Autism Genetics: Recent Advances in Candidate Gene Search

Vandana Rai

Reader, Human Molecular Genetics Laboratory
Department of Biotechnology, V.B.S. Purvanchal University
Jaunpur, U.P., India

Abstract

Autism is a developmental disorder characterized by impaired social interaction and communication as well as repetitive behaviors and restricted interests. The consequences of this disorder for everyday life adaptation are extremely variable. Family studies have shown that autism runs in families and twin studies indicate that the basis of that familial aggregation is genetic. Despite compelling evidence from twin and family studies indicating a strong genetic involvement in the etiology of autism, the unequivocal detection of autism susceptibility genes remains an elusive goal. This paper will review the literature to date summarizing the results of candidate gene studies with a focus on recent progress.

Keywords: Autism, Neurodevelopment, Candidate Gene, Genome-wide screen, Linkage, Association study

Introduction

The term autism has been made public in the middle of the last century by Leo Kanner [1] and Hans Asperger [2] describing a very specific psychopathology recognized in children. Autism (OMIM 209850) is a serious neurodevelopmental disorder characterized by impairments in social interaction, abnormalities in verbal and nonverbal communication, and restricted, stereotyped interests and behaviors [3]. The population prevalence of autism is debated; recent reports indicate that ~1 in 500 individuals have autism [4, 5]. The symptoms of autism are discernible in the first 3 years of an affected infant's life and manifest throughout the life span. Autistic individuals exhibit a wide spectrum of cognitive abilities, with about 30% displaying severe developmental delay and a significant proportion also experiencing epileptic seizures.
Diagnostic tools for autism have been internationally standardized for reliable diagnoses in view of the worldwide genetic studies. In addition to the diagnostic criteria of DSM-IV3 or ICD-104, this has been accomplished by the release and actual development of the Autism Diagnostic Interview-Revised (ADI-R) [6], a parents or caregivers questionnaire, and the Autism Diagnostic Observation Schedule- Generic (ADOS-G) [7], a direct testing tool of the patients' current behavioral pattern. The causes of autism are still unknown; however, considerable evidence exists for the involvement of genetic factors. The evidence is based on (i) higher concordance rate among monozygotic compared with dizygotic twins [8], (ii) Strong familial aggregation of autism and (iii) higher sibling recurrence risk (6–8%) [9].

Despite the evidence of a genetic component, however, genetics does not explain the whole picture. Autism is most likely a "multifactorial" disorder. Wide-ranging phenotypes, both within families as well as within the monozygotic twins studied, suggest that simple modes of inheritance are not operative. Furthermore, case reports of autism associated with environmental factors, such as rubella virus, valporic acid, and thalidomide exposure during pregnancy, lead many researchers to postulate that nongenetic mechanisms may also produce an autistic syndrome [10].

Postmortem examination of the brains of persons with autism finds consistent evidence for abnormalities in size, density, and dendritic arborization of neurons in the limbic system, including the amygdala, hippocampus, anterior cingulate, and mammillary bodies. There is a stunting of neuronal processes and an increased neuronal packing density, suggesting a curtailment of normal development. These affected regions are strongly interconnected, and together they comprise the majority of the limbic system. The limbic system, especially the amygdala, is part of a neural structure that supports social and emotional functioning. These postmortem findings, therefore, are often heralded as the first good entrance points for understanding the pathobiology of the autism spectrum disorders. There is supportive evidence for an amygdala theory of autism from the experimental monkey work of Bachevalier and colleague [11]. She has produced an animal model of autism by lesioning the amygdalae of monkeys shortly after birth.

The other genetic disorders in which secondary autism is observed include fragile X syndrome [12], tuberous sclerosis [13], Rett syndrome [14], Möbius syndrome [15], Sotos, Neurofibromatosis I, phenylketonuria, Joubert syndrome [16], and Smith-Lemli-Opitz syndrome [17].

Compared with schizophrenia, bipolar disorder and other neuropsychiatric disorders, genetic research in autism is relatively new, but progress has been rapid. However, there remain several unresolved issues in the genetic epidemiology of the disorder. The issue most likely to lead to significant obstacles in identifying susceptibility genes is that of genetic heterogeneity the possibility that two or more independent genetic mechanisms might lead to the disorder. Autism is heterogeneous both in its phenotypic expression and its etiology. The search for genes associated with autism and the neurobiological mechanisms that underlie its behavioral symptoms has been hampered by this heterogeneity. No single gene has yet been specifically linked to autism with replicability, but the disorder is believed to be polygenic.
Candidate genes

Despite the advances that have been made, and reports of some positive findings from both linkage and association genetic studies, thus far not a single susceptibility gene for autism has been identified. The current view is that each locus identified in these studies contains genes with only small or moderate effects on the etiology of autism. These small effect sizes make the identification of specific genes significantly more difficult, especially the fact that it involves both phenotypic and genetic heterogeneity. Numerous researchers have argued that new approaches, which go beyond the standard methods, will be needed for real advances to be made in finding genes for autism over the next few years [18-20].

There are two approaches to identifying genetic contributors to disease. The first is a genome wide search in which linkage or association analysis is used to identify regions of the genome that may contain autism susceptibility genes. The second is the candidate gene approach, which investigates a specific gene or genes for involvement in autism risk.

Several genome-wide scans have been performed for autism and evidence in favor of linkage has been reported for the majority of the chromosomes. However, in most cases, this evidence has not reached statistical significance [19, 21-39]. Genome-wide linkage scans are an unbiased approach to localize genetic factors and have identified several chromosomal regions as promising locations for autism genes, including peaks on chromosomes 2q, 7q, 15q, and 17q.

In the candidate gene approach, genes are chosen for study based on either what is known about the gene’s function, its location (for example in a recognized linkage peak), or a combination of both. In pursuing specific candidates, most studies have focused on genes expressed predominantly in the brain [40, 41]. Several candidate genes are hypothesized to be involved in autism; however, no single candidate gene has consistently emerged as involved in autism risk.

In candidate gene search investigators use various experimental techniques and pathophysiologic models of autism to identify candidate genes. The relevance of these genes to autism pathogenesis is determined by the use of experimental methods to assess the biological activity, expression, and allelic associations in populations with autism and their families. Candidate genes that are involved in the cause of autism are genes whose product is known to play a role in brain development or to be associated with brain structures, neurotransmitters, or neuromodulators implicated in autism on the basis of previous research findings. Once candidate genes have been identified, affected individuals and age-, gender-, and ethnically matched control subjects are tested for the presence of mutations in the gene sequence or relative levels of expressed protein.

Different criteria make a gene or genetic locus eligible for association studies or further screening for variants or mutations. The gene or gene product (i) is thought to be of relevance for behavior in humans, (ii) belongs to a neurodevelopment pathway in the brain by expression in fetal brain tissues, (iii) has been implicated through studies of animal models, (iv) has been identified through a chromosomal abnormality, and (v) has been located positionally by linkage studies.
Approximately 50 genes (Fig. 1 and 2) and many more variants have been investigated for association with autism. Genes tested as candidates for involvement in autism include genes involved in neurotransmission (i.e. HTT, TH, GABA receptors), genes involved in brain development, genes contributing to related disorders (i.e., FMR-1, the fragile X syndrome (FXS) gene, and NF-1, a neurofibromatosis gene), and genes thought to be involved in speech language (i.e., EN2, HRAS-1). Several genes are under investigation like- Neurexin1 [42], MeCP2 [43], GRM8 [44], PTEN [45-47], Reelin [40, 48-53], SLC25A12 [54], HRAS [55,56], EN2 [57], WNT2 [58], HOXA1 [59], CNTNAP2 [60,61], SHANK3 [62], MET [63], AUTS1 [24], UBE3A [64], GABAR [65-67], NLGN3 and NLGN4 [68-71], 5HTT [41,67,72-78]. Though positive results have occasionally been reported, replication has been limited. In this review the main focus is only on thirteen of these genes due to the repeated reports of association of these genes with autism by several investigators- SLC6A4(5-HTT), NRXN, GABAR, MeCP2, EN2, MET, RELN, CNTNAP2, WNT2, FOXP2, PTEN, NLGN3, NLGN4. These genes have been studied by several groups, and both positive and negative results have been reported. These candidates are selected based on known involvement in pathways related to neurodevelopment and/or evidence from pharmacological interventions that implicate specific biomolecular pathways.

Figure1. A glimpse of genes/chromosomal sites implicated for autism based on linkage studies. Black circle denotes centromere. Red colour denotes location of candidate genes.
**Autism Genetics: Recent Advances in Candidate Gene Search**

**Figure 2.** A glimpse of genes/chromosomal sites implicated for autism based on linkage studies. Black circle denotes centromere. Red colour denotes location of candidate genes.

**Solute carrier family 6 gene (SLC6A4 or 5-HTT)**

Genes involved in the physiological pathway of serotonin are strong candidates for autism as serotonin serves as neurotransmitter in the brain responsible for a couple of cognitive functions. This may be a candidate gene because of hyperserotoninemia observed in approximately 25% of patients with autism \[72-75\]. SLC6A4 gene lies at 17q11-q12. Serotonin (5-hydroxytryptamine; (5-HT) is a neurotransmitter in the central and peripheral nervous systems. Following release, 5-HT is actively cleared from synaptic spaces by SLC6A4, a high-affinity, Na\(^+\)- and Cl\(^-\)-dependent transporter localized in presynaptic neuronal membranes. Serotonin transporter mediates reuptake of serotonin from the synapses. Interest in this gene and its protein products derives from a plausible role for serotonin in the repetitive behaviors observed frequently in patients. Impaired function of serotonin system may result in depression, epilepsy, obsessive-compulsive behaviors and affective disorders. Cook et al., \[72\], for example, reported preferential transmission of the small allele of the HTT promoter polymorphism to autistic probands. Klauck et al., \[73\] attempting replication, reported preferential transmission of the larger allele, and the IMGSAC reported no association with either allele \[79\].

**Neurexins gene (NXN)**

Neurexins encode a highly polymorphic family of neuronal proteins that interact with neuroligins to promote synaptic functioning. The three neurexin genes (NRXN1, NRXN2 and NRXN3, located at the human chromosome loci 2q32, 11q13 and...
14q24.3-q31.1, respectively) have two independent promoters, which determine two mRNA classes: long mRNAs, encoding for α-neurexins and short mRNAs, encoding for β-neurexins \[80\]. Evidence for neurexin involvement in autism comes from a number of recent investigations \[39, 80-84\]. Feng et al., \[81\] screened 3 neurexin beta genes in 72 individuals with autism and 535 controls, followed by sequencing of exon 1 of NRXN1β in an additional 192 additional cases. They identified two heterozygous missense mutations (S14L and T40S) and a GG insertion in position 26 of the corresponding protein. People with autism carrying the S14L mutations apparently have seizures and facial dysmorphisms. Missense mutations were found in 4 individuals with autism and in-frame deletions, and insertions were detected in 9 additional cases. No such mutations were reported in controls but they also occurred in first-degree relatives (parents and/or unaffected siblings), who displayed very heterogeneous phenotypes ranging from hyperactivity, depression and/or learning problems to apparently normal behavior.

In another study, a de novo heterozygous 300 kb deletion in the coding exons of the NRXN1 gene was found in two autistic sisters, one was non-verbal, whereas the other had mild language regression \[82\]. In a recent case–control study by Yan et al., \[84\], five rare patient-specific variants were identified in 116 autistic individuals whereas only one additional variant (G28A) was also found in controls. Finally, Kim et al., \[83\] recently identified a number of rare coding variants in a scan of NRXN1 coding exons in 57 individuals with autism. The presence of deletions or other mutations also in unaffected siblings of patients with autism and even in controls \[83, 84\] suggest that such mutations could confer vulnerability to autism rather than causing the disease.

**GABA receptor genes (GABRA,GABRB,GABRG)**

Multiple lines of evidence, including alterations in levels of GABA and GABA receptors in autistic patients, indicate that the GABAergic system, which is responsible for synaptic inhibition in the adult brain, may be involved in autism. GABA is the major inhibitory neurotransmitter in the mammalian CNS, and acts by binding to the GABA-A receptor. Hussman \[85\] suggested that autism is the result of an imbalance of the excitatory glutamatergic and inhibitory GABAergic pathways, resulting in over stimulation in the brain and inability to filter out excess stimuli from environmental and intrinsic sources. This theory is supported by multiple lines of evidence- (i) histological, biochemical, and molecular approaches have demonstrated altered levels and distribution of GABA and GABA receptors in peripheral blood and plasma, as well as in the brain, including decreased GABA-A receptors and benzodiazepine binding sites in the hippocampal formation \[86,87\], (ii) reported alterations in GABAergic neurons, as demonstrated by the increased packing density of GABAergic interneurons in the CA3 and CA1 subfields, and by the decreased numbers and reduced size of cerebellar GABAergic Purkinje cells \[88,89\], (iii) chromosomal abnormalitites like duplications and isodicentric chromosomes in the region containing the three clustered GABA receptor subunits on chromosome 15q have been associated with autism \[90\], (iv) evidence for both linkage and allelic
Autism Genetics: Recent Advances in Candidate Gene Search

association have been reported for this same GABA gene cluster \[^{66,79,91,92}\], and (v) mutations have been reported in multiple GABA receptor genes in families with epilepsy. Investigation of association of GABA receptor subunits outside of the chromosome 15 region has been limited \[^{35}\]. There are 19 known GABA receptor subunits arranged in clusters throughout the genome. Functional pentamers formed by various combinations of these subunits result in receptors of varying properties and sensitivities. The amounts and functional capabilities of individual receptor subunits that form a specific pentamer can affect the amount and quality of signaling in different parts of the brain.

The γ-aminobutyric acid (GABAA) receptor gene cluster located at 15q11-13 (which contains genes for three of the receptor’s subunits: GABRB3, GABRA5, and GABRG3) is strongly implicated in the pathogenesis of autism, given its involvement in the inhibition of excitatory neural pathways and its expression in early development \[^{93}\]. Any malfunction of these genes may have implications for the inhibition of excitatory neural pathways as well as during early brain development and therefore pathological for autism. A couple of linkage and association studies reported limited evidence for involvement of the GABAA receptors, where the most common positive linkage finding was within the GABRB3 gene \[^{8,35,66,67,94-96}\].

Methy-CpG-binding protein 2 gene (MeCP2)

MeCP2 binds to methylated CpG sites and associates with chromatin modifying factors histone deacetylase and DNA methyltransferase. This is a transcriptional repressor that binds to methylated CpG dinucleotides generally located at gene promoters and recruits HDAC1 and other proteins involved in chromatin repression. MeCP2 gene located on chromosome Xq28 and de novo mutations occur in 80% of female patients with Rett syndrome, a pervasive developmental disorder generally characterized by regression, autism, microcephaly, stereotyped behaviors, epilepsy and breathing problems, whereas in males mutations are generally lethal \[^{43}\]. The gene encoding brain derived neurotropic factor has been identified as a target of MeCP2 transcriptional repression. Although MeCP2 was predicted to be a global transcriptional repressor, a paucity of MeCP2 target genes have been identified in MeCP2-null mouse brain, RTT brain or cell lines by gene expression profiling. Lam et al. \[^{97}\] and Carney et al. \[^{98}\] identified mutations in the MECP2 gene in sporadic cases of autism, whereas no mutations in the MECP2 gene were found in a sample of 59 autistic individuals by Vourch et al. \[^{99}\]. Several groups have screened the MECP2 gene for mutations in patients with autism \[^{97-108}\] and reported frameshift, nonsense, missense mutations and de novo splice variant in intron 2 (IVS2 + 2delTAAG). Importantly, female patients with autism carrying MECP2 mutations appear mentally retarded, but do not display any clinical trait resembling Rett syndrome.

Engrailed 2 gene (EN2)

EN2 is located at 7q36.1 and is a homeobox gene that regulates development of the
cerebellum. It has attracted attention as a result of the fact that cerebellum abnormalities are among the most consistent findings from pathological and neuroimaging studies in autism. Evidence for EN2 involvement in autism comes from a number of recent investigations [109-112].

**MET proto oncogene (MET)**

MET is also a strong functional candidate for involvement in autism because it encodes a receptor tyrosine kinase involved in neuronal growth and organization, as well as immunological and gastrointestinal functioning; these are systems in which abnormalities have been suggested in autism. Met gene lies at 7q31. Variants in the MET promoter region show strong association with autism. In particular, Campbell et al., [113] found significant over transmission of the common C allele in autism cases in multiple samples. Case-control comparisons found significant overrepresentation of the C allele in autism, with a relative risk of 2.27. In a separate study, significantly decreased MET protein levels were found in autopsied cortical tissue from individuals with autism [114]. The C risk allele is believed to be a functional regulator of the MET gene. Campbell et al., [113] also found that mouse cells transfected with human MET promoter variants showed a 2-fold decrease in MET promoter activity associated with the C allele.

**Reelin gene (RELN)**

The reelin (RELN) gene, which localizes to a site of chromosomal translocation at 7q22, encodes a large secreted glycoprotein that controls intracellular interactions [115,116]. It is of particular interest given that it binds to neuronal receptors and that the pathology of autism can include migration cell defects [117]. Alterations in RELN protein affect cortical and cerebellar development, and the cerebellar neuronal abnormalities are among the more robust pathologic findings in autism [118]. Both family-based and population-based association studies also indicate that variations in RELN may confer risk to autism. RELN maps to the 7q22 chromosomal region, where suggestive or significant linkage to autism has been reported in several studies. Both family- and population-based association studies also indicate that variations in RELN may confer risk to autism. In particular, a large polymorphic trinucleotide repeat in the 5′ UTR of the RELN gene has been implicated in autism in several studies [50, 53]. Preferential transmission of the large repeat polymorphisms to autistic versus unaffected siblings has also been reported [40,119]. A contribution of RELN in autism is further supported by studies of mutant reeler mice, which carry a large deletion in RELN and show atypical cortical organization similar to the cerebral abnormalities documented in postmortem studies in autism [117].

**Contactin-associated protein 2 gene (CNTNAP2)**

Contactin-associated protein is a member of the neurexin superfamily, a group of transmembrane proteins that mediate cell-cell interactions in the nervous system.
CNTNAP2 is consistently expressed at high levels in the prefrontal and anterior temporal cortex, as well as in the dorsal thalamus, caudate, putamen, and amygdala. Three recent studies further support a role of CNTNAP2 in autism [60, 61, 120]. Arking et al., [60] detected significant linkage at 7q35 (which covers the CNTNAP2 locus), and a follow-up association study in 72 multiplex families found significant over transmission of the T allele in a common polymorphism residing in the intron between exons 2 and 3 of CNTNAP2. Notably, this result was replicated in an independent sample of 1,295 parent-child trios. In another report, Bakkaloglu et al., [120] resequenced CNTNAP2 in a cohort of 635 individuals with autism and 942 controls, finding several rare variants in individuals with autism that were not present in controls. Alarcon et al., [61] reported further evidence implicating CNTNAP2 as an autism susceptibility gene and specifically investigated the association of CNTNAP2 with an autism language phenotype. In a 2-stage association study, investigators found significant association between variants in CNTNAP2 and an index of language delay in autistic children. In addition, a microdeletion in CNTNAP2 was identified in 1 proband and his father but was not seen in 1,000 controls. An independent expression study of fetal brain development was performed, with results indicating preferential expression of CNTNAP2 in the language centers of the brain (i.e. frontal and anterior temporal lobes). Collectively, these studies provide compelling evidence that CNTNAP2 mutations could be associated with autism and perhaps particularly the language endophenotypes of autism. CNTNAP2 is one of the largest genes in the human genome (2.3 million bases or ~1.5% of chromosome 7), and future studies will therefore be important to tease out specific variants that underlie these associations [121].

**Wingless-type mmtv integration site family, member 2 gene (WNT2)**

WNT2 is involved in CNS development and interacts with DVL1. Wassink et al., [122] examined WNT2 as a candidate gene for autism for the following reasons: first, the WNT family of genes influences the development of numerous organs and systems, including the central nervous system; second, WNT2 is located in the 7q31-q33 region linked to autism and is adjacent to a chromosomal breakpoint in an individual with autism; third, a mouse knockout of the dishevelled-1 (DVL1) gene, a member of a gene family essential for the function of the WNT pathway, exhibits a behavioral phenotype characterized primarily by diminished social interaction. Evidence for WNT2 involvement in autism comes from a number of recent investigations.

**Forkhead box P 2 gene (FOXP2)**

FOXP2 gene is located at 7q31.1 and directly regulates expression of the CNTNAP2 gene, encoding a neurexin expressed in developing human cortex, by binding to a regulatory sequence in intron 1. Both FOXP2 and CNTNAP2 are involved in developmental speech and language disorders [123-128]. Several researchers reported association between FOXP2 gene and autism [129, 130].
Phosphatase and tensin homologue gene (PTEN)
PTEN is a tumor suppressor gene located on human chromosome 10q23, influencing G1 cell-cycle arrest and apoptosis. In the central nervous system, PTEN inactivation results in excessive dendritic and axonal growth with increased numbers of synapses [131]. Germline mutations resulting in PTEN haploinsufficiency thus facilitate cell-cycle progression and oncogenesis, leading to macrocephaly/macrosony and to cancer development, respectively [132]. Interestingly, genetic syndromes due to PTEN germline haploinsufficiency, in addition to excessive, dysplastic or neoplastic growth, are often characterised by autism or MR and progressive macrocephaly, as initially reported by Goffin et al., [133]. A study by Butler et al., [134] identified three de novo heterozygous PTEN germline mutations in 18 (16.6%) people with macrocephaly and autism. The three newly identified mutations all lead to missense changes in evolutionarily conserved amino acid residues. Others found PTEN mutations in different percentages of macrocephalic patients with autism [135-137]. Seventy cases of autism with PTEN mutations are invariably characterized by severe to extreme macrocephaly. Macrocephaly has been consistently found in approximately 20% of patients with autism recruited in independent samples [138-142]. Head circumference in these patients with autism is typically normal at birth and an overgrowth seemingly develops over the first few years of life [143]. Interestingly, the majority of macrocephalic patients are actually macrosomic, as head size is significantly correlated with excessive height and weight in this subgroup of patients with autism. The incidence of PTEN de novo mutations in these macrocephalic/macrosonic patients with autism can be estimated at 4.7%. Importantly, these patients are at increased risk of developing a variety of PTEN-related cancers during adulthood.

Neuroligin gene (NLGN)
Thomas et al., [144] have reported deletion in X chromosome of three autistic females, and this report leads to identification of neuroligin3 gene at Xp 22.3 and neuroligin 4 gene atXq13. Neuroligins are cell-adhesion molecules localised postsynaptically in both glutamatergic (NLGN1, NLGN3, NLGN4, NLGN4Y) and gamma-aminobutyric acid-ergic (NLGN2) Synapses [5,145], with important function in synaptogenesis during brain development and in connection of pre- and post-synaptic membraenes. Extensive genetic screens conducted by several groups have confirmed the low frequency of neuroligin mutations among patients with autism [14, 68-71]. A frameshift mutation in NLGN4 and a missense mutation in NLGN3 in two separate families have been found [14] leading to functional inactivation of neuroligins [68]. Laumonnier et al., [69] found a frameshift mutation (D429X in NLGN4) in 13 affected male members of a single pedigree. Yan et al., [70] reported four other NLGN4 missense mutations in patients with autism (G99S, K378R, V403M and R704C). More recently, Lawson-Yuen et al., [71] reported exonic deletion in NLGN4 in a family affected with autism, and a range of other learning and psychiatric disorders. The NLGN4 product has not been studied extensively, but it shares about 70% homology with the rest of the neuroligin family that complexes with neurexin to facilitate synaptic formation and function.
Conclusion
After several decades of halting progress, the entire field of autism genetics is moving forward at a remarkable pace. Over just the past couple of years, a specific genetic mutation in NLGN4 has been identified as being responsible for rare cases of mental retardation and/or pervasive developmental disabilities, EN2 has emerged as a strong candidate for association with the autism phenotype, and a linkage region on chromosome 17q has been confirmed in independent samples using rigorous statistical criteria. These are just a handful of the exciting recent findings in the field that offer avenues for real progress. Of course, the identification of risk alleles or rare causative mutations is just one important step in unraveling the biology of autism, and effort that will require the combined contributions of a variety of fields including geneticists, clinical researchers, developmental neurobiologists and neuroimagers. Though the final goal is to develop new treatments and to reveal strategies for prevention is still over the horizon.

References


Autism Genetics: Recent Advances in Candidate Gene Search


Autism Genetics: Recent Advances in Candidate Gene Search


