AUTISM AS A PARADIGMATIC COMPLEX GENETIC DISORDER

Jeremy Veenstra-VanderWeele¹, Susan L. Christian¹, and Edwin H. Cook, Jr.¹,²
Departments of Psychiatry¹, Pediatrics, and Human Genetics², The University of Chicago, Illinois 60637; email:jveenstr@yoda.bsd.uchicago.edu, schrist@bsd.uchicago.edu, ed@yoda.bsd.uchicago.edu

Key Words autism, chromosomal abnormalities, epigenetic, linkage, association

Abstract Autism is one of the most heritable complex disorders, with compelling evidence for genetic factors and little or no support for environmental influence. The estimated prevalence of autism has increased since molecular genetic studies began, owing to loosening of diagnostic criteria and, more importantly, to more complete ascertainment strategies. This has led to a reduction in the sibling relative risk, but strong heritability estimates remain. It is essential to recognize that genetics is the only current approach to understanding the pathophysiology of autism in which there is not the usual concern about whether one is studying a consequence rather than a cause. There are hundreds, if not thousands, of patients with autism spectrum disorder with documented single-gene mutations or chromosomal abnormalities. Autism may be one of the most complex, yet strongly genetic, disorders in which chromosomal disorders, relatively rare highly penetrant mutations, and multiplicative effects of common variants all have support in different cases and families. The field of complex genetics is replete with many researchers and reviewers who want to promote their overly focused interest in one method at the exclusion of others. However, it is essential that the restricted interests of patients with autism not be reflected in overly restrictive genetic approaches if we are to better understand the genetics of autism in the most expeditious and thorough manner.

INTRODUCTION

Autistic disorder [OMIM 209850 (Online Mendelian Inheritance of Man, http://www.ncbi.nlm.nih.gov/entrez/dispmim.cgi?id=209850)] is a neurodevelopmental disorder characterized by three areas of abnormality: impairment in social interaction, impairment in communication, and restricted and repetitive patterns of interest or behavior (2). In the current diagnostic scheme, autistic disorder is one of a few diagnoses within the category of pervasive developmental disorders (PDDs). This category also includes PDD not otherwise specified (NOS), Rett syndrome,
and Asperger syndrome, which is characterized by relatively spared communication despite deficits in the other two areas. As is true for many disorders in medicine, autistic disorder is likely not a distinct, categorical disorder but instead represents one extreme of a spectrum of social and communication impairment and behavioral restriction. Recent research has examined autistic traits in a population of twins and found that social impairment follows a unimodal distribution without a clear demarcation to separate cases of the disorder (32). Future discussion of autism may more appropriately refer to autism spectrum disorder (ASD).

Evidence suggests that ASD has a large genetic component with complex inheritance. The most recent estimate places the prevalence at 0.1% to 0.2% (24) for narrow diagnosis of autistic disorder and 0.6% for ASD. The prevalence of autistic disorder is approximately four times higher in males than in females, and the gender differential is even higher in milder forms of ASD. There is concern about a possible increase in prevalence, but changes in diagnostic methodology and ascertainment strategy complicate comparisons across time (49). Twin studies show a 60% to 91% concordance rate in monozygotic twins, depending on whether a narrow or broad phenotype is considered, in contrast to no observations of concordance in dizygotic twins under narrow phenotypic definition and 10% concordance under broader phenotypic definition (9). Sibling recurrence rate is estimated to be 4.5% (67). The pattern of relative risk in the narrow phenotype of autistic disorder is consistent with multiplicative inheritance, with multiple variants at unlinked loci interacting to lead to the phenotype.

Animal models are emerging that may provide clues to gene and protein systems important in relevant behaviors, but they are unlikely to truly parallel the human disorder. Likewise, less complex genetic diseases with some shared symptoms, such as Rett syndrome or Fragile X syndrome, may result from disruption of gene or protein systems that may also be disrupted in the complex genetic syndrome of ASD. Direct approaches include three overlapping methodologies to identify genes or regions of interest in autism: chromosomal methods, such as karyotyping and fluorescent in situ hybridization (FISH); linkage studies, such as genome screens in affected sibling pairs; and gene association studies, including candidate gene studies.

Within the different approaches to ASD genetics, areas of study that are emphasized are either: (a) likely to produce insight into the underlying neurobiology of ASDs within the next several years, or (b) illustrative of the process whereby genetic variants are identified and the way genetic methodology is used to resolve conflicting findings. Readers should note the importance of heterogeneity within this syndrome that may eventually be parsed into subgroups and symptom clusters using approaches such as quantitative trait analysis. Heterogeneity may be important in thinking about divergent findings between studies using simplex families versus those using multiplex families because the latter are uncommon and may reflect a different grouping of gene variants without having a divergent clinical picture (40). We also highlight epigenetic inheritance, a sometimes neglected area of study that is crucial in understanding one subsyndrome within ASD, maternal
chromosome 15q11–q13 duplication syndrome, and may also explain diverse chromosomal abnormalities. Finally, because gene variants in complex genetic disorders are likely both subtle and diverse (105), we emphasize the importance of fully characterizing gene variation and using haplotype analysis to clarify and narrow regions of interest.

ANIMAL MODELS

Appropriate animal models are difficult to identify in ASD because the behavioral phenotype does not easily convert to lower mammals. Many transgenic mice demonstrate abnormalities in social behavior or behavioral inhibition in various experimental conditions, but it is difficult to equate these simple behaviors to the ASD phenotype. The most appropriate models to date have been knockout mouse models of Rett syndrome and Fragile X syndrome, both discussed in the corresponding subsections below. A few other transgenic mouse lines have particularly interesting social phenotypes. Mice lacking Dvl1, a gene important in determining cell polarity, show abnormal sensory gating and reduced social interaction (81). Mice lacking the oxytocin gene (Oxt) fail to change their behavior on re-exposure to a known cage mate, thus reflecting an apparent lack of social memory (46). Conversely, mice that have had their arginine vasopressin receptor 1A gene (avpr1a) promoter region replaced by prairie vole avpr1a promoter show increased affiliative behavior (141). Male avpr1a knockout mice had reduced anxiety-like behavior and a profound impairment in social recognition (14).

LOCUS OR ALLELIC HETEROGENEITY

A couple of definitions are crucial for understanding studies of complex genetics. Locus heterogeneity means that defects in different loci, or genes, may cause the same phenotype. The concept arose in relatively simple genetics, where, for example, defects in one of a few different genes can cause hereditary nonpolyposis colorectal cancer (HNPCC) (96). This idea becomes more complicated in complex genetic disorders, where different (likely overlapping) groups of genetic variants may cause susceptibility to disease. The large range of symptoms across individuals with ASD suggests that locus heterogeneity will probably be very common.

Allelic heterogeneity means that different defects in the same gene may lead to the same or different patterns of genetic disease. This concept also arises in simple genetic disorders, where different mutations in the same gene can lead to different phenotype patterns. For example, many different mutations in a single gene (CFTR) have been described in cystic fibrosis, some of which are associated with mild forms of the disorder or may lead to more pancreatic than lung findings (126). Similar patterns are emerging at the methyl-CpG binding protein 2 gene (MECP2) in Rett syndrome (144). As rare and common variants in relevant genes [e.g., the
serotonin transporter (\textit{SLC6A4}) are better understood, allelic heterogeneity may also become relevant in complex genetic disorders in psychiatry.

**Potential Endophenotypes and Quantitative Trait Analysis**

ASD also exhibits corresponding clinical heterogeneity. Recent advances in diagnosis, such as the Autism Diagnostic Interview (ADI-R) (84) and Autism Diagnostic Observation Schedule (ADOS) (83), may reduce uncertainty in diagnosis, but considerable variability remains. Symptoms and signs, rather than etiologies, comprise psychiatric syndromes such as autism, as well as other medical syndromes, including diabetes, hypertension, and lupus. Each gene variant may make a different contribution to the disorder, with gene variant A important in social cognition and gene variant B important in language acquisition. When clustering of risk alleles reaches a certain threshold, an individual is at increased risk of developing the disorder. A subthreshold number of risk alleles may result in the broader autism phenotype identified in family members of patients with autism (103, 104).

Endophenotypes are measurable components of a syndrome, including anatomical, biochemical, neuropsychological, and other measures (55). A measurable, quantitative trait offers many potential advantages over the diagnostic category. These traits might be measurable with more reliability and validity than is possible for the disorder. Inheritance of a trait may correspond to fewer genes because the trait may be inherited separately from several other traits that become a “disorder” only when they overlap. A trait may capture the true nature of a disorder as a spectrum and could even expand the number of subjects by tracing the trait through family members not affected with the full disorder. Finally, statistical methods are more powerful for quantitative than for dichotomous variables. One example of the endophenotype approach is in schizophrenia, where abnormalities in the P50 auditory-evoked potential response are associated with potentially functional promoter polymorphisms in the alpha7 neuronal nicotinic acetylcholine receptor subunit gene (\textit{CHRNA7}) (78).

Initial attempts in ASD have primarily considered the endophenotype of general language impairment. A number of other potential quantitative traits are of particular interest in autism: level of intellectual functioning, degree of social or communication impairment, presence of seizure disorder, dysmorphology, savant abilities, restrictive and repetitive behaviors, and—two that deserve particular attention—head circumference and whole blood serotonin.

An increased rate of macrocephaly in autistic disorder was mentioned in the initial description (68). More recent observers note that most children with autism are born with normal head circumference and show an increased rate of growth during early childhood (77); however, only about 20% of children with autistic disorder meet criteria for macrocephaly (50). The increased rate of growth in head circumference is most dramatic in the first year of life and corresponds to increased growth of the cerebral cortex as measured by MRI (38). A number of
questions follow logically from these observations, including whether there is a corresponding behavioral subgroup. Because head circumference is monitored in most countries during pediatrician visits during the first year of life, this trait could quickly be incorporated into genetic analyses in ASD.

Since the first description of hyperserotonemia in autistic disorder (115), numerous studies have identified elevated whole blood or platelet serotonin (5-HT) in about 25% of patients (34). Initial functional imaging studies in autistic disorder suggest that brain serotonin synthesis capacity may be increased in children with autism (28, 29). More research to characterize the serotonin system, particularly in the periphery, could yield important information about the abnormalities found in ASD. One example of this approach is an association mapping study of whole blood serotonin independent of psychiatric disorder, revealing evidence for association at the integrin beta 3 gene (ITGB3) on chromosome 17q (135).

LINKAGE FINDINGS

Genome-wide linkage studies can identify large chromosomal regions segregating genes within families with a given phenotype. Based on experience in adult psychiatric disease, investigators generally favor nonparametric approaches that make no assumptions about disease transmission and typically use sibling pairs to identify increased sharing of alleles among affected family members (8).

There have been several genome-wide scans with generally small sample sizes. The scattered and varied linkage findings both within and across studies suggest the involvement of 15 or more genes in the complex and heterogeneous syndrome of ASD (110). Much larger sample sizes than are currently available will need to be analyzed to produce adequate power to identify most of these genes.

A few regions of interest are emerging across the reported linkage studies. Significant linkage findings have been reported on chromosomes 2q (62) and 3q (6). Several other regions of interest with suggestive linkage evidence have emerged in more than one study, including regions on chromosomes 7q, 13q, 16p, and 17q. Several other regions of interest have been identified in one or more linkage samples but haven’t emerged consistently, including regions on chromosomes 1p, 1q, 5p, 6q, 8q, 15q, 19p, and Xq (6, 21, 30, 62, 82, 102, 110, 120). Some of the first-pass genome screen studies incorporated different definitions of affection, essentially drawing the line at a different point along the autism spectrum in the hope that a more homogeneous sample would emerge (62). Others considered parent-of-origin effects, an approach that may prove fruitful once it is applied to larger samples (82).

Now that first-pass genome screen data is available in multiple samples, investigators are trying to identify subgroups or quantitative traits that may refine linkage findings in particular regions. To date, the primary subgrouping considered has been phrase speech delay (PSD) past 36 months, which occurs in about 50% of autistic disorder sibling pairs. Linkage to chromosome 2q increased in two
samples when restricted to sibling pairs with PSD (21, 119). Linkage to chromosome 7q also increased in one PSD-restricted sample (18). A novel extension of this linkage analysis considered parents as affected or unaffected based on their own histories of speech and language difficulties (18). Furthermore, the region on chromosome 7q was also implicated when age at first word and repetitive behavior were separately mapped as quantitative traits within autism sibling pairs (1). In one sample, a suggestive linkage signal on chromosome 13q reflected only the subset of the sample with PSD (18). Regional linkage analyses in candidate gene regions on 15q and 17q also incorporated subphenotypes, and these studies are integrated into the discussion of candidate gene regions below.

Association and Candidate Gene Studies

With a few exceptions, tests of genetic association have been used to study candidate genes in disease. Association methods are much more powerful than linkage methods at a given locus, allowing genes of weaker effect to be detected, but the polymorphisms tested must be much closer to the susceptibility variant. These tests compare the observed alleles in individuals with a disease to those expected by chance. Family-based association testing is now the gold standard in avoiding population stratification bias in a disorder such as ASD in which parents are usually available for genotyping. Recently, methods that properly control for potential population differences (Genome Control and STRAT) were developed for case-control studies (7, 106). Genome-wide association studies will soon offer the potential to detect genes of small effect without requiring hypotheses about pathophysiology or chromosomal location. Coupling this approach with layered analysis of multiple samples may allow investigators to efficiently avoid both type I and type II error.

Family structure and ascertainment differs between association and linkage studies. Most association studies are conducted in samples of autism trios, with an affected proband and two parents. Linkage studies require families with multiple affected members that are uncommon in ASD because the sibling recurrence risk is only about 4.5%. These multiplex families may be biased toward families with less complex inheritance than seen in the overall disorder, despite a clinically indistinguishable phenotype (40). Some hypothesize that such enriched linkage samples are more likely to contain multiple rare polymorphisms within a single gene of interest, complicating identification of more common genetic variation important in the overall population (105). In contrast, families ascertained in association studies may more closely reflect the typical disorder. Except where specifically noted, the association studies discussed in detail below were conducted in samples of autistic disorder trios.

Numerous candidate gene studies have been conducted in ASD on the basis of limited knowledge of neuropharmacology, developmental neuropathological abnormalities, or chromosomal anomalies. Because we have little knowledge of pathophysiology and pharmacology in ASD, a limited number of genes represent
primary candidates for the disorder and deserve more scrutiny. After studying primary candidate genes, some justification could be concocted for studying almost any gene that is expressed in the brain. Further, investigators and reviewers struggle to identify the appropriate level of statistical correction for association tests, with the correct approach lying somewhere between a Bonferroni correction for each gene tested in a given sample to correction for every polymorphism in the human genome (109).

Most association studies have considered segregation of disease with alleles at single polymorphisms; however, a haplotype approach may be more powerful. The presence and size of haplotype blocks is highly variable and must be directly assessed by marker-to-marker linkage disequilibrium. Only rarely have association studies in ASD even considered haplotypes or linkage disequilibrium across a gene of interest, but failing to do so will likely produce type II error or result in inconsistent results across studies (143). Multiple functional variants may also exist within a given gene or gene region, thereby producing more than one haplotype that shows positive association with disease. For a more detailed discussion of haplotypes and association studies, please consult Reference 130.

CHROMOSOME 15Q11–Q13 REGION The chromosome 15q11–q13 region has been the focus of multiple association studies in ASD based on the maternally inherited 15q11–q13 duplication syndrome. This syndrome could reflect an underlying genetic mechanism in this region even in patients who do not have the duplication, but this remains to be demonstrated. If the general ASD population is affected by variation in this region, there might be at least two mechanisms behind the variants involved. Because duplication is predicted to increase expression of the relevant gene(s), variation is expected to similarly affect gene expression. However, in parallel to the Angelman and Prader-Willi syndromes, a second type of variation may affect genetic imprinting or chromatin structure of the entire region. Researchers will likely need to study the methylation patterns that regulate genomic imprinting in patients with ASD to clarify the intersection of primary DNA sequence and epigenetic factors in causing disease in individuals without duplications.

Initial results in this region have been varied, with positive association at several different markers with only rare replication. An initial study identified significant association at 155CA-2, a polymorphism within the gamma-amino butyric acid (GABA) receptor A β3 subunit (GABRB3) gene (37). This finding was replicated in one study (20) but not in others. A microsatellite marker simply known as GABRB3,3′ to the GABRB3 gene, was also implicated in one association study (88). A study from the same research group also found significant evidence for maternal linkage to this polymorphism in subjects sharing high scores on an “insistence on sameness” factor derived from the Autism Diagnostic Interview (ADI-R) (118).

In addition to the cluster of findings around GABRB3, investigation has also focused on the UBE3A and ATP10C genes that lie within the maternal expression domain. One sibling pair study found association at D15S122, a microsatellite
marker located in the 5′ end of UBE3A (94). A follow-up study found association at multiple SNPs within ATP10C (93), with greater evidence within trios than within sibling pairs.

A number of linkage studies have also focused on this region. One study in a sibling pair sample revealed modest evidence for linkage at D15S511, another microsatellite marker 3′ to GABRB3 (95). This evidence increased into the suggestive range in a subgroup of sibling pairs with savant skills, a factor derived from the ADI-R that corresponds to a relative sparing of one area of cognitive function (125). Modest evidence for linkage without subject subgrouping was also detected in this region (102, 120).

SEROTONIN TRANSPORTER GENE (SLC6A4) The serotonin transporter gene (SLC6A4) is a primary candidate gene based on hyperserotonemia in autism, increased platelet serotonin uptake in hyperserotonemic first-degree relatives of probands with ASD, responsiveness of obsessive-compulsive disorder (OCD) related symptoms to potent serotonin transporter inhibitors (36), and linkage findings in the region. A variable number tandem repeat (VNTR) polymorphism in the SLC6A4 promoter 5-HTTLPR affects transcription and is associated with neuroticism or anxiety (79), as well as liability to depression in the face of adverse life events (22). Another VNTR in SLC6A4 intron 2 may also affect transcription (47, 85).

In autistic disorder, most but not all studies have nominally significant evidence of transmission disequilibrium for SLC6A4 polymorphisms. The initial study found association with 5-HTTLPR but stronger association with haplotypes including the short form of the 5-HTTLPR and the 12-repeat form of the intron 2 polymorphism (36). An independent sample by the same investigators replicated preferential transmission of the short 5-HTTLPR/12 repeat intron 2 VNTR haplotype, but not of the 5-HTTLPR as a single marker (71). Subsequent studies have focused primarily on the 5-HTTLPR and have been inconsistent. Some have replicated association with the short 5-HTTLPR allele (31, 90), whereas others have found association with the long form (73, 128, 139). One study found overall preferential transmission of the long 5-HTTLPR allele, and a subset with more severe social impairment showed preferential transmission of the short allele (128). Several family-based studies have found no significant association at this polymorphism (39, 86, 100). One study found a positive 5-HTTLPR/intron 2 VNTR haplotype association in the trio portion of the sample despite no significant findings when including sibling pairs (13).

Because these initial results were contradictory, inquiries have expanded to include other polymorphisms within or near SLC6A4. Kim and colleagues (71) found more significant association with a number of 5′ single-nucleotide polymorphisms (SNPs) than with 5-HTTLPR. Another trio study found most significant association with a haplotype including the short form of the 5-HTTLPR and two of the SNPs associated in the previous study (31). A sibling pair study failed to show any association in a haplotype analysis containing multiple SNPs 3′ to 5-HTTLPR (90).
Linkage studies are also beginning to show evidence in this region (62, 140). One study reports suggestive linkage in a region of 17q overlapping with SLC6A4 and found more evidence for linkage in a subgroup of families with elevated scores on a rigid/compulsive factor derived from the ADI-R (90).

Investigators have also sought to understand the relationships between these polymorphisms and whole blood serotonin levels. Again, results are inconsistent. Three studies found a nonsignificant trend for association between elevated whole blood serotonin and homozygosity for the long form of the 5-HTTLPR polymorphism in patients that would reflect only a very small effect size (4, 12, 101). Despite no association with autistic disorder, Coutinho and colleagues reported significant association between elevated whole blood serotonin and haplotypes of the 5-HTTLPR and intron 2 polymorphisms, as well as each marker individually (39). We have observed similar elevations with the same haplotypes (E.H. Cook, unpublished observation). A previous study of these two polymorphisms found nonsignificant trends for association between elevated whole blood serotonin and homozygosity for the alleles corresponding to the most significant haplotype, but analysis by haplotype was not performed (12).

Enough evidence has accumulated to suggest that this gene is likely involved in ASD. Studies of trios are more likely to detect association within the 5′ region of the gene, but sibling pair studies have identified linkage in this region as well. The linkage samples may contain some families with rare mutations that may not correspond to the association pattern observed in the overall population (105). Locus heterogeneity could also be at work with different variants having effects on different aspects of the core (e.g., rigid/compulsive behavior) or associated (e.g., aggression) phenotypes. Conversely, increased significance with haplotype analyses suggests that the responsible variant or variants have not yet been identified but may be in partial linkage disequilibrium with polymorphisms currently being studied. Polymorphisms within the coding region are rare (54), but a rare I425V missense variant investigated leads to constitutive activation of the serotonin transporter and cosegregates with susceptibility to social phobia, OCD, and Asperger syndrome (70, 98). However, the regulatory region of the gene remains incompletely characterized and may contain important polymorphisms, including variation within the 5-HTTLPR (91), as well as conserved noncoding sequence upstream of the 5-HTTLPR. This region merits more extensive study, including completion of resequencing of the gene and flanking region, haplotype analysis in a larger sample of ASD families, and further exploration of the relationship among multiple haplotypes and several phenotypes related to autism.

**RELN gene** The reelin gene (RELN) lies in the chromosome 7q linkage region and specific mutations in RELN have been identified that cause autosomal recessive lissencephaly, a disorder of failed neuronal migration (58). Abnormalities in expression of the corresponding protein, reelin, are being investigated in postmortem brain studies of autistic disorder (45). One study in autistic disorder reported preferential transmission of “long” alleles of a trinucleotide repeat polymorphism in
the 5′ region of RELN (99). A replication study found that overall transmission disequilibrium was not significant but that “long” alleles were preferentially transmitted (145). Three additional studies in large samples found no evidence for association with the RELN trinucleotide repeat (16, 43, 76).

NEUROLIGEN GENES NLGN3 AND NLGN4 The neuroligin genes NLGN3 and NLGN4, both located on the X chromosome, encode cell-adhesion proteins important in synapse formation (123). NLGN4 is located near the junction of the pseudoautosomal region on chromosome Xp, where deletions have been identified in a few patients with autism (127). It has a counterpart NLGN4Y on the Y chromosome that is expressed in a similar pattern within the male brain (65), but oligomerization appears crucial for neuroligin activity, so mutation in just one version of the protein could potentially interfere with binding among copies of the intact version of the protein (41). A frameshift mutation in NLGN4 was identified in two brothers with ASD as well as in their unaffected mother (65).

Based on the findings at NLGN4, a paralogous gene, NLGN3, was also sequenced, revealing a mutation that causes an amino acid change within a highly conserved region in two brothers with ASD (65). Unfortunately, no other family members were available for sequencing, so it is difficult to assess cosegregation between the variant and disease. No other mutations were identified in NLGN3 or NLGN4 within 154 patients with ASD or 200 controls. Association studies will be crucial to evaluate whether these genes are involved in a larger group of patients with ASD. Evaluating this protein system within human cells or animal models may clarify the role of these variants in disease. Even if neuroligin gene variants are important in only one family, studying this protein system may still increase our understanding of the pathophysiology of ASD.

OTHER CANDIDATE GENES Other candidate genes have been associated with ASD in single studies that await replication. One of the strongest findings is a recent association with SNPs in the mitochondrial aspartate/glutamate carrier SLC25A12 gene in the chromosome 2 linkage region (108). Mice lacking the glutamate receptor subunit GluR6 (GRIK2) gene show abnormalities in long-term potentiation within the hippocampus (33), and an abnormal response at a sister receptor was identified in mouse models of Fragile X syndrome (59). Jamain and colleagues (64) reported an association between autistic disorder and maternal transmission of alleles at three polymorphisms within GRIK2, which lies under a suggestive linkage peak on chromosome 6p21. The WNT2 gene on chromosome 7q encodes a member of the same signaling pathway as DVL1 (see “Animal Models,” above). Mutations leading to amino acid changes were detected in WNT2 in two families with autism, with one unaffected parent transmitting the mutation to two affected children. A modest association was also reported for a 3′ polymorphism in a sample of ASD sibling pairs (133). The arginine–vasopressin receptor 1A gene (AVPR1A) is important in regulating basic social functions in animal models, as discussed
above. Two studies of AVPR1A in ASD found nominal association for a different one of the two microsatellites in the 5’ flanking region in each study (72, 133a). HOXA1 belongs to a gene family that determines pattern formation within the central nervous system. Although one positive family-based association was reported at HOXA1 (60), attempts to replicate this finding in larger samples show no supporting evidence (42, 51, 80, 114).

RELATIVELY LESS COMPLEX GENETIC AND CHROMOSOMAL DISORDERS

Several relatively less complex genetic or chromosomal syndromes, including Rett, Fragile X, Prader-Willi, Angelman, Smith-Lemli-Opitz, and Turner syndromes, maternal chromosome 15q11–q13 duplications and triplications and tuberous sclerosis, frequently include prominent symptoms relevant to ASD. Because the etiology of many of these disorders is partially known, investigators hope to gain insight into the more complex genetics of ASD by studying the molecular and brain system abnormalities both in humans with these diseases and in corresponding animal models. More single-gene or chromosomal disorders will likely be identified from within the heterogeneous syndrome of ASD; however, simple genetic disorders will not likely account for the majority of affected children.

Rett Syndrome

Rett syndrome, one of the Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition (DSM-IV) PDDs, is an X-linked disorder with an incidence of approximately 1 in 10,000–15,000 girls, with 99.5% of cases being sporadic. Girls with classic Rett syndrome develop normally until approximately 6–18 months of age, when they lose speech and purposeful hand movements and subsequently progress to develop microcephaly, seizures, ataxia, stereotypic hand movements, abnormal breathing patterns, and autistic behavior. Mutations in MECP2 were identified as the cause of Rett syndrome (3). The MeCP2 protein normally binds methylated CpG dinucleotides, forming part of the Sin3A/HDAC complex that deacetylates histones to repress gene expression. Therefore, Rett syndrome must be a result of inappropriate gene overexpression at some time and location during development. Heterozygous female mecp2 knockout mice develop symptoms at about 9 months of age, after reaching maturity, which raises the possibility that the genetic defect could affect brain stability rather than development (25, 57), and is consistent with the finding that the MeCP2 protein is expressed after neurons reach a certain level of maturity (116). Male mice heterozygous for an mecp2 truncation mutation have symptoms developing at 6 weeks and eventually manifest many of the core features of Rett syndrome, even including stereotyped forelimb movements, less exploration of novel environments (open-field test), and social dysfunction (117).
Cytogenetic Studies of Autism

Eleven studies have been reported where standard cytogenetic analysis was performed on individuals with autism (Table 1; follow the Supplemental Material link from the Annual Reviews home page at http://www.annualreviews.org). Overall, 1826 subjects were karyotyped, of which 78 showed chromosomal abnormalities for a cumulative rate of 4.3%. Of these 78 cases, 24 had Fragile X syndrome, corresponding to a fragile site at Xq27. Separating the Fragile X cases left 54 (3.0%) with other abnormalities. Twelve cases with fragile sites other than Fragile X were excluded from these calculations (53, 132), as was one case with trisomy 13 (75). The types of abnormalities that would cause chromosomal imbalances include deletions, duplications, supernumerary marker chromosomes, unbalanced translocations, ring chromosomes, trisomies, monosomies, and isochromosomes, whereas inversions, fragile sites, and balanced translocations would disrupt single genes but not alter gene dosage.

Chromosomal Abnormalities and Autism

There have been many individual case reports of chromosomal abnormalities and autism. Table 2 (follow the Supplemental Material link from the Annual Reviews home page at http://www.annualreviews.org) provides a compilation of all reported cases, including individual case reports and abnormalities identified as part of the screening listed in Table 1. Overall, chromosomal abnormalities with autism as part or all of the phenotype have been reported for all 22 autosomes plus both sex chromosomes. Abnormalities that cause chromosomal imbalances include 19 out of 22 autosomes plus both sex chromosomes. These data demonstrate the heterogeneity of chromosomal abnormalities that have been observed in autism and ASDs.

Certain chromosomal abnormalities have been reported with sufficient frequency in autistic subjects to be subgrouped with a specific disorder or syndrome. One example is the Fragile X syndrome at Xq27. In a cytogenetic survey performed by Wassink and colleagues (132), 2.2% of the subjects karyotyped (6 out of 278) had the Fragile X syndrome. All individuals were ascertained through the University of Iowa Child and Adolescent Psychiatry Clinic and met DSM-III, DSM-IIIR, or DSM-IV criteria for autism. In another study, an examination of 25 multiplex families for autism and related disorders identified 3 out of 25 (12%) with Fragile X syndrome (56). In a prospective study, 57 boys with Fragile X syndrome were tested using the Childhood Autism Rating Scale (CARS). Of these, 14 (25%) met the criteria for autism, 12 out of 14 were in the mild to moderately autistic range, and 2 were severe (10). In another study, 24 children with Fragile X were examined using ADI-R, ADOS-G, and DSM-IV autistic diagnostic systems. Of these, 8 (33%) met the criteria for autism (113). With the identification of the FMR-1 gene that causes Fragile X syndrome, molecular testing is now available to directly diagnose these individuals.
DUPLICATIONS OF CHROMOSOME 15Q11-Q13 Duplications of chromosome 15q11–q13 with autistic features were first reported in 6 boys with a large supernumerary marker chromosome later found to be derived from chromosome 15 (52, 89). These marker chromosomes are also referred to as “inv dup” (15), for inverted duplication of chromosome 15, or “psu dic” (15), for pseudodicentric chromosome 15, where 1 of the 2 centromeres is inactivated. Some physical abnormalities noted in these patients include kyphosis, epicanthic folds, seizure disorder, and hypotonus (52). They also demonstrate moderate to severe mental retardation and epilepsy. Duplications of this region may also arise within the chromosome as an interstitial duplication. Since then, many cases of autism associated with either a maternal interstitial duplication of 15q11–q13 or the presence of a maternal large marker 15 chromosome have been reported (Table 2). In a study of 140 probands with autism, two had a maternal interstitial duplication of 15q11–q13, giving a prevalence of ~1% (37). The extra chromosomal material observed in autism involves the same region that, when deleted, causes Prader-Willi syndrome if paternally inherited or Angelman syndrome when maternally inherited. The differences in these disorders are related to the presence of imprinted genes within the 15q11–q13 region that are expressed from only one parental chromosome. Deleting the single active copy of a gene results in total loss of expression of the affected gene or genes producing the disease phenotype. In familial cases of duplication of the 15q11–q13 region, an increased risk of ASD and other developmental impairments is associated with inheritance from the maternal chromosome (35). Inheritance from the paternal chromosome typically has no effect or at least much lower penetrance (15, 35). Given the incomplete penetrance of maternally inherited interstitial duplications of 15q11–q13 and the range of severity of cases of inv dup (15) cases, it is likely that the phenotype of duplications is modified by other loci either within 15q11–q13 or in other chromosomal regions.

DELETIONS OF THE 15Q11-Q13 REGION Deletions of the 15q11–q13 region were also reported in two cases with autism not associated with either Prader-Willi or Angelman syndrome (132). However, there is no molecular confirmation of this result. The region proximal of 15q11 contains large stretches of DNA ranging in size from 1–20 Mb that contain multiple pseudogenes (44, 111). Routine cytogenetics alone cannot distinguish between deletions/duplications of this proximal region that is associated with a normal phenotype from deletions/duplications of the Prader-Willi/Angelman/autism region that has an abnormal phenotype (19). Without additional molecular testing using FISH, these deletions of 15q11–q13 reported in individuals with autism may represent a coincidental finding of a deletion of the more proximal region.

TURNER SYNDROME Turner syndrome affects females who possess only one copy of the X chromosome (45, X). The typical clinical features include short stature, webbing of the posterior neck, increased carrying angle of the arms, sternal deformities, and streak ovaries. Some individuals with Turner syndrome also show
social and cognitive impairments and have an elevated risk of ASD with attention-deficit/hyperactivity disorder. Analysis of parental origin of the X chromosome revealed that girls who received an X chromosome from their mother (45, Xm) (and therefore did not receive one from their father) had significantly worse social cognition, behavioral inhibition, and verbal and visuospatial memory than those who received their X chromosome from their father (45, Xp) (121). The parent-of-origin effect suggests an imprinted locus (a gene on the X chromosome expressed only when inherited from the father) that could lead to differences between male and female psychosocial development because males do not receive an X chromosome from their father. Such an effect could be partially responsible for male predominance of ASD.

22Q11 DELETION SYNDROME There have been several cases of autism with deletion of the 22q11–q12 region (Table 2). The features of this deletion have been grouped into multiple syndromes including the DiGeorge syndrome, the Velo-cardio-facial syndrome (VCFS), or more generally the 22q11 deletion syndrome. Examination of 32 children and young adults with a confirmed deletion of 22q11 found that 56% had a neuropsychiatric disorder that warranted further assessment (92). The phenotypes varied from 31% with an ASD, 44% with attention-deficit/hyperactivity disorder, and 16% with both. In a different study, 103 subjects diagnosed using ADI criteria for classic autism were tested for the presence of the 22q11 deletion and none were found (97). Overall, some features of the ASDs are observed in individuals with VCFS, including poor social competence and social withdrawal.

Other abnormalities that have multiple cases reported and may represent additional subgroups include deletion 2q37 and deletion 22q13. Table 2 lists eight cases that include a deletion of the chromosome 2q37 region. In a pilot study of 10 subjects with autism, one was identified with a deletion of 2q37 using a scan involving only the telomeres (137). Four cases have been reported that have a deletion of the 22q13 region, including 2 ring chromosomes (Table 2).

Balanced Translocations

Balanced translocations produce an exchange of genomic material between two chromosomes that may or may not cause an abnormal phenotype, depending on the breakage sites. Balanced translocations have been reported for 19 cases of autism or Asperger syndrome (Table 2). Of these, 7 are also within regions showing linkage to autism (Table 3; follow the Supplemental Material link from the Annual Reviews home page at http://www.annualreviews.org). Disrupting genes at the translocation breakpoints in patients with autism quickly identifies potential susceptibility genes as a complement to fine mapping of a large genomic region identified by linkage.

Currently, the breakpoints for four apparently balanced translocations observed in autistic individuals have been cloned. The neurobeachin (NBEA) gene located at 13q13.2 is disrupted in a translocation between chromosomes 5 and 13 [t(5;13)(q12.1;q13.2)] in a patient with autism (23). To add support for
neurobeachin as a susceptibility gene for autism, one individual with autistic disorder had a deletion of 13q12–q13 including the NBEA gene (122), and two cases of autistic disorder with deletions including 13q13 have been reported: del(13)(q12–q14) and del (13)(q12–q22) (112, 124).

The GRPR gene on Xp22.13 is disrupted by t(X;8)(p22.13;q22.1) (63) a region that has also been reported in three cases of deletions of Xp22 and autism (127).

Two other genes encoding neureligins NLGN3 and NLGN4 are located within Xp22. As mentioned above, rare mutations within these genes have been found in individual families (65).

An apparently balanced translocation, t(2;8)(q35;q21.2), in a child with autism had a cryptic deletion of between 4.34 and 4.41 Mb that disrupted the PAX3 gene. The 8q21.1 breakpoint was located within another gene, MMP16, that encodes matrix metalloproteinase 16 (17).

Chromosome 7 has been a hotspot of interest for autism genes. A translocation breakpoint between chromosomes 7 and 13, t(7;13)(q31.3;q21)mat has been cloned and identified the disrupted 7q31 gene as “RAY1” or “FAM4A1” (129). Two inversions, inv(7)(q22q31.2)mat and inv(7)(p12.2q31.3), and another translocation, t(2;7)(p23;q31.3), also occur within this same region and could potentially disrupt this same gene (5, 131).

Another case involving a cryptic deletion associated with an apparently balanced translocation is a t(15;16)(p12?,p13.3) that disrupts A2BP1 (87).

**GENOMIC IMBALANCES**

Several of the chromosomal regions presented in Table 2, including 15q11–q13, 17p11, 17q11, and 22q11, contain multiple copies of nearly identical segments of genomic DNA termed segmental duplications (27) that comprise ~3.6% of the human genome (11, 26). Because of the high degree of similarity between a series of these segmental duplications, identifying clones for sequencing that can be accurately mapped has been difficult. One region that is particularly difficult to map and sequence is the 15q11–q13 region, which leads to a high risk for ASD when duplications of maternal origin occur (26a; S.L. Christian, personal observation). The presence of these segmental duplications makes these regions particularly vulnerable to chromosomal rearrangements that may result in an abnormal phenotype.

The 17q11 region, which contains the neurofibromatosis 1 (NF1) gene, also rearranges to cause deletions that produce a NF1 microdeletion syndrome. These 1.5-Mb deletions contain at least 11 genes, including the serotonin transporter gene, and are present in ~5% to 20% of patients with NF1 (66). Several studies show an estimated frequency of autism in patients with neurofibromatosis, ranging from 0.2% to 14% (48). In another study, 74 patients with NF1 were assessed to determine if an association was present with autism. Three patients (4%) had autism as an additional diagnosis (136). It is unknown whether NF1 due to microduplication has a higher risk for autism than cases with other mutations.

The primary method for detecting most chromosomal rearrangements is through routine cytogenetic analysis that has a minimal resolution of ~2–5 Mb of DNA.
If a region looks potentially abnormal using cytogenetics, then fluorescence in situ hybridization (FISH) analysis can be performed to confirm the cytogenetic result. Another method currently available is a FISH-based approach that studies only the telomeres (74). One study involving telomere FISH on 10 autism patients identified 1 with a deletion of 2q37 (137). Microarray comparative genomic hybridization (array CGH) is emerging as a new technology that will allow the simultaneous screening of the entire human genome for chromosomal imbalances in a single hybridization. This microarray technique will thus allow, for the first time, a cost-effective method for screening patients with autism for chromosomal abnormalities, including abnormalities too small to be detected by conventional techniques.

**CORRELATION OF AUTISM LINKAGE REGIONS WITH CYTOGENETIC ABNORMALITIES**

Whole genome scans have identified multiple regions with suggestive linkage to autism, indicating complex inheritance and genetic heterogeneity of autism and ASDs. Table 3 lists the cytogenetic location for the microsatellite marker with the maximal multipoint lod score (MMLS) identified from each study. The cytogenetic location was determined using the July 2003 release of the human genome sequence comparing the location of the marker with the cytogenetic band that corresponds with the sequence (UCSC human genome database, http://genome.ucsc.edu/) (69). The actual region of linkage may extend a significant distance in both proximal and distal directions. A comparison of the linkage regions with known chromosomal abnormalities observed in autism is presented in the last column on the right in Table 3. In these cases, chromosomal abnormalities that are within one band length of the putative linkage region (i.e., chromosomal abnormalities at 3q25 would be linked to a microsatellite marker in 3q26) have been added from Table 2 to reflect an overlap near the peak most likely within the linkage region.

In multiple studies, some regions show suggestive linkage to autism, increasing support for pursuing further studies to identify susceptibility genes at these locations. The additional identification of a chromosomal abnormality within some of these same regions provides another method to support the presence of autism susceptibility genes within these regions, and, in the case of balanced rearrangements, a breakpoint that may involve an autism susceptibility gene.

**EPIGENETICS AND ALTERNATIVE MODES OF INHERITANCE**

Prior to the development of the field of molecular genetics, “epigenetics” was considered the developmental processes that connect the genotype to the phenotype or a gene-environment interaction. The term “epigenetics” now refers to the temporal
and spatial processes that regulate gene expression (63a). The most common modifications to primary DNA sequence affecting gene expression include methylation and histone acetylation. Methylation of the cytosine residues located in promoter CpG islands prevents binding of transcription factors and shuts down transcription. These modifications allow regulation of both temporal and tissue-specific gene expression. Hypermethylation of the trinucleotide repeat region in *FMR1* causes Fragile X syndrome where 30% of affected individuals have ASD. Parent-of-origin effects (i.e., genomic imprinting) depend on DNA methylation and histone modification of a single parental allele and underlie the maternal chromosome 15q11–q13 duplication syndrome. Histone acetylation is an epigenetic modification that influences chromatin structure. One ASD that results from aberrant chromatin structure is Rett syndrome. The Rett syndrome gene, *MECP2*, encodes a protein that in its normal state promotes chromatin condensation at particular points in the genome.

In more complex genetic disorders such as ASD, there could be more subtle variation in genes that may encode proteins or RNA molecules that act as regulators of gene expression and thereby perturb multiple systems. Yu and colleagues (142) hypothesized that chromatin structure or function may be abnormal in families of patients with ASD, thereby leading to a wide variety of chromosomal abnormalities that reflect rather than cause the primary defect. It is possible that intervention shifting epigenetic effects in one direction or another may eventually provide an alternative treatment route for some patients with ASD.

**Environment**

Two lines of work are worth mentioning as proofs of principle of environmental influences on gene expression. First, a recent study demonstrated that methylation could be exogenously modified in adult humans, with a resulting change in gene expression. Patients on hemodialysis that have elevated homocysteine show impaired DNA methylation and have a resulting shift in expression of imprinted genes that are inappropriately transcribed from both chromosomes. Supplementation with folic acid reduces homocysteine levels, favors DNA methylation, and corrects the pattern of imprinted gene expression (61).

The second line of work centers on the agouti mouse, whose coat color varies from yellow to black in a specific genotype of the *agouti* gene. Litters of genetically identical heterozygous agouti mice show remarkable variation in coat pattern that corresponds to variable methylation of the gene promoter. However, females with a yellow coat tend to have litters composed of yellow and mottled mice, whereas females with a black coat have 20% black offspring (107). A recent study found that supplementation with folic acid and other vitamins early in pregnancy made female mice more likely to bear offspring with yellow coats (134), which suggests that vitamin supplementation can cause transmissible, lifelong alterations in phenotype.
CONCLUSIONS AND FUTURE DIRECTIONS

Autism is a disorder with a strong genetic basis. It is unclear whether the less than 10% of variance that is not genetic is due to stochastic factors, measurement error, environmental influence, mitochondrial genomic variation, or other non-Mendelian genetic influences. ASDs include a small percentage of cases with chromosomal or single-gene variations contributing strongly to susceptibility, but a high percentage of cases relate to multiplicative actions of more common variants. One challenge in autism is how to systematically combine evidence for a given gene or region when some data arises by chromosomal abnormalities, rare variants, association with common variants, and/or linkage findings. This integration may not be feasible, or the failure to develop such approaches may simply demonstrate that some of those investigating autism (possibly even investigators in general) are sometimes too rigid and compulsive to make the creative leaps necessary to solve the puzzle of autism most efficiently and rapidly.

There is no doubt that there is a feeling of disappointment in the field of autism genetics because definitive answers have not come more easily. There are several reasons to put this aside and push forward: (a) autism remains strongly genetic, even if possibly more complex than originally estimated; (b) aside from rare single-gene and chromosomal causes of autism, there are no other cases of strict autism with defined etiology; (c) although each variant in the cases in which multiple relatively common alleles conspire to cause autism may contribute weakly to risk for autism, the potential for understanding autism and developing more effective treatments is not necessarily proportional. An excellent example is type I diabetes: the insulin locus was originally implicated in type I diabetes by association studies, but more than 600 families were needed to confirm by linkage. Despite only weak contribution to genetic susceptibility, administration of insulin continues to be the definitive treatment. In autism, the serotonin transporter may be analogous, given that it is a candidate based on the only pharmacological treatment that can address the core restricted and repetitive behavior domains of autism. Either better understanding of that locus or others may lead to development of treatments with the efficacy analogous to insulin for the full syndrome of autism. Although multiplicative inheritance leads to weaker individual locus susceptibility effects that may be difficult to detect and confirm, a single one of the multiple genetic variants may allow development of novel therapeutic approaches.

Given that it is imperative to move on to full elucidation of the genetics of autism, the main future direction should be to apply innovative technologies, such as CGH, genome-wide association studies beginning with efficient chip-based genotyping currently at the level of greater than 100,000 SNPs per chip, and other innovative strategies. Much larger sample sizes are being collected with measures, such as the ADI-R, which will allow rational association and linkage approaches given phenotypic heterogeneity. Investigators should pay more attention to collecting data on whole blood (platelet) serotonin levels, associated features (e.g., aggression
and precipitants of aggression if present; for example, seizures), and diagnostic features.

ACKNOWLEDGMENTS

This work was supported in part by separate awards to S.L.C. and E.H.C. from the National Alliance for Autism Research and NIH U19 HD35482 (EC).

The Annual Review of Genomics and Human Genetics is online at http://genom.annualreviews.org

LITERATURE CITED

in social recognition and reduction in anxiety-like behavior in vasopressin V1a receptor knockout mice. 

Neuropsychopharmacology 29:483–93


Mol. Psychiatry 8:885–92


Am. J. Hum. Genet. 61:1342–52


Am. J. Hum. Genet. 61:1342–52


Mol. Psychiatry 7:311–16


Science 301:386–89


Nat. Genet. 27:327–31


Genome Biol. 4:R25


Hum. Mol. Genet. 8:1025–37


Synapse 28:33–43


Ann. Neurol. 42:666–69


Mol. Psychiatry 9:587–93


Genome browser at UCSC. *Genome Res.* 12:996–1006


87. Martin CL, Ilkin Y, Powell C, Rao K,


140. Yonan AL, Alarcon M, Cheng R,


