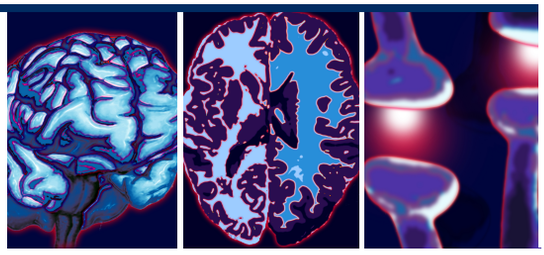
REVIEW



Genetic pathways to autism spectrum disorders

Sunil Q Mehta[1] & Erika L Nurmi*[1]

Practice points

- Autism spectrum disorder (ASD) is a clinically heterogeneous set of illnesses with behavioral deficits in three core domains: language, social communication and cognitive flexibility.
- While there are examples of environmental factors contributing to risk, there is a strong genetic component in the development of ASD.
- Some cases of ASD are part of a larger genetic syndrome ('syndromic' ASD), while the majority of cases are idiopathic.
- Genetic testing is indicated after a clinical diagnosis of ASD in order to help identify whether a case of ASD is syndromic or idiopathic. This testing can allow clinicians to screen for potential medical comorbidities associated with a syndrome and can also have important implications for future family planning.
- Current recommendations from the American Academy of Pediatrics and American College of Medical Genetics include ordering chromosomal microarray testing and fragile X testing for new cases of ASD. It is estimated that between 10 and 20% of cases will have a positive result using chromosomal microarray testing.
- Recent sequencing studies on simplex families with ASD reveal that there may be two different genetic architectures for ASD cases. *De novo* mutations may be responsible for some sporadic cases of ASD, resulting in simplex families (where only one child is affected). Inherited mutations may contribute more to familial ASD, which can result in multiplex families (more than one sibling is affected).
- Challenges remain in determining which mutations contribute to the development of ASD and which mutations comprise normal human variation. Once these challenges are overcome, whole-genome sequencing of patients with ASD will probably replace chromosomal microarrays as the first-line genetic test in ASD.

[1]Division of Child & Adolescent Psychiatry, Department of Psychiatry & Biobehavioral Sciences, UCLA Semel Institute for Neuroscience & Human Behavior, David Geffen School of Medicine, University of California at Los Angeles, CA, USA
*Author for correspondence: enurmi@ucla.edu

Future Medicine part of fsg

SUMMARY Over the past several decades, progress in understanding the genetic basis of autism spectrum disorder (ASD) has dramatically altered our conception of its genetic architecture. Once believed to be an oligogenic disorder of common susceptibility variants, autism is now considered to be a collection of distinct 'autisms' marked by profound genetic heterogeneity. While twin and family studies have demonstrated a strong genetic etiology, genome-wide linkage and association studies have been limited by the extreme underlying heterogeneity. Genome-wide association studies have identified a few variants with small effects on ASD risk, but no common variants that clearly explain the few replicated linkage signals have been identified, suggesting that common variation is unlikely to play a central role. Recent successes in characterizing genetic risk have been driven by technological advances permitting the identification of *de novo* variants, both single-nucleotide variants and copy number variants, occurring in sporadic autism. The power to detect modest, rare inherited effects has been achieved through growing sample sizes through large collaborations; however, the inherited risk of ASD remains largely uncharted. Several hundred risk genes for ASD have been proposed, many linked via shared physiologic pathways. While many investigators now estimate that the number of autism risk genes will reach the thousands, pathway analysis will facilitate the understanding of ASD pathophysiology, the identification of novel risk genes and the development of clinically actionable targets. Molecular diagnosis has become possible for many ASD subtypes and will continue to expand. Targeted interventions will be developed and individualized based on diagnostic data and the growing appreciation of the biology of autism.

Autism spectrum disorder (ASD) encompasses a range of developmental disorders that share dysfunction in specific core domains, including deficits in social interaction, impaired or atypical language development, and restricted or repetitive interests and behaviours [1]. Prevalence estimates have increased by an order of magnitude over the past several decades, and ASD is now estimated to range from 0.6 to 1.1% [2]. A great deal of the increased frequency can be attributed to improved standardization of criteria, and increased recognition and referral to specialized evaluators by families, schools and primary care physicians [3,4]. Additionally, the autism spectrum now encompasses both higher and lower functioning individuals who in the past would proabably not have received an autism diagnosis [5]. Furthermore, as the availability of community and state services for autism increase, greater incentives for referral and diagnosis exist [6]. Finally, cultural factors, such as older average age of parenthood, may also contribute to a true increase in prevalence [7–9]. ASD is more common in males, with a 4:1 ratio, which may be explained by sexual dimorphism in brain development resulting in a higher threshold to manifest an ASD phenotype in females [10,11].

Complex genetic etiology

ASD can be associated with single-gene deficits, such as Rett syndrome, fragile X syndrome and tuberous sclerosis (TSC), or large chromosomal aberrations such as proximal chromosome 15q duplications [12]. It has, therefore, become evident that there are a multitude of genetic pathways to ASD. ASD is believed to represent a collection of disorders with a complex genetic etiology, ranging from one gene of large effect, to multiple genes with additive or interacting effects, to gene–environment interplay. Based on data from family and twin studies [13], it is clear that ASD has a strong genetic contribution. Early twin studies, using a narrow autism definition, found a high concordance among monozygotic (MZ) twins and almost no concordance between dizygotic (DZ) twins; giving a heritability of approximately 90% [14–16]. Larger twin studies using a broader phenotype have since shown lower MZ (47–77%) and higher DZ concordance (14–31%), resulting in a heritability estimate of 38–80% [17,18]. While previous estimates of sibling recurrence risk (the odds of ASD development in a sibling of an affected child) ranged from 3 to 10% [19–21], a recent large-scale, multisite prospective study observed a rate of 20% [22]. Given the similar risk to both siblings and DZ twins, shared *in utero* environmental factors are unlikely to contribute broadly to ASD etiology. In fact, heritability and recurrence risk in ASD ranks among the greatest of any complex genetic disorders.

While relatively high, the imperfect concordance between MZ twins suggests that the environment plays a role in whether or to what

extent ASD risk genes are manifested. In the 1960s, an association was found between autism and maternal viral infection (e.g., rubella) during pregnancy [23,24]. Prenatal exposure to the antiepileptic drug sodium valproate has also been shown to increase ASD risk [25]. Given a population prevalence of 0.8%, the ASD odds ratios (OR) given prenatal rubella infection or valproate exposure would be at least 9.9 and 6.0, respectively, significantly greater than common genetic risk factors. Although these are now extremely rare exposures worldwide, this historical data underscores the importance of the *in utero* environment. A more recent population survey found a mild increase in risk for ASD associated with maternal influenza infection and prolonged fevers during pregnancy [26]. Advanced parental age at conception predicts increased ASD risk in offspring, ranging from odds ratios of 1.3–1.6 for a single older parent to 3.1 when both parents are greater than 40 years old [7–9]. An inverse correlation between ASD and the proximity of maternal address from major highways during pregnancy [27], with the greatest effect (OR: 2.2) during the third trimester, suggests that airborne pollutants may be an environmental risk factor. This study was limited by sampling from a single geographic area (CA, USA) and the use of proximity as a surrogate measure of exposure. Therefore, replication will be important for confirmation. While the potential influence of MMR vaccination has been an enduring concern among parents, in-depth investigation has proven this link without basis [28–30].

ASD is multifactorial and complicated by extreme heterogeneity. As genetic factors are dissected and pathways are revealed, they provide a scaffold upon which to explore environmental factors. Importantly, heritability estimates cannot identify relatively invariant or pervasive environmental factors that may be acting in concert with genetic factors. Undoubtedly, additional modifiers, both positive and negative, will be confirmed and can serve in the development of therapeutic interventions.

Over the past 20 years, advancing technology has enabled the appreciation of a diverse genetic architecture underlying ASD. Two putative pathways to ASD have emerged: a sporadic form caused by microscopic or submicroscopic chromosomal events of large effect, and an inherited form arising from many interacting genes of small individual effects. The various technologies presented below are each suited to exploring certain types of variants and, therefore, have uncovered different pieces of the ASD puzzle.

Structural chromosomal variation & syndromic ASD

Structural variation, rearrangements of large segments of chromosomes, often visible with standard cytogenetic methods, is common (~6–7%) in ASD [31]. The ability to identify and analyze copy number variations (CNVs), the deletion or duplication of chromosomal intervals, has been revolutionized by microarray technology [32]. *De novo* CNVs, those not present in parents but acquired during gamete or zygote development, are an important cause of sporadic ASD (~5–10%) [31,33]. A number of studies have reported recurrent events leading to ASD phenotypes including 7q11.23, 16p11.2 and 15q11–13 **(Table 1)** [34–36]. As large regions spanning many genes are impacted by this type of variation, single causal genes are not necessarily identifiable. In several cases, however, CNV studies have provided clear candidate genes for ASD, including *BZRAP1, MGDA2* [37], *SHANK2, SYNGAP1, DLGAP2* and *PTCHD1* [38]. Pathway analysis performed in a recent investigation [39] confirmed many of the candidate genes previously identified, and implicated pathways involved in synapse formation and dendritic morphogenesis [40]. Although individually these types of variation are rare (none account for more than 1% of all ASD cases), most confer considerable ASD risk **(Table 1)**. ASD-associated CNVs often show variable expressivity, however, and are commonly seen in other neuropsychiatric phenotypes including schizophrenia, intellectual disability without ASD and epilepsy [41]. As studies have failed to observe an excess of transmitted CNVs in ASD, large CNVs are unlikely to play a central role in inherited ASD [39].

Linkage, endophenotypes & quantitative traits

While initial attempts at linkage screens failed to produce signals achieving genome-wide significance, more recent large collaborative studies have identified significant linkage to chromosomes 3q25–27 [42], 2q, 7q [43], 17q11–q21 [44] and 20p13 [45], but only those on 7q and 17q have been replicated [46,47]. Fine mapping efforts have failed to localize single common variants accounting for the linkage signals [48,49]. The

Table 1. Syndromes and structural variation with an autism spectrum disorder-related phenotype.

Chromosomal interval	Cases with ASD (%)	Ref.
1q21 duplication	50	[108,109]
3p deletion/duplication	<50	[110–112]
7q35–q36 (cortical dysplasia focal epilepsy)	70	[113–115]
10q23 (Cowden/BRRS)	20	[116,117]
11p15 maternal deletion (Beckwith–Wiedemann)	~7	[118]
11q13 (Smith–Lemli–Opitz)	50	[119]
12p13 (Timothy)	60–80	[120,121]
15q11–q13 maternal duplication†	>50	[122,123]
15q13 deletion	<50	[124,125]
15q11–q13 maternal deletion (Angelman)	40–80	[126,127]
15q11–q13 paternal deletion (Prader–Willi)	20–25	[128]
16p11.2 deletion/duplication†	55	[31,36,92,93]
17p11 (Potocki–Lupski)	66–90	[129,130]
22q11 deletion (DiGeorge/VCFS)	15–50	[131–133]
22q13 deletion (Phelan-McDermid/*SHANK3*)	>50	[134,135]
21 trisomy (Down's)	6–18	[136,137]
Xq27 (fragile X)†	6 females, 25 males	[138]

†The majority of these abnormalities are found in less than 1% of ASD samples; however, higher rates are observed for fragile X mutation (2–14%) [139], proximal 15q duplication (1–3%) [12] and 16p11.2 deletion (1%) [93].
ASD: Autism spectrum disorder; BBRS: Bannayan–Riley–Ruvalcaba syndrome; VCFS: Velocardiofacial syndrome.

failure of linkage analysis, even in large samples, to find regions of overlap or common variations explaining individual signals is probably due to the heterogeneity present in ASD and the prominent role of rare genetic variation.

The main approaches to tackling the problem of heterogeneity have been to subset subjects into more homogeneous groups based on clinical features or to measure a more quantitative trait or endophenotype than a categorical ASD diagnosis. Subsets of ASD have included the presence of epilepsy, macrocephaly, specific psychiatric comorbitidities, intellectual disability and language delay. Given the male bias in autism, subsetting by gender has been fruitful, identifying a male-specific linkage peak at 17q11 in the Autism Genetic Research Exchange (AGRE) sample that reached genome-wide significance [50] and was subsequently replicated [47]. Suggestive peaks at 7q and 21q were present in AGRE cases with developmental regression [51]. Including only cases with Asperger's disorder revealed suggestive linkage peaks at 5q21 and 15q22 [52]. Subsetting based on the presence of preserved ('splinter') or enhanced ('savant') skills improved linkage to 15q11–q13 [53]. Stratification by IQ in the Autism Genome Project (AGP) dataset did not yield significant evidence for linkage, but did find distinct, mutually exclusive patterns of linkage between high- and low-IQ groups [54]. Regions of interest were highlighted for specific subsets, but no clear candidate loci emerged, perhaps because they generally reduce the sample size examined and thus power.

Quantitative trait loci mapping in ASD may improve the power to localize signals since this method leverages continuous data from all subjects. Instead of considering ASD as a qualitative diagnosis, investigators seek to quantify specific ASD-related traits among unaffected family members and healthy controls [55–58]. The presence of subclinical features of ASD in family members of patients and the existence of a 'broad autism phenotype' argue that ASD falls at the extreme of a continuum [59]. In the past decade, considerable progress has been made in defining those traits, determining their heritability, and developing psychometric instruments to quantitatively measure them (reviewed in [60]). Of ASD-related traits examined, social motivation and range of interest/cognitive flexibility demonstrated the highest heritability [61]. Robinson and colleagues have shown that such traits are stable across childhood [62] and that genetic factors that influence these traits in the general population are likely to be affected in patients with ASD [63]. Recent genome-wide quantitative trait loci scans have used older measures such as IQ [64],

Autism Diagnostic Interview-Revised (ADI-R) subscales [65] and language delay [66] to identify loci of interest, but none have reached genome-wide significance. Wheelwright and colleagues have developed an Autism Spectrum Quotient questionnaire that may be useful to quantitate ASD traits once validated [67]. Genome-wide scans based on quantitative, heritable traits are more likely to be successful than previous studies, which relied on standard collected measures unlikely to stem from a common genetic source, such as age of first word on ADI-R or IQ. Large-scale studies of the segregation of specific cognitive and behavioral traits in ASD and related psychiatric illness are warranted.

Association

■ Candidate gene studies

The failure of linkage studies to uncover risk alleles in complex genetic diseases such as ASD prompted the field to turn to association mapping, which is less sensitive to locus heterogeneity and can detect variants of smaller effect. Early association studies targeted candidate genes identified either by proximity to suggestive linkage peaks or chromosomal rearrangements known to cause ASD (positional candidates) or due to their known biological function relevant to the ASD phenotype (functional candidates). As with most psychiatric disorders, candidate gene studies have reported positive association of ASD phenotypes with numerous genes, although with varying degrees of supporting evidence [32]. As many of these utilized underpowered, multiple-ethnicity samples, tested a biased selection of genes, and failed to correct for multiple testing, the majority of these genes probably represent false positive associations or type I error. As the field has grown to appreciate these issues, genetic studies have turned to large-scale genome-wide approaches with ethnically matched samples. Nevertheless, some of the replicated candidate risk genes may represent true associations, especially as repeatedly identified pathways have emerged **(Table 2)**. Several public databases have compiled the available evidence linking the currently known candidate genes to ASD cases [201].

A popular approach to validating gene candidates into a meaningful ASD model employs gene ontology databases to organize genes into common biological processes, and then construct gene networks based on different models of interaction [38,68]. As the list of ASD candidates is large and since most proteins participate in physiological interactions, virtually all biological processes undertaken by neurons have been implicated. Furthermore, the basis for these network connections are highly variable, ranging from empirically validated interactions to comention of genes in the scientific literature.

Key pathways with multiple ASD candidates include synaptic proteins, DNA methylation and chromatin remodeling factors, splicing and post-transcriptional regulators, and mTOR and MAPK signaling pathways **(Table 2)**. Other areas for further exploration include actin regulators (SYNGAP1, CYFIP1, ARHGEF6 and NHS), microtubule regulators (LIS1, YWHAE and DCX), ion channels and their regulators (SCN1A, PRSS12, CACNA1C, CACNA1F, SLC9A6 and GRIA3), transcription factors (MEF2C, HOXA1, SATB2, FOXP1, FOXG1, RAI1, ARX, NFIX, TBX1, ZNF81 and ZNF674), and proteins involved in vesicular transport (OCRL, AHI1, VPS13B, SYN1

Table 2. Pathways common to autism spectrum disorder-associated genes.

Biochemical pathway or complex	ASD-linked genes†	Ref.
mTOR signaling pathway	(*TSC1; TSC2*); *PTEN*; *NF1*; *MID1*	[140,141]
MAPK/MEK/ERK signaling pathway	*BRAF*; *HRAS*; *KRAS*; *MEK1*; *CREBBP*	[142,143]
Synapse assembly/maintenance complex	(*NRXN1; NLGN3*); *NLGN4; IQSEC2;* (*SHANK2; SHANK3; HOMER; MGLUR5*); *CASK; DMD*	[144–151]
Splicing/post-transcriptional regulators	*AFF2; CDKL5; UPF3B; PQBP1*; (*FMRP; CYPFIP1*); (*WNK3; A2BP1*)	[79,152–158]
DNA methylation/chromatin remodeling factors; mediator complex	*NSD1; JARID1C; PHF8; CHD7; EHMT1;* (*MECP2; ATRX; SMC1A; NIPBL; MED5; MED12*)	[159–165]
Transcription factors	*MEF2C; ENGRAILED2; CUX1; NFIB; HOXA1; SATB2; FOXP1; RAI1; ARX; TBX1; GNB1L*	[160–177]

Genes within parentheses are known to form complexes.
†For references showing that these genes are linked to ASD, see [32].
ASD: Autism spectrum disorder.

and RAB39B). It remains unclear whether these pathways converge on final common 'ASD pathways' [69,70]; however, recent brain expression studies in ASD suggest that this may indeed be the case [71].

Genome-wide association study

The genome-wide association study (GWAS) design provides an unbiased screen of common variation across the genome, but because of the huge number of independent tests performed, it requires large sample sizes to provide the power to correct for multiple comparisons. A stringent correction for genome wide significance of $p < 5 \times 10^8$ is traditionally applied. Several GWASs have been performed in ASD, and while a few loci have met genome-wide significance thresholds, even some with limited follow-up replications, independent GWASs have not produced overlapping results. In 2009, two large, high-density single nucleotide polymorphism (SNP) GWASs were performed, each relying heavily on the AGRE families, but yielding disparate results. Since neither group identified significant loci in the AGRE families alone, both added additional samples to improve power. Wang and colleagues combined data from 780 AGRE families and a case–control cohort totaling 2503 affected individuals and 6491 control subjects [72]. The single genome-wide significant signal mapped to a linkage disequilibrium block in an intergenic region between the cell adhesion genes *CDH9* and *CDH10* [72]. The association was replicated in two independent samples, yet no functional effect on expression of either gene was found. Weiss *et al.* analyzed 1553 affected individuals, combining the 780 AGRE families with 258 families from an NIMH sample [45]. One common SNP upstream of the axon guidance factor *SEMA5A* demonstrated a protective effect (OR: 0.4) against ASD [45]. The protective effect was replicated in an independent sample ($p < 2.5 \times 10^7$) and demonstrated functional effects on *SEMA5A* expression in human cell lines, blood and brain. A third study by the AGP Consortium failed to replicate either group's findings [68], but reported genome-wide significance at an intronic SNP in the *MACROD2* gene [73]. Similar to the outcome of linkage studies, the fruits of these family-based GWAS have been highly sample and method dependent, in contrast to *de novo* CNVs of large effect that have proven more consistent across samples. While the lack of overlap between GWAS in ASD is discouraging, lessons from successful nonpsychiatric complex disease GWAS have taught us that tens to hundreds of thousands of subjects are required to extract the signals of smaller effect that might be seen across datasets. Furthermore, since the microarrays used for GWAS until recently have not included markers capable of detecting signal from variants with low allele frequencies, rare and private mutations are probably missed.

Sequencing

In the past 5 years, the ability to sequence whole genomes or exomes (the protein-coding portion of the genome) has proven scalable and affordable. The challenge that remains is interpreting the vast amount of data generated, especially noncoding variation. Since each genome sequenced reveals over 3 million total variants, many of which are novel [74,75], the task of separating signal from noise is daunting. While 80% of noncoding DNA has been estimated by the ENCODE project to be biologically active [76], the function of these regions remains poorly understood. Since protein-damaging mutations are generally interpretable, early sequencing research has focused on exploring exomes, but even a single exome sequence contains approximately 20,000 total and 100 loss-of-function variants [77]. The lowest hanging fruit are *de novo* variants, with about 74 easily identified new genomic variants per generation [78], and these are especially relevant to simplex ASD families.

Over the past few years, a number of exome sequencing studies in ASD families have been published. Three of these analyzed *de novo* single nucleotide variants or small insertions/deletions (indels) that were present in affected patients from the Simons Simplex collection of ASD families, but not their unaffected parents or siblings [79–81]. Although they performed similar analyses on a partially overlapping set of families, the three studies found deleterious mutations in different sets of genes. Sanders *et al.* reported that one gene, *SCN2A* (a voltage-gated sodium channel subunit associated with epilepsy [81]), harbored mutations frequently enough to consider it an ASD risk gene [81]. While no recurrent deleterious mutations were detected by Iossifov and colleagues, many of the mutations they discovered affected genes encoding proteins that interact with the fragile X mental retardation protein, itself a candidate gene due to the frequency of ASD seen in fragile X syndrome [76]. O'Roak *et al.* identified recurrent

protein-altering mutations in *CHD8* (a chromatin remodeling gene) and *NTNG1* (a cell surface protein involved in axon guidance) [80]. All three studies found that the majority of *de novo* events occurred in DNA of paternal origin and that the frequency of *de novo* mutations was correlated with parental age. This correlation was confirmed in a fourth exome sequencing project by Neale and colleagues, who also revealed recurrent mutations in *CHD8* and *SCN2A* in ASD [82]. A recent meta-analysis combining data from all four papers argues that genes involved in transcriptional regulation and chromatin remodeling play a prominent role in the pathogenesis of ASD [83]. Using statistical modeling of their datasets, the existence of between 350 and 400 or 384 and 821 ASD risk genes were predicted by Iossifov *et al.* [76] and O'Roak *et al.* [80], respectively. If these predictions are correct, it is perhaps not surprising that there was relatively little overlap in identified genes in the three studies, which sequenced only 150–350 families.

Taken together, these studies paint a picture of *de novo* mutations in at least several hundred loci contributing to the genetic risk in *simplex* ASD, each accounting for only a small-to-moderate portion of ASD [80–84]. They suggest that while the overall mutation rate in ASD patients is the same as the general population, ASD patients are far more likely to have deleterious mutations (truncations, frameshift or splice-site mutations) than unaffected family members. However, since exome sequencing only covers approximately 1% of the human genome and ignores critical regions involved in gene regulation (promoters, introns and untranslated regions), these studies do not exclude mutations in regulatory regions as a major factor in the genetic risk for ASD. Although it has been estimated that 85% of disease-causing mutations are in exons [85], this estimate stems from known Mendelian genetic disorders; there are multiple lines of evidence ruling out a Mendelian mode of inheritance for the vast majority of ASD. It is interesting to note that SNPs associated with ASD primarily localize to intergenic regions [45,72,86], reminiscent of the pattern seen with many risk alleles influencing nonpsychiatric complex disease.

Searching for inherited risk alleles in multiplex families with ASD is a far more analytically challenging proposition since *de novo* events are not expected to play a major role in genetic risk in these families. Chahrour *et al.* sequenced the exomes of 16 probands from multiplex families from the AGRE collection that showed excess regions of homozygosity and identity by descent, indicating a shared common ancestor [87]. They discovered a completely different set of genes (expression of which is influenced by neuronal activity) affected by deleterious mutations than the simplex sequencing studies. Although this sample size was much smaller than the simplex exome sequencing studies, it suggests that there may be a difference in the physiological underpinnings and pathogenesis of simplex and multiplex ASD. At present, sequencing studies in ASD have been largely focused on *de novo* variants; more studies applying sequencing to inherited ASD should emerge in the coming years.

Figure 1 illustrates a model for conceptualizing how different risk factors may combine to cause ASD. An assumption in this model is that sporadic ASD is more common in simplex families and inherited ASD risk alleles predominate in multiplex families. The multiplex families segregate a certain amount of inherited genetic risk across generations that will vary by family (unpatterned marbles). During gametogenesis, both simplex and multiplex parents will accrue *de novo* mutations (CNV and SNV marbles) in their sperm and ova, although the mutations in the simplex families will confer larger effects to result in ASD. Finally, during fetal development (and possibly later); environmental factors will also play a role (striped marbles). Gene–gene and gene–environment interactions will also contribute. While in both types of families, an additive or multi-hit risk model is illustrated (empirical evidence for such a model comes from analyses of simplex and multiplex CNVs [88–91]), studying each type of family will probably yield different information about the genetic architecture of ASD.

Clinical implications

While ASD is diagnosed clinically, array-based comparative genomic hybridization and chromosomal microarray (CMA) genetic testing are now available, affordable, and often relevant to clinical and family decision-making. Identification of a known cytogenetic abnormality or gene mutation may provide prognostic information and alert clinicians to any of the medical comorbidities that are commonly associated with syndromic ASD. Although the strict definition of syndromic ASD is a clinical one, many children are evaluated for ASD in toddlerhood, while the full features of some syndromes are not apparent

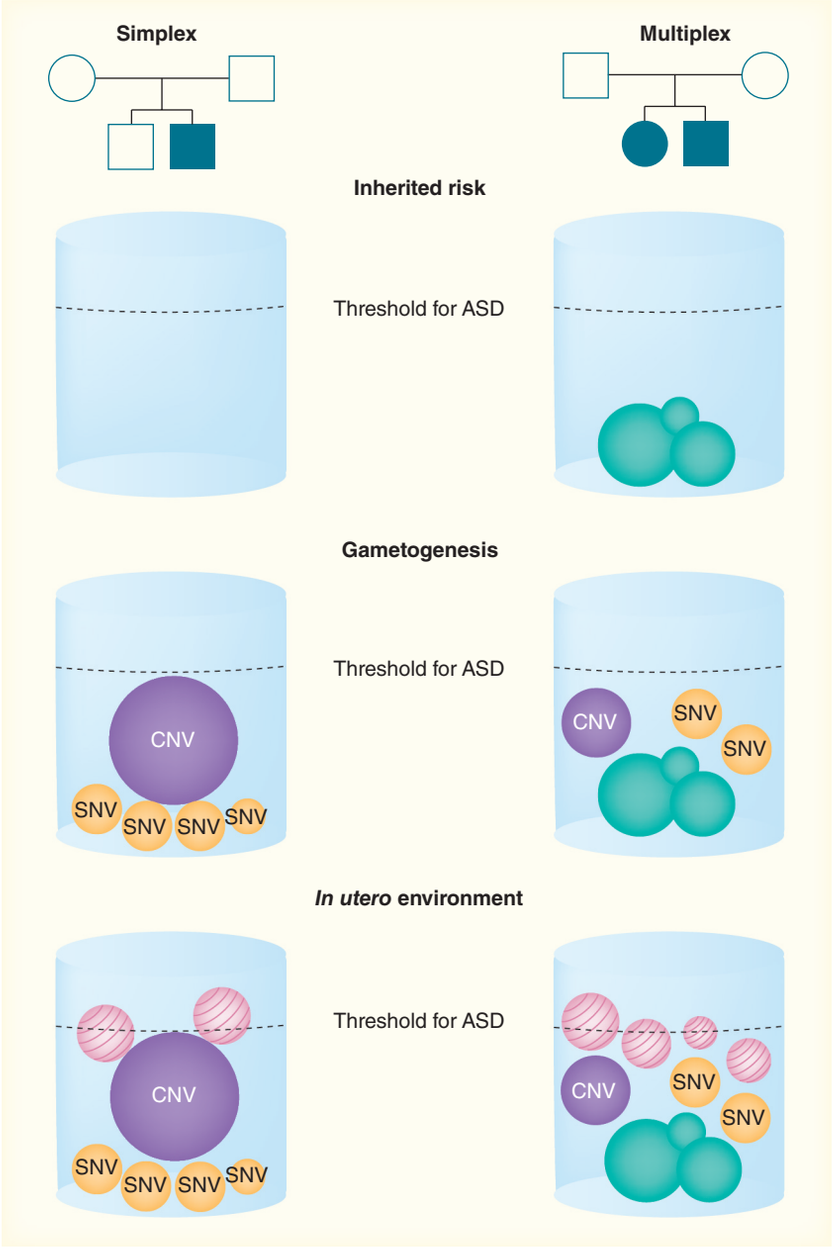


Figure 1. Possible differences in the genetic architecture of simplex and multiplex autism spectrum disorders. Example pedigrees of simplex and multiplex families are given at the top of the figure. Solid shapes represent affected cases while hollow shapes are unaffected family members. Jars represent a hypothetical risk burden that can be tolerated due to resilience until a threshold is reached and ASD manifests. Marbles indicate risk factors that act together to produce ASD when the threshold is exceeded. The unpatterned marbles represent background inherited mutations that are insufficient by themselves to cause ASD. CNV and SNV marbles represent *de novo* risk mutations that arise during gametogenesis. The striped marbles represent environmental effects present in the *in utero* environment or later. The size of the marble indicates relative effect size. Gene–gene and gene–environment interactions, while probably important, are not shown here for the sake of simplicity. ASD: Autism spectrum disorder; CNV: Copy number variable; SNV: Single nucleotide variant.

until late adolescence or adulthood. Identifying a genetic abnormality associated with an ASD-related syndrome can raise clinical suspicion for a syndromic form of ASD, so associated comorbidities of that syndrome are assessed. Some of these comorbidities are highly actionable, including increased frequency of sudden cardiac death observed in patients with 15q11–q13 duplication and of astrocytomas and renal cell tumors associated with TSC. Family planning decisions may be influenced by the profound difference in recurrence risk between *de novo* versus inherited mutations. However, even within inherited forms of ASD, penetrance is highly variable. A salient example is the reduced penetrance for ASD and broader phenotypes observed with proximal 16p duplication versus deletion [36,92–94].

The American College of Medical Genetics and the American Academy of Pediatrics recommend CMA testing as a first-line test in the evaluation of ASD [95,96]. Studies have shown that CMA has a substantially greater resolution and clinical yield than the G-band karyotyping, which was the previous standard of care. Guidelines of the American Academy of Neurology and the American Academy of Child and Adolescent Psychiatry have not yet been revised [97,98]. Between 10 and 20% of patients with ASD will have ASD-related duplications or deletions that can be detected by CMA tests [95]. Limitations of this technology include missed events due to gaps in genomic coverage, the inability to detect balanced chromosomal rearrangements and the fact that many events will be reported with unknown significance. While not all pathological CNVs are known, CNVs are part of the normal genomic landscape and most do not adversely impact health. Diagnostic yield is improved by ordering CMA testing for patients with ASD plus intellectual disability, dysmorphic features or congenital abnormalities [95]. A complete diagnostic work-up recommended by the American Academy of Pediatrics includes an audiological evaluation to rule out hearing impairment as a cause of language delay and a thorough physical examination focusing on possible dysmorphic features, congenital abnormalities and skin pigmentation (revealed by Woods' lamp examination) to help identify syndromic causes of ASD that might particularly benefit from genetic testing [99]. It is important for clinicians to understand that most genetic causes of ASD will not be detected by CMA,

and a normal or negative CMA result does not imply the absence of a genetic etiology.

Clinical sequencing is now being employed throughout medicine; however, CMA testing remains better suited to analyze the types of variation (CNVs) that have large effects in ASD. Although more complete information is garnered from sequencing, the signal-to-noise challenge present with CMAs is magnified in clinical sequencing. As more SNVs conferring risk for ASD are discovered, clinical sequencing will offer increasing value. While commercial enterprises and academic medical centers currently offer genome sequencing to the public, ASD-relevant information is extremely limited.

As the different kinds of ASD risk variants are identified and molecular pathways understood, our ability to provide molecular diagnoses for ASD will expand. This will enable early detection and facilitate long-term outcome studies that can better describe the widely varying prognoses and trajectories of ASD. The targeting of specific interventions can then be tailored based on genetic information. Pharmacological algorithms can be optimized and personalized, and the design of new treatments can be informed by biology. Research on candidate genes and syndromic forms of ASD has led to trials of drugs that target biochemical pathways in those syndromes, such as mTOR inhibitors for TSC [100,101] and mGluR5 antagonists for fragile X syndrome [102]. Within the next decade, clinicians should be able to use genomic data to offer individualized behavioral and pharmacological treatments to many of their patients with ASD.

Conclusion & future perspective

The clinical and genetic heterogeneity of ASD have led to the suggestion that ASD encompasses hundreds of unique autisms [103]. Currently, a few rare *de novo* variants have been solidly established as ASD risk alleles, accounting for only a small proportion of the total genetic risk, although often with a large or causal impact. By contrast, it is likely that common variants confer only small individual effects in ASD, especially given the diminished resulting reproductive success (any variant greatly reducing reproductive success as seen in ASD would remain rare in a population); however, the population-attributable risk may be large given the high population prevalence. Identifying inherited risk alleles of smaller effect is a major goal for the next decade, and will undoubtedly be fueled by ever improving molecular sequencing and computational approaches. Within the next few years, whole genome sequencing will become more economical than exome sequencing, and an even greater challenge will be presented in trying to make sense of noncoding genomic regions. With a greater understanding of the noncoding RNAs and regulatory regions that comprise the majority of human DNA, we can appreciate more subtle ASD risk variants that may be especially important in multiplex families.

The proportion of ASD cases that are simplex versus multiplex is unknown. Although epidemiological studies have placed the sibling recurrence rate at 20% [22], this does not account for families that choose to have only one child. While sequencing families to identify causative *de novo* mutations is productive from a research perspective, it is unlikely to have much diagnostic yield in the clinic, as a clinician will rarely know the genetic architecture of a given child's ASD (particularly when parents bring in a firstborn). Although the cost of obtaining sequence data is rapidly decreasing, the major barrier to using sequence data to help patients with ASD is informatic (i.e., knowing which genetic changes are relevant to the child's condition and which changes are benign genetic variation). As research progresses, we will be able to define an 'ASD network' of genes that capture the majority of idiopathic ASD cases.

Recent genetic studies cement the case that ASD is a complex polygenic disorder and argue that ASD may be due to many possible genetic lesions all leading to a common behavioral phenotype. One interesting pattern that is emerging from these studies is the prevalence of candidates involved in regulating other gene networks (either at the transcriptional or translational level) among the list of ASD risk loci. It is possible that, in order to produce the pathology seen in ASD, many genes must be affected simultaneously, and this can be accomplished either by having multiple mild mutations in neural development genes or having more severe mutations in the genes that regulate the neural development genes. One could imagine that in multiplex families carrying subthreshold mutations in neural development genes, one more *de novo* mutation in a related neural development gene would be enough to cause the offspring to have an ASD. In a simplex family, a single *de novo* mutation in a neural development gene might

not be enough to cause an ASD, but a *de novo* event that impacts a gene network might. As further sequencing results accumulate, it will be fascinating to observe whether regulatory risk genes predominate in simplex families, compared with multiplex families where more subtle variation in genes directly involved in neural development and function might be key.

As genetic pathways are better understood, environmental factors can be systematically assessed. If many genes must be affected in order to cross an ASD threshold, how might environmental factors accomplish this? Ever since the discovery of imprinted syndromes (Angelman, Prader–Willi, maternal proximal chromosome 15q duplication and Beckwith–Wiedemann Syndromes) that have ASD as a component, there has been considerable interest in epigenetic factors in ASD. However, while there have been highly suggestive data in animal models showing a link between *in utero* environment and the methylation of neural development genes (see [104]), as well as human genome studies suggesting that autism risk genes are likely to be regulated by DNA methylation [105], data examining epigenetic changes in cases of ASD versus controls have been inconsistent. Using a histone trap technique, Shulha *et al.* found differences in the methylation of neural development genes in the prefrontal cortex of patients with ASD compared with controls [106]. However, using a whole-genome methylation microarray, Ginsberg *et al.* found no difference between cases and controls in the occipital cortex [107]. This inconsistency can be explained by the different experimental techniques or brain regions studied, which further underscores the challenge of posing epigenetic questions. Epigenetic studies rely on post-mortem brain tissue, which necessitates small sample sizes and introduces variability in collection, storage and treatment of tissue. While the development of standardized post-mortem tissue collections will improve the reliability of epigenetic data, access to tissue is an obstacle to epigenetic studies of ASD. All of the current sequencing studies have used whole blood from patients and families as the source of DNA. However, the level of genetic or epigenetic mosaicism in the brains of patients with ASD is unknown. Solving the technical hurdles associated with epigenetic studies will allow investigation into the spatial and temporal regulation of genes in the brain that may play a crucial role in the development of ASD. While our current appreciation of the intricacies of epigenetic regulation is rudimentary, it is clear that this mechanism provides a point of communication between the genes and environment that together underlie risk for ASD.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as:
- ■ of interest
- ■■ of considerable interest

1 American Psychiatric Association. *Diagnostic and Statistical Manual Of Mental Disorders (Revised 4th Edition)*. American Psychiatric Association, Washington, DC, USA (2000).

2 Autism and Developmental Disabilities Monitoring Network, Centers for Disease Control. Prevalence of autism spectrum disorders – Autism And Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. *MMWR Surveill. Summ.* 61(3), 1–19 (2012).

3 Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators, Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders – Autism and Developmental Disabilities Monitoring Network, United States, 2006. *MMWR Surveill. Summ.* 58(10), 1–20 (2009).

4 Fombonne E, Quirke S, Hagen A. Prevalence and interpretation of recent trends in rates of pervasive developmental disorders. *McGill J. Med.* 12(2), 73 (2009).

5 Croen LA, Grether JK, Hoogstrate J, Selvin S. The changing prevalence of autism in California. *J. Autism Dev. Disord.* 32(3), 207–215 (2002).

6 Hertz-Picciotto I, Delwiche L. The rise in autism and the role of age at diagnosis. *Epidemiology* 20(1), 84–90 (2009).

7 Croen LA, Najjar DV, Fireman B, Grether JK. Maternal and paternal age and risk of autism spectrum disorders. *Arch. Pediatr. Adolesc. Med.* 161(4), 334–340 (2007).

8 Durkin MS, Maenner MJ, Newschaffer CJ *et al.* Advanced parental age and the risk of autism spectrum disorder. *Am. J. Epidemiol.* 168(11), 1268–1276 (2008).

9 Shelton JF, Tancredi DJ, Hertz-Picciotto I. Independent and dependent contributions of advanced maternal and paternal ages to autism risk. *Autism Res.* 3(1), 30–39 (2010).

10 Baron-Cohen S, Knickmeyer RC, Belmonte MK. Sex differences in the brain: implications for explaining autism. *Science* 310(5749), 819–823 (2005).

■ **Good general review of twin studies in autism.**

11 Baron-Cohen S, Lombardo MV, Auyeung B, Ashwin E, Chakrabarti B, Knickmeyer R. Why are autism spectrum conditions more prevalent in males? *PLoS Biol.* 9(6), e1001081 (2011).

12 Sutcliffe JS, Nurmi EL, Lombroso PJ. Genetics of childhood disorders: XLVII. Autism, part 6: duplication and inherited susceptibility of chromosome 15q11–q13 genes in autism. *J. Am. Acad. Child Adolesc. Psychiatry* 42(2), 253–256 (2003).

13 Ronald A, Hoekstra RA. Autism spectrum disorders and autistic traits: a decade of new twin studies. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 156B(3), 255–274 (2011).

14 Bailey A, Le Couteur A, Gottesman I *et al.* Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol. Med.* 25(1), 63–77 (1995).

- **Article that reached different conclusions from previous twin studies and generated some controversy in the field.**

15 Ritvo ER, Freeman BJ, Mason-Brothers A, Mo A, Ritvo AM. Concordance for the syndrome of autism in 40 pairs of afflicted twins. *Am. J. Psychiatry* 142(1), 74–77 (1985).

16 Steffenburg S, Gillberg C, Hellgren L *et al.* A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. *J. Child Psychol. Psychiatry* 30(3), 405–416 (1989).

17 Hallmayer J, Cleveland S, Torres A *et al.* Genetic heritability and shared environmental factors among twin pairs with autism. *Arch. Gen. Psychiatry* 68(11), 1095–1102 (2011).

18 Lichtenstein P, Carlstrom E, Rastam M, Gillberg C, Anckarsater H. The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. *Am. J. Psychiatry* 167(11), 1357–1363 (2010).

19 Chakrabarti S, Fombonne E. Pervasive developmental disorders in preschool children: confirmation of high prevalence. *Am. J. Psychiatry* 162(6), 1133–1141 (2005).

- **Describes a large-scale study that significantly changed the genetic counseling that parents receive and our thinking about at-risk siblings of children with autism spectrum disorder (ASD).**

20 Icasiano F, Hewson P, Machet P, Cooper C, Marshall A. Childhood autism spectrum disorder in the Barwon region: a community based study. *J. Paediatr. Child Health* 40(12), 696–701 (2004).

21 Lauritsen MB, Pedersen CB, Mortensen PB. Effects of familial risk factors and place of birth on the risk of autism: a nationwide register-based study. *J. Child Psychol. Psychiatry* 46(9), 963–971 (2005).

22 Ozonoff S, Young GS, Carter A *et al.* Recurrence risk for autism spectrum disorders: a baby siblings research consortium study. *Pediatrics* doi:10.1542/peds.2010-2825 (2011).

23 Chess S. Autism in children with congenital rubella. *J. Autism Child Schizophr.* 1(1), 33–47 (1971).

24 Deykin EY, Macmahon B. Viral exposure and autism. *Am. J. Epidemiol.* 109(6), 628–638 (1979).

25 Rasalam AD, Hailey H, Williams JH *et al.* Characteristics of fetal anticonvulsant syndrome associated autistic disorder. *Dev. Med. Child Neurol.* 47(8), 551–555 (2005).

26 Atladottir HO, Henriksen TB, Schendel DE, Parner ET. Autism after infection, febrile episodes, and antibiotic use during pregnancy: an exploratory study. *Pediatrics* 130(6), e1447–e1454 (2012).

- **Excellent review of the controversy regarding ASD and vaccinations and how it was settled.**

27 Volk HE, Hertz-Picciotto I, Delwiche L, Lurmann F, McConnell R. Residential proximity to freeways and autism in the CHARGE Study. *Environ. Health Perspect.* 119(6), 873–877 (2011).

28 Baird G, Pickles A, Simonoff E *et al.* Measles vaccination and antibody response in autism spectrum disorders. *Arch. Dis. Child* 93(10), 832–837 (2008).

29 Deer B. How the case against the MMR vaccine was fixed. *BMJ* 342, c5347 (2011).

30 Uno Y, Uchiyama T, Kurosawa M, Aleksic B, Ozaki N. The combined measles, mumps, and rubella vaccines and the total number of vaccines are not associated with development of autism spectrum disorder: the first case-control study in Asia. *Vaccine* 30(28), 4292–4298 (2012).

31 Marshall CR, Noor A, Vincent JB *et al.* Structural variation of chromosomes in autism spectrum disorder. *Am. J. Hum. Genet.* 82(2), 477–488 (2008).

32 Betancur C. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res.* 1380, 42–77 (2011).

33 Sebat J, Lakshmi B, Malhotra D *et al.* Strong association of *de novo* copy number mutations with autism. *Science* 316(5823), 445–449 (2007).

34 Sanders SJ, Ercan-Sencicek AG, Hus V *et al.* Multiple recurrent *de novo* CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* 70(5), 863–885 (2011).

- **A very thorough analysis of copy number variations in ASD.**

35 Fischbach GD, Lord C. The Simons Simplex Collection: a resource for identification of autism genetic risk factors. *Neuron* 68(2), 192–195 (2010).

36 Kumar RA, Karamohamed S, Sudi J *et al.* Recurrent 16p11.2 microdeletions in autism. *Hum. Mol. Genet.* 17(4), 628–638 (2008).

37 Bucan M, Abrahams BS, Wang K *et al.* Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genet.* 5(6), e1000536 (2009).

38 Pinto D, Pagnamenta AT, Klei L *et al.* Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 466(7304), 368–372 (2010).

39 Levy D, Ronemus M, Yamrom B *et al.* Rare *de novo* and transmitted copy-number variation in autistic spectrum disorders. *Neuron* 70(5), 886–897 (2011).

40 Gilman SR, Iossifov I, Levy D, Ronemus M, Wigler M, Vitkup D. Rare *de novo* variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron* 70(5), 898–907 (2011).

41 Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu. Rev. Med.* 61, 437–455 (2010).

42 Auranen M, Vanhala R, Varilo T *et al.* A genomewide screen for autism-spectrum disorders: evidence for a major susceptibility locus on chromosome 3q25–27. *Am. J. Hum. Genet.* 71(4), 777–790 (2002).

43 International Molecular Genetic Study of Autism Consortium. A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am. J. Hum. Genet.* 69(3), 570–581 (2001).

44 Yonan AL, Alarcon M, Cheng R *et al.* A genomewide screen of 345 families for autism-susceptibility loci. *Am. J. Hum. Genet.* 73(4), 886–897 (2003).

45 Weiss LA, Arking DE, Daly MJ, Chakravarti A. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 461(7265), 802–808 (2009).

46 Badner JA, Gershon ES. Regional meta-analysis of published data supports linkage of autism with markers on chromosome 7. *Mol. Psychiatry* 7(1), 56–66 (2002).

47 Cantor RM, Kono N, Duvall JA *et al.* Replication of autism linkage: fine-mapping peak at 17q21. *Am. J. Hum. Genet.* 76(6), 1050–1056 (2005).

48 Stone JL, Merriman B, Cantor RM, Geschwind DH, Nelson SF. High density SNP association study of a major autism linkage region on chromosome 17. *Hum. Mol. Genet.* 16(6), 704–715 (2007).

49 Strom SP, Stone JL, Ten Bosch JR *et al.* High-density SNP association study of the 17q21 chromosomal region linked to autism identifies CACNA1G as a novel candidate gene. *Mol. Psychiatry* 15(10), 996–1005 (2010).

50 Stone JL, Merriman B, Cantor RM *et al.* Evidence for sex-specific risk alleles in autism spectrum disorder. *Am. J. Hum. Genet.* 75(6), 1117–1123 (2004).

51 Molloy CA, Keddache M, Martin LJ. Evidence for linkage on 21q and 7q in a subset of autism characterized by developmental regression. *Mol. Psychiatry* 10(8), 741–746 (2005).

52 Salyakina D, Ma DQ, Jaworski JM *et al.* Variants in several genomic regions associated with asperger disorder. *Autism Res.* 3(6), 303–310 (2010).

53 Nurmi EL, Dowd M, Tadevosyan-Leyfer O, Haines JL, Folstein SE, Sutcliffe JS. Exploratory subsetting of autism families based on savant skills improves evidence of genetic linkage to 15q11-q13. *J. Am. Acad. Child Adolesc. Psychiatry* 42(7), 856–863 (2003).

54 Vieland VJ, Hallmayer J, Huang Y *et al.* Novel method for combined linkage and genome-wide association analysis finds evidence of distinct genetic architecture for two subtypes of autism. *J. Neurodev. Disord.* 3(2), 113–123 (2011).

55 Chasson GS, Timpano KR, Greenberg JL, Shaw A, Singer T, Wilhelm S. Shared social competence impairment: another link between the obsessive–compulsive and autism spectrums? *Clin. Psychol. Rev.* 31(4), 653–662 (2011).

56 Delorme R, Gousse V, Roy I *et al.* Shared executive dysfunctions in unaffected relatives of patients with autism and obsessive–compulsive disorder. *Eur. Psychiatry* 22(1), 32–38 (2007).

57 Rommelse NN, Geurts HM, Franke B, Buitelaar JK, Hartman CA. A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neurosci. Biobehav. Rev.* 35(6), 1363–1396 (2011).

58 Scott-Van Zeeland AA, Abrahams BS, Alvarez-Retuerto AI *et al.* Altered functional connectivity in frontal lobe circuits is associated with variation in the autism risk gene CNTNAP2. *Sci. Transl. Med.* 2(56), 56ra80 (2010).

59 Losh M, Adolphs R, Poe MD *et al.* Neuropsychological profile of autism and the broad autism phenotype. *Arch. Gen. Psychiatry* 66(5), 518–526 (2009).

60 Constantino JN. The quantitative nature of autistic social impairment. *Pediatr. Res.* 69(5 Pt 2), 55R–62R (2011).

61 Sung YJ, Dawson G, Munson J, Estes A, Schellenberg GD, Wijsman EM. Genetic investigation of quantitative traits related to autism: use of multivariate polygenic models with ascertainment adjustment. *Am. J. Hum. Genet.* 76(1), 68–81 (2005).

62 Robinson EB, Munir K, Munafo MR, Hughes M, Mccormick MC, Koenen KC. Stability of autistic traits in the general population: further evidence for a continuum of impairment. *J. Am. Acad. Child Adolesc. Psychiatry* 50(4), 376–384 (2011).

63 Robinson EB, Koenen KC, Mccormick MC *et al.* Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Arch. Gen. Psychiatry* 68(11), 1113–1121 (2011).

64 Chapman NH, Estes A, Munson J *et al.* Genome-scan for IQ discrepancy in autism: evidence for loci on chromosomes 10 and 16. *Hum. Genet.* 129(1), 59–70 (2011).

65 Liu XQ, Georgiades S, Duku E *et al.* Identification of genetic loci underlying the phenotypic constructs of autism spectrum disorders. *J. Am. Acad. Child Adolesc. Psychiatry* 50(7), 687e613–696e613 (2011).

66 Cho SC, Yoo HJ, Park M *et al.* Genome-wide association scan of korean autism spectrum disorders with language delay: a preliminary study. *Psychiatry Investig.* 8(1), 61–66 (2011).

67 Wheelwright S, Auyeung B, Allison C, Baron-Cohen S. Defining the broader, medium and narrow autism phenotype among parents using the Autism Spectrum Quotient (AQ). *Mol. Autism* 1(1), 10 (2010).

68 Anney RJ, Kenny EM, O'Dushlaine C *et al.* Gene-ontology enrichment analysis in two independent family-based samples highlights biologically plausible processes for autism spectrum disorders. *Eur. J. Hum. Genet.* 19(10), 1082–1089 (2011).

69 Bill BR, Geschwind DH. Genetic advances in autism: heterogeneity and convergence on shared pathways. *Curr. Opin. Genet. Dev.* 19(3), 271–278 (2009).

70 Geschwind DH. Autism: many genes, common pathways? *Cell* 135(3), 391–395 (2008).

71 Voineagu I, Wang X, Johnston P *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* 474(7351), 380–384 (2011).

72 Wang K, Zhang H, Ma D *et al.* Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* 459(7246), 528–533 (2009).

73 Anney R, Klei L, Pinto D *et al.* A genome-wide scan for common alleles affecting risk for autism. *Hum. Mol. Genet.* 19(20), 4072–4082 (2010).

■ **Directly compares the diagnostic yields of karyotyping and chromosomal microarrays.**

74 Pelak K, Shianna KV, Ge D *et al.* The characterization of twenty sequenced human genomes. *PLoS Genet.* 6(9), e1001111. (2010).

75 Abecasis GR, Auton A *et al.*; The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422), 56–65 (2012).

76 Iossifov I, Ronemus M, Levy D *et al. De novo* gene disruptions in children on the autistic spectrum. *Neuron* 74(2), 285–299 (2012).

77 Pennisi E. Genomics. ENCODE project writes eulogy for junk DNA. *Science* 337(6099), 1159–1161 (2012).

78 Macarthur DG, Tyler-Smith C. Loss-of-function variants in the genomes of healthy humans. *Hum. Mol. Genet.* 19(R2), R125–R130 (2010).

79 Kong A, Frigge ML, Masson G *et al.* Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* 488(7412), 471–475 (2012).

80 O'Roak BJ, Vives L, Girirajan S *et al.* Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* 485(7397), 246–250 (2012).

81 Sanders SJ, Murtha MT, Gupta AR *et al.* *De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485(7397), 237–241 (2012).

82 Neale BM, Kou Y, Liu L *et al.* Patterns and rates of exonic *de novo* mutations in autism spectrum disorders. *Nature* 485(7397), 242–245 (2012).

83 Ben-David E, Shifman S. Combined analysis of exome sequencing points toward a major role for transcription regulation during brain development in autism. *Mol. Psychiatry* doi:10.1038/mp.2012.148 (2012) (Epub ahead of print).

84 Shi X, Yasumoto S, Kurahashi H *et al.* Clinical spectrum of *SCN2A* mutations. *Brain Dev.* 34(7), 541–545 (2012).

85 Choi M, Scholl UI, Ji W *et al.* Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc.*

Natl Acad. Sci. USA 106(45), 19096–19101 (2009).

86 Hu VW, Addington A, Hyman A. Novel autism subtype-dependent genetic variants are revealed by quantitative trait and subphenotype association analyses of published GWAS data. *PLoS One* 6(4), e19067 (2011).

87 Chahrour MH, Yu TW, Lim ET *et al.* Whole-exome sequencing and homozygosity analysis implicate depolarization-regulated neuronal genes in autism. *PLoS Genet.* 8(4), e1002635 (2012).

88 Girirajan S, Rosenfeld JA, Cooper GM *et al.* A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat. Genet.* 42(3), 203–209 (2010).

89 Klei L, Sanders SJ, Murtha MT *et al.* Common genetic variants, acting additively, are a major source of risk for autism. *Mol. Autism* 3(1), 9 (2012).

90 Leblond CS, Heinrich J, Delorme R *et al.* Genetic and functional analyses of *SHANK2* mutations suggest a multiple hit model of autism spectrum disorders. *PLoS Genet.* 8(2), e1002521 (2012).

91 Schaaf CP, Sabo A, Sakai Y *et al.* Oligogenic heterozygosity in individuals with high-functioning autism spectrum disorders. *Hum. Mol. Genet.* 20(17), 3366–3375 (2011).

92 Hanson E, Nasir RH, Fong A *et al.* Cognitive and behavioral characterization of 16p11.2 deletion syndrome. *J. Dev. Behav. Pediatr.* 31(8), 649–657 (2010).

93 Weiss LA, Shen Y, Korn JM *et al.* Association between microdeletion and microduplication at 16p11.2 and autism. *N. Engl. J. Med.* 358(7), 667–675 (2008).

94 Mccarthy SE, Makarov V, Kirov G *et al.* Microduplications of 16p11.2 are associated with schizophrenia. *Nat. Genet.* 41(11), 1223–1227 (2009).

95 Miller DT, Adam MP, Aradhya S *et al.* Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* 86(5), 749–764 (2010).

96 American Academy of Pediatrics. *Autism Caring for Children With Autism Spectrum Disorder.A Resource Toolkit for Clinicians (2nd Edition).* American Academy of Pediatrics, IL, USA (2013).

97 Mcgrew SG, Peters BR, Crittendon JA, Veenstra-Vanderweele J. Diagnostic yield of chromosomal microarray analysis in an autism primary care practice: which guidelines to implement? *J. Autism Dev. Disord.* 42(8), 1582–1591 (2012).

98 Shen Y, Dies KA, Holm IA *et al.* Clinical genetic testing for patients with autism spectrum disorders. *Pediatrics* 125(4), e727–e735 (2010).

99 Johnson CP, Myers SM. Identification and evaluation of children with autism spectrum disorders. *Pediatrics* 120(5), 1183–1215 (2007).

100 Bissler JJ, Mccormack FX, Young LR *et al.* Sirolimus for angiomyolipoma in tuberous sclerosis complex or lymphangioleiomyomatosis. *N. Engl. J. Med.* 358(2), 140–151 (2008).

101 Franz DN, Belousova E, Sparagana S *et al.* Efficacy and safety of everolimus for subependymal giant cell astrocytomas associated with tuberous sclerosis complex (EXIST-1): a multicentre, randomised, placebo-controlled Phase 3 trial. *Lancet* 381(9861), 125–132 (2013).

102 Berry-Kravis E, Hessl D, Coffey S *et al.* A pilot open label, single dose trial of fenobam in adults with fragile X syndrome. *J. Med. Genet.* 46(4), 266–271 (2009).

103 Geschwind DH. Advances in autism. *Annu. Rev. Med.* 60, 367–380 (2009).

104 Grissom NM, Reyes TM. Gestational overgrowth and undergrowth affect neurodevelopment: similarities and differences from behavior to epigenetics. *Int. J. Dev. Neurosci.*, (2012).

105 Wang T, Pan Q, Lin L *et al.* Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. *Hum. Mol. Genet.* 21(26), 5500–5510 (2012).

106 Shulha HP, Cheung I, Whittle C *et al.* Epigenetic signatures of autism: trimethylated H3K4 landscapes in prefrontal neurons. *Arch. Gen. Psychiatry* 69(3), 314–324 (2012).

107 Ginsberg MR, Rubin RA, Falcone T, Ting AH, Natowicz MR. Brain transcriptional and epigenetic associations with autism. *PLoS ONE* 7(9), e44736 (2012).

108 Mefford HC, Sharp AJ, Baker C *et al.* Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N. Engl. J. Med.*359(16), 1685–1699 (2008).

109 Szatmari P, Paterson AD, Zwaigenbaum L *et al.* Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat. Genet.* 39(3), 319–328 (2007).

110 Fernandez T, Morgan T, Davis N *et al.* Disruption of contactin 4 (*CNTN4*) results in developmental delay and other features of 3p deletion syndrome. *Am. J. Hum. Genet.* 74(6), 1286–1293 (2004).

111 Glessner JT, Wang K, Cai G *et al.* Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 459(7246), 569–573 (2009).

112 Roohi J, Montagna C, Tegay DH *et al.* Disruption of contactin 4 in three subjects with autism spectrum disorder. *J. Med. Genet.* 46(3), 176–182 (2009).

113 Jackman C, Horn ND, Molleston JP, Sokol DK. Gene associated with seizures, autism, and hepatomegaly in an Amish girl. *Pediatr. Neurol.* 40(4), 310–313 (2009).

114 Poot M, Beyer V, Schwaab I *et al.* Disruption of *CNTNAP2* and additional structural genome changes in a boy with speech delay and autism spectrum disorder. *Neurogenetics* 11(1), 81–89 (2010).

115 Strauss KA, Puffenberger EG, Huentelman MJ *et al.* Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. *N. Engl. J. Med.* 354(13), 1370–1377 (2006).

116 Orrico A, Galli L, Buoni S, Orsi A, Vonella G, Sorrentino V. Novel *PTEN* mutations in neurodevelopmental disorders and macrocephaly. *Clin. Genet.* 75(2), 195–198 (2009).

117 Varga EA, Pastore M, Prior T, Herman GE, Mcbride KL. The prevalence of *PTEN* mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly. *Genet. Med.* 11(2), 111–117 (2009).

118 Kent L, Bowdin S, Kirby GA, Cooper WN, Maher ER. Beckwith Weidemann syndrome: a behavioral phenotype–genotype study. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 147B(7), 1295–1297 (2008).

119 Tierney E, Nwokoro NA, Porter FD, Freund LS, Ghuman JK, Kelley RI. Behavior phenotype in the RSH/Smith–Lemli–Opitz syndrome. *Am. J. Med. Genet.* 98(2), 191–200 (2001).

120 Splawski I, Timothy KW, Sharpe LM *et al.* Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell* 119(1), 19–31 (2004).

121 Splawski I, Yoo DS, Stotz SC, Cherry A, Clapham DE, Keating MT. *CACNA1H* mutations in autism spectrum disorders. *J. Biol. Chem.* 281(31), 22085–22091 (2006).

122 Cook EH Jr, Lindgren V, Leventhal BL *et al.* Autism or atypical autism in maternally but not paternally derived proximal 15q duplication. *Am. J. Hum. Genet.* 60(4), 928–934 (1997).

123 Depienne C, Moreno-De-Luca D, Heron D *et al.* Screening for genomic rearrangements

and methylation abnormalities of the 15q11–q13 region in autism spectrum disorders. *Biol. Psychiatry* 66(4), 349–359 (2009).

124 Ben-Shachar S, Lanpher B, German JR *et al.* Microdeletion 15q13.3: a locus with incomplete penetrance for autism, mental retardation, and psychiatric disorders. *J. Med. Genet.* 46(6), 382–388 (2009).

125 Sharp AJ, Mefford HC, Li K *et al.* A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat. Genet.* 40(3), 322–328 (2008).

126 Bonati MT, Russo S, Finelli P *et al.* Evaluation of autism traits in Angelman syndrome: a resource to unfold autism genes. *Neurogenetics* 8(3), 169–178 (2007).

127 Peters SU, Beaudet AL, Madduri N, Bacino CA. Autism in Angelman syndrome: implications for autism research. *Clin. Gen.* 66(6), 530–536 (2004).

128 Descheemaeker MJ, Govers V, Vermeulen P, Fryns JP. Pervasive developmental disorders in Prader–Willi syndrome: the Leuven experience in 59 subjects and controls. *Am. J. Med. Genet.* A 140(11), 1136–1142 (2006).

129 Potocki L, Bi W, Treadwell-Deering D *et al.* Characterization of Potocki-Lupski syndrome (dup[17][p11.2p11.2]) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype. *Am. J. Hum. Genet.* 80(4), 633–649 (2007).

130 Treadwell-Deering DE, Powell MP, Potocki L. Cognitive and behavioral characterization of the Potocki–Lupski syndrome (duplication 17p11.2). *J. Dev. Behav. Pediatr.* 31(2), 137–143 (2010).

131 Fine SE, Weissman A, Gerdes M *et al.* Autism spectrum disorders and symptoms in children with molecularly confirmed 22q11.2 deletion syndrome. *J. Autism Dev. Disord.* 35(4), 461–470 (2005).

132 Niklasson L, Rasmussen P, Oskarsdottir S, Gillberg C. Autism, ADHD, mental retardation and behavior problems in 100 individuals with 22q11 deletion syndrome. *Res. Dev. Disabil.* 30(4), 763–773 (2009).

133 Vorstman JA, Morcus ME, Duijff SN *et al.* The 22q11.2 deletion in children: high rate of autistic disorders and early onset of psychotic symptoms. *J. Am. Acad. Child Adolesc. Psychiatry* 45(9), 1104–1113 (2006).

134 Manning MA, Cassidy SB, Clericuzio C *et al.* Terminal 22q deletion syndrome: a newly recognized cause of speech and language disability in the autism spectrum. *Pediatrics* 114(2), 451–457 (2004).

135 Moessner R, Marshall CR, Sutcliffe JS *et al.* Contribution of *SHANK3* mutations to autism spectrum disorder. *Am. J. Hum. Genet.* 81(6), 1289–1297 (2007).

136 Diguiseppi C, Hepburn S, Davis JM *et al.* Screening for autism spectrum disorders in children with Down syndrome: population prevalence and screening test characteristics. *J. Dev. Behav. Pediatr.* 31(3), 181–191 (2010).

137 Lowenthal R, Paula CS, Schwartzman JS, Brunoni D, Mercadante MT. Prevalence of pervasive developmental disorder in Down's syndrome. *J. Autism Dev. Disord.* 37(7), 1394–1395 (2007).

138 Hatton DD, Sideris J, Skinner M *et al.* Autistic behavior in children with fragile X syndrome: prevalence, stability, and the impact of FMRP. *Am. J. Med. Genet. A* 140A(17), 1804–1813 (2006).

139 Winarni TI, Utari A, Mundhofir FE, Hagerman RJ, Faradz SM. Fragile X syndrome: clinical, cytogenetics and molecular screening among autism spectrum disorder children in Indonesia. *Clin. Genet.* doi:10.1111/cge.12095 (2013) (Epub ahead of print).

140 Ehninger D, Silva AJ. Rapamycin for treating tuberous sclerosis and autism spectrum disorders.*Trends Mol. Med.* 17(2), 78–87 (2011).

141 Liu E, Knutzen CA, Krauss S, Schweiger S, Chiang GG. Control of mTORC1 signaling by the Opitz syndrome protein MID1. *Proc. Natl Acad. Sci. USA* 108(21), 8680–8685 (2011).

142 Ji H, Wang Z, Perera SA *et al.* Mutations in *BRAF* and *KRAS* converge on activation of the mitogen-activated protein kinase pathway in lung cancer mouse models. *Cancer Res.* 67(10), 4933–4939 (2007).

143 Bandyopadhyay S, Chiang CY, Srivastava J *et al.* A human MAP kinase interactome. *Nat. Methods* 7(10), 801–805 (2010).

144 Choi YB, Li HL, Kassabov SR *et al.* Neurexin-neuroligin transsynaptic interaction mediates learning-related synaptic remodeling and long-term facilitation in aplysia. *Neuron* 70(3), 468–481 (2011).

145 Kim JE, O'Sullivan ML, Sanchez CA *et al.* Investigating synapse formation and function using human pluripotent stem cell-derived neurons. *Proc. Natl Acad. Sci. USA* 108(7), 3005–3010 (2011).

146 Shoubridge C, Walikonis RS, Gecz J, Harvey RJ. Subtle functional defects in the Arf-specific guanine nucleotide exchange factor IQSEC2 cause non-syndromic X-linked intellectual disability. *Small GTPases* 1(2), 98–103 (2010).

147 Grabrucker AM, Knight MJ, Proepper C *et al.* Concerted action of zinc and ProSAP/Shank in synaptogenesis and synapse maturation. *EMBO J.* 30(3), 569–581 (2011).

148 Atasoy D, Schoch S, Ho A *et al.* Deletion of CASK in mice is lethal and impairs synaptic function. *Proc. Natl Acad. Sci. USA* 104(7), 2525–2530 (2007).

149 Daoud F, Candelario-Martinez A, Billard JM *et al.* Role of mental retardation-associated dystrophin-gene product Dp71 in excitatory synapse organization, synaptic plasticity and behavioral functions. *PLoS ONE* 4(8), e6574 (2009).

150 Ronesi JA, Collins KA, Hays SA *et al.* Disrupted Homer scaffolds mediate abnormal mGluR5 function in a mouse model of fragile X syndrome. *Nat. Neurosci.* 15(3), 431–440, S431 (2012).

151 Auerbach BD, Osterweil EK, Bear MF. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature* 480(7375), 63–68 (2011).

152 Melko M, Douguet D, Bensaid M *et al.* Functional characterization of the AFF (AF4/FMR2) family of RNA-binding proteins: insights into the molecular pathology of FRAXE intellectual disability. *Hum. Mol. Genet.* 20(10), 1873–1885 (2011).

153 Lee JA, Tang ZZ, Black DL. An inducible change in Fox-1/A2BP1 splicing modulates the alternative splicing of downstream neuronal target exons. *Genes Dev.* 23(19), 2284–2293 (2009).

154 Ricciardi S, Kilstrup-Nielsen C, Bienvenu T, Jacquette A, Landsberger N, Broccoli V. CDKL5 influences RNA splicing activity by its association to the nuclear speckle molecular machinery. *Hum. Mol. Genet.* 18(23), 4590–4602 (2009).

155 Chamieh H, Ballut L, Bonneau F, Le Hir H. NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nat. Struct. Mol. Biol.* 15(1), 85–93 (2008).

156 Tapia VE, Nicolaescu E, Mcdonald CB *et al.* Y65C missense mutation in the WW domain of the Golabi-Ito-Hall syndrome protein PQBP1 affects its binding activity and deregulates pre-mRNA splicing. *J. Biol. Chem.* 285(25), 19391–19401 (2010).

157 Schenck A, Bardoni B, Moro A, Bagni C, Mandel JL. A highly conserved protein family interacting with the fragile X mental retardation protein (FMRP) and displaying selective interactions with FMRP-related

proteins FXR1P and FXR2P. *Proc. Natl Acad. Sci. USA* 98(15), 8844–8849 (2001).

158 Fogel BL, Wexler E, Wahnich A *et al.* RBFOX1 regulates both splicing and transcriptional networks in human neuronal development. *Hum. Mol. Genet.* 21(19), 4171–4186 (2012).

■ **Largest genome-wide association study on patients with ASD published and found one single nucleotide polymorphism that reached genome-wide significance (although this was not replicated).**

159 Lucio-Eterovic AK, Singh MM, Gardner JE, Veerappan CS, Rice JC, Carpenter PB. Role for the nuclear receptor-binding SET domain protein 1 (NSD1) methyltransferase in coordinating lysine 36 methylation at histone 3 with RNA polymerase II function. *Proc. Natl Acad. Sci. USA* 107(39), 16952–16957 (2010).

160 Iwase S, Lan F, Bayliss P *et al.* The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell* 128(6), 1077–1088 (2007).

161 Kleine-Kohlbrecher D, Christensen J, Vandamme J *et al.* A functional link between the histone demethylase PHF8 and the transcription factor ZNF711 in X-linked mental retardation. *Mol. Cell* 38(2), 165–178 (2010).

162 Srinivasan S, Dorighi KM, Tamkun JW. *Drosophila* Kismet regulates histone H3 lysine 27 methylation and early elongation by RNA polymerase II. *PLoS Genet.* 4(10), e1000217 (2008).

163 Kramer JM, Kochinke K, Oortveld MA *et al.* Epigenetic regulation of learning and memory by *Drosophila* EHMT/G9a. *PLoS Biol.* 9(1), e1000569 (2011).

164 Kagey MH, Newman JJ, Bilodeau S *et al.* Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467(7314), 430–435 (2010).

165 Kernohan KD, Jiang Y, Tremblay DC *et al.* ATRX partners with cohesin and MeCP2 and contributes to developmental silencing of imprinted genes in the brain. *Dev. Cell* 18(2), 191–202 (2010).

166 Mikhail FM, Lose EJ, Robin NH *et al.* Clinically relevant single gene or intragenic deletions encompassing critical neurodevelopmental genes in patients with developmental delay, mental retardation, and/or autism spectrum disorders. *Am. J. Med. Genet.* A 155A(10), 2386–2396 (2011).

■■ **Describes the results of exome sequencing efforts on simplex ASD families.**

167 Morrow EM, Yoo SY, Flavell SW *et al.* Identifying autism loci and genes by tracing recent shared ancestry. *Science* 321(5886), 218–223 (2008).

168 Gharani N, Benayed R, Mancuso V, Brzustowicz LM, Millonig JH. Association of the homeobox transcription factor, ENGRAILED 2, 3, with autism spectrum disorder. *Mol. Psychiatry* 9(5), 474–484 (2004).

169 Choi J, Ababon MR, Matteson PG, Millonig JH. Cut-like homeobox 1 and nuclear factor I/B mediate ENGRAILED2 autism spectrum disorder-associated haplotype function. *Hum. Mol. Genet.* 21(7), 1566–1580 (2012).

170 Tischfield MA, Bosley TM, Salih MA *et al.* Homozygous *HOXA1* mutations disrupt human brainstem, inner ear, cardiovascular and cognitive development. *Nat. Genet.* 37(10), 1035–1037 (2005).

171 Talkowski ME, Rosenfeld JA, Blumenthal I *et al.* Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. *Cell* 149(3), 525–537 (2012).

172 O'Roak BJ, Deriziotis P, Lee C *et al.* Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations. *Nat. Genet.* 43(6), 585–589 (2011).

173 Laje G, Morse R, Richter W, Ball J, Pao M, Smith AC. Autism spectrum features in Smith-Magenis syndrome. *Am. J. Med. Genet. C Semin. Med. Genet.* 154C(4), 456–462 (2010).

174 Marsh E, Fulp C, Gomez E *et al.* Targeted loss of Arx results in a developmental epilepsy mouse model and recapitulates the human phenotype in heterozygous females. *Brain* 132(Pt 6), 1563–1576 (2009).

175 Turner G, Partington M, Kerr B, Mangelsdorf M, Gecz J. Variable expression of mental retardation, autism, seizures, and dystonic hand movements in two families with an identical *ARX* gene mutation. *Am. J. Med. Genet.* 112(4), 405–411 (2002).

176 Chen YZ, Matsushita M, Girirajan S *et al.* Evidence for involvement of GNB1L in autism. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 159B(1), 61–71 (2012).

177 Paylor R, Glaser B, Mupo A *et al.* Tbx1 haploinsufficiency is linked to behavioral disorders in mice and humans: implications for 22q11 deletion syndrome. *Proc. Natl Acad. Sci. USA* 103(20), 7729–7734 (2006).

■ Website

201 SFARI gene home. http://gene.sfari.org