A New Genetic Mechanism for Autism

Julie Gauthier and Guy A. Rouleau
Université de Montréal
Canada

1. Introduction
It has been almost half a century since Leo Kanner first described the clinical phenotype associated with Autism\(^1\). Since Kanner’s descriptions, much effort has been devoted to understanding and identifying the factors which may contribute to Autism. It was not until the early 80’s that compelling evidence started to accumulate suggesting that Autism is a disorder of abnormal brain development. It is now generally accepted that both genetic and environmental factors are implicated in the etiology of this intriguing condition. In the current chapter, we will focus on the role of genetic factors involved in the pathogenesis of Autism. While multiple hypotheses have been proposed to account for the genetic origin of Autism, some have received more empirical support than others. While it is likely that more than one genetic mechanisms is involved in the pathogenesis of this disorder, this chapter will focus on a recent hypothesis implicating \textit{de novo} mutations in synaptic genes. This hypothesis is based on the proposition that rare, highly penetrant mutations affecting any of many different genes which code for synaptic molecules, and which are specific to single families, predispose to Autism. Empirical lines of evidence for this hypothesis will be presented, along with examples, some of which are derived from work by our group.

1.1 The role of genetics in autism
Over the last two decades many studies aimed at identifying the genetic causes of Autism have shown that genetic factors play a predominant role in the genesis of this disorder. Twin studies predict that the heritability (i.e. the degree to which a given trait is controlled by inheritance) of Autism is between 70-90\% (Bailey et al., 1995; Marco & Skuse, 2006; Lichtenstein et al., 2010). The relatively low concordance rate in dizygotic twins, the sharp decrease in recurrence risk of Autism in second- and third-degree relatives of autistic subjects (0.18% and 0.12\% respectively) as well as a low risk in first-degree relatives of autistic subjects (3-7\%) (Chakrabarti & Fombonne, 2001; Muhle et al., 2004) predicts two different genetic scenarios: 1) Autism could be explained by the co-inheritance in one individual of multiple disease-predisposing alleles, each with a small but additive effect, resulting in disease, or 2) by \textit{de novo} mutations (i.e. not inherited from either parent). The first scenario is a polygenic inheritance, where the effect of multiple risk genes acting additively or multiplicatively results in disease. The second, very recently considered for Autism, argues that a fraction of the cases would result from incompletely penetrant new

\(^1\) In this chapter the term Autism refers to the autism spectrum disorders which include Autism disorder, Asperger syndrome and Pervasive developmental disorder not otherwise specified
mutations or *de novo* mutations, such that only identical twins would share the genetic predisposition to Autism, hence the much higher monozygotic concordance than dizygotic concordance. Familial aggregation studies demonstrate that the risk of developing Autism is greater in offspring than in the parents. There are many possible explanations, one of which is that new mutations cause a fraction of Autism cases.

Dozens of genome-wide genetic linkage studies have been conducted in Autism kindred, identifying genetic signals of various significance levels on almost every arm of every human chromosome. There are few instances of consistent replication of linkage to any one site. Fine-mapping of these loci has been difficult, mainly because of genetic heterogeneity, so researchers have frequently opted to look at candidate genes directly based on the linkage results and/or on relevant gene functions. Association studies have also been used to search for genes in Autism. Although, in general, most studies have not been replicated, a few have been yielding a crop of possible susceptibility genes. Although the few replicated positive association studies are promising, it is surprising that no causative mutations or sequence variants have been identified in any of the loci associated with the disorder. In the absence of such mutations, the role of these genes in Autism remains unproven.

A number of genes have been strongly associated with Autism. For example, the X-linked neuroligin genes, *NLGN3* and *NLGN4*, two synaptic molecules, have been found to be associated with Autism. In the original report, two families with affected brothers (one with autism disorder and the other with Asperger syndrome) have a frameshifting and a missense mutation within the coding region of *NLGN4* and *NLGN3*, respectively (Jamain et al., 2003). Interestingly, both were new mutations that occurred in the mother of the affected brothers. Other mutations have since been described in these genes further supporting their role in the pathogenesis of Autism. *NLGN3* and *NLGN4* belong to the neuroligin family of postsynaptic cell adhesion molecules that are widely expressed in the brain (Philibert et al., 2000). The products of these genes are involved in late steps of synaptogenesis, mediating the specific recruitment of pre- and postsynaptic proteins to the site of initial synaptic contact. Two independent studies have shown *in vitro* that the frameshifting and the missense mutations described in humans alter the formation of presynaptic terminals (Chih et al., 2004; Comoletti et al., 2004). Since this discovery, mutations in other genes have been linked to Autism (discussed below). These findings support our hypothesis, which will be discussed in section below, that Autism is mainly a synaptic disorder largely caused by *de novo* mutations in synaptic genes.

### 2. The concept of “de novo” mutations

Recent studies on the direct measurement of human mutation rate have revealed that in any single conceptus there is approximately $1.1 \times 10^{-8}$ ($0.76 \times 10^{-8}$ to $2.2 \times 10^{-8}$) mutation per base per generation (Awadalla et al., 2010; Lynch, 2010; Roach et al., 2010). A newborn is thought to have acquired about sixty new mutations in his/her genome. Among these, approximately 0.86 new deleterious mutation will lead to an altered amino acid, which corresponds to an average of about 1 new coding mutation per conceptus (Eyre-Walker & Keightley, 1999; Giannelli et al., 1999; Crow, 2000). *De novo* or spontaneous germline mutations can lead to serious clinical consequences, such as a disease, when affecting critical genes.

#### 2.1 Common disease and common variants

Classical linkage and association studies, as mentioned earlier, have largely failed to identify predisposing genes for Autism as well as a number of other psychiatric disorders. The main
reason for this lack of success is likely to be allelic and non-allelic genetic heterogeneity, with dozens to perhaps hundreds of genes predisposing to Autism, with each gene having many allelic variants. Such heterogeneity would require an enormous sample size to detect predisposing genes using population genetic approaches. It is likely that this heterogeneity results mostly from our limited ability to sub-phenotype brain disorders, particularly behavioural disabilities. The diagnosis of most psychiatric disorders remains largely based on clinical criteria, which define broad categories of dysfunction that may or may not be biologically linked. To date, there is no consistent, biologically validated method for defining these sub-phenotypes. Simply stated, Autism, as currently defined, probably result from so many different genes and alleles that classical genetic methods will prove inefficient in the identification of susceptibility genes for this disorder.

The hypothesis that a common disease may be caused by common variants was the favoured model for the genetic architecture of Autism until recently. Indeed, the constellation of published association studies reflects the widespread belief of the involvement of common variants in Autism. This hypothesis was appealing to many investigators since the common variants should be identifiable using methods such as linkage disequilibrium (Reich & Lander, 2001). Unfortunately, there are very few examples to support this hypothesis, particularly for brain disorders. Clinicians argue that Autism is a highly heterogeneous group of disorders, and none of them can be explained by single or even a few common variants. If this were the case, the plethora of genetic studies performed over the years should already have identified some of these variants. The widely distributed linkage and association positive signals scattered all over the genome rejects the existence of one or a few major predisposing common variants in this disorder. Furthermore, the few genes that have been found to definitely predispose to Autism explain only a small fraction of cases. This is not to say that some common variants will not be found for some predisposing genes, but this mechanism is unlikely to explain all the genetics of this condition.

It has recently been recognized that many complex disorders may result from a mix of common and rare variants. Let us consider breast cancer as an example (Nathanson et al., 2001): BRCA1 and BRCA2 genes contribute to a relatively common genetic disorder, but have many different rare mutations, even in a founder population (the Ashkenazim). For a complex trait such as Autism, the occurrence of many rare variants in many different disease predisposing genes seems to better predict the genetic architecture of the disorders (Pritchard, 2001; Pritchard & Cox, 2002; Smith & Lusis, 2002).

As a final point on the common disease common variant hypothesis, most studies looking at disease-causative mutations for Autism report mutations that are not recurrent, i.e. not observed more than once and specific to one individual. Again this suggests that mutations at many different loci may contribute to Autism, a result consistent with the failure to find common heritable variants with a major effect on disease risk. Lack of recurrence of mutations may in fact reflect the possibility that autistic traits can result from many different genetic defects.

2.2 Rare variants and new mutations
While not all amino acid substitutions will be deleterious, a significant fraction will be and may lead to disease. Therefore, for a disorder that may result from dysfunction in any one of hundreds of different genes, new mutations may be responsible for a significant fraction of cases. For example, should dysfunction in any of 100 different genes potentially lead to
Autism, and assuming amino acid changes lead to gene dysfunction in one fifth of instances, new mutations could cause Autism in one case out of 1,000 births, which would correspond to over 10% of cases based on the overall population incidence.

Looking at simple Mendelian traits, we can see that new mutations are common. For example, 1 in 6,000 live births harbour a novel mutation causing neurofibromatosis type 1 (Stephens et al., 1992; Grimm et al., 1994; Hudson et al., 1997). The frequency of new point mutations in Duchenne Muscular Dystrophy is similar, 1 in 10,500 live births (Grimm et al., 1994). One can argue that these are large genes, allowing for high mutation rate. Nonetheless, these are surprisingly high numbers of novel deleterious mutations. Let us consider Rett syndrome, which is closely related to Autism, and results from mutations in the relatively small MECP2 gene (4 exons, 498 amino acids). The incidence of Rett syndrome is one in 10,000-15,000 females. Because 99-99.5% of all cases are sporadic with new mutations, this represents a new mutation rate of one in 5,000-7,500 live births for this small gene of 498 amino acids (Hagberg, 1985; Hagberg & Hagberg, 1997; Van den Veyver & Zoghbi, 2001). This example clearly shows that, for neurodevelopmental disorders, new mutations can act dominantly and can occur with a high enough frequency to explain the relatively high incidence of Autism. A high rate of new mutations can in part explain why genetic studies have so far failed to identify many Autism genes, and why diseases have been identified for a mere 3% of genes in the human genome. Mutations in genes leading to severe outcome where there is a strong negative selection against the phenotype, such as lethality in embryonic stages or reduced reproductive fitness, will not be transmitted to multiple family members, and therefore will not be detected by linkage gene mapping.

3. The role of de novo mutation in autism

Though we predict that de novo mutations will be a frequent cause of Autism, we do not think that it will be the only genetic explanation. The alternative genetic hypothesis for complex traits, mentioned previously, predicts that disease results from a combination or pattern of genotypes at different susceptibility loci. In recent years, statisticians have developed analytical methods that capture contributions from multiple susceptibility loci, and provide evidence for the localization of disease genes on human chromosomes (Sherriff & Ott, 2001; Hoh & Ott, 2003; Carlson et al., 2004). However, such analyses are very complex and yield few successful examples, even considering the simplest scenarios (Tiret et al., 1994; Bolk et al., 2000; Zetterberg et al., 2003). It seems that genome-wide searches using realistic sample sizes may not have the power to detect potential multi-gene interactions. The existence of new mutations, which contribute to this heterogeneity, makes classical genetic approaches even more difficult. Thus, the failure of conventional linkage and genome-wide association studies to identify but a few causative Autism genes is most likely due to two main confounding factors: phenotypic and genetic heterogeneity. Phenotypic heterogeneity is due to the inability to distinguish closely related clinical subtypes in the autism spectrum of behavioural disturbances. Genetic heterogeneity refers to fact that many different genes (and/or alleles of the same genes) lead to the same phenotype.

3.1 Monozygotic and dizygotic concordance

De novo mutations in identical twins would result in their sharing the same genetic predisposition to Autism. These alleles would be highly but not completely penetrant; hence
the high monozygotic concordance and the low dizygotic concordance, as in the latter case the unaffected twin would not share the novel disease predisposing allele. Instances of non-penetrance would explain the fact that monozygotic twin concordance is not 100%. A de novo or spontaneous mutation can arise from different mechanisms and in different periods in the development of an individual. This kind of mutation can occur in the gametes (sperm or eggs), very early in the developing foetus or later in life as observed in cancer. The partial phenotypic concordance in monozygotic twin could also be in part explained by the occurrence of a de novo mutation early in the development of one twin, but not the other.

3.2 Reduced reproductive fitness
In the general population, the mutational load can be thought of as a balance between selection against a deleterious gene and its acquisition of new mutations. Lower rates of reproduction constitute a negative selection factor that should reduce the number of mutant alleles in the population, ultimately leading to decreased disease prevalence. These selective pressures tend to be of different intensity in different environments. In the case of Autism, only rarely do individuals with Autism have children, particularly the more severely affected individuals (Nicolson & Szatmari, 2003). Thus, Autism has a lower reproductive fitness (which is the ability to pass on genes by having offspring) due to an early age of onset and severely impaired cognitive and social functions. This observation should influence the disorders incidence and prevalence; but this is not what we observe. Autism incidence and prevalence seems to be relatively constant worldwide..

Studies of monogenic diseases\(^2\) indicate that rare diseases with strong negative selection generally exhibit very large allelic diversity, hence many different mutations (Smith & Lusis, 2002). One exception to this pattern of high allelic diversity occurs when disease alleles also provide protection from negative environmental selective pressures. One example is the thalassemias, which confer resistance to malaria. Once arisen, the strong positive selective pressure conferred by these alleles allowed their relatively rapid spread through a specific population. However, such phenomena are usually regional, in response to specific regional environmental pressures. There are no examples of such phenomena occurring with equal strength in all cultural and geographical parts of the world, which needs to be the case to have a uniform incidence of Autism throughout the world. De novo mutations could explain this relatively uniform high incidence of disease, as new disease predisposing alleles will continually be introduced at a similar rate in all parts of the world.

3.3 Effects of paternal age
The male-to-female ratio of de novo mutations is estimated at about 4–6:1, presumably due to a higher number of germ-cell divisions with age in males (Crow, 2000). Therefore, one would predict that de novo mutations would more frequently come from males, particularly older men (Li et al., 2002). At the genetic level, increased risk for a disease with increasing paternal age can be explained by spermatogonial stem cell divisions that occur over the lifetime of males contributing to higher mutational rates in the sperm of older men. A higher paternal origin of de novo mutations has been shown for many diseases, including Apert syndrome (Moloney et al., 1996), Crouzon syndrome (Glaser et al., 2000), Multiple endocrine neoplasia type II (Carlson et al., 1994) and neurofibromatosis type 1 (Jadayel et al., 1990).

\(^2\) Disorders caused by the inheritance of a single defective gene
Rett syndrome, a neurodevelopmental disease closely related to Autism, results almost entirely from new mutations which are exclusively of paternal origin (Girard et al., 2001; Trappe et al., 2001). A role for de novo mutations in Autism would predict that the incidence of disease should increase with increasing paternal age. Indeed, multiple recent studies reported advancing paternal age as a significant risk factor for Autism (Miller, 2006; Cantor et al., 2007; Croen et al., 2007; Puleo et al., 2008). Some authors have predicted a high incidence of male-derived novel mutations in many mental disorders (Preuss et al., 2004). Similar observation has been reported for schizophrenia (Malaspina et al., 2001) and intellectual disabilities (Malaspina et al., 2005), two conditions phenotypically related to Autism. These observations provide strong evidence that accumulation of de novo mutations in paternal sperm contributes to the overall risk of Autism.

3.4 Worldwide incidence
Data from a worldwide amalgam of studies show that the incidence of Autism has been maintained at a constant, relatively high prevalence in the worldwide population across a wide range of cultures and countries (McDonald & Paul, 2010). This occurs despite a strong negative selection against this condition. Indeed and with the exception of variants which date back to speciation, one would expect that common variants would result in a detectable uneven disease incidence across different populations due to migration, different population growth and isolation. This is not the case for Autism. In addition, this is not what one would predict in diseases with reduced reproductive fitness like Autism, unless there was a high new mutation rate. These observations emphasize the importance of de novo mutations in the pathogenicity of Autism.

Taken together, the high prevalence, the high monozygotic twin concordance, the predicted high level of allelic and non-allelic genetic heterogeneity, the uniform worldwide high incidence despite significantly reduced reproductive fitness, constitute evidences that Autism may result at least in part, from de novo mutations.

4. De novo mutations in genes associated with autism
The fact that a growing number of studies, several from our group, report the association of rare genetic variants with Autism constitutes strong evidence for the de novo hypothesis. Indeed, among causal genes identified for Autism, Rett syndrome and intellectual disability (three closely related disorders), the predisposing mutations, whether they be copy number variations, insertions/deletions or point mutations, are very frequently of de novo origin. A good example of such a gene is SHANK3 encoding a synaptic scaffolding protein. Two of the first three mutations reported in the first manuscript linking SHANK3 to Autism were actually of de novo origin; one a deletion of the terminal 22q13 and the other a G insertion leading to a frameshift that was carried by two affected brothers (Durand et al., 2007). None of these mutations were found in the parents. Many subsequent reports on mutation screening of SHANK3 gene in Autism also find novel de novo mutations (Moessner et al., 2007; Gauthier et al., 2009). Other examples, such as the neuroligins genes, NLGN3 and NLGN4, also clearly demonstrate the importance of de novo mutations in Autism. Jamain et al. found a single nucleotide insertion in two affected brothers, one with typical autism and the other with Asperger that arose de novo in the mother (Jamain et al., 2003). The NRXN1 gene has also been found to harbour de novo pathogenic mutations in persons with Autism,
as well as in intellectual disabilities and in schizophrenia (Ching et al., 2010). Ching et al. found twelve deletions in NRXN1 in patients with Autism and four were de novo copy number variations not identified in either parent (Ching et al., 2010). SYNGAP1 (Hamdan et al.) and IL1RAPL1 (Piton et al., 2008) are two other examples where we found de novo mutations in individuals with Autism and/or intellectual disability. Actually, de novo mutations are also a common cause of intellectual disability (Hamdan et al.; Vissers et al., 2010). As expected and as recently observed in Autism, de novo mutations have all been identified in different genes. For example, our group found six de novo deleterious mutations in females individuals with intellectual disability in SYNGAP1 gene (encoding synaptic Ras GTPase activating protein 1) (Hamdan et al.; Piton et al., 2008; Hamdan et al., 2009).

In our recent study on the direct measurement of the de novo mutation rate in Autism and schizophrenia, we found a significant excess of potentially deleterious de novo mutations in individuals with Autism and schizophrenia (Awadalla et al., 2010). In this study, we examined variants identified by direct re-sequencing of 401 genes in a cohort of 285 autistic or schizophrenic individuals and for a subset of these genes in population control individuals. For the analysis, we distinguished functional from non-functional sites based on the effect of a mutation on the transcription or translation of the protein at a given position. Among trios without family history of Autism, we observed a significant enrichment of functional de novo mutations (p = 0.003 in one-tail binomial test; p = 0.022 Fisher’s exact test). Using a binomial test, our observed number of missense to nonsense de novo mutations was also significantly higher than the neutral expectation (p = 0.04), suggesting that some of the mutations are likely to be pathogenic. All of our reported observations suggest an excess of potentially disease-predisposing de novo mutations in the Autism and schizophrenia cohorts. Indeed, in this study, from sequencing only 8% of genes of the human genome, functional de novo mutations were found in 5% of individuals with no family history of Autism, exhibiting a wide range of clinical phenotypes. These few examples and many others recently published collectively provide strong evidence for a major role of de novo mutations in Autism.

5. Altered synaptic connectivity in autism

The synapse is the locus of neural communication which is critical for human brain function. Defects in synaptic transmission are thought to underlie many common developmental brain disorders that are characterized by grossly normal brain structure (Zoghbi, 2003; Levitt et al., 2004). At a cellular level, there are presynaptic nerve endings specialized for the activity-dependent release of transmitter into the synaptic cleft, which is encapsulated by glial cells and contains adhesive molecules that keep presynaptic endings in register with postsynaptic specializations (“densities”) on neural cell bodies and branches. In the mature nervous system these structures signal by chemical transmission and thus integrate and propagate the electrical signals that communicate through the brain. Synapses are thought to form in the embryo largely by genetically pre-programmed, activity-independent and evolutionarily conserved mechanisms (Goodman & Shatz, 1993). During post-natal development, which is the period during which many developmental brain diseases start to manifest themselves, synaptic activity is required to select, refine and stabilize mature connectivity patterns (Katz & Shatz, 1996). Thus cells that fire together wire together.

3 A trio constitutes an affected individual and both his/her biological parents
Multiple indirect lines of evidence support the hypothesis of altered synaptic connectivity in Autism. These come in part from brain-imaging and neuropathological studies showing numerous alterations to both gross and microscopic structures of the brain of autistic individuals. For instance, an increased brain volume (Piven et al., 1996), increased brain weight (Bailey et al., 1998), abnormal neuronal morphology, with decreased complexity of dendritic branching and underdeveloped neuronal arbor (Bauman & Kemper, 1985; Raymond et al., 1996) have all been observed in autistic individuals. An abnormal neuronal density in the cerebellar hemispheres has also been observed (Bauman & Kemper, 1985).

Notably, several components of the limbic system, including the amygdala (Lotspeich & Ciaranello, 1993) and the hippocampus (Raymond et al., 1996), have been found to be abnormal at the microscopic level. Cytoarchitectural features that are frequently abnormal include reduced numbers of Purkinje neurons in the cerebellum and vermis and small tightly packed neurons in regions of the limbic system, especially in the entorhinal cortex and in the medially placed nuclei of the amygdala. The reduced neuronal size and shortened dendritic pattern found in post-mortem studies are consistent with synaptic alterations. This synaptic deficiency hypothesis has been also proposed for schizophrenia, a neurodevelopmental disorder that is also characterized by marked disruptions of information processing and cognition (Glantz & Lewis, 2000).

More recently, in an effort to directly determine if spine densities, or the synaptic connectivity, are altered in autistic subjects, Hussler and Zhang examined the structural microcircuitry within the cerebral cortex i.e. dendritic spine densities on cortical pyramidal cells from autistic subjects and age-matched control cases, on neurons located within both the superficial and deep cortical layers of frontal (BA 9), temporal (BA 21), and parietal lobe (BA 7). They observed several alterations in spine density in autistic subjects; for example the average spine densities in Autism were higher than those found in control cases, supporting altered synaptic connectivity and plasticity in the brains of individuals affected with Autism (Hutsler & Zhang, 2009).

Other evidence suggesting impaired synaptic function in autistic individuals includes the discovery of mutations in different synaptic genes, such as the neurexins, the neurexins and SHANK3 (see examples below in section 5.1). As mentioned earlier, Rett syndrome shares many features with Autism. Mutations in the coding region of the MECP2 gene (a transcription repressor factor expressed by neurons and preferentially abundant in mature neurons) are known to be implicated in this severe disorder, with the vast majority of cases resulting from new mutations. Mutations in MECP2 gene have also been identified in 3 females who meet the full diagnostic criteria for Autism, underscoring the similarity of these diseases (Carney et al., 2003). While the target genes for MECP2 protein remain unknown, the small brain size and the reduced neuronal and dendrite sizes in Rett Syndrome patients suggest that MECP2 may play a role in synaptic processes (Shahbazian et al., 2002; Balmer et al., 2003). Recent findings using the olfactory system as a model to study MECP2 expression during development suggest that it may be involved in the formation of synaptic contacts (Cohen et al., 2003). These data further support the possibility that Autism results mainly from synaptic dysfunction.

### 5.1 Synaptic genes as candidates for autism

At a molecular level, synapses are organized as macromolecular “machines” (Grant, 2003). These synaptic machines consist of a presynaptic release apparatus and a signalling device at the postsynaptic density held together in quasi-crystalline registry at the adhesive cleft.
Many of the proteins constituting these various components have been identified by decades of synaptic biochemistry and physiological genetics, and their macromolecular assemblies have been characterized by proteomic analysis. The presynaptic release apparatus consists of proteins that include those for the structural cytoskeleton, vesicular membrane and trafficking components, vesicle fusion grid and nerve terminal membrane (Phillips et al., 2001; Blondeau et al., 2004). The postsynaptic density consists of structural proteins as well as signalling components such as tyrosine kinases and phosphatases, while both pre- and post-synaptic membranes contain fast voltage-gated channels and neurotransmitter-gated receptors, channels, transporters and G-protein coupled receptors mediating neuromodulation (Walikonis et al., 2000; Satoh et al., 2002). Rapid and selective communication across the synapse is ensured by the firm adhesion of each compartment at the cleft by cell surface as well as secreted extracellular matrix components (Huber et al., 2003). It is therefore not surprising that synaptic genes constitute the largest class of genes associated to developmental brain disorders – with many more to be discovered. Likewise, since many of these proteins are exposed at the extracellular surface, they could provide excellent “druggable” targets.

The discovery of genes clinically relevant to Autism is accelerating, with many involved in the synapse including several neuroligands, as well as genes involved in the glutamatergic pathway (Betancur et al., 2009). Of particular interest is the example of the synaptic cell adhesions and associated molecules including the neuroligins-neurexins-SHANK3 genes. Mutations in the X-linked neuroligin-3 (NLGN3) and neuroligin-4 (NLGN4X and NLGN4Y) genes have been identified in brothers with autism. Laumonnier et al. identified a two base-pair deletion in NLGN4 in 12 affected members of a French family with X-linked mental retardation, some of whom were also autistic (Laumonnier et al., 2004). Jamain et al. identified a C-to-T transition in the NLGN3 gene, in two brothers, one with autism and the other with Asperger syndrome (Jamain et al., 2003). The SHANK3 gene, which codes for a synaptic protein that binds directly to neuroligins, seems crucial for the development of language and social cognition. SHANK3 mutations and small cytogenetic rearrangements have been implicated with the Autism phenotype (Durand et al., 2007; Gauthier et al., 2009). Other genes involved in this pathway have been found to be mutated in autistic individuals. Indeed, variants in SHANK2 and LRRTM1 are reported in schizophrenia and Autism (Francks et al., 2007; Berkel et al., 2010). Other synaptic molecules implicated in Autism are the protocadherin family genes, which have been shown to be associated with Autism. Marshall et al. detected causative copy number variations in PCDH9 gene and a de novo translocation deleting the CDH18 genes in Autism (Marshall et al., 2008). Moreover, in a study of consanguineous families of Autism, a large homozygous deletion implicating PCDH10 was detected (Morrow et al., 2008). All of these examples emphasized the role of impaired synaptic pathways in the pathogenesis of Autism.

6. Similar genetic architecture in other neurodevelopmental disorders

Autism, schizophrenia, and intellectual disability are all severe neurodevelopmental disorders that have childhood or early adulthood onset with a lifetime disability. Clinical manifestations of these disorders are diverse and complex, and include abnormalities in neuronal excitability, processing of complex information, as well as behaviors such as anxiety and impaired social interactions. Pathological studies, neuroimaging and other clinical observations predict that these disorders result from disrupted neurodevelopment
caused by genetic and environmental factors (Lewis & Levitt, 2002). There is a significant overlap in clinical manifestations in these mental disorders, such as episodic psychosis and/or seizures, impaired cognitive functions, and language problems. Fifteen to thirty percent of Autism patients present with seizures and 20% of psychotic patients were diagnosed as having pervasive developmental disorders (Matese et al., 1994). Also, there is no clear clinical or neurobiological distinction between childhood schizophrenia, pervasive developmental disorder and autism (Mouridsen et al., 2000). Furthermore, these neurodevelopmental disorders can be included within the allelic spectrum of the same candidate gene. These observations strongly suggest that Autism, schizophrenia and intellectual disability may share similar pathogenic pathways and, thus, potential candidate genes. In addition, Autism, schizophrenia and intellectual disability have also a high prevalence, a high monozygotic twin concordance, a predicted high level of allelic and non-allelic genetic heterogeneity and a uniform worldwide high incidence despite significantly reduced reproductive fitness. All of these observations support the notion that the de novo mutations model be involved in all these disorders.

6.1 One gene, three phenotypic conditions
We believe that Autism, schizophrenia and intellectual disability, all severe neurodevelopmental disorders, can be studied in a similar manner, which focuses on the synaptic gene de novo mutation model. In addition, in some instances mutations in the same gene can lead to Autism, intellectual disability or schizophrenia, three clinically distinct phenotypes, as defined in the Diagnostic and Statistical Manual of Mental Disorders, the reference manual for the classification of mental disorders. A recent example is our findings with the SHANK3 gene. Disruption of this synaptic gene was originally associated with the 22q13.3 deletion syndrome [OMIM 606232] characterized by neonatal hypotonia, global developmental delay, normal to accelerated growth, absent to severely delayed speech, autistic behavior (OMIM 209850), and minor dysmorphic features. In 2007, Durand et al. and Moessner et al., showed that abnormal gene dosage of SHANK3 is associated with Autism (Durand et al., 2007; Moessner et al., 2007). In addition, we identified a de novo splicing mutation in SHANK3 in a patient with non-syndromic intellectual disability without Autism (Hamdan et al.). We also found deleterious de novo mutations in the SHANK3 gene in a patient diagnosed with schizophrenia plus cognitive impairment. Similarly, mutations in the NRNX1 gene can lead to rare forms of Autism and schizophrenia (Kim et al., 2008; Rujescu et al., 2009). This phenomenon has been observed in other diseases and, as stressed by Zoghbi and Warren in their recent paper (Zoghbi & Warren, 2010), other examples include the ARX gene which causes X-linked lissencephaly, agenesis of corpus callosum with abnormal genitalia, cognitive deficits with or without seizures, or cognitive deficits, dystonia, and seizures; LMNA gene, which can lead to a diversity of disorders including Emery-Dreifuss muscular dystrophy Type 2, Charcot-Marie-Tooth axonal neuropathy limb girdle muscular dystrophy Type 1B, Hutchinson-Gilford progeria syndrome, and many other different clinical manifestations.

Altogether, these findings suggest that the neuroanatomical and physiological disturbances resulting from dysfunction of mutant genes may be influenced by the effect of genetic modifiers, the nature of the gene’s role in the human brain and the effect of environmental experiences of the affected individuals, leading to different clinical outcomes in different patients. Differences in the mutation types (for example, point mutation vs. large gene disruptions) must certainly also contribute to the phenotypic variability. Although this
observation is intriguing, multiple phenotypic manifestations from mutations of the same single gene have been described for many other diseases. Finally, the observation that one gene can lead to many phenotypes raised the question of whether Autism, schizophrenia and intellectual disability are different entities or part of a same phenotypic continuum.

7. Gene hunting approaches and the impact of the development of new technologies

In the last few years, a new generation of technologies, referred to as the next-generation DNA sequencing technologies have been developed which allow screening of the entire genome (i.e. > 20,000 genes) of single individuals within a matter of days. This new technology has revolutionized genetic research and has allowed new approaches in the search for diseases-causing genes. Before the advent of the next-generation DNA sequencing technologies, gene screening for the identification of disease-causing mutation was done one gene at a time. Next-generation DNA sequencing enable the parallel sequencing of all the 20,000 genes, leading to faster identification of mutations. These technologies therefore constitute the ideal method of screening for rare causative variants in all genes simultaneously. In the context of Autism and of the above mentioned de novo mutation genetic mechanism, the major advantage of these technologies is to make possible the identification of very rare de novo mutations, by comparing the genetic variants in an affected subject to those in both of his/her parents (a family trio). Given sufficient coverage and quality in next-generation DNA sequencing datasets, identifying de novo mutations in trios is highly feasible.

Sequencing of entire genomes is still rather expensive, so many groups now focus on the sequencing of the entire “exome” (i.e., the various coding regions of the genome) of an individual. Focusing on the exome is a reasonable approach as the vast majority of disease-causing mutations identified to date disrupt the protein-coding regions of genes. Such mutations include nonsense, small insertion/deletions, frameshifts, splicing and missense mutations, whose consequences can be predicted in silico based on well-annotated reference datasets (e.g. the human consensus CDS (CCDS) subset of the NCBI RefSeq database includes 23,339 consistently annotated protein-coding transcripts). These coding regions constitute less than 3% of the entire genome. Oligonucleotide hybridization-based methods that permit the capture and amplification of virtually all human exons at once, i.e. the human “exome”, are now commercially available (NimbleGen, Agilent, Illumina). Limiting the analysis to the exome significantly increases the number of samples that can be sequenced, and is more likely to identify causative genes. Since it is likely that highly penetrant alleles (i.e. protein-truncating or missense) in different autistic cases will result from mutations in dozens of different genes, it is preferable to concentrate on re-sequencing only the coding regions of the genome, or the “exome”, using targeted microarray capture followed by next generation sequencing. In addition mutations in coding regions are most easily interpreted thus making the link to the disorder easier to establish. The availability of these technologies is accelerating all aspects of the gene hunting process e.g. increasing number of genes that can be now analysed in a shorter period of time and the number of subjects being studied.

7.1 Challenges for the next-generation approaches

As next-generation DNA sequencing technologies improve, and as it becomes possible to rapidly produce detailed lists of variants per individual genome, the challenge will be to
discern the pathogenic variants from the benign ones and establish the link to the disorder. Three major challenges can be identified. The first and most technical one is the ability to handle very large datasets in the order of tens or hundreds of terabytes in size, and access to powerful computing platforms that can process these datasets, and adequate resources for storage, retrieval and archiving. The second is the capacity to develop robust yet comprehensive methods to identify variants from next-generation DNA sequencing datasets. Choosing the correct sequence coverage and quality filters will ensure a maximum of true variants to be identified, with as few false positives or false negatives as possible. Several programs are available that can align short sequence reads to a reference genomic sequence and call potential homozygous or heterozygous variants. However, the total number of variants identified, even using different parameters within the same program, can vary largely. Intuitive paradigms and empirically determined cut off points need to be implemented. Furthermore, the accurate annotation of genomic variants is critical to classifying different variants for their potential impact on transcription, splicing or translation. This step must be comprehensive so that potential protein-truncating variants are not missed given that alternate splicing can lead to different transcripts with different open reading frames within a single gene. Finally, developing an experimental design based on the known or anticipated genetic mechanisms underlying the disease or condition, and on high quality diagnostic procedures, requires that affected and unaffected individuals be carefully selected from family groups to help prioritize variants for further analysis, and to maximize the chances of finding causative genes. In our case, identifying de novo mutations by analysing trios can very quickly lead to the identification of causative mutations and risk genes. Successful mastery of these key strategic and technical competencies is necessary to identify potentially pathogenic variants.

7.2 Copy number variations and autism

The use of microarray approaches for the detection of copy number variations, which its continuously improving resolution, provides additional evidence for the occurrence of de novo genomic events in the pathogenesis of Autism. In the last decade, studies linking copy number variation, and Autism have revealed that de novo and inherited copy number variations, including deletions and duplications, translocations, and inversions of chromosomes, all may significantly contribute to the pathogenesis of Autism, usually as penetrant rare variants (Sebat et al., 2007; Walsh et al., 2008). For example, Sebat et al. showed that de novo copy number variations are more common in autistic patients than in non-autistic individuals (Sebat et al., 2007). They found that 10% of their patients with sporadic Autism (i.e. individuals with no history of the disorder) harboured a de novo copy number variation, while the frequency was only 3% in patients with an affected first-degree relative and 1% in controls. Other similar examples include the study from Marshall et al. (Marshall et al., 2008), Christian et al. (Christian et al., 2008) and Szatmari et al. (Szatmari et al., 2007) and more recently the report of Bremer et al. (Bremer et al.), which are all consistent with the hypothesis that de novo or weakly recurrent copy number variations seem to be significant contributing factor in the pathogenesis of Autism. Interestingly, based on the results of their copy number variations analyses, Marshall et al. concluded that “structural de novo were found in sufficiently high frequency in Autism subjects suggesting that cytogenetic and microarray analyses be considered in routine clinical workup” (Marshall et al., 2008). Although this is not the focus of this chapter, the genes identified by copy number variation
analysis also support the notion that there are shared biological pathways in Autism, intellectual disability and schizophrenia (Guilmatre et al., 2009).

8. Conclusion and directions for future studies

As outlined throughout this chapter, several lines of evidence support the role of \textit{de novo} mutations in the pathogenesis of Autism. \textit{De novo} mutations are a well-established genetic mechanism for the development of a number of disorders such as Rett Syndrome and certain types of cancers but have been poorly explored for common diseases like Autism. In general, the development of technologies often brings new challenges, but mostly allows research to accelerate. Indeed, this technological progress is already starting to provide data supporting the role of \textit{de novo} mutations in Autism. This hypothesis is gaining acceptance in the scientific community, as reflected in the growing number of recent publications on this subject. Although the focus of this chapter is on the role of \textit{de novo} mutations in Autism, we acknowledge the fact that many other genetic or non-genetic mechanisms certainly contribute to this disorder.

The accessibility of next-generation DNA sequencing methodologies have enabled researchers to analyse a large amount of DNA and has had an important impact on gene hunting strategies, which have shifted from a tendency to look at single genes, one at the time, to multiple genes simultaneously. One interesting consequence of next-generation DNA sequencing is that it recently permitted to directly estimate the rate of \textit{de novo} germline base substitution mutations in humans (Awadalla et al., 2010; Durbin et al., 2010). Based on these data, the challenge will be to determine if the observed \textit{de novo} mutation rate detected in a disease is greater than the baseline rate.

In the last few years, researchers have identified several genes contributing to Autism, and most encode for proteins that are part of the synaptic machinery. An important concern for future research, where there will be rapid identification of many potential Autism gene mutations, will be to determine if they have a functional relevance to the disorder. Indeed, this question needs to be judiciously examined for most of the variants discovered. This will require to study model organism systems as proposed in our current Synapse to disease project (S2D; http://www.synapse2disease.ca), a large-scale medical research project launched in 2006 aiming to identify genes involved in several neurological and psychiatric diseases caused by defects in the development and functioning of the brain and nervous system. This project’s philosophy is that once the base changes are discovered and considered likely to be “pathogenic mutations”, biological validation must be conducted \textit{in vitro} and \textit{in vivo} in different model organism (e.g. fly, worm, fish, etc.) to determine their functional effects. Biological validation is an essential step often missing from most genetic studies and has thus severely limited data interpretation in the context of disease pathology. This validation will consequently help understanding pathways in neurodevelopmental disorders and ultimately give insights for the development of targeted therapeutic strategies.

Another challenge for future research in the field is the issue of whether genomic variants beyond the coding regions of a gene contribute to the etiology of the disorder. As mentioned earlier, the majority of mutations identified in Autism are located within coding regions but it should not be forgotten that variants in the non-coding regions, particularly the regulatory gene region, can also lead to disease.
Finally, the accessibility of the next-generation DNA sequencing technologies is facilitating the gene hunting process for researchers, whereas its application for clinical diagnostic testing seems to be inevitable, particularly as the cost per base continues to decrease. Although, the clinical tests based on these technologies represent particular challenges and will need careful validation, the connection between research findings in the genetics of Autism or any other neurodevelopmental disorders and clinical applications is closer than ever. All these research and technological advancements are for the greatest benefit of families.

9. Acknowledgments

We wish to thank Anna Bonnel and Fadi F. Hamdan for their critical reading of this manuscript.

10. References


Autism Spectrum Disorders: The Role of Genetics in Diagnosis and Treatment


Estimated prevalence rates of autism spectrum disorders (ASDs) have increased at an alarming rate over the past decade; current estimates stand as high as 1 in 110 persons in the population with a higher ratio of affected males to females. In addition to their emotional impact on the affected persons and their family members (in fact, the latter are often unrecognized unaffected â€œpatientsâ€ themselves), the economic and social impacts of ASDs on society are staggering. Persons with ASDs will need interdisciplinary approaches to complex treatment and life planning, including, but not limited to, special education, speech and language therapy, vocational skills training and rehabilitation, social skills training and cognitive remediation, in addition to pharmacotherapy. The current book highlights some of the recent research on nosology, etiology, and pathophysiology. Additionally, the book touches on the implications of new research for treatment and genetic counseling. Importantly, because the field is advancing rapidly, no book can be considered the final word or finished product; thus, the availability of open access rapid publication is a mechanism that will help to assure that readers remain current and up-to-date.

**How to reference**
In order to correctly reference this scholarly work, feel free to copy and paste the following:
