

## Genetics of autism spectrum disorder

GOKUL RAMASWAMI<sup>1</sup> AND DANIEL H. GESCHWIND<sup>1,2,3\*</sup>

<sup>1</sup>Program in Neurogenetics, Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, CA, United States

<sup>2</sup>Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California, Los Angeles, CA, United States

<sup>3</sup>Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, United States

### Abstract

Autism spectrum disorder (ASD) is a prevalent neurodevelopmental disorder characterized by impaired social interaction and stereotyped behaviors. ASD has a strong and complex genetic component, with multiple familial inheritance patterns and an estimate of up to 1000 genes potentially implicated. Over the past decade, genomic technologies have enabled rapid progress in the identification of risk genes for ASD. In this chapter, we review the delineation of ASD disease genes starting from traditional genetic studies such as linkage and association, and then focusing on more recent studies utilizing genomic technologies, such as high-throughput genotyping and exome sequencing.

### INTRODUCTION

Autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental disorder mainly characterized by impaired social interaction and restricted and stereotyped behaviors (American Psychiatric Association, 2013). ASD has an incidence of about 1 in 68 individuals and a 3:4 ratio of affected males to females (Elsabbagh et al., 2012). Associated comorbid phenotypes include other neurodevelopmental conditions such as intellectual disability, neurologic disorders such as epilepsy, soft neurologic motor signs, and nonneurologic medical conditions such as gastrointestinal and cardiac problems (Geschwind, 2009; Aldinger et al., 2015). There exists a large degree of clinical heterogeneity and developmental trajectories underlying ASD (Szatmari et al., 2015), even between monozygotic twins (Kolevzon et al., 2004), suggesting roles for both genetic and environmental factors.

ASD has a large heritable component. Infants born to families with an affected older sibling have a 20% chance

of developing ASD (Ozonoff et al., 2011). Concordance rates among monozygotic twins, dizygotic twins, and siblings are 30–99%, 0–65%, and 3–30%, respectively, with an estimated overall heritability of 0.7–0.8 (Bailey et al., 1995; Rosenberg et al., 2009; Hallmayer et al., 2011; Colvert et al., 2015). However, recent population-based studies in Scandinavia show a lower estimated heritability of 0.5–0.6 (Sandin et al., 2014), which is not surprising given the wide confidence intervals in the published heritability estimates based on twins. Additionally, over the last decade the discovery of a significant role for genetic, but nonheritable, de novo mutations in ASD risk has emerged. The contribution from de novo mutations is estimated to be between 15% and 25% (Ronemus et al., 2014), indicating that in total genetic factors account for a large fraction of ASD risk in the population. Environmental factors are thought to play a minor role; however, they have not been well characterized (Newschaffer et al., 2007), with only a few examples such as prenatal exposure to valproate (Christensen et al., 2013) linked to increased risk of ASD.

\*Correspondence to: D.H. Geschwind, 2506A Gonda Bldg, 695 Charles E. Young Dr. South, Los Angeles CA 90095-1761, United States. Tel: +1-310-794-6570, E-mail: dhg@mednet.ucla.edu

Although the overall genetic liability for ASD is high, its genetic architecture is extremely diverse with contributions from alleles of varying frequencies (i.e., common, rare, very rare), inheritance patterns (i.e. dominant, recessive, X-linked, de novo), and variant type (i.e., large chromosomal rearrangements, copy number variants (CNV), small insertions/deletions (indels), and single-nucleotide variants (SNVs)) (De Rubeis and Buxbaum, 2015; de la Torre-Ubieta et al., 2016). Early studies using classic genetic mapping techniques such as linkage, candidate gene screening, and common variant association provided modest results, mainly because they were underpowered to detect causal variants in light of the high genetic heterogeneity in ASD. Over the past decade, advances in DNA sequencing technologies along with the collection of large patient cohorts and widespread data sharing have enabled enormous progress in the identification of genes implicated in ASD (Geschwind and State, 2015). Here, we summarize the current knowledge of genetic variants in the etiology of ASD.

### GENETIC LINKAGE

The first studies on genetic factors in ASD were genetic linkage studies in multiplex families aiming to identify chromosomal regions that are co-inherited by affected individuals (Geschwind and State, 2015). Only a few studies discovered loci that met genome-wide significance thresholds due to the small sample sizes and sparse marker sets used. Furthermore, most of the loci identified were not replicated in subsequent studies (Abrahams and Geschwind, 2008). Only two loci have been successfully replicated: one mapping to chromosome 20p13 (Weiss et al., 2009) and one mapping to chromosome 7q35 (Alarcon et al., 2002) which may be associated with the gene *CNTNAP2* (Alarcon et al., 2008). Linkage studies have also been used to map chromosomal regions associated with ASD endophenotypes. Quantitative phenotypes such as language delay and social responsiveness (Constantino et al., 2003) have been linked to *CNTNAP2* on chromosome 7q35 (Alarcon et al., 2002; Geschwind, 2008) and more recently two genome-wide significant peaks on chromosome 8 (Lowe et al., 2015), respectively.

### COMMON VARIANTS

The contribution of common variants to ASD has been mainly assessed using genome-wide association studies (GWAS). GWAS utilize high-throughput genotyping of common variants, defined as genetic polymorphisms

present in at least 5% of the population. Common variants are estimated to play an important role, accounting for 40–60% of the overall liability of ASD (Klei et al., 2012; Gaugler et al., 2014). However, the effect size of each individual variant is quite small (Anney et al., 2012), necessitating large sample sizes to achieve sufficient detection power to identify specific risk alleles. To date, all of the GWAS studies in ASD have been underpowered, with sample sizes on the order of a few thousand cases (Ma et al., 2009; Wang et al., 2009; Weiss et al., 2009; Anney et al., 2010, 2012; Chaste et al., 2015). In contrast, GWAS of other prevalent diseases such as schizophrenia have sample sizes on the order of 30,000 cases (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). For ASD, only two loci have reached genome-wide significance: intergenic variants between *CDH9* and *CDH10* on chromosome 5p14.1 (Wang et al., 2009) and variants within the *MACROD2* gene on chromosome 20p12.1 (Anney et al., 2010). Neither has been replicated in an independent study. For robust detection of common ASD risk alleles, future GWAS studies will require much larger sample sizes to be sufficiently powered.

### RARE INHERITED ALLELES

Over 100 rare Mendelian syndromes have been associated with ASD (Abrahams and Geschwind, 2008; Betancur, 2011), indicating that, at least in part, ASD is a grouping of rare disorders. In contrast to the strong male bias in idiopathic ASD, syndromic ASD cases have a 1:1 male-to-female ratio (Yoo, 2015). Classic neurodevelopmental syndromes such as fragile X syndrome (caused by mutations in *FMRI*), Rett syndrome (caused by mutations in *MECP2*), tuberous sclerosis (caused by mutations in *TSC1* and *TSC2*), Timothy syndrome (caused by mutations in *CACNA1C*), and dup15q syndrome all display partial comorbidity with ASD. Mutation analysis in multiplex families has been a fruitful strategy to identify genes underlying syndromic ASD. Rare X-linked mutations in two members of the neuroligin family *NLGN3* and *NLGN4* were identified in males diagnosed with ASD and mental retardation from multiple families (Jamain et al., 2003). Using linkage analysis or homozygosity mapping in consanguineous families, rare recessive mutations were identified in *CNTNAP2* from Amish families as well as in *SLC9A9* and *BCKDK* from Middle Eastern families in individuals with ASD and epilepsy (Strauss et al., 2006; Morrow et al., 2008; Novarino et al., 2012). Rare likely damaging inherited variants have also been identified in three members of

a family of scaffolding proteins, *SHANK1*, *SHANK2*, and *SHANK3*, as well as two members of the neurexin family, *NRXN1* and *NRXN3* (Moessner et al., 2007; Kim et al., 2008; Berkel et al., 2010, 2012; Sato et al., 2012; Vaags et al., 2012), indicating a potential role of these genes in ASD.

The emergence of exome sequencing (DNA sequencing of all protein-coding regions of the genome) has enabled high-throughput assessment of protein-disrupting mutations in thousands of individuals. New syndromes associated with ASD continue to be identified via exome sequencing, such as a combination of facial dysmorphism and ASD caused by mutations in *ANDP* (Helsmoortel et al., 2014). Exome sequencing studies of large ASD cohorts have helped elucidate the contribution of inherited alleles to the risk for ASD by identifying an increased rate of rare inherited truncating mutations in genes that are infrequently mutated in the population (Lim et al., 2013; Krumm et al., 2015). Exome sequencing has also been applied to consanguineous families to identify inherited recessive variants in the genes *AMT*, *PEX7*, *SYNE1*, *VPS13B*, *PAH*, and *POMGNT1* (Yu et al., 2013). Future exome and whole-genome sequencing (WGS) of multiplex families, such as those in the Autism Genetic Resource Exchange, will substantially expand knowledge of rare inherited alleles in ASD (Lajonchere and Consortium, 2010).

## DE NOVO VARIATION

De novo variants arising in the parental germline are the underlying cause for many cases of ASD. The first de novo events observed in ASD were large CNVs. Using cytogenetic techniques, large deletions or duplications, such as duplication of 15q (Bundey et al., 1994), deletion of 22q11.2 (Fine et al., 2005), deletion of Xp22.3 (Thomas et al., 1999), and duplication or deletion of 16p11.2 (Weiss et al., 2008), were identified as genetic risk factors for ASD. During the past decade, oligonucleotide arrays for high-throughput genotyping have enabled detection of submicroscopic CNVs. Using array comparative genomic hybridization, Sebat and colleagues (2007) demonstrated that de novo submicroscopic CNVs were strongly associated with ASD. It quickly became apparent that ASD patients harbor significantly more de novo CNVs than control subjects (Levy et al., 2011) and that multiple recurrent de novo CNVs, such as duplication of 7q11.23 and microdeletion of 16p11.2, are prevalent in ASD patients (Kumar et al., 2008; Sanders et al., 2011). De novo CNVs exhibit vast heterogeneity in size – some are small enough to affect a single gene

or large enough to encompass many genes – and clinical presentation; patients with multiple de novo CNVs usually have more severe phenotypes (Marshall et al., 2008; Pinto et al., 2010). More recent studies with larger cohorts, such as the Simons Simplex Collection, a collection of ASD families with a single affected child and usually one or more unaffected siblings, have found that genes within disease-associated de novo CNVs converge upon particular cellular functions such as neuronal signaling, synaptic function, and chromatin remodeling (Pinto et al., 2014; Sanders et al., 2015).

Exome sequencing studies in simplex cohorts, such as the Simons Simplex Collection, have identified numerous de novo variants implicated in ASD. In 2012, four groups published studies using exome sequencing on trios to identify many novel ASD susceptibility genes harboring de novo loss of function or likely gene-disrupting SNVs (Iossifov et al., 2012; Neale et al., 2012; O’Roak et al., 2012b; Sanders et al., 2012). Some of the recurrently hit genes were *DYRK1A*, *POGZ*, *CHD8*, *NTNG1*, *GRIN2B*, *KATNAL2*, and *SCN2A* (Chen et al., 2015). All of these studies found that de novo SNVs have a predominantly paternal origin and the rate of de novo mutation frequency increases with paternal age, concordant with other studies on de novo mutation frequencies (Kong et al., 2012). Subsequent exome sequencing studies have identified a number of genes implicated in ASD (Table 21.1) (De Rubeis et al., 2014; Iossifov et al., 2014). Although the genes identified so far are functionally heterogeneous, many of them are significantly involved in synaptic formation, transcriptional regulation, and chromatin remodeling (De Rubeis et al., 2014; O’Roak et al., 2012a, 2014). It is estimated that there are 400–1000 ASD risk genes (Geschwind and State, 2015).

WGS is starting to be utilized to comprehensively identify all de novo variants in ASD families. In contrast to exome sequencing, WGS can identify mutations in noncoding regulatory elements as well as coding mutations. Although it is currently challenging to interpret the functional impact of noncoding mutations, information such as evolutionary conservation, functional genomics, chromatin state, sequence motifs, and molecular quantitative trait loci is currently being integrated to facilitate annotation of noncoding genomic regions (Ward and Kellis, 2012). Recently, the first study using WGS to identify potentially disease-relevant noncoding variants in simplex ASD families found a significant enrichment of de novo and private disruptive mutations in DNase I hypersensitivity sites located within 50 kb of genes previously implicated in ASD by exome studies (Turner et al., 2016).

Table 21.1

Shared biologic pathways for genes with two or more de novo or rare inherited variants in probands with autism spectrum disorder (ASD)<sup>a</sup>

Biologic process	Genes
Histone modification	<i>KMT2E, KMT2C, WAC, KMT5B, KDM5B, KDM6B, PHF2</i>
Methylation	<i>KMT2E, KMT2C, KMT5B</i>
Demethylation	<i>KDM5B, KDM6B, PHF2</i>
Regulation of gene expression, epigenetic	<i>ARID1B, TNRC6B, PHF2</i>
Wnt signaling pathway	<i>CUL3, TNRC6B, CHD8, TCF7L2</i>
Negative regulation of intracellular signal transduction	<i>CUL3, DYRK1A, SYNGAP1</i>
Multicellular organismal signaling	<i>SCN2A, ANK2, CACNA2D3</i>
Regulation of ion transmembrane transport	<i>SCN2A, ANK2, CACNA2D3</i>
Single-organism behavior	<i>ADNP, TBR1, SYNGAP1, GRIN2B</i>
Cognition	<i>ADNP, TBR1, SYNGAP1, GRIN2B</i>
Rhythmic process	<i>ADNP, KDM5B, DYRK1A</i>
Regulation of cell morphogenesis	<i>RIMS1, ADNP, BCL11A, TBR1, DSCAM, SYNGAP1</i>
Regulation of cellular component size	<i>NCKAP1, ADNP, DSCAM</i>

Modified from Geschwind DH, State MW (2015) Gene hunting in autism spectrum disorder: on the path to precision medicine. *Lancet Neurol* 14: 1109–1120.

<sup>a</sup>The list of genes associated with ASD was compiled from two original research papers that present a comprehensive statistical analysis of essentially all published ASD exomes sequenced to date (De Rubeis et al., 2014; Iossifov et al., 2014), which also includes data from earlier work (O’Roak et al., 2011, 2012b; Iossifov et al., 2012; Neale et al., 2012; Sanders et al., 2012; Willsey et al., 2013; Dong et al., 2014). The list includes genes that harbor two or more rare, de novo splice-site, nonsense, or frameshift variants in ASD probands or those with a false discovery rate of less than 0.05, calculated using the TADA method (De Rubeis et al., 2014), which combines data on rare inherited or de novo damaging missense, splice-site, nonsense, or frameshift variants. Note that KDM5B has an equal count of de novo variants in unaffected siblings (Iossifov et al., 2014). The gene ontology (GO) terms for these genes – standardized descriptions of gene products that indicate associated biologic processes – were obtained using gProfiler (Reimand et al., 2016). Common GO terms were identified as those shared by three or more genes. The resulting list was shortened using REVIGO (Supek et al., 2011) to obtain a representative subset of GO terms using an algorithm that clusters by semantic similarities, avoiding redundancy.

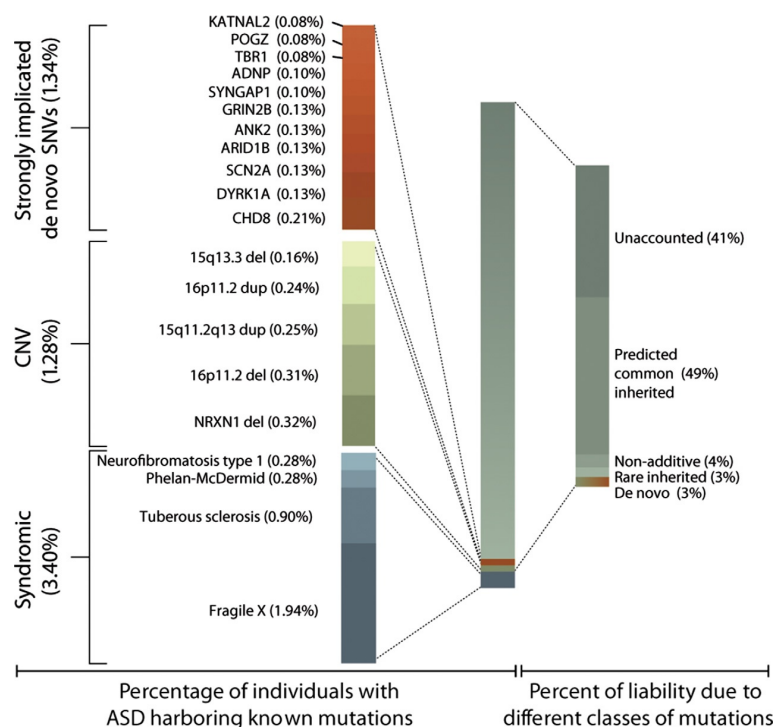
## GENETIC CONTRIBUTIONS TO ASD IN THE POPULATION

The contributing genetic alleles have not been delineated in a majority of patients with ASD. Although hundreds of ASD risk genes have been identified, the contribution of each particular gene within the ASD population is very small, with none found in more than 2% of patients. The largest class of genetic risk for ASD, around 50% of the total, is estimated to derive from common variants of an additive effect, almost all of which have yet to be identified (Gaugler et al., 2014). A further 5–10% of genetic risk is estimated from rare inherited and de novo variants, leaving 40% of the genetic risk currently unaccounted for (Fig. 21.1).

## CLINICAL GENETIC TESTING

Currently, yields for detection of large structural abnormalities by clinical microarray testing in all patients is

5–10% and fragile X testing in males is 2%, making these tests part of the standard of practice (Schaefer et al., 2013; Jeste and Geschwind, 2014; Tammimies et al., 2015; de la Torre-Ubieta et al., 2016). Even higher yields for clinical diagnosis using exome sequencing are expected based on the published research studies (Iossifov et al., 2014) as well as reports of clinical cohorts (Lee et al., 2014). In neurodevelopmental disorders, the yield of exome sequencing is about 30% (Lee et al., 2014), making it among the most powerful diagnostic tests available. Moreover, in particular instances, genetic diagnosis by exome sequencing has directly informed therapeutic strategies. In branched-chain keto acid dehydrogenase kinase deficiency (OMIM 614923), a syndromic form of ASD, identification of inherited recessive truncating mutations in *BCKDK*, a negative regulator of branched-chain amino acid (BCAA) degradation, implicated that a deficiency of BCAAs in these patients contributed to their



**Fig. 21.1.** Genetic contributions to autism spectrum disorder (ASD) population. The percentage of individuals with ASD harboring known mutations, as well as the percentage of liability from different classes of mutations, taken from [Gaugler et al. \(2014\)](#). The percentage variance in liability measures the contribution of a particular variant or class of variants relative to the population variance in a theoretic variable called liability. Liability is a continuous and normally distributed latent variable that represents each individual's risk (both genetic and environmental) for developing a disease. (Reproduced from de la Torre-Ubieta L, Won H, Stein JL, et al. (2016) Advancing the understanding of autism disease mechanisms through genetics. *Nat Med* 22: 345–361.)

diagnosis of ASD. Dietary supplementation of *BCKDK* knockout mice with increased BCAAs led to an amelioration of neurologic phenotypes, suggesting a promising therapy for ASD patients with *BCKDK* mutations ([Novarino et al., 2012](#)). Thus, genetic testing first by microarray, followed by exome sequencing or WGS, should be integrated into the standard of practice for ASD ([Fig. 21.2](#)).

## CONCLUSIONS

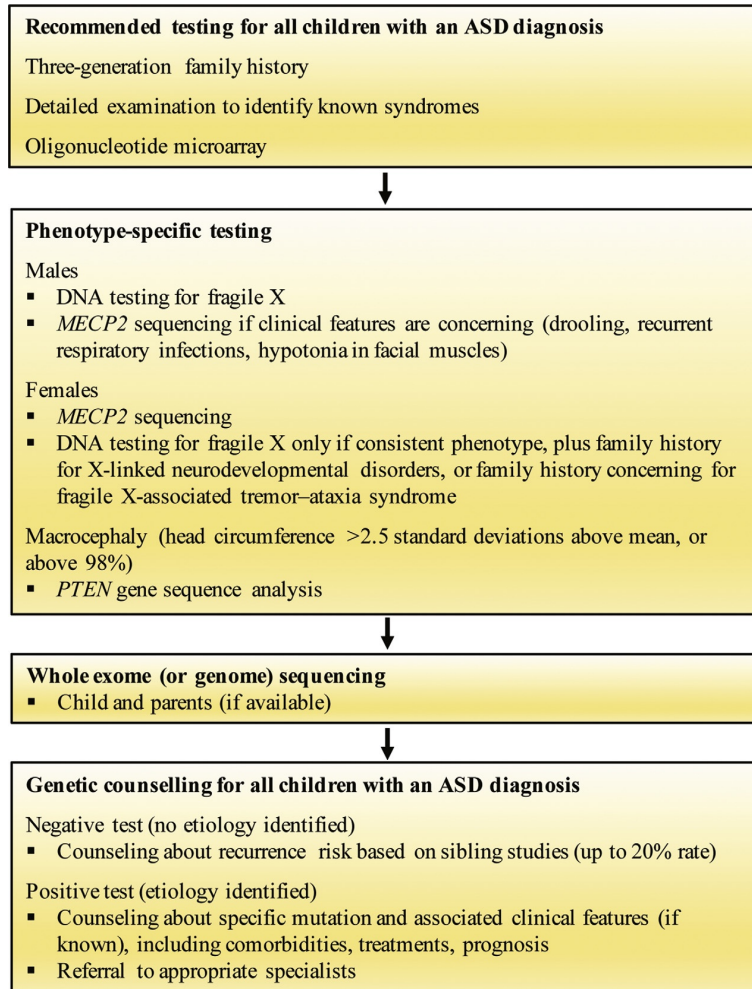
Empowered by genomic technologies and large data-sharing efforts, human genetic studies are making rapid progress in identifying risk genes for ASD. Despite the large clinical and genetic heterogeneity underlying ASD, genotype–phenotype correlations are starting to be identified by clustering patients based

on their genetic background, a so-called genotype-first approach ([Stessman et al., 2014](#)). The integration of genetics with clinical phenotypes and other functional genomic techniques such as transcriptomics and epigenomics will lead to a better understanding of the molecular mechanisms involved in ASD and ultimately inform clinical care.

## ACKNOWLEDGMENTS

We thank Laura Perez-Cano and Elizabeth Ruzzo for critical reading of the manuscript. Funding was provided by grants from National Institutes of Health (5R01 MH100027; 5P50 HD055784-09; 1R01 MH094714-05; 5R01 NS073871-04; 5R01 MH100900-02) and California Institute for Regenerative Medicine (GC1R-0667-A).





**Fig. 21.2.** Recommendations for clinical genetic testing in children with autism spectrum disorder (ASD). Genetic testing in ASD should be undertaken in stepwise fashion to integrate clinical history with relevant tests in an efficient manner. (Modified from Jeste SS, Geschwind DH (2014) Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nat Rev Neurol* 10: 74–81.)

## REFERENCES

- Abrahams BS, Geschwind DH (2008). Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet* 9: 341–355.
- Alarcon M, Cantor RM, Liu J et al. (2002). Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. *Am J Hum Genet* 70: 60–71.
- Alarcon M, Abrahams BS, Stone JL et al. (2008). Linkage, association, and gene-expression analyses identify *CNTNAP2* as an autism-susceptibility gene. *Am J Hum Genet* 82: 150–159.
- Aldinger KA, Lane CJ, Veenstra-VanderWeele J et al. (2015). Patterns of risk for multiple co-occurring medical conditions replicate across distinct cohorts of children with autism spectrum disorder. *Autism Res* 8: 771–781.
- American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders. American Psychiatric Association, Washington, DC.
- Anney R, Klei L, Pinto D et al. (2010). A genome-wide scan for common alleles affecting risk for autism. *Hum Mol Genet* 19: 4072–4082.
- Anney R, Klei L, Pinto D et al. (2012). Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum Mol Genet* 21: 4781–4792.
- Bailey A, Le Couteur A, Gottesman I et al. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* 25: 63–77.
- Berkel S, Marshall CR, Weiss B et al. (2010). Mutations in the *SHANK2* synaptic scaffolding gene in autism spectrum disorder and mental retardation. *Nat Genet* 42: 489–491.

- Berkel S, Tang W, Trevino M et al. (2012). Inherited and de novo SHANK2 variants associated with autism spectrum disorder impair neuronal morphogenesis and physiology. *Hum Mol Genet* 21: 344–357.
- Betancur C (2011). Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res* 1380: 42–77.
- Bunday S, Hardy C, Vickers S et al. (1994). Duplication of the 15q11-13 region in a patient with autism, epilepsy and ataxia. *Dev Med Child Neurol* 36: 736–742.
- Chaste P, Klei L, Sanders SJ et al. (2015). A genome-wide association study of autism using the Simons Simplex Collection: does reducing phenotypic heterogeneity in autism increase genetic homogeneity? *Biol Psychiatry* 77: 775–784.
- Chen JA, Penagarikano O, Belgard TG et al. (2015). The emerging picture of autism spectrum disorder: genetics and pathology. *Annu Rev Pathol* 10: 111–144.
- Christensen J, Gronborg TK, Sorensen MJ et al. (2013). Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA* 309: 1696–1703.
- Colvert E, Tick B, McEwen F et al. (2015). Heritability of autism spectrum disorder in a UK population-based twin sample. *JAMA Psychiatry* 72: 415–423.
- Constantino JN, Davis SA, Todd RD et al. (2003). Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *J Autism Dev Disord* 33: 427–433.
- de la Torre-Ubieta L, Won H, Stein JL et al. (2016). Advancing the understanding of autism disease mechanisms through genetics. *Nat Med* 22: 345–361.
- De Rubeis S, Buxbaum JD (2015). Genetics and genomics of autism spectrum disorder: embracing complexity. *Hum Mol Genet* 24: R24–R31.
- De Rubeis S, He X, Goldberg AP et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* 515: 209–215.
- Dong S, Walker MF, Carriero NJ et al. (2014). De novo insertions and deletions of predominantly paternal origin are associated with autism spectrum disorder. *Cell Rep* 9: 16–23.
- Elsabbagh M, Divan G, Koh YJ et al. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Res* 5: 160–179.
- Fine SE, Weissman A, Gerdes M et al. (2005). Autism spectrum disorders and symptoms in children with molecularly confirmed 22q11.2 deletion syndrome. *J Autism Dev Disord* 35: 461–470.
- Gaugler T, Klei L, Sanders SJ et al. (2014). Most genetic risk for autism resides with common variation. *Nat Genet* 46: 881–885.
- Geschwind DH (2008). Autism: many genes, common pathways? *Cell* 135: 391–395.
- Geschwind DH (2009). Advances in autism. *Annu Rev Med* 60: 367–380.
- Geschwind DH, State MW (2015). Gene hunting in autism spectrum disorder: on the path to precision medicine. *Lancet Neurol* 14: 1109–1120.
- Hallmayer J, Cleveland S, Torres A et al. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry* 68: 1095–1102.
- Helsmoortel C, Vulto-van Silfhout AT, Coe BP et al. (2014). A SWI/SNF-related autism syndrome caused by de novo mutations in ADNP. *Nat Genet* 46: 380–384.
- Iossifov I, Ronemus M, Levy D et al. (2012). De novo gene disruptions in children on the autistic spectrum. *Neuron* 74: 285–299.
- Iossifov I, O’Roak BJ, Sanders SJ et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 515: 216–221.
- Jamain S, Quach H, Betancur C et al. (2003). Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet* 34: 27–29.
- Jeste SS, Geschwind DH (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nat Rev Neurol* 10: 74–81.
- Kim HG, Kishikawa S, Higgins AW et al. (2008). Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet* 82: 199–207.
- Klei L, Sanders SJ, Murtha MT et al. (2012). Common genetic variants, acting additively, are a major source of risk for autism. *Mol Autism* 3: 9.
- Kolevzon A, Smith CJ, Schmeidler J et al. (2004). Familial symptom domains in monozygotic siblings with autism. *Am J Med Genet B Neuropsychiatr Genet* 129B: 76–81.
- Kong A, Frigge ML, Masson G et al. (2012). Rate of de novo mutations and the importance of father’s age to disease risk. *Nature* 488: 471–475.
- Krumm N, Turner TN, Baker C et al. (2015). Excess of rare, inherited truncating mutations in autism. *Nat Genet* 47: 582–588.
- Kumar RA, KaraMohamed S, Sudi J et al. (2008). Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* 17: 628–638.
- Lajonchere CM, Consortium A (2010). Changing the landscape of autism research: the autism genetic resource exchange. *Neuron* 68: 187–191.
- Lee H, Deignan JL, Dorrani N et al. (2014). Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 312: 1880–1887.
- Levy D, Ronemus M, Yamrom B et al. (2011). Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron* 70: 886–897.
- Lim ET, Raychaudhuri S, Sanders SJ et al. (2013). Rare complete knockouts in humans: population distribution and significant role in autism spectrum disorders. *Neuron* 77: 235–242.
- Lowe JK, Werling DM, Constantino JN et al. (2015). Social responsiveness, an autism endophenotype: genome-wide significant linkage to two regions on chromosome 8. *Am J Psychiatry* 172: 266–275.
- Ma D, Salyakina D, Jaworski JM et al. (2009). A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann Hum Genet* 73: 263–273.

- Marshall CR, Noor A, Vincent JB et al. (2008). Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 82: 477–488.
- Moessner R, Marshall CR, Sutcliffe JS et al. (2007). Contribution of SHANK3 mutations to autism spectrum disorder. *Am J Hum Genet* 81: 1289–1297.
- Morrow EM, Yoo SY, Flavell SW et al. (2008). Identifying autism loci and genes by tracing recent shared ancestry. *Science* 321: 218–223.
- Neale BM, Kou Y, Liu L et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485: 242–245.
- Newschaffer CJ, Croen LA, Daniels J et al. (2007). The epidemiology of autism spectrum disorders. *Annu Rev Public Health* 28: 235–258.
- Novarino G, El-Fishawy P, Kayserili H et al. (2012). Mutations in BCKD-kinase lead to a potentially treatable form of autism with epilepsy. *Science* 338: 394–397.
- O’Roak BJ, Deriziotis P, Lee C et al. (2011). Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* 43: 585–589.
- O’Roak BJ, Vives L, Fu W et al. (2012a). Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338: 1619–1622.
- O’Roak BJ, Vives L, Girirajan S et al. (2012b). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485: 246–250.
- O’Roak BJ, Stessman HA, Boyle EA et al. (2014). Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat Commun* 5: 5595.
- Ozonoff S, Young GS, Carter A et al. (2011). Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. *Pediatrics* 128: e488–e495.
- Pinto D, Pagnamenta AT, Klei L et al. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 466: 368–372.
- Pinto D, Delaby E, Merico D et al. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet* 94: 677–694.
- Reimand J, Arak T, Adler P et al. (2016). g:Profiler – a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res* 44: W83–W89.
- Ronemus M, Iossifov I, Levy D et al. (2014). The role of de novo mutations in the genetics of autism spectrum disorders. *Nat Rev Genet* 15: 133–141.
- Rosenberg RE, Law JK, Yenokyan G et al. (2009). Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Arch Pediatr Adolesc Med* 163: 907–914.
- Sanders SJ, Ercan-Sencicek AG, Hus V et al. (2011). Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* 70: 863–885.
- Sanders SJ, Murtha MT, Gupta AR et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485: 237–241.
- Sanders SJ, He X, Willsey AJ et al. (2015). Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron* 87: 1215–1233.
- Sandin S, Lichtenstein P, Kuja-Halkola R et al. (2014). The familial risk of autism. *JAMA* 311: 1770–1777.
- Sato D, Lionel AC, Leblond CS et al. (2012). SHANK1 deletions in males with autism spectrum disorder. *Am J Hum Genet* 90: 879–887.
- Schaefer GB, Mendelsohn NJ, Professional P et al. (2013). Clinical genetics evaluation in identifying the etiology of autism spectrum disorders: 2013 guideline revisions. *Genet Med* 15: 399–407.
- Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511: 421–427.
- Sebat J, Lakshmi B, Malhotra D et al. (2007). Strong association of de novo copy number mutations with autism. *Science* 316: 445–449.
- Stessman HA, Bernier R, Eichler EE (2014). A genotype-first approach to defining the subtypes of a complex disease. *Cell* 156: 872–877.
- Strauss KA, Puffenberger EG, Huentelman MJ et al. (2006). Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. *N Engl J Med* 354: 1370–1377.
- Supek F, Bosnjak M, Skunca N et al. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6: e21800.
- Szatmari P, Georgiades S, Duku E et al. (2015). Developmental trajectories of symptom severity and adaptive functioning in an inception cohort of preschool children with autism spectrum disorder. *JAMA Psychiatry* 72: 276–283.
- Tammimies K, Marshall CR, Walker S et al. (2015). Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. *JAMA* 314: 895–903.
- Thomas NS, Sharp AJ, Browne CE et al. (1999). Xp deletions associated with autism in three females. *Hum Genet* 104: 43–48.
- Turner TN, Hormozdiari F, Duyzend MH et al. (2016). Genome sequencing of autism-affected families reveals disruption of putative noncoding regulatory DNA. *Am J Hum Genet* 98: 58–74.
- Vaags AK, Lionel AC, Sato D et al. (2012). Rare deletions at the neurexin 3 locus in autism spectrum disorder. *Am J Hum Genet* 90: 133–141.
- Wang K, Zhang H, Ma D et al. (2009). Common genetic variants on 5p14.1 associated with autism spectrum disorders. *Nature* 459: 528–533.
- Ward LD, Kellis M (2012). Interpreting noncoding genetic variation in complex traits and human disease. *Nat Biotechnol* 30: 1095–1106.
- Weiss LA, Shen Y, Korn JM et al. (2008). Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 358: 667–675.



- Weiss LA, Arking DE, Gene Discovery Project of Johns H et al. (2009). A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 461: 802–808.
- Willsey AJ, Sanders SJ, Li M et al. (2013). Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* 155: 997–1007.
- Yoo H (2015). Genetics of autism spectrum disorder: current status and possible clinical applications. *Exp Neurobiol* 24: 257–272.
- Yu TW, Chahrour MH, Coulter ME et al. (2013). Using whole-exome sequencing to identify inherited causes of autism. *Neuron* 77: 259–273.