The Genetics of Autism Spectrum Disorders

Guillaume Huguet, Marion Benabou, and Thomas Bourgeron

Abstract In the last 30 years, twin studies have indicated a strong genetic contribution to Autism Spectrum Disorders (ASD). The heritability of ASD is estimated to be 50%, mostly captured by still unknown common variants. In approximately 10% of patients with ASD, especially those with intellectual disability, de novo copy number or single nucleotide variants affecting clinically relevant genes for ASD can be identified. Given the function of these genes, it was hypothesized that abnormal synaptic plasticity and failure of neuronal/synaptic homeostasis could increase the risk of ASD. In parallel, abnormal levels of blood serotonin and melatonin were reported in a subset of patients with ASD. These biochemical imbalances could act as risk factors for the sleep/circadian disorders that are often observed in individuals with ASD. Here, we review the main pathways associated with ASD, with a focus on the roles of the synapse and the serotonin-NAS-melatonin pathway in the susceptibility of ASD.

G. Huguet • M. Benabou
Institut Pasteur, Human Genetics and Cognitive Functions Unit, Paris, France
CNRS UMR3571 Genes, Synapses and Cognition, Institut Pasteur, Paris, France
Sorbonne Paris Cité, Human Genetics and Cognitive Functions, University Paris Diderot, Paris, France

T. Bourgeron (*)
Institut Pasteur, Human Genetics and Cognitive Functions Unit, Paris, France
CNRS UMR3571 Genes, Synapses and Cognition, Institut Pasteur, Paris, France
Sorbonne Paris Cité, Human Genetics and Cognitive Functions, University Paris Diderot, Paris, France
FondaMental Foundation, Créteil, France
Gillberg Neuropsychiatry Centre, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
e-mail: thomasb@pasteur.fr

© The Author(s) 2016
P. Sassone-Corsi, Y. Christen (eds.), A Time for Metabolism and Hormones, Research and Perspectives in Endocrine Interactions, DOI 10.1007/978-3-319-27069-2_11
Introduction

Autism Spectrum Disorders (ASD) are a group of neuropsychiatric disorders characterized by problems in social communication as well as the presence of restricted interests and stereotyped and repetitive behaviors (Kanner 1943; Asperger 1944; Coleman and Gillberg 2012). Epidemiological studies estimate that more than 1% of the population could receive a diagnosis of ASD (Elsabbagh et al. 2012; Developmental Disabilities Monitoring Network Surveillance Year Principal 2014). Individuals with ASD can also suffer from other psychiatric and medical conditions, including intellectual disability (ID), epilepsy, motor control difficulties, Attention-Deficit Hyperactivity Disorder (ADHD), tics, anxiety, sleep disorders, epilepsy, depression or gastrointestinal problems (Gillberg 2010; Moreno-De-Luca et al. 2013). The term ESSENCE, for ‘Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations,’ was coined by Christopher Gillberg to take into account this clinical heterogeneity and syndrome overlap (Gillberg 2010). There are four to eight times more males than females with ASD (Elsabbagh et al. 2012), but the sex ratio is more balanced in patients with ID and/or dysmorphic features (Miles et al. 2005). Autism can be studied as a category (affected vs. unaffected) or as a quantitative trait using auto- or hetero-questionnaires such as the Social Responsiveness Scale (SRS) or the autism quotient (AQ) (Ronald et al. 2006; Skuse et al. 2009; Constantino 2011). Using these tools, autistic traits seem to be normally distributed in clinical cases as well as in the general population (Ronald et al. 2006; Skuse et al. 2009; Constantino 2011).

The causes of autism remain largely unknown, but twin studies have constantly shown a high genetic contribution to ASD. Molecular genetics studies have identified more than 100 ASD risk genes carrying rare and penetrant deleterious mutations in approximately 10–25% of patients (Huguet et al. 2013; Gaugler et al. 2014; Bourgeron 2015). In addition, quantitative genetics studies have shown that common genetic variants could capture almost all the heritability of ASD (Huguet et al. 2013; Gaugler et al. 2014). The genetic landscape of ASD is shaped by a complex interplay between common and rare variants and is most likely different from one individual to another (Gardener et al. 2011; Hallmayer et al. 2011; Bourgeron 2015). Remarkably, the susceptibility genes seem (Huguet et al. 2013) to converge in a limited number of biological pathways, including chromatin remodeling, protein translation, actin dynamics and synaptic functions (Bourgeron 2009; Toro et al. 2010; Huguet et al. 2013; Bourgeron 2015). In addition, several studies have pointed to a dysfunction of the serotonin-NAS-melatonin pathway in patients with ASD. Abnormalities of this pathway might increase the risk of circadian/sleep disorders often observed in patients with ASD.

In this chapter, we will detail the advances in the genetics of ASD (Fig. 1) with a focus on the role of both synapses and biological rhythms in the susceptibility of ASD (Abrahams and Geschwind 2008; Bourgeron 2009; Toro et al. 2010; Devlin and Scherer 2012; Huguet et al. 2013).
Twin and Family Studies in ASD

Based on more than 13 twin studies published between 1977 and 2015, researchers have estimated the genetic and environmental contribution to ASD (Fig. 2). In 1977, the first twin study of autism by Folstein and Rutter (1977) reported on a cohort of 11 monozygotic (MZ) twins and 10 dizygotic (DZ) twins. This study showed that MZ twins were more concordant for autism—36 % (4/11)—compared with 0 % (0/10) for DZ twins. When a “broader autism phenotype” was used, the concordance increased to 92 % for MZ twins and to 10 % DZ twins (Bailey et al.1995). Since this first small scale study, twin studies have constantly reported a higher concordance for ASD in MZ compared with DZ (Ritvo et al.1985; Steffenburg et al. 1989; Bailey et al. 1995; Le Couteur et al. 1996). Between 2005 and 2009, three twin studies with relatively large groups of twins (285–3419) have reported high concordances for ASD in MZ twins (77–95 %) compared with DZ twins (31 %; Ronald et al.2005; Taniai et al.2008; Rosenberg et al. 2009). Notably, MZ concordances were similar to those reported in the previous studies, but DZ concordances were higher. In 2010, Lichtenstein et al. reported a relatively low concordance for ASD in 39 % of the MZ twins compared with other studies (the concordance for DZ twins in this study was 15 %). However, as previously indicated by studies using the “broader autism phenotype,” all discordant MZ twins of this cohort had symptoms of ESSENCE (e.g., ID, ADHD, language delay, etc.). A significant proportion of the genetic contribution to ASD was shown to be shared with other neurodevelopmental disorders such as ADHD (>50 %) and learning disability (>40 %; Lichtenstein et al. 2010; Ronald et al. 2010; Lundstrom et al. 2011; Ronald and Hoekstra 2011). In summary, when all twin studies are taken into account, concordance for ASD is roughly 45 % for
MZ twins and 16% for DZ twins (Ritvo et al. 1985; Steffenburg et al. 1989; Bailey et al. 1995; Le Couteur et al. 1996).

Family studies also showed that the recurrence of having a child with ASD increases with the proportion of the genome that the individual shares with one affected sibling or parent (Constantino et al. 2010; Risch et al. 2014; Sandin et al. 2014). In a population-based sample of 14,516 children diagnosed with ASD (Sandin et al. 2014), the relative risk for ASD (compared to the general population) was estimated to be 153.0 [95% confidence interval (CI): 56.7–412.8] for MZ twins, 8.2 (3.7–18.1) for DZ twins; 10.3 (9.4–11.3) for full siblings, 3.3 (95% CI, 2.6–4.2) for maternal half siblings, 2.9 (95% CI, 2.2–3.7) for paternal half siblings, and 2.0 (95% CI, 1.8–2.2) for cousins.

Heritability is the proportion of the phenotypic variation in a trait of interest, measured in a given studied population and in a given environment, that is co-varying with genetic differences among individuals in the same population. In 1995, based on a twin study, Bailey et al. estimated the heritability of autism to be 91–93%. Since then, the estimation of heritability has differed from one study to another, but the genetic variance has accounted for at least 38% and up to 90% of the phenotypic variance (Hallmayer et al. 2011; Ronald and Hoekstra 2011; Sandin et al. 2014). Using a large cohort of 14,516 children diagnosed with ASD Sandin et al. (2014), estimated the heritability to be 0.50 (95% CI, 0.45–0.56) and the non-shared environmental influence was also 0.50 (95% CI, 0.44–0.55).

Fig. 2 The main twins studies in ASD. A total of 13 twins studies and 17 heritability studies are depicted. Means of concordance and heritability weighted by sample size are presented on the right of the figure (Adapted from Huguet and Bourgeron 2016)
Surprisingly, only the additive genetic component and the non-shared environment seemed to account for the risk of developing ASD (Sandin et al. 2014).

In summary, epidemiological studies provide crucial information about the heritability of ASD. However, they do not inform us about the genes involved or the number and frequency of their variants. In the last 15 years, candidate genes and whole-genome analyses have been performed to address these questions.

From Chromosomal Rearrangements to Copy Number Variants in ASD

The first genetic studies that associated genetic variants with ASD used observations from cytogenetic studies (Gillberg and Wahlstrom 1985). However, because of the low resolution of the karyotypes (several Mb), it was almost impossible to associate a specific gene to ASD using this approach. The prevalence of large chromosomal abnormalities is estimated to be less than 2 % (Vorstman et al. 2006). Thanks to progress in molecular technologies such as Comparative Genomic Hybridization (CGH) or SNP arrays, the resolution in the detection of genomic imbalances has dramatically increased. Depending on the platforms, Copy Number Variants (CNVs) of more than 50 kb are now robustly detected (Pinto et al. 2011). Since the first articles published in 2006, a very large number of studies have investigated the contribution of CNVs to ASD (Jacquemont et al. 2006; Sebat et al. 2007). Several studies using the Simons Simplex Collection could even provide an estimation of the frequency of the de novo CNVs in patients with ASD compared with their unaffected siblings (Sanders et al. 2011). All together, de novo CNVs are present in 4–7 % of the patients with ASD compared to 1–2 % in the unaffected siblings and controls (Glessner et al. 2009; Sanders et al. 2011; Pinto et al. 2014). The studies have also indicated that de novo CNVs identified in patients are most likely altering genes and most especially genes associated with synaptic functions and/or regulated by FMRP, the protein responsible for the fragile X syndrome (Pinto et al. 2010, 2014). Beyond ASD, large CNVs (>400 kb) affecting exons are present in 15 % of patients with Developmental Delay (DD) or ID (Cooper et al. 2011). Most of the CNVs are private to each individual, but some are recurrently observed in independent patients. For example, three loci on chromosomal regions 7q11, 15q11.2–13.3, and 16p11.2 have been strongly associated with ASD (Ballif et al. 2007; Kumar et al. 2008; Weiss et al. 2008; Szafranski et al. 2010; Sanders et al. 2011; Leblond et al. 2014).

In summary, large chromosomal rearrangements and CNVs increase the risk of having ASD in 5–10 % of the individuals (Vorstman et al. 2006; Pinto et al. 2010, 2014). To go further in the identification of the ASD risk genes, candidate genes and whole exome/genomes studies were performed.
From Candidate Genes to Whole Exome/Genome Sequencing Studies in ASD

The first approach to associate a gene with ASD was to select specific candidate genes based on data coming from functional or genetic studies or a combination of the two. This approach was successful in identifying several synaptic genes associated with ASD such as NLGN3, NLGN4X, SHANK3 and NRXN1 (Jamain et al. 2003; Durand et al. 2007; Szatmari et al. 2007). Thanks to the advance in Next Generation Sequencing (NGS), we can now interrogate all genes of the genome in an unbiased manner using Whole Exome/Genome Sequencing (WES, WGS).

To date, more than 18 WES studies of sporadic cases of ASD (O’Roak et al. 2011, 2012a; Chahrour et al. 2012; Iossifov et al. 2012; Neale et al. 2012; Sanders et al. 2012; He et al. 2013; Lim et al. 2013; Liu et al. 2013, 2014; Willsey et al. 2013; Yu et al. 2013; An et al. 2014; De Rubeis et al. 2014; Iossifov et al. 2014; Samocha et al. 2014; Chang et al. 2015; Krumm et al. 2015) have been performed, comprising altogether more than >4000 families (Table 1). In almost all these studies, the authors have especially focused their analysis on the contribution of de novo Single Nucleotide Variants (SNVs) in ASD. All together, the average number of de novo coding SNVs per individual (including missense, splicing, frameshift, and stop-gain variants) is estimated to be approximately 0.86 in female patients, 0.73 in male patients, and 0.60 in unaffected male and female siblings (Krumm et al. 2014; Ronemus et al. 2014). Interestingly, de novo SNVs were three times more likely to be on the paternal chromosome than on the maternal one (Kong et al. 2012; O’Roak et al. 2012a) with an increase of almost two de novo mutations per year and doubled every 16.5 years (Kong et al. 2012).

Based on these studies (Iossifov et al. 2012; Neale et al. 2012; O’Roak et al. 2012a; Sanders et al. 2012), 3.6–8.8 % of the patients were shown to carry a de novo causative mutation (Iossifov et al. 2012) with a twofold increase of deleterious mutations in the patients compared with their unaffected siblings. In a meta-analysis, using more than 2500 families, Iossifov et al. (2014) found that de novo Likely Gene Disrupting (LGD) mutations (frameshift, nonsense and splice site) were more frequent in patients with ASD compared with unaffected siblings (P = 5 × 10^{-7}). The carriers of these de novo LGDs were more likely diagnosed with a low non-verbal IQ. The de novo LGDs are significantly enriched in genes involved in chromatin modeling factors (P = 4 × 10^{-6}) and in genes regulated by the FMRP complex (p = 4 × 10^{-7}). Following these whole exome studies, targeted re-sequencing studies of the most compelling candidate genes were performed (O’Roak et al. 2012b). All together, 10 genes carrying de novo mutations were significantly associated with ASD: CHD8, Dyrk1A, GRIN2B, KATNAL2, RIMS1, SCN2A, POGZ, ADNP, ARID1B and TBR1.

Only a few studies have analyzed the contribution of inherited SNVs in ASD. In 2013, Lim et al. analyzed whole exome sequencing of 933 cases (ASD) and 869 controls for the presence of rare complete human knockouts (KO) with homozygous or compound heterozygous loss-of-function (LoF) variants (≤5 %
Table 1  Summary of the main whole exome/genome sequencing studies in ASD

<table>
<thead>
<tr>
<th>Studies</th>
<th>Tech</th>
<th>#ASD analyzed in the study</th>
<th># ASD specific to this study</th>
<th>#ASD coming from other studies</th>
<th>#Controls</th>
<th>#Unaffected sibling</th>
<th>#Parents</th>
<th>Analysis of de novo variants</th>
<th>Analysis of inherited variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Roak et al. (2011)</td>
<td>WES</td>
<td>20</td>
<td>20</td>
<td>–</td>
<td>–</td>
<td>20</td>
<td>38</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>O’Roak et al. (2012a)</td>
<td>WES</td>
<td>229</td>
<td>209</td>
<td>20 (O’Roak et al. 2011)</td>
<td>–</td>
<td>50</td>
<td>418</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>Neale et al. (2012)</td>
<td>WES</td>
<td>175</td>
<td>175</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>350</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>Sanders et al. (2012)</td>
<td>WES</td>
<td>238</td>
<td>238</td>
<td>–</td>
<td>–</td>
<td>200</td>
<td>476</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>Iossifov et al. (2012)</td>
<td>WES</td>
<td>343</td>
<td>343</td>
<td>–</td>
<td>–</td>
<td>343</td>
<td>686</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Chahrour et al. (2012)</td>
<td>WES</td>
<td>16</td>
<td>16</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>X</td>
</tr>
<tr>
<td>Lim et al. (2013)</td>
<td>WES</td>
<td>1496</td>
<td>1004</td>
<td>492 (Sanders et al. 2012)</td>
<td>5474</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Studies</th>
<th>Tech</th>
<th>#ASD analyzed in the study</th>
<th># ASD specific to this study</th>
<th>#ASD coming from other studies</th>
<th>#Controls</th>
<th>#Unaffected sibling</th>
<th>#Parents</th>
<th>Analysis of de novo variants</th>
<th>Analysis of inherited variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samocha et al. (2014)</td>
<td>WES</td>
<td>1078</td>
<td>−</td>
<td>1078 (Asperger 1944; Coleman and Gillberg 2012; Developmental Disabilities Monitoring Network Surveillance Year Principal 2014; Elsabbagh et al. 2012; Gillberg 2010; Constantino 2011)</td>
<td>−</td>
<td>343</td>
<td>2156</td>
<td>X</td>
<td>−</td>
</tr>
<tr>
<td>An et al. (2014)</td>
<td>WES</td>
<td>40</td>
<td>40</td>
<td>−</td>
<td>−</td>
<td>8</td>
<td>80</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Study</td>
<td>Total Samples</td>
<td>WGS Coverage</td>
<td>SNV Detection</td>
<td>Indel Detection</td>
<td>SNP Detection</td>
<td>RNA Detection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------------</td>
<td>--------------</td>
<td>---------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kong et al. (2012)</td>
<td>40</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michaelson et al. (2012)</td>
<td>20</td>
<td>20</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shi et al. (2013)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jiang et al. (2013)</td>
<td>32</td>
<td>32</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yuen et al. (2015)</td>
<td>85</td>
<td>85</td>
<td>170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nemirovsky et al. (2015)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
frequency). They observed a significant twofold increase in complete KOs in patients with ASD compared to controls. They estimated that such complete KO mutations could account for 3% of the patients with ASD. For the X chromosome, there was a significant 1.5-fold increase in complete KO in affected males compared to unaffected males that could account for 2% of males with ASD (Lim et al. 2013). The same year, Yu et al. (2013) analyzed 104 consanguineous families including 79 families with a single child with ASD (simplex families) and 25 families with more than one affected individual (multiplex families) collected by the Homozygosity Mapping Collaborative for Autism (HMCA). They identified biallelic mutations in AMT, PEX7, SYNE1, VPS13B, PAH, and POMGNT1. Finally, a very recent study by Krumm et al. (2015) ascertained the relative impact of inherited and de novo variants (CNVs or SNVs) on ASD risk in 2377 families. Inherited truncating variants were enriched in probands (for SNV odds ratio = 1.14, \(P = 0.0002\); for CNV odds ratio = 1.23, \(P = 0.001\)) in comparison to unaffected siblings (Krumm et al. 2015). Interestingly, they also observed a significant maternal transmission bias of inherited LGD to sons. New ASD-risk genes were also identified such as RIMS1, CUL7 and LZTR1.

To date, few whole genome sequencing studies have been published for ASD (Table 1). Michaelson et al. (2012) analyzed 40 WGS of monozygotic twins concordant for ASD and their parents. They proposed that ASD-risk genes could be hot spots of mutation in the genome and confirmed the association between ASD and de novo mutations in GPR98, KIRREL3 and TCF4. Shi et al. (2013) analyzed a large pedigree with two sons affected with ASD and six unaffected siblings, focusing on inherited mutations. They identified ANK3 as the most likely candidate gene. In 2015, Yuen et al. analyzed 85 families with two children affected with ASD. This study represents the largest published WGS data set in ASD. They identified 46 ASD-relevant mutations present in 36 of 85 (42.4%) families. Only 16 ASD-relevant mutations of 46 (35%) identified were de novo. Very interestingly, for more than half of the families (69.4%; 25 of 36), the two affected siblings did not share the same rare penetrant ASD risk variant(s).

Whole genome sequencing is also very efficient to identify mutations in regions of the human genome that are wrongly annotated and in regions that are highly GC rich. For example, mutations on the SHANK3 gene were rarely identified using whole exome sequencing, given its high GC content (Leblond et al. 2014). In contrast, whole genome sequencing could successfully identify SHANK3 mutations (Nemirovsky et al. 2015; Yuen et al. 2015).

The Common Variants in ASD

In the general population, one individual carries on average three million genetic variants in comparison to the reference human genome sequence (Xue et al. 2012; Fu et al. 2013; Genome of the Netherlands and Genome of the Netherlands 2014). The vast majority of the variants (>95%) are the so-called common variants shared
with more than 5 % of the human population (Xue et al. 2012; Fu et al. 2013; Genome of the Netherlands and Genome of the Netherlands 2014). While there is not a clear border between common and rare variants, it is nevertheless interesting to estimate the role of the genetic variants found in the general population in the susceptibility to ASD.

Using quantitative genetics, Klei et al. (2012) estimated that common variants were contributing to a high proportion of the liability of ASD: 40 % in simplex families and 60 % in multiplex families. In 2014, a study by Gaugler et al. used the same methodology (Yang et al. 2011) and provided an estimation of the heritability (52.4 %) that is almost exclusively due to common variation, leaving only 2.6 % of the liability to the rare variants. The contribution of common variants is therefore important, but unfortunately the causative SNPs still remain unknown since they are numerous (>1000) and each is associated with a low risk. To date, the largest genome wide association studies (GWAS) performed on <5000 families with ASD were underpowered to identify a single SNP with genome wide significance (Anney et al. 2012; Cross-Disorder Group of the Psychiatric Genomics 2013).

The recruitment of larger cohorts of patients with dimensional phenotypes is therefore warranted to better ascertain the heritability of ASD and to identify the genetic variants, which explain most of the genetic variance.

The Genetic Architecture of ASD

Based on the results obtained from epidemiological and molecular studies, it is now accepted that the genetic susceptibility to ASD can be different from one individual to another with a combination of rare deleterious variants (R) and a myriad of low-risk alleles [also defined as the genetic background (B)]. Most of the inherited part of ASD seems to be due to common variants observed in the general population, with only a small contribution from rare variants (Fig. 3). Importantly, while the \textit{de novo} mutations are considered per se as genetic factors, they do not contribute to the heritability since they are only present in the patient (with the relatively rare exception of germinal mosaicismisms present in one of the parental germlines and transmitted to multiple children). These \textit{de novo} events could therefore be considered as “environmental causes” of ASD but acting on the DNA molecule. It is currently estimated that more than 500–1000 genes could account for these “monogenic forms” of ASD (Iossifov et al. 2012; Sanders et al. 2012), confirming the high degree of genetic heterogeneity.

The interplay between the rare or \textit{de novo} variants R and the background B will also influence the phenotypic diversity observed in the patients carrying rare deleterious mutations. In some individuals, a genetic background B will be able to buffer or compensate the impact of the rare genetic variations R. In contrast, in some individuals, the buffering capacity of B will not be sufficient to compensate the impact of R and they will develop ASD (Rutherford 2000; Hartman et al. 2001). In the R’n’B model, ASD can be regarded as a collection of many genetic forms of
“autisms,” each with a different etiology ranging from monogenic to polygenic models.

The presence of multiple hits of rare CNVs, SNVs or indels in a single individual also illustrates the complexity of the genetic landscape of ASD (Girirajan et al. 2010, 2012; Leblond et al. 2012). In addition, the analysis of the whole genome sequence of multiplex families also indicates that clinically relevant mutations can be different from one affected sib to another even in a single family (Yuen et al. 2015). It is therefore still difficult to ascertain robust genotype-phenotype relationships based on our current knowledge.

Fortunately, although the ASD-risk genes are numerous, they seem to converge in a limited number of biological pathways that are currently scrutinized by many researchers.

The Biological Pathways Associated with ASD

Unbiased pathway analyses indicated that ASD-risk genes seem to be enriched in groups of proteins with specific functions (Voineagu et al. 2011; De Rubeis et al. 2014; Ronemus et al. 2014; Uddin et al. 2014; Hormozdari et al. 2015). Pinto et al. (2014) analyzed the burden of CNVs in 2446 individuals with ASD and

---

**Fig. 3** Relative contribution of the genetics and environment in ASD. Based on twin and familial studies, it is estimated that the genetic and environmental contributions to ASD are approximately 50/50%. Most of the heritable part seems to be due to common variants observed in the general population, with a small contribution of rare variants. Importantly, the *de novo* mutations are genetic causes of ASD but do not contribute to the heritability since there are only present in the patient. These *de novo* events are therefore considered to be “environmental causes” of ASD, but acting on the DNA molecule (Adapted from Huguet and Bourgeron 2016).
2640 controls and found enrichment in genes coding post-synaptic density proteins and FMRP targets. Ronemus et al. (2014) reviewed the results of four whole exome sequencing studies and showed an enrichment of mutated genes in chromatin modifier genes ($P = 4 \times 10^{-6}$) and FMRP targets ($P = 7 \times 10^{-6}$). Protein-protein interactions (PPI) analyses of the genes carrying LGD mutations also showed enrichment in proteins involved in neuronal development and axon guidance, signaling pathways and chromatin and transcription regulation. De Rubeis et al. (2014) also used the PPI network and showed enrichment in clusters of proteins involved in the cell junction TGF beta pathway, cell communication and synaptic transmission, neurodegeneration and transcriptional regulation.

In parallel to the genetic studies, several transcriptomic analyses were performed using post-mortem brain of individuals with ASD (Voineagu et al. 2011; Gupta et al. 2014). Several genes were differentially expressed or correlated between brain regions. Two network modules were identified. The first module was related to interneurons and to genes involved in synaptic function (downregulated in brains from ASD patients compared to controls). The second module was related to immunity and microglia activation (upregulated in brains from ASD patients compared to controls).

Based on these results, neurobiological studies using cellular and animal models have been performed to identify the main mechanisms leading to ASD. Remarkably, several studies showed that neuronal activity seems to regulate the function of many of the ASD-risk genes, leading to the hypothesis that abnormal synaptic plasticity and failure of neuronal/synaptic homeostasis could play a key role in the susceptibility to ASD (Belmonte and Bourgeron 2006; Auerbach et al. 2011; Toro et al. 2010). Here, we will only depict four main biological pathways associated with ASD: chromatin remodeling, protein synthesis, protein degradation, and synaptic function (Fig. 4). In parallel, several biochemical studies have indicated a dysfunction in the serotonin-NAS-melatonin pathway.

**Chromatin Remodeling** Mutations in genes encoding key regulators of chromatin remodeling and gene transcription (e.g., MECP2, MEF2C, HDAC4, CHD8 and CTNNB1) have been reported in individuals with ASD (Fig. 4). Remarkably, a subset of these genes is regulated by neuronal activity and influences neuronal connectivity and synaptic plasticity (Cohen et al. 2011; Sando et al. 2012; Ebert et al. 2013).

**Protein Synthesis** The level of synaptic proteins can be influenced by neuronal activity through global and local synaptic mRNA translation (Ma and Blenis 2009). Several genes involved in such activity-driven regulation of synaptic proteins have been found to be mutated in individuals with ASD (Kelleher and Bear 2008). For example, the mTOR pathway controls global mRNA translation and its deregulation causes diseases associated with increased cell proliferation and loss of autophagy, including cancer (Ma and Blenis 2009), but also increases the risk for ASD. Remarkably, mutations in the repressor of the mTOR pathway such as NF1, PTEN and SynGAP1 cause an increase of translation in neurons and at the synapse (Auerbach et al. 2011). Mutations of the FMRP–EIF4E–CYFIP1 complex cause the
Fig. 4 Examples of the biological pathways associated with ASD. The ASD-risk genes code for proteins involved in chromatin remodeling, transcription, protein synthesis and degradation, cytoskeleton dynamics, and synaptic functions. Proteins associated with ASD are circled in red (Adapted from Huguet and Bourgeron 2016)
fragile X syndrome and increase the risk of ASD (Budimirovic and Kaufmann 2011). This protein complex controls local translation of mRNA at the synapse and acts downstream of the Ras-ERK signaling pathway. This complex regulates the translation of more than 1000 specific genes, many of which are ASD risk genes (De Rubeis et al. 2013; Fernandez et al. 2013; Gkogkas et al. 2013; Santini et al. 2013). An alteration of this FMRP–EIF4E–CYFIP1 complex should therefore create an imbalance in the level of many synaptic proteins that are associated with ASD.

Protein Degradation The Ubiquitin-Proteasome System (UPS) is central for the degradation of the proteins and, consequently, for the regulation of synapse composition, assembly and elimination (Mabb and Ehlers 2010). The UBE3A gene encodes a ubiquitin ligase, is mutated in patients with Angelman syndrome and is duplicated on the maternal chromosome 15q11 in individuals with ASD. Neuronal activity increases UBE3A transcription through the MEF2 complex and regulates excitatory synapse development by controlling the degradation of ARC, a synaptic protein that decreases long-term potentiation by promoting the internalization of AMPA receptors (Greer et al. 2010).

Synaptic Functions Many proteins encoded by ASD-risk genes participate in different aspects of neuronal connectivity, such as glutamatergic (e.g., GRIN2B), GABAergic (e.g., GABRA3 and GABRB3) and glycineric (e.g., GLRA2) neurotransmission, neuritogenesis (e.g., CNTN), the establishment of synaptic identity (e.g., cadherins and protocadherins), neuronal conduction (CNTNAP2) and permeability to ions (CACNA1, CACNA2D3 and SCN1A). Some of these proteins are directly involved in activity-driven synapse formation, such as the neurexins (NRXNs) and neuroligins (NLGNs). Some are scaffolding proteins involved in the positioning of cell-adhesion molecules and neurotransmitter receptors at the synapse (Sheng and Kim 2011; Choquet and Triller 2013). For example, deletions, duplications and coding mutations in the three SHANK genes (SHANK1, SHANK2 and SHANK3) have been recurrently reported in individuals with ASD (Leblond et al. 2014). SHANK proteins assemble into large molecular platforms in interaction with glutamate receptors and actin-associated proteins (Grabrucker et al. 2011). In vitro, SHANK3 mutations identified in individuals with ASD reduce actin accumulation in spines affecting the development and morphology of dendrites as well as the axonal growth cone motility (Durand et al. 2012). In mice, mutations in SHANK3 decrease spine density in the hippocampus but also increase dendritic arborizations in striatal neurons (Peca et al. 2011). Mice mutated in SHANK present with behavior resembling autistic features in humans. Shank1 knockout mice display increased anxiety, decreased vocal communication, decreased locomotion and, remarkably, enhanced working memory, but decreased long-term memory (Hung et al. 2008; Silverman et al. 2011; Wohr et al. 2011). Shank2 knockout mice present hyperactivity, increased anxiety, repetitive grooming, and abnormalities in vocal and social behaviors (Schmeisser et al. 2012; Won et al. 2012). Shank3 knockout mice show self-injurious repetitive
grooming, and deficits in social interaction and communication (Bozdagi et al. 2010; Peca et al. 2011; Wang et al. 2011; Yang et al. 2012).

The Serotonin-NAS-Melatonin Pathway

In parallel to the genetic investigations pointing to the biological pathways described above, several biological abnormalities have been reported in individuals with ASD, including neurochemical, immunological, endocrine or metabolic traits (Lam et al. 2006; Rossignol and Frye 2012), which may provide insights into the pathophysiology of ASD. Among these, elevated blood serotonin is one of the most replicated findings (Pagan et al. 2012) (Fig. 5). A deficit in melatonin (which chemically derives from serotonin) in the blood or urine of individuals with ASD has also been described in several studies (Tordjman et al. 2005; Melke et al. 2008) and is associated with increased peripheral N-acetylserotonin (NAS), the intermediate metabolite between serotonin and melatonin. Several mutations affecting this pathway were identified but the mechanisms of these biochemical impairments remain mostly unexplained. Melatonin is a neurohormone mainly synthesized in

![Diagram of the serotonin-NAS-melatonin pathway in ASD](image)

**Fig. 5** The serotonin-NAS-melatonin pathway in ASD. (a) The serotonin-NAS-melatonin synthesis pathway consists of two enzymatic steps involving first the N-acetylation of serotonin to N-acetylserotonin (NAS) by the rate-limiting enzyme AANAT and the methylation of NAS by the ASMT (also known as HIOMT). Different alterations such as higher serotonin or low melatonin levels were observed in the blood of patients with ASD. The enzymes are represented in gray and metabolites are in white. Alterations of the biochemical parameters are shown with red arrows. (b) A schematic view of the pineal gland with the pinealocytes that contain the “melatoninosome.” This complex includes at least four proteins: AANAT, ASMT, 14-3-3 and S-antigen. The immunofluorescence confocal image of AANAT (green) and 14-3-3ε (red) in the pinealocytes is adapted from Maronde et al. (2011). The structure of the 14-3-3ζ homodimer binding to AANAT is adapted from Obsil et al. (2001). (c) Main sources of serotonin (blue) and melatonin (green) and the symptoms or comorbidities of ASD relevant to alterations in serotonin and melatonin levels observed in ASD (Adapted from Pagan et al. 2012)
the pineal gland during the night. It is a biological signal of light/dark cycles and is considered to be a major circadian synchronizer. It is also a modulator of metabolism, immunity, reproduction and digestive function. Furthermore, it displays antioxidant and neuroprotective properties and can directly modulate neuronal networks (Bourgeron 2007). Melatonin appears as a therapeutic target of the frequently reported sleep disorders associated with ASD (Andersen et al. 2008; Wright et al. 2011; Malow et al. 2012). NAS displays intrinsic biological properties: it is an agonist of the TrkB receptor and may thus share the neurotrophic properties of brain-derived neurotrophic factor (BDNF), the canonical TrkB ligand (Jang et al. 2010; Sompol et al. 2011). Serotonin conversion into melatonin involves two sequential enzymatic steps: N-acetylation of serotonin into NAS by arylalkylamine N-acetyltransferase (AANAT, EC: 2.3.1.87) followed by methylation by acetylserotonin O-methyltransferase (ASMT, also called hydroxyindole O-methyltransferase HIOMT, EC: 2.1.1.4) (Fig. 5a). We previously showed that deleterious mutations of the ASMT gene could disrupt melatonin synthesis in a subset of patients with ASD. Nevertheless, the frequency of such a deleterious mutation is too low (2 % of the cases) to explain the relatively high frequency of melatonin deficit in ASD (>50 % of the patients, taking as a threshold the fifth percentile of the controls). More recently, we observed a low level of the 14-3-3 proteins both in the blood platelet and pineal gland of patients with ASD (Pagan et al. 2014). These ubiquitous chaperone proteins are known to form a protein complex, the ‘melatoninosome,’ involving AANAT and ASMT in pinealocytes (Obsil et al. 2001; Maronde et al. 2011). This interaction between 14-3-3 and AANAT and/or ASMT might be essential for the production of melatonin and an adequate balance of the serotonin-NAS-melatonin pathway. Indeed, a low level of 14-3-3 could eventually lead to a deficit in enzyme activity, contributing to the global disruption of the serotonin-NAS-melatonin pathway observed in ASD. Studies investigating the regulation of the complex 14-3-3/ASMT/AANAT in ASD and controls are in progress.

Perspectives

In the last 30 years, very significant progress has been made in the genetics of ASD. We now have a better knowledge on the genetic architecture of this heterogeneous syndrome and some of the biological pathways have been investigated using different approaches such as cellular and animal models. There are, however, many aspects of the genetics of ASD that remain largely unknown.

The first challenge concerns the role of the common variants. These variants are most likely playing a key role in the susceptibility to ASD and in the severity of the symptoms. But, because the impact of each single SNP is very low, it is currently impossible to identify the risk alleles using conventional GWAS. In human quantitative traits such as height, neuroanatomical diversity or intellectual quotient, very
large cohorts of many thousands of individuals are necessary to identify the main causative SNPs (Toro et al. 2014; Yang et al. 2010; Deary et al. 2012).

The second challenge concerns the stratification of the patients and the role of the ASD-risk genes during brain development/function. Based on our current knowledge, the genetic architecture of ASD seems to be different from one individual to another, with possibly contrasting impact on when and where neuronal connectivity could be atypical compared to the general population. For example, in animal models, several mutations lead to higher connectivity whereas other mutations alter synaptic density. It is therefore crucial to increase our knowledge from a basic research perspective about the biological roles of the ASD-risk genes and their partners.

Finally, while we all agree that biological research is necessary to improve the quality of life of the patients and their families (for example, to alleviate the comorbidities associated with ASD like sleep and gastrointestinal problems), progress should also be made toward better recognition and inclusion of people with neuropsychiatric conditions in our societies (no mind left behind). Hopefully, increasing knowledge in genetics, neurology and psychology should allow for better diagnosis, care for and integration of individuals with autism.

Acknowledgments We thank Roberto Toro for critical reading of the manuscript. This work was funded by the Institut Pasteur, the Bettencourt-Schueller foundation, Centre National de la Recherche Scientifique, University Paris Diderot, Agence Nationale de la Recherche (SynDiv-ASD), the Conny-Maeva Charitable Foundation, the Cognacq Jay Foundation, the Orange Foundation, and the Fondamental Foundation.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution-Noncommercial 2.5 License (http://creativecommons.org/licenses/by-nc/2.5/) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

The images or other third party material in this chapter are included in the work’s Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work’s Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

References


dysfunction in autism spectrum disorder and intellectual disability (in press)
Rev Genomics Hum Genet 14:191–213
Hung AY, Futai K, Sala C, Valttschanoff JG, Ryu J, Woodworth MA, Kidd FL, Sung CC,
synaptic transmission, but enhanced spatial learning in mice lacking Shank1. J Neurosci
28:1697–1708
Leotta A, Kendall J, Grabowska E, Ma B, Marks S, Rodgers L, Stepansky A, Troge J,
Fulton RS, Magrini VJ, Ye K, Darnell JC, Darnell RB, Mardis ER, Wilson RK, Schatz MC,
Iossifov I, O’Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon
KT, Vives L, Patterson KE, Smith JD, Paeper B, Nickerson DA, Dea J, Dong S, Gonzalez LE,
Mandell JD, Mane SM, Murtha MT, Sullivan CA, Walker MF, Waqar Z, Wei L, Willsay AJ,
MC, Ye K, McCombie WR, Shendure J, Eichler EE, State MW, Wigler M (2014) The
contribution of de novo coding mutations to autism spectrum disorder. Nature 515:216–221
Jacquemont ML, Sanlaville D, Redon R, Raoul O, Cormier-Daire V, Lyonnet S, Amiel J, Le
Merrer M, Heron D, de Blois MC, Prieur M, Vekemans M, Carter NP, Munich A, Colleaux L,
Philippe A (2006) Array-based comparative genomic hybridisation identifies high frequency of
cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. J
Med Genet 43:843–849
Jamain S, Quach H, Betancur C, Rastam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B,
neuroligins NLGN3 and NLGN4 are associated with autism. Nat Genet 34:27–29
activates TrkB receptor in a circadian rhythm. Proc Natl Acad Sci USA 107:3876–3881
Nalpathamkalam T, Thiruvathkinapuranam B, Thompson A, Uddin M, Walker S, Luo J,
Anagnostou E, Zwaigenbaum L, Ring RH, Wang J, Lajonchere C, Wang J, Shih A,
genetic variants in autism spectrum disorder by whole-genome sequencing. Am J Hum Genet
93:249–263
Klei L, Sanders SJ, Murtha MT, Hus V, Lowe JK, Willsey AJ, Moreno-De-Luca D, Yu TW,
Fombonne E, Geschwind D, Grice DE, Ledbetter DH, Lord C, Mane SM, Martin CL, Martin
DM, Morrow EM, Walsh CA, Melhem NM, Chaste P, Sutcliffe JS, State MW, Cook EH Jr,
Roeder K, Devlin B (2012) Common genetic variants, acting additively, are a major source of
risk for autism. Mol Autism 3:9
Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA,
Sigrudsson A, Jonasdottir A, Wong WS, Sigrudsson G, Walters GB, Steinberg S,
Hegason H, Thorleifsson G, Gudbjartsson DF, Helgason A, Magnusson OT,
Thorsteinsdottir U, Stefansson K (2012) Rate of de novo mutations and the importance of
father’s age to disease risk. Nature 488:471–475
and molecular neuroscience. Trends Neurosci 37:95–105


to behaviour management strategies: a randomised controlled crossover trial. J Autism Dev Disord 41:175–184