Genetic Studies of Autism and Autistic-Like Traits

Lina Jonsson

2015

UNIVERSITY OF GOTHENBURG

Department of Pharmacology
Institute of Neuroscience and Physiology,
Sahlgrenska Academy at the University of Gothenburg, Sweden

lina.jonsson@neuro.gu.se

Printed by Ineko AB, Gothenburg, Sweden

ISBN: 978-91-628-9269-2
hdl.handle.net/2077/37522
ISBN: 978-91-628-9270-8 (electronic version)

Genetic Studies of Autism and Autistic-Like Traits

Lina Jonsson
Department of Pharmacology, Institute of Neuroscience and Physiology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden

ABSTRACT

Autism spectrum disorder (ASD) is characterized by impairment in social interaction, language impairment and repetitive behavior with varying degrees of severity. ASD represents the lower end on a continuously distributed measure of autistic-like traits (ALTs). Although a strong genetic component has repeatedly been identified in ASD, the genetic cause of ASD is still unknown for the majority of ASD cases.

One of the main interests in this thesis is the neurobiology of melatonin, this interest is based on findings indicating lower levels of melatonin in children with ASD. In our investigations of rare mutations in melatonin related genes in subjects with ASD, we identified a previously reported mutation that has been shown to decrease the activity of one of the enzymes involved in the melatonin synthesis: the acetylserotonin O-methyltransferase (ASMT) (paper I). In the analysis of five common variations in the *ASMT* gene in relation to ALTs in the general population we found association between a single nucleotide polymorphism and social interaction impairment in girls (paper II).

To broaden the analysis of genetic influences on ALTs, we have performed association analyses between ALTs in the general population and common variation in genes previously found to be associated with ASD (*RELN*, *CNTNAP2*, *SHANK3* and *CDH9/10* region) (paper III). Although these regions have previously been suggested to be strong ASD candidate regions, our results do not suggest a major influence of the investigated common variations on ALTs.

In the final paper, rare inherited genetic variations were investigated in a large family with autism and language disorders. In this study, we used several techniques, including whole exome sequencing and copy number variation analysis (paper IV). In the family, several rare genetic variations which may partly explain the genetic etiology for autism in this family were identified. We performed functional analyses for a mutation identified in the *CYP11A1* gene, indicating a gain of function mutation. The *CYP11A1* gene encodes the first enzyme in the steroid hormone biosynthesis, thus our results may be in line with previous findings that have shown an elevated prenatal steroidogenic activity in ASD.

In conclusion, we have identified both common and rare genetic variation that may increase the genetic susceptibility for ASD. Our analyses have highlighted the importance of taking both rare and common genetic susceptibility factors, as well as different symptoms of the disorders, into account when elucidating the complex inheritance of ASDs.

Keywords: ASD, ALT, Genetics, Melatonin
ISBN: 978-91-628-9269-2

LIST OF PAPERS

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. **Lina Jonsson**, Elin Ljunggren, Anna Bremer, Christin Pedersen, Mikael Landén, Kent Thuresson, MaiBritt Giacobini and Jonas Melke. 2010. Mutation screening of melatonin-related genes in patients with autism spectrum disorders. *BMC Med. Genomics* 3, 10.

II. **Lina Jonsson**, Henrik Anckarsäter, Anna Zettergren, Lars Westberg, Hasse Walum, Sebastian Lundström, Henrik Larsson, Paul Lichtenstein, Jonas Melke. 2014. Association between *ASMT* and autistic-like traits in children from a Swedish nationwide cohort. *Psychiatr. Genet.* 24(1):21-7

III. **Lina Jonsson**, Anna Zettergren, Erik Pettersson, Daniel Hovey, Henrik Anckarsäter, Lars Westberg, Paul Lichtenstein, Sebastian Lundström and Jonas Melke. 2014. Association study between autistic-like traits and polymorphisms in the autism candidate regions *RELN*, *CNTNAP2*, *SHANK3*, and *CDH9/10*. *Molecular Autism*, 5:55

IV. **Lina Jonsson**, Carmela Miniscalco, Mats Johnson, Christopher Gillberg, Tommy Martinsson, Jonas Melke. Screening for rare inherited variation in a multiplex family with autism and language disorders. *Manuscript*.

CONTENT

ABBREVIATIONS

AA-NAT	Arylalkylamine N-Acetyltransferase
ALT	Autistic-Like Trait
ASD	Autism Spectrum Disorder
ASMT	Acetylserotonin O-Methyltransferase
A-TAC	Autism-Tics, AD/HD And Other Comorbidities Inventory
CATSS	The Child And Adolescent Twin Study In Sweden
CDH10	Cadherin 10
CDH9	Cadherin 9
CNTNAP2	Contactin Associated Protein-2
CNV	Copy Number Variation
CYP11A1	Cytochrome P450 11A1
DSM	Diagnostic and Statistical Manual Of Mental Disorders
GPR-50	G Protein-Coupled Receptor 50
GWAS	Genome Wide Association Study
MTNR1A	Melatonin Receptor 1A
MTNR1B	Melatonin Receptor 1B
NDP	Neurodevelopmental Problem
P450scc	Cholesterol Side Chain Cleavage Enzyme
PDD	Pervasive Developmental Disorder
RELN	Reelin
SHANK3	SH3 and Multiple Ankyrin Repeat Domains Protein 3
SNP	Single Nucleotide Polymorphism
WES	Whole Exome Sequencing

INTRODUCTION

AUTISM

Characteristics and prevalence

Autism spectrum disorder (ASD), hereafter referred to as autism, is a neuro-developmental disorder characterized by impairments in social interaction, language impairments and restricted and repetitive behavior and interests[1]. The symptoms of autism most often emerge during early childhood[2]. Social interaction impairments refers to non-verbal communication, such as making eye-contact, initiating and responding to smiling or initiating and responding to physical contact by greetings or waving good-bye. These non-verbal communications can be identified in children very early in life, before language has developed. The social communication domain refers to the ability to communicate, or converse, both verbally and non-verbally with another person. One of the earliest forms of non-verbal communication during development is joint attention, which can be a mutual attention towards an object. The communication through spoken language is often delayed in autism[3]. When children with autism learn to talk, their language can be characterized by stereotypic speech that in some cases may involve echolalia and unusual intonations[4]. Autism is a behavioral based diagnosis and several diagnostic instruments are used clinically (*see below*).

For a long time, autism was regarded as a very rare disorder with a prevalence of about 1-5 cases per 10,000 subjects in the population[5]. However, the prevalence has increased during the last decades with current prevalence estimates of approximately 1% in the population[6-8]. Autism is about four times more prevalent in boys compared to girls[7].

Autism has a high comorbidity with other neurodevelopmental and neuropsychiatric disorders[9,10]. Disorders often co-existing with autism are for example attention deficit hyperactivity disorder (ADHD), tic disorder, oppositional defiant disorder and bipolar disorder.

Table 1. DSM-IV[1] Criteria for Autistic Disorder

(I) A total of six (or more) items from (A), (B), and (C), with at least two from (A), and one each from (B) and (C)

A. Qualitative impairment in social interaction, as manifested by at least two of the following:

1. Marked impairments in the use of multiple nonverbal behaviors such as eye-to-eye gaze, facial expression, body posture, and gestures to regulate social interaction
2. Failure to develop peer relationships appropriate to developmental level
3. A lack of spontaneous seeking to share enjoyment, interests, or achievements with other people
4. Lack of social or emotional reciprocity

B. Qualitative impairments in communication as manifested by at least one of the following:

1. Delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime)
2. In individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others
3. Stereotyped and repetitive use of language or idiosyncratic language
4. Lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level

C. Restricted repetitive and stereotyped patterns of behavior, interests and activities, as manifested by at least two of the following:

1. Encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus
2. Apparently inflexible adherence to specific, nonfunctional routines or rituals
3. Stereotyped and repetitive motor mannerisms (*e.g* hand or finger flapping or twisting, or complex whole-body movements)
4. Persistent preoccupation with parts of objects

(II) Delays or abnormal functioning in at least one of the following areas, with onset prior to age 3 years:

A. Social interaction. **B.** Language as used in social communication. **C.** Symbolic or imaginative play

(III) The disturbance is not better accounted for by Rett's Disorder or Childhood Disintegrative Disorder

DSM-IV=Diagnostic and Statistical Manual Of Mental Disorders 4th edition

The historical development of autism diagnostics

In 1943, Leo Kanner first described autism, which he would call early infantile autism[11]. The eleven children included in Kanner's report were described as children without a predisposition to be social and with significant problems when faced with change in the non-social world, which he termed "resistance to change" or "insistency on sameness". Kanner also noted that the majority of the children had language problems. In 1944, Hans Asperger described the first cases of Asperger's syndrome. The children described in his report displayed several similarities with infantile autism. However, compared with the infantile autism, individuals with Asperger's syndrome did not have significant delays or difficulties with language. Although Asperger defined the syndrome in 1944, it did not gain much attention until it was given the name Asperger's syndrome in 1981[12]. Since the description of autism by Kanner, several diagnostic manuals have been developed, such as the Diagnostic and Statistical Manual of Mental Disorders (DSM), Autism Diagnostic Interview-Revised (ADI-R) and Autism

Table 2. DSM-5[13] Criteria for Autism Spectrum Disorder

A. Persistent deficits in social communication and social interaction across multiple contexts, as manifested by the following, currently or by history:

- **1.** Deficits in social-emotional reciprocity
- **2.** Deficits in nonverbal communicative behaviors used for social interaction
- **3.** Deficits in developing, maintaining, and understanding relationships

B. Restricted, repetitive patterns of behavior, interests, or activities, as manifested by at least two of the following, currently or by history:

- **1.** Stereotyped or repetitive motor movements, use of objects, or speech
- **2.** Insistence on sameness, inflexible adherence to routines, or ritualized patterns or verbal nonverbal behavior
- **3.** Highly restricted, fixated interests that are abnormal in intensity or focus
- **4.** Hyper- or hyporeactivity to sensory input or unusual interests in sensory aspects of the environment

C. Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities, or may be masked by learned strategies in later life).

D. Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning.

E. These disturbances are not better explained by intellectual disability (intellectual developmental disorder) or global developmental delay. Intellectual disability and autism spectrum disorder frequently co-occur; to make comorbid diagnoses of autism spectrum disorder and intellectual disability, social communication should be below that expected for general developmental level.

DSM-5=Diagnostic and Statistical Manual Of Mental Disorders 5th edition

Diagnostic Observation Schedule (ADOS). All diagnostic manuals are behavior-based and involve evaluation of the level of the core impairments of autism.

The DSM offers standard criteria for the classification of mental disorders and is used by mental health professionals, researchers and drug regulation agencies. When infantile autism was included as one of the disorders under the term pervasive developmental disorder (PDD) in the DSM-III in 1980, the diagnosis became officially recognized. When the DSM-IV was released in 1994, infantile autism was renamed to autistic disorder (table 1) and the diagnoses included under the umbrella term PDD were autistic disorder, Asperger's syndrome, pervasive developmental disorders not otherwise specified (PDD-NOS), childhood disintegrative disorder and Rett's disorder. In DSM-IV, the inclusion criteria were broadened compared to DSM-III. In 2013, the fifth edition of the diagnostic manual was released (DSM-5)[13]. In this version of the manual, the diagnoses previously assembled under the umbrella term PDD have been merged into the single diagnosis of Autism Spectrum Disorder (table 2). Thus, DSM-5 canceled Asperger's disorder as a separate diagnosis and homogenized it under autism spectrum disorder, with severity measures within the broader diagnosis[14]. Another major change was that the triad of symptoms for ASD diagnosis were merged into two domains, the social interaction and communication domains were merged into a single domain. It has previously been shown to be difficult to separate these two domains and some symptoms could be found within both domains[14]. In addition, when these two domains were combined, the language specific criterion in the

communication domain such as "delay in, or total lack of, the development of spoken language" was removed from the ASD criteria in DSM-5. Instead a new diagnosis of "social (pragmatic) communication disorders" was defined to identify individuals whose symptoms could better be explained by this disorder than the ASD criteria. The restricted and repetitive behavior domain remained as the second domain for ASD in DSM-5.

With the new DSM-5, there was a conceptual change of the diagnostics. Whereas the ASD diagnostics previously regarded the included disorders under PDD as discrete disorders, within DSM-5 these discrete disorders have been replaced with one single ASD diagnosis with varying degrees of severity. This concept was first identified decades ago[15,16], suggesting a continuum of symptoms that are also present in the general population.

Autistic-like traits

The view of autism as a spectrum forming a continuum[15,17] means that individuals who do not meet the diagnostic criteria for ASD may display milder phenotypes related to the autism characteristics; these traits are referred to as autistic-like traits (ALTs). Initially described as milder manifestations of psychopathology, ALTs now represent the boundary between autism and normality. Both a phenotypic and genetic overlap between the lower and upper extreme ends on the continuum has been shown in population based studies[18,19]. In family based studies, it is recognized that there is an aggregation of ALTs in close relatives to subjects with autism[20]. The ALTs can be categorized within the same triad as the core characteristics for autism: social interaction, communication and restricted and repetitive behavior. The three domains have been shown to be partly separate symptoms with small overlaps between the symptoms. Since individuals with ALTs do not always express problems on all of the three ALT domains, the correlation between these symptoms has been shown to be lower for ALTs as compared to when these symptoms are measured in autism[21].

A few instruments for ALT measurements have been developed such as the Autism-Tics, AD/HD and other Comorbidities inventory (A-TAC)[22] and the Autism-Spectrum Quotient (AQ)[23]. One of the main differences between these measures is whether they are designed to capture a broad normal variation of ALTs or to capture ALTs more closely related to the diagnostic criteria used in for example the DSM. The A-TAC is based on DSM-IV-criteria while, for example, the AQ is designed to capture a broader normal variation of ALTs in the general population.

What causes autism?

At present, no single cause for autism has been identified. A large impact of genetic factors has been identified, which will be discussed in more detail in a separate chapter in this thesis. The importance of a heritable component was suspected early on by Leo Kanner, however the view of the importance of genetics has varied over time. During the decades after the study by Kanner, mothers to children with autism were blamed for being cold and many people believed their coldness was the cause of their children's diagnosis. This view prevailed until the 1960's when the biological basis for autism was reviewed[24]. The importance of genetic factors was strengthened by the first twin study of autism[25] and thereby gained broad recognition. At present, a few environmental risk factors have been identified, such as prenatal exposure to substances such as valproic acid[26], thalidomide and misoprostol[27] as well as maternal infections, *e.g.* rubella[28].

The prevalence of autism has increased drastically during the last decades. Several explanations for this increase have been suggested, including conceptual changes of autism, change in diagnostic criteria over time and environmental factors. As mentioned in a previous section, the diagnostic manuals have changed over time, which indeed has affected the diagnostics of autism. In addition, when autism was first identified, it was regarded as a discrete disorder; however, it is now regarded as continuum under the term of ASD. The awareness of autism in the community and the availability of diagnostic services in many countries has probably also contributed to increased prevalence[29]. In addition to a few environmental factors that have been identified, gene-environment interactions can also occur. One example of a gene-environment interaction has been shown for Parkinson's disease, where the risk for disease due to exposure to an environmental toxin increased drastically for people carrying a genetic susceptibility variant[30]. To study the possibility of interactions between environmental and genetic factors has been difficult due to small samples sizes and the investigation of a limited amount of polymorphisms. However, the new genome wide approaches have been suggested to offer new strategies to analyze this interaction in several neuropsychiatric disorders[31].

Treatments for autism

There is no cure for autism and the current treatments available are used to reduce symptoms associated with autism. Several attempts have been made to try to find pharmacological treatments with only limited success. Autism associated symptoms such as aggression and stereotypical behaviors can be ameliorated by pharmacological treatments with neuroleptics, such as risperidone and aripiprazole; however these

drugs have not been shown to suppress the core symptoms of autism[32,33]. There are some promising studies that investigate therapies based on genetic studies in autism that occur in known genetic syndromes, such as tuberous sclerosis and fragile X syndrome. Functional analyses of the genes involved in these genetic syndromes have led to the identification of possible pharmacological targets. One example is a drug called rapamycin that has been shown to rescue behavioral deficits in animal models for the genetic syndrome tuberous sclerosis[34]. Rapamycin targets the so called mTOR signaling pathway, which is important for several neurological phenotypes. Several promising results for drugs targeting the mTOR pathway are being evaluated for possible future treatments of autism and related phenotypes[35].

Another example of symptoms occurring in autism is sleep disorders. The use of oral treatment of melatonin in these children has been shown to be beneficial[36]. Studies of melatonin will be discussed in a separate chapter of this thesis (*see "Melatonin"*).

DNA AND GENETIC VARIATION

DNA and genes

The double helix structure of deoxyribonucleic acid (DNA) was first described by Watson and Crick in 1953[37]. The DNA helix structure consists of four different nucleotides or so called bases: adenine (A), cytosine (C), guanine (G) and thymine (T). These nucleotides bind to each other forming the ladder-like DNA double-helix structure. The human genome contains approximately 20,000 genes, which are the functional units coding for a polypeptide/protein or for an RNA chain in the organism. Approximately one percent of the 3.2 billion base-pairs in the human DNA code for these genes (*i.e.* the exome). Of the entire DNA, approximately 1% are exons, 24% are introns (*i.e.* between exons) and 75% are intergenic (*i.e.* between genes).

In the 1950's it was established that humans have 23 chromosome pairs[38], of which 22 pairs are autosomal and one pair is the sex chromosomes: the X and Y chromosomes. Females have two X-chromosomes while men have one X-chromosome and one Y-chromosome. However, there are so called pseudo-autosomal regions (PAR) on the sex chromosomes. These regions are homologous on both the X and Y-chromosomes[39], thus any genes located within them are inherited just like any autosomal genes. Two main PARs have been identified; PAR1 comprises a large region, 2.6 megabases (Mb), at the tip of the short arm of both sex chromosomes, while the PAR2 comprises 320 kilobases (kb) at the tip of the long arms of the sex chromosomes. The PAR1 contains several genes, for example the gene encoding the acetylserotonin O-methyltransferase

(ASMT), the last enzyme in the synthesis of the hormone melatonin, which is investigated in this thesis.

Genetic variation

Genetically, all humans are approximately 99.9% similar to any other human, and the remaining 0.1% variation, together with environmental factors, is what makes individuals different from each other[40,41]. This variation between humans can be used to search for genetic variation that could contribute or cause different phenotypes or diseases.

Genetic variation has arisen through the introduction of mutations, which are changes in DNA sequence. For each DNA locus there are two copies; one on each chromosome. Genetic sequences may differ among individuals as a result of mutations; the different sequences are referred to as an allele. The combination of alleles at loci that differ between individuals is called a genotype. For example, a person carrying the same alleles on both chromosomes has a homozygous genotype while a person carrying two different alleles has a heterozygous genotype. A locus that has two or more alleles that are more common than 1% in the population is defined as a polymorphism.

A mutation can be a single base-pair mutation, insertion/deletion or structural variation. Although the terms are overlapping, different forms of genetic variation can be distinguished based on size (table 3). Single nucleotide polymorphisms (SNPs), originating from a point mutation, are changes in one single base-pair. A mutation occurring in the coding sequences, *i.e.* the exome, can lead to a change of the amino acid (non-synonymous), introduce stop codon (nonsense) or have no effect on amino acid (synonymous). The function of genetic variation that occurs outside of the coding sequences are generally not known, however it can for example alter transcription of genes located in a nearby location. There has been a rapid development of new methods to analyze the DNA sequence during the recent decades. In 1970s, Sanger

Table 3. Different types of genetic variation.

Type of genetic variation	*Size*
Single base-pair	1 bp
Small deletion/insertion (and microsatellites)	base-pairs
Copy number variation (CNV)	> 1 kb
Chromosomal aberration	> Mega base-pairs
Trisomy/monosomy	Whole chromosome

Bp=base-pair. Kb=kilobase-pairs. Mb=mega base-pair.

developed a method for sequencing DNA that was used to sequence the first gene[42]. However, the first whole genomes were not sequenced until 2001 using a Sanger sequencing based method called shotgun sequencing[43,44]. The relatively high workload and cost for Sanger sequencing methods have pushed the development towards new high-throughput methods for sequencing. There was a drastic methodological change when the so called next generation sequencing (NGS), or massive parallel sequencing was introduced, making it possible to sequence the whole genome quickly and at a substantially lower cost compared to Sanger sequencing. As the prices have decreased even more for NGS, it is now being used for sequencing in large populations.

Structural variations are larger segments of DNA, which are duplicated or deleted, causing a so called copy number variation (CNV, defined as larger than 1 kb)[40,45]. Since the correct number of chromosomes in humans was assessed in the 1950's, cytogenetic methods have been used in research and diagnostics. However, with the development of methods with increased resolution, such as comparative genomic hybridization (CGH) and SNP-arrays, smaller submicroscopic CNVs can also be identified. In 2004, two important studies found that these submicroscopic CNVs were highly abundant across the human genome[45,46]. The first-generation map of CNVs in humans was constructed in 2006 (the HapMap collection)[47]. The large chromosomal aberrations are usually rare and have large effects on the phenotype, while smaller submicroscopic CNVs constitute a large proportion of the genetic variability between humans.

GENETICS OF COMPLEX DISORDERS

Heritability

Heritability is the proportion of observed differences on a trait in a population that are due to genetic factors. Estimates of heritability are hence always relative to the genetic and environmental factors in a given population, and are not absolute measurements of the contributing factors to a specific phenotype. Heritability can be estimated for both binary traits, *i.e.* cases-controls, and continuous traits using several different methods[48]. One of the most commonly used heritability estimation methods is the twin study, in which the impact of genetic and environmental factors is identified by comparing the concordance between monozygotic (MZ) and dizygotic (DZ) twin pairs. A high concordance means that both subjects in a twin pair have the same phenotype/disease. If monozygotic twins have twice as high concordance rate compared to the dizygotic twins, this is interpreted as a high heritability for the specific disease or continuous trait, *i.e.* a large genetic impact for the disease. These

comparisons can be made since MZ twins share 100% of their genetic information, while DZ twins only share approximately 50%.

Linkage disequilibrium

At the beginning of meiosis, there is an exchange of genetic material between homologous chromosomes. This results in greater genetic diversity being inherited from the parents to the child due to a recombination of genes. The probability that a recombination occurs between two alleles increases with the distance between the alleles. Two alleles in close vicinity that are associated non-randomly in a population, is termed linkage disequilibrium (LD). Essentially all types of genetic studies (see below), use LD-information to ascertain to what extent genetic variation may influence a complex trait or disease.

Common and rare genetic variation in complex disorders

When a genetic impact for a trait or disease has been established by heritability estimation, different approaches are needed to understand which genetic variations are implicated. In the human genome, both common and rare genetic variation can be involved in disease and there are two main hypotheses on how genetic influence may impact common complex disease. The dominating hypothesis has been that many common variations with small effects cause common disease; this is also known as the "common disease - common variation" (CD/CV) hypothesis. However, an alternative hypothesis suggests that the genetic causes of a common disease do not necessarily have to be common in the population, but that few rare variants with a large effect on the phenotype may cause the disease. This "common disease – rare variant" (CD/RV) hypothesis hence states that individually rare mutations could accumulate in the population and account for a significant proportion of a common disease. In reality, the truth is probably somewhere in between these two hypotheses for complex disorders[49]. There are several genetic approaches to investigate both common and rare genetic variation. Genetic studies can be performed for all types of variation in the genome, such as SNPs and CNVs. However, the alleles of CNVs are generally much less frequent than those of SNPs. Thus, CNVs have generally been more important in study designs investigating the possible influence of rare variants (*see below*).

Linkage analyses

Family based studies have a great advantage since it is possible to identify inheritance and origin of transmitted alleles that can be used for so called linkage analysis, which is a method to identify a chromosomal region where the disease gene is located. This analysis uses information from polymorphic markers across the genome to search for

loci segregating with disease. This approach has been shown to be most useful for phenotypes or diseases caused by highly penetrant genetic variants. For complex diseases, the linkage results are generally conflicting since it is unlikely that every family included in the study will carry the same collection of underlying genetic factors

Association studies

In general, an association study investigates the co-occurrence, more often explained by chance, of two or more traits in a population. Genetic association analyses can for example be case-control, family-based, and quantitative trait association studies. In these studies, the allele frequencies or genotypes are compared between groups or for continuous measures. Association analyses have been used in candidate gene studies since the 1990's and with the introduction of the genome wide SNP-genotyping arrays in the 2000's, it became possible to perform genome wide association studies, so called GWASs.

Screening for novel or rare variants

To identify rare or novel mutations in the genome, methods developed to screen the genome are used, such as sequencing. As compared to association analyses of known common genetic variation, these methods are designed to identify novel or rare mutations. Up until very recently, mutation screening studies have been hypothesis-driven candidate gene studies, in which genes pointed out by linkage studies and/or biological hypothesis have been investigated. Several of the major autism candidate genes described in the next chapter have initially been identified by a combination of linkage and mutation screening studies.

GENETICS OF AUTISM AND AUTISTIC-LIKE TRAITS

The genetic cause for autism has been identified in a small proportion of cases where autism co-occurs with a known genetic syndrome, also known as syndromic autism. The identification of the syndromic forms of autism was the first time that the genetic cause for autism was found. However, for the majority of autism cases the genetic cause has not been identified, also known as idiopathic autism.

In addition to the clinical heterogeneity within autism described in the first chapter of this thesis, there is also a large heterogeneity for the underlying genetics for autism, thus, rare variants in the same gene (allelic heterogeneity) or multiple genes (locus heterogeneity) may converge to cause autism. As the genetic techniques have developed over time, new refined strategies for genetic investigation have emerged. Hundreds of genes have so far been implicated in autism based on a large number of genetic

studies[49-52]. Genetic analyses of autistic-like traits have so far been focusing on common genetic variation that may increase the susceptibility for autism and related traits.

Heritability of autism

Autism has been identified as the most genetic neuropsychiatric disorder and the first heritability study for autism in the late 1970s found a large genetic impact for autism[25]. Since then, several studies have also shown a high heritability for autism[53-58]. However, the exact heritability estimates from these studies range between 45 to 90%, probably reflecting differences in phenotype assessment across studies and the use of different mathematical models for heritability estimates[59]. In addition to the twin based approaches, large studies have shown that the risk for autism increased with genetic relatedness in a Swedish sample of more than two million individuals[58] and that full siblings to autism cases have an increased risk for autism compared to half siblings[60,61]. Thus, these studies also support a strong genetic influence in autism etiology.

Moderate to high heritability estimates have also been shown for the dimensional measures of ALTs[18,20,62,63]. In addition, a shared genetic etiology has been suggested between ALTs and autism based on the identification of similar heritability estimates at the extreme ends of the autism continuum[19,64].

A recent twin study by Lundstrom *et al.*[65] has shown that within monozygotic twin pairs where one twin has autism, the co-twin displays autism or different neurodevelopmental problems in 9 out of 10 cases. These co-existing problems were shown to be lower in dizygotic twins, indicating that there are substantial shared genetics between neurodevelopmental problems. A shared genetic etiology has also been shown between several neuropsychiatric disorders[66-68].

Autism genes

A complex genetic etiology for autism has been identified for several years, however with emerging results from new genome wide technologies, this complexity is even more apparent. The large number of genes implicated in autism has been identified based on both biological and genetic findings. During the first years of autism genetic research, the main types of genetic analyses were linkage studies and candidate gene based association studies. As genome wide technologies with increased resolution have been developed, hypothesis-free approaches such as GWAS, CNV and whole exome/genome sequencing analyses have become prevailing strategies to identify genetic variation associated with autism

Several genome wide studies have identified a major impact of *de novo* mutations in autism[69,70]. These *de novo* mutations have been shown to have a larger impact in families with one affected individual, *i.e.* simplex families, compared to multiplex families with several affected individuals[71-73]. However, this finding is not consistent in all studies[74]. On a behavioral basis, autistic-like traits are less abundant in family members in the simplex families compared to multiplex families, suggesting that there indeed are differences in the genetic etiology of simplex and multiplex families[75]. In addition, common variations have been estimated to have a larger influence in multiplex families[76].

To present findings from genetic research in autism, there are several ways to categorize genetic variants. In the following paragraphs, the genetic variants will be classified based on their frequency and genetic impact in autism: causal or high impact variants and susceptibility variants for autism. For clarification regarding these categories, causal variants can alone contribute to a diagnosis, while a high impact variant requires a combination with other high impact or susceptibility variants. Generally, high impact variants are identified at low frequencies in the population, compared to the causal variants which are regarded as rare and susceptibility variations which are common. It should be noted that several genetic variants are located somewhere in between these categories, which are not strict entities or standard classifications. In addition, several of the identified risk variants of autism are not specific to autism, and have been identified in other disorders such as schizophrenia, bipolar disorder and ADHD[66,68].

Causal or high impact variants in autism

The first genetic causes in autism were identified for syndromic autism. In these cases, the autism diagnosis is secondary to a known genetic syndrome, such as fragile X syndrome[77], Rett disorder[78] and tuberous sclerosis[79]. Several of the genetic syndromes are so called single-gene disorders, meaning that disruption of a single gene is causal for disease. For example, the fragile X syndrome is caused by mutations in the *FMR1* gene and Rett disorder by mutations in the *MECP2*. Although the disrupted gene is known, different causal mutations can occur in these genes. Depending on the location of the mutation in the gene, this may affect the protein function differently. Thus, the clinical presentation of the genetic syndromes can be highly heterogeneous, highlighting the complex relationship between genotype and phenotype. In autism research, these genetic disorders are used as "model disorders" for idiopathic autism. Several of the affected biological pathways are beginning to be clarified by thorough analyses of these genes. Hence, if we can understand how the loss of function of the affected genes can lead to autistic behaviors such as social communication deficits, it

will help us to better understand the underlying biology of autism. The causal genes for the single-gene disorders, such as fragile X syndrome, are well characterized functionally; however, for many of the CNVs that have been identified in autism, the most crucial gene(s) are still not known (*e.g* 22q11.2 and 16p11.2). Thus, further genomic studies of CNVs associated with autism will greatly increase our neurobiological understanding of autism and related disorders.

The strong evidence of an impact of CNVs in autism identified in 2007[71,80] has been confirmed by several recent large studies[72,74,81-83]. CNVs that have been identified to have a causal or high impact effect on autism are, for example, the recurrent CNVs 15q11-q13 duplication[84-86], *SHANK2* deletion[87] and the 16p11.2 deletion[88,89]. Extensive studies of the chromosomal region 15q11-q13 have shown that this region is subject to regulation by genomic imprinting, which is an epigenetic process that can lead to the expression of genes from only one parent. Maternally derived duplications in the 15q11-q13 are strongly implicated in autism[85], while maternally derived deletions in this region cause Angelman syndrome, mainly due to the loss of expression of a gene called *UBE3A* in the brain[90]. Paternally derived deletions of the same region cause Prader-Willi syndrome[91].

Point mutations have also been identified as causal or high impact variants for autism. These point mutations have, for example, been identified in genes involved in synaptic functioning such as neuroligin 4 (*NLGN4*)[92] and contactin associated protein-2 (*CNTNAP2*)[93]. Several of the synaptic genes have been implicated in autism[94]; both causal/high impact and susceptibility genetic variants have been identified in autism. Synaptic proteins are encoded by a range of different genes: the *SHANK* genes, neuroligin genes and neurexin genes (including *CNTNAP2*)[95,96]. Shortly described, the neurexin genes (NRXNs) codes for a family of synaptic adhesion proteins located on the presynaptic membrane and bind to their postsynaptic counterpart, the neuroligin proteins (NLGNs). The SH3 and multiple ankyrin repeat domains (SHANKs) are scaffolding proteins that bind NLGN-NRXN complexes at the postsynaptic density.

Genetic variants that have a high impact, but with variable penetrance, have also been identified in autism. This has lead to the suggestion that a "second-hit" or "multiple-hits" are required for passing the threshold for autism diagnosis or worsen the symptoms for autism. There are several ways that these combinations can occur, which may be one of the explanations for the phenotypical heterogeneity in autism. Identified high impact CNVs is, for example, 16p12.1 deletion[97]. It has been shown that carriers of 16p12.1 were more likely to carry additional large CNVs in accordance with a multiple-hit model for autism[97].

In 2012, a large number of rare sequence variants was identified when the exomes were sequenced in a large number of autism cases[51,52,98-100]. The majority of the *de novo* or rare mutations found in these studies were identified in separate genes. However, a few *de novo* mutations have been identified in the same genes in the whole exome studies (for example the gens *CHD8*, *KATNAL2*, *SCN2A*, *DYRK1A*, and *POGZ*)[49], supporting that disruptions of these genes is of importance in autism.

Susceptibility genes for autism and ALTs

Since autism has a prevalence of approximately one percent in the population, common variation was initially thought to have a major influence in autism in line with the CD/CV hypothesis. However, only a few common variants have been identified in autism and some of these will be mentioned in the following paragraphs. These variants have been difficult to replicate, probably due to the small effect sizes for these variants. However, recent studies have suggested that there is indeed a large influence of a combination of common variations in autism[76,101]. In addition, common variation has been shown to contribute to the phenotypic variance for ALTs[102].

Several candidate gene studies have suggested a number of susceptibility genes in autism. Although some of these candidate gene hypotheses have been strengthened in the large genome-wide analyses, such as involvement of synaptic genes in autism, some of the findings from candidate gene studies have been difficult to replicate. Several candidate gene studies have been based on both genetic and non-genetic studies, such as biological findings showing, for example, differences in melatonin levels or testosterone levels. These examples are mentioned based on genes investigated in this thesis, however several other hypotheses have also been identified. Autism has been shown to be approximately four times more prevalent in boys compared to girls [6]. The investigation of genes related to the sex hormones[103-105] is partly based on biological findings showing association between elevated testosterone levels and autistic-like traits[106-109]. Several of the other steroid hormones have also been found to be increased in autism, suggesting an increased steroidogenic activity in autism[110]. It has also been suggested that anti-androgen pharmacological treatments are beneficial in some cases with severe phenotypes related to increased androgen levels[111]. Taken together, genes related to the sex hormones have been identified as susceptibility genes for autism and related traits.

When the genome wide genotyping arrays were introduced, three large autism GWASs were performed during a short period of time[112-114]. Two genome wide significant SNPs have been found in these studies: rs4307059, in the intergenic region between the cadherin 9 and 10 genes (*CDH9* and *CDH10*)[112], and rs4141463, in an intronic region

of the MACRO domain containing 2 (*MACROD2*) gene[114]. However, the largest GWAS meta-analysis only identified association between the gene Astrotactin 2 (*ASTN2*) and individuals with autism of European ancestry[115].

CNVs have also been shown to increase the susceptibility for ASD. For example, several recurrent CNVs were mentioned in the causal variants paragraph, such as the 15q11.2-q13 and 16p11.2. Within some of these complex chromosomal regions, common CNVs increasing the susceptibility for autism have also been identified, such as 15q11.2 duplication[116].

A GWAS has also been performed for a broader psychiatric phenotype including autism, ADHD, bipolar disorder, major depressive disorder and schizophrenia[117]. In this study, the main finding was association between the two brain expressed genes coding for L-type voltage-gated calcium-channel subunits (*CACNA1C* and *CACNB2* genes). It should be mentioned that mutations in the *CACNA1C* gene cause a genetic syndrome called Timothy syndrome, in which autism often occurs[118]. Thus, both susceptibility and causal genetic variants have been identified in the *CACNA1C* gene[117,118].

GWASs have also been performed for ALTs[102,119-121]. The majority of these studies did not find any genome wide significances between common variations and ALTs[102,119,120]; however, one of the studies indentified association between social communication difficulties and the regions 3p22.2 and 20p12.3[121]. Candidate gene association studies have also identified association between specific autistic-like traits, such as specific language problems in the general population and genetic variation in genes such as *CNTNAP2* and *FOXP2*[122].

MELATONIN

In two of the included papers in this thesis, the focus was to investigate the influence of genetic variation in melatonin-related genes on autism and autistic-like traits. Our analyses were based on findings showing altered levels of melatonin in autism and that genetic variation in genes related to melatonin has been associated with autism.

Function and synthesis of melatonin

Melatonin function

The neurohormone melatonin is well known for its role in circadian sleep-wake rhythm. Since the identification of melatonin[123], the hormone has been extensively studied with regards to its biosynthesis and biological actions. It has been shown to be involved in several physiological functions, such as sleep induction, circadian rhythm regulation and immune response[124]. Furthermore, melatonin has been suggested to be an important regulator of embryonic neurodevelopment[125-127] and is one of few hormones that are able to pass the placenta during pregnancy[128]. During the first trimester, melatonin receptors have been found to be present in the placenta, where it has been suggested to act as a local regulator of placental function[129]. Abnormal melatonin signaling has been shown to be a risk factor for several medical conditions such as diabetes and psychiatric disorders[130].

Melatonin acts by binding to the G-protein coupled melatonin receptors MTNR1A and MTNR1B. Although a third melatonin receptor called e GPR-50 has been identified, this receptor does not have an affinity for melatonin; it inhibits melatonin action by binding to the MTNR1A receptor[131].

Melatonin synthesis

Melatonin is mainly synthesized in the pineal gland in the brain[124]. The primary function of the pineal gland is to transduce light and dark information to the whole body by releasing the hormone melatonin directly into the blood stream. The melatonin levels are low during daytime, while there is a peak in melatonin production during the night.

Melatonin is synthesized from tryptophan, which is taken up from the blood stream. Tryptophan is subsequently converted to serotonin by two enzymatic reactions in the pineal gland (figure 1). The conversion from serotonin to melatonin includes two important enzymes: arylalkylamine N-acetyltransferase (AA-NAT) and acetylserotonin O-methyltransferase (ASMT)[132,133]. The ASMT is regarded as the rate-limiting enzyme for melatonin production during the night when the important melatonin peak occurs[124].

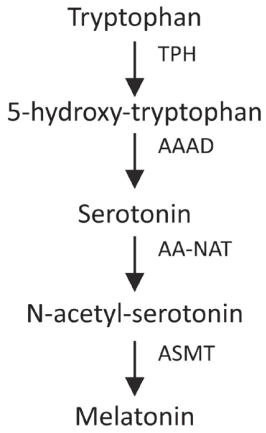

Figure 1. Melatonin synthesis in the pineal gland. AA-NAT=Aryl-Alkyl-amine N-Acetyl-Transferase, AAAD= Aromatic L-Amino Acid Decarboxy-lase. ASMT=Acetyl-Serotonin O-Methyl-Transferase. TPH=Tryptophan Hydroxylase.

There are several complex mechanisms involved in this circadian regulation of melatonin. The main regulation of circadian rhythms is located within the suprachiasmatic nuclei (SCN) in the hypothalamus. Within the SCN, the circadian rhythms are regulated by both external factors and the internal clock genes, which are part of an autonomous circadian rhythm in the SCN that also exists without external cues such as light and dark. One of the outputs from the SCN is the rhythmic regulation of melatonin production, thus sending a circadian message to the whole body.

When melatonin has been synthesized, it is immediately released into the systemic circulation where it reaches the peripheral and central targets. Melatonin has a short half-life of approximately 20 minutes[134]. Melatonin levels can be measured directly in plasma or indirectly by saliva measurements. The melatonin production over time can be measured in urine by measuring its inactive metabolite: 6-sulphatoxymelatonin.

Melatonin and autism/ALTs

Sleep disorders are highly prevalent in subjects with autism, ranging from 40% to 80%[135]. In addition, melatonin is often successfully used as treatment in autism[136] where it has not only been shown to improve sleep impairments but also daytime behavior[137]. All studies investigating melatonin levels in autism have found abnormal levels in cases compared to controls[138-145]. In several of the studies, a decreased nighttime melatonin level was associated with autism compared to controls[138,139,141-144]. Although the results were not completely consistent, they strongly indicate that lower nocturnal levels of melatonin are often observed in autism. These lower melatonin levels may be of importance in a subgroup of children with autism. Interestingly, in the

study by Tordjman *et al.*[138] the decreased nocturnal level of the melatonin metabolite 6-SM has been correlated with increased severity of verbal language, imitative social play and repetitive use of objects.

Several studies have identified elevated whole-blood levels of serotonin in autism[146,147]. It has been suggested that these elevated levels may be associated with the lowered melatonin synthesis since serotonin is a precursor to melatonin[140,142,148]. In addition, the serotonin-N-acetylserotonin-melatonin pathway has been suggested as a biomarker for autism spectrum disorders[146].

Genetic studies of melatonin-related genes

The regulation of melatonin secretion has been shown to have a large genetic component in a twin study[149]. The melatonin-related genes investigated have primarily been the genes coding for the two synthesis enzymes (*AA-NAT* and *ASMT*) and the melatonin receptors (*MTNR1A*, *MTNR1B* and *GPR50*). Both common and rare genetic variations in melatonin-related genes have been implicated in autism and also ALTs.

Melatonin synthesis enzymes

Of the two main enzymes for the biosynthesis for melatonin (AA-NAT and ASMT) the *ASMT* gene, encoding the nocturnally rate-limiting enzyme in melatonin synthesis, has been more extensively studied in autism genetic studies. However, two mutations in the *AA-NAT* have been shown to alter melatonin production in a mutation screening study in ADHD cases and controls[150].

The crystal structure has recently been identified for *ASMT*, making it easier to interpret the effects of genetic variation in the gene[151]. Several rare mutations in the *ASMT* gene have been identified in autism[142,151-154]. Some of these mutations have been shown to have effects on the enzymatic function of ASMT, which has been shown both by *in vitro* investigations and by measuring melatonin production in subjects carrying the mutation[142,151]. One of the identified rare mutations is the splice site mutation IVS5+2T>C[142,152,153], causing decreased enzymatic function of the ASMT.

Common variations in the *ASMT* gene have also been investigated in autism. Two SNPs (rs4446909 and rs5989681) in the promoter region of the *ASMT* gene have been associated with autism[142]. In addition, the more common allele of these two SNPs in autism compared to controls were also associated with lower ASMT transcript levels[142], suggesting that these polymorphisms may lower the melatonin production. However the association between these SNPs and autism could not be replicated in a Chinese

Han population[154] or in another European population[153]. A common CNV (microduplication), spanning exon 2-8 of the *ASMT* gene, has been associated with autism in a case-control study[155].

Melatonin receptors

The *MTNR1A, MTNR1B* and *GPR-50* are located in 4q35.2, 11q14.3 and Xq26, respectively. A mutation screening study of these three genes has identified several deleterious mutations in a few of the 300 included cases with autism, and also in controls[156]. This study did not find clear associations between autism and the melatonin receptors, although they found several interesting functional mutations, such as the p.I49N-mutation in *MTNR1A* which showed a complete loss of melatonin binding and signaling capacity. One of the mutations altering the function of the MTNR1B (p.V124I) has been identified in two subjects with autism and was not reported in any controls[152,156]. In addition, this study has provided functional studies of genetic variations in the melatonin receptor genes that are of importance to future studies of these receptors, not only in autism. Genetic variation in *MTNR1B* has previously been mainly associated with diabetes type 2[157,158].

AIMS

The broad aim of this project is to identify genetic factors underlying autism spectrum disorder (ASD) and autistic-like traits (ALTs).

The specific aims:

1. To screen for rare genetic variation in genes encoding the melatonin receptors (*MTNR1A*, *MTNR1B* and *GPR-50*) and the two melatonin synthesis enzymes (*ASMT* and *AA-NAT*) (paper I).

2. To assess the influence of common genetic variations in the *ASMT* gene on dimensional measures of autistic-like traits (paper II).

3. To assess the influence common genetic variations in autism candidate regions (*CNTNAP2*, *RELN*, *SHANK3*, and rs4307059) on autistic-like traits (paper III).

4. To screen the whole genome/exome for rare inherited copy number variations (CNVs) and point mutations in a multiplex family where several members have autism and language problems (paper IV).

SUBJECTS AND METHODS

POPULATIONS AND MEASUREMENTS

The Child and Adolescent Twin Study in Sweden

The Child and Adolescent Twin Study in Sweden (CATSS) is an ongoing longitudinal study since 2004 focusing on both somatic and neurodevelopmental problems in children and adolescents[159]. All twins born after 1992 are identified through the Swedish twin registry and asked to participate. To date, approximately 20,000 subjects have participated in the study. The CATSS is a longitudinal study and investigations at different ages are included. The children were first asked to participate when they were 9 or 12 years old. During the first three years of the study, both 9 and 12 year old twins were included (born between 1992 and 1995). However, after the first three years of the study, only 9-year old children were included. This age has been chosen since most of the child psychiatric problems have been established by this age, but the children have not yet reached puberty. To assess measures of somatic and mental health in the children, several different instruments have been used in the study[159].

In papers II and III, two CATSS subpopulations were included: 1,747 subjects (paper II) and 12,319 subjects (paper III). Individuals with known genetic syndromes and brain damages were excluded from the analyses. Within the CATSS population, approximately a third of the twins are monozygotic twins and two thirds are dizygotic twins.

Measurements

Information regarding autistic-like traits has been collected using the Autism-Tics, AD/HD and other Comorbidities inventory (A-TAC)[22,160,161]. The A-TAC inventory, which is a parental questionnaire, has been designed for large scale epidemiological studies to capture neurodevelopmental problems in children when they are 9 or 12 years old. The problems captured in the A-TAC cover the clinical diagnoses ASD, ADHD, tic disorders, developmental coordination disorder, and learning disorders. There are 96 questions included in the A-TAC, of which 17 questions are related to three different domains of ALTs: social interaction impairment (6 questions), restricted and repetitive behavior (5 questions) and language impairments (6 questions). Each of the questions can be answered as "No" (score 0.0), "Yes, to some extent" (score 0.5) and "Yes" (score 1.0). These questions are based on the diagnostic criteria used in the DSM-IV, thus the A-TAC captures ALTs related to the diagnosis of

ASD. The total scores for all questions, as well as within each module, were used in the association analyses. Thus, seventeen is the highest score on the total ALT score. These scores were also used as a proxy for ASD diagnosis, using a cut-off score at 8.5, in the case-control analyses in paper III. This cut-off has been shown to have high sensitivity and specificity for ASD[22,161].

In addition to investigations of ALTs, genetic factors for the overlap of neurodevelopmental problems (NDPs) have been identified using the A-TAC[162]. This study investigated if the large degree of observed overlap between NDPs could be attributed to specific genetic or environmental factors. A general genetic factor, and three specific genetic subfactors, with 100% heritability by design have been identified by Pettersson *et al.*[162], based on 53 of the items in the A-TAC. These 53 items have different genetic loads onto each of the identified genetic factors. All the included A-TAC items loaded onto the general genetic factor, while subsets of items loaded onto the specific genetic subfactors. Thus, the general genetic factor did not represent specific neurodevelopmental problems and has been suggested to be indicative of severity. The three genetic subfactors were shown to tap three different problem categories: tics and autism, hyperactivity and learning problems. It should be noted that no trait has 100% heritability, however the identified genetic factors capture the genetic overlap between the neurodevelopmental problems.

DNA collection

Saliva samples have continuously been collected from the twins in CATSS using Oragene Self Collection kit (Oragene). Twin zygosity was measured using a panel of 47 SNPs[163].

ASD population

In paper I, mutation screening of melatonin-related genes was performed on 109 individuals with ASD from two populations. Sixty-five cases were identified at the Mölndal Hospital (Gothenburg, Sweden) and 44 cases were recruited from several locations within Stockholm County (Sweden). The subjects with ASD were diagnosed using the DSM-IV or Autism Diagnostic Interview-Revised (ADI-R). Subjects with known genetic syndromes, such as Fragile X syndrome and tuberous sclerosis, were excluded from the study.

Mölndal population

Thirty-two of the 65 ASD cases met the criteria for autistic disorder while the other 33 subjects met the criteria for pervasive developmental disorders not otherwise specified (PDD-NOS). Mental retardation was also identified in 24 of the subjects with autistic

disorder and in 28 of the subjects with PDD-NOS. This population included 39 males and 26 females.

Stockholm population

This population included 44 subjects with ASD, who have been recruited from several locations in Stockholm County. In this population, 26 subjects met the criteria for autistic disorder, 26 subjects met the criteria for PDD-NOS and 5 subjects met the criteria for Asperger's syndrome. Mental retardation was also identified to range from mild to moderate in 21 ASD cases.

Control population

The control population included 188 subjects. These subjects were included in the study to investigate if the identified rare variants could be found in controls. The subjects in this population were not matched with the ASD populations with regards to age and gender.

ASD multiplex family

In paper IV, a multiplex family comprising thirteen family members was included. Two brothers with autistic disorder were diagnosed using the DSM-IV and ICD-10. These brothers were identified through a language and autism screening for the project *AUtism Detection and Intervention in Early life (AUDIE)* in Gothenburg, Sweden[164,165]. They belong to a large family in which there is a high occurrence of ASD and language problems. Both children have speech sound disorder and language disorder. There were no pregnancy or delivery complications noted for any of the children.

Within the extended family there are several cases of autism and language problems. The mother to the index children has dyslexia. One of the children's uncles has a classic autism diagnosis and also mental retardation. One aunt is suspected to have an Asperger's syndrome diagnosis (or another autism spectrum disorder). At a young age she was also reported to have language problems. The index children have two cousins, a boy and a girl, of whom one has speech and language problems.

Blood or saliva samples have been collected from thirteen family members. DNA was prepared from blood samples in nine individuals, using QIAmp® DNA Blood Mini kit (Qiagen). From four individuals, saliva samples were collected using Oragene Self Collection kit (Oragene). DNA from these samples was prepared using PrepIT (DNA Genotek, Ottawa, Canada).

GENOTYPING

Different methods have been used for genotyping in this thesis. Genotyping of SNPs in the large CATSS population (paper II and III) was performed using the KASP™ genotyping chemistry (KBiosciences, LGC). The quantitative real-time PCR (q-PCR) method was used to for genotyping CNVs (paper II and IV). In addition, variants identified from the sequencing projects (paper I and IV) were genotyped using the Pyrosequencing method.

SNP genotyping

In papers II and III, the genotyping of SNPs in the CATSS population was performed using the KASP™ genotyping chemistry (KBiosciences, LGC). The KASP genotyping assay contains the DNA sequence of interest, two competitive, allele specific forward primers and one common reverse primer. Allele specific primer means that each primer only codes for one of the nucleotides in the SNP. These primers also contain an additional tail sequence that corresponds with one of two universal FRET (fluorescent resonance energy transfer) cassettes present in the KASP Master mix. The fluorescence is measured to determine the genotype in each subject; if there are equal amounts of fluorescence in the reaction indicates a heterozygous genotype since both allele specific primers are producing PCR product in the PCR reaction.

CNV genotyping

The quantitative real-time polymerase chain reaction (q-PCR) was used for genotyping a microdeletion in the *ASMT* gene (paper II) and for validation of CNVs identified from the Affymetrix 6.0 SNP array (paper IV). TaqMan® Copy Number Assays (Applied Biosystems, Life Technologies) were used for these analyses.

In general, PCR is used to amplify selected DNA sequences that can be used for downstream PCR based analyses. The q-PCR method quantifies the produced PCR product by real time recordings of the PCR process in 7900HT Sequence Detection System (Applied Biosystems, Life Technologies). To be able to detect the PCR products, the primers have incorporated fluorescence markers. The measured fluorescence reflects the amount of produced PCR product. Thus, this analysis gives information regarding DNA quantity for each cycle in the PCR. For the CNV genotyping, a TaqMan® Copy Number Assay for the CNV of interest is run together with a TaqMan® Copy Number Reference Assay. The cycle threshold (C_T) values were obtained for both the TaqMan® Copy Number Assays and TaqMan® Copy Number Reference Assay in the Sequence detection system (SDS, Applied biosystems, Life

Technologies). These results were analyzed by the relative quantitation method using CopyCaller™ Software to determine copy number state for the investigated CNVs.

Pyrosequencing

Pyrosequencing (Qiagen) was used for genotyping the variants identified in melatonin-related genes in a control population (paper I). In addition, this method was also used for validation of rare variants identified from the whole exome sequencing (paper IV). Pyrosequencing is a method based on the "sequencing by synthesis" principle. For pyrosequencing, a single stranded DNA sequence of interest is used as the template for the sequencing primer. For the sequencing, a predetermined order for adding nucleotides is specified. If an added nucleotide is incorporated in the DNA synthesis, this causes emission of light by a luciferase-mediated conversion of luciferin to oxyluciferi. An added nucleotide that is not incorporated in the synthesis does not emit light, thus the predetermined order of nucleotide addition can be used to determine the sequence of interest.

STATISTICAL ANALYSES

The association analyses between SNPs and phenotype measures in the CATSS population were performed using SAS 9.3 (SAS Institute Inc.). For analyses of continuous traits (papers II and III), the mixed procedure (PROC MIXED) was used in the SAS 9.3. For case-control association analyses (paper III), the glimmix procedure (PROC GLIMMIX) was used. In our statistical analyses, both twins in each twin pair were included in the analyses. Since monozygotic twin pairs are genetic copies and dizygotic twins share approximately 50% of their genomic sequences, this was corrected for in the statistical analyses by specifying two separate variance-covariance matrices for monozygotic twins and dizygotic twins in the mixed effects models.

SEQUENCING

Two different sequencing methods were used in papers I and IV included in this thesis. Whole exome sequencing was used in the exploratory investigation of an ASD multiplex family (paper IV). Whole exome sequencing is a high throughput method, also known as next generation sequencing, for sequencing of all exons simultaneously. The Sanger sequencing method was used in the candidate gene study screening melatonin-related genes (paper I). Using this method, each exon of the investigated genes was sequenced separately.

Whole exome sequencing

The two index patients, their parents and two brothers to the mother in the investigated family in paper IV were selected for whole exome sequencing using 3 µg DNA. The SureSelect kit v. 3 (Agilent) was used for the exome sequencing and was run on the HiScanSQ (Illumina) at the Genomics Core Facility, University of Gothenburg. The alignment and variant calling was performed using the GATK software (Broad Institute) at the Bioinformatics Core Facility, University of Gothenburg.

In whole exome sequencing, the genetic sequence of almost all human genes is identified. This results in a large amount of data; in our six analyzed subjects a total of 716,454 genetic variants were identified. All of these genetic variations are not implicated in autism, thus, several filtration steps are applied to narrow down the list of autism candidate variants.

For variant filtration and selection of variants for further inspection in our study, we used the Ingenuity variant analysis (IVA) software. In IVA, several different criteria for variant filtrations can be used. In our analyses, we used three main filtering criteria based on quality of the sequences, variant frequency reported in public databases and predicted functionality of the variants according to *in silico* predictions. In addition, the variants were manually filtered based on inheritance patterns. Selected variants were also inspected in the Integrative Genome Browser (Broad Institute).

In our analyses, the first approach was to search for variants in genes previously reported as autism candidate genes in the SFARI gene database (Banerjee-Basu and Packer, 2010). The SFARI database contains information regarding genes that have been implicated in autism. For the second approach, we searched for novel variants not previously associated with autism. The prevalence of identified rare variants was obtained from publicly available databases such as the Exome Variant Server (EVS) from the NHLBI Exome Sequencing Project.

Sanger sequencing

Mutation screening of the melatonin-related genes *AA-NAT* (arylalkylamine N-acetyltransferase), *ASMT*, *MTNR1A*, *MTNR1B* (melatonin receptor 1A and 1B) and *GPR-50* (G protein-coupled receptor 50) was performed using the Sanger sequencing chain-termination method (paper I). In a regular PCR, the nucleotides will be added on the 3'-OH on the most recently incorporated nucleotide. In the Sanger sequencing method, regular nucleotides and a few nucleotides lacking the 3'-OH, *i.e.* termination

nucleotide, will generate DNA sequences of different lengths. The fragments of different lengths can thereafter be separated with capillary electrophoresis using an ABI 3730 genetic analyzer (Applied Biosystems, Foster City, CA). The sequencing results were analyzed in Genalys© in our study.

COPY NUMBER VARIATION

The Affymetrix 6.0 SNP-array (AROSAB, Aarhus, Denmark) was used for the copy number variation (CNV) analyses in thirteen subjects from one family. The SNP-array contains approximately two million markers across the genome. The identification of CNVs is based on measuring the signal intensities of the markers in the array to identify increased or lower intensities compared to a pooled reference genome. To obtain reliable CNV measures, a CNV was called when several consecutive markers indicate increased or decreased intensities, corresponding to a duplication or a deletion, respectively.

The genotyping console software (GTC v4.1, Affymetrix) was used for quality control and CNV calling. The CN5 algorithm in the GTC was used for CNV calling. An in-house reference file (provided by AROSAB, Aarhus, Denmark), that was based on 44 individuals (16 males and 28 females) carrying no or very few chromosomal changes, was used in the GTC. In addition, only CNVs that were also identified by a second CNV calling algorithm were included in downstream analyses. The Birdseye algorithm in the Birdsuite software (Broad Institute) was used as the second CNV calling algorithm.

The first approach for CNV analysis was to search for CNVs previously associated with autism. CNVs were regarded as autism risk variants if they were identified in 11 regions where an excess of CNVs have been identified in autism. These 11 regions have been identified in a previous study evaluating almost 5,000 CNVs in 2,373 ASD subjects[166]. In addition, we used the SFARI gene database[167] to investigate if the CNVs previously have been identified in autism. Our second approach was to search for rare CNVs identified in regions with few previously reported CNVs in two large studies investigating CNVs in the general population[168,169] or in the database for genomic variants[170].

FUNCTIONAL ANALYSIS OF *CYP11A1*

In paper IV, a rare mutation in the *CYP11A1* (p.T369M) gene coding for the P450 side chain cleavage enzyme (P450scc) was identified. This enzyme catalyzes the first step

of the steroid hormone biosynthesis by converting cholesterol to pregnenolone. The enzyme function was measured in cell culture experiments where a plasmid containing the P450scc enzyme was transfected into COS-1 cells.

The identified mutation (p.T369M) in *CYP11A1* coding for the P450scc was recreated in the P450 moiety of the F2-plasmid expressing the fusion protein NH2-P450scc–Ferredoxin–Reductase–Ferredoxin-COOH[171]. Mutagenesis was performed for the T369M-mutation in *CYP11A1* and two previously identified mutations, F215S[172] and A269V[173], using the QuickChange® site-directed mutagenesis kit at the Mammalian Protein Expression core facility, University of Gothenburg. These plasmids were transfected into COS-1 cells, which are a fibroblast-like cell line derived from African green monkey kidney tissue.

For the measurement of P450scc enzymatic function in COS-1 cells, the enzymatic substrate cholesterol was added to the cell culture medium. The enzymatic function was determined by measuring the enzymatic product, *i.e.* pregnenolone, at different time points after the addition of the cholesterol. The pregnenolone concentrations in the cell culture medium were measured using the enzyme-linked immunosorbent assay (ELISA) for pregnenolone. To control for transfection efficiency of the plasmid, we co-transfected the cells with a control plasmid called Renilla Luciferase reporter plasmid (pRL-CMV, Promega). The amount of Renilla Luciferase was measured using a Renilla Luciferase assay (Promega). These measurements were used to normalize the concentrations of pregnenolone for transfection efficiency.

RESULTS AND DISCUSSION

GENETIC VARIATION IN MELATONIN-RELATED GENES

Low melatonin levels are repeatedly identified in autism indicating a possible importance for melatonin-related genes in the etiology of autism. To further investigate the involvement of these genes in autism, we investigated rare variation in genes coding for melatonin synthesis enzymes and the major melatonin receptors (paper I). In addition, we investigated the association between a broader phenotype, the continuously distributed autistic-like traits in the general population, and common variation in the *ASMT* gene, coding for the last enzyme in the melatonin synthesis (paper II).

Autism and rare mutations in melatonin genes (paper I)

For analyses of rare genetic variation in melatonin-related genes in autism, we sequenced the genes encoding the two major enzymes for melatonin synthesis (*ASMT* and *AA-NAT*) and three melatonin receptors (*MTNR1A*, *MTNR1B* and *GPR50).* Our analyses were based on previous findings of altered melatonin levels in subjects with autism and previous identification of rare genetic variation causing decreased levels of melatonin in mutation carriers[142].

The five melatonin-related genes were sequenced in individuals with autism. Six rare mutations were identified, of which three mutations were identified in *ASMT*, one in *MTNR1A* and two in *MTNR1B*. The main finding was the identification of yet another carrier of the previously reported splice site mutation (IVS5+2T>C) in *ASMT*. This mutation has been found to cause drastically lower melatonin levels in carriers[142]. Prior to the publication of our study, this splice site mutation had been identified in 6 of 640 autism cases (0.9%) and in only one of the 916 control subjects (0.1%)[142,153]. At present, the mutation has also been identified in 21 of the 6,499 control subjects (0.3%) included in the publicly available database Exome Variant Server (EVS) (from NHLBI Exome Sequencing Project). Taken together with our results, the data indicates that the mutation occurs more often in autism cases compared to controls. The mutation has also been identified in a subject with ADHD and coexisting autistic traits, language delay and sleeping problems[150] as well as in a subject with intellectual disability and autistic features without any major sleep problems[174]; thus these studies found two carriers in two separate populations of ADHD (N=101) and intellectual disability (N=377), but the mutation was not identified in any controls in these studies

(N=660 for both studies). Thus, the splice site mutation (IVS5+2T>C) may not only be related to autism, but also other phenotypes that could be related to abnormal melatonin synthesis.

Of the variants affecting protein sequence in our study, only the mutation (V124I) in the *MTNR1B* gene caused an amino acid change. This mutation was identified in one person with autism, but was absent in our comparison group. In functional analyses, this mutation has been shown to alter the function of the melatonin receptor[156].

Autistic-like traits and common variation in *ASMT* (paper II)

The rare variants identified in *ASMT* (and other melatonin genes) only explain a small portion of the low melatonin levels observed in autism patients. Hence, common variation in these genes may also be associated with autism related phenotypes and/or melatonin levels. In paper II, we wanted to explore the influence of common variation in this gene on milder phenotypes, *i.e.* autistic-like traits (ALTs). The investigation of continuous ALTs makes it possible to analyze the severity of symptoms. In addition, it also opens up for separate analyses of the three core domains of autism, which are impairments in social interaction, language impairments and restricted behavior. Thus, we also investigated if genetic variation affects all three domains or only a specific ALT domain.

To investigate common variation in the *ASMT* gene, we selected six SNPs covering the gene and also the promoter region, where several SNPs previously associated with autism are located. The association analyses were performed between these SNPs and ALTs measured in 1771 subjects from the general population (CATSS population). In addition, association analyses were investigated for a common CNV in the *ASMT* gene based on a previous study showing an increased prevalence of this CNV in autism[155].

The main finding in this study was the association between social communication impairments in girls and a SNP in the last intron of the *ASMT* gene (rs5949028). The location in the last intron of the *ASMT* gene does not implicate a major functional effect on gene regulation. This mutation could, however, be in linkage disequilibrium with other yet unknown functional variants. To relate our finding with previous genome wide studies of ALTs, the resolution of the genome wide arrays is of importance. Although the SNP-arrays have been developed to have a high resolution, some of the early genome wide arrays do not capture genetic variation in the *ASMT* gene. Thus far, four genome wide association studies for ALTs have been published[102,119-121], of which three were published after our paper II was accepted for publication. Of these studies, only the study by Jones *et al.*[120] covers the *ASMT* gene.

However, this study did not find association between the *ASMT* gene and ALTs. Overall, the genome wide studies for ALTs have identified very few associations between common variation and ALTs. In the ALT GWASs, different measures of ALTs have been used. We measured ALTs using the A-TAC, whereas, for example, the Autism Spectrum Quotient (AQ)[23] was used in the study by Jones *et al.*[120]. A major difference between these two measures of ALTs is that the A-TAC is a DSM-based questionnaire, thus it captures traits related to the diagnostic criteria, while the AQ is designed to capture a broader definition of ALTs.

Since the publication of paper II, we have also analyzed three SNPs (rs5949028, rs5989681 and rs4446909) in *ASMT* and one SNP in *MTNR1B* (rs4753426) in a larger CATSS sample (>12,000 subjects) (unpublished data). However, we could not replicate the association between *ASMT* and ALTs. We did find a significant association between the investigated SNP in *MTNR1B* (rs4753426) and total ALT scores in girls (p=0.002). Overall, replication of associations between common variation and autism or ALTs has been shown to be difficult and further investigations of common variation in *ASMT* and *MTNR1B* are warranted.

The CNV spanning exon 2-8 in *ASMT* was analyzed for a possible influence on ALTs in our study; however the CNV was not shown to be associated with continuous measures of ALTs. To be able to perform case-control analysis comparable to the first study by Cai *et al.*[155], we would have needed a larger sample. Since the investigated CATSS population included subjects from the general population, this population resembles the prevalence of autism in the general population.

Conclusion paper I and II

Although none of the identified melatonin related mutations display full penetrance, we and others have indeed shown that variants in these genes are associated with higher risk for autism. We identified an association between common variation in the *ASMT* and ALTs, however this could not be replicated in our second cohort. The strongest evidence has been shown for the splice site mutation (IVS5+2T>C) in the *ASMT* gene, which disrupts the function of the enzyme leading to dramatically decreased melatonin levels in the carriers. Indeed, it has also recently been suggested that disruption in the melatonin synthesis pathway (serotonin-NAS-melatonin pathway) could be used as a biomarker for autism in a subgroup of cases[146]. Taken together, our results support that *ASMT* is a susceptibility gene for autism.

AUTISTIC-LIKE TRAITS AND COMMON VARIATION IN AUTISM CANDIDATE GENES (PAPER III)

In the third study, we further investigated the influence of common genetic variation on autistic-like traits in a large sample from the general population. For this study, we selected five SNPs in genes that previously have been strongly implicated in autism. In addition to investigating ALTs, we also investigated the association between the selected SNPs and genetic factors for neurodevelopmental problems (NDPs).

Although hundreds of genes have been implicated as risk genes for autism, various penetrance and effect sizes have been shown for genetic variation in these genes. The genes selected for investigation in our study have repeatedly been shown to be of importance in autism, both in human genetic studies and in functional studies. Two of the genes (*CNTNAP2* and *SHANK3*) are involved in synaptic functioning, which is one of the major pathways associated with autism. Furthermore, reelin (encoded by *RELN*) has been shown to have critical functions during neurodevelopment and is involved in the maintenance of synaptic function in adulthood[175]. Decreased reelin expression has been found in several neurodevelopmental disorders such as schizophrenia, bipolar disorder and autism[176-178]. Finally, we selected a SNP (rs4307059) in the region between the cadherin 9 (*CDH9*) and cadherin 10 (*CDH10*) genes. This SNP has been associated with both autism[112] and autistic-like traits[179]. Although these are strong candidate genes for autism, the main finding in our study was that we did not find any significant association between the investigated SNPs and measures of ALTs or NDPs in our large sample from the general population.

The SNP rs4307059 is one of the few common variations that has been associated in both an autism case-control study[112] and a study investigating the ALT social communication domain[179]. Although our analyses did not show any association between rs4307059 and ALTs, one of the differences between our study and the study by Pourcain *et al.*[179] was that different measures of ALTs were used in these studies.

CNTNAP2 is one of the largest mammalian genes and spans more than 3.3 Mb in the chromosomal region 7q35. A rare homozygous mutation has been identified in the *CNTNAP2* gene in Amish children with mental retardation and language impairments[93]. In addition, mouse models lacking *CNTNAP2* have been shown to display features similar to the core domains of behavior and cognition in autism[180]. In our study, we chose to investigate two SNPs in the large *CNTNAP2* gene: rs7794745 (between exon 2 and 3) and rs2710102 (between exon 13 and 14). The SNP rs7794745 has been associated with autistic disorder[181]. In addition, carriers of the risk

allele in the general population have been shown to have altered brain volumes, as measured by structural magnetic resonance imaging (MRI), in regions previously associated with autism[182]. The other investigated SNP (rs2710102), and several other common variations in the region between exon 13 and 14, have been associated with language problems[183], such as age at first word in children with autism[184]. In addition, the SNP rs2710102 has been associated with language development in the general population[122]. However, negative association studies between *CNTNAP2* and autism have also been published[185]. Taken together, previous findings indicate that the selected SNPs were strong candidates for ALTs association analyses.

SHANK3 was the first of the SHANK genes to be associated with autism, by the identification of three autism families carrying deletions or a single nucleotide insertion in the *SHANK3* gene[186]. Rare mutations and deletions have also thereafter been identified in autism[187,188], although the association analyses between common variations in autism case-control studies have shown some inconsistent results for most of the SNPs in *SHANK3*. In our study, a non-synonymous SNP (rs9616915, p.I245T) in the *SHANK3* gene was investigated.

The SNP analyzed in the *RELN* gene is a non-synonymous SNP rs362691 (p.L997V) that has been associated with autism in a recent meta-analysis[189], which included five association studies[190-194]. Other frequently studied common variations in *RELN*, such as a trinucleotide repeat (GGC) and rs736707, was not shown to have a major impact for autism in the meta-analysis by Wang *et al.*[189].

In addition to the analyses of ALTs, we also investigated the influence of the selected SNPs on genetic factors for NDPs. These genetic factors have been identified in the CATSS population in a study by Pettersson *et al.*[162] with the aim to understand whether the extensive overlap among NDPs was primarily genetic or environmental. Since previous research indicates a large overlap between the genetic factors for different neurodevelopmental disorders, we wanted to investigate if our selected SNPs were associated with this genetic overlap between NDPs measured in the general population. Although we did not find any association between the genetic factors for NDPs and the selected SNPs in this study, further investigations are warranted to understand the genetic susceptibility for NDPs.

In general, it has been difficult to replicate associations between common variations and autism or ALTs. Since common variation usually has very small effect sizes, the approach to investigate continuous measures of ALTs rather than case-control autism analyses is a useful strategy. Although, the results in paper III do not suggest a major

importance of the investigated common variations on ALTs or the genetic factors for NDPs, it can not be excluded that other genetic variation in the investigated genes may have an influence on ALTs.

SCREENING FOR RARE VARIATIONS IN A MULTIPLEX FAMILY (PAPER IV)

Encouraged by findings identifying several genetic variations not only specific for autism, but also other neuropsychiatric disorders, we used the latest genetic technologies to investigate a large family with several members with autism and language problems (figure 2). We analyzed CNV and rare sequence mutations in a family consisting of two index children with autism and eleven family members.

The genetic heterogeneity for autism suggests that multiple genetic loci may be involved in the etiology for autism. However the relationship between genotype and phenotype is not well understood. For the investigation of rare CNVs and point mutations in this study, the first approach was to search for previously identified high impact variants in autism. The second approach was to search for novel variants with possible implications as risk variants for autism and language problems.

The main finding in this study was the identification of several variants suggested to influence the genetic etiology for autism, and also language problems (figure 3). Three very rare paternally inherited CNVs were identified in the index children: 6q15, 11q21, and 4q34.3. In addition, we also identified a paternally inherited deletion in the 15q11.2 region. The whole exome sequencing study revealed several rare mutations, of which 18 mutations were validated and investigated in the whole family. Several of these mutations are located in genes previously associated with autism or other neuropsychiatric disorders. However, we also found novel variants not previously

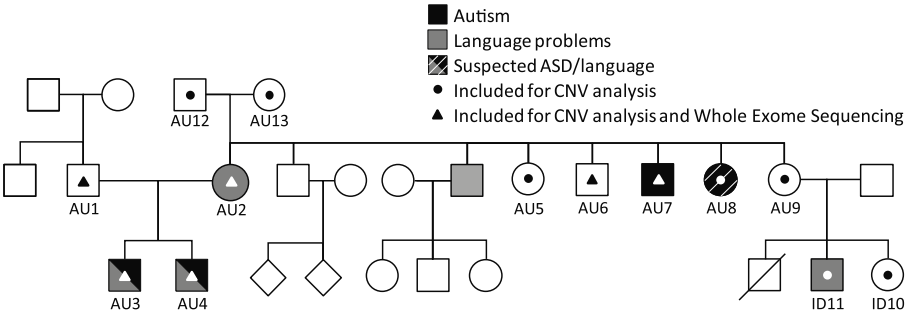

Figure 2. Pedigree of the multiplex ASD family investigated in paper IV.

identified as susceptibility genes for autism or related phenotypes.

For one of the identified mutations we investigated the functional properties in cell culture. This mutation (p.T369M) is located in the *CYP11A1* gene, coding for the first enzyme in the steroid hormone biosynthesis: the P450 side chain cleavage enzyme (P450scc). Our analyses suggest that this mutation causes a gain of function of the enzyme. Since the P450scc is the first step in all steroid hormone biosynthesis, this may have implications on the downstream production of both sex hormones, such as androgens, and also on cortisol and mineralocorticoids, such as aldosterone. In our study, we did not measure hormone levels in the carriers of this mutation; however, this would be useful to determine if the mutation has any effects on hormone levels in carriers. Alterations in androgen levels have previously been associated with autism[106-109]. In addition, the recent report of elevated levels of all steroid hormones in autism[110], indicates a potential role for the identified mutation in the *CYP11A1*.

Taken together, our study investigated the phenotypic and molecular characteristics in a multiplex family with autism and related phenotypes. We did not identify CNVs or sequence mutations with high probability of being pathogenic enough to be causal for autism. However, we present several genetic variations with possible influence on the genetic etiology of autism and related phenotypes in this family. To understand the possible implications of the identified variants, functional analyses of these variants are warranted.

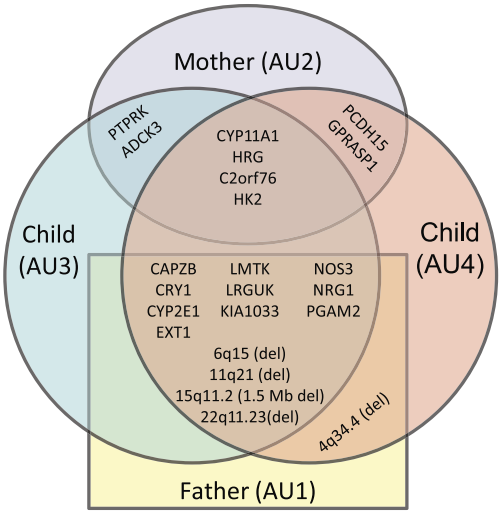

Figure 3. Inheritance of the point mutations and CNVs in the mother, father and the two index children. Identified rare point mutations are indicated by the gene they are located in. CNVs are described by the chromosomal region they are located in.

CONCLUDING REMARKS

Autism is characterized by varying degrees of difficulties in social communication and repetitive behavior. Within autism, there is a large clinical heterogeneity of the core characteristics ranging from mildly to profoundly disabling. The varying degrees within the spectrum have been shown to be on a continuum, where milder forms of these symptoms are displayed as autistic-like traits (ALTs) in the general population. A person can display ALTs related to one or more of the core domains that may be viewed as normally occurring traits in the general population.

In addition to the clinical heterogeneity of autism, the genetic etiology of autism has also been shown to be highly heterogeneous. A large number of genes have been identified as high impact or susceptibility genes for autism with varying confidence. The majority of the identified rare high impact variants for autism have each only been identified in a few cases. However, many of the disrupted genes have been identified in shared biological pathways. In addition to the rare variants identified, numerous common variations have been associated with autism and also ALTs. Although each common variation usually has a small effect size, they have been shown to have a big impact *en masse*.

Since it has been difficult to replicate common variant findings in autism case-control studies, investigations of endophenotypes may be another strategy to gain insight into the influences of common variations in autism. We investigated common variations in relation to the dimensionally distributed autistic-like traits in our analyses, where we selected common variants in genes that have been repeatedly associated with autism. In our studies, we have identified association between common variation in the *ASMT* gene and ALTs; however, the selected variants in other autism candidate genes were not associated with ALTs. Either these variants do not have an effect on the ALTs, or the effect sizes are too small to be identified in our sample.

Investigations of rare genetic variation were performed in case-control or family based studies in this thesis. Functional investigations of identified mutations are generally needed to understand their impact on the disorder. One of our identified rare mutations in the *ASMT* gene has previously been shown to drastically decrease melatonin levels in carriers in functional studies. In addition, in the family based study we could identify several rare variations that could potentially increase the risk for autism-related problems. Of these identified variants, our functional analysis of a mutation in the *CYP11A1* gene was implied to increase the function of the enzyme

P450scc (encoded by the *CYP11A1* gene). However, to understand the possible clinical implications of the other identified mutations in our study, functional studies of these genetic variations are warranted.

Taken together, in our studies we have identified several genetic variations acting as risk variants for autism and related phenotypes. Our analyses of rare genetic variation in *ASMT* may provide insight for the use of melatonin treatments in autism. Investigations of rare inherited variation may gain further insight on the underlying biological mechanism for autism. Overall, the field of autism genetics has identified a large number of genes in different biological pathways that can be potential targets for autism treatment strategies in the future.

SAMMANFATTNING PÅ SVENSKA

Personer med autism har problem med social kommunikation, språkproblem och stereotypa beteenden. Genetiska faktorer har visat sig vara av stor betydelse för autismspektrumtillstånd (används synonymt med autism här), men exakt vilka gener som är orsakande är till stor del fortfarande okänt. I den här avhandlingen ingår fyra studier där vi har fokuserat på att både fördjupa kunskapen om gener som tidigare har visat sig vara av betydelse för autism och att hitta nya gener som kan vara av betydelse för autism. Vi har även undersökt geners inverkan på autismliknande drag i normalbefolkningen.

I två av de ingående studierna (*paper I and II*) har vi fokuserat på gener som är relaterade till hormonet melatonins funktion och syntes, med ett särskilt fokus på genen *ASMT* som kodar för ett protein som reglerar tillverkningen av melatonin. Melatonin är ett hormon som produceras under natten och är bland annat inblandat i sömn- och vakenhetsregleringen. Tidigare studier har visat sänkta nivåer av melatonin hos barn med autism. Melatonin har också ofta positiva effekter när det används farmakologiskt till barn med autism. Genom att undersöka ovanliga mutationer i melatonin-relaterade gener har vi i *ASMT*-genen identifierat en mutation som tidigare har visats orsaka en försämrad produktion av melatonin (*paper I*). Vi har även funnit association mellan en vanlig genetisk variation, så kallad ”single nucleotide polymorphism” (SNP), och autismliknande drag i normalbefolkningen (*paper II*). Sammantaget stärker våra studier ett samband mellan melatoninrelaterade gener och autism.

Ett stort antal gener har identifierats som riskgener för autism och vi har valt att undersöka om genetisk variation i fem gener (*CNTNAP2*, *RELN*, *SHANK3* och rs4307059) ökar risken för autismliknande drag i normalbefolkningen (*paper III*). En av de studerade SNParna (rs4307059), som är lokaliserad mellan två gener som heter cadherin 9 (CDH9) och cadherin 10 (CDH10), har tidigare associerats med både autism och problem med social kommunikation. De andra generna är några av de gener som har en tydlig association till autism, och är inblandade i neuroutveckling (genen *RELN*) och synapsfunktion (*CNTNAP2* och *SHANK3*). Våra analyser tyder inte på att de studerade genetiska variationerna i de här generna ökar risken för autismliknande drag i normalbefolkningen.

För att förutsättningslöst undersöka genetiska faktorer för autism i en stor familj som har flera individer med autism och/eller språkproblem har vi använt oss av så kallade

helgenomsundersökningar, där man kan screenar hela genomet för ovanliga mutationer. I vår studie har vi studerat ovanliga punktmutationer i exomet, som är DNA-sekvenser som kodar för proteiner, via så kallad exomsekvensering. Vi har även studerat så kallade kopienummervariationer (via SNP-arrays). Kopienummervariationer (CNVs) innebär att sekvenser i arvsmassan finns i för många eller för få kopior eller saknas helt. Vi identifierade ett flertal ovanliga CNVs och punktmutationer, som kan vara av betydelse för autism i den studerade familjen. Vi valde att studera funktionen av en av de identifierade punktmutationerna i en gen som heter *CYP11A1*, som kodar för ett protein som reglerar det första steget i steroidhormonsyntesen där slutprodukterna bland annat är steroidhormonerna kortisol, testosteron och östrogen. Våra analyser indikerar att mutationen kan orsaka en ökad funktion hos proteinet som *CYP11A1* kodar för, men för att veta om det har en betydelse i individer som bär på mutationen så behöver vi också mäta steroidhormonnivåer i de individerna.

Sammantaget har vi sett en inblandning av melatoninrelaterade gener vid autism och vi har identifierat en rad olika gener och CNVs som kan vara riskgener för autism. Vår forskning bidrar till en ökad kunskap om gener som kan vara inblandade vid autism, vilket ökar förståelsen för bakomliggande orsaker och kan leda till bättre behandlingsstrategier i framtiden.

ACKNOWLEDGEMENT

I would like to express my gratitude to all the people who have contributed, in one way or another, to this thesis. Special thanks to:

My supervisor Jonas Melke for introducing me to scientific research and for all the support during these years. Thanks for all the discussions about everything from genetics to Rolands 505:a. It has been a great pleasure to work with you.

Lars Westberg for all the research discussions and the help with this thesis.

My co-supervisor Tommy Martinsson for the help with genetic analyses guidance.

Daniel Hovey, Sara Karlsson and Anna Zettergren for the great discussions and coffee/tea breaks in our room. Thanks to Daniel and Sara for all the help with writing this thesis.

Collaborators Paul Lichtenstein, Sebastian Lundström, Henrik Anckarsäter, Erik Pettersson, Carmela Miniscalco, Mats Johnson, Christopher Gillberg, Mikael Landén, Kent Thuresson and Carina Algede for all your expertise and good collaborations.

Gunilla Bourghardt, Ann-Christin Reinhold and Inger Oscarsson for all the help with the work in the lab.

The genomics core facility staff for all the help with my genetic analyses.

Michael, Daniel, Martin and Marie for the help with my cell cultures.

Colleagues and friends at the department of pharmacology for the many social activities, political discussions and lunches at lyktan: Erik, Martin, Jakob, Robert, Nina, Daniel H, Sara, Camilla, Ghazaleh, Anna, Melker, Daniel E, Richard, Patrik, Fredrik, Alana, Caroline, Elisabeth and Petra.

Britt-Marie Benbow and Anna-Lena Carlred for the administrative help.

All co-workers at the department of pharmacology for making it a great workplace.

Thanks to all my friends and large family for all the support, it would not have been possible without you. Sara, Gärda and Emma for always being there. Kim Kelly is my friend: Helene, Helena and Malin for all the great times in repan. Steve for all your english help.

Margareta, Ingemar, Sara and Jacob for always supporting and inspiring me.

REFERENCES

1. APA (2000) *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. American Psychiatric Association, Washington, DC.
2. Zwaigenbaum, L., Bryson, S. & Garon, N. (2013) Early identification of autism spectrum disorders. *Behav Brain Res* **251** 133-146.
3. Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver, L., Chawarska, K., Constantino, J., Dawson, G., Dobkins, K., Fein, D., Iverson, J., Klin, A., Landa, R., Messinger, D., Ozonoff, S., Sigman, M., Stone, W., Tager-Flusberg, H., *et al* (2009) Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. *Pediatrics* **123** 1383-1391.
4. Tager-Flusberg, H. & Joseph, R.M. (2003) Identifying neurocognitive phenotypes in autism. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **358** 303-314.
5. Wing, L. & Potter, D. (2002) The epidemiology of autistic spectrum disorders: is the prevalence rising? *Ment Retard Dev Disabil Res Rev* **8** 151-161.
6. Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D. & Charman, T. (2006) Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet* **368** 210-215.
7. Fernell, E. & Gillberg, C. (2010) Autism spectrum disorder diagnoses in Stockholm preschoolers. *Res Dev Disabil* **31** 680-685.
8. Elsabbagh, M., Divan, G., Koh, Y.J., Kim, Y.S., Kauchali, S., Marcin, C., Montiel-Nava, C., Patel, V., Paula, C.S., Wang, C., Yasamy, M.T. & Fombonne, E. (2012) Global prevalence of autism and other pervasive developmental disorders. *Autism Res* **5** 160-179.
9. Gillberg, C. (2010) The ESSENCE in child psychiatry: Early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations. *Res Dev Disabil* **31** 1543-1551.
10. Simonoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T. & Baird, G. (2008) Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *J Am Acad Child Adolesc Psychiatry* **47** 921-929.
11. Kanner, L. (1968) Autistic disturbances of affective contact. *Acta paedopsychiatrica* **35** 100-136.
12. Wing, L. (1981) Asperger's syndrome: a clinical account. *Psychol Med* **11** 115-129.
13. APA (2013) *Diagnostic and statistical manual of mental disorders: DSM-5. American Psychiatric Association.*
14. Hazen, E.P., McDougle, C.J. & Volkmar, F.R. (2013) Changes in the diagnostic criteria for autism in DSM-5: controversies and concerns. *J Clin Psychiatry* **74** 739-740.
15. Wing, L. (1988) The continuum of autistic characteristics. In Schopler, E. & Mesibov, G. (eds), *Diagnosis and Assessment in Autism*. Plenum, New York, pp. 91-110.
16. Wing, L. & Gould, J. (1979) Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. *J Autism Dev Disord* **9** 11-29.
17. Constantino, J.N. & Todd, R.D. (2003) Autistic traits in the general population: a twin study. *Arch Gen Psychiatry* **60** 524-530.
18. Ronald, A., Happe, F., Price, T.S., Baron-Cohen, S. & Plomin, R. (2006) Phenotypic and genetic overlap between autistic traits at the extremes of the general population. *J Am Acad Child Adolesc Psychiatry* **45** 1206-1214.

19. Lundstrom, S., Chang, Z., Rastam, M., Gillberg, C., Larsson, H., Anckarsater, H. & Lichtenstein, P. (2012) Autism Spectrum Disorders and Autisticlike Traits Similar Etiology in the Extreme End and the Normal Variation. *Arch Gen Psychiat* **69** 46-52.
20. Constantino, J.N. & Todd, R.D. (2005) Intergenerational transmission of subthreshold autistic traits in the general population. *Biol Psychiatry* **57** 655-660.
21. Geschwind, D.H. (2008) Autism: many genes, common pathways? *Cell* **135** 391-395.
22. Larson, T., Anckarsater, H., Gillberg, C., Stahlberg, O., Carlstrom, E., Kadesjo, B., Rastam, M. & Lichtenstein, P. (2010) The autism--tics, AD/HD and other comorbidities inventory (A-TAC): further validation of a telephone interview for epidemiological research. *BMC Psychiatry* **10** 1.
23. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. (2001) The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J Autism Dev Disord* **31** 5-17.
24. Rutter, M. (1968) Concepts of autism: a review of research. *J Child Psychol Psychiatry* **9** 1-25.
25. Folstein, S. & Rutter, M. (1977) Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry* **18** 297-321.
26. Bromley, R.L., Mawer, G.E., Briggs, M., Cheyne, C., Clayton-Smith, J., Garcia-Finana, M., Kneen, R., Lucas, S.B., Shallcross, R. & Baker, G.A. (2013) The prevalence of neurodevelopmental disorders in children prenatally exposed to antiepileptic drugs. *Journal of neurology, neurosurgery, and psychiatry* **84** 637-643.
27. Bandim, J.M., Ventura, L.O., Miller, M.T., Almeida, H.C. & Costa, A.E. (2003) Autism and Mobius sequence: an exploratory study of children in northeastern Brazil. *Arquivos de neuro-psiquiatria* **61** 181-185.
28. Berger, B.E., Navar-Boggan, A.M. & Omer, S.B. (2011) Congenital rubella syndrome and autism spectrum disorder prevented by rubella vaccination--United States, 2001-2010. *BMC Public Health* **11** 340.
29. Gillberg, C., Cederlund, M., Lamberg, K. & Zeijlon, L. (2006) Brief report: "the autism epidemic". The registered prevalence of autism in a Swedish urban area. *J Autism Dev Disord* **36** 429-435.
30. Dutheil, F., Beaune, P., Tzourio, C., Loriot, M.A. & Elbaz, A. (2010) Interaction between ABCB1 and professional exposure to organochlorine insecticides in Parkinson disease. *Arch Neurol* **67** 739-745.
31. McGrath, J.J., Mortensen, P.B., Visscher, P.M. & Wray, N.R. (2013) Where GWAS and epidemiology meet: opportunities for the simultaneous study of genetic and environmental risk factors in schizophrenia. *Schizophrenia bulletin* **39** 955-959.
32. Marrus, N., Underwood-Riordan, H., Randall, F., Zhang, Y. & Constantino, J.N. (2014) Lack of effect of risperidone on core autistic symptoms: data from a longitudinal study. *J Child Adolesc Psychopharmacol* **24** 513-518.
33. Ching, H. & Pringsheim, T. (2012) Aripiprazole for autism spectrum disorders (ASD). *Cochrane Database Syst Rev* **5** CD009043.
34. Ehninger, D., Han, S., Shilyansky, C., Zhou, Y., Li, W., Kwiatkowski, D.J., Ramesh, V. & Silva, A.J. (2008) Reversal of learning deficits in a Tsc2+/- mouse model of tuberous sclerosis. *Nat Med* **14** 843-848.
35. Ehninger, D. & Silva, A.J. (2011) Rapamycin for treating Tuberous sclerosis and Autism spectrum disorders. *Trends in molecular medicine* **17** 78-87.
36. Doyen, C., Mighiu, D., Kaye, K., Colineaux, C., Beaumanoir, C., Mouraeff, Y., Rieu, C., Paubel, P. & Contejean, Y. (2011) Melatonin in children with autistic spectrum disorders: recent and practical data. *Eur Child Adolesc Psychiatry* **20** 231-239.

37. Watson, J.D. & Crick, F.H. (1953) The structure of DNA. *Cold Spring Harb Symp Quant Biol* **18** 123-131.
38. Tjio, J.H. & Levan, A. (1956) The Chromosome Number of Man. *Hereditas* **42** U1-6.
39. Helena Mangs, A. & Morris, B.J. (2007) The Human Pseudoautosomal Region (PAR): Origin, Function and Future. *Current genomics* **8** 129-136.
40. Feuk, L., Carson, A.R. & Scherer, S.W. (2006) Structural variation in the human genome. *Nat Rev Genet* **7** 85-97.
41. Kidd, J.M., Cooper, G.M., Donahue, W.F., Hayden, H.S., Sampas, N., Graves, T., Hansen, N., Teague, B., Alkan, C., Antonacci, F., Haugen, E., Zerr, T., Yamada, N.A., Tsang, P., Newman, T.L., Tuzun, E., Cheng, Z., Ebling, H.M., Tusneem, N., *et al* (2008) Mapping and sequencing of structural variation from eight human genomes. *Nature* **453** 56-64.
42. Sanger, F., Nicklen, S. & Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74** 5463-5467.
43. Venter, J.C. & Adams, M.D. & Myers, E.W. & Li, P.W. & Mural, R.J. & Sutton, G.G. & Smith, H.O. & Yandell, M. & Evans, C.A. & Holt, R.A. & Gocayne, J.D. & Amanatides, P. & Ballew, R.M. & Huson, D.H. & Wortman, J.R. & Zhang, Q. & Kodira, C.D. & Zheng, X.H. & Chen, L., *et al* (2001) The sequence of the human genome. *Science* **291** 1304-1351.
44. Lander, E.S. & Linton, L.M. & Birren, B. & Nusbaum, C. & Zody, M.C. & Baldwin, J. & Devon, K. & Dewar, K. & Doyle, M. & FitzHugh, W. & Funke, R. & Gage, D. & Harris, K. & Heaford, A. & Howland, J. & Kann, L. & Lehoczky, J. & LeVine, R. & McEwan, P., *et al* (2001) Initial sequencing and analysis of the human genome. *Nature* **409** 860-921.
45. Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M., Chi, M., Navin, N., Lucito, R., Healy, J., Hicks, J., Ye, K., Reiner, A., Gilliam, T.C., Trask, B., Patterson, N., *et al* (2004) Large-scale copy number polymorphism in the human genome. *Science* **305** 525-528.
46. Iafrate, A.J., Feuk, L., Rivera, M.N., Listewnik, M.L., Donahoe, P.K., Qi, Y., Scherer, S.W. & Lee, C. (2004) Detection of large-scale variation in the human genome. *Nat Genet* **36** 949-951.
47. Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W., Cho, E.K., Dallaire, S., Freeman, J.L., Gonzalez, J.R., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J.R., *et al* (2006) Global variation in copy number in the human genome. *Nature* **444** 444-454.
48. Tenesa, A. & Haley, C.S. (2013) The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet* **14** 139-149.
49. Berg, J.M. & Geschwind, D.H. (2012) Autism genetics: searching for specificity and convergence. *Genome biology* **13** 247.
50. Betancur, C. (2011) Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Res* **1380** 42-77.
51. Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.H., Narzisi, G., Leotta, A., Kendall, J., Grabowska, E., Ma, B., Marks, S., Rodgers, L., Stepansky, A., Troge, J., Andrews, P., Bekritsky, M., *et al* (2012) De novo gene disruptions in children on the autistic spectrum. *Neuron* **74** 285-299.
52. Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., Walker, M.F., Ober, G.T., Teran, N.A., Song, Y., El-Fishawy, P., Murtha, R.C., Choi, M., Overton, J.D., Bjornson, R.D., *et al* (2012) De novo

mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485** 237-241.

53. Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E. & Rutter, M.L. (1995) Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* **25** 63-77.
54. Steffenburg, S., Gillberg, C., Hellgren, L., Andersson, L., Gillberg, I.C., Jakobsson, G. & Bohman, M. (1989) A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. *J Child Psychol Psychiatry* **30** 405-416.
55. Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., Miller, J., Fedele, A., Collins, J., Smith, K., Lotspeich, L., Croen, L.A., Ozonoff, S., Lajonchere, C., Grether, J.K. & Risch, N. (2011) Genetic heritability and shared environmental factors among twin pairs with autism. *Arch Gen Psychiatry* **68** 1095-1102.
56. Rosenberg, R.E., Law, J.K., Yenokyan, G., McGready, J., Kaufmann, W.E. & Law, P.A. (2009) Characteristics and concordance of autism spectrum disorders among 277 twin pairs. *Arch Pediatr Adolesc Med* **163** 907-914.
57. Lichtenstein, P., Carlstrom, E., Rastam, M., Gillberg, C. & Anckarsater, H. (2010) The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. *Am J Psychiatry* **167** 1357-1363.
58. Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Larsson, H., Hultman, C.M. & Reichenberg, A. (2014) The familial risk of autism. *JAMA* **311** 1770-1777.
59. Ronald, A. & Hoekstra, R.A. (2011) Autism spectrum disorders and autistic traits: a decade of new twin studies. *Am J Med Genet B Neuropsychiatr Genet* **156B** 255-274.
60. Constantino, J.N., Todorov, A., Hilton, C., Law, P., Zhang, Y., Molloy, E., Fitzgerald, R. & Geschwind, D. (2013) Autism recurrence in half siblings: strong support for genetic mechanisms of transmission in ASD. *Mol Psychiatry* **18** 137-138.
61. Gronborg, T.K., Schendel, D.E. & Parner, E.T. (2013) Recurrence of autism spectrum disorders in full- and half-siblings and trends over time: a population-based cohort study. *JAMA pediatrics* **167** 947-953.
62. Hoekstra, R.A., Bartels, M., Verweij, C.J. & Boomsma, D.I. (2007) Heritability of autistic traits in the general population. *Arch Pediatr Adolesc Med* **161** 372-377.
63. Constantino, J.N., Hudziak, J.J. & Todd, R.D. (2003) Deficits in reciprocal social behavior in male twins: evidence for a genetically independent domain of psychopathology. *J Am Acad Child Adolesc Psychiatry* **42** 458-467.
64. Robinson, E.B., Koenen, K.C., McCormick, M.C., Munir, K., Hallett, V., Happe, F., Plomin, R. & Ronald, A. (2011) Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Arch Gen Psychiatry* **68** 1113-1121.
65. Lundstrom, S., Reichenberg, A., Melke, J., Rastam, M., Kerekes, N., Lichtenstein, P., Gillberg, C. & Anckarsater, H. (2014) Autism spectrum disorders and coexisting disorders in a nationwide Swedish twin study. *J Child Psychol Psychiatry*.
66. Cristino, A.S., Williams, S.M., Hawi, Z., An, J.Y., Bellgrove, M.A., Schwartz, C.E., Costa Lda, F. & Claudianos, C. (2014) Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. *Mol Psychiatry* **19** 294-301.
67. Lee, S.H. & Ripke, S. & Neale, B.M. & Faraone, S.V. & Purcell, S.M. & Perlis, R.H. & Mowry, B.J. & Thapar, A. & Goddard, M.E. & Witte, J.S. & Absher, D. & Agartz, I. & Akil, H. & Amin, F. & Andreassen, O.A. & Anjorin, A. & Anney, R. & Anttila, V. & Arking, D.E., *et al* (2013) Genetic

relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* **45** 984-994.

68. Gratten, J., Wray, N.R., Keller, M.C. & Visscher, P.M. (2014) Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nat Neurosci* **17** 782-790.
69. Stein, J.L., Parikshak, N.N. & Geschwind, D.H. (2013) Rare inherited variation in autism: beginning to see the forest and a few trees. *Neuron* **77** 209-211.
70. Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., Smith, J.D., Paeper, B., Nickerson, D.A., Dea, J., Dong, S., Gonzalez, L.E., Mandell, J.D., Mane, S.M., Murtha, M.T., *et al* (2014) The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **515** 216-221.
71. Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., Leotta, A., Pai, D., Zhang, R., Lee, Y.H., Hicks, J., Spence, S.J., Lee, A.T., Puura, K., Lehtimaki, T., *et al* (2007) Strong association of de novo copy number mutations with autism. *Science* **316** 445-449.
72. Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y., Thiruvahindrapduram, B., Fiebig, A., Schreiber, S., Friedman, J., Ketelaars, C.E., Vos, Y.J., Ficicioglu, C., Kirkpatrick, S., Nicolson, R., *et al* (2008) Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* **82** 477-488.
73. Zhao, X., Leotta, A., Kustanovich, V., Lajonchere, C., Geschwind, D.H., Law, K., Law, P., Qiu, S., Lord, C., Sebat, J., Ye, K. & Wigler, M. (2007) A unified genetic theory for sporadic and inherited autism. *Proc Natl Acad Sci U S A* **104** 12831-12836.
74. Pinto, D. & Pagnamenta, A.T. & Klei, L. & Anney, R. & Merico, D. & Regan, R. & Conroy, J. & Magalhaes, T.R. & Correia, C. & Abrahams, B.S. & Almeida, J. & Bacchelli, E. & Bader, G.D. & Bailey, A.J. & Baird, G. & Battaglia, A. & Berney, T. & Bolshakova, N. & Bolte, S., *et al* (2010) Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466** 368-372.
75. Bernier, R., Gerdts, J., Munson, J., Dawson, G. & Estes, A. (2012) Evidence for broader autism phenotype characteristics in parents from multiple-incidence autism families. *Autism Res* **5** 13-20.
76. Klei, L., Sanders, S.J., Murtha, M.T., Hus, V., Lowe, J.K., Willsey, A.J., Moreno-De-Luca, D., Yu, T.W., Fombonne, E., Geschwind, D., Grice, D.E., Ledbetter, D.H., Lord, C., Mane, S.M., Martin, C.L., Martin, D.M., Morrow, E.M., Walsh, C.A., Melhem, N.M., *et al* (2012) Common genetic variants, acting additively, are a major source of risk for autism. *Mol Autism* **3** 9.
77. Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F.P. & et al. (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65** 905-914.
78. Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U. & Zoghbi, H.Y. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* **23** 185-188.
79. Consortium, T.E.C.T.S. (1993) Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell* **75** 1305-1315.
80. Szatmari, P. & Paterson, A.D. & Zwaigenbaum, L. & Roberts, W. & Brian, J. & Liu, X.Q. & Vincent, J.B. & Skaug, J.L. & Thompson, A.P. & Senman, L. & Feuk, L. & Qian, C. & Bryson, S.E. & Jones, M.B. & Marshall, C.R. & Scherer, S.W. & Vieland, V.J. & Bartlett, C. & Mangin, L.V., *et al* (2007)

Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* **39** 319-328.

81. Sanders, S.J., Ercan-Sencicek, A.G., Hus, V., Luo, R., Murtha, M.T., Moreno-De-Luca, D., Chu, S.H., Moreau, M.P., Gupta, A.R., Thomson, S.A., Mason, C.E., Bilguvar, K., Celestino-Soper, P.B., Choi, M., Crawford, E.L., Davis, L., Wright, N.R., Dhodapkar, R.M., DiCola, M., *et al* (2011) Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* **70** 863-885.
82. Levy, D., Ronemus, M., Yamrom, B., Lee, Y.H., Leotta, A., Kendall, J., Marks, S., Lakshmi, B., Pai, D., Ye, K., Buja, A., Krieger, A., Yoon, S., Troge, J., Rodgers, L., Iossifov, I. & Wigler, M. (2011) Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron* **70** 886-897.
83. Pinto, D. & Delaby, E. & Merico, D. & Barbosa, M. & Merikangas, A. & Klei, L. & Thiruvahindrapuram, B. & Xu, X. & Ziman, R. & Wang, Z. & Vorstman, J.A. & Thompson, A. & Regan, R. & Pilorge, M. & Pellecchia, G. & Pagnamenta, A.T. & Oliveira, B. & Marshall, C.R. & Magalhaes, T.R., *et al* (2014) Convergence of Genes and Cellular Pathways Dysregulated in Autism Spectrum Disorders. *Am J Hum Genet*.
84. Dykens, E.M., Sutcliffe, J.S. & Levitt, P. (2004) Autism and 15q11-q13 disorders: behavioral, genetic, and pathophysiological issues. *Ment Retard Dev Disabil Res Rev* **10** 284-291.
85. Hogart, A., Wu, D., LaSalle, J.M. & Schanen, N.C. (2010) The comorbidity of autism with the genomic disorders of chromosome 15q11.2-q13. *Neurobiol Dis* **38** 181-191.
86. Bundey, S., Hardy, C., Vickers, S., Kilpatrick, M.W. & Corbett, J.A. (1994) Duplication of the 15q11-13 region in a patient with autism, epilepsy and ataxia. *Dev Med Child Neurol* **36** 736-742.
87. Leblond, C.S., Heinrich, J., Delorme, R., Proepper, C., Betancur, C., Huguet, G., Konyukh, M., Chaste, P., Ey, E., Rastam, M., Anckarsater, H., Nygren, G., Gillberg, I.C., Melke, J., Toro, R., Regnault, B., Fauchereau, F., Mercati, O., Lemiere, N., *et al* (2012) Genetic and functional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders. *PLoS Genet* **8** e1002521.
88. Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M.A., Green, T., Platt, O.S., Ruderfer, D.M., Walsh, C.A., Altshuler, D., Chakravarti, A., Tanzi, R.E., Stefansson, K., Santangelo, S.L., Gusella, J.F., *et al* (2008) Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* **358** 667-675.
89. Kumar, R.A., KaraMohamed, S., Sudi, J., Conrad, D.F., Brune, C., Badner, J.A., Gilliam, T.C., Nowak, N.J., Cook, E.H., Jr., Dobyns, W.B. & Christian, S.L. (2008) Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* **17** 628-638.
90. Kishino, T., Lalande, M. & Wagstaff, J. (1997) UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet* **15** 70-73.
91. Prader, A., Labhart, A. & Willi, H. (1956) Ein Syndrom von Adipositas, Kleinwuchs, Kryptorchismus und Oligophrenie nach Myatonieartigem Zustand im Neugeborenenalter. *Schweiz Med Wschr* **86** 1260-1261.
92. Jamain, S., Quach, H., Betancur, C., Rastam, M., Colineaux, C., Gillberg, I.C., Soderstrom, H., Giros, B., Leboyer, M., Gillberg, C. & Bourgeron, T. (2003) Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet* **34** 27-29.
93. Strauss, K.A., Puffenberger, E.G., Huentelman, M.J., Gottlieb, S., Dobrin, S.E., Parod, J.M., Stephan, D.A. & Morton, D.H. (2006) Recessive symptomatic focal epilepsy and mutant contactin-associated protein-like 2. *N Engl J Med* **354** 1370-1377.
94. Bourgeron, T. (2009) A synaptic trek to autism. *Curr Opin Neurobiol* **19** 231-234.

95. Huguet, G., Ey, E. & Bourgeron, T. (2013) The genetic landscapes of autism spectrum disorders. *Annual review of genomics and human genetics* **14** 191-213.
96. Chen, J., Yu, S., Fu, Y. & Li, X. (2014) Synaptic proteins and receptors defects in autism spectrum disorders. *Frontiers in cellular neuroscience* **8** 276.
97. Girirajan, S., Rosenfeld, J.A., Cooper, G.M., Antonacci, F., Siswara, P., Itsara, A., Vives, L., Walsh, T., McCarthy, S.E., Baker, C., Mefford, H.C., Kidd, J.M., Browning, S.R., Browning, B.L., Dickel, D.E., Levy, D.L., Ballif, B.C., Platky, K., Farber, D.M., *et al* (2010) A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet* **42** 203-209.
98. O'Roak, B.J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J.J., Girirajan, S., Karakoc, E., Mackenzie, A.P., Ng, S.B., Baker, C., Rieder, M.J., Nickerson, D.A., Bernier, R., Fisher, S.E., Shendure, J. & Eichler, E.E. (2012) Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet* **44** 471.
99. Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., Lin, C.F., Stevens, C., Wang, L.S., Makarov, V., Polak, P., Yoon, S., Maguire, J., Crawford, E.L., Campbell, N.G., Geller, E.T., Valladares, O., Schafer, C., Liu, H., *et al* (2012) Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **485** 242-245.
100. Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., Wong, W.S., Sigurdsson, G., Walters, G.B., Steinberg, S., Helgason, H., Thorleifsson, G., Gudbjartsson, D.F., Helgason, A., Magnusson, O.T., Thorsteinsdottir, U., *et al* (2012) Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488** 471-475.
101. Gaugler, T., Klei, L., Sanders, S.J., Bodea, C.A., Goldberg, A.P., Lee, A.B., Mahajan, M., Manaa, D., Pawitan, Y., Reichert, J., Ripke, S., Sandin, S., Sklar, P., Svantesson, O., Reichenberg, A., Hultman, C.M., Devlin, B., Roeder, K. & Buxbaum, J.D. (2014) Most genetic risk for autism resides with common variation. *Nat Genet* **46** 881-885.
102. St Pourcain, B., Whitehouse, A.J., Ang, W.Q., Warrington, N.M., Glessner, J.T., Wang, K., Timpson, N.J., Evans, D.M., Kemp, J.P., Ring, S.M., McArdle, W.L., Golding, J., Hakonarson, H., Pennell, C.E. & Smith, G.D. (2013) Common variation contributes to the genetic architecture of social communication traits. *Mol Autism* **4** 34.
103. Chakrabarti, B., Dudbridge, F., Kent, L., Wheelwright, S., Hill-Cawthorne, G., Allison, C., Banerjee-Basu, S. & Baron-Cohen, S. (2009) Genes related to sex steroids, neural growth, and social-emotional behavior are associated with autistic traits, empathy, and Asperger syndrome. *Autism Res* **2** 157-177.
104. Zettergren, A., Jonsson, L., Johansson, D., Melke, J., Lundstrom, S., Anckarsater, H., Lichtenstein, P. & Westberg, L. (2013) Associations between polymorphisms in sex steroid related genes and autistic-like traits. *Psychoneuroendocrinology* **38** 2575-2584.
105. Henningsson, S., Jonsson, L., Ljunggren, E., Westberg, L., Gillberg, C., Rastam, M., Anckarsater, H., Nygren, G., Landen, M., Thuresson, K., Betancur, C., Leboyer, M., Eriksson, E. & Melke, J. (2009) Possible association between the androgen receptor gene and autism spectrum disorder. *Psychoneuroendocrinology* **34** 752-761.
106. Auyeung, B., Taylor, K., Hackett, G. & Baron-Cohen, S. (2010) Foetal testosterone and autistic traits in 18 to 24-month-old children. *Mol Autism* **1** 11.
107. Spratt, E.G., Nicholas, J.S., Brady, K.T., Carpenter, L.A., Hatcher, C.R., Meekins, K.A., Furlanetto, R.W. & Charles, J.M. (2012) Enhanced cortisol response to stress in children in autism. *J Autism Dev Disord* **42** 75-81.

108. Knickmeyer, R., Baron-Cohen, S., Fane, B.A., Wheelwright, S., Mathews, G.A., Conway, G.S., Brook, C.G. & Hines, M. (2006) Androgens and autistic traits: A study of individuals with congenital adrenal hyperplasia. *Horm Behav* **50** 148-153.
109. Knickmeyer, R., Baron-Cohen, S., Raggatt, P. & Taylor, K. (2005) Foetal testosterone, social relationships, and restricted interests in children. *J Child Psychol Psychiatry* **46** 198-210.
110. Baron-Cohen, S., Auyeung, B., Norgaard-Pedersen, B., Hougaard, D.M., Abdallah, M.W., Melgaard, L., Cohen, A.S., Chakrabarti, B., Ruta, L. & Lombardo, M.V. (2014) Elevated fetal steroidogenic activity in autism. *Mol Psychiatry*.
111. Geier, D.A., Kern, J.K., King, P.G., Sykes, L.K. & Geier, M.R. (2012) An evaluation of the role and treatment of elevated male hormones in autism spectrum disorders. *Acta neurobiologiae experimentalis* **72** 1-17.
112. Wang, K., Zhang, H., Ma, D., Bucan, M., Glessner, J.T., Abrahams, B.S., Salyakina, D., Imielinski, M., Bradfield, J.P., Sleiman, P.M., Kim, C.E., Hou, C., Frackelton, E., Chiavacci, R., Takahashi, N., Sakurai, T., Rappaport, E., Lajonchere, C.M., Munson, J., *et al* (2009) Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **459** 528-533.
113. Weiss, L.A., Arking, D.E., Daly, M.J. & Chakravarti, A. (2009) A genome-wide linkage and association scan reveals novel loci for autism. *Nature* **461** 802-808.
114. Anney, R. & Klei, L. & Pinto, D. & Regan, R. & Conroy, J. & Magalhaes, T.R. & Correia, C. & Abrahams, B.S. & Sykes, N. & Pagnamenta, A.T. & Almeida, J. & Bacchelli, E. & Bailey, A.J. & Baird, G. & Battaglia, A. & Berney, T. & Bolshakova, N. & Bolte, S. & Bolton, P.F., *et al* (2010) A genome-wide scan for common alleles affecting risk for autism. *Hum Mol Genet* **19** 4072-4082.
115. Anney, R., Ripke, S., Arking, D.E., Cook, E., Robinson, E.B., Anttila, V., Gill, M., Santangelo, S.L., Sutcliffe, J., Daly, M. & Devlin, B. (2013) Meta-analysis of European ancestry individuals with autism spectrum disorder reveals strong association 3' of the Astrotactin 2 (ASTN2) gene locus on chromosome 9 *XXIst World Congress of Psychiatric Genetics: Redefining Mental Illness Through Genetics*. Boston.
116. Stefansson, H., Meyer-Lindenberg, A., Steinberg, S., Magnusdottir, B., Morgen, K., Arnarsdottir, S., Bjornsdottir, G., Walters, G.B., Jonsdottir, G.A., Doyle, O.M., Tost, H., Grimm, O., Kristjansdottir, S., Snorrason, H., Davidsdottir, S.R., Gudmundsson, L.J., Jonsson, G.F., Stefansdottir, B., Helgadottir, I., *et al* (2014) CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature* **505** 361-366.
117. Consortium, C.-D.G.o.t.P.G. (2013) Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381** 1371-1379.
118. Splawski, I., Timothy, K.W., Sharpe, L.M., Decher, N., Kumar, P., Bloise, R., Napolitano, C., Schwartz, P.J., Joseph, R.M., Condouris, K., Tager-Flusberg, H., Priori, S.G., Sanguinetti, M.C. & Keating, M.T. (2004) Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell* **119** 19-31.
119. Ronald, A., Butcher, L.M., Docherty, S., Davis, O.S., Schalkwyk, L.C., Craig, I.W. & Plomin, R. (2010) A genome-wide association study of social and non-social autistic-like traits in the general population using pooled DNA, 500 K SNP microarrays and both community and diagnosed autism replication samples. *Behav Genet* **40** 31-45.
120. Jones, R.M., Cadby, G., Melton, P.E., Abraham, L.J., Whitehouse, A.J. & Moses, E.K. (2013) Genome-wide association study of autistic-like traits in a general population study of young adults. *Frontiers in human neuroscience* **7** 658.

121. St Pourcain, B., Skuse, D.H., Mandy, W.P., Wang, K., Hakonarson, H., Timpson, N.J., Evans, D.M., Kemp, J.P., Ring, S.M., McArdle, W.L., Golding, J. & Smith, G.D. (2014) Variability in the common genetic architecture of social-communication spectrum phenotypes during childhood and adolescence. *Mol Autism* **5** 18.
122. Whitehouse, A.J., Bishop, D.V., Ang, Q.W., Pennell, C.E. & Fisher, S.E. (2011) CNTNAP2 variants affect early language development in the general population. *Genes Brain Behav* **10** 451-456.
123. Lerner, A.B., Case, J.D. & Takahashi, Y. (1960) Isolation of melatonin and 5-methoxyindole-3-acetic acid from bovine pineal glands. *J Biol Chem* **235** 1992-1997.
124. Simonneaux, V. & Ribelayga, C. (2003) Generation of the melatonin endocrine message in mammals: a review of the complex regulation of melatonin synthesis by norepinephrine, peptides, and other pineal transmitters. *Pharmacol Rev* **55** 325-395.
125. de Faria Poloni, J., Feltes, B.C. & Bonatto, D. (2011) Melatonin as a central molecule connecting neural development and calcium signaling. *Functional & integrative genomics* **11** 383-388.
126. de Borsetti, N.H., Dean, B.J., Bain, E.J., Clanton, J.A., Taylor, R.W. & Gamse, J.T. (2011) Light and melatonin schedule neuronal differentiation in the habenular nuclei. *Dev Biol* **358** 251-261.
127. Bellon, A., Ortiz-Lopez, L., Ramirez-Rodriguez, G., Anton-Tay, F. & Benitez-King, G. (2007) Melatonin induces neuritogenesis at early stages in N1E-115 cells through actin rearrangements via activation of protein kinase C and Rho-associated kinase. *J Pineal Res* **42** 214-221.
128. Klein, D.C. (1972) Evidence for the placental transfer of 3 H-acetyl-melatonin. *Nature: New biology* **237** 117-118.
129. Iwasaki, S., Nakazawa, K., Sakai, J., Kometani, K., Iwashita, M., Yoshimura, Y. & Maruyama, T. (2005) Melatonin as a local regulator of human placental function. *J Pineal Res* **39** 261-265.
130. Barnard, A.R. & Nolan, P.M. (2008) When clocks go bad: neurobehavioural consequences of disrupted circadian timing. *PLoS Genet* **4** e1000040.
131. Levoye, A., Dam, J., Ayoub, M.A., Guillaume, J.L., Couturier, C., Delagrange, P. & Jockers, R. (2006) The orphan GPR50 receptor specifically inhibits MT1 melatonin receptor function through heterodimerization. *EMBO J* **25** 3012-3023.
132. Weissbach, H., Redfield, B.G. & Axelrod, J. (1960) Biosynthesis of melatonin: enzymic conversion of serotonin to N-acetylserotonin. *Biochim Biophys Acta* **43** 352-353.
133. Axelrod, J., Shein, H.M. & Wurtman, R.J. (1969) Stimulation of C14-melatonin synthesis from C14-tryptophan by noradrenaline in rat pineal in organ culture. *Proc Natl Acad Sci U S A* **62** 544-549.
134. Gibbs, F.P. & Vriend, J. (1981) The half-life of melatonin elimination from rat plasma. *Endocrinology* **109** 1796-1798.
135. Cortesi, F., Giannotti, F., Ivanenko, A. & Johnson, K. (2010) Sleep in children with autistic spectrum disorder. *Sleep Med* **11** 659-664.
136. Guenole, F., Godbout, R., Nicolas, A., Franco, P., Claustrat, B. & Baleyte, J.M. (2011) Melatonin for disordered sleep in individuals with autism spectrum disorders: systematic review and discussion. *Sleep Med Rev* **15** 379-387.
137. Malow, B., Adkins, K.W., McGrew, S.G., Wang, L., Goldman, S.E., Fawkes, D. & Burnette, C. (2011) Melatonin for Sleep in Children with Autism: A Controlled Trial Examining Dose, Tolerability, and Outcomes. *J Autism Dev Disord*.
138. Tordjman, S., Anderson, G.M., Bellissant, E., Botbol, M., Charbuy, H., Camus, F., Graignic, R., Kermarrec, S., Fougerou, C., Cohen, D. & Touitou, Y. (2012) Day and nighttime excretion of 6-sulphatoxymelatonin in adolescents and young adults with autistic disorder. *Psychoneuroendocrinology*.

139. Tordjman, S., Anderson, G.M., Pichard, N., Charbuy, H. & Touitou, Y. (2005) Nocturnal excretion of 6-sulphatoxymelatonin in children and adolescents with autistic disorder. *Biol Psychiatry* **57** 134-138.
140. Mulder, E.J., Anderson, G.M., Kemperman, R.F., Oosterloo-Duinkerken, A., Minderaa, R.B. & Kema, I.P. (2010) Urinary excretion of 5-hydroxyindoleacetic acid, serotonin and 6-sulphatoxymelatonin in normoserotonemic and hyperserotonemic autistic individuals. *Neuropsychobiology* **61** 27-32.
141. Kulman, G., Lissoni, P., Rovelli, F., Roselli, M.G., Brivio, F. & Sequeri, P. (2000) Evidence of pineal endocrine hypofunction in autistic children. *Neuro Endocrinol Lett* **21** 31-34.
142. Melke, J., Goubran Botros, H., Chaste, P., Betancur, C., Nygren, G., Anckarsater, H., Rastam, M., Stahlberg, O., Gillberg, I.C., Delorme, R., Chabane, N., Mouren-Simeoni, M.C., Fauchereau, F., Durand, C.M., Chevalier, F., Drouot, X., Collet, C., Launay, J.M., Leboyer, M., *et al* (2008) Abnormal melatonin synthesis in autism spectrum disorders. *Mol Psychiatry* **13** 90-98.
143. Leu, R.M., Beyderman, L., Botzolakis, E.J., Surdyka, K., Wang, L. & Malow, B.A. (2011) Relation of melatonin to sleep architecture in children with autism. *J Autism Dev Disord* **41** 427-433.
144. Nir, I., Meir, D., Zilber, N., Knobler, H., Hadjez, J. & Lerner, Y. (1995) Brief report: circadian melatonin, thyroid-stimulating hormone, prolactin, and cortisol levels in serum of young adults with autism. *J Autism Dev Disord* **25** 641-654.
145. Ritvo, E.R., Ritvo, R., Yuwiler, A., Brothers, A., Freeman, B.J. & Plotkin, S. (1993) Elevated Daytime Helatonin Concentrations in Autism: A Pilot Study. *Eur Child Adolesc Psychiatry* **2** 75-78.
146. Pagan, C., Delorme, R., Callebert, J., Goubran-Botros, H., Amsellem, F., Drouot, X., Boudebesse, C., Le Dudal, K., Ngo-Nguyen, N., Laouamri, H., Gillberg, C., Leboyer, M., Bourgeron, T. & Launay, J.M. (2014) The serotonin-N-acetylserotonin-melatonin pathway as a biomarker for autism spectrum disorders. *Translational psychiatry* **4** e479.
147. Schain, R.J. & Freedman, D.X. (1961) Studies on 5-hydroxyindole metabolism in autistic and other mentally retarded children. *J Pediatr* **58** 315-320.
148. Janusonis, S. (2008) Origin of the blood hyperserotonemia of autism. *Theoretical biology & medical modelling* **5** 10.
149. Hallam, K.T., Olver, J.S., Chambers, V., Begg, D.P., McGrath, C. & Norman, T.R. (2006) The heritability of melatonin secretion and sensitivity to bright nocturnal light in twins. *Psychoneuroendocrinology* **31** 867-875.
150. Chaste, P., Clement, N., Botros, H.G., Guillaume, J.L., Konyukh, M., Pagan, C., Scheid, I., Nygren, G., Anckarsater, H., Rastam, M., Stahlberg, O., Gillberg, I.C., Melke, J., Delorme, R., Leblond, C., Toro, R., Huguet, G., Fauchereau, F., Durand, C., *et al* (2011) Genetic variations of the melatonin pathway in patients with attention-deficit and hyperactivity disorders. *J Pineal Res* **51** 394-399.
151. Botros, H.G., Legrand, P., Pagan, C., Bondet, V., Weber, P., Ben-Abdallah, M., Lemiere, N., Huguet, G., Bellalou, J., Maronde, E., Beguin, P., Haouz, A., Shepard, W. & Bourgeron, T. (2013) Crystal structure and functional mapping of human ASMT, the last enzyme of the melatonin synthesis pathway. *J Pineal Res* **54** 46-57.
152. Jonsson, L., Ljunggren, E., Bremer, A., Pedersen, C., Landen, M., Thuresson, K., Giacobini, M. & Melke, J. (2010) Mutation screening of melatonin-related genes in patients with autism spectrum disorders. *BMC Med Genomics* **3** 10.
153. Toma, C., Rossi, M., Sousa, I., Blasi, F., Bacchelli, E., Alen, R., Vanhala, R., Monaco, A.P., Jarvela, I. & Maestrini, E. (2007) Is ASMT a susceptibility gene for autism spectrum disorders? A replication study in European populations. *Mol Psychiatry* **12** 977-979.
154. Wang, L., Li, J., Ruan, Y., Lu, T., Liu, C., Jia, M., Yue, W., Liu, J., Bourgeron, T. & Zhang, D. (2013) Sequencing ASMT identifies rare mutations in Chinese Han patients with autism. *PLoS One* **8** e53727.

155. Cai, G., Edelmann, L., Goldsmith, J.E., Cohen, N., Nakamine, A., Reichert, J.G., Hoffman, E.J., Zurawiecki, D.M., Silverman, J.M., Hollander, E., Soorya, L., Anagnostou, E., Betancur, C. & Buxbaum, J.D. (2008) Multiplex ligation-dependent probe amplification for genetic screening in autism spectrum disorders: efficient identification of known microduplications and identification of a novel microduplication in ASMT. *BMC Med Genomics* **1** 50.

156. Chaste, P., Clement, N., Mercati, O., Guillaume, J.L., Delorme, R., Botros, H.G., Pagan, C., Perivier, S., Scheid, I., Nygren, G., Anckarsater, H., Rastam, M., Stahlberg, O., Gillberg, C., Serrano, E., Lemiere, N., Launay, J.M., Mouren-Simeoni, M.C., Leboyer, M., *et al* (2010) Identification of pathway-biased and deleterious melatonin receptor mutants in autism spectrum disorders and in the general population. *PLoS One* **5** e11495.

157. Lyssenko, V., Nagorny, C.L., Erdos, M.R., Wierup, N., Jonsson, A., Spegel, P., Bugliani, M., Saxena, R., Fex, M., Pulizzi, N., Isomaa, B., Tuomi, T., Nilsson, P., Kuusisto, J., Tuomilehto, J., Boehnke, M., Altshuler, D., Sundler, F., Eriksson, J.G., *et al* (2009) Common variant in MTNR1B associated with increased risk of type 2 diabetes and impaired early insulin secretion. *Nat Genet* **41** 82-88.

158. Bonnefond, A., Clement, N., Fawcett, K., Yengo, L., Vaillant, E., Guillaume, J.L., Dechaume, A., Payne, F., Roussel, R., Czernichow, S., Hercberg, S., Hadjadj, S., Balkau, B., Marre, M., Lantieri, O., Langenberg, C., Bouatia-Naji, N., Charpentier, G., Vaxillaire, M., *et al* (2012) Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. *Nat Genet* **44** 297-301.

159. Anckarsater, H., Lundstrom, S., Kollberg, L., Kerekes, N., Palm, C., Carlstrom, E., Langstrom, N., Magnusson, P.K.E., Halldner, L., Bolte, S., Gillberg, C., Gumpert, C., Rastam, M. & Lichtenstein, P. (2011) The Child and Adolescent Twin Study in Sweden (CATSS). *Twin Research and Human Genetics* **14** 495-508.

160. Larson, T., Lundstrom, S., Nilsson, T., Selinus, E.N., Rastam, M., Lichtenstein, P., Gumpert, C.H., Anckarsater, H. & Kerekes, N. (2013) Predictive properties of the A-TAC inventory when screening for childhood-onset neurodevelopmental problems in a population-based sample. *BMC Psychiatry* **13** 233.

161. Hansson, S.L., Svanstrom Rojvall, A., Rastam, M., Gillberg, C. & Anckarsater, H. (2005) Psychiatric telephone interview with parents for screening of childhood autism - tics, attention-deficit hyperactivity disorder and other comorbidities (A-TAC): preliminary reliability and validity. *Br J Psychiatry* **187** 262-267.

162. Pettersson, E., Anckarsater, H., Gillberg, C. & Lichtenstein, P. (2013) Different neurodevelopmental symptoms have a common genetic etiology. *J Child Psychol Psychiatry* **54** 1356-1365.

163. Hannelius, U., Gherman, L., Makela, V.V., Lindstedt, A., Zucchelli, M., Lagerberg, C., Tybring, G., Kere, J. & Lindgren, C.M. (2007) Large-scale zygosity testing using single nucleotide polymorphisms. *Twin Res Hum Genet* **10** 604-625.

164. Kantzer, A.K., Fernell, E., Gillberg, C. & Miniscalco, C. (2013) Autism in community pre-schoolers: developmental profiles. *Res Dev Disabil* **34** 2900-2908.

165. Nygren, G., Sandberg, E., Gillstedt, F., Ekeroth, G., Arvidsson, T. & Gillberg, C. (2012) A new screening programme for autism in a general population of Swedish toddlers. *Res Dev Disabil* **33** 1200-1210.

166. Menashe, I., Larsen, E.C. & Banerjee-Basu, S. (2013) Prioritization of Copy Number Variation Loci Associated with Autism from AutDB-An Integrative Multi-Study Genetic Database. *PLoS One* **8** e66707.

167. Banerjee-Basu, S. & Packer, A. (2010) SFARI Gene: an evolving database for the autism research community. *Disease models & mechanisms* **3** 133-135.

168. Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., Fitzgerald, T., Hu, M., Ihm, C.H., Kristiansson, K., Macarthur, D.G., Macdonald, J.R., Onyiah, I., Pang, A.W., Robson, S., *et al* (2010) Origins and functional impact of copy number variation in the human genome. *Nature* **464** 704-712.
169. McCarroll, S.A., Kuruvilla, F.G., Korn, J.M., Cawley, S., Nemesh, J., Wysoker, A., Shapero, M.H., de Bakker, P.I., Maller, J.B., Kirby, A., Elliott, A.L., Parkin, M., Hubbell, E., Webster, T., Mei, R., Veitch, J., Collins, P.J., Handsaker, R., Lincoln, S., *et al* (2008) Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* **40** 1166-1174.
170. MacDonald, J.R., Ziman, R., Yuen, R.K., Feuk, L. & Scherer, S.W. (2014) The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res* **42** D986-992.
171. Harikrishna, J.A., Black, S.M., Szklarz, G.D. & Miller, W.L. (1993) Construction and function of fusion enzymes of the human cytochrome P450scc system. *DNA and cell biology* **12** 371-379.
172. Tee, M.K., Abramsohn, M., Loewenthal, N., Harris, M., Siwach, S., Kaplinsky, A., Markus, B., Birk, O., Sheffield, V.C., Parvari, R., Hershkovitz, E. & Miller, W.L. (2013) Varied clinical presentations of seven patients with mutations in CYP11A1 encoding the cholesterol side-chain cleavage enzyme, P450scc. *J Clin Endocrinol Metab* **98** 713-720.
173. Sahakitrungruang, T., Tee, M.K., Blackett, P.R. & Miller, W.L. (2011) Partial defect in the cholesterol side-chain cleavage enzyme P450scc (CYP11A1) resembling nonclassic congenital lipoid adrenal hyperplasia. *J Clin Endocrinol Metab* **96** 792-798.
174. Pagan, C., Botros, H.G., Poirier, K., Dumaine, A., Jamain, S., Moreno, S., de Brouwer, A., Van Esch, H., Delorme, R., Launay, J.M., Tzschach, A., Kalscheuer, V., Lacombe, D., Briault, S., Laumonnier, F., Raynaud, M., van Bon, B.W., Willemsen, M.H., Leboyer, M., *et al* (2011) Mutation screening of ASMT, the last enzyme of the melatonin pathway, in a large sample of patients with intellectual disability. *BMC Med Genet* **12** 17.
175. Folsom, T.D. & Fatemi, S.H. (2013) The involvement of Reelin in neurodevelopmental disorders. *Neuropharmacology* **68** 122-135.
176. Fatemi, S.H., Earle, J.A. & McMenomy, T. (2000) Reduction in Reelin immunoreactivity in hippocampus of subjects with schizophrenia, bipolar disorder and major depression. *Mol Psychiatry* **5** 654-663, 571.
177. Fatemi, S.H., Stary, J.M. & Egan, E.A. (2002) Reduced blood levels of reelin as a vulnerability factor in pathophysiology of autistic disorder. *Cellular and molecular neurobiology* **22** 139-152.
178. Fatemi, S.H., Stary, J.M., Halt, A.R. & Realmuto, G.R. (2001) Dysregulation of Reelin and Bcl-2 proteins in autistic cerebellum. *J Autism Dev Disord* **31** 529-535.
179. Pourcain, B.S., Wang, K., Glessner, J.T., Golding, J., Steer, C., Ring, S.M., Skuse, D.H., Grant, S.F.A., Hakonarson, H. & Smith, G.D. (2010) Association Between a High-Risk Autism Locus on 5p14 and Social Communication Spectrum Phenotypes in the General Population. *Am J Psychiat* **167** 1364-1372.
180. Penagarikano, O., Abrahams, B.S., Herman, E.I., Winden, K.D., Gdalyahu, A., Dong, H., Sonnenblick, L.I., Gruver, R., Almajano, J., Bragin, A., Golshani, P., Trachtenberg, J.T., Peles, E. & Geschwind, D.H. (2011) Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell* **147** 235-246.
181. Arking, D.E., Cutler, D.J., Brune, C.W., Teslovich, T.M., West, K., Ikeda, M., Rea, A., Guy, M., Lin, S., Cook, E.H. & Chakravarti, A. (2008) A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *Am J Hum Genet* **82** 160-164.

182. Tan, G.C., Doke, T.F., Ashburner, J., Wood, N.W. & Frackowiak, R.S. (2010) Normal variation in fronto-occipital circuitry and cerebellar structure with an autism-associated polymorphism of CNTNAP2. *NeuroImage* **53** 1030-1042.
183. Penagarikano, O. & Geschwind, D.H. (2012) What does CNTNAP2 reveal about autism spectrum disorder? *Trends in molecular medicine* **18** 156-163.
184. Alarcon, M., Abrahams, B.S., Stone, J.L., Duvall, J.A., Perederiy, J.V., Bomar, J.M., Sebat, J., Wigler, M., Martin, C.L., Ledbetter, D.H., Nelson, S.F., Cantor, R.M. & Geschwind, D.H. (2008) Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *Am J Hum Genet* **82** 150-159.
185. Sampath, S., Bhat, S., Gupta, S., O'Connor, A., West, A.B., Arking, D.E. & Chakravarti, A. (2013) Defining the contribution of CNTNAP2 to autism susceptibility. *PLoS One* **8** e77906.
186. Durand, C.M., Betancur, C., Boeckers, T.M., Bockmann, J., Chaste, P., Fauchereau, F., Nygren, G., Rastam, M., Gillberg, I.C., Anckarsater, H., Sponheim, E., Goubran-Botros, H., Delorme, R., Chabane, N., Mouren-Simeoni, M.C., de Mas, P., Bieth, E., Roge, B., Heron, D., *et al* (2007) Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet* **39** 25-27.
187. Moessner, R., Marshall, C.R., Sutcliffe, J.S., Skaug, J., Pinto, D., Vincent, J., Zwaigenbaum, L., Fernandez, B., Roberts, W., Szatmari, P. & Scherer, S.W. (2007) Contribution of SHANK3 mutations to autism spectrum disorder. *Am J Hum Genet* **81** 1289-1297.
188. Gauthier, J., Spiegelman, D., Piton, A., Lafreniere, R.G., Laurent, S., St-Onge, J., Lapointe, L., Hamdan, F.F., Cossette, P., Mottron, L., Fombonne, E., Joober, R., Marineau, C., Drapeau, P. & Rouleau, G.A. (2009) Novel de novo SHANK3 mutation in autistic patients. *Am J Med Genet B Neuropsychiatr Genet* **150B** 421-424.
189. Wang, Z., Hong, Y., Zou, L., Zhong, R., Zhu, B., Shen, N., Chen, W., Lou, J., Ke, J., Zhang, T., Wang, W. & Miao, X. (2014) Reelin gene variants and risk of autism spectrum disorders: An integrated meta-analysis. *Am J Med Genet B Neuropsychiatr Genet* **165** 192-200.
190. Sharma, J.R., Arieff, Z., Gameeldien, H., Davids, M., Kaur, M. & van der Merwe, L. (2013) Association analysis of two single-nucleotide polymorphisms of the RELN gene with autism in the South African population. *Genetic testing and molecular biomarkers* **17** 93-98.
191. Dutta, S., Sinha, S., Ghosh, S., Chatterjee, A., Ahmed, S. & Usha, R. (2008) Genetic analysis of reelin gene (RELN) SNPs: no association with autism spectrum disorder in the Indian population. *Neurosci Lett* **441** 56-60.
192. Li, H., Li, Y., Shao, J., Li, R., Qin, Y., Xie, C. & Zhao, Z. (2008) The association analysis of RELN and GRM8 genes with autistic spectrum disorder in Chinese Han population. *Am J Med Genet B Neuropsychiatr Genet* **147B** 194-200.
193. Serajee, F.J., Zhong, H. & Mahbubul Huq, A.H. (2006) Association of Reelin gene polymorphisms with autism. *Genomics* **87** 75-83.
194. Bonora, E., Beyer, K.S., Lamb, J.A., Parr, J.R., Klauck, S.M., Benner, A., Paolucci, M., Abbott, A., Ragoussis, I., Poustka, A., Bailey, A.J. & Monaco, A.P. (2003) Analysis of reelin as a candidate gene for autism. *Mol Psychiatry* **8** 885-892.