIFT	A prediction tool of the SNV impact based on the evolutional conservation of the protein's amino acid sequence	These tools score human variants and are usuful to estimate how deleterious a given variants will be to protein function. All of them can be applied to predict the impact of variants with amino-acid substitutions. CADD can be also used for indels.
PolyPhen2	A prediction tool of SNV impact based on the protein sequence and structure.	
CADD	A prediction tool of the impact of SNVs and short indels. It is an integrative metric built from diverse genetic features such as evolutionary constraint, epigenetic status and the score of other prediction tools including SIFT and PolyPhen2.	