Recent Advances in the Genetics of Autism

Abha R. Gupta and Matthew W. State

Autism is a strongly genetic disorder, with an estimated heritability of greater than 90%. Nonetheless, its specific genetic etiology remains largely unknown. Over the past several years, the convergence of rapidly advancing genomic technologies, the completion of the human genome project, and successful collaborative efforts to increase the number of deoxyribonucleic acid samples available for study have led to the first solid clues regarding the genetic origins of autism spectrum disorders. This article addresses the obstacles that have confronted gene discovery efforts and reviews recent linkage, cytogenetic, and candidate gene association studies relevant to autism spectrum disorders. In addition, promising avenues for future research and the potential contribution of emerging genomic technologies are considered.

Key Words: Association, autism spectrum disorder, candidate genes, cytogenetics, genetics, linkage

The search for genes contributing to autism spectrum disorders (ASDs), including autism, Asperger syndrome, and pervasive developmental disorder not otherwise specified (PDD-NOS), has been a long and arduous one. Despite several decades of evidence supporting a significant genetic contribution, the specific variants causing or increasing the risk for these debilitating disorders remain largely elusive. Fortunately, in the last several years, the prospects for success have improved dramatically as a consequence of the sequencing of the human genome, the increasing speed and sensitivity and decreasing costs of genomic technologies, and a successful collaborative effort to expand the number of well-characterized patients widely available for study. As a result, the first insights into the cellular and molecular mechanisms underlying autism likely have already been made, and it is reasonable to expect even more dramatic advances in the near future.

This review will address the challenges facing investigators interested in the genetic underpinnings of autism, with a particular emphasis on the issue of genetic heterogeneity. It then will provide a summary of the recent findings from genomewide linkage studies as well as cytogenetic and candidate gene analyses. The discussion will conclude with a consideration of the role that new technologies and analytic approaches may play in accelerating the already impressive rate of progress in the field.

Given the breadth and depth of autism genetics research over more than a decade, this brief review cannot hope to be exhaustive. In addition, there are multiple related topics that will not be addressed here, including the epidemiology of ASDs, the possible contribution of environmental factors and epigenetics (i.e., heritable changes that are not coded for in the sequence of the DNA), the potential value of animal model systems, and the role that neuroimaging may play in genetic investigations. The present discussion will focus instead on highlighting recent progress in gene identification and illustrating the alternative approaches to discovery that now are coalescing to shed light on the molecular and cellular origins of ASDs.

From the Child Study Center (ARG, MWS), and Departments of Pediatrics (ARG) and Genetics (MWS), Yale University School of Medicine, New Haven, Connecticut.

Address reprint requests to Matthew W. State, M.D., Ph.D., Genetics and Child Psychiatry, Yale University School of Medicine, 230 South Frontage Road, New Haven, CT 06520; E-mail: matthew.state@yale.edu.

Received March 20, 2006; revised June 19, 2006; accepted June 20, 2006.

The Challenge of Heterogeneity and the Genetic Architecture of Autism

Certain aspects of the genetics of autism generate little debate. For example, autism is acknowledged to be among the most heritable of neuropsychiatric disorders. Monozygotic twins have been found to share the diagnosis far more than dizygotic twins (Bailey et al 1995), suggesting that the observed familial clustering largely is explained by genetic factors. Moreover, observations about the patterns of inheritance among families, including studies of twins, first-degree relatives, and more distantly related individuals, all suggest that ASDs are not caused by the action of a single gene transmitted in a simple dominant, recessive, or X-linked fashion. Indeed, the number of contributing loci, or genetic regions, has been estimated to be in the neighborhood of 15 (Risch et al 1999), though it is conceivable that even this may be a marked underestimate.

Despite this general agreement, key questions remain regarding the specifics of the underlying genetic architecture of ASDs. A major unresolved issue is whether this group of relatively common disorders, estimated collectively to affect 60 in 10,000 individuals (Fombonne 2005), is largely the consequence of genetic variations that also are common in the general population. A second important issue is to what extent the study of individuals or families that have the condition as a result of "simple" genetics may contribute to understanding this genetically complex spectrum of disorders.

There are reasons to suspect that common variation will be found to play a large role in the etiology of ASDs. Within the normal human genome, most of the variation in a given population is accounted for by polymorphisms that are present in more than 1% of individuals. These, by definition, are considered common alleles. It would be logical then to hypothesize that a disease that is common would reflect this overall architecture. This would hold true as long as natural selection against these alleles did not play a significant role.

The fact that fundamental impairments in social reciprocity could have a large effect on reproductive fitness might suggest that ASDs violate this basic assumption. However, the overall pattern of inheritance across families and the observation that relatives of affected individuals may show subtle symptoms suggest that common forms of ASD may be the result of a conspiracy of multiple genes, each contributing small increments of risk to the individual. In this case, each polymorphism might not be subject to marked negative selection. In fact, it is conceivable that individually, some of the genetic changes contributing to ASDs could possess selective advantages. These variants might then be expected to be present at high population frequencies.

Although such a model certainly is plausible, there also is both theoretical and empirical evidence that rare genetic varia-

0006-3223/07/$32.00
doi:10.1016/j.biopsych.2006.06.020

tion contributes to ASDs. Although it often is taken as such, common disease is, in fact, not synonymous with common variation. There are several reasonable alternatives that could alone or in combination account for the relatively high population frequency of ASDs as well as for the large observed difference between the concordance rates for monozygotic and dizygotic twins. These might include a high rate of de novo mutation and the contribution of marked locus and allelic heterogeneity, that is, the presence of multiple rare alleles affecting multiple genes capable of increasing the risk for the phenotype. Indeed, there already is considerable experimental evidence supporting a role for rare variation contributing to ASDs, as will be discussed in more detail.

The question of whether ASDs are likely to be the consequence of common or rare variation is of more than academic interest. For instance, those methods and study designs that may be most appropriate to identify one type of risk may have little ability to identify the other. A widely used methodology known as candidate gene association typically involves an investigation of known common genetic polymorphisms in or near a gene of interest. The alleles are evaluated to determine whether they are correlated with risk in an affected population. This approach is quite practical, because the relatively high frequency of the variation being studied helps ensure that a sufficient number of occurrences may be observed in a reasonably sized sample to offer the potential for a statistically significant conclusion.

However, this strategy may be blind to rare variations. For example, more than a dozen studies have investigated the contribution of the long and short alleles of the serotonin transporter to autism (reviewed in Bacchelli and Maestrini 2006). Typically, such studies would not be able to determine whether multiple rare functional genetic changes also were present in affected individuals. This possibility only could be evaluated by using a method, such as direct DNA sequencing, that is able to query the DNA sequence of the gene in an unbiased fashion. When this type of analysis recently was undertaken in a group of affected families, several potentially important rare alleles were identified (Sutcliffe et al 2005).

The search for multiple rare variants, however, carries its own liabilities. For example, the costs of direct sequencing typically are many times greater than are the costs of genotyping a small number of known polymorphisms. In addition, the infrequency with which one observes any given allele may pose significant challenges, including designing and funding studies with sufficient numbers of patients and controls to support a statistically significant conclusion.

A second question with regard to the issue of common versus rare variation is whether there are useful examples of Mendelian genetics at work in the context of what widely is acknowledged to be a complex genetic disease. In fact, there are many patients already identified who have symptoms indistinguishable from idiopathic autism who also have chromosomal abnormalities or well-described genetic syndromes that are transmitted in a Mendelian fashion (reviewed in Filipek 2005; Vorstman et al 2006). These observations would appear to put this question to rest. Nonetheless, there remains an ongoing debate regarding the extent to which investigation of either genetic syndromes or so-called outlier cases are of value in the quest to understand common forms of the illness. This issue will be considered in more depth subsequently in the section entitled Cytogenetic Analyses with regard to the recent findings regarding the gene *NLGN4X* (neuroligin 4 on the X chromosome).

Linkage Studies

When the pathophysiology of a condition is poorly understood, gene discovery methods that rely on basic genetic principles, as opposed to etiologic hypotheses, may be particularly valuable. Linkage studies are one such approach, often referred to as positional cloning. These analyses seek to determine whether the transmission of a chromosomal segment from one generation to another coincides with the presence of the phenotype of interest. If every chromosome is evaluated simultaneously, the study is referred to as genomewide linkage.

In situations in which there is unlikely to be Mendelian inheritance, nonparametric linkage studies may be undertaken that do not require a starting hypothesis regarding the specific nature of the genetic transmission. A particularly popular version of this approach is the affected sib-pair study, in which siblings with ASDs are evaluated to determine whether they share any regions of the genome more frequently than would be expected by chance.

Multiple genomewide scans have been published evaluating ASDs, and evidence in favor of linkage has been reported for the majority of chromosomes (Table 1). As a general proposition, this evidence has not reached statistical significance, a threshold that typically is quantified by a logarithm of the odds (LOD) score. This statistic represents the logarithm of the likelihood ratio of observing the data under a model of linkage compared with observing the data under a model of free recombination (no linkage). According to the most widely accepted criteria, a LOD score of 3.6 in a sib-pair analysis is taken as evidence for significant linkage (Lander and Kruglyak 1995). A LOD score of 2.2 is considered suggestive evidence, and a LOD score of 5.4 is thought to be highly significant. Translated into a more tangible form, one would expect to see, by chance, a suggestive linkage peak once in every genome scan or a significant peak approximately once every 20 scans.

Despite increasing sample sizes and considerable methodologic sophistication, there has been a vexing absence of direct agreement among the more than dozen genomewide linkage scans published in the past decade. Nonetheless, some interesting patterns have emerged. For instance, multiple investigators have identified regions on chromosomes 7 and 17 showing suggestive or significant linkage to ASD (Table 1).

Chromosome 7q is one of the most frequently implicated regions in genomewide studies. These findings are a Rorschach test of sorts for geneticists in the field, generating either considerable optimism or much consternation. Although many groups have found hints that the long arm of this chromosome may contain an autism allele, no study has replicated any other at precisely the same locus. In addition, in the largest genomewide nonparametric study published to date, no significant or suggestive evidence for linkage was identified across this region (Yonan et al 2003).

Still, 7q remains an area of interest for a number of reasons. First, as noted, multiple suggestive linkage signals have been reported, and it was the only region supported by a meta-analysis of four published studies (Badner and Gershon 2002). It is not uncommon for peaks in complex genetic disorders to be wide and shift between studies (Rutter 2005). This may be because of the presence of multiple risk loci on 7q that differ across study samples. Second, a number of chromosomal rearrangements involving this interval have been identified in patients with ASDs (reviewed in Vorstman et al 2006). Finally, numerous brain-expressed transcripts map to 7q and have

Table 1. Linkage Studies in Autism

Chromosome	Marker	LOD Score	Reference	Genes Tested[a]
1p13.2	D1S1675	2.63	Auranen et al (2002)	
1q21.3	D1S498	2.32	Auranen et al (2002)	
1q22	D1S2721	2.88	Ylisaukko-oja et al (2004)	
1q23.3	D1S484	3.58	Ylisaukko-oja et al (2004)	
1q42.2	D1S1656	3.06	Buxbaum et al (2004)	
2q31.1	D2S2188	4.80	International Molecular Genetic Study of Autism Consortium (2001)	ATF2, cAMP-GEFII, CHN1, DLX1, DLX2, GAD1, INPP1, NEUROD1, SLC11A3, SLC25A12
2q31.1	D2S335	3.32	Buxbaum et al (2001)	ATF2, cAMP-GEFII, CHN1, DLX1, DLX2, GAD1, INPP1, NEUROD1, SLC11A3, SLC25A12
3p24.1	D3S2432	3.32	Ylisaukko-oja et al (2004)	
3p25.3	**D3S3691**	**2.22**	**McCauley et al (2005)**	
3q22.1	D3S3045–D3S1763	3.10	Alarcon et al (2005)	
3q26.32	D3S3715, D3S3037	4.81	Auranen et al (2002)	
4q23	**D4S1647**	2.87	**Buxbaum et al (2004)**	
4q27	D4S3250	2.73	Buxbaum et al (2004)	
4q32.3	D4S2368	2.82	Ylisaukko-oja et al (2004)	TDO2
5p13.1	**D5S2494**	**2.55**	**Liu et al (2001)**	
5p13.1	**D5S2494**	**2.54**	**Yonan et al (2003)**	
6q14.3	D6S1270	2.61	Buxbaum et al (2004)	GABRR1, GABRR2
6q16.3	**D6S283**	**2.23**	**Philippe et al (1999)**	**GABRR1, GABRR2, GRIK2**
7q21.2	**D7S1813**	**2.2**	**Barrett et al (1999)**	**DLX6, PCLO, PON1**
7q22.1	D7S477	3.55	International Molecular Genetic Study of Autism Consortium (2001)	CUTL1, PAI1, PIK3CG, RELN, SRPK2, SYPL
7q32.1–34	D7S530–D7S684	3.55	International Molecular Genetic Study of Autism Consortium (1998)	COPG2, CPA1, CPA5, FOXP2, GRM8, KIAA0716, LAMB1, LRRN3, NRCAM, PEG1/MEST, PTPRZ1, RAY1, UBE2H, WNT2
7q34–36.2	D7S1824–D7S3058	2.98	Alarcon et al (2002)	EN2
7q36.1	**D7S483**	**3.7**	**Molloy et al (2005)**	**EN2**
9p22.2	D9S157	3.11	International Molecular Genetic Study of Autism Consortium (2001)	
9q34.3	D9S1826	3.59	International Molecular Genetic Study of Autism Consortium (2001)	DBH, TSC1
11p11.2–13	**D11S1392–D11S1993**	**2.24**	**Yonan et al (2003)**	**BDNF**
13q12.3	**D13S217–D13S1229**	**2.3**	**Barrett et al (1999)**	
13q22.1	**D13S800**	**3.0**	**Barrett et al (1999)**	
13q32.1–32.3	D13S793–D13S1271	2.86	Ylisaukko-oja et al (2004)	
15q21.2	CYP19	2.21	International Molecular Genetic Study of Autism Consortium (2001)	DYX1C1
16p13.13	**D16S3102**	**2.93**	**International Molecular Genetic Study of Autism Consortium (2001)**	**A2BP1, ABAT, ABCC1, BFAR, CREBBP, EMP2, GRIN2A, MRTF-B, PRKCB1, SSTR5, TSC2, UBN1**
16p13.2	D16S407	2.22	International Molecular Genetic Study of Autism Consortium (2001)	A2BP1, ABAT, ABCC1, BFAR, CREBBP, EMP2, GRIN2A, MRTF-B, PRKCB1, SSTR5, TSC2, UBN1
17p11.2	D17S1298–D17S1299	2.22	Alarcon et al (2005)	
17q11.2	D17S1294–D17S798	4.3	Stone et al (2004)	NF1, OMGP62, SLC6A4
17q11.2	**D17S1294**	**2.85**	**McCauley et al (2005)**	**NF1, OMGP62, SLC6A4**
17q11.2	**D17S1800**	**2.83**	**Yonan et al (2003)**	**NF1, OMGP62, SLC6A4**
17q11.2	**HTTINT2**	**2.34**	**International Molecular Genetic Study of Autism Consortium (2001)**	**NF1, OMGP62, SLC6A4**
17q21.2	**D17S1299**	**2.26**	**McCauley et al (2005)**	**HOXB1**
17q21.32	D17S2180	4.1	Cantor et al (2005)	HOXB1
17q24.3	D17S1290–D17S1301	2.84	Alarcon et al (2005)	
19p13.11	**D19S930**	**2.77**	**McCauley et al (2005)**	**RAB3A**
19p13.12	D19S714	2.53	Liu et al (2001)	RAB3A
19p13.12	D19S714	2.31	Buxbaum et al (2004)	RAB3A
21q21.1	**D21S1437**	**3.4**	**Molloy et al (2005)**	
Xq21.33	**DXS6789**	**2.54**	**Shao et al (2002b)**	
Xq25	DXS1047	2.67	Liu et al (2001)	

Results from genomewide linkage analyses are given with a focus on intervals generating LOD scores of ≥2.2 (threshold for suggestive linkage as per Lander and Kruglyak [1995]). Three groups report several additional loci with LOD of ≥2.2; only loci for which markers were published are included in the table (Alarcon et al 2002, 2005; McCauley et al 2005). Boldface results indicate that the entire patient set was analyzed; not boldface results were obtained from a subset of patients who were stratified by various diagnostic criteria, which often differ between studies. Candidate genes that map within one chromosomal band of the linkage peak and that have been investigated in at least one published study are listed.

LOD, logarithm of the odds.

[a]References for genes available upon request.

known functions that plausibly could be involved in the pathophysiology of ASDs. These include *EN2* (engrailed 2), discussed in more detail in Candidate Gene Studies.

In light of the difficulties encountered in attempts to identify an autism gene by using linkage analysis, many investigators have turned to experimenting with alternative approaches to phenotypic assignment. The rationale is that more phenotypically homogenous subgroups may cull out more genetically homogenous samples, resulting in more consistency in the data. One example is the study of a phenotype involving the acquisition of phrased speech at a late age (older than 36 months). A study of 95 families reported a LOD score of 2.39 at 2q31.3. When a subset of 49 families meeting a "narrow" diagnosis of autism and having phrase speech delay (PSD) was analyzed, the evidence in favor of linkage improved, with a LOD score of 3.32 (Buxbaum et al 2001). A second group performed focused linkage analysis of chromosome 2q and found a score of 1.12 at 2q33 in 99 families. This subsequently improved to 2.86 in a subset of 45 families with PSD (Shao et al 2002a).

These examples suggest that there is potential value in stratifying patients as well as in identifying endophenotypes, that is, measurable heritable traits that are present along the biological path from gene to syndrome. This intuitively is an attractive notion. If autism is not a single entity but a collection of overlapping phenotypes resulting from the combined action of multiple risk alleles, it appears logical that an approach that parses the clinical presentation into biologically relevant components might be more powerful than one that relies on categorical diagnoses. Of course, any time that one engages in multiple comparisons, the risk for false-positive results is increased. Consequently, one must be cautious about interpreting an initially unremarkable linkage result that improves after subsequent analyses. The multiple comparisons should be taken into account in establishing an appropriate statistical threshold, or some other form of validation should be used, such as identification and confirmation of the functional disease variant.

Among the most exciting recent linkage findings are those identifying a putative autism locus at 17q. A study of 345 multiplex families from the Autism Genetic Resource Exchange (AGRE) database (Geschwind et al 2001) yielded its highest LOD score, 2.83, for the region 17q11.2 (Yonan et al 2003). Similarly, a study of 158 multiplex families produced its highest LOD score, 2.85, for 17q11.2 (McCauley et al 2005). Given the male predominance in ASDs, estimated to be approximately 4:1 for narrowly defined autism and even higher for broad-spectrum phenotypes, it has been hypothesized that focusing on families with male-only transmission might identify a distinctive genetic subgroup. Two hundred fifty-seven AGRE families from the initial sample (AGRE_1) were subdivided into male-only (MO) versus female-containing sibships. Genomewide linkage analysis yielded a LOD score of 4.3 in the MO families (Stone et al 2004). The same research group then undertook a genomewide scan in an independent sample of 109 sibling pairs from 91 additional AGRE families (AGRE_2; Cantor et al 2005). Initially, this scan yielded no evidence for significant linkage or for replication of the AGRE_1 analysis. However, when 56 MO sibling pairs were evaluated, the resulting LOD score was 4.1 at 17q21. Investigation of allele sharing in this region showed statistical significance at the same marker in both the AGRE_1 and AGRE_2 MO subgroups, reaching a widely accepted threshold for replication (Lander and Kruglyak 1995). Additional genotyping of the combined AGRE_1 and AGRE_2 MO samples revealed a maximum LOD score somewhat telomeric to both of the peaks from the individual scans. Nonetheless, the data are among the strongest linkage evidence to date. As discussed in Candidate Gene Studies, efforts are now focusing on both mutation screening and candidate gene association studies within this interval.

Cytogenetic Analyses

It has been appreciated for some time that children with autism carry chromosomal abnormalities at a greater frequency than the typically developing population. These findings have been catalogued extensively (Bugge et al 2000; Veenstra-Vanderweele et al 2004; Vorstman et al 2006; Wassink et al 2001). There is general agreement that there is a wide distribution of abnormalities, that no single rearrangement is likely to account for a major fraction of individuals with ASDs, and that some clustering of abnormalities occurs.

The most frequent site of chromosomal abnormalities found in ASD patients is 15q11–13. This region is of major interest as well because of its relationship with other developmental and behavioral syndromes. Deletion of the maternally inherited copy of this interval, as well as mutations in the gene *UBE3A* mapping therein, lead to Angelman syndrome, a disorder that has some overlap with severe autism phenotypes. Deletion of the paternally inherited copy leads to Prader-Willi syndrome, a disorder with prominent behavioral features including perseveration, obsessive-compulsive phenomena, and impulsive behavior. With respect to ASDs, duplications of this chromosomal segment, predominantly though not exclusively involving the maternal copy, have been reported numerous times (reviewed in Veenstra-Vanderweele et al 2004 and Vorstman et al 2006). Several candidate genes map to this interval, such as *GABRB3* (GABA A receptor beta-3) and *UBE3A* (ubiquitin protein ligase E3A), but mutations have not been identified in cytogenetically normal patients with ASDs, and no association of common alleles has been conclusively demonstrated.

Clusters of chromosomal abnormalities have been reported in other areas of the genome as well, including but not limited to the long arms of chromosomes 2 and 7 and the short arm of chromosome 5. A recent accounting by Vorstman and colleagues (2006) presents a very useful summary of candidate regions suggested by this approach.

In addition, multiple genes have been identified that are disrupted physically by chromosomal rearrangements, including *RAY1* (also known as *ST7,* suppression of tumorigenicity 7), *AUTS2* (autism susceptibility gene 2), *PAX3* (paired box gene 3), *MMP16* (matrix metalloproteinase 16), *NBEA* (neurobeachin), *GRPR* (gastrin-releasing peptide receptor), *A2BP1* (ataxin 2-binding protein 1; reviewed in Veenstra-Vanderweele et al 2004), and *GABRG1* (GABA-A receptor gamma-1; Vincent et al 2006). For these genes, neither association of common alleles nor rare functional mutations yet have been reported.

Given the findings to date, chromosomal abnormalities are not expected to lead to the identification of a common autism risk allele. However, the value of cytogenetic approaches has been brought into sharp focus by recent findings with regard to the neuroligins, a gene family encoding proteins involved in synaptogenesis.

NLGN4X, a member of this family located on the X chromosome, was identified as having a potential role in ASDs because it resides at Xp22.3, a region found to be deleted in three girls with autistic features (Thomas et al 1999). Jamain and colleagues (2003) mutation-screened this gene in 158 probands and found a patently deleterious frameshift mutation in two affected brothers, one with autism and one with Asperger syndrome, as well as in

their unaffected mother. Given that the gene is present on the X chromosome, it was argued that the sons would lose essentially all expression of this transcript, whereas the unaffected mother would retain protein function as a result of the presence of one normal copy of the gene. The investigators also evaluated *NLGN3* at Xq13. A single-nucleotide substitution at a highly conserved amino acid was found in two affected brothers, one with autism and one with Asperger syndrome, as well as their unaffected mother. The findings, particularly with respect to the frameshift found in *NLGN4X*, represented the first demonstration of a clearly functional point mutation associated with ASD in the absence of other features of a genetic syndrome.

Notably, soon after the publication of these findings, a second group reported on a parametric linkage analysis of a large pedigree with 13 affected males, 6 of whom were available for study: 4 with mental retardation (MR), 1 with autism and MR, and 1 with PDD (Laumonnier et al 2004). After finding support for linkage at Xp22.3, the investigators sequenced *NLGN4X* and discovered a frameshift mutation that co-segregated in the family. This independent replication provided strong evidence that rare mutations in this gene contribute to developmental delay and ASD.

Subsequent studies suggest that the frequency of *NLGN* mutations involved in ASDs is low. No *NLGN4X* coding mutations were identified in a total of 416 patients in three studies (Blasi et al 2006; Gauthier et al 2005; Vincent et al 2004). Another group found four missense mutations in *NLGN4X* among 148 patients, but three did not clearly segregate with ASD when family members were evaluated (Yan et al 2005). No *NLGN3* coding mutations have been identified in a total of 512 patients in four studies (Blasi et al 2006; Gauthier et al 2005; Vincent et al 2004; Yan et al 2005).

These findings highlight several important questions confronting autism researchers. The first is whether the identification of rare variants in a small number of families has significant value for the field. As noted in the prior paragraph, the contribution of *NLGNs* to the overall population burden of individuals with ASDs appears to be quite small. Nonetheless, the discovery may still provide a critical starting point for intensive investigations of the molecular and cellular mechanisms underlying ASD biology. For instance, researchers already have begun to study the consequences of the *NLGN* mutations identified in affected individuals. Expression of truncated NLGN4X in cultured nonneuronal cells leads to retention of the mutant protein in the endoplasmic reticulum rather than transport to the plasma membrane. Expression of mutant NLGN3 also was found to lead to intracellular retention, though to a lesser degree (Chih et al 2004). In vitro experiments have confirmed that NLGNs are involved in synaptogenesis. The expression of *NLGN1* and *2* in cultured nonneuronal cells triggers presynaptic differentiation, indicating that NLGNs are sufficient to induce presynaptic development (Scheiffele et al 2000). NLGNs also promote postsynaptic differentiation in cultured hippocampal neurons; inhibition of *NLGN1, -2,* and *-3* expression decreases excitatory and inhibitory synapses, especially inhibitory synaptic function (Chih et al 2005). It is important to note that presynaptic development associated with expression of normal *NLGN3* and *4* in cultured hippocampal neurons was affected severely by the ASD-related mutations (Chih et al 2004). Similar consequences of mutant *NLGN3* also were reported by two other groups (Chubykin et al 2005; Comoletti et al 2004).

Although the relationship between these observations and the specific etiology of ASDs remains to be conclusively demonstrated, findings with regard to the genetics of other complex disorders suggest that these results may turn out to be quite important. There are multiple examples in which extremely rare genetic mutations have provided the first clues to critical disease pathways. Such discoveries in Parkinson's disease, Alzheimer's disease, breast cancer, and hypertension, to name a few, have represented the first tangible insights into the pathophysiology of these conditions, in some cases leading to novel therapeutic insights. Intensive investigation of *NLGN* mutations, even if relevant for only a handful of cases of ASD, may offer similar hope.

A second important question highlighted by the foregoing discussion is that of the overlap between genes that contribute to MR and autism. Although there is tremendous interest in identifying genes that carry risk solely for social disability, to date the majority of the data regarding ASDs involves genetic mutations that may lead either to cognitive or social disability, or both. In the case of *NLGN*4X, it appears that a single change in a single transcript may result in MR, ASDs, or a combination of the two. This is consistent with the data emerging from studies of genetic syndromes. For instance, with respect to fragile X syndrome, it is now well accepted that a substantial number of patients present with symptoms indistinguishable from autism, whereas others demonstrate nonautistic cognitive delay, in addition to the other defining features of the syndrome (Hagerman et al 2005).

This type of overlap between MR and ASD, especially in the context of known genetic syndromes, has led some researchers to discount the relevance of these findings as nonspecific and, thus, as not truly relevant for the study of autism. In fact, investigations of the phenomena of phenotypic overlap between MR and ASDs may turn out to be fundamentally important. For instance, the evidence to date suggests that genes influencing certain fundamental developmental processes in the central nervous system, such as synapse formation, synaptic plasticity, and axon pathfinding, could play a role in multiple forms of developmental neuropsychiatric disorders. Determining what types of cellular, molecular, epigenetic, or environmental mechanisms might dictate one outcome versus another could be tremendously valuable in illuminating the pathophysiology of social disability.

Candidate Gene Studies

Studies of candidate genes have largely been divided into two types: those that seek to determine whether common variants confer small increments of risk for disease and those that seek to identify rare functional mutations in a gene of major effect. Although the former approach largely has been favored, investigators are increasingly likely to conduct both types of analyses, particularly when studying candidate genes implicated by their location within a linkage interval or by their correspondence to a chromosomal abnormality.

Although there has been steady proliferation of sophisticated statistical approaches to identifying genetic association, at heart such studies are variations on the standard case–control design that is common to medical research. Genetic variation is the identified risk factor, and one seeks to determine whether there is a statistically significant relationship between exposure to the variants and increased (or decreased) population risk for the phenotype.

Numerous genes have been evaluated for association with ASDs, with multiple positive results. However, replication has been the rare exception rather than the rule. As the use of common variant association strategies has grown in popularity

across all areas of medicine, the propensity for false-positive or nonreplicable findings has been documented clearly (Hirschhorn et al 2002). The reasons for this are the subject of debate, and an in-depth discussion is beyond the scope of this review. Briefly, however, two issues are known to be particularly problematic. The first is population stratification. Because genetic variation is highly dependent on ethnic origin, matching of cases and controls on this variable is essential to avoid spurious results. This may be quite difficult in practice because differences that are sufficient to lead to erroneous results often are not overtly observable. Several useful approaches have been developed to guard against this type of error. In childhood disorders, one may take advantage of family-based tests of association by using virtual controls that are matched perfectly for ethnicity. When family samples are not available, one may use what is referred to as a genomic control, in which polymorphic markers that are not hypothesized to be related to the disorder are genotyped and compared in cases and controls to assess and correct for stratification.

A second issue involves the question of multiple comparisons. On the one hand, as the number of alleles and phenotypic categorizations multiplies within a given study, a statistical threshold of .05 quickly becomes inappropriate to guard against type 1 error. On the other hand, when one investigates, as is now possible, hundreds of thousands of single-nucleotide polymorphisms (SNPs) simultaneously, one risks missing important hypothesis-generating findings by holding to an overly strict correction for multiple comparisons. A number of useful strategies are emerging to address this issue. However, at this point, it is clear that replication in an independent sample, both within an individual laboratory and, more important, across laboratories, is the standard to which such studies must be held to ultimately confirm the relationship between a common variant and a disorder of interest.

In the case of autism, it also is highly likely that both phenotypic and genetic heterogeneity contribute to the difficulty in detecting and replicating significant association. As noted above, it may well be that several overlapping sets of disease genes may be causative among different populations. It also is possible that subtle, as of yet unknown phenotypic differences may result from the interplay of slightly different sets of genes. Under these circumstances, it is easy to imagine that one group of investigators may by chance identify a subset of patients carrying a particular risk which may not be present in the independent sample subsequently tested by another group seeking to replicate the findings. These factors may turn out to explain the variably replicating associations that are highlighted by recent reports for the *SLC25A12* (mitochondrial aspartate/glutamate carrier; Rabionet et al 2006) and *GABRB3* (GABA-A receptor beta-3; McCauley et al 2004) genes. The recruitment of very large samples of patients may be the only means to reconcile conflicting findings.

Among the many interesting and biologically plausible candidate genes presented in the literature, the evidence for the serotonin transporter and *engrailed 2* are illustrative of complementary gene discovery approaches.

Serotonin Transporter

A candidate gene with a particularly long and venerable history is the serotonin transporter, *SLC6A4*. Initially, interest in this locus was driven by the finding that nearly a third of patients with ASD have platelet hyperserotonemia. After more than a dozen studies of common polymorphisms in and around the transcript, results remain difficult to interpret, with some groups confirming and others rejecting significant association (reviewed in Bacchelli and Maestrini 2006).

However, interest in this gene has not waned, in part because of the recent results of positional cloning efforts. The largest genomewide nonparametric study to date (Yonan et al 2003) showed its maximum LOD score at a marker close to *SLC6A4*. A subsequent analysis focusing on this interval in 341 families resulted in a LOD score of 5.8 that increased to 8.0 for the 202 male-only families (Sutcliffe et al 2005). The investigators evaluated several common variants in *SLC6A4*, found only nominal evidence for association at two SNPs, and concluded that these could not account for the linkage peak observed. They then elected to mutation screen the gene and regulatory regions in 120 families with the highest family-specific LOD scores, with the result that four rare sequence variants at highly conserved amino acids were identified. In each case, segregation data were inconclusive but generally supported a relationship between the allele and affected status. Moreover, the coding variants were associated with increased severity of rigid-compulsive behaviors. These data raised the prospect that multiple rare variants of *SLC6A4* contribute to ASDs and demonstrate the potential value of combining intensive mutation screening with common allele-association strategies when investigating intervals identified in linkage studies.

Engrailed 2

Evidence of a common genetic variant contributing to ASDs has been found in recent investigations of the gene *EN2* (*engrailed 2*) at 7q36.3. *EN2* is a homeodomain gene involved in development of the cerebellum. Knockout mice lacking the EN2 protein develop an abnormal cerebellar foliation pattern (Millen et al 1994, 1995) and a decrease in the number of cells in cerebellar circuits (Kuemerle et al 1997). Transgenic mice that overexpress the protein ectopically in Purkinje cells develop a hypoplastic cerebellum with reduced Purkinje cell numbers (Baader et al 1998) and disruption of banding patterns (Baader et al 1999). These defects are reminiscent of results from histopathologic studies of ASD in human beings. Cerebellar abnormalities, such as a decrease in the number of Purkinje cells, are among the most consistent findings from such studies (Palmen et al 2004). In addition, the *EN2* chromosomal locus is one of the several regions on 7q that have arisen in linkage studies (reviewed in Bartlett et al 2005). Therefore, *EN2* has emerged as both a functional and positional candidate gene.

Although results of initial association studies were inconclusive (Petit et al 1995; Zhong et al 2003), a more recent study investigated four *EN2* SNPs and found significant association with two intronic markers, rs1861972 and rs1861973, among 167 AGRE families (Gharani et al 2004). A subsequent study by the same investigators analyzed the identical SNPs and 14 additional SNPs spanning the entire gene. Significant association with the two intronic SNPs was detected in 222 different AGRE families and 129 families from the National Institute of Mental Health (NIMH) repository (Benayed et al 2005). In the AGRE data set, the p value for rs1861972 was .0834 under a narrow diagnosis but was a significant .0296 under a broad diagnosis. The p values for rs1861973 were .0268 and .0121 under narrow and broad diagnoses, respectively. The p values also were significant for both SNPs under both diagnostic categories in the NIMH data set. Therefore, the association was replicated in multiple, independent study populations. The total set of 518 families (2336 individuals) is one of the largest association studies completed to

date in ASD. The *p* value for the haplotype (i.e., a group of closely spaced or contiguous alleles on a single chromosome) containing the two SNPs was .00000035, providing significant evidence that *EN2* is a susceptibility gene for ASD. The individual relative risk conferred by the haplotype was estimated to be 1.40, which is consistent with the notion that multiple alleles, each contributing relatively small increments of risk, may contribute to ASDs. The high frequency of the haplotype in the sample (approximately 67%) yielded a population-attributable risk of 40%; that is, sequence variations in *EN2* may influence as many as 40% of cases of ASD.

These results for the *EN2* SNPs are promising. However, they await replication by an independent research group as well as the identification of the functional alleles responsible for the observed association. If and when these data are forthcoming, the findings will be of enormous significance in the quest to uncover common genetic contributions to ASD.

Future Directions

The foregoing discussion has highlighted both the obstacles confronting autism genetics researchers as well as the dramatic progress of late. It appears clear that the field is poised to deliver on the promise of identifying multiple autism alleles. In addition to the types of approaches already noted, several developments have made a notable contribution to the recent achievements, and they promise to drive progress in gene discovery.

The Availability of Biomaterials

A critically important advance over the last decade has been the increased availability of DNA and cell lines from well-characterized patients. One only needs to look at the number of articles that acknowledge the Autism Genetic Resource Exchange to get a sense of the impact of the widespread availability of high-quality phenotypic data and biological samples. Similar resources available through the NIMH, as well as its strong support for sample sharing, have helped create an environment in which researchers new to the field are able to readily test their hypotheses. In addition, groups long dedicated to autism genetics research have been able to substantially increase their subject numbers and to use these publicly available samples for replication sets, both critical aspects of successful gene-discovery efforts. In concert with advancing genomic technologies, there is little doubt that the increasingly large numbers of well-characterized biological samples from individuals with autism and their families will be the key to future progress.

Advancing Genomic Technologies

The development of high-throughput, low-cost genomic technologies in the areas of genotyping and sequencing are having a major impact on the field. For example, a mainstay of both linkage and association studies is the evaluation of known polymorphic DNA markers. Recent advances in microarray technologies (in which many thousands of spots of DNA can be arrayed on a single microscope slide) allow researchers to query hundreds of thousands of markers in a single reaction at comparatively low cost. This capacity allows small laboratories to conduct genomewide linkage analyses quickly and has opened the door to a new type of analysis known as whole-genome association.

As noted above, the common case–control association study involves the specification of a hypothesis about genes that are believed to play a role in autism. To be successful, such studies must choose a marker that is quite close to the genetic change that actually leads to the phenotype. With the recent development of microarray-based SNP genotyping platforms that carry 300,000 to 500,000 markers evenly spaced throughout the genome, there now is likely to be sufficient coverage to look for association without having to choose candidate genes in advance, an approach known as genomewide association. The power of this approach was demonstrated in the recent identification of the role of complement factor H in age-related macular degeneration (Klein et al 2005). There is tremendous excitement in the field about leveraging such methods to perform positional cloning in ASD.

Advances in Cytogenetic Technologies

Microarray technology also is transforming the identification of chromosomal deletions or duplications. There are several new techniques available to accomplish this; one in widespread use is array-based comparative genomic hybridization. This method uses patient DNA and control DNA, each labeled with a fluorescent tag (typically either red or green). Equal amounts of genetic material from patient and control are hybridized to known regions of the human genome that are pre-arrayed on a slide. If patient and control have equal copy numbers at a given locus, the colors are represented in equal measure (and the spot appears yellow). If the patient has lost (deleted) a locus, only the control color is visualized. Conversely, if the patient has an extra copy at a locus (duplication), the patient color predominates.

With currently available microarrays, copy number changes from several hundred to several thousand nucleotides can be identified, which is 100–1000 times more sensitive than conventional cytogenetics, and the resolution is increasing at a rapid pace. As these technologies have developed, it has become clear that there is considerably more structural variation to the normal human genome than previously has been suspected (Eichler 2006; Iafrate et al 2004; Sebat et al 2004). Consequently, it no longer is possible to draw conclusions that, for example, a small deleted chromosomal region in a patient with ASD is related to disease, even if the interval contains interesting, brain-expressed candidate genes. Indeed, there are many instances of copy number losses at important brain-expressed genes that have been found commonly in apparently normal individuals (Conrad et al 2006; Hinds et al 2006; McCarroll et al 2006).

Despite having to address these unexpected complexities, the ability to identify submicroscopic chromosomal changes holds tremendous promise. It is likely that in certain instances, formerly cryptic deletions or duplications will point to a rare mutation, perhaps in a gene of major effect that is only relevant for a fraction of patients, such as is the case with *NLGN*. It also appears increasingly likely, however, that common copy-number variations may contribute to disease risk in a fashion analogous to other types of common genomic variation. It will be both challenging and exciting to sort out these new possibilities.

Conclusions

After more than a decade of halting progress, research into the genetics of autism now is moving forward at a remarkable pace. Over just the past couple of years, *EN2* has emerged as a strong candidate for association with ASDs, a linkage region on chromosome 17q has been replicated in independent samples with rigorous statistical criteria, and the findings of rare mutations in the *NLGN*s are providing novel insights into the potential molecular and cellular mechanisms underlying social disability. Of course, the identification of common risk alleles or rare causative mutations is just one important step in unraveling the biology of

ASDs, an effort that will require the combined contributions of geneticists, clinical researchers, developmental neurobiologists, and neuroimagers. Although the ultimate goal of understanding the pathophysiology of these disorders in a way that allows for the development of effective treatments and prevention strategies is still over the horizon, the field now clearly has taken the first steps in that direction.

This work was supported by National Institutes of Health grants T32 MH18268 (to ARG) and K23 RR16118 (to MWS).

Alarcon M, Cantor RM, Liu J, Gilliam TC, Geschwind DH (2002): Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. *Am J Hum Genet* 70:60–71.

Alarcon M, Yonan AL, Gilliam TC, Cantor RM, Geschwind DH (2005): Quantitative genome scan and Ordered-Subsets Analysis of autism endophenotypes support language QTLs. *Mol Psychiatry* 10:747–757.

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

Baader SL, Sanlioglu S, Berrebi AS, Parker-Thornburg J, Oberdick J (1998): Ectopic overexpression of engrailed-2 in cerebellar Purkinje cells causes restricted cell loss and retarded external germinal layer development at lobule junctions. *J Neurosci* 18:1763–1773.

Baader SL, Vogel MW, Sanlioglu S, Zhang X, Oberdick J (1999): Selective disruption of "late onset" sagittal banding patterns by ectopic expression of engrailed-2 in cerebellar Purkinje cells. *J Neurosci* 19:5370–5379.

Bacchelli E, Maestrini E (2006): Autism spectrum disorders: Molecular genetic advances. *Am J Med Genet* 142C:13–23.

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

Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M (1995): Autism as a strongly genetic disorder: Evidence from a British twin study. *Psychol Med* 25:63–77.

Barrett S, Beck JC, Bernier R, Bisson E, Braun TA, Casavant TL, et al (1999): An autosomal genomic screen for autism. Collaborative linkage study of autism. *Am J Med Genet* 88:609–615.

Bartlett CW, Gharani N, Millonig JH, Brzustowicz LM (2005): Three autism candidate genes: A synthesis of human genetic analysis with other disciplines. *Int J Dev Neurosci* 23:221–234.

Benayed R, Gharani N, Rossman I, Mancuso V, Lazar G, Kamdar S (2005): Support for the homeobox transcription factor gene *ENGRAILED 2* as an autism spectrum disorder susceptibility locus. *Am J Hum Genet* 77:851–868.

Blasi F, Bacchelli E, Pesaresi G, Carone S, Bailey AJ, Maestrini E, International Molecular Genetic Study of Autism Consortium (2006): Absence of coding mutations in the X-linked genes neuroligin 3 and neuroligin 4 in individuals with autism from the IMGSAC collection. *Am J Med Genet B Neuropsychiatr Genet* 141:220–221.

Bugge M, Bruun-Petersen G, Brondum-Nielsen K, Friedrich U, Hansen J, Jensen G, et al (2000): Disease associated balanced chromosome rearrangements: A resource for large scale genotype-phenotype delineation in man. *J Med Genet* 37:858–865.

Buxbaum JD, Silverman J, Keddache M, Smith CJ, Hollander E, Ramoz N, Reichert JG (2004): Linkage analysis for autism in a subset families with obsessive-compulsive behaviors: Evidence for an autism susceptibility gene on chromosome 1 and further support for susceptibility genes on chromosome 6 and 19. *Mol Psychiatry* 9:144–150.

Buxbaum JD, Silverman JM, Smith CJ, Kilifarski M, Reichert J, Hollander E, et al (2001): Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. *Am J Hum Genet* 68:1514–1520.

Cantor RM, Kono N, Duvall JA, Alvarez-Retuerto A, Stone JL, Alarcon M, et al (2005): Replication of autism linkage: Fine-mapping peak at 17q21. *Am J Hum Genet* 76:1050–1056.

Chih B, Afridi SK, Clark L, Scheiffele P (2004): Disorder-associated mutations lead to functional inactivation of neuroligins. *Hum Mol Genet* 13:1471–1477.

Chih B, Engelman H, Scheiffele P (2005): Control of excitatory and inhibitory synapse formation by neuroligins. *Science* 307:1324–1328.

Chubykin AA, Liu X, Comoletti D, Tsigelny I, Taylor P, Sudhof TC (2005): Dissection of synapse induction by neuroligins: Effect of a neuroligin mutation associated with autism. *J Biol Chem* 280:22365–22374.

Comoletti D, De Jaco A, Jennings LL, Flynn RE, Gaietta G, Tsigelny I, et al (2004): The Arg451Cys-neuroligin-3 mutation associated with autism reveals a defect in protein processing. *J Neurosci* 24:4889–4893.

Conrad DF, Andrews TD, Carter NP, Hurles ME, Pritchard JK (2006): A high-resolution survey of deletion polymorphisms in the human genome. *Nat Genet* 38:75–81.

Eichler EE (2006): Widening the spectrum of human genetic variation. *Nat Genet* 38:9–11.

Filipek P (2005): Medical aspects of autism. In: F Volkmar, R Paul, A Klin, D Cohen, eds. *Handbook of Autism and Pervasive Developmental Disorders, vol 1. 3rd ed.* Hoboken, NJ: Wiley, 534–578.

Fombonne E (2005): Epidemiological studies of pervasive developmental disorders. In: F Volkmar, R Paul, A Klin, D Cohen, eds. *Handbook of Autism and Pervasive Developmental Disorders, vol 1. 3rd ed.* Hoboken, NJ: Wiley, 42–69.

Gauthier J, Bonnel A, St-Onge J, Karemera L, Laurent S, Mottron L, et al (2005): NLGN3/NLGN4 gene mutations are not responsible for autism in the Quebec population. *Am J Med Genet B Neuropsychiatr Genet* 132:74–75.

Geschwind DH, Sowinski J, Lord C, Iversen P, Shestack J, Jones P, et al (2001): The autism genetic resource exchange: A resource for the study of autism and related neuropsychiatric conditions. *Am J Hum Genet* 69:463–466.

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

Hagerman RJ, Ono MY, Hagerman PJ (2005): Recent advances in fragile X: A model for autism and neurodegeneration. *Curr Opin Psychiatry* 18:490–496.

Hinds DA, Kloek AP, Jen M, Chen X, Frazer KA (2006): Common deletions and SNPs are in linkage disequilibrium in the human genome. *Nat Genet* 38:82–85.

Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K (2002): A comprehensive review of genetic association studies. *Genet Med* 4:45–61.

Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, et al (2004): Detection of large-scale variation in the human genome. *Nat Genet* 36:949–951.

International Molecular Genetic Study of Autism Consortium (1998): A full genome screen for autism with evidence for linkage to a region on chromosome 7q. *Hum Mol Genet* 7:571–578.

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

Jamain S, Quach H, Betancur C, Rastam M, Colineaux C, Gillberg IC, et al (2003): Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet* 34:27–29.

Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al (2005): Complement factor H polymorphism in age-related macular degeneration. *Science* 308:385–389.

Kuemerle B, Zanjani H, Joyner A, Herrup K (1997): Pattern deformities and cell loss in Engrailed-2 mutant mice suggest two separate patterning events during cerebellar development. *J Neurosci* 17:7881–7889.

Lander E, Kruglyak L (1995): Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nat Genet* 11:241–247.

Laumonnier F, Bonnet-Brilhault F, Gomot M, Blanc R, David A, Moizard MP, et al (2004): X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am J Hum Genet* 74:552–557.

Liu J, Nyholt DR, Magnussen P, Parano E, Pavone P, Geschwind D, et al (2001): A genomewide screen for autism susceptibility loci. *Am J Hum Genet* 69:327–340.

McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, et al (2006): Common deletion polymorphisms in the human genome. *Nat Genet* 38:86–92.

McCauley JL, Li C, Jiang L, Olson LM, Crockett G, Gainer K, et al (2005): Genome-wide and Ordered-Subset linkage analyses provide support for autism loci on 17q and 19p with evidence of phenotypic and interlocus genetic correlates. *BMC Med Genet* 6:1.

McCauley JL, Olson LM, Delahanty R, Amin T, Nurmi EL, Organ EL, et al (2004): A linkage disequilibrium map of the 1-Mb 15q12 GABA-A receptor subunit cluster and association to autism. *Am J Med Genet* 131:51–59.

Millen KJ, Hui C-C, Joyner AL (1995): A role for *En-2* and other murine homologues of *Drosophila* segment polarity genes in regulating positional information in the developing cerebellum. *Development* 121:3935–3945.

Millen KJ, Wurst W, Herrup K, Joyner AL (1994): Abnormal embryonic cerebellar development and patterning of postnatal foliation in two mouse Engrailed-*2* mutants. *Development* 120:695–706.

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

Palmen SJ, van Engeland H, Hof PR, Schmitz C (2004): Neuropathological findings in autism. *Brain* 127:2572–2583.

Petit E, Herault J, Martineau J, Perrot A, Barthelemy C, Hameury L, et al (1995): Association study with two markers of a human homeogene in infantile autism. *J Med Genet* 32:269–274.

Philippe A, Martinez M, Guilloud-Bataille M, Gillberg C, Rastam M, Sponheim E, et al (1999): Genome-wide scan for autism susceptibility genes. Paris Autism Research International Sibpair Study. *Hum Mol Genet* 8:805–812.

Rabionet R, McCauley JL, Jaworski JM, Ashley-Koch AE, Martin ER, Sutcliffe JS, et al (2006): Lack of association between autism and SLC25A12. *Am J Psychiatry* 163:929–931.

Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, et al (1999): A genomic screen of autism: Evidence for a multilocus etiology. *Am J Hum Genet* 65:493–507.

Rutter M (2005): Genetic influences and autism. In: F Volkmar, R Paul, A Klin, D Cohen, eds. *Handbook of Autism and Pervasive Developmental Disorders, vol 1. 3rd ed.* Hoboken, NJ: Wiley, 425–452.

Scheiffele P, Fan J, Choih J, Fetter R, Serafini T (2000): Neuroligin expressed in nonneuronal cells triggers presynaptic development in contacting axons. *Cell* 101:657–669.

Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, et al (2004): Large-scale copy number polymorphism in the human genome. *Science* 305:525–528.

Shao Y, Raiford KL, Wolpert CM, Cope HA, Ravan SA, Ashley-Koch AA, et al (2002a): Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. *Am J Hum Genet* 70:1058–1061.

Shao Y, Wolpert CM, Raiford KL, Menold MM, Donnelly SL, Ravan SA, et al (2002b): Genomic screen and follow-up analysis for autistic disorder. *Am J Med Genet* 114:99–105.

Stone JL, Merriman B, Cantor RM, Yonan AL, Gilliam TC, Geschwind DH, Nelson SF (2004): Evidence for sex-specific risk alleles in autism spectrum disorder. *Am J Hum Genet* 75:1117–1123.

Sutcliffe JS, Delahanty RJ, Prasad HC, McCauley JL, Han Q, Jiang L, et al (2005): Allelic heterogeneity at the serotonin transporter locus (SLC6A4) confers susceptibility to autism and rigid-compulsive behaviors. *Am J Hum Genet* 77:265–279.

Thomas NS, Sharp AJ, Browne CE, Skuse D, Hardie C, Dennis NR (1999): Xp deletions associated with autism in three females. *Hum Genet* 104:43–48.

Veenstra-Vanderweele J, Christian SL, Cook EH Jr (2004): Autism as a paradigmatic complex genetic disorder. *Annu Rev Genomics Hum Genet* 5:379–405.

Vincent JB, Horike SI, Choufani S, Paterson AD, Roberts W, Szatmari P, et al (2006): An inversion inv(4)(p12-p15.3) in autistic siblings implicates the 4p GABA receptor gene cluster. *J Med Genet* 43:429–434.

Vincent JB, Kolozsvari D, Roberts WS, Bolton PF, Gurling HM, Scherer SW (2004): Mutation screening of X-chromosomal neuroligin genes: No mutations in 196 autism probands. *Am J Med Genet B Neuropsychiatr Genet* 129:82–84.

Vorstman JAS, Staal WG, van Daalen E, van Engeland H, Hochstenbach PFR, Franke L (2006): Identification of novel autism candidate regions through analysis of reported cytogenetic abnormalities associated with autism. *Mol Psychiatry* 11:18–28.

Wassink TH, Piven J, Patil SR (2001): Chromosomal abnormalities in a clinic sample of individuals with autistic disorder. *Psychiatr Genet* 11:57–63.

Yan J, Oliveira G, Coutinho A, Yang C, Feng J, Katz C, et al (2005): Analysis of the neuroligin 3 and 4 genes in autism and other neuropsychiatric patients. *Mol Psychiatry* 10:329–332.

Ylisaukko-oja T, Nieminen-von Wendt T, Kempas E, Sarenius S, Varilo T, von Wendt L, et al (2004): Genome-wide scan for loci of Asperger syndrome. *Mol Psychiatry* 9:161–168.

Yonan AL, Alarcon M, Cheng R, Magnusson PK, Spence SJ, Palmer AA, et al (2003): A genomewide screen of 345 families for autism-susceptibility loci. *Am J Hum Genet* 73:886–897.

Zhong H, Serajee FJ, Nabi R, Huq AH (2003): No association between the EN2 gene and autistic disorder. *J Med Genet* 40:e4.