

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Autism Spectrum Disorders: Insights from Genomics

---

John J. Connolly and Hakon Hakonarson

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/54357>

---

## 1. Introduction

The field of genetics has made considerable scientific progress in the past several years and continues to evolve at a rapid pace. This progress parallels developments in genomic technology, where instrumentation and methodology are becoming increasingly sophisticated and cost-effective. Here, we review recent developments in understanding autism spectrum disorder (ASD) from a genomics perspective. A large catalog of common and rare variants has now been associated with ASD, and we are beginning to see some of these discoveries translate into pharmacogenomic intervention. This review provides an overview of genome-wide association studies (GWAS) and common genetic variants, followed by an overview of the status of rare variant research, which have risen to prominence with the proliferation of next-generation sequencing and techniques for identifying copy number variants. While these approaches need not be mutually exclusive, they provide a useful structure for organizing relevant genetic factors. Although there is much work to be done before these discoveries will enter the clinic, the past decade has seen us make major inroads in elucidating the causes of ASD and making tentative steps towards developing treatments.

### 1.1. Defining the autism phenotype

Autism is known to be highly heterogeneous, and this phenomenon has made definitions of the phenotype somewhat problematic. The American Psychiatric Association recently proposed revisions to its Diagnostic and Statistical Manual of Mental Disorders V (DSM-5) criteria for ASD (see Wing *et al.*, 2011) [1], acknowledging the long-observed overlap between social and communication dimensions (previously separate). Thus, ASD will be defined by 1) persistent deficits in social communication and social interaction across contexts, and 2) restricted, repetitive patterns of behavior, interests, or activities. These should impair every-

day functioning, not be accounted for by general developmental delays, and be present from early childhood.

For large-scale genome analyses, DSM criteria have been considered insufficiently precise, and cases are often selected using scores from the Autism Diagnostic Interview (ADI-R) [2], Autism Diagnostic Observation Schedule (ADOS) [3], and/or Social Responsiveness Scale (SRS) [4]. These instruments offer a more robust psychometric platform, and cases defined as “autism” are required to meet strict threshold criteria (e.g. all sub-dimensions of the ADI-R and ADOS). Individuals not quite meeting these criteria may be subsumed under the “broader” autism phenotype, which also typically includes Asperger syndrome, childhood disintegrative disorder and pervasive developmental disorder not otherwise specified. A diagnosis of Rett syndrome—which has a reportedly distinct pathophysiology, clinical course, and diagnostic strategy (Levy, Mandell & Schultz, 2009) [5] and will likely be removed in the impending publication of DSM-V—is typically exclusionary. Intellectual impairment, which is often co-morbid with ASD (Dawson *et al.*, 2007; Bölte *et al.*, 2009) [6,7] is not an exclusionary criterion, but is co-varied in statistical analyses. Given the broad range of IQ tests and their associated psychometric properties, this requires considerable finesse.

Standardization of diagnostic criteria has facilitated the accumulation of large ASD sample-sets, where institutions can share (de-identified) data. In this vein, initiatives such as the Autism Genome Project include data from several thousand ASD individuals, greatly increasing statistical power of relevant analyses.

## 1.2. Heritability of ASD

Although Skuse (2007) [8] cautions that heritability estimates of ASD may have been skewed by the co-inheritance of (low) intelligence or other variables, there is little doubt that genetic factors play a key role in autism. In the most widely-cited twin study, Bailey *et al.* (1995) [9] report that monozygotic twins are 92% concordant on a broad spectrum of cognitive or social abnormalities, compared with only 10% for dizygotic twins. Parents and siblings of individuals with ASD often exhibit subsyndromal levels of impairment (Piven *et al.*, 1997) [10], and having an affected sibling is the single biggest risk factor for developing an ASD. In an analysis of 943,664 Danish children (Lauritsen *et al.*, 2005) [11], the strongest predictors of autism were siblings with ASD, who conferred a 22-fold increased risk, while Fombonne (2005) [12] suggested that this risk may be even greater.

## 1.3. Early genetic studies – insights from Rett syndrome and Fragile X

Early efforts to identify the genetic causes of ASD utilized linkage and association approaches. Linkage studies, more prominent in the 1980s and 1990s, typically focus on families or larger pedigrees and are well powered to identify rarer genetic variants. The most common linkage approach is the affected sib-pair design (see O’Roak & State, 2008) [13], which examines the transmission of genomic segments through generations. Linkage studies helped define the locus containing *FMR1*, which is mutated in fragile X syndrome (e.g. Richards *et al.*, 1991) [14], Approximately 30% of children with fragile X syndrome meet cri-

teria for autism (Rogers *et al.*, 2001; Harris *et al.*, 2008) [15,16]. Similarly, linkage studies have been important to identifying *MECP2* as the major cause of Rett syndrome (e.g. Curtis *et al.*, 1993) [17].

Association studies take a different approach. Rather than track transmission of specific genomic regions through generations, association studies scan the breadth of the genome. Here, the goal is to determine *post-hoc* whether identified variants are more or less common in affected individuals. Early association studies (i.e. pre HapMap) were complementary with the linkage approach, and in many designs, linkage primed target loci for this more fine-grained analysis.

These early insights have played an important role in shaping our current understanding of ASD. Functional studies of *FMR1* and *MECP2* have highlighted the importance of synaptic dysfunction (Ramocki & Zoghbi, 2008) [18] as a unifying factor that could extend into the more common forms of autism and, as discussed below, remain highly relevant to our understanding of the broader ASD phenotype.

## 2. Genome wide association and common variants

Aside from notable successes with fragile X and Rett syndrome, early linkage and association studies have been inconsistent in resolving more complex genetic correlates of ASD, and candidate genes have often not been replicated between studies. These challenges may in part be accounted-for by their relatively low resolution/coverage, making it difficult to detect candidate loci other than those of major effect. A shift in technology was required to get beyond such challenges, which was engendered by the introduction of high-resolution single nucleotide polymorphism (SNP) arrays. SNP arrays provided coverage of many thousand (now several million) common SNPs, which could be examined at a relatively low cost across large sample sets.

Genome-wide association studies (GWAS) examine the frequency of SNPs in case *versus* control populations, and can adopt either a case-control or family-based design. The former allows researchers to avoid the often complex process of acquiring diagnostic/phenotype data from a patient's family, and can incorporate very large numbers of control datasets that may be more readily available. The latter controls for the often confounding phenomenon of population stratification, where variants more common to specific racial groups may either be erroneously identified as causal, or obscure actual causal variants. A major caveat with family-based designs is the often unfounded assumption that unaffected family members do not share causal variants.

GWAS test for common variants (>1% population frequency), with the assumption that ASD are at least in part caused by the coinheritance of multiple risk variants, each of small individual effect (odds ratios typically between ~1:1 and ~1:5). This assumption is known as the common disease-common variant (CDCV) model (Risch & Merikangas, 1996) [20].

## 2.1. The 5p14.1 locus

A 2009 paper from our laboratory (Wang *et al.*, 2009 [21]) was the first to identify common variants for ASD on a genome-wide scale. Our group examined 780 families (3,101 individuals) with affected children, a second, independent group of 1,204 affected individuals, and 6,491 controls. All were of European ancestry. We identified six genetic markers on chromosome 5 in the 5p14.1 region that confirmed susceptibility to ASD. This locus has been replicated in two additional independent cohort studies (Ma *et al.*, 2009; St Pourcain *et al.*, 2010) [22, 23], lending further support for 5p14 as associated with ASD risk.

As shown in Figure 1, the region straddles two genes, *CDH9* and *CDH10*. Both genes encode type II classical cadherins, transmembrane proteins that promote cell adhesion. Cadherins represent a large family of transmembrane proteins that mediate calcium-dependent cell-cell adhesion, and have been shown to generate synaptic complexity in the developing brain (Redies, Hertel & Hübner, 2012) [24]. The association of cadherins is consistent with the cortical-disconnectivity model of autism (e.g. Gepner & Féron, 2009) [25], which postulates that ASD may result from an increase or decrease in functional connectivity and neuronal synchronization in relevant neural pathways. While this hypothesis may yet be confirmed, a recent study by Kerin *et al.* (2012) [26] suggests a more complex mechanism to explain association between ASD and the 5p14.1 locus.

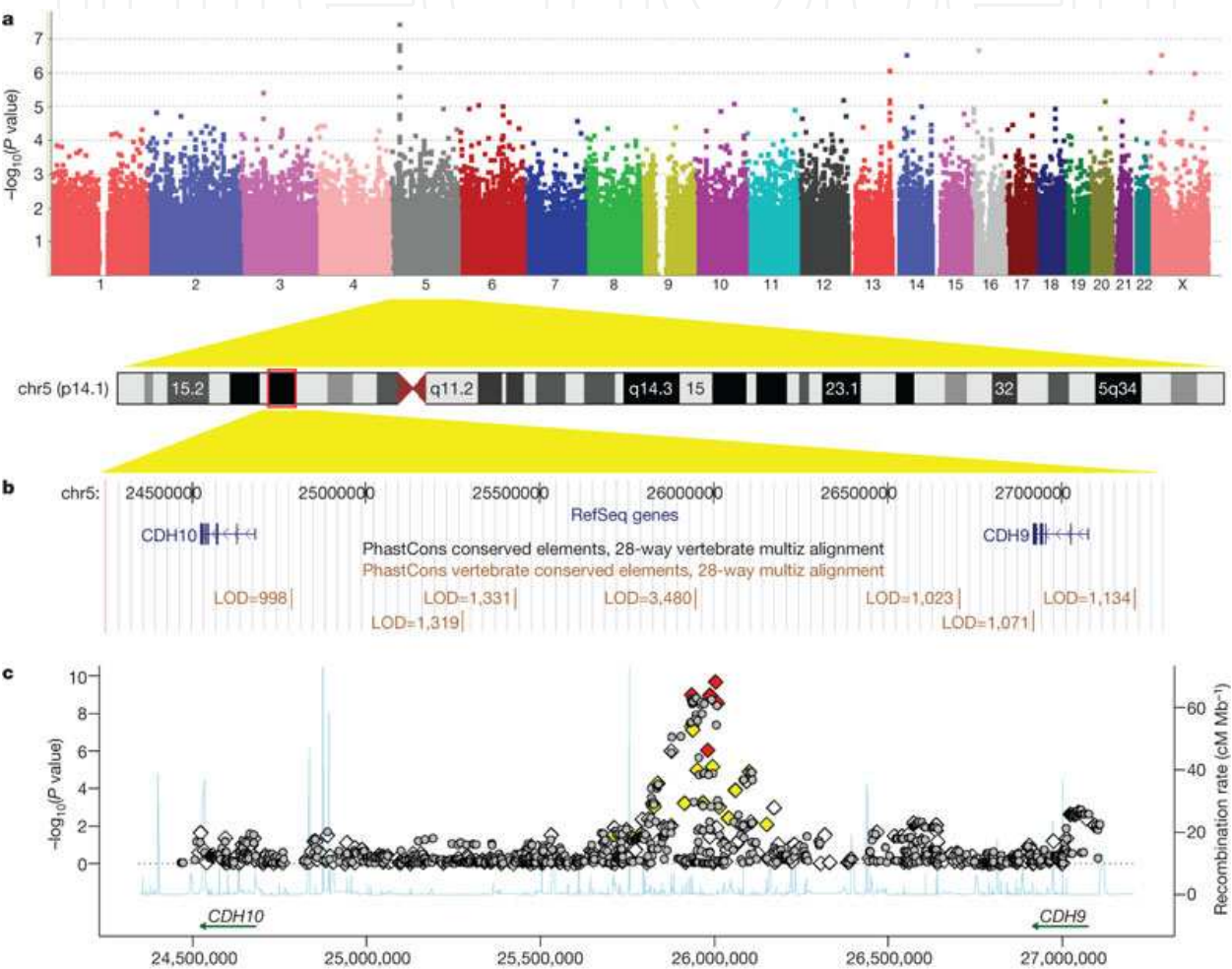
Basing their analyses on the genomic region surrounding the rs4307059 locus, the authors used a bioinformatics approach (i.e. Genome Browser) to examine relevant expressed sequence tags (ESTs) and RNA (by Tiling Array within the 100-kb linkage disequilibrium at the GWAS peak). Only one functional element—a single noncoding RNA—was located. The 3.9-kb RNA corresponded to moesin pseudogene 1 (*MSNP1*), and has 94% sequence identity to the mature mRNA of the protein-coding gene *MSN*. Located on the X chromosome (Xq11.2), *MSN* spans 74 kb and contains 13 exons. It produces a 4-kb mRNA, and encodes the 577-amino acid moesin protein. The noncoding RNA at 5p14.1 was encoded by the opposite (antisense) strand of *MSNP1*, and was therefore named moesin pseudogene 1, antisense – *MSNP1AS*.

Follow-up analyses by the group largely confirm that *MSNP1AS* is expressed in the brain, providing important functional validation. Using custom TaqMan Gene Expression assays to target the region, they showed that while *MSN* was widely expressed in all tissues tested, *MSNP1AS* was expressed variably. Sites of greatest expression were the adult temporal cerebral cortex, adult peripheral blood, and fetal heart, as well as three immortalized cell lines. Moreover, postmortem analyses (qPCR on total RNA) of fresh-frozen, superior temporal gyri of ASD-control pairs (n=10) found a 12.7-fold increase of *MSNP1AS* expression in the temporal cortex of individuals with ASD. Individuals with ASD also showed a 2.4-fold increase in *MSN* expression in the same region. Interestingly, there was no evidence of increased expression for either *CDH9* or *CDH10*.

The group next used genotype—determined from three associated SNPs from the original Wang *et al.* paper—as the independent variable in expression analyses. All three SNPs, rs7704909, rs12518194, and rs4307059, have a high degree of linkage disequilibrium (LD)



( $r^2>0.98$ ). Using resequencing to compare relevant genotypes, they identified highly significant differences in *MSNP1AS* expression. Thus, the T/T genotype at rs7704909 corresponded to a 23.3 fold increase in *MSNP1AS* RNA compared to the C/C genotype. For the rs4307059, the T/T *versus* C/C genotype corresponded to a 22.0-fold increase in *MSNP1A* expression. For rs12518194, the A/A *versus* G/G genotype corresponded to a 10.8 fold increase in *MSNP1A* expression. Again, there was no evidence of increased/decreased expression differences for *CDH9* or *CDH10* in relation to genotype or case/control status.



**Figure 1.** Genome-wide association results at the 5p14.1 region. **a**, A Manhattan plot shows the  $\log_{10}(p)$  values of SNPs from their combined association analysis. **b**, The 5p14.1 region in the UCSC Genome Browser, with conserved genomic elements in the PhastCons track. **c**, Genotyped (diamonds) and imputed (grey circles) SNPs are plotted with their combined  $p$  values. Genotyped SNPs are colored on the basis of their correlation with rs4307059 (red:  $r^2 \geq 0.5$ ; yellow:  $0.2 \leq r^2 < 0.5$ ; white:  $r^2 < 0.2$ ). Estimated recombination rates from HapMap are plotted to reflect local linkage disequilibrium. Adapted from Wang et al., 2009 [21]. Reprinted with permission from Nature Publishing Group.

Although Western blot analyses did not identify significant differences in moesin protein levels between cases and controls, overexpressing *MSNP1AS* in human cell lines was shown to significantly reduce levels of the moesin protein. The authors speculated that relevant alterations in moesin may occur only during specific development landmarks, which may im-

pact neurodevelopment. This would explain why moesin levels are not elevated in the ASD samples *per se*, in spite of the marked differences in *MSNP1AS* expression. Further work is needed to confirm this hypothesis, and quantification of moesin protein levels at key developmental stages would certainly contribute in this respect.

Taken as a whole, these results provide compelling support for 5p14.1 as a risk locus for ASD. Although sample sizes for some analyses were small (10 ASD-control pairs for post-mortem studies), this quite rigorous series of experiments draws a clear path from GWAS result through functional validation. As such, these results help allay criticism of the GWAS approach as a means of candidate discovery. Thus, a 2010 review by McClellan and King (2010) [27] singled out the 5p14.1 locus as an example of the “perils of cryptic population stratification”. These comments seemed somewhat misguided in the light of rigorous methodologies developed by the GWAS community for controlling stratification (e.g. EigenStrat) [28], replication [22, 23], and now functional validation by the Kerin *et al.* group [26].

Similarly, replication/validation of the 5p14.1 locus provides an important demonstration of the legitimacy of associations in intergenic regions. Again, McClellan and King had disputed the utility of such results, questioning how “genome-wide association studies come to be populated by risk variants with no known function?” It is important to emphasize that the GWAS approach typically does not tag the disease variant, but rather its approximate location—through linkage disequilibrium, this is typically 100kb or less. Moreover, as in the Kerin *et al.* paper, the significant SNP may be tagging an intergenic regulatory element, which has functional consequences far beyond the associated region, in this case the *MSN* locus on the X-chromosome.

Finally, these expression analyses provide a reminder about the capabilities of different genomic technologies. In the past twelve months, a number of high-profile next-generation sequencing (NGS) studies have been able to examine genomic correlates of ASD with unprecedented resolution. These types of studies—reviewed in greater depth below—have been interpreted as the future of ASD genetics and, to a large extent, this may be true. However, we note that DNA sequencing in the 5p14.1 region would not have identified the non-coding RNA at this locus. Thus, although NGS platforms used for RNA-sequencing are becoming increasingly sophisticated (Ozsolak & Milos, 2011) [29], microarray studies retain a place in guiding genomic discovery.

## 2.2. Other replicated common variants from candidate gene studies

A number of other common variants from candidate gene studies have been proposed as ASD risk factors. These include Contactin Associated Protein 2 (*CNTNAP2*), located on chromosome 7q35, which has been identified as a candidate for the age at first word endophenotype (Alarcón *et al.*, 2002) [30]. A follow-up by the same group (Alarcón *et al.*, 2008) [31] using linkage, association, and gene-expression analyses, found *CNTNAP2* to be the only autism-susceptibility gene to reach significance across all approaches. An independent linkage analysis by Arking *et al.* (2008) [32] also highlighted *CNTNAP2* as a significant ASD candidate gene. *CNTNAP2* is part of the neurexin family, which have repeatedly been associated with autism (see below). Interestingly, Vernes *et al.* (2008) [33] showed that

*CNTNAP2* binds to *FOXP2*, which is a well-established correlate of language and speech disorders (Lai *et al.*, 2001) [34], and are commonly observed in ASD.

Another locus indentified by the candidate gene approach is Engrailed 2 (*EN2*), a homeobox gene that is critical to the development of the midbrain and cerebellum. Like other homeobox genes, it regulates morphogenesis. *EN2* is a human homolog of the engrailed gene, which is found in *Drosophila*. *En2* mouse mutants have anatomic phenotypes in the cerebellum that resemble cerebellar abnormalities reported in autistic individuals (Cheng *et al.*, 2010) [35]. In three separate datasets, Benayed *et al.* (2005, 2009) [36, 37] have reported and replicated a significant association between *EN2* and both broad and narrow ASD phenotypes. Wang *et al.* (2008) [38] also found an association between *EN2* and ASD in a Chinese Han sample, although Zhong *et al.* (2003) [39] failed to find evidence of an underlying association.

The oncogene *MET* is also strongly linked to ASD etiology, having been supported by a number of studies in the past decade (e.g. IMGSA, 2001; Campbell *et al.*, 2006, 2008; Sousa *et al.*, 2009) [40-43]. Recently, Eagleson *et al.* (2011) [44] reported a role for *Met* signaling in cortical interneuron development in vitro in a mouse model.

### 2.3. Unexplained variance

For the most significant discovery SNP identified in the Wang *et al.* study above (rs4307059), the risk allele frequency was 0.65 in cases with an odds ratio of 1.19. This is comparable with common variant discoveries in other psychiatric disorders including schizophrenia (Glessner *et al.*, 2009) [45], bipolar disorder (Ferreira, 2008) [46], and attention-deficit/hyperactivity disorder (Arcos-Burgos *et al.* 2010) [47]. While it is important not to undermine the significance of these findings, it should be noted that the predictive value of such ratios is relatively low (Dickson *et al.* 2010) [48], often explaining less than 5% of the total risk (review at <http://www.genome.gov/gwastudies>). However, it is also possible that these common SNPs may be tagging a rarer causative variant (i.e. synthetic association), where the effect sizes may be markedly underestimated by the GWAS variant as we recently reported (Dickson *et al.* 2010) [48]. In one example, Wang *et al.* (2010) [49] examined the *NOD2* locus as a cause of Crohn's disease. Using resequencing data, they found that three causal variants explain > 5% of the genetic risk, where GWAS had estimated the risk at ~1%. This finding has two potentially important implications. First of all, it highlights the need for careful phenotyping of cohorts, which is important to ensure that the phenotypes produced by rare-variants are not being "filtered-out" and missed as a consequence. A long range haplotype analysis of the GWAS data at respective loci is recommended in an attempt to enrich for individuals with rare-causative variants, who can be selected from the cohort and sequenced for confirmation (Wang *et al.*, 2010) [49].

Second, the results of our Crohn's disease study suggest that in certain circumstances, there may be an explicit relationship between tagged variants and underlying rare variants. Thus, the distinction between loci harboring common *versus* rare variants is not necessarily concrete. Indeed, the same locus may harbor both common and rare variants (Anderson *et al.*,



2011). In recent years, we have seen an increased emphasis on the former, which is reflected in an upsurge in the number of copy number variation (CNV) and NGS studies.

### 3. Rare variants – CNVs and next-generation sequencing

#### 3.1. Copy number variation in ASD

CNVs are insertions, deletions, or translocations in the human genome that are universal in the general population (e.g. Pinto *et al.*, 2010) [50]. CNVs can be detected by the same SNP arrays used in GWAS, and vary in length from many megabases to 1 kilobase or smaller. They are often not associated with any observable phenotype.

One of the most widely-known CNVs is Down syndrome, which is characterized by an extra chromosome 21. Rett syndrome is also caused by a CNV, which includes a deletion in *MECP2*. CNVs can be inherited or occur *de novo*, the cause of which is thus far unknown. Common disease-causing CNVs are infrequent but rare CNVs, with a frequency of less than 1%, have been identified for a range of disorders including ADHD (e.g. Williams *et al.*, 2010) [51], schizophrenia (e.g. Glessner *et al.*, 2010; Levinson *et al.*, 2011) [52, 53], bipolar disorder (e.g. Chen *et al.*, 2010) [54] and many others.

Sebat *et al.* (2007) [55] provided some early insights into the genomic features of CNVs in ASD. Firstly, they noted that *de novo* CNVs were individually rare – from 118 ASD cases, none of the identified variants were observed more than twice, with the majority seen just once. This confirmed the widely-held assumption that many different loci can contribute to the same ASD phenotype. The sheer volume of loci identified by this approach (multiple loci on 20 chromosomes) affirms the extraordinarily complexity of ASD.

A number of subsequent studies have greatly expanded the number of candidate loci using the CNV approach. Our laboratory (Bucan *et al.*, 2009) [56] reported 150+ CNVs in 912 ASD families that were not found in 1,488 controls. Critically, 27 of these loci were replicated in an independent cohort of 859 ASD cases and 1,051 controls. Some of the rare variants we identified had previously been associated with autism, including *NRXN1* and *UBE3A*, (Guilmatre *et al.*, 2009) [57]. Samaco *et al.* (2005) [58] previously identified significant deficits in *ube3a* expression in *mecp2*-deficient mice, suggesting a shared pathological pathway with Rett syndrome (as well as Angelman syndrome, and autism). Similarly, Kim *et al.* (2008) [59] associated *NRXN1* with a balanced chromosomal abnormality at chromosome 2p16.3 in two unrelated ASD individuals. Rare variants in the coding region included two missense changes.

Glessner *et al.* (2009) [52] identified and reported CNVs in two major gene networks, including neuronal cell adhesion molecules (such as *NRXN1*) and the ubiquitin gene family (such as *UBE3A*). Interestingly, four of the most prominent genes enriched by CNVs in ASD cases (*UBE3A*, *PARK2*, *RFWD2* and *FBXO40*) – all of which were uncovered independently – are part of the ubiquitin gene family. Ubiquitination can alter protein function after translation, and degrade target proteins in conjunction with proteasomes. The ubiquitin–proteasome

system operates at pre- and post-synapses, whose functions includes regulating neurotransmitter release, recycling synaptic vesicles in pre-synaptic terminals, and modulating changes in dendritic spines and post-synaptic density (Yi & Ehlers, 2005) [60]. As well as implicating an ASD-ubiquitination network, we also identified a second pathway involving *NRXN1*, *CNTN4*, *NLGN1*, and *ASTN2*. Genes in this group mediate neuronal cell-adhesion, and contribute to neurodevelopment by facilitating axon guidance, synapse formation and plasticity, and neuron–glial interactions. We also note that ubiquitins are involved in recycling cell-adhesion molecules, which is a possible mechanism by which these two networks are cross linked.

In a similar approach, Pinto *et al.* (2010) [50] further confirmed the importance of rare CNVs as causal factors for ASD. The group did not observe a significant difference between cases and controls in terms of raw number of CNVs or estimated CNV size. However, the number of CNVs in genic regions was significantly greater in ASD cases. Again, loci enriched for CNVs include a number of genes known to be important for neurodevelopment and synaptic plasticity, such as *SHANK2*, *SYNGAP1*, and *DLGAP2*. Between 5.5% and 5.7% of ASD cases have at least one *de novo* CNV, further confirming the significance of *de novo* genetic events as risk factors for autism. Similar to the Glessner study, the Pinto group mapped CNVs to a series of networks involved in the development and regulation of the central nervous system functions. Implicated networks include neuronal cell adhesion, GTPase regulation (important for signal transduction and biosynthesis), and GTPase/Ras signaling, also involved in ubiquitination.

Finally, Gai *et al.* (2011) [61] took a slightly different approach, focusing exclusively on inherited CNVs. While underlying loci were not necessarily common to those identified by the Glessner and Pinto groups, enrichment in pathways involving central nervous system development, synaptic functions and neuronal signaling processes was again confirmed. The Gai *et al.* study also emphasized the role of glutamate-mediated neuronal signals in ASD.

Collectively, these CNV studies suggest that certain hotspots on the genome are particularly vulnerable to ASD, which include loci on chromosomes 1q21, 3p26, 15q11-q13, 16p11, and 22q11. These hotspots are part of large gene networks that are important to neural signaling and neurodevelopment and have additionally been associated with other neuropsychiatric disorders. In particular, a number of CNV studies in schizophrenia have highlighted structural mutations incorporating chromosomes 1q21, 15q13, and 22q11 (e.g. Glessner *et al.*, 2010) [62], which are significantly enriched in cases versus controls, with *NRXN1* being a standout in this regard. From a phenotype perspective, autism and schizophrenia seem very different, both in behavioral manifestation and age of onset, and it may seem counter-intuitive that associated loci should overlap. Some authors have addressed this peculiarity by proposing that schizophrenia and autism may in fact be different poles of the same spectrum. Thus, Crespi and Braddock (2008) [63] suggest that social cognition is underdeveloped in ASD and over-developed in the psychotic spectrum, with a similar polarization of language and behavioral phenotypes. Although speculative, this hypothesis has gained some traction. In the next several years, genomic, imaging, and model-systems approaches will

likely shed further light on the relationship between autism, schizophrenia and other neuropsychiatric disorders.

### 3.2. Sequencing familial forms of ASD

To this point, we have focused primarily on the complex interactions of polygenic networks as the major cause of ASD. However, this is not exclusively the case. Paralleling the recent spate of CNV studies is a renewed focus on rare disorders. These include familial forms of complex diseases that are potentially monogenic or with less complex inheritance pattern. At the outset of this chapter, we emphasized the overlap with fragile X syndrome, where one third of cases are co-morbid for ASD. As mentioned, fragile X is caused by a failure to express the protein coded by *FMR1*. However, mutations in *FMR1* do not always result in fragile X and can result in a phenotype more representative of ASD. Thus, Muhle *et al.* (2004) [64] found that 7-8% of idiopathic ASD cases may have mutations at the *FMR1* locus. Likewise, although mutations in *MECP2* are the common cause of Rett syndrome, certain mutations at the same locus have been associated with idiopathic autism (Carney *et al.* (2003).

X-linked genes encoding neuroligins *NLGN3*, *NLGN4* and *SHANK3* (a neuroligin binding partner) are other prominent examples of distinct rare genetic causes. A parallel can be drawn between these studies and studies of mental retardation and epilepsy, which include many rare syndromes that collectively account for a substantial proportion of the two disorders (Morrow *et al.*, 2008). Indeed it is perhaps more than coincidence that autism is heavily co-morbid with these two conditions, with ~40% of ASD cases meeting diagnostic criteria for mental retardation and epilepsy respectively (Bölte *et al.*, 2009; Danielsson *et al.*, 2005) [7,65]. It is also noteworthy that many of these monogenic-related genes are also major players in neurodevelopment and synapse activity. Other prominent examples include *TSC1*, *TSC2* (Osborne *et al.*, 1991; Franz, 1998) [66, 67], *NF1*, and *UBE3A* (see Morrow *et al.*, 2008) [68].

The identification of monogenic or possibly oligogenic autisms is likely to accelerate in the next several years as NGS becomes more widely available. In our group, we recently encountered a family of two parents, six healthy siblings, and two siblings with severe autism suggestive of autosomal recessive inheritance. Unsuccessful attempts using linkage and CNV approaches failed to identify a causal locus, but whole-exome sequencing at 20x coverage identified four genes, including one with a non-synonymous SNP in the protocadherin alpha 4 isoform1 precursor (*PCDHA4*) gene, which presents a strong candidate gene, currently under validation. Protocadherins are part of the cadherin family that facilitates neuronal cell adhesion and this discovery is consistent with the functional properties of the PCDH family.

Known syndromes with ASD features include fragile X, neurofibromatosis type 1, down syndrome, tuberous sclerosis, neurofibromatosis (which confers a 100-fold increased risk for ASD Li *et al.* (2005) [69], Angelman, Prader-Willi and related 15q syndromes, and at least several dozen others (see Zafeiriou *et al.*, 2007, for a comprehensive review) [70]. Table 1 from Volkmar *et al.* (2005) [71] lists the most commonly associated syndromes with median rate and range. It is likely that many more unidentified rare syndromes with Mendelian causes have ASD phenotypes. As of September 2012, the Online Mendelian Inheritance in

Man (OMIM) database listed over 7,000 known or suspected Mendelian diseases (MD), with ~3,500 (~50%) of these having an identified molecular basis (<http://omim.org/statistics/entry>). Since OMIM derives its data from published data, these figures likely under-represent rare disorders, which may go unreported. As such, there may be several times more Mendelian disorders that have no defined genetic etiology to date. Given the large-representation of autism phenotypes in known syndromes, we can assume a similar trend in unreported ASD syndromes.

The proportion of ASD accounted for by rare variants remains to be determined. Irrespective, as with many other aspects of scientific inquiry, the study of these events will continue to play an important role in explicating the pathogenesis of ASD. El-Fishawy and State (2010) [72] point to hypercholesterolemia and hypertension (Brown, 1974; Lifton *et al.*, 2001) [73,74] as examples where rare mutations have been successful in driving a molecular understanding of the disease as opposed to identifying risk factors in the general population. Rare mutations, particularly when they are Mendelian, carry large effects and are typically located in genic regions. These characteristics make the resolution of underlying networks distinctly less complex and, moreover, are amenable to modeling in other systems.

Recent groundbreaking studies by Marchetto *et al.* (2010) [75] and Muotri *et al.* (2010) [76], who created a cell culture model of Rett syndrome, are potentially exciting developments in this regard. Here, the researchers used skin biopsies from four Rett syndrome patients, each carrying a different *MECP2* mutation, to culture induced pluripotent stem cells (iPS). Once the iPS cells developed into neurons, they showed a decreased number of neurons and dendritic spines, consistent with neurodevelopmental disruptions. Intervention with insulin-like growth factor 1 (*IGF1*), which is known to regulate neurodevelopment, was subsequently shown to reverse Rett-like symptoms in a mouse model of the disease. This innovative approach is an exciting model of how rare gene approaches can stimulate our understanding of the pathophysiology and potential reversibility of ASD.

Syndrome	Number of Studies	Median Rate	Range %
Tuberous sclerosis	11	1.1	0–3.8
Fragile X	9	0.0	0–8.1
Down syndrome	12	0.7	0–16.7
Neurofibromatosis 1	6	0	0–1.4

**Table 1.** Associated disorders and their rate in autism (from Volkmar *et al.*, 2005 in Zafeiriou *et al.* 2007) [70,71]

3.3. Large-scale next-generation sequencing

In April 2012, *Nature* simultaneously published three papers that used exome sequencing to probe genomic correlates of ASD. This represented something of a landmark for both ASD and NGS research, as it demonstrated the viability of NGS on a large scale – the three studies combined examined 600 trios (parents and offspring), plus a 935 further ASD cases. Col-



lectively, these papers suggest that several hundred or more genes may be considered autism candidates, and again highlight the staggering complexity of the phenotype.

O’Roak *et al.* (2012) [77] sequenced 677 individual exomes from 209 families – primarily from the Simons Simplex Collection [78]. In 189 new probands, they validated 120 severely disruptive *de novo* mutations, 39% of which occur in a highly interconnected b-catenin/chromatin remodeling protein network. The group observed a strong paternal bias (41:10) in the rate of *de novo* mutations, which supports the hypothesis that the germline mutation rate in coding regions is markedly more prominent among males. These *de novo* events were more common in older fathers, marking paternal age as a significant risk factor for ASD.

Among the identified *de novo* loci, 62 were identified as top candidate mutations based on severity and/or supporting evidence from the literature. Interestingly, probands with these mutations were broadly distributed in terms of IQ score, with only a modest (non-significant) association with intellectual impairment. Recurrent protein-disruptive mutations were identified in two genes: netrin G1 (*NTNG1*) and chromodomain helicase DNA binding protein 8 (*CHD8*). *NTNG1* is known to play a role in axon guidance and dendritic organization (Nishimura-Akiyoshi *et al.*, 2007) [79]. *CHD8* regulates  $\beta$ -catenin and p53 signaling, and has not previously been associated with ASD. This gene was emphasized as particularly noteworthy, after follow-up protein-protein interaction (PPI) analyses, showed that  $\beta$ -catenin and p53 signaling may be features of an ASD-relevant network. In total 49 of proteins in the PPI network were highly interconnected, with a number of underlying genes also previously associated with neurodevelopment.

Neale *et al.* (2012) [80] exome-sequenced 175 trios and also focused on *de novo* mutations. As per the O’Roak study, there was a correlation between paternal age and *de novo* events for offspring ( $P < 0.0001$ ), and also for maternal age ( $P = 0.000365$ ). Across the sample set, the group observed 161 point mutations, of which 101 were missense, 50 silent, and 10 nonsense. Two conserved splice site rare single nucleotide variants and six frameshift insertions/deletions (indels) were also observed. Three genes were found to harbor two *de novo* mutations: *BRCA2* (two missense), *FAT1* (two missense) and *KCNMA1* (one missense, one silent).

The group next performed PPI analyses to determine whether interactions between genes associated with *de novo* mutations, as well as existing ASD candidates, was of etiological importance. This pathway approach, which additionally incorporated data from Sanders *et al.* study (below) [81], found that the distribution of functional *de novo* mutations is not random. The average distance for non-synonymous variants was significantly larger for controls *versus* cases (3.78 *vs.* 3.66;  $P = .033$ ). This suggests that a proportion of these *de novo* events contribute to autism. A model whereby *de novo* variants in up to 20% of cases, confer a 10- to 20-fold increased risk was supported.

In the third of these Nature papers, Sanders *et al.* (2012) [81] performed exome sequencing on 238 families, including 200 quartets (parents, 1 affected and 1 unaffected sibling) from the Simons Simplex Collection [78]. Comparing *de novo* non-synonymous single nucleotide variants (SNVs) between affected and unaffected siblings, the group observed a significantly

( $P=.01$ ) higher proportion among the probands (125 total) *versus* their unaffected sibling (87 total). From simulations, the authors concluded that two or more *de novo* nonsense/splice-site mutations should be considered significant. The gene sodium channel, voltage-gated, type II,  $\alpha$  subunit gene (*SCN2A*) was the only such gene – with two ASD individuals found to harbor relevant nonsense mutations. Mutations in *SCN2A* have been associated with epilepsy (Kamiya *et al.*, 2004; Ogiwara *et al.*, 2009) [82, 83] and idiopathic ASD in multiplex families (Weiss *et al.*, 2003) [84]. Neither of the probands has a history of seizures.

Combining the exomes from their study with those from O’Roak *et al.* ( $n$  for probands = 414), the groups identified two additional genes that each contained two loss-of-function mutations: the katanin p60 subunit A-like 2 (*KATNAL2*) and chromodomain helicase DNA binding protein 8 (*CHD8*). O’Roak *et al.* also evaluated these three novel candidates using exome sequencing on 935 cases and 870 controls. Three additional loss-of-function mutations each were observed in *KATNAL2* and *CHD8* in individuals with ASD, while none were identified in controls.

It is important to note, however, that for *de novo* events in general, there was no evidence to support the hypothesis that multiple events in any individual conferred an increased risk of ASD. As such, the ‘two *de novo* hit’ hypothesis is not supported.

In a fourth independent exome sequencing study involving 343 families from the Simons Simplex Collection Iossifov *et al.* (2012) [85] also reported a relatively equal distribution of *de novo* mutations in cases and controls. Again however, loss-of-function mutations—nonsense, splice site, and frame shifts—were more common in individuals with ASD (59 *versus* 28). Of the 59 “likely gene disruptions (LGD)” in ASD cases, none occurred more than once, although two—*NRXN1* and *PHF2*—had been identified in a previous CNV study by the same group (Gilman *et al.*, 2011) [86]. Intriguingly, the 59-strong LGD shared considerable overlap with a set of 842 proteins that interact with the fragile X protein, FMRP. In total, 14 of the 59 appeared on the FMRP list ( $P=.006$ ). Furthermore, 13 of 72 CNV candidates from the group’s previous CNV paper were also on the list ( $P=.0004$ ), meaning 26 of the combined 129 total were FMRP-related ( $P<1\times10^{-13}$ ).

The authors subsequently screened for *de novo* mutations in upstream targets of *FMR1*. One was identified – a deletion in *GRM5* that removes a single amino acid and causes an additional substitution at the same site. *GRM5* encodes the glutamate receptor mGluR5 (Bear *et al.*, 2004) [87] and, as noted below, mGluR5 antagonists are currently in clinical trial (Jacquemont *et al.*, 2011) [88] having indicated success in mouse models (Dölen *et al.*, 2007) [89]. Further elucidating the relationship between *FMR1*/FMRP and these ASD candidates is clearly an important next step in maximizing the impact of these findings. These are discussed further in the section below.

Collectively, all four of these exome sequencing studies converge upon the conclusion that ASD is highly heterogeneous, with several hundred or more loci potential risk variants. Simulations by the Neale *et al.* group confirm the statistical implausibility that hundreds of variants with high penetrance are possible, and a model where *de novo* variants in up to 20% of cases, confer ~10- to 20-fold increased risk is supported. The studies also converge on the

conclusion that paternal age (and possibly maternal age) is a significant ASD risk factor, but the frequency and size of *de novo* mutations *per se* is not. Evidence for three candidate genes—*CHD8*, *KATNAL2*, and *SCN2A*—would seem quite strong, though further functional studies are needed to help define pathogenesis. Perhaps most exciting is the association between *GRM5* and existing/novel candidates. As we have learned from GWAS, larger sample sets are clearly needed to fully harness the power of NGS in relation to such a complex phenotype. While these studies have been important in proposing novel candidates and confirming existing hypotheses of ASD, we await with anticipation results from the sequencing of all 2,648 families from the Simon Simplex Collection.

#### 4. Toward a treatment?

Ultimately, the primary goal of genome research should be to propose targets for intervention. As mentioned above, a number of translational studies have begun to probe the metabotropic glutamate receptor, mGluR5, as a potential target for fragile X syndrome treatment. These studies have a theoretical basis in the hypothesis that protein-synthesis-dependent functions of metabotropic receptors are exaggerated in fragile X syndrome (Bear, Huber & Warren, 2004) [87]. Thus, the fragile X protein, FMRP, is thought to work in functional opposition to mGluR5 (and mGluR1). Where FMRP is absent, mGluR-dependent protein synthesis becomes over-activated, resulting in neurological and behavioral abnormalities.

Dölen *et al.* (2007) [89] crossed *Fmr1* mutant with *Grm5* mutant to produce *Fmr1* knockouts who also had a selective reduction in mGluR5 expression. They found that a 50% reduction in mGluR5 gene dosage rescued a range of deficits in *Fmr1* mutants. Relevant measures included protein synthesis in hippocampus, density of dendritic spines (layer 3 pyramidal neurons), visual responsiveness, and cognitive performance (inhibitory avoidance – a hippocampus-dependent memory). This provides confirmation that mGluR5 and FMRP are functionally oppositional. Moreover, it suggests possible pharmacological avenues by which this genetic disease may be treated.

A range of translational studies have begun to target this pathway. These include efforts to inhibit the activity of individual mGluR5 (Jacquemont *et al.*, 2011; Berry-Kravis *et al.*, 2009) [88,90], and FMRP-regulated proteins (Paribello *et al.*, 2010) [91], NMDA (Wei *et al.*, 2012) [92], and GSK3 $\beta$  (lithium, Berry-Kravis *et al.*, 2008) [93], which have shown promise in open label and (in some instances) clinical trials (see Berry-Kravis *et al.*, 2011 for review) [94]. Moreover, these compounds may have clinical application to the broader ASD phenotype. Silverman *et al.* (2012) [95] recently reported that the mGluR5 antagonist, GRN-529, decreased ASD-related symptoms of autism in two different mouse models of the disease (repetitive grooming/repetitive jumping). In addition to the Iossifov *et al.* (2012) [85] sequencing study discussed above, Kelleher *et al.* (2012) [96] recently showed that idiopathic autism cases may have higher burden of mGluR5 variants. The group found that in 209 idiopathic cases, there was significant enrichment for rare functional variants in the mGluR5 pathway—namely the genes *TSC1*, *TSC2* and *SHANK3*, and *HOMER1*—relative to controls

(n=300). It is likely that drugs targeting the mGluR5 pathway, if/when approved for fragile X syndrome, will lead to human clinical trials for ASD. This translational approach – which delineates a direct route from gene discovery, through functional validation to treatment, is clearly the blueprint by which genome research can have tangible clinical impact.

## 5. Conclusions

ASD are clearly highly heritable disorders and advances in gene-finding technology in the past decade have rapidly accelerated gene discovery. As is typically the case, successive developments have made the problem more complex such that there are huge numbers of candidate genes, most of which remain to be replicated. In spite of this complexity, we can observe a number of patterns beginning to unfold 1) the relative scarcity of causal common variants, 2) the growing list of causal rare variants, and 3) the emergence of monogenic disorders with primary and secondary ASD phenotypes.

The monogenic autisms are particularly interesting from a treatment perspective, as they provide a mechanism for studying ASD phenotypes in model systems and are an obvious target for drug intervention. They are also amenable to clinical testing and the decreasing cost of research technologies means that this capacity is more widely available to clinicians. In fact, as the resolution of clinical instruments becomes more sophisticated, it is likely that the clinic will become a primary workplace for syndromic discovery.

A key requirement in driving gene discovery is the necessity of high-quality phenotype data. ASDs are notoriously heterogeneous, and are fractionated in terms of symptoms and trajectory. Mandy & Skuse (2008) [97] reviewed seven factor analysis studies of ASD symptoms, and found that all but one dissociated social and non-social factors. In a non-clinical sample of 3,000 twin pairs, Happé *et al.* (2006) [98] examined autistic-like traits and found consistently low correlations ( $r = 0.1-0.4$ ) between each of the core deficits on the autism spectrum. Endophenotypes, sub-components or sub-processes of the broader phenotype, may provide a productive avenue to disentangling some of this complexity. By filtering out all but a few discrete measures, we can theoretically increase the signal-to-noise ratio in genotype-phenotype associations. A number of endophenotypes for ASD have been associated with disease genes, including head circumference (associated with the *HOXA1* A218G polymorphism, Conciatori *et al.*, 2004) [99]; age at first word (associated with a quantitative trait locus on 7q35, Alarcón *et al.* 2005) [100]; delayed magnetoencephalography evoked responses to auditory stimuli (Roberts *et al.*, 2010) [101]; and enhanced perception (Mottron *et al.*, 2006) [102]. The endophenotype approach is arguably more consistent with rare-/mono-genic discovery, where a mutated network may not yield a diagnosis of autism *per se*, but nevertheless cause associated abnormalities. Note, this approach does not diminish the pleiotropic effects of genes involved in neurodevelopment, and only serves to make the point that the relevant genotype may associate with some but not all ASD features.

The converse, of course, is also true, as a large number of candidate genes contribute to the majority of known ASD. With ~80% of genes expressed in the brain it is likely that this num-



ber will continue to grow, and here again careful phenotyping is critical to identifying functional consequences. Ultimately, the primary goal is not to determine the frequency of variation/mutation in cases versus controls, but to determine the pathway(s) and gene networks that lead to pathology. We will also need to identify other major biological players such as epigenetic factors, RNA regulatory elements, and environmental exposures, which are critical components of the ASD equation. While daunting, the elucidation of these elements will doubtlessly take us closer to developing effective treatments for ASD. Given the current rate of progress, we have cause for cautious optimism in this regard.

## Author details

John J. Connolly<sup>1</sup> and Hakon Hakonarson<sup>1,2</sup>

<sup>1</sup> Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, USA

<sup>2</sup> Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia, USA

## References

- [1] Wing L, Gould J, Gillberg C. Autism spectrum disorders in the DSM-V: better or worse than the DSM-IV? *Res Dev Disabil.* 2011 Mar-Apr;32(2):768-73.
- [2] Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord.* 1994 Oct;24(5):659-85.
- [3] Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, Mawhood L, et al. Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *J Autism Dev Disord.* 1989 Jun;19(2):185-212.
- [4] Constantino JN, Davis SA, Todd RD, Schindler MK, Gross MM, Brophy SL, et al. 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.* 2003 Aug;33(4):427-33.
- [5] Levy SE, Mandell DS, Schultz RT. Autism. *Lancet.* 2009 Nov 7;374(9701):1627-38.
- [6] Dawson M, Soulières I, Gernsbacher MA, Mottron L. The level and nature of autistic intelligence. *Psychol Sci.* 2007 Aug;18(8):657-62.
- [7] Bolte S, Dziobek I, Poustka F. Brief report: The level and nature of autistic intelligence revisited. *J Autism Dev Disord.* 2009 Apr;39(4):678-82.
- [8] Skuse DH. Rethinking the nature of genetic vulnerability to autistic spectrum disorders. *Trends Genet.* 2007 Aug;23(8):387-95.

- [9] Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, et al. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med*. 1995 Jan;25(1):63-77.
- [10] Piven J, Palmer P, Jacobi D, Childress D, Arndt S. Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. *Am J Psychiatry*. 1997 Feb;154(2):185-90.
- [11] 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*. 2005 Sep;46(9):963-71.
- [12] Fombonne E. Epidemiology of pervasive developmental disorders. *Pediatr Res*. 2009 Jun;65(6):591-8.
- [13] O'Roak BJ, State MW. Autism genetics: strategies, challenges, and opportunities. *Autism Res*. 2008 Feb;1(1):4-17.
- [14] Richards RI, Holman K, Kozman H, Kremer E, Lynch M, Pritchard M, et al. Fragile X syndrome: genetic localisation by linkage mapping of two microsatellite repeats FRAXAC1 and FRAXAC2 which immediately flank the fragile site. *J Med Genet*. 1991 Dec;28(12):818-23.
- [15] Rogers SJ, Wehner DE, Hagerman R. The behavioral phenotype in fragile X: symptoms of autism in very young children with fragile X syndrome, idiopathic autism, and other developmental disorders. *J Dev Behav Pediatr*. 2001 Dec;22(6):409-17.
- [16] Harris SW, Hessel D, Goodlin-Jones B, Ferranti J, Bacalman S, Barbato I, et al. Autism profiles of males with fragile X syndrome. *Am J Ment Retard*. 2008 Nov;113(6):427-38.
- [17] Curtis AR, Headland S, Lindsay S, Thomas NS, Boye E, Kamakari S, et al. X chromosome linkage studies in familial Rett syndrome. *Hum Genet*. 1993 Jan;90(5):551-5.
- [18] Ramocki MB, Zoghbi HY. Failure of neuronal homeostasis results in common neuropsychiatric phenotypes. *Nature*. 2008 Oct 16;455(7215):912-8.
- [19] Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005 Apr 15;308(5720):385-9.
- [20] Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science*. 1996 Sep 13;273(5281):1516-7.
- [21] Wang K, Zhang H, Ma D, Bucan M, Glessner JT, Abrahams BS, et al. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*. 2009 May 28;459(7246):528-33.
- [22] Ma D, Salyakina D, Jaworski JM, Konidari I, Whitehead PL, Andersen AN, et al. A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann Hum Genet*. 2009 May;73(Pt 3):263-73.
- [23] St Pourcain B, Wang K, Glessner JT, Golding J, Steer C, Ring SM, et al. Association between a high-risk autism locus on 5p14 and social communication spectrum phenotypes in the general population. *Am J Psychiatry*. 2010 Nov;167(11):1364-72.

- [24] Redies C, Hertel N, Hubner CA. Cadherins and neuropsychiatric disorders. *Brain Res.* 2012 Aug 27;1470:130-44.
- [25] Gepner B, Feron F. Autism: a world changing too fast for a mis-wired brain? *Neurosci Biobehav Rev.* 2009 Sep;33(8):1227-42.
- [26] Kerin T, Ramanathan A, Rivas K, Grepo N, Coetzee GA, Campbell DB. A noncoding RNA antisense to moesin at 5p14.1 in autism. *Sci Transl Med.* 2012 Apr 4;4(128):128ra40.
- [27] McClellan J, King MC. Genetic heterogeneity in human disease. *Cell.* 2010 Apr 16;141(2):210-7.
- [28] Li Q, Yu K. Improved correction for population stratification in genome-wide association studies by identifying hidden population structures. *Genet Epidemiol.* 2008 Apr;32(3):215-26.
- [29] Ozsolak F, Milos PM. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet.* 2011 Feb;12(2):87-98.
- [30] Alarcon M, Cantor RM, Liu J, Gilliam TC, Geschwind DH. Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. *Am J Hum Genet.* 2002 Jan;70(1):60-71.
- [31] Alarcon M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM, et al. Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *Am J Hum Genet.* 2008 Jan;82(1):150-9.
- [32] Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, et al. A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *Am J Hum Genet.* 2008 Jan;82(1):160-4.
- [33] Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, et al. A functional genetic link between distinct developmental language disorders. *N Engl J Med.* 2008 Nov 27;359(22):2337-45.
- [34] Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature.* 2001 Oct 4;413(6855):519-23.
- [35] Cheng Y, Sudarov A, Szulc KU, Sgaier SK, Stephen D, Turnbull DH, et al. The Engrailed homeobox genes determine the different foliation patterns in the vermis and hemispheres of the mammalian cerebellum. *Development.* 2010 Feb;137(3):519-29.
- [36] Benayed R, Choi J, Matteson PG, Gharani N, Kamdar S, Brzustowicz LM, et al. Autism-associated haplotype affects the regulation of the homeobox gene, ENGRAILED 2. *Biol Psychiatry.* 2009 Nov 15;66(10):911-7.
- [37] Benayed R, Gharani N, Rossman I, Mancuso V, Lazar G, Kamdar S, et al. Support for the homeobox transcription factor gene ENGRAILED 2 as an autism spectrum disorder susceptibility locus. *Am J Hum Genet.* 2005 Nov;77(5):851-68.
- [38] Wang L, Jia M, Yue W, Tang F, Qu M, Ruan Y, et al. Association of the ENGRAILED 2 (EN2) gene with autism in Chinese Han population. *Am J Med Genet B Neuropsychiatr Genet.* 2008 Jun 5;147B(4):434-8.

- [39] Zhong H, Serajee FJ, Nabi R, Huq AH. No association between the EN2 gene and autistic disorder. *J Med Genet*. 2003 Jan;40(1):e4.
- [40] Campbell DB, Sutcliffe JS, Ebert PJ, Militerni R, Bravaccio C, Trillo S, et al. A genetic variant that disrupts MET transcription is associated with autism. *Proc Natl Acad Sci U S A*. 2006 Nov 7;103(45):16834-9.
- [41] Campbell DB, Li C, Sutcliffe JS, Persico AM, Levitt P. Genetic evidence implicating multiple genes in the MET receptor tyrosine kinase pathway in autism spectrum disorder. *Autism Res*. 2008 Jun;1(3):159-68.
- [42] Sousa I, Clark TG, Toma C, Kobayashi K, Choma M, Holt R, et al. MET and autism susceptibility: family and case-control studies. *Eur J Hum Genet*. 2009 Jun;17(6):749-58.
- [43] A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am J Hum Genet*. 2001 Sep;69(3):570-81.
- [44] Eagleson KL, Campbell DB, Thompson BL, Bergman MY, Levitt P. The autism risk genes MET and PLAUR differentially impact cortical development. *Autism Res*. 2011 Feb;4(1):68-83.
- [45] Glessner JT, Hakonarson H. Common variants in polygenic schizophrenia. *Genome Biol*. 2009;10(9):236.
- [46] Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet*. 2008 Sep;40(9):1056-8.
- [47] Arcos-Burgos M, Jain M, Acosta MT, Shively S, Stanescu H, Wallis D, et al. A common variant of the latrophilin 3 gene, LPHN3, confers susceptibility to ADHD and predicts effectiveness of stimulant medication. *Mol Psychiatry*. 2010 Nov;15(11):1053-66.
- [48] Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. *PLoS Biol*. 2010 Jan;8(1):e1000294.
- [49] Wang K, Dickson SP, Stolle CA, Krantz ID, Goldstein DB, Hakonarson H. Interpretation of association signals and identification of causal variants from genome-wide association studies. *Am J Hum Genet*. 2010 May 14;86(5):730-42.
- [50] Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*. 2010 Jul 15;466(7304):368-72.
- [51] Williams NM, Zaharieva I, Martin A, Langley K, Mantripragada K, Fossdal R, et al. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. *Lancet*. 2010 Oct 23;376(9750):1401-8.
- [52] Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature*. 2009 May 28;459(7246):569-73.
- [53] Levinson DF, Duan J, Oh S, Wang K, Sanders AR, Shi J, et al. Copy Number Variants in Schizophrenia: Confirmation of Five Previous Findings and New Evidence for



- 3q29 Microdeletions and VIPR2 Duplications. *Am J Psychiatry*. 2011 Mar;168(3):302-16.
- [54] Chen X, Li X, Wang P, Liu Y, Zhang Z, Zhao G, et al. Novel association strategy with copy number variation for identifying new risk Loci of human diseases. *PLoS One*. 2010;5(8):e12185.
- [55] Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, et al. Strong association of de novo copy number mutations with autism. *Science*. 2007 Apr 20;316(5823):445-9.
- [56] Bucan M, Abrahams BS, Wang K, Glessner JT, Herman EI, Sonnenblick LI, et al. Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genet*. 2009 Jun;5(6):e1000536.
- [57] Guilmatre A, Dubourg C, Mosca AL, Legallic S, Goldenberg A, Drouin-Garraud V, et al. Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation. *Arch Gen Psychiatry*. 2009 Sep;66(9):947-56.
- [58] Samaco RC, Hogart A, LaSalle JM. Epigenetic overlap in autism-spectrum neurodevelopmental disorders: MECP2 deficiency causes reduced expression of UBE3A and GABRB3. *Hum Mol Genet*. 2005 Feb 15;14(4):483-92.
- [59] Kim HG, Kishikawa S, Higgins AW, Seong IS, Donovan DJ, Shen Y, et al. Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet*. 2008 Jan; 82(1):199-207.
- [60] Yi JJ, Ehlers MD. Ubiquitin and protein turnover in synapse function. *Neuron*. 2005 Sep 1;47(5):629-32.
- [61] Gai X, Xie HM, Perin JC, Takahashi N, Murphy K, Wenocur AS, et al. Rare structural variation of synapse and neurotransmission genes in autism. *Mol Psychiatry*. 2011 Mar 1.
- [62] Glessner JT, Reilly MP, Kim CE, Takahashi N, Albano A, Hou C, et al. Strong synaptic transmission impact by copy number variations in schizophrenia. *Proc Natl Acad Sci U S A*. 2010 Jun 8;107(23):10584-9.
- [63] Crespi B, Badcock C. Psychosis and autism as diametrical disorders of the social brain. *Behav Brain Sci*. 2008 Jun;31(3):241-61; discussion 61-320.
- [64] Muhle R, Trentacoste SV, Rapin I. The genetics of autism. *Pediatrics*. 2004 May; 113(5):e472-86.
- [65] Danielsson S, Gillberg IC, Billstedt E, Gillberg C, Olsson I. Epilepsy in young adults with autism: a prospective population-based follow-up study of 120 individuals diagnosed in childhood. *Epilepsia*. 2005 Jun;46(6):918-23.
- [66] Osborne JP, Fryer A, Webb D. Epidemiology of tuberous sclerosis. *Ann N Y Acad Sci*. 1991;615:125-7.
- [67] Franz DN. Diagnosis and management of tuberous sclerosis complex. *Semin Pediatr Neurol*. 1998 Dec;5(4):253-68.

- [68] Morrow EM, Yoo SY, Flavell SW, Kim TK, Lin Y, Hill RS, et al. Identifying autism loci and genes by tracing recent shared ancestry. *Science*. 2008 Jul 11;321(5886):218-23.
- [69] Li W, Cui Y, Kushner SA, Brown RA, Jentsch JD, Frankland PW, et al. The HMG-CoA reductase inhibitor lovastatin reverses the learning and attention deficits in a mouse model of neurofibromatosis type 1. *Curr Biol*. 2005 Nov 8;15(21):1961-7.
- [70] Zafeiriou DI, Ververi A, Vargiami E. Childhood autism and associated comorbidities. *Brain Dev*. 2007 Jun;29(5):257-72.
- [71] Volkmar FR. *Handbook of Autism and Pervasive Developmental Disorders*. 3 ed. New Jersey: John Wiley & Sons; 2005.
- [72] El-Fishawy P, State MW. The genetics of autism: key issues, recent findings, and clinical implications. *Psychiatr Clin North Am*. 2010 Mar;33(1):83-105.
- [73] Brown MS, Goldstein JL. Expression of the familial hypercholesterolemia gene in heterozygotes: mechanism for a dominant disorder in man. *Science*. 1974 Jul 5;185(4145):61-3.
- [74] Lifton RP, Gharavi AG, Geller DS. Molecular mechanisms of human hypertension. *Cell*. 2001 Feb 23;104(4):545-56.
- [75] Marchetto MC, Carromeu C, Acab A, Yu D, Yeo GW, Mu Y, et al. A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell*. 2010 Nov 12;143(4):527-39.
- [76] Muotri AR, Marchetto MC, Coufal NG, Oefner R, Yeo G, Nakashima K, et al. L1 retrotransposition in neurons is modulated by MeCP2. *Nature*. 2010 Nov 18;468(7322):443-6.
- [77] O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*. 2012 Apr 4.
- [78] Fischbach GD, Lord C. The Simons Simplex Collection: a resource for identification of autism genetic risk factors. *Neuron*. 2010 Oct 21;68(2):192-5.
- [79] Nishimura-Akiyoshi S, Niimi K, Nakashiba T, Itohara S. Axonal netrin-Gs trans-neuronally determine lamina-specific subdendritic segments. *Proc Natl Acad Sci U S A*. 2007 Sep 11;104(37):14801-6.
- [80] Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature*. 2012 Apr 4.
- [81] Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature*. 2012 Apr 4.
- [82] Kamiya K, Kaneda M, Sugawara T, Mazaki E, Okamura N, Montal M, et al. A non-sense mutation of the sodium channel gene SCN2A in a patient with intractable epilepsy and mental decline. *J Neurosci*. 2004 Mar 17;24(11):2690-8.

- [83] Ogiwara I, Ito K, Sawaishi Y, Osaka H, Mazaki E, Inoue I, et al. De novo mutations of voltage-gated sodium channel  $\alpha$ II gene SCN2A in intractable epilepsies. *Neurology*. 2009 Sep 29;73(13):1046-53.
- [84] Weiss LA, Escayg A, Kearney JA, Trudeau M, MacDonald BT, Mori M, et al. Sodium channels SCN1A, SCN2A and SCN3A in familial autism. *Mol Psychiatry*. 2003 Feb; 8(2):186-94.
- [85] Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, et al. De novo gene disruptions in children on the autistic spectrum. *Neuron*. 2012 Apr 26;74(2):285-99.
- [86] 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*. 2011 Jun 9;70(5):898-907.
- [87] Bear MF, Huber KM, Warren ST. The mGluR theory of fragile X mental retardation. *Trends Neurosci*. 2004 Jul;27(7):370-7.
- [88] Jacquemont S, Curie A, des Portes V, Torrioli MG, Berry-Kravis E, Hagerman RJ, et al. Epigenetic modification of the FMR1 gene in fragile X syndrome is associated with differential response to the mGluR5 antagonist AFQ056. *Sci Transl Med*. 2011 Jan 5;3(64):64ra1.
- [89] Dolen G, Osterweil E, Rao BS, Smith GB, Auerbach BD, Chattarji S, et al. Correction of fragile X syndrome in mice. *Neuron*. 2007 Dec 20;56(6):955-62.
- [90] Berry-Kravis E, Hessel D, Coffey S, Hervey C, Schneider A, Yuhas J, et al. A pilot open label, single dose trial of fenobam in adults with fragile X syndrome. *J Med Genet*. 2009 Apr;46(4):266-71.
- [91] Paribello C, Tao L, Folino A, Berry-Kravis E, Tranfaglia M, Ethell IM, et al. Open-label add-on treatment trial of minocycline in fragile X syndrome. *BMC Neurol*. 2010;10:91.
- [92] Wei H, Dobkin C, Sheikh AM, Malik M, Brown WT, Li X. The therapeutic effect of memantine through the stimulation of synapse formation and dendritic spine maturation in autism and fragile X syndrome. *PLoS One*. 2012;7(5):e36981.
- [93] Berry-Kravis E, Sumis A, Hervey C, Nelson M, Porges SW, Weng N, et al. Open-label treatment trial of lithium to target the underlying defect in fragile X syndrome. *J Dev Behav Pediatr*. 2008 Aug;29(4):293-302.
- [94] Berry-Kravis E, Knox A, Hervey C. Targeted treatments for fragile X syndrome. *J Neurodev Disord*. 2011 Sep;3(3):193-210.
- [95] Silverman JL, Smith DG, Rizzo SJ, Karras MN, Turner SM, Tolu SS, et al. Negative allosteric modulation of the mGluR5 receptor reduces repetitive behaviors and rescues social deficits in mouse models of autism. *Sci Transl Med*. 2012 Apr 25;4(131):131ra51.
- [96] Kelleher RJ, 3rd, Geigenmuller U, Hovhannisyan H, Trautman E, Pinard R, Rathmell B, et al. High-throughput sequencing of mGluR signaling pathway genes reveals enrichment of rare variants in autism. *PLoS One*. 2012;7(4):e35003.

- [97] Mandy WP, Skuse DH. Research review: What is the association between the social-communication element of autism and repetitive interests, behaviours and activities? *J Child Psychol Psychiatry*. 2008 Aug;49(8):795-808.
- [98] Happe F, Ronald A, Plomin R. Time to give up on a single explanation for autism. *Nat Neurosci*. 2006 Oct;9(10):1218-20.
- [99] Conciatori M, Stodgell CJ, Hyman SL, O'Bara M, Militerni R, Bravaccio C, et al. Association between the HOXA1 A218G polymorphism and increased head circumference in patients with autism. *Biol Psychiatry*. 2004 Feb 15;55(4):413-9.
- [100] Alarcon M, Yonan AL, Gilliam TC, Cantor RM, Geschwind DH. Quantitative genome scan and Ordered-Subsets Analysis of autism endophenotypes support language QTLs. *Mol Psychiatry*. 2005 Aug;10(8):747-57.
- [101] Roberts TP, Khan SY, Rey M, Monroe JF, Cannon K, Blaskey L, et al. MEG detection of delayed auditory evoked responses in autism spectrum disorders: towards an imaging biomarker for autism. *Autism Res*. 2010 Feb;3(1):8-18.
- [102] Mottron L, Dawson M, Soulières I, Hubert B, Burack J. Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *J Autism Dev Disord*. 2006 Jan;36(1):27-43.



