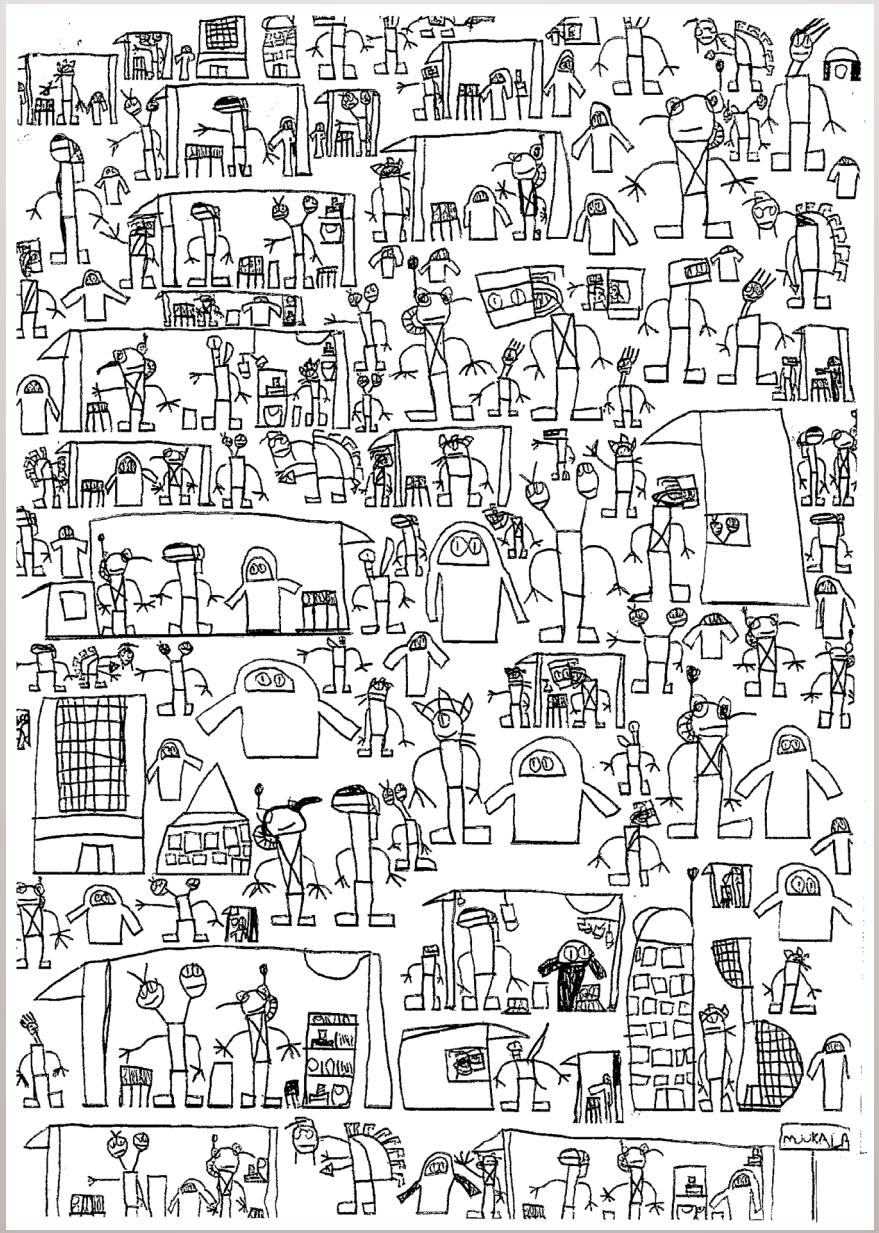
Exploring Genetic Susceptibility to Autism Spectrum Disorders

Katri Kantojärvi



UNIVERSITY OF HELSINKI
FACULTY OF MEDICINE

EXPLORING GENETIC SUSCEPTIBILITY TO AUTISM SPECTRUM DISORDERS

Katri Kantojärvi

Department of Medical Genetics

Haartman Institute

University of Helsinki

ACADEMIC DISSERTATION

To be publicly discussed, with the permission of The Faculty of Medicine, University of Helsinki, in the lecture hall 2, Biomedicum Helsinki, on the 26th of April, at 12 o'clock noon.

HELSINKI 2013

The artwork on the cover page is drawn by a boy with childhood autism.

Published with the permission of the Department of Child Neurology, Helsinki University Hospital and the parents of the child.

ISBN 978-952-10-8694-6 (paperback)
ISBN 978-952-10-8695-3 (PDF)
Unigrafia OY
Helsinki, Finland 2013

Supervised by

Docent Irma Järvelä
Department of Medical Genetics
Haartman Institute
University of Helsinki
Helsinki, Finland

and

Docent Päivi Onkamo
Department of Biosciences
University of Helsinki
Helsinki, Finland

Reviewed by

MD, PhD Jacob Vorstman
Department of Psychiatry
University Medical Center Utrecht
Netherlands

and

Docent Liisa Myllykangas
Department of Pathology
University of Helsinki and HUSLAB
Helsinki, Finland

Opponent

Professor Leena Haataja
Department of Child Neurology
Turku University Hospital
Turku, Finland

Everyone is a genius. But if you judge
a fish on its ability to climb a tree, it will
live its whole life believing that it is stupid.

-Albert Einstein

ABSTRACT

Autism spectrum disorders (ASD) are neuropsychiatric disorders characterized by restricted repetitive behavior and abnormalities in communication and social interaction. Although the heritability of the trait has been estimated to be relatively high, the model of inheritance of ASD seems to be very complex and probably involves multiple interacting genes. Also, environmental factors together with genetic risk factors may increase the risk for autism.

The purpose of this study was to detect genetic variants predisposing to ASD. We initiated the study with fine mapping of chromosomal region Xq11.1-q21.33 in 99 Finnish ASD families where linkage has been observed in earlier studies. The highest multipoint LOD score was obtained with the marker DXS1225 at Xq21.1 (NPL_{all}=3.43). However, Sanger sequencing of five candidate genes in the linkage region did not reveal any disruptive mutations. Despite that, in Study IV we detected significant association in a family-based genome-wide association (GWA) scan at Xq21 in the *PCDH11X* gene. There are several candidate genes in the linkage region and exome sequencing of the X chromosome might reveal disruptive mutations in ASD.

Aberrant glutamate metabolism has been observed in both ASD and obsessive compulsive disorder (OCD). In Study II we analyzed if single nucleotide polymorphisms (SNPs) previously associated in OCD in the glutamate transporter gene *SLC1A1* at 9p24 are also associated in ASD. We further analyzed whether the SNPs reported earlier in the largest linkage study of ASD by the Autism Genome Project (2007) at chromosomal areas 9p24 (rs1340513 and rs722628) and 11p12-13 (rs1358054 and rs1039205) are associated with ASD in a sample of 175 Finnish ASD cases and 216 controls. SNP rs1340513, in *JMJD2C*, the neighboring gene of *SLC1A1,* showed significant association with ASD in this study. In this thesis, the GWA scan (Study IV) and copy number variants (CNVs) with a different sample set also supported the association at locus 9p24 with *SLC1A1, JMJD2C* and *PTPRD*.

Synaptic defects have been suggested to be the mechanism underlying autism. Disruptive mutations in synaptic genes *NRXN1, NLGN3/4X* and *SHANK2/3* have been detected in ASD patients. In Study III we continued the mutation screen of the *SHANK2* gene in 455 European

ASD families. Several harmful mutations were detected and in functional analyses we observed that they reduced synaptic density *in vitro*. We also detected deletions in the *SHANK2* gene in ASD samples. We noticed that patients who had a *SHANK2* deletion carried additional CNVs on chromosomal region 15q11-q13, which has previously been associated to ASD and several other neuropsychiatric disorders. This supports a multiple hit model for ASD. Additional studies are warranted to analyze how many mutations are sufficient to contribute to ASD and what kind of combination of genetic defects will be detected in individual ASD families.

In this thesis, we performed a genome-wide scan with a novel sample set of 83 Finnish ASD families and 750 controls (a cohort of Health 2000 study). We detected the strongest association at chromosome 16p13.2 with the *RBFOX1* gene, which regulates tissue-specific splicing of several autism related genes. Preliminary evidence of epistatic interactions was obtained between SNPs in *NRXN1* and *UBA52,* as well as *RBFOX1* and *SCN1A,* and *DLG2* and *RBFOX3*. In addition, we performed promoter analyses for ASD candidate genes and discovered a possible transcriptional regulatory site in the promoter area of *AVPR1A*, which might partly explain the associations with autism observed with the promoter polymorphisms of this gene. Furthermore, a transcription factor (TF) binding site for early growth response (EGR) was enriched in autism candidate genes. Finally, we performed CNV analyses and detected large (>400 kb) CNVs in chromosomal regions 15q13.3, 16p11.2, 17q12 and 22q11. CNVs in the same regions have been detected in ASD and other neuropsychiatric disorders such as ADHD, epilepsy, schizophrenia and intellectual disability in earlier studies. We also observed CNVs in known ASD candidate genes for example *DISC1, FOXG1, ASMT, PCDH11X,* and *PRODH.*

In conclusion, the results obtained in this thesis show that several genetic risk variants predispose to ASD and epistasis between ASD candidate genes play an important role in these disorders. More studies are warranted to explore the combination and interaction of genetic risk variants and their pathways and environmental triggers which all together could contribute to ASD.

TIIVISTELMÄ

Autismikirjon häiriöt ovat neuropsykiatrisia sairauksia, joille tunnusomaista ovat vaikeudet sosiaalisessa vuorovaikutuksessa ja kommunikaatiossa sekä elämää hallitsevat rutiinit ja rituaalit. Perintötekijöillä on havaittu olevan osuutta autismikirjon sairauksien synnyssä, mutta periytymismalli on monimutkainen ja luultavasti useiden geenimuutosten yhteisvaikutusten aiheuttama. Todennäköisesti myös ympäristötekijät yhdessä geneettisten riskitekijöiden kanssa kasvattavat riskiä autismin puhkeamiseen.

Tämän tutkimuksen tarkoitus oli tunnistaa autismikirjon sairauksille altistavia geenimuutoksia suomalaisista potilaista ja kontrolleista koostuvissa aineistoissa. Aloitimme tutkimuksen hienokartoittamalla jo aiemmissa suomalaisissa tutkimuksissa autismikirjon häiriöihin kytkeytyvää kromosomialuetta Xq11.1-q21.33. Aineistona oli 99 autismikirjoon kuuluvaa perhettä. Paras kytkentäluku (LOD score) havaittiin markkerilla DXS1225 alueella Xq21.1 (NPL_{all}=3.43). Alueelta sekveksointiin 5 autismin kandidaattigeeniä, mutta tautia aiheuttavia mutaatioita ei havaittu tutkituissa perheissä. Osatyössä IV havaitsimme kuitenkin vahvan assosiaation lähellä parasta kytkentäaluetta keskushermoston toimintaan liittyvässä protokadheriini-geenissä *PCDH11X* kromosomialueella Xq21. Kytkentäalueella on useita autismin ehdokasgeenejä, joten koko alueen sekvensointi olisi tarpeellista.

Autismikirjon häiriöissä ja pakko-oireisessa häiriössä on havaittu glutamaatti aineenvaihdunnan poikkeavuuksia. Osatyössä II analysoimme yleisiä polymorfioita (SNPs) glutamaatti-transportteri geenissä *SLC1A1* kromosomialueella 9p24, joiden on aiemmissa tutkimuksissa havaittu liittyvän pakko-oireiseen häiriöön. Aineistona oli 175 autismikirjoon kuuluvaa henkilöä ja 216 kontrollia. Tutkimme myös 4 SNPiä kromosomialueilta 9p24 (rs1340513 ja rs722628) ja 11p12-13 (rs1358054 ja 1039205), joiden on aiemmin havaittu liittyvän autismiin laajassa kansainvälisessä tutkimuksessa (Autism Genome Project 2007). Tässä tutkimuksessa SNP rs1340513 assosioitui autismiin (P=0.007) metylointiin liittyvässä geenissä *JMJD2C*, joka sijaitsee geenin *SLC1A1* lähellä. Lisäksi väitöskirjan osatyössä IV havaittiin assosiaatio eri autismimateriaalilla kromosomipaikkaan 9p24 geeneihin *SLC1A1, JMJD2C* ja *PTPRD*.

Häiriöiden synapsien toiminnassa on epäilty liittyvän autismiin. Autisteilla on aiemmin havaittu haitallisia muutoksia synapsien syntyyn vaikuttavissa geeneissä *NRXN1*, *NLGN3/4X* ja *SHANK2/3.* Osatyössä III analysoimme *SHANK2* geenin mutaatioita 455 eurooppalaisessa autismiperheessä. Suomalaisia perheitä oli mukana 99. Tulokset osoittivat, että autismista kärsivillä oli selvästi verrokkeja enemmän sellaisia *SHANK2*-geenin haitallisia muutoksia, jotka olivat yhteydessä alentuneeseen synapsien tiheyteen. *SHANK2*-geenin deleetio ja autismissa aiemmin löydetty kromosomin 15q11-13 kopioluvun muutos löytyivät samoilta potilailta. Tulos tukee käsitystä että autismin puhkeamiseen vaaditaan useita haitallisia muutoksia perimässä.

Neljännessä osatyössä teimme koko perimän laajuisen assosiaatioanalyysin uudessa suomalaisessa 83 autismiperheen ja 750 normaalin hengen kontrollimateriaalissa (Terveys2000-aineisto). Vahvimman assosiaation havaitsimme kromosomialueella 16p13.2 geenissä *RBFOX1*, joka säätelee useiden autismiin liitettyjen geenien ilmentymistä ja silmukointia. Geenien välistä yhteistoimintaa havaittiin geenien *NRXN1* ja *UBA52*, *RBFOX1* ja *SCN1A* sekä *DLG2* ja *RBFOX3* välillä. Teimme myös promoottorianalyysin autismin kandidaattigeeneille ja havaitsimme mahdollisen transkription säätelyalueen geenissä *AVPR1A*, mikä voisi osittain selittää autismiin aiemmin liitetyn promoottorialueen polymorfian tässä geenissä. Lisäksi havaitsimme että sitoutumiskohta transkriptio faktorille EGR oli rikastunut autismin ehdokasgeeneissä verrokkigeeneihin verrattuna. Tutkimme tässä aineistossa myös kopioluvun muutoksia ja tunnistimme suuria (>400 kb) muutoksia kromosomialueilta 15q13.3, 16p11.2, 17q12 and 22q11. Kopioluvun muutoksia näillä alueilla on aiemmissa tutkimuksissa havaittu autismin lisäksi myös muissa neuropsykiatrisissa sairauksissa. Havaitsimme kopioluvun muutoksia myös autismiin aiemmin liitetyissä geeneissä *DISC1, FOXG1, ASMT, PCDH11X* ja *PRODH.*

Tässä väitöskirjassa havaitut tulokset osoittavat että *RBFOX1* geenin muutokset liittyvät autismiin suomalaisessa perhemateriaalissa ja vahvistavat käsitystä, että useat geneettiset muutokset ja niiden yhteisvaikutukset altistavat autismille. Lisätutkimukset ovat tarpeellisia selvittämään mitkä geneettisten variaatioiden yhdistelmät, yhdessä ympäristötekijöiden kanssa vaaditaan autismin puhkeamiseen.

CONTENTS

ABBREVIATIONS

22q11DS	22q11 deletion syndrome
aCGH	array comparative genomic hybridization
ADHD	attention deficit-hyperactivity disorder
ADI-R	Autism Diagnostic Interview - Revised
ADOS	Autism Diagnostic Observation Schedule
AGP	Autism Genome Project
AGRE	Autism Genetic Resource Exchange
APA	American Psychiatric Association
AS	Asperger syndrome
ASD	autism spectrum disorders
ASDI	Asperger Syndrome Diagnostic Interview
ASSQ	Asperger Syndrome Screening Questionnaire
AVP	arginine vasopressin
AVPR1A	arginine vasopressin receptor 1A
bp	base pair
CARS	Childhood Autism Rating Scale
CNV	copy number variation
CY-BOCS	Children´s Yale-Brown Obsessive-Compulsive Scale
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, 4th Edition
EGR	early growth response
EIBI	early intensive behavioral intervention
GWAS	genome-wide association study
HFA	high function autism
HFAP	human alpha-fetoprotein
ICD-10	International Classification of Diseases, 10th Revision
ID	intellectual disability

IMGSAC	International Molecular Genetic Study of Autism Consortium
JMJD2C	jumonji domain containing 2C
kb	kilobase
LOD	logarithm of the odds
LD	linkage disequilibrium
MAF	minor allele frequency
Mb	megabase
MDD	major depressive disorder
NPL	non-parametric LOD
OCD	obsessive compulsive disorder
PCDH11X	protocadherin 11 X-linked
PDD-NOS	pervasive developmental disorder not otherwise specified
PGC	Psychiatric Genomic Consortium
PSD	postsynaptic density
RBFOX1	RNA binding protein, fox-1 homolog (C. elegans) 1
SHANK2	SH3 and multiple ankyrin repeat domains 2
SLC1A1	solute carrier family 1, member 1
SNP	single nucleotide polymorphism
SRS	Social Responsiveness Scale
STR	short tandem repeat
TDT	transmission disequilibrium test
TF	transcription factor
WHO	World Health Organization
XLID	X-linked Intellectual disability
Zmax	maximum LOD score

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original articles referred to in the text by their Roman numerals. In addition, some unpublished data are also presented.

I. **Kantojärvi K**,* Kotala I*, Rehnström K, Ylisaukko-oja T, Vanhala R, Nieminen- von Wendt T, Järvelä I. Fine mapping of Xq11.1-21.33 and mutation screening of *RPS6KA6, ZNF711*, *ACSL4*, *DLG3* and *IL1RAPL2* for autism spectrum disorders (ASD). Autism Res 4:228-332, 2011.

II. **Kantojärvi K**, Onkamo P, Alen R, Nieminen-von Wendt T, von Wendt L, Vanhala R, Järvelä I. Analysis of 9p24 and 11p12-13 regions in autism spectrum disorders: rs1340513 in the *JMJD2C* gene is associated with ASDs in Finnish sample. Psychiatric Genetics 20:102-108, 2010.

III. Leblond CS, Heinrich J, Delorme R, Proepper C, Betancur C, Huguet G, Konyukh M, Chaste P, Ey E, Rastam M, Anckarsäter H, Nygren G, Ståhlberg O, Gillberg IC, Melke J, Toro R, Regnault B, Fauchereau F, Mercati O, Lemière N, Skuse D, Poot M, Holt R, Järvelä I, **Kantojärvi K**, Vanhala R, Curran S, Collier D, Bolton P, Chiocchetti A, Klauck SM, Poustka F, Freitag CM, Bacchelli E, Minopoli F, Maestrini E, Mazzone L, Ruta L, Sousa I, Vicente A, Oliveira G, Pinto D,Scherer S, Zelenika D, Delepine M, Lathrop M, Guinchat V, Devillard F, Assouline B, Mouren MC, Leboyer M, Gillberg C, Boeckers TM ,Bourgeron T. Genetic and functional analyses of *SHANK2* mutations provide evidence for a multiple hit model of autism spectrum disorders. PLoS Genet 8(2):e1002521, 2012.

IV. **Katri Kantojärvi*,** Jaana Oikkonen*, Seppo Koskinen, Ilona Kotala, Jenni Kallela, Raija Vanhala, Liisa Holm, Irma Järvelä, Päivi Onkamo. GWA and in silico promoter analysis in autism strengthen the role of *RBFOX1* and EGR transcription factors and depict a closely interacting network of proteins. Submitted.

*These authors contributed equally to this work.

1 INTRODUCTION

Autism spectrum disorders (ASD) are characterized by abnormalities in communication and social interaction. Restricted repetitive and stereotyped behavior with unusual interests and activities is also typical. ASD affect an estimated 0.62 % of the population. ASD prevalence has been increasing, which might, in part, be a result of changing diagnostic criteria and more frequent reporting of ASD cases (Newschaffer et al. 2007). ASDs are more common in men than in women, with an average male to female ratio of 4.3:1 (Fombonne 2005).

The word “autism” originates from the Greek word “autos”, which means “self”. It describes a situation in which individuals become absorbed in their own world and lose contact with other people. Eugen Bleuler was a Swiss psychiatrist who first used the term “autistic psychopathy” to describe the desire of schizophrenic patients to withdraw from public places (Bleuler 1916). The distinction between autism and schizophrenia remained controversial for a long time and autistic disorders were first introduced under the diagnostic criterion of childhood schizophrenia (DSM-II) (American Psychiatric Association 1968).

American psychiatrist Leo Kanner described 11 children, mostly boys with intellectual disability, repetitive and ritualistic interests and activities, delayed early language development and severe social isolation (Kanner 1943). Leo Kanner´s paper “Autistic Disturbances of Affective Contact” is the basis for the modern definition and today´s diagnostic criteria for infantile autism. Kanner also reported that an early cold relationship between the mother and the child might be the cause of autism. Bruno Bettelheim was an Austrian-American child psychologist and writer who popularized the "Refrigerator Mother" hypothesis of autism (Bettelheim 1967). Contrary to Kanner, he ignored that these same mothers had other children who were not autistic. The hypothesis was later rejected by scientists when the importance of the genetic background of autism became clear.

A year after Kanner´s publication an Austrian paediatrician, Hans Asperger, described four boys with autistic psychopathy but normal intellectual abilities (Asperger 1944). He called

them “little professors” who talked about their own interests but had difficulties with non-verbal communication such as understanding other peoples´ facial expressions. He observed that the onset of this condition seemed to be later than in autism and it almost never became evident before three years of age. The combination of these problems were generally referred to as Asperger syndrome (AS) and knowledge of this condition began to expand after the English review by Lorna Wing (Wing 1981).

ASDs have one of the highest heritability estimates of neuropsychiatric disorders and studies on the genetic component have been conducted since the 1980´s. The idea of genetic causes of autism has changed in the last few years. Autism has been suggested to be a polygenic disorder caused by multiple genetic risk factors, each with a weak effect. It has also been thought to be a group of disorders caused by heterogeneous genetic risk factors influencing common neuronal pathways (Bourgeron 2009, Toro et al. 2010) wherein only a single highly penetrant mutation could cause autism in a limited number of patients (Betancur 2011). Girirajan et al. (2010) proposed a two-hit model, wherein a second CNV together with 16p12.1 microdeletion resulted in more severe clinical manifestation in developmental disorders. In this thesis, molecular genetic studies of mostly Finnish ASD cases are presented with several approaches including the study of candidate genes, genome-wide association (GWA) and copy number variants (CNVs).

2 REVIEW OF THE LITERATURE

2.1 AUTISM SPECTRUM DISORDERS

2.1.1 Clinical features and diagnostics

Autism spectrum disorders (ASD) are a group of early onset developmental disabilities which all express difficulties in three core areas; communication, social skills and behavioral flexibility. The disorders express varying severity. According to the International Classification of Diseases (ICD-10) diagnostic criteria, ASDs include childhood autism, Asperger syndrome (AS), childhood disintegrative disorder, pervasive developmental disorder not otherwise specified (atypical autism) and Rett syndrome (World Health Organization 1993).

Impairment of the communication exists across a broad spectrum and it affects both verbal and nonverbal communication (Cashin et al. 2009). In most children with ASD, communication skills are severely delayed. The majority of preschool-age children with childhood autism presents with little to no functional speech and must be taught to apply alternative forms of communication such as pictures, sign language, and speech generating devices. The use of intonation in speech and interpretation of other people´s use of it is also impaired. Non-verbal communication is restricted as well as understanding other people´s facial expressions.

The impairment of social skills includes difficulties in interpreting communication and a lack of ability to form a theory about what other people are thinking or feeling (theory of mind) (Cashin et al. 2009). Impaired behavioral flexibility manifests as a restricted and repetitive range of interests and activities and, at times, in ritualistic manners. The restricted range of interests often arises as obsessions. These obsessions appear as a resource to mediate anxiety (Grandin & Scariano 1996). Compared to obsessive compulsive disorders (OCD) there is less fear of catastrophic consequences if the obsessive behaviors are not performed (Baron-Cohen & Wheelwright 1999).

Patterns of symptoms appear in childhood autism before the age of 3. It is characterised by delayed language development, unusual behavior, and social and communication difficulties as well as in some cases by delay in intellectual development. Only about 10 % of individuals with classical autism are able to live relatively independently; most require lifelong assistance (Howlin et al. 2004). Asperger syndrome (AS) exhibits no delay in language or cognitive development and has milder behavioral signs. AS is usually recognized later than classical autism, at preschool age (Khouzam et al. 2004). The detailed diagnostic criteria for childhood autism and Asperger syndrome are shown in Table 1.

Pervasive developmental disorder-not otherwise specified (PDD-NOS) also called “atypical autism” has fewer and milder symptoms compared to classical autism. Childhood disintegrative disorder has normal development for about 2 years, in terms of acquisition of communication skills, non-verbal behavior, and motor functioning and skills (such as toy play), and then substantial regression or loss of functioning appears (after 2 years).

Rett syndrome is a rare neurodevelopmental disorder in females that is classified to ASD. It characterized by severe intellectual disability, hand-wringing stereotypes, hyperventilation, loss of purposive hand movements and autistic features. Rett syndrome is caused by mutations in the *MECP2* (methyl-CpG binding protein 2) gene (Amir et al. 1999). In less than 5 % of the cases mutations in genes *CDKL5* (cyclin-dependent kinase-like 5) or *FOXG1* (forkhead box G1) are noticed to cause Rett syndrome (Ariani et al. 2008, Russo et al. 2009).

15-47 % of children with ASD express autistic regression which occurs when a child appears to develop typically, but around the second year of life begin to lose the previously acquired speech and social skills, accompanied with onset of autistic features (Stefanatos et al. 2008). Autism with and without regression is a continuous spectrum of behaviors with some children expressing late losses and some early delays of skills (Ozonoff et al. 2008).

Table 1. Diagnostic criteria for childhood autism and Asperger syndrome according to the ICD-10 classification (World Health Organization 1993).

Childhood autism F84.0	
A	**Presence of abnormal or impaired development before the age of three years, in at least one out of the following areas:** 1. receptive or expressive language as used in social communication 2. the development of selective social attachments or of reciprocal social interaction 3. functional or symbolic play
B	**Qualitative abnormalities in reciprocal social interaction, manifest in at least one of the following areas:** 1. failure adequately to use eye-to-eye gaze, facial expression, body posture and gesture to regulate social interaction 2. failure to develop (in a manner appropriate to mental age, and despite ample opportunities) peer relationships that involve a mutual sharing of interests, activities and emotions 3. A lack of socio-emotional reciprocity as shown by an impaired or deviant response to other people's emotions; or lack of modulation of behavior according to social context, or a weak integration of social, emotional and communicative behaviors.
C	**Qualitative abnormalities in communication, manifest in at least two of the following areas:** 1. a delay in, or total lack of development of spoken language that is not accompanied by an attempt to compensate through the use of gesture or mime as alternative modes of communication (often preceded by a lack of communicative babbling) 2. relative failure to initiate or sustain conversational interchange (at whatever level of language skills are present) in which there is reciprocal to and from responsiveness to communications of the other person 3. Stereotyped and repetitive language or idiosyncratic use of words or phrases 4. abnormalities in pitch, stress, rate, rhythm and intonation of speech
D	**Restricted repetitive, and stereotyped patterns of behavior, interests and activities, manifest in at least two of the following areas:** 1. an encompassing preoccupation with one or more stereotyped and restricted patterns of interest that are abnormal in content or focus; or one or more interests that are abnormal in their intensity and circumscribed nature although not abnormal in their content or focus. 2. apparently compulsive adherence to specific, non-functional, routines or rituals 3. stereotyped and repetitive motor mannerisms that involve either hand or finger flapping or twisting, or complex whole body movements 4. preoccupations with part-objects or non-functional elements of play materials (such as their odour, the feel of their surface, or the noise or vibration that they generate) 5. distress over changes in small non-functional, details of environment
E	**The clinical picture is not attributable to other varieties of pervasive developmental disorder; specific developmental disorder of receptive language (F80.2) with secondary socio-emotional problems; reactive attachment disorder (F94.1) or disinhibited attachment disorder (F94.2); mental retardation (F70-72) with some associated emotional or behavioral disorder; schizophrenia (F20) of unusually early onset; and Rett syndrome (F84.2).**

Asperger syndrome F84.5	
A.	**A lack of any clinically significant general delay in spoken or receptive language or cognitive development. Diagnosis requires that single words should have developed by 2 years of age or earlier and that communicative phrases be used by 3 years of age or earlier. Self-help skills, adaptive behavior, and curiosity about the environment during the first 3 years should be at a level consistent with normal intellectual development. However, motor milestones may be somewhat delayed and motor clumsiness is usual (although not a necessary diagnostic feature). Isolated special skills, often related to abnormal preoccupations, are common, but are not required for the diagnosis.**
B.	**Qualitative abnormalities in reciprocal social interaction (criteria as for autism). Diagnosis requires demonstrable abnormalities in at least two out of the following four areas:** 1. Failure adequately to use eye-to-eye gaze, facial expression, body posture and gesture to regulate social interaction. 2. Failure to develop (in a manner appropriate to mental age, and despite ample opportunities) peer relationships that involve a mutual sharing of interests, activities and emotions. 3. Lack of socio-emotional reciprocity as shown by an impaired or deviant response to other people's emotions, and/or lack of modulation of behavior according to social context, and/or a weak integration of social, emotional and communicative behaviors 4. Lack of spontaneous seeking to share enjoyment, interests or achievements with other people (e.g. lack of showing, bringing or pointing out to other people objects of interest to the individual).
C	**An unusually intense circumscribed interests, or restricted, repetitive, and stereotyped patterns of behavior, interests, and activities (criteria as for autism, however it would be less usual for these to include either motor mannerisms or preoccupations with part-objects or non-functional elements of play materials). Diagnosis requires demonstrable abnormalities in at least two out of the following four areas:** 1. An encompassing preoccupation with one or more stereotyped and restricted patterns of interests that is abnormal in context or focus; or one or more interests that are abnormal in their intensity and circumscribed nature though not in their content or focus. 2. Apparently compulsive adherence to specific, non-functional routines or rituals. 3. Stereotyped and repetitive motor mannerisms that involve either hand/finger flapping or twisting, or complex whole body movements. 4. Preoccupation with part-objects or non-functional elements of play materials.
D	**The disorder is not attributable to other varieties of pervasive developmental disorder; schizotypal disorder (F21); simple schizophrenia (F20.6); reactive and disinhibited attachment disorders of childhood (F94.1 and .2); obsessional personality disorder (F60.5); obsessive compulsive disorder (F42).**

The classification will soon (in year 2013) be altered by the fifth edition of the Diagnostic and Statistical Manual of Mental disorders (DSM-V). It will combine previously used subtypes, including Asperger syndrome, and the disorders will be treated as a continuum of phenotypes. It is also likely that Rett syndrome will be excluded from the group (Hebebrand and Buitelaar 2011). Differences in the ICD–10 and DSM–IV definitions for the same disorder

have impeded international communication and research efforts. The American Psychiatric Association (APA) and the World Health Organization (WHO) aim to harmonize the upcoming DSM–V and ICD–11 classifications by minimizing or eliminating differences between these two classification systems (First 2009).

In addition to the core symptoms several other clinical findings, which are presented in Table 2, are observed in a significant proportion of individuals with ASD. These impairments include sensory abnormalities, motor signs, sleep disturbance, gastrointestinal disturbance, epilepsy and comorbid psychiatric diagnosis (Geschwind 2009). 10 - 15 % of individuals with autistic features have a medical syndrome which is usually caused by a single gene disorder (Folstein and Rosen-Sheidley 2001). The most common of these are Fragile-X syndrome, tuberous sclerosis, neurofibromatosis, phenylketonuria, Rett syndrome and Angelman´s syndrome.

The diagnosis of autism and intellectual disability is challenging because there is significant overlap between them. Intellectual disability is present in approximately 67 % of individuals with autism and 28 % of patients with intellectual disability have autistic features (Kaufman et al. 2010). Both in autism and in intellectual disability without autism the majority of the cases are males, suggesting a X-chromosomal effect. In fact, some molecular evidence has been obtained for a common genetic background for ASD and X-linked intellectual disability (XLID) (Laumonnier et al. 2004, Betancur 2011) (See chapter 2.4.3).

Approximately 22 % of individuals with autism develop epilepsy later in life. Seizures begin after the age of 10 in the majority of the patients. In some cases the seizures do not start until adulthood (Bolton et al. 2011). Epilepsy in autism is associated with intellectual disability, poorer verbal abilities and it is more common in females than males (Bolton et al. 2011). Autistic individuals with epilepsy have an increased number of relatives with broader autism phenotype which indicate that familial liability to autism raises the risk for epilepsy in the proband (Bolton 2011).

Table 2. Domains of impairment in ASD. Adapted from (Geschwind 2009).

Domain	Autism	Asperger	PDD-NOS	ASD
Social communication	required	required	required	
Language	required	-	variable	
Repetitive, restrictive behaviors	required	required	variable	
Sensory abnormalities	>90 %	80 %	variable	94 %
Developmental regression[a]	15 % - 40 %	?	?	15 % - 40 %
Motor signs[b]	60 % -80 %	60 %	60 %	60 % -80 %
Gross motor delay	10 %	?	?	5 % -10 %
Sleep disturbance	55 %	5 % -10 %	40 %	50 %
Gastrointestinal disturbance[c]	45 %	4 %	50 %	4 % - 50 %
Epilepsy[d]	10 % -60 %	0 % - 5 %	5 % -40 %	6 % -60 %
Comorbid psychiatric diagnosis[e]	70 %	60 %	> 25 %	25 % -70 %

[a]Loss of function in either or both the language or social skills domain.
[b]Motor signs include hypotonia, gait problems, toe walking, and apraxia.
[c]Six months or more of diarrhea, constipation, reflux, or bloating.
[d]The range of epilepsy estimates reflects the presence of other comorbid features, such as concurrent intellectual disability or intellectual disability and cerebral palsy, which significantly increase epilepsy risk (25 %–30 % and 60 %, respectively).
[e]Mood disorders, conduct disorders, aggression, and attention deficit/hyperactivity disorder (ADHD). ADHD symptoms are observed in ∼25 % of children with ASD.

ASDs often show obsessive repetitive symptoms that are characteristic to obsessive-compulsive disorders. OCD (MIM 164230) belongs to the anxiety disorders and is characterized by intrusive senseless thoughts and impulses (obsessions) and repetitive intentional behaviors (compulsions). Individuals with ASD may be more susceptible to high anxiety because of their unique social, behavioral, communicative and sensory difficulties (Bellini 2006). Studies which used Children´s Yale-Brown Obsessive-Compulsive Scale (CY-BOCS; Scahill et al. 1997) have observed that large subgroups of children with ASD engage in significant obsession and compulsions (Scahill et al. 2006, Zandt et al. 2007). 37 per cent of people diagnosed with ASD have comorbid obsessive compulsive disorder (Leyfer et al.

2006). There is a difference in obsessive compulsive symptoms between ASD and OCD. In ASD repetitive thoughts about a specific interest produce largely positive, affective emotions while OCD related obsessions are unpleasant and they are performed to reduce perceived threat (Spiker et al. 2012).

Leyfer et al. (2006) reported that the prevalence rates for other comorbid psychiatric disorders in children with autism are 44 % for specific phobias, 31 % for ADHD, 24 % for major depression and less than 2 % for bipolar disorder. None of the patients in that study met the diagnostic criteria for schizophrenia. Over 10 % of children with autism have a phobia of loud noises. Leyfer et al. (2006) emphasized that it is challenging to determine if a child´s difficulties are due to autism-related manifestations or comorbid psychiatric disorders.

A Finnish study reported co-morbid psychiatric disorders in an Asperger/High functioning (HFA) autism sample set. They observed common (prevalence 74 %) and often multiple comorbid psychiatric disorders in AS/HFA; behavioral disorders were shown in 44 %, anxiety disorders in 42 % and tic disorders in 26 %. Major depressive disorder (MDD) and anxiety disorders as comorbid conditions indicated significantly lower levels of functioning (Mattila et al. 2010).

Several studies have indicated a strong association between familial mood and anxiety disorders, and autism (Piven et al. 1991, Smalley et al. 1995, DeLong 2004, Cohen and Tsiouris 2006). Cohen and Tsiouris (2006) observed that 36% of mothers and 17% of fathers of an autistic child had a lifetime history of a major mood disorder. All of the mothers with recurrent MDD or Bipolar Disorder had a first episode prior to the birth of their autistic child. This study hypothesized that shared risk alleles between autism and depression genes are accompanied by epistatic interactions among these genes that, in turn, modify the expression of each disorder. Depending upon the number and type of these risk alleles individuals are likely to have a specific autism or depression phenotype (Figure 1) (Cohen and Tsiouris 2006).

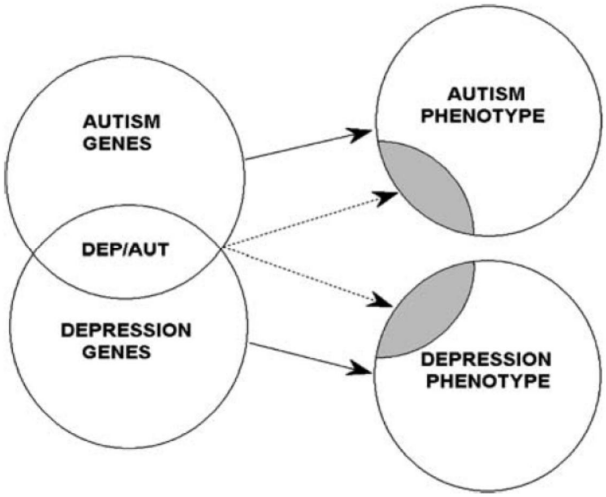


Figure 1. This figure presents a Venn diagram of the hypothesized overlap of risk alleles (DEP/AUT alleles) that are common to both autism and recurrent mood disorders. The dashed arrows emerging from the DEP/AUT overlap represent this modifier gene effect. The shaded regions represent DEP/AUT modified sub-groups of autism and depression. Reprinted with permission. Cohen and Tsiouris, J. Autism Dev. Disord. 2006. ABBREVIATIONS: DEP=depression, AUT=autism.

Concerning anatomical differences to normal children, autism often involves early brain and head overgrowth. Abnormal brain and head overgrowth begins at 9-18 months of age, often concurrently or soon after emergence of clinical signs of autism (Courchesne et al. 2011). Overgrowth and neural dysfunction are evident in multiple brain regions that are involved in higher-order social, emotional, communication, and cognitive development (Courchesne et al. 2011). Figure 2 presents the parts of the brain which are affected by autism.

Male children with autism have on average 67 % more prefrontal neurons than those in the control group. The autistic group also has greater than average brain weight. Pathological increase in neuron numbers may be a key contributor to brain overgrowth in autism. Genes located within copy number variation regions in autism may cause abnormal proliferation of neuronal cells during development (Courchesne et al. 2011). Apoptotic mechanisms during the third trimester and early postnatal life normally remove subplate neurons, which comprise about half the neurons produced in the second trimester and are present only

during cortical development (Kanold 2009). A failure of that early developmental process could also create a pathological excess of cortical neurons (Courchesne et al. 2011).

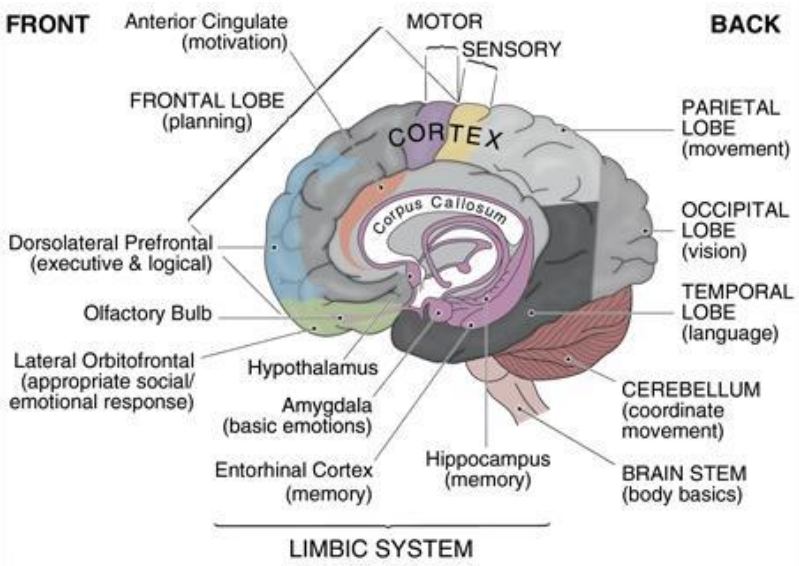


Figure 2. Parts of the brain affected by autism (http://www.autisminformationservice.com/specialreport.html).

2.1.2 Prevalence

The initial assessment for the rate of childhood autism was 4 in 10000 children (Wing et al. 1976). Over time diagnostic criteria have embraced a broader phenotype of ASD and estimates for ASD prevalence have increased from 1/1000 (Bryson et al. 1988) to the current 1/150-1/200 for ASD and 1/500 for the narrow diagnosis of strict autism (Geschwind 2009). In a Finnish epidemiological study the prevalence of ASD was 8.4 in 1000 and that of autism was 4.1 in 1000 (Mattila et al. 2011). A recent review of global prevalence reported the median prevalence of ASD to be 6.2/1000 (Elsabbagh et al. 2012). There were no significant differences between populations. The evidence reviewed did not support a strong impact of ethnic, cultural or socioeconomic factors. However, rates from low- and middle-income countries are mainly lacking (Elsabbagh et al. 2012). Several factors influencing this rise of

prevalence have been proposed including the broadening of diagnostic criteria, better service availability and awareness of ASD in both the lay and professional public. All ASDs are generally more common in males than in females (ratio ~4:1). In high-functioning children with childhood autism, the male-to-female ratio can be as high as 8:1 (Fombonne 2005).

2.1.3 Concordance rates and recurrence risks

ASDs have the highest heritability compared to other neuropsychiatric disorders. In earlier studies monozygotic twins showed significantly higher concordance rates (36–95 %) for autism than did dizygotic twins (0–23 %), which confirmed the strong genetic component of autism (Folstein & Rutter 1977a, Folstein & Rutter 1977b, Steffenburg et al. 1989, Bailey et al. 1995). A recent twin study with a larger study sample estimated slightly lower concordance rates, 58 % in monozygotic pairs and 21 % for dizygotic pairs in males for strict autism. For female twins the concordance rate was 60 % for monozygotic pairs and 27 % for dizygotic pairs. For ASD, the concordance rate was 77 % for monozygotic pairs and 31 % for dizygotic pairs in males. For females these rates were 50 % for monozygotic pairs and 36 % for dizygotic pairs. The study proposed that environmental factors explain 55 % of the liability to autism (Hallmayer et al. 2011).

Earlier studies estimated that the sibling recurrence risk for ASD is 3-10 % (Chakrabarti & Fombonne 2001, Icasiano et al. 2004, Lauritsen et al. 2005). A recent study with large sample and prospective longitudinal design reported that even 18.7 % of infants with at least one older sibling with ASD developed the disorder (Ozonoff et al. 2011). If there are 2 autistic children in the family the recurrence risk for the third child is 25 %. The risk of having one or more features of ASD is 30 % in adult siblings (Folstein and Rosen-Sheidley 2001). Detailed genetic findings in ASD are reported in Chapter 2.4.

2.1.4 Environmental factors

Environmental factors may increase the risk for autism (Newschaffer et al. 2002, Hallmayer et al. 2011). A comprehensive meta-analysis examined over 60 prenatal and neonatal risk factors for autism (Gardener et al. 2011). Factors associated with autism risk were abnormal presentation, fetal stress, umbilical-cord complications, birth injury or trauma, multiple birth, maternal hemorrhage, summer birth, low birth weight, small gestational age, low 5-minute Apgar score, feeding difficulties, congenital malformation, meconium aspiration, neonatal anemia, ABO or Rh-incompatibility and hyperbilirubinemia. Factors not associated in autism risk included assisted vaginal delivery, anesthesia, post-term birth, high birth weight and head circumference (Gardener et al. 2011).

A recent genome-wide sequencing study revealed that paternal age is associated with an increasing rate of *de novo* mutations (Kong et al. 2012). These mutations increase the risk for ASD and schizophrenia, thus advanced paternal age is a risk factor (Kong et al. 2012). Pre-conceptual exposure to environmental mutagens also causes *de novo* mutations. Mercury, cadmium, nickel, trichloroethylene and vinyl chloride are suggested mutagens with increased risk for autism (Kinney et al. 2010). A recent study reported that exposure to traffic-related air pollution during gestation and the first year of life is associated with autism (Volk et al. 2013).

It has been noticed that autism prevalence is higher in towns with a prospering IT industry, like Eindhoven and Silicon Valley, than in other towns of similar size (Buchen 2011, Roelfsema et al. 2012). Baron-Cohen (2012) reported that scientists and engineers have an increased risk of having a child with autism. His theory suggests that mating of people with technical minds passes down linked groups of genes that confer useful cognitive talent, but also increase the child´s chances of developing autism. These genetic variants underlying autism may persist because they are co-inherited with variants underlying certain cognitive talents typical to both autism and the technical mind (Baron-Cohen 2012). Other factors, like higher education and the older age of parents could also explain the high prevalence of autism in these IT centers (Buchen 2011).

Inflammations are a possible trigger to autism. Congenital cytomegalovirus infection and rubella have been reported to be associated in autism (King 2011). On the other hand, the postulated risk of the MMR (measles, mumps and rubella) vaccinate and vaccines containing the preservative thiomersal to increase autism has been disproven by many studies (Farrington et al. 2001, Parker et al. 2004, Fombonne et al. 2006, Richler et al. 2006).

Maternal factors in pregnancy linked to increased autism risk include valproic acid, thalidomide, alcohol, depression, schizophrenia, OCD, autoimmune disease, stress, allergic reaction, and hypothyroidism (King 2011). All of these factors may initiate the expression of genes which are sensitive to retinoid acid and/or estradiol, whether by directly increasing (promotion) or reducing the production of human alpha-fetoprotein (HFAP) (King et al. 2011). Impaired HFAP level has a significant effect on the foetal brain.

Other environmental factors which are associated with increased risk for autism include residence in regions that are urbanized, located in higher latitudes or experience high levels of precipitation. These are all associated with decreased sun exposure and increased risk for vitamin D deficiency (Kinney et al. 2010). Prenatal lack of vitamin D increases the risk of autism (Eyles et al. 2012). Respectively, children with ASD have been reported to have significantly lower levels of D vitamin in their plasma than controls (Eyles et al. 2012). D vitamin is synthesized in the skin triggered by sunlight and plays an important role in repairing DNA damage and protecting against oxidative stress, which is a key cause of DNA damage. Vitamin D deficiency contributes to higher mutation rates and impaired repair of DNA (Kinney et al. 2010). Skin pigmentation is also known to directly decrease the actinic production of vitamin D3 (Abrams 2002). Accordingly ASD prevalence increases in infants from migrant mothers with dark skin compared to offspring from lighter skinned migrants (Dealberto 2011).

A higher prevalence of autism has been reported in children with conception in November in California (Mazumdar et al. 2012). The reason for this was unclear but might be caused by seasonal flu, which in California generally peaks in two weeks of January. A similar occurrence in schizophrenia was associated with exposure to influenza during gestation (Mednick et al. 1994). Mazumdar and colleagues (2012) reported that with children who were conceived in November in California, the second trimester coincided with the pollen

season. This could be explained by maternal asthma in the second trimester which is associated with an increased risk of autism. Another study reviewed more than 6.5 million births in California and reported that children conceived in winter are more likely to develop autism than those conceived in other months of the year (Zerbo et al. 2011). Environmental factors have been proposed to increase the prevalence of ASD 10-fold over the past 4 decades. The reason for this increased prevalence rate of ASD may also be due to broader diagnostic criteria and more frequent reporting of ASD (Massing-Courtney et al. 2013).

2.1.5 Treatment

Longitudinal studies indicate that the long-term prognosis of children diagnosed with autism is poor, with most of them not obtaining independent status (Geschwind 2009). At present, there are no curative or psychopharmacological therapies to effectively treat all symptoms of the disorder. 45 % of children with ASD are treated with psychotropic medication. Currently only risperidone is FDA-approved for the treatment of autism (Geschwind 2009). Risperidone is a medicine which is used as an asymptomatic treatment for irritability.

Early intensive behavioral intervention (EIBI), a treatment based on the principles of applied behavior analysis delivered for multiple years is one of the more well-established treatments for ASD (Reichow et al. 2012). It aims to lessen the impact of symptoms in ASD like behavioral challenges and cognitive difficulties. The method breaks behaviors down into subcategories and teaches each specific subcategory through repetition, prompts and positive reinforcement. Several publications and meta-analyses indicate that EIBI may improve the quality of life and level of functioning for children with ASD (Howlin et al. 2009). However it does not produce significant changes in all areas of a child's functioning and it is not effective for all children with ASD (Howlin et al. 2009).

The need for functional medication is crucial. Also, the identification of genetic and other biomarkers for specific subtypes and their relation to treatment response constitute critical areas of research. It is also essential to gather more efficient pharmacologic and cognitive-

behavioral therapies and a better notion of which therapy is most appropriate for which child.

2.2 STRUCTURE OF THE HUMAN GENOME

The human genome is stored within the 23 chromosome pairs of the cell nucleus and the mitochondrion. Chromosome pairs comprise 22 pairs of autosomes and one pair of sex chromosomes (XX in females and XY in males). The total length of the genome (22 autosomal chromosomes and the X chromosome) is 3615 cM (Kong et al. 2002).

The international Human Genome Sequencing Consortium (Lander et al. 2001) and a company, Celera Genomics (Venter et al. 2001), produced the first complete sequences of the human genomes. These sequences covered about 90 % in 2001 and were later refined to 99 % (International Human Genome Sequencing Consortium 2004). Since then the International Hapmap Project Consortium 2010, the 1000 Genomes Project Consortium 2010 and many other studies have updated and complemented these studies (Lander 2011).

The haploid human genome contains approximately 20,500 protein-coding genes (Clamp et al. 2007). Protein-coding sequences account for only a very small fraction of the genome (approximately 1.5 %), and the rest is associated with regulatory DNA sequences, introns, non-coding RNA molecules and sequences with indeterminate functions (Lander 2011).

The sequences of any two human beings are about 99.9 % identical (Lander 2011). The human genome consists of just over three billion nucleotides and 0.1 percent of difference represents altogether three million variants between any pair of individuals. The vast majority of these variants have no functional significance. Sequence variations in the human genome include single nucleotide polymorphisms (SNPs), microsatellites and structural variations. The variations can be used as genetic markers to study differences between individuals and populations. On average there are 60 new mutations in every individual, compared to his/her parents (Conrad et al. 2011). Even monozygotic twins have genetic differences due to mutations occurring during development, e.g. with differences in copy number variation profiles (Bruder et al. 2008).

It is estimated that a total of 10 to 30 million SNPs exist in the human population of which at least 1 % are functional. About 9 to 10 million are common SNPs with minor allele frequency (MAF) above 0.05 (International HapMap Consortium et al. 2007). The density of common SNPs is roughly about one per 300 bases and it varies along the genome. The mutation rate of the SNP is low, approximately 1.0 to 2.5 x 10^{-8} mutations per nucleotide per generation (Nachman & Crowell 2000, 1000 Genomes Project Consortium 2010). SNPs are useful markers for association studies because most SNPs have likely arisen in a single mutation event in the past.

Mechanisms behind genome variations are mutations and recombination. Mutations can be classified as spontaneous or induced mutations. Spontaneous mutations are caused by internal factors such as occasional random errors in DNA replication during cell division. Induced mutations are a result of external factors such as mutagenic radiation or chemicals. Based on the effect on structure mutations are classified as point mutations, insertions or deletions. They can also be classified as loss-of-function, gain-of-function or neutral mutations based on the effect they have on gene function. Genetic recombination breaks and rejoins DNA strands to form new molecules of DNA encoding a novel set of genetic information. Recombination can occur between similar molecules of DNA, as in homologous recombination, usually occurring during mitosis. It can also occur between dissimilar molecules, as in non-homologous end joining in meiosis which creates novel combinations of the parental genomes which are then transmitted to the offspring.

Microsatellites are also known as short tandem repeats (STRs) and are repeated sequences of 2-6 base pairs of DNA which are flanked by unique sequences on both sides. The number of repeats can vary between individuals. The human genome contains 150,000 microsatellites. The mutation rate is much higher compared to SNPs, in the range of 1.5 x 10^{-3} per locus per generation (Butler 2006). Microsatellites are highly polymorphic which have made them a useful tool in forensics and gene mapping.

Copy-number variations (CNVs) are alterations that result in an abnormal number of copies of one or more sections of DNA. CNVs correspond to relatively large regions of the genome that have been deleted or duplicated in certain chromosomes (Figure 3). Locus specific mutation rates for genomic rearrangement range between 10^{-4} and 10^{-5}, and thus are more

frequent than point mutations (Lupski 2007). To date, 11,700 CNVs overlapping over 1000 genes have been revealed (Redon et al. 2006, Conrad et al. 2006).

CNVs account for roughly 13 % of human genomic DNA and each variation ranges from about one kilobase (1000 nucleotide bases) to several megabases in size (Stankiewicz & Lupski 2010). It is estimated that approximately 0.4 % of the genomes of unrelated people differ because of copy number variations (Kidd et al. 2008).

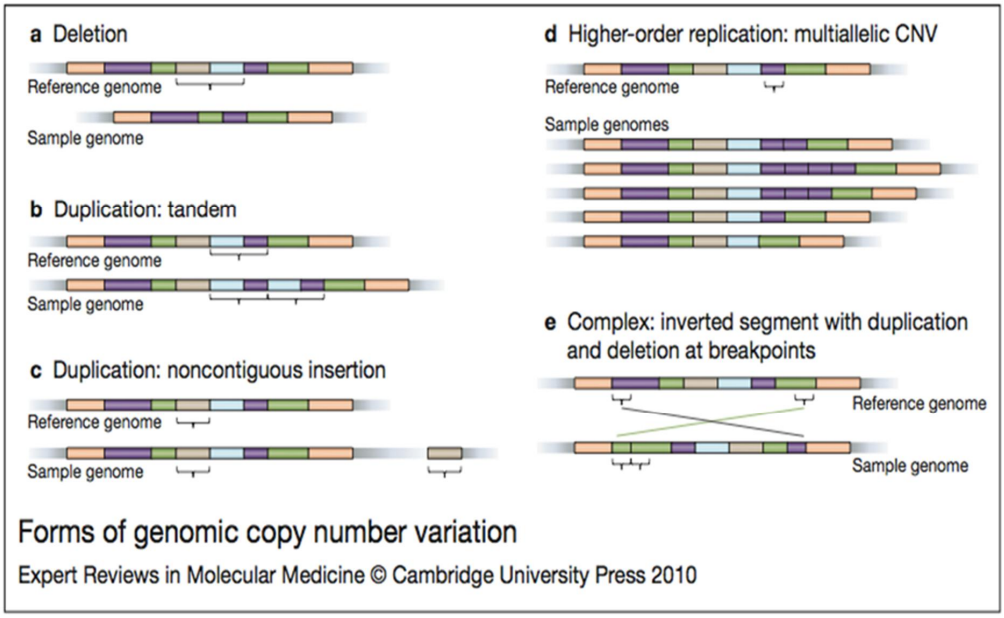


Figure 3. Forms of genomic copy number variation. Reprinted with permission. Lee and Scherer, Expert Rev. Mol. Med. 2010.

CNVs can be inherited or *de novo*. Large *de novo* CNVs are likely to be more often causative for diseases, especially if dosage-sensitive genes or regulatory sequences are affected by the genomic rearrangement (Stankiewicz & Lupski 2010). Mechanisms for the formation of copy number variations are non-allelic homologous recombination, non-homologous end joining and replication-error mechanisms (Stankiewicz & Lupski 2010).

CNVs have been observed to be enriched in genes encoding secreted, olfactory, and immunity associated proteins. The overrepresentation of immunity and chemo-sensation genes in human CNVs could imply that they might have been selectively favored in recent

evolution in fighting infection and sensing our environment (Nguyen et al. 2006). Rare, large, often *de novo* CNVs have been linked to varying phenotypes including autism, developmental disorders and schizophrenia (de Vries et al. 2005, Sebat et al. 2007, Stefansson et al. 2008).

2.3 GENE MAPPING STRATEGIES

In this chapter an overview of the most essential methods is given, with emphasis on the methods used in this thesis. The purpose of genetic mapping is to identify a genetic variant which influences the phenotype of interest. The most commonly used markers are microsatellites and SNPs. Microsatellites are informative due to their high polymorphism content. SNPs compensate for their higher density in the genome which allows numbers of them to be genotyped.

Figure 4 illustrates the relationship between the frequency of genetic variants and disease susceptibility. Higher penetrance, lower frequency variants that are not detected by current GWA approaches are amenable to new high-throughput sequencing efforts and might provide valuable information for genetics of complex disorders (McCarthy et al. 2008).

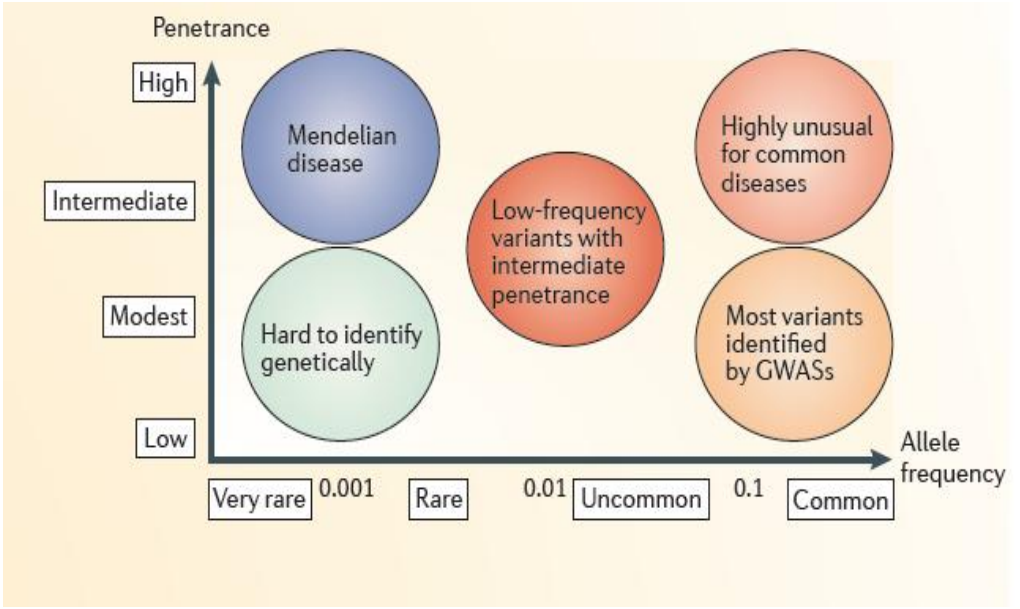


Figure 4. Genetic variants and disease susceptibility. Reprinted with permission. McCarthy et al. Nat.Rev.Genet. 2008.

2.3.1 Linkage analysis

Linkage is the tendency for a causal variant and a genetic marker to be inherited together because of their location near one another on the same chromosome. The aim of linkage analysis is to reveal those chromosomal loci, which harbor genetic variant(s) predisposing to a certain phenotype. For this purpose, the genome is covered by genotyping a number of genetic markers, and analyzed for co-segregation of the phenotype and the markers. Parametric linkage analysis is suitable for phenotypes where the inheritance pattern is known, whereas nonparametric linkage analysis is applicable also for phenotypes with less clear patterns of inheritance. Parametric linkage analysis follows co-segregation of two genetic factors, the marker and the phenotype, at specific loci in pedigrees using the frequency of meiotic recombination as an estimate of genetic distance. Parametric linkage analysis produces a logarithm of the odds (LOD) score that statistically estimates whether the marker and postulated disease loci are likely to lie near each other on a chromosome. A LOD score of 3 or more is generally taken to indicate that the two loci are linked.

Nonparametric linkage analysis detects markers in which affected relatives share alleles more often than expected by chance. Linkage analysis benefits from large pedigrees because they provide information from many meioses. On the other hand, the accuracy of the method is limited due to the requirement of informative meiosis: it can only identify rather large genomic regions, which is why other methods are warranted to narrow down the region of interest.

2.3.2 Association analyses

Association analyses test whether single locus allele frequencies differ between cases and controls. Individuals who share a disease mutation through common descent are likely to also share a haplotype of alleles surrounding the mutation. Typically, one does not find the direct causal allele but more probably markers in linkage disequilibrium (LD) with it. Thus, LD is a term for non-random association of alleles at two or more loci. In the genome LD exist as haplotype blocks in varying length and consist of regions with low recombination rates.

Especially with complex traits, genome-wide association studies (GWAS) are an effective tool to detect disease associated genes. In GWAS, allele frequencies of hundreds of thousands of SNPs are compared between cases and controls. Furthermore, the number of study subjects tends to be high: even in the order of thousands. Though rather labor-intensive and costly, the total number of GWA studies published to date is approaching 1459 (www.genome.gov/GWAStudies, accessed 11/12/2012).

One problem with the case-control design is that genotype and haplotype frequencies vary between ethnic or geographic populations. If the case and control populations are not well matched for ethnicity or geographic origin then false positive association can occur because of these differences. Family-based association designs aim to avoid this issue by using the parents as “pseudo”- controls for the case. With family-based association, the most commonly used test is the transmission disequilibrium test (TDT). TDT measures association by transmission of alleles from parents to affected offspring. If an allele increases the risk of having a disease then that allele is expected to be transmitted from parent to affected offspring more often than other alleles.

2.3.3 Scanning methodologies for CNV

Numerous array-based platforms for CNV detection exist utilizing the technologies based on Array -Comparative Genomic Hybridization (aCGH) and SNP genotyping or both (Haraksingh et al. 2011). In this thesis we utilized Illumina the SNP array in both the GWA study and CNV detection.

Array comparative genomic hybridization (aCGH) is a molecular method for the discovery of genomic imbalances (Figure 5). To identify copy number gain or loss from DNA, samples from 2 different individuals, usually from the test and a control individual, are labeled with distinctive fluorophores. The samples are simultaneously hybridized with a set of hybridization targets, typically long oligonucleotides, *in situ* to the same well in the array platform. After this, the platform is scanned and the scan gives the intensity ratio of the two fluorescence signals that represent the average copy number ratio between the two genomic DNA samples. The fluorescence signals are analyzed by computer software, and a gain or loss of fluorescence signal intensity from the test DNA implies a gain or loss of test DNA copy numbers (Choy et al. 2010).

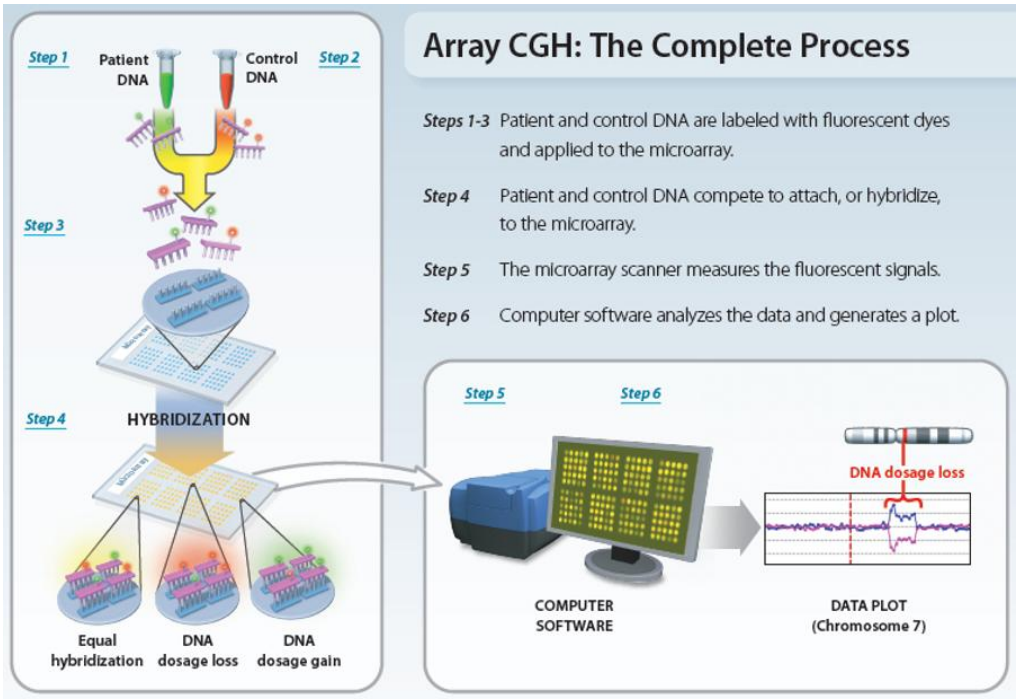


Figure 5. Diagram of the microarray-based comparative genomic hybridization (aCGH) process (Theisen 2008).

High-throughput array technologies for identifying SNPs can also be used to identify CNVs. The SNP array consists of probes for the detection of both SNPs and CNVs. First, the sample DNA is labeled with fluorescent dye and then the sample is hybridized with a single stranded genetic probe of interest. In the SNP genotyping array only one sample is hybridized per microarray, therefore in this method control DNA is not needed. Instead of comparing intensities from two samples, the $\log_2$ ratios are generated by clustering the intensities measured at each probe across many samples (Alkan et al. 2011). The intensities from each microarray spot of the platform are generated by the difference in hybridization between the sample DNA and probe. The hybridization differs because of the presence or absence of different SNP's in the sample DNA. Statistical algorithms have been developed to analyze the data of the intensity ratios, with which the CNV statuses of the tested individual can be determined.

2.3.4 Sequencing-based methods

Traditional Sanger sequencing (Sanger and Coulson 1975) has typically been used for the mutation search in candidate genes identified through linkage or association analysis, or for genes of interest based on functional studies. The aim of the sequencing is to identify the actual risk variants present in cases but not in controls. Recent development of high-throughput sequencing methods has enabled genome wide sequencing of large cohorts. Whole genome resequencing produces a massive amount of data, and needs to be followed by bioinformatic analyses to filter out risk variants from benign polymorphisms. With costs rapidly decreasing it is now possible to sequence exomes or entire genomes of individuals for the purpose of research.

2.4 MOLECULAR GENETIC STUDIES IN ASD

Many experimental methods have been used to identify ASD associated genes, including the earlier linkage analyses and candidate gene association or experimental studies in animal models as well as the more recent genome-wide association studies, genome-wide CNV studies and expression profiling. It is estimated that the genetic cause can be identified in up to 20 % of ASD cases (Vorstman & Ophoff 2013). The results of autism studies have been compiled in e.g. AutismKB (http://autismkb.cbi.pku.edu.cn/), which is an evidence-based knowledgebase of ASD genetics. In this chapter the major molecular genetic findings in ASD are presented. Figure 6 illustrates ASD related loci reported in linkage, GWA, CNV and candidate gene studies.

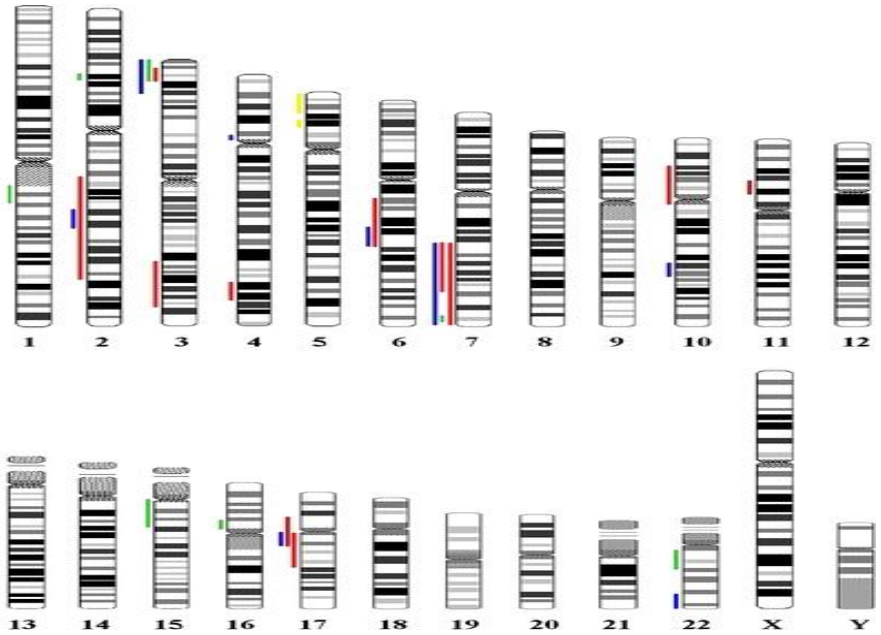


Figure 6. Replicated findings of linkage (red bars), Genome-wide association (yellow bars), copy number variation (green bars) and candidate gene (blue bars) studies. Reprinted with permission. Freitag et al. Eur. Child Adolesc. Psychiatry, 2010.

2.4.1 Linkage studies

The first genome-wide screen in autism was published in 1998 showing the highest linkage peak at chromosome 7q22-q31 (IMGSAC 1998). Since that as many as 158 linkage regions have been reported to associate with ASD, comprising all the chromosomes (Xu et al. 2012). Replicated linkage in ASD, as detected by at least two independent studies has been obtained in regions 2q21–33, 3q25–27, 3p25, 4q32, 6q14–21, 7q22, 7q31–36, 11p12–13 and 17q11–21 (Freitag et al. 2010). A meta-analysis confirmed the region 7q22–32, and reported suggestive evidence for linkage to 10p12–q11.1 and 17p11.2–q12 (Trikalinos et al. 2006). Several genome-wide linkage studies in ASD have been performed with limited concordance of linked loci, reflecting either numerous genes of weak effect and/or sample heterogeneity. The summary of the best replicated linkage results in ASD are shown in Table 3.

The largest linkage study today is by the Autism Genome Project Consortium (AGP) (Szatmari et al. 2007), which reported the linkage and CNVs from altogether 1181 families with at least 2 affected individuals each. Chromosomal region 11p12-13 reached suggestive but not genome wide significance. This locus has not been implicated in previous linkage scans. Locus 9p24 also showed suggestive linkage. Both of these chromosomal regions contain glutamate transporter genes which are candidate genes in ASD. *SLC1A1* (solute carrier family 1, member 1) locates at 9p24 and *SLC1A2* (solute carrier family 1, member 1) at 11p12-13. At 11p12-13, stratification of the sample set into families with only male cases and families with also female cases yielded more significant peaks in the female containing families. Based on the CNV analyses, further stratification of families was performed to decrease heterogeneity, and suggestive linkage evidence was observed also for 15q23–25.3, in addition to 11p12–p13.

The lack of genome-wide significant linkage findings is probably due to the genetic heterogeneity of the studied samples. Genetic and diagnostic heterogeneity is an underlying problem specifically in international collaborations where hundreds of samples are collected from many different countries. To decrease genetic heterogeneity, study samples have been divided into endophenotypes. Endophenotype is a term which is used to parse behavioral symptoms into more stable phenotypes with a clear genetic connection. This approach, however, reduces the power of the analysis due to smaller sample size. Stratification studies using a clinical sub-phenotype include two Finnish studies concerned with Asperger pedigrees only (Ylisaukko-oja et al. 2004, Rehnström et al. 2006). Other subgroups in ASD linkage studies comprise, for instance, sex (Szatmari et al. 2007), age at first word (Alarcon et al. 2002, Alarcon et al. 2005, Schellenberg et al 2006), obsessive compulsive behavior (Buxbaum et al. 2004) and social responsiveness (Duvall et al. 2007).

The original Finnish genome-wide linkage study in autism revealed the strongest linkage in locus 3q25-27 (Auranen et al. 2002). Other putative linkage findings were observed at 1q21-23 and Xq11. Based on this first linkage study, fine mapping within the best loci were performed. Fine mapping confirmed linkage peaks at chromosomal regions 1q23 (Kilpinen et al. 2009) and at Xq13.1 (Ylisaukko-oja et al. 2005). In genome-wide studies with extended AS

families the most solid linkage finding was observed at 3p14-24 (Ylisaukko-oja et al. 2004, Rehnström et al. 2006).

In many studies the same family material has been included to the linkage and association analyses and in most of these studies linkage peaks did not overlap with the association peak. This suggests that the linkage peaks contain rare, highly penetrant risk alleles while associated alleles detected in GWA studies are common variants with low risk. SNP data used in association analyses should be utilized in linkage analyses by characterizing genome wide SNP data through homozygosity and haplotype sharing analysis within families.

Table 3. Summary of linkage peaks in autism spectrum disorders with support from independent studies. Listed are loci obtaining a logarithm of odds (LOD) score > 3 in at least one study, and a LOD > 2 in at least one additional study. Table adapted from Abrahams and Geschwind (2008).

Locus	Phenotype	Cohort	Best marker	References
1q21-q23	ASD	FIN	D1S1653	Auranen et al. (2002)
	AS	FIN	D1S848	Ylisaukko-oja et al. (2004)
2q24-q31	ASD	IMGSAC	D2S2188	IMGSAC (2001)
	ASD	US	D2S364	Buxbaum et al. (2001)
3q25-27	ASD	FIN	D3S3037	Auranen et al. (2002)
	ASD	UTAH	rs1402229	Coon et al. (2005)
5p13-p14	ASD	AGP	S1968011	Szatmari et al. (2007)
	ASD	AGRE	D5S2494	Liu et al. (2001)
	ASD	AGRE	D5S1473	Yonan et al. (2003)
7q22-q31	ASD	IMGSAC	D7S477	IMGSAC (1998)
	ASD	IMGSAC	D7S477	IMGSAC (2001)
7q34-36	Age at first word	AGRE	D7S1824 and D7S3058	Alarcon et al. (2002)
	Age at first word	AGRE	D7S2426	Alarcon et al. (2005)
9q33—q34	Age at first word	CPEA	S9S164	Schellenberg et al. (2006)
	ASD	AGP	rs536861	Szatmari et al. (2007)
11p12-13	ASD	AGP	rs1358054	Szatmari et al. (2007)
	SRS score	AGRE	ATA34E08	Duvall et al. (2007)
17q11-21	ASD	AGRE	D17S1294	Stone et al. (2004)
	ASD	AGRE	D17S2180	Cantor et al. (2005)
	ASD	AGRE	D17S1800	Yonan et al. (2003)

ABBREVIATIONS: FIN=Finland, IMGSAC=International Molecular Genetic Study of Autism Consortium, AGRE=Autism Genetic Resource Exchange, CPEA=Collaborative Programs of Excellence in Autism Network at the National Institute of Health, AGP=Autism Genome Project, SRS=Social Responsiveness Scale.

2.4.2 Genome-wide association studies

Four GWA studies to date (December 2012) have reached genome-wide significant associations with ASD. Wang and colleagues (2009) performed a GWA study for a cohort of 780 families (3101 subjects) from the Autism Genetic Resource Exchange (AGRE) and a case-control cohort of 1204 affected individuals and 6491 controls. The combined sample set consisted of more than 10000 subjects of European ancestry. The most significant SNP, rs4307059 on chromosome 5p14.1, reached a P-value of 2.1×10^{-10} with the combined dataset. This SNP locates between genes cadherin 9 (*CDH9*) and cadherin 10 (*CDH10*) which both encode neuronal cell adhesion molecules. A recent study identified functional noncoding RNA *MSNP1AS* (moesin pseudogene 1, antisense) which is transcribed within the linkage peak at 5p14.1 (Kerin et al. 2012). It is 94 % identical and antisense to the X chromosomal *MSN* gene that encodes the moesin protein which regulates neuronal architecture. Individuals with ASD who carry a T allele of SNP rs4307059 showed increased expression of *MSNP1AS* in postmortem samples of the temporal cortex. *MSNP1AS* noncoding RNA binds to *MSN* and could regulate the level of moesin in human cell lines and thus contribute to ASD risk.

Weiss and colleagues (2009) reported the results of a genome wide linkage and association scan of 1031 multiplex autism families with altogether 1553 affected individuals. They identified a suggestive association with ASD on chromosomes 6q27 and 20p13. Genotyping of the top hits with additional families revealed significant association of autism with a SNP on chromosome 5p15 between *SEMA5A* (semaphoring 5A) and *TAS2R1* (taste receptor type 2 member 1) ($P=2 \times 10^{-7}$). Expression analyses with brain tissue samples of 20 cases and 10 controls showed that expression of the *SEMA5A* gene was reduced in the brains of autistic patients. This study had partly overlapping samples with Wang et al. (2009) but both studies also included unique datasets. In contrast to Wang et al. (2009), Weiss et al. (2009) did not detect a significant association of SNP rs4307059 and ASD on chromosome 5p14.1.

The Autism Genome Project (AGP) performed a genome wide scan by genotyping 1 million SNPs in 1558 ASD families (Anney et al. 2010). The AGP consortium represents more than 50 centers in North America and Europe. In this study SNP rs4141463 at the *MACROD2* (MACRO

domain containing 2) gene on chromosome 20p12.1 reached genome wide association significance threshold $P < 5 \times 10^{-8}$ within strict diagnosis and European ancestry. After this first stage, the Autism Genome Project added 1301 ASD families at Stage 2 and performed a GWA study with altogether 2705 ASD families (Anney et al. 2012). The strongest association was now detected at SNP rs1718101 in European individuals with higher IQ ($P=7.8 \times 10^{-9}$). This SNP locates in the *CNTNAP2* (contactin associated protein-like 2) gene on chromosome 7q35. Association to the *MACROD2* gene was not detected at Stage 2.

The GWA studies described here did not replicate each other´s findings. Currently, common gene variants explain less than 5 % of the risk of ASD (Connolly et al. 2012). Obviously, the genetic complexity and heterogeneity of ASD individuals hampers GWA studies. Connolly and colleagues (2012) aimed to reduce this complexity by targeting to endophenotypes. This GWA study of 2165 ASD individuals examined the association between genomic loci and a specific endophenotype of ASD. Assessment tools were the ADI-R (Autism Diagnostic Interview-Revised), ADOS (Autism Diagnostic Observation Schedule) and SRS (Social Responsiveness Scale). They identified a number of loci with significant associations. Genes showing association of specific endophenotypes were *KCND*2 (potassium voltage-gated channel, Shal-related subfamily, member 2) in overly serious facial expression, *NOS2A* (nitric oxide synthase 2, inducible) in loss of motor skills and *NELL1* (nel-like 1 isoform 1 precursor) in faints, fits, or blackouts. The replication of this strategy is warranted to confirm the findings.

The Psychiatric Genomics Consortium (PGC) which conducts meta-analyses of genome wide genetic data for psychiatric diseases detected genome-wide significant associations with schizophrenia for seven loci (Schizophrenia Psychiatric Genome wide Association Study (GWAS) Consortium 2011) in a very large material (>20 000) size. The PGC is also incorporating ASD samples and it is assumed that new genome-wide significant loci in ASD will be detected with large enough study material.

2.4.3 Candidate gene studies

According to AutismKB, 434 genes related to autism have been detected (Xu et al. 2012). These genes are mapped to pathways including neuroactive ligand-receptor interaction, synapse transmission and axon guidance. Positional and functional evidence to identify autism candidate genes is gained from GWA studies, genome wide CNV studies, linkage analyses, low-scale genetic association studies, expression profiling, mouse models and other experimental studies. In addition, the results between studies vary, -mainly because of the differences in sample sizes, ethnic groups and phenotypes. Some studies include all ASD individuals while others focus on infantile autism only. This thesis focuses on the autism candidate genes *SHANK2*, *SLC1A1*, *AVPR1A* and also a few known genes underlying intellectual disability. These genes are reviewed in the following section.

The genetic causes of ASD are diverse (Betancur 2011), but the main category of genes associated with the disorder is related to the development and function of neuronal circuits (Toro et al. 2011, Gilman et al. 2011). Rare mutations in *NLGN3* (neuroligin 3)*, NLGN4X* (neuroligin 4, X-linked) (Jamain et al. 2003), *SHANK3* (SH3 and multiple ankyrin repeat domains 3) (Durand et al. 2007), *SHANK2* (SH3 and multiple ankyrin repeat domains 2) (Berkel et al. 2010) and *NRXN1* (neurexin 1) (Ching et al. 2010) have revealed pathways affecting synaptic functions in autism (Figure 7). These proteins play a crucial role in the formation and stabilization of synapses (Südhof 2008), as well as in synaptic homeostasis (Yu and Goda 2009). *SHANK2* and *SHANK3* code for scaffolding proteins located in the postsynaptic density (PSD) of glutamatergic synapses. Deletions of *SHANK3* on chromosome 22q13 are one of the major genetic abnormalities in neurodevelopmental disorders (Cooper et al. 2011): mutations of *SHANK3* have been identified in patients with ASD, intellectual disability (ID) and schizophrenia (Durand et al. 2007, Moessner et al. 2007, Gauthier et al. 2010, Hamdan et al. 2011). Mutations of *SHANK2* have also recently been reported in both ASD and ID (Pinto et al. 2010, Berkel et al. 2010). Differences in the clinical outcome of mutation carriers has been attributed to the presence of still uncharacterized additional genetic, epigenetic and/or environmental factors (Persico and Bourgeron 2006). Genetic epistasis between the NRXN-NLGN-SHANK pathway is suggested to have a significant role in

ASD. Figure 7 illustrates the ASD associated genes involved in neurodevelopment and synaptic functions.

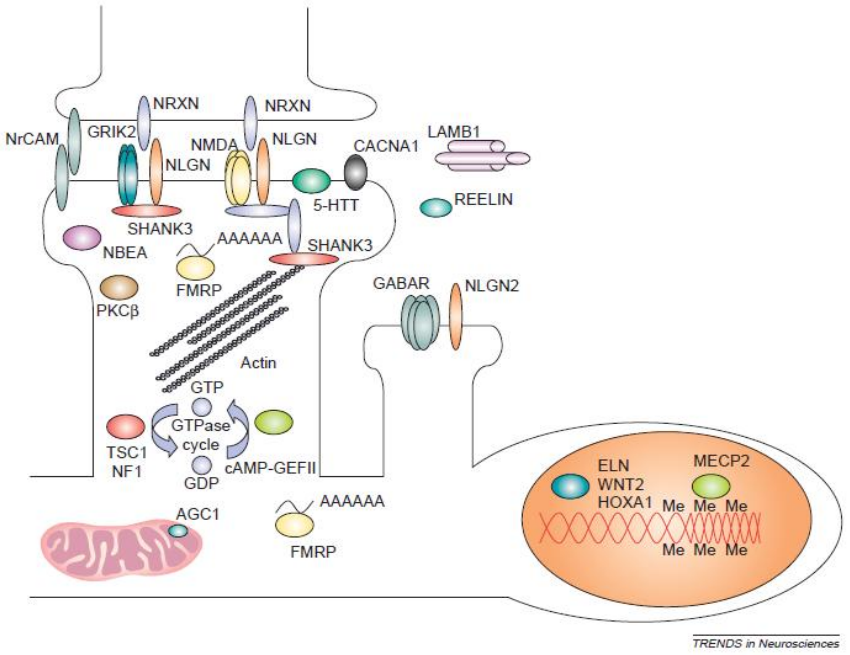


Figure 7. Proteins that are altered in function or amount in ASD. Reprinted with permission. Persico and Bourgeron, Trends Neurosci. 2006.

The 9p24 locus is a common candidate for obsessive compulsive disorder and autism (Hanna et al. 2002, Willour et al. 2004, Szatmari et al. 2007). This region contains the *SLC1A1* gene which encodes a neuronal glutamate transporter EAAC1 (excitatory amino acid carrier 1). *SLC1A1* is expressed in the cortex, striatum and thalamus. Neuroimaging studies have found that neurotransmitter concentrations are lower in the anterior cingulate and greater in the caudate in OCD patients compared to controls. Glutamatergic levels in the caudate decrease to control level following treatment with serotonin re-uptake inhibitors (Rosenberg et al. 2000, Rosenberg et al. 2004). Several studies of OCD have detected a significant association at single nucleotide polymorphisms (SNPs) covering the *SLC1A1* gene region (Arnold et al. 2006, Dickel et al. 2006, Stewart et al. 2007, Shugart et al. 2009, Wendland et al. 2009, Samuels et al. 2011). As stated in Chapter 2.1.1 ASD patients often manifest repetitive symptoms similar to OCD patients (Bodfish et al. 2000, Russell et al. 2005, Leyfer et al. 2006).

In OCD, an autistic subtype with social deficits has been suggested (Bejerot et al. 2001, Bejerot 2007). Recently, some evidence for the association of *SLC1A1* in male patients with autism (Brune et al 2008) and anxiety in autism (Gadow et al. 2010) has been reported.

The receptor coded by *AVPR1A* (arginine vasopressin receptor 1A) gene, V1aR, mediates the influences of the arginine vasopressin (AVP) hormone in the brain, which has a prominent role in higher cognitive functions, such as memory and learning (Fink et al. 2007). *AVPR1A* mediates many social, emotional and behavioral traits, including pair bonding, parenting, sibling relationships, altruism (Donaldson and Young 2008) and musical aptitude (Ukkola et al. 2009, Ukkola-Vuoti et al. 2011). Therefore, it has been suggested as an important candidate gene for ASD and a number of studies have observed the association with autism (Kim et al. 2002, Wassink et al. 2004, Yirmiya et al. 2006, Yang et al. 2010a, Yang et al. 2010b, Tansey et al. 2011). Notably, practically all associations found with the *AVPR1A* gene are located in the promoter region of the gene. In this thesis the promoter of the *AVPR1A* gene was studied in detail.

At least 50 genes identified are associated with syndromic X-linked ID and over 40 genes with non-syndromic X-linked ID (Betancur 2011). Several genes first reported in syndromic conditions were later detected in subjects with non-syndromic forms e.g. *ARX* (aristaless related homeobox)*, CASK* (calcium/calmodulin-dependent serine protein kinase)*, JARID1C* (jumonji, AT rich interactive domain 1C) and *ATRX* (alpha thalassemia/mental retardation syndrome X-linked) (Betancur 2011). Several distinct, individually rare genetic causes in ASD have been detected which resemble the genetic architecture of ID. In fact all known genetic variants associated with ASD are also causes of ID, indicating a common genetic basis for these two neurodevelopmental disorders (Betancur et al. 2011). Many well recognized ID genes can also predispose to ASD, with or without ID. Genes found both in ID and autism are e.g. *PTCHD1* (patched domain containing 1), *JARID1C, GRIK2* (glutamate receptor, ionotropic, kainate 2), *IL1RAPL1* (interleukin 1 receptor accessory protein-like 1)*, SHANK3 and NLGN4X* (Kaufman et al. 2010, Betancur 2011). Known X-linked genes associated with ID and/or ASD are presented in Figure 8.

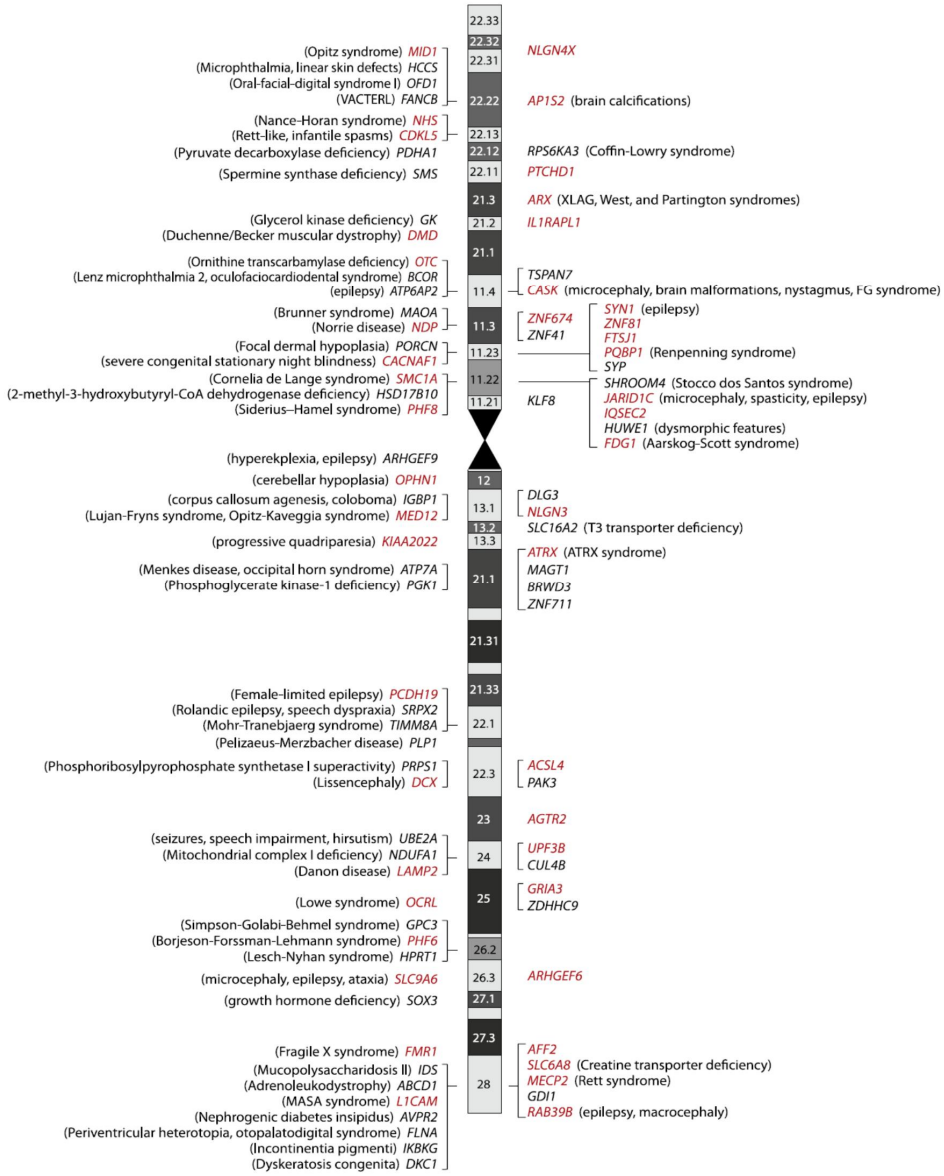


Figure 8. Genes implicated in syndromic and/or non-syndromic forms of X-linked intellectual disability (XLID) and their localization on the X chromosome. Genes that have been reported to be mutated in ASD are highlighted in red. Genes that cause syndromic forms of XLID are shown on the left; those that can cause non-syndromic forms are on the right. The distinction between syndromic and non-syndromic genes is not precise, and for several genes on the right, mutations have been reported in families with syndromic as well as non-syndromic XLID; the syndromic presentation is indicated in parentheses. Reprinted with permission. Betancur, Brain. Res. 2011.

The most common single gene disorder in ASD is Fragile X syndrome caused by mutation in the *FMR1* (fragile X mental retardation 1) gene, which is present in about 2 % of ASD cases. Other monogenic disorders (mutated gene in brackets) described in ASD are Rett syndrome (*MECP2, CDKL5, FOXG1*), tuberous sclerosis (*TSC1, TSC2*) and neurofibromatosis (*NF1*) (Betancur 2011).

2.4.4 Structural variations and CNVs

Microscopically detectable chromosomal aberrations have been estimated to account for ~5 % of ASD cases (Betancur 2011). The most common cytogenetic abnormality in ASDs, accounting for approximately 1-3 % of cases is maternal duplication of the imprinted 15q11-q13 region (Stankiewicz & Lupski 2010). This region overlaps with the deletion causing Angelman syndrome and Prader-Willi syndrome, both of which also have overlapping clinical symptoms with ASDs (Veltman et al. 2004, Bonati et al. 2007). Other common cytogenetic abnormalities in ASD are deletions of 2q37, 7q22q31, 18q21q23, 22q13.3, and Xp22; duplications of 7q11.23 and 17p11.2 and aneuploidy of chromosome Y (Stankiewicz & Lupski 2010).

The first systematic study of submicroscopic chromosomal rearrangements in idiopathic ASD revealed an excess of *de novo* CNVs in individuals with autism (Sebat et al. 2007). The *de novo* CNVs were particularly enriched in families with only one affected individual. The overlap between identified CNVs was small; most CNVs were detected only in a single individual. The AGP linkage study was the first autism study where genome wide SNP data was assessed for variation in copy number of submicroscopic regions (Szatmari et al. 2007). Thereafter evidence for the role of CNVs in the multilayered etiology of ASD has been obtained from several studies (Marshall et al. 2008, Glessner et al. 2009, Pinto et al. 2010, van der Zwaag et al. 2010, Levy et al. 2011, Salyakina et al. 2011, Sanders et al. 2011). Recurrent microdeletions and microduplications in ASD have been identified on chromosome 16p11.2 (Kumar et al. 2008, Weiss et al. 2008). A known recurrent microdeletion at 22q11.2 has been associated with increased rates of ASD (Vorstman et al. 2006).

The submicroscopic deletions and duplications may affect ASD risk genes and/or their regulatory elements that function in multiple cellular processes, like in neurodevelopmental pathways (Kakinuma & Sato 2008, Bourgeron 2009, Pinto et al. 2010, Sanders et al. 2011, Berkel et al 2012). Added to this, CNVs overlapping with the ones observed in autism have been found in other neuropsychiatric disorders including schizophrenia, intellectual disability, attention deficit hyperactivity disorder and epilepsy syndromes suggesting a partially shared genetic background (Guilmatre et al. 2009, Betancur 2011, Cooper et al. 2011, Elia et al. 2011, Glessner et. 2012). These CNVs have been discovered more often in patients than in the normal control population (Marshall et al. 2008, Glancy et al. 2009, Sanders et al. 2011).

There is evidence that *de novo* copy number variants (CNV) contribute to autism risk in 5-15 % of cases in families with one affected individual (Sebat et al. 2007, Marshall et al. 2008, Glessner et al. 2009, Pinto et al. 2010, Levy et al. 2011, Sanders et al. 2011). Inherited CNVs are found in up to 50 % of ASD subjects for whom one of the presumable normal parents also has the CNV. These familiar CNVs may include candidate genes relevant to ASD when they are rare in the normal population (Li et al. 2012). Most of the reported CNVs in autism have been large (over 400kb) (Sebat 2007, Marshall 2008, Glessner 2009). The recent development of single nucleotide polymorphism platforms has improved the resolution and enabled researchers to detect smaller CNVs (Haraksingh et al. 2011).

2.4.5 Exome sequencing

The recent development of sequencing technologies provides an effective tool to detect basically all variants in the human genome. Exome sequencing is a powerful tool to detect genetic variants in the coding region of the human genome that comprises 1.5 % of the total genomic sequence. Sequencing of the whole genome will identify variants in introns and intergenic regions containing regulatory elements. Exome sequencing is considered particularly useful for detecting rare mutations specifically in monogenic disorders. It can also be applied to find rare genetic variants in common complex disorders like ASD. In this chapter I summarize the most significant results from exome sequencing studies in ASD.

O´Roak and colleagues (2012a) performed exome sequencing in 209 ASD families. 39 % of the disruptive mutations detected in ASD were located in genes related to β-catenin-chromatin remodeling protein networks. Recurrent protein altering mutations were observed in two genes *CDH8* (chromodomain helicase DNA binding protein 8) and *NTNG1* (netrin G1). In addition to that, a mutation screen of candidate genes in larger patient material revealed protein altering *de novo* mutations in genes *GRIN2B* (glutamate [NMDA] receptor subunit epsilon-2 precursor), *LAMC3* (Homo sapiens laminin, gamma 3) and *SCN1A* (Homo sapiens sodium channel, voltage-gated, type I, alpha subunit).

Sanders and colleagues (2012) studied the exomes of 238 families, including 200 quartets. Non-synonymous *de novo* mutations occurred > 40 % more frequently in affected individuals than in unaffected siblings. Disruptive *de novo* point mutations in the *SCN2A* (sodium channel protein type 2 subunit alpha) gene showed significant association with ASD. When the authors combined the sequencing data with an earlier study (O´Roak et al. 2012a) two more genes *KATNAL2* (katanin p60 subunit A-like 2) and *CDH8* (chromodomain helicase DNA-binding protein 8) also met this criterion for an ASD risk gene.

In the exome sequencing study of Neale and colleagues (2012), 175 ASD trios were investigated. 46 % of the cases had a missense or nonsense *de novo* variant. However, this rate was only modestly higher than the mutation rate expected in unaffected individuals. In contrast, proteins encoded by the genes that harboured the *de novo* missense or nonsense mutations showed connectivity to ASD related genes supporting protein-protein interaction as an important mechanism in ASD. Again, the genes associated with ASD were *CDH8* and *KATNAL2.* A new *de novo* splice site mutation in the *SCN2A* gene was subsequently detected in one additional ASD case.

The studies mentioned here demonstrate significant association of three genes with ASD; *de novo* mutations at the *CDH8, KATNAL2,* and *SCN2A* genes contribute to risk for ASD. They also suggest a polygenic model for ASD. *De novo* point mutations were overwhelmingly often located on the paternal chromosome, and were associated with paternal age. Children of older fathers had a modestly increased risk to develop ASD (Neale et al. 2012, O´Roak et al. 2012a, Sanders et al. 2012). Maternal age had positive correlation with ASD risk in one of the studies (Neale et al. 2012). A recent Icelandic next generation sequencing study

confirmed the observation that the father´s age is associated with the rate of *de novo* mutations of the child (Kong et al. 2012). This is in line with the mechanism of human spontaneous mutation; in males the mutation rate is much higher than in females mainly because there are more germ-cell divisions in sperm than in egg, and the difference increases with age (Crow et al. 2000).

In addition to the previously mentioned studies, a recent exome sequencing study in 343 ASD families reported that many of the disrupted genes are associated with the fragile X protein, FMRP. This strengthens the links between autism and synaptic plasticity (Iossifov et al. 2012). Also, gene-disrupting mutations (nonsense, splice site and frame shifts) were detected twice as frequently in affected versus unaffected children.

Targeted sequencing of limited regions in the genome with a large sample set is a cost-effective approach to detect rare variants in ASD. O´Roak and colleagues (2012b) performed targeted sequencing for 44 candidate genes in 2,446 ASD probands. They reported that recurrent disruptive mutations in six genes; *CDH8, DYRK1A* (dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A)*, GRIN2B, TBR1* (T-box, brain, 1)*, PTEN* (phosphatase and tensin homolog) and *TBL1XR1* (transducin (beta)-like 1 X-linked receptor 1) may contribute to 1 % of sporadic ASD.

The results of exome sequencing studies together with CNV data indicate extensive locus heterogeneity but also provide targets for future diagnostics and therapeutics. There are many challenges and limitations with current methods of direct functional testing of rare variants and identifying alleles with modest effects. Ongoing next-generation sequencing studies by large international Autism consortiums might reveal disruptive mutations at regulatory regions. It likely will also give new information about small and disruptive structural variations in intergenic regions. It is challenging to distinguish pathogenic from benign *de novo* mutations when the mutation is located outside of the coding region. Systematic collecting and sharing of mutation and phenotype data internationally will help to understand the non-coding part of the genome (Veltman and Brunner 2012).

3 AIMS OF THE STUDY

The purpose of this thesis was to detect genetic variants predisposing to autism spectrum disorders in Finnish ASD families. Several different approaches were used in this thesis including linkage and candidate gene analysis, GWA study, CNV detection and promoter analyses.

The specific aims of the study were:

1. To fine map Xq11.1-q21.33 and sequence the coding areas of the *RPS6KA6, ZNF711, ACSL4, DLG3,* and *IL1RAPL2* genes in the region (I).
2. To analyze if SNPs previously detected in ASD (AGP 2007) in chromosomal regions 9p24 and 11p12-13 and SNPs earlier reported in association with OCD at glutamate transporter gene *SLC1A1* at 9p24 are associated with ASD in Finnish sample (II).
3. To analyze the role of ProsSAP1A (isoform of SHANK2 gene) as a predisposing gene in autism (III).
4. To perform a genome-wide scan, promoter and pathway analyses in novel set of Finnish ASD families (IV).
5. To identify CNVs in Finnish data set (unpublished).

4 MATERIALS AND METHODS

4.1 PATIENT MATERIAL

The original family material in Studies I and III was obtained mainly via Helsinki University Hospital, Jyväskylä Central Hospital and Kuopio University Hospital. The second set of families, in Studies II and IV were collected through the Department of Child Neurology, Helsinki University Hospital in 2007-2008. Diagnostic evaluations were made by a multi-disciplinary group of clinicians at the neurological department of each of the hospitals.

Initially, patients were referred from outpatient clinics to the child neurological departments for a one to two weeks observation period. Thorough clinical and medical examinations were performed including neurological examinations, assessment of developmental history as well as psychological and neuropsychological examinations. Final diagnoses were based on ICD-10 (World Health Organization 1993) and DSM-IV (American Psychiatric Association 1994) diagnostic nomenclatures. Childhood Autism Rating Scale (CARS) (Chopler et al. 1980), Asperger Syndrome Screening Questionnaire (ASSQ) (Ehlers and Gillberg 1993), and Asperger´s Syndrome Diagnostic Interview (ASDI) (Gillberg et al. 2001) were used as screening instruments. Families with associative medical conditions such as Fragile X syndrome, chromosomal aberrations, neurocutaneous syndromes and profound mental retardation were excluded.

Because the diagnostic tools for autism vary between different countries, the diagnostic methods used in Finland have been validated. 115 subjects who were originally recruited for the first Finnish molecular genetic study in autism (Auranen et al. 2002) were included in the validation study (Lampi et al. 2010). Diagnoses of subjects with infantile autism were re-assessed using the international Autism Diagnostic Interview – Revised (ADI-R) questionnaire (Lord et al. 1994), the gold standard of the field. Overall, ADI-R showed 96 % consistency with the original diagnoses of the Finnish autism sample (Lampi et al. 2010).

Families and individuals included in this study are summarized in Table 4. The different study samples are described in more detail below, and in the original articles (I-IV).

Table 4. Summary of family sets and individuals used in this study.

Study	Families	Individuals	Cases	Controls	Autism	AS
I	99[a]	356	x	100	117	15
II	x	x	175	216	139	36
III[b]	x	x	455	431	x	x
III[c]	x	x	851	1090	x	x
IV	81	257	81	750	78	12
IV[d]	205	737	x	x	241	0

[a]Families as in the study of Ylisaukko-oja et al. (2005)

[b]Finnish samples as in Study I

[c]Results were integrated with those reported by Berkel et al. (2010)

[d]the *AVPR1A* gene was studied in a larger set of families in Study IV

AS = Asperger syndrome

In Study I, fine mapping of the X-chromosome was performed in 99 Finnish ASD families, each containing one to three affected members with ASD. Families consisted from those in the publication of Ylisaukko-oja et al. (2005). The total number of individuals included in the study was 356, of which 117 were affected with infantile autism. Individuals with a diagnosis of autism were classified into liability class 1, the narrow diagnostic category, and LC2 a broader phenotype, which includes cases with both infantile autism and Asperger syndrome. All families were Finnish, except one father who is of Turkish origin. Forty-two affected males were selected for candidate gene analysis. In the families contributing to linkage at Xq21-24, we chose affected individuals sharing a haplotype not present in unaffected individuals of the family. All individuals included in the candidate gene analysis were males. The control sample set was comprised of 100 anonymous Finnish blood donors.

In Study II, the patient material consisted of a novel set of Finnish families comprising 175 patients with ASD. 139 individuals had infantile autism and 36 individuals had AS. 138 of the ASD patients were men and 37 were women. The control material consisted of 216 anonymous Finnish blood donors.

In Study III, the *SHANK2* gene was sequenced in 455 patients with ASD and in 431 controls. The study material was recruited by the Paris (Paris Autism Research International Sibpair) study at specialized clinical centres disposed in France, Sweden, Germany, Finland and UK. The Finnish samples are the same as in Study I. 100 anonymous Finnish blood donors were included in the control material.

In Study IV, a total of 83 Finnish families participated in the study, including 257 individuals in total. The sample set was included in Study II as well. The families contained 1 to 3 affected members, 78 affected with infantile autism and 12 with Asperger syndrome. Within affected individuals the male-female ratio was 5:1. The families were analyzed using the IlluminaHumanOMNI 12 v.1.0 panel consisting of 733 000 single nucleotide polymorphisms (SNPs). The candidate gene *AVPR1A* in Study IV was genotyped in a larger set of families with autism (N=205) consisting altogether of 737 individuals.

No phenotype data was available for the anonymous blood donors used as controls in Studies I-III. Thus they are used as population controls. In Study IV, the population based control material consisted of 750 unrelated subjects from a cohort of the Health 2000 Study, where genetic risk factors of metabolic syndrome (MetS) are studied (Kristiansson et al 2012). In the control subjects DSM-IV mental disorders; alcohol use, depression and anxiety were excluded using the Composite International Diagnostic Interview (CIDI) (Pirkola et al. 2005). Psychotic disorders were excluded using the research version of the Structured Clinical Interview for DSM-IV (SCID-1) (Perälä et al. 2007).

The studies have been approved by the Ethical Committee of the Hospital of Children and Adolescents, Helsinki University Central Hospital. The Ethical Committee of the Finnish Red Cross Blood Transfusion Services in Helsinki approved the use of the DNA of anonymous blood donors. The Board of Health 2000, The National Institute for Health and Welfare has given permission for control samples in Study IV. Informed written consent was obtained from the subjects and/or their parents.

4.2 METHODS

All the methods used in this study have been described in detail in the original articles (Table 5). An overview of the methods is also given below.

Table 5. Methods used in this study.

Method	Reference	Publication(s)
Experimental procedures		
DNA extraction	Gentra systems, Minneapolis, MN, US	I, II, III, IV
Polymerase Chain Reaction (PCR)	Kleppe et al. (1971)	I, II, III
Electrophoresis, ABI 3730	Applied Biosystems, Foster City, CA, US	I,II,III
7500 fast real-time PCR System	Applied Biosystems, Foster City, CA, US	II
Genome wide SNP genotyping; Illumina Infinium Human OmniExpress-12v1 bead chip	Illumina, San Diego, CA, US	IV
Illumina Infinium HD Human610-Quad BeadChip	Illumina, San Diego, CA, US	IV
qPCR: CFX96^{TM} Real-Time PCR System	Bio-Rad Laboratories, Hercules, CA, US	IV
Analysis programs		
Genemapper 4.0	Applied Biosystems, Foster City, CA, US	IV
Ingenuity Pathway Analyses	IPA; Ingenuity systems	IV
Promoter analysis; Transfac	Matys et al. (2003)	IV
Sequencher 4.7	Gene Codes, Ann Arbor, MI, US	I, II, III
Statistical methods and software		
FBAT	Horvath et al. (2001)	II, IV
PEDCHECK 1.1	O'Connell and Weeks (1998)	I, II, IV
MLINK	Lathrop et al. (1984), Lathrop et al. (1986)	I
HOMOG	Ott (1986)	I
MERLIN 1.0.1 MINX	Abecasis et al. (2002)	I
PLINK	Purcell et al. (2007)	II, IV
Blosum62	Styczynski et al. (2008)	I
PolyPhen	Ramensky et al. (2002)	I
SIFT	Ng and Henikoff (2003)	I
SPSS 15.0	SPSS Inc., Chicago, IL, US	II

4.2.1 DNA isolation and genotyping

Genomic DNA from the samples used in Studies I and III was extracted from EDTA-treated peripheral blood samples with the Puregene DNA purification system (Gentra Systems, Minneapolis, MN, US) according to the manufacturer´s protocol or using a phenol-chloroform protocol modified from Vandenplas and colleagues (1984). DNA from blood samples of new sample material collected in years 2007-2008 were extracted using the HUSLAB protocol MP025 non-enzymatic DNA extraction method with a slight modification; Igepal was used to disrupt cell membranes instead of nonident.

In Studies I, II and III primer sequences were designed using the Primer3 program. The DNA of the study subjects was amplified by polymerase chain reaction (PCR). The polymerase chain reaction conditions were as follows: 10 or 5 min at 94°C followed by 35 cycles of the denaturation step: 30 s at 94°C, annealing step: 30 s at temperature specific for each primer (50-65°C), the elongation step: 30 s at 72°C and final extension for 10 min at 72°C. Sequencing was performed using cycle sequencing with the Big Dye Terminator kit (ABI, Foster City, CA, USA) and reactions were run on an ABI 3730 capillary sequencer. SNP genotyping in Study II was performed with a 7500 fast real-time PCR system (Applied Biosystems, Foster City, CA, US) according to manufacturer´s instructions.

Microsatellite markers in Studies I and IV were selected from the Marshfield Medical Research Foundation map and Primer sequences were obtained from UCSC Genome Browser. Forward primers were labeled at the 5'-end with FAM or PET fluorescent dye. PCR-products were pooled and electrophoresed on an ABI3730 automatic DNA sequencer. Genotypes were assigned using Genemapper 4.0 software (ABI, Foster City, CA, USA). All microsatellite and SNP markers used in this study (Studies I, II, IV) were checked and corrected for Mendelian errors prior to analysis using the Pedcheck program (O'Connell and Weeks 1998).

Genome wide SNP genotyping of ASD families was performed using the Illumina Infinium Human OmniExpress-12v1 bead chip (www.illumina.com) consisting of 733 202 markers. The genotyping of Health 2000 control material was performed with the Illumina Infinium HD Human610-Quad BeadChip.

We used quantitative PCR to validate the CNVs detected in Study IV. SYBR green-based genomic real-time qPCR analysis was performed on a CFX96™ Real-Time PCR System (Bio-Rad Laboratories, Hercules, CA, US) by using SsoFast™ EvaGreen® Supermix (Bio-Rad Laboratories, Hercules, CA, US) according to the cycling protocol. Primers were developed by the Primer3 program (http://frodo.wi.mit.edu/) (Rozen and Skaletsky 2000). qPCR quantifications were performed in duplicate on 1 ng/µl DNA and included a water control. Copy numbers were measured relative to the reference gene, zinc finger protein 80 *(ZNF80)*.

4.2.2 Analysis Programs

In Study IV the networks and functional analyses were generated through the use of IPA (Ingenuity® Systems, www.ingenuity.com). IPA is a web-based software application that transforms a list of genes into a set of networks based on the Ingenuity Pathways Knowledge Base, a large dataset which contains biological and chemical relationships extracted from scientific literature. IPA enables user to analyse, integrate and understand the experimental data derived from SNP microarrays, gene-expression, microRNA, proteomics and other experiments which produce a list of genes or chemicals. The goal of IPA is to discover and visualize hidden causal connections of the dataset of interest (www.ingenuity.com).

In Study IV we used Transfac® to predict transcription factor (TF) binding sites. Transfac® is a knowledge-base containing published data on eukaryotic transcription factors, their target genes and regulatory binding sites (Matys et al. 2003). The goal of the transcription factor binding analyses is to better understand the mechanism of regulation and to characterize the mutations disrupting the regulatory mechanisms. In this thesis neurospecific TF binding sites were predicted with the algorithm Match in the Transfac® database (www.biobase.de).

4.2.3 Statistical genetic analyses

In Study I linkage analysis was performed using the MLINK (Lathrop et al. 1984, Lathrop et al. 1986), HOMOG (Ott 1986) and MERLIN 1.0.1 MINX (Abecasis et al. 2002) programs. In Studies II and IV allelic association analyses were done by FBAT (Horvath et al. 2001) or PLINK (Purcell et al. 2007). In Study IV all quality checks of the genome wide SNP data were performed using PLINK software. Genotypes were checked for correct Mendelian transmission when family information was available. To exclude sample swaps and contamination, gender checks were performed, and the samples monitored for identical-by-descent (IBD) sharing and mean heterozygosity. For the analysis of the genome-wide scan, markers were pruned for too low minor allele frequency (MAF <0.05), missing genotype frequency (if over 5 % of marker genotypes were missing) and HWE (p-value < 0.001 in controls). Subjects were pruned out if they had more than 5% of their genotypes missing and if they were of non-Finnish origin.

In Study IV we used PLINK for both the TDT analyses within the autism families and for the case-control association analysis. The rationale for using these two approaches for association is that the methods are concerned with slightly different aspects of heredity; the TDT measures linkage in the presence of association, observing transmissions of alleles from heterozygous parents to affected offspring, whereas case-control association simply measures whether a particular allele is more common in the cases than in the controls. While TDT is not sensitive for population stratification contrary to the case-control setting, the case-control study is better at finding common risk alleles. Thus, the approaches complement each other.

CNV calls and analysis is based on build GRCh37/hg19 of the human genome reference sequence (UCSC Genome Browser on Human Feb. 2009). Two different algorithms PennCNV (Wang et al. 2007) and QuantiSNP (Colella et al. 2007) were used for the identification of the CNVs. Only the calls consistent between both the algorithms were used for further analysis, to minimize the rate of false discoveries (Dellinger et al. 2010, Pinto et al. 2010). As the impact of rare CNVs on ASD predisposition has been established (Pinto et al. 2010), we chose to analyze only rare, possibly pathogenic CNVs. Rare CNVs were extracted by excluding all the common CNVs based on the CNV frequencies in the Finnish control population (> 1% frequency). Controls were the same as in study IV. The AutismKB (http://autismkb.cbi.pku.edu.cn/) database was used to determine the novelty of ASD candidate CNVs.

5 RESULTS AND DISCUSSION

5.1 Fine mapping and sequencing of candidate genes at Xq11.1-q21.33 (I and unpublished data)

In a previous genome-wide scan of 369 microsatellite markers in 38 Finnish families with ASD (Auranen et al. 2002), one of the positive loci was observed at the X chromosome, the best marker DXS7132 locating at Xq11.1. Further analysis of the region in a larger sample set resulted in the highest LOD score (Zmax=2.39) at DXS7117 at Xq13.1 (Ylisaukko-oja et al. 2005). As many of the patients with ASD have intellectual disability and the linked region contained several causative genes for ID we performed fine mapping of Xq11.1-q21.33 with a total of 26 microsatellite markers.

A maximum two-point LOD score of 3.00 was obtained at DXS1225 on Xq21.1 using a broad phenotype (LC2) and dominant model (Figure 9). The highest multipoint LOD score was obtained with the same marker DXS1225 (HLOD=2.94, NPL_{all}=3.43 corresponding to P=0.0004). In concordance with the two-point analysis, the dominant inheritance model yielded the best results. Thus, the region of interest shifted distally from DXS7132 at Xq11.1 to DXS1225 at Xq21.1 (hg 18). There were numerous genes in the linked region that are related to XLID. Of them *RPS6KA6* (ribosomal protein S6 kinase alpha-6) and *ZNF711* (zinc finger protein 711) (Yntema et al. 1999, Tarpey et al. 2009) reside at the peak region, *ACSL4* (acyl-CoA synthetase long-chain family member 4) and *DLG3* (discs, large homolog 3), which have previously been known to cause XLID (Tarpey et al. 2004, Renieri et al. 2005, Zanni et al. 2010) and *IL1RAPL2* (interleukin 1 receptor accessory protein-like 2), a homologous gene for *IL1RAPL1* that is mutated in autism (Piton et al. 2008, Piton et al. 2011) and intellectual disability with autism (Bhat et al. 2008).

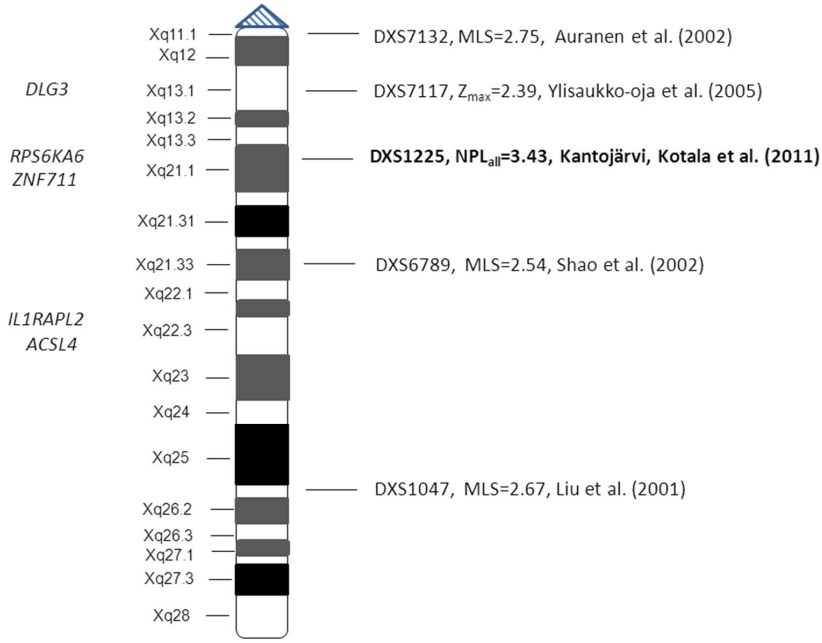


Figure 9. The locations of the linkage peaks on chromosome Xq in this and earlier studies. The locations of candidate genes sequenced in this study are also presented.

We sequenced the coding regions and splice sites of the aforementioned five candidate genes in 42 male patients with infantile autism from families contributing to the linkage. We could not establish any functional mutations in these patients who carried a haplotype not shared by unaffected individuals in the family. One non-synonymous variant was detected at the *RPS6KA6* gene. Prediction analysis showed that it was a harmless variant. This implies that the coding sequence variants of these five genes may not explain the linkage observed in our data set. The current study design cannot exclude the existence of rare mutations in these genes in the Finnish sample. Recently, Piton et al. (2011) performed a systematic re-sequencing of X-chromosomal synaptic genes in ASD (N=142) and schizophrenia (N=143). The gene set included the *DLG3*, *IL1RAPL2* and *RPS6KA6* genes. None of the variants they detected in these three genes were likely to be involved in ASD (Piton et al. 2011).

Causative mutation in neuroligin3 (*NLGN3*) at Xq13 has been reported in autism (Jamain et al. 2003). In previous Finnish scan of Xq (Ylisaukko-oja et al. 2005) *NLGN3* was sequenced in the same data set used in this study. Disruptive mutations were not detected. There are several other genes in the linkage area which are highly expressed in the brain. Sequencing of the whole Xq11.1-q21.33 region is necessary to reveal autism related X-chromosomal mutations in our dataset. Sequencing of the coding region of the X-chromosome has revealed a number of novel causative mutations for XLID. Altogether 105 genes have been found underlying XLID (http://www.ggc.org/research/molecular-studies/xlid.html) (November 2012). Based on the current knowledge of the partially shared genetic background of autism and ID, these genes serve as candidate genes for autism.

Chromosomal aberrations at Xq21 were analyzed from genome wide SNP data used in Study IV. Two of the patients had a CNV at the Xq21.31 region containing *PCDH11X/Y* (protocadherin 11 X/Y-linked gene) (Figure 10). *PCDH11X/Y* gene belongs to the protocadherin family. *PCDH11X/Y* genes are present on both the X and Y chromosomes, and are expressed in the cortex of the human brain. In our study a male patient (3076) with infantile autism had a 580 kb paternally inherited duplication at the X chromosome and a 446 kb duplication at the Y chromosome which both contain parts of the *PCDH11X/Y* gene. The male patient also had a 115 kb intronic X-chromosomal *de novo* duplication in the same gene region. Female patient (1872) with infantile autism had a 137 kb intronic *de novo* deletion in the *PCDH11X* gene. In earlier studies of autism, two CNVs at *PCDH11X* have been detected (Pinto et al. 2010). Disruption of the *PCDH11X* gene has also been described in language delay (Speevak and Farrell 2011). Association studies of the *PCDH11X* gene in late-onset Alzheimer disease have been conflicting (Carrasquillo et al. 2009, Beecham et al. 2010). The Y chromosomal CNVs we detected were novel according to the AutismKB.

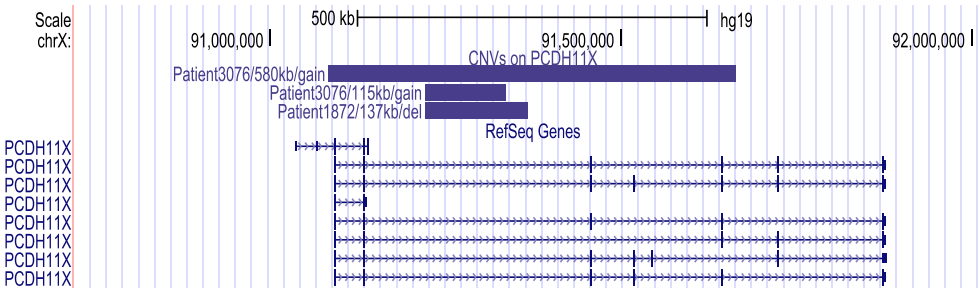


Figure 10. CNVs detected in patients with autism on gene *PCDH11X*.

The *PCDH11X/Y* genes, like the other members of the protocadherin family, play a major role in synaptogenesis, the formation of functional synapses (Kim et al. 2011). *PCDH11Y* is one of the rare genes on the Y chromosome and it is specific to humans (Blanco et al. 2000). *PCDH11Y* may play a role in human specific cognitive functions, such as language; therefore the defects in this gene could predispose to neuropsychiatric disorders (Durand et al. 2006).

Previous studies of ID have reported that not all of the truncating mutations identified are causative. Tarpey et al. (2009) observed that 1 % of the protein-truncating mutations on the X chromosome were found to be unrelated to disease. Furthermore, Najmabadi et al. (2011) reported that 12 of the observed inactivating mutations did not co-segregate with ID. More experimental studies are warranted to see if the detected mutations have an effect for the phenotype studied. Considering autism, presumably the interaction of the disrupted genes predispose to autism, while one single mutation is not sufficient.

5.2 Analysis of 9p24 and 11p12-13 regions in ASD (II and unpublished data)

In Study II we analyzed whether the most significant SNPs reported in the largest linkage study of ASD (Szatmari et al. 2007) at chromosomal areas 9p24 (rs1340513 and rs722628) and 11p12-13 (rs1358054 and rs1039205) are associated with ASD in the Finnish sample. Furthermore, considering the common genetic background of ASD and obsessive compulsive disorders (OCD), we studied SNPs in the *SLC1A1* gene at 9p24, which has been previously associated with OCD (Dickel et al. 2006, Stewart et al. 2007).

The SNPs associated with ASD at the 9p24 region are presented in Figures 11 and 12. SNPs rs1358054 and rs1039205 at the chromosomal region 11p12-13 did not show association with ASD in the Finnish sample. The strongest association was detected with SNP rs1340513 in the *JMJD2C* alias *KDM4C* gene at 9p24.1 (P=0.007, corrected P=0.011). A recent large GWA study in OCD detected association in the *JMJD2C* gene, although none of the studied SNPs reached genome-wide significance at 9p24 (Steward et al. 2012). In this thesis, we observed that one female subject (1492) with Asperger syndrome had an inherited, homozygous 4.7 kb deletion which located 15 kb from *JMJD2C* to the 5` direction that might affect the regulation of this gene. In the Finnish controls, CNVs in this region were not detected. Another CNV has been detected in an earlier study in *JMJD2C* in autism (Pinto et al. 2010) (Figure 12).

The product of the *JMJD2C* gene is a histone demethylating enzyme, which converts trimethylated histone residues to the dimethylated form and thus affects chromatin structure and gene transcription (Cloos et al. 2006). *JMJD2C* is highly expressed in the brain (Nagase et al. 1998) and aberration of its expression level has been detected in autistic patients compared to controls (Hu et al. 2009). Genes affecting histone demethylation and chromatin structure have been reported in ASD, for example *MECP2* underlying Rett syndrome.

We observed a possible gene x gene epistatic interaction between SNPs rs2228622 at *SLC1A1* and rs1340513 at *JMJD2C* with SPSS logistic regression (P=0.002). We also noticed that SNP rs2228622 in the *SLC1A1* gene was also associated with heavy routines and rituals (P=0.0026). The study was however limited by the small sample size of individuals with known endophenotypes, warranting a replication of the result in bigger material. In another recent study, SNP rs301430 located in the *SLC1A1* gene was associated with anxiety but not in repetitive behavior in children with ASD (Gadow et al. 2010). This is in line with our study where this SNP did not show association with repetitive behavior. In addition to these, some evidence for association of rs301979 in *SLC1A1* in male patients with autism has been reported (Brune et al. 2008). Several SNPs have shown association with OCD (Arnold et al. 2006, Dickel et al. 2006, Stewart et al. 2007, Shugart et al. 2009, Wendland et al. 2009, Samuels et al. 2011). The strongest association with OCD in the *SLC1A1* region has been

found at rs301443 (P=0.000067) residing 7.5 kb outside at the 3´ end of the *SLC1A1* gene (Shugart et al. 2009). Variations reported to be associated with ASD at *SLC1A1* in this thesis and by others or with OCD are presented in Figure 11. In this thesis one male subject (1946) with autism had paternally inherited a 168 kb duplication on the glutamate transporter *SLC1A1* gene. He had childhood autism and ritualistic behavior.

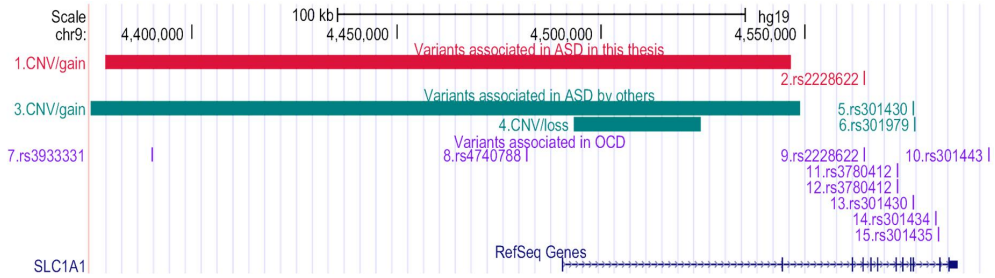


Figure 11. Locations of variants associated with ASD or OCD are presented: 1. In this thesis patient 1946 had a paternally inherited 168 kb duplication on the *SLC1A1* gene, 2. SNP rs2228622 was associated in heavy routines and rituals in ASD (P=0.0026), 3. 420 kb inherited duplication in autism (Bremer et al. 2011), 4. 31 kb inherited deletion in autism (Pinto et al. 2010), 5. SNP rs301430 was associated in anxiety in ASD (P=0.01) (Gadow et al. 2010), 6. rs301979 was associated in ASD (P=0.01) (Brune et al. 2008), SNPs associated in OCD are 7. rs3933331 in hoarding (P=0.005) (Wendland et al. 2009), 8. rs4740788 (P=0.003) (Samuels et al. 2011), 9. rs2228622, males only (P=0.045) (Steward et al. 2007), 10. rs301443 (P=0.000067) (Shugart et al. 2010), 11. rs3780412, males only (P=0.002) (Dickel et al. 2006), 12. rs3780412, males only (P=0.045) (Steward et al. 2007), 13. rs301430 (P=0.03) (Dickel et al. 2006), 14. rs301434 (P=0.006) and rs301435 (P=0.03) (Arnold et al. 2006).

Interestingly, we also detected a significant association of two SNPs to ASD at chromosomal region 9p23-24 in Study IV. There, the most significant results were obtained at the 9p23 region at the *PTPRD* gene which is located only 1.2 Mb from *JMJD2C* (Figure 12). The strongest associations were found with SNP rs16926616 ($5.7x10^{-5}$, assoc) locating 292 kb from the *PTPRD* gene and SNP rs324484, ($P=1.1x10^{-3}$, TDT) at *PTPRD*. In an earlier, large GWA study in autism a significant association was detected between genes *JMJD2C* and *PTPRD* (Weiss et al. 2009). The *PTPRD* gene has been associated with restless legs syndrome (Schormair et al. 2008), ADHD (Elia et al. 2010) and OCD (Steward et al. 2012). Furthermore, the postsynaptic *IL1RAPL1* gene that is associated with intellectual disability and autism

mediates synapse formation by trans-synaptic interaction with pre-synaptic PTPRD (Yoshida et al. 2011). Figure 13 illustrates the overview of the locations of the best SNPs at 9p24 which were associated with ASD in this thesis.

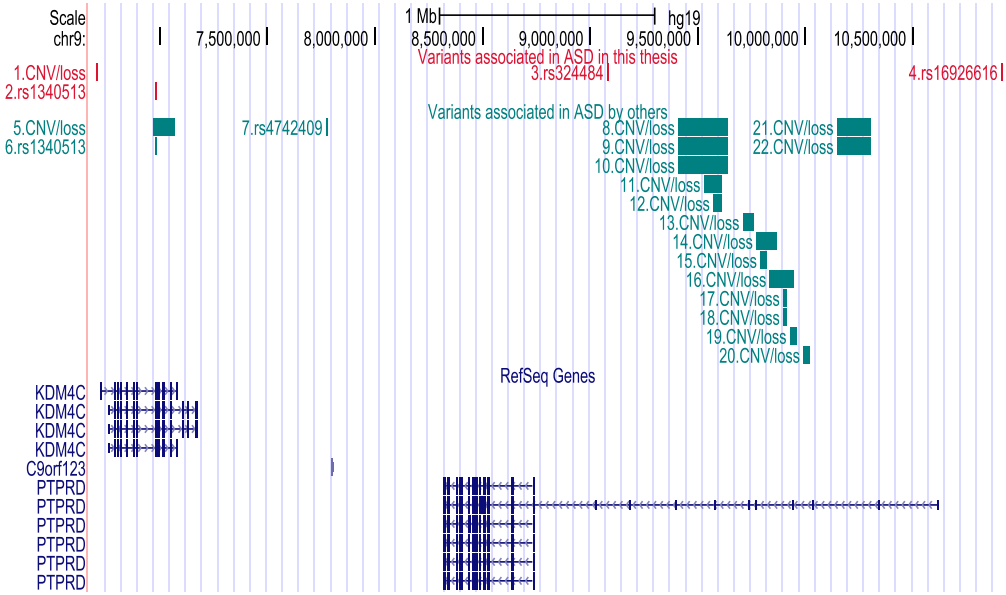


Figure 12. Variants associated with ASD in this thesis or by others in the *KDM4C/PTPRD* region. 1. One patient with Asperger (1492) had inherited a homozygous 4.7 kb deletion from both parents 15 kb from *JMJD2C* to the 5`direction, 2. rs1340513 at *KDM4C* was associated in ASD (P=0.007) in Study II, 3. rs324484 at *PTPRD* and 4. rs16926616 near *PTPRD* were associated in ASD in Study IV ($P=1.1 \times 10^{-3}$ and $P=5.7 \times 10^{-5}$ respectively). Previous studies reported 5. one deletion in autism at *KDM4C* (Pinto et al. 2010), 6. rs1340513 associated with ASD (P=0.0007) (Szatmari et al. 2007), 7. rs4742409 associated with ASD ($P=1.6 \times 10^{-5}$) (Weiss et al. 2009), 8.-22. Several deletions at the *PTPRD* gene (Pinto et al. 2010, Gai et al. 2012, Levy et al. 2011, Sanders et al. 2011).

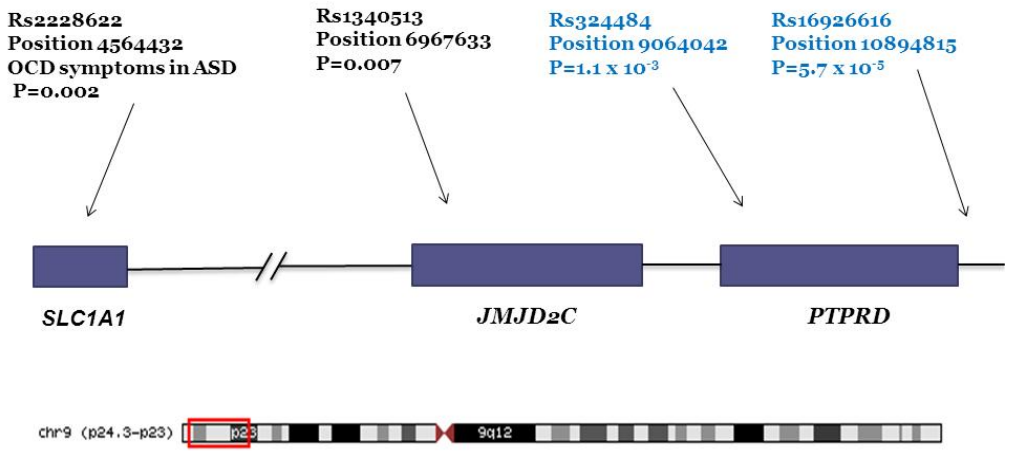


Figure 13. SNPs associated in ASD in Study I (black) and Study IV (blue).

In addition to the *SLC1A1* gene, we identified several other glutamate genes containing CNVs in ASD patients. Four of the patients (1490, 1869, 1890 and 2923) had a 7.8 kb deletion at 11q14.3 in the *GRM5* gene (Homo sapiens glutamate receptor, metabotropic 5); three of them had autism and one had Asperger syndrome. Based on AutismKB this CNV has not been detected in ASD before. L-glutamate is the major excitatory neurotransmitter in the central nervous system and activates both ionotropic and metabotropic glutamate receptors. Glutamatergic neurotransmission is involved in most aspects of normal brain function and can be perturbed in many neuropathologic conditions. The metabotropic glutamate receptors belong to a family of G protein-coupled receptors that affect signal transduction (Ozawa et al. 1998). Group I includes GRM1 and GRM5 and these receptors have been shown to activate phospholipase C. Group II contains GRM2 and GRM3 while Group III includes GRM4, GRM6, GRM7 and GRM8. Group II and III receptors are linked to the inhibition of the cyclic AMP cascade but differ in their agonist selectivity. Deletions in *GRM5* were discovered earlier in patients with attention-deficit hyperactivity disorder (ADHD) (Elia et al. 2010, Elia et al. 2011).

Other glutamate related genes located in CNV regions observed in this study were *GRID1* (glutamate receptor, ionotropic, delta 1), *GRIK1* (glutamate receptor, ionotropic, kainate 1) and *GRIN2A* (glutamate receptor, ionotropic, N-methyl D-aspartate 2A). One of the patients

had a 4 kb deletion at 10q23.2 located at the *GRID1* gene. This gene encodes a subunit of the glutamate receptor channels that mediate most of the fast excitatory synaptic transmission in the central nervous system and play a key role in synaptic plasticity. A 4.5 kb deletion at 21q21.3 located at *GRIK1* was detected in one of the patients (2994). In another case-control and family-based study, nominal significant association was observed between single nucleotide polymorphisms of *GRIK1* and schizophrenia (Hirata et al. 2012). In this thesis a 111 kb deletion at 16p13.2 was detected in one of the patients (2964), located 139 kb from the gene *GRIN2A* to the 5´direction. The GRIN2A receptor is found at excitatory synapses throughout the brain and it has an important role in neuronal activity and development. Mutations in this gene are involved in intellectual disability and learning difficulties (Endele et al. 2010). *GRIN2A* is also associated with autism (Barnby et al. 2005, De Krom et al. 2009).

5.3 Genetic and functional analyses of *SHANK2* mutations provide evidence for a multiple hit model of ASD (III)

In order to better understand the role of the NRXN-NLGN-SHANK pathways in ASD, we screened for *SHANK2* CNVs and coding mutations in ASD cases and controls collected from France, Sweden, Germany, Finland and UK. Furthermore, we analyzed co-occurrent CNVs in patients who had *de novo SHANK2* deletions in this and earlier studies (Berkel et al. 2010, Pinto et al. 2010).

The human *SHANK2* gene (NM_012309.3) contains 25 exons and spans 621.8 kb. Transcription of S*HANK2* produces 4 isoforms: *SHANK2E* (*AB208025*), *ProSAP1A* (*AB208026*), *ProSAP1* (*AB208027*) and *AF141901* from three distinct promoters. In this study, these isoforms were validated by RT-PCR and sequencing. We noticed that *SHANK2* isoforms are differentially expressed in human tissues. Inter-individual differences in the relative amount of mRNA between isoforms of the gene have been previously reported in other synaptic genes such as *NLGN1-4Y*, *PCDH11X/Y* and S*HANK3* (Jamain et al. 2003, Durand et al. 2006, Durand et al. 2007). Exons 19, 20 and 23 were found to be expressed only in the brain in all individuals tested. This kind of brain specific splicing has been observed earlier in exon 18 of *SHANK3* (Durand et al. 2007), which is similar to exons 19 and 20 in *SHANK2*.

For the detection of CNVs affecting *SHANK2* we genotyped an independent sample of 260 patients with the Illumina 1 M Duo SNP array. A 41.2 kb *de novo* deletion was found in one patient with autism and moderate ID. The deletion covered twelve exons (5-16) and altered all SHANK2 isoforms. In an earlier study, deletions of *SHANK2* were found in one patient with ASD and in another with intellectual disability (Berkel et al. 2010). Furthermore, one additional *SHANK2 de novo* deletion was detected by the Autism Genome Project (Pinto et al. 2010). *SHANK2* deletions were absent in more than 5000 controls and it is not listed in the Database of Genomic Variants (http://projects.tcag.ca/variation/). Figure 14 illustrates the locations of the CNVs detected in the SHANK2 protein.

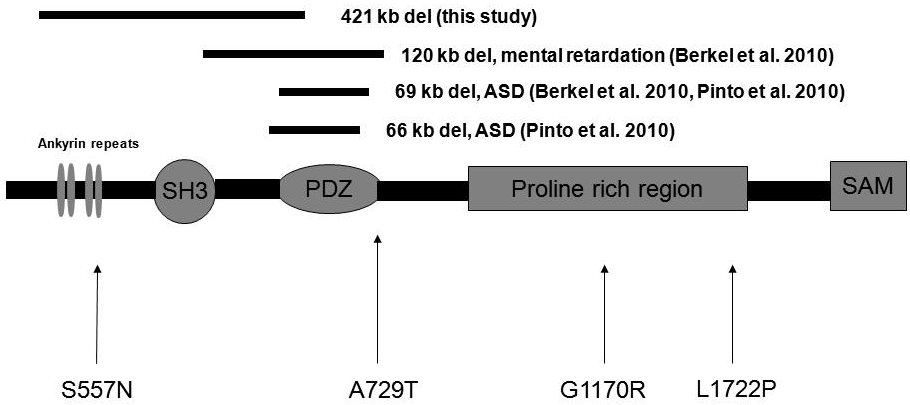


Figure 14. The locations of the CNVs disrupting the SHANK2 protein are presented. The longest SHANK2 isoform (SHANK2E, AB208025) contains ankyrin (ANK) repeats, SRC homology 3 (SH3) domain, postsynaptic density 95/Discs large/zona occludens-1 (PDZ) homology domain, Proline rich region and sterile alpha motif (SAM) domain. Proline rich region contains HOMER binding motif, dynamin-2 binding motif and contactin binding motif. The locations of conserved variants which reduced the number of synapses significantly are also shown. See text below.

To screen additional mutations in *SHANK2* we sequenced all exons of the longest *SHANK2E* isoform in 230 patients with ASD and 230 controls. We also sequenced the *ProSAP1A* isoform that corresponds to the major *SHANK2* isoform in the brain in an additional 225 patients and 201 controls.

Three patients with ASD had a mutation within the *SHANK2E* isoform: two independent patients had R174C listed in dbSNP and the third one R185Q. No variants were identified in the control samples. Within the *ProSAP1A* isoform, we identified 24 non-synonymous variations. When our results were integrated with the results of Berkel et al. (2010), a total of 40 variants of ProSAP1A were identified. Three of them have previously been reported in dbSNP. There was no enrichment of rare variants of *SHANK2* (MAF <1%) in patients with ASD compared with controls. Only two variants, Y967C and R569H, with MAF >1% were observed. Table 6 presents the variants detected in the ASD patients only. These were not detected in 1090 controls. 8 variants were detected in both ASD patients and controls, and 17 variants in controls only.

Table 6. *ProSAP1A/SHANK2* variations identified in ASD patients only.

Exon	Nucleotide	Amino Acid	Conservation in SHANK proteins	Study
E11	G70344397A	R405W	Yes (SHANK1 & SHANK3)	Berkel et al. 2010
E11	G70344284A	R443C	Yes (SHANK3)	This study
E13	G70322214A	P587S	No	Berkel et al. 2010
E14	C70222501A	R598L	Yes (SHANK3)	This study
E17	C70026597A	V717F	Yes (SHANK1 & SHANK3)	This study
E17	C70026561T	A729T	Yes (SHANK3)	This study
E22	G70014059A	R841X	Yes (SHANK3)	Berkel et al. 2010
E24	C70010562T	E1162K	No	This study
E24	C70010538T	G1170R	Yes (SHANK1 & SHANK3)	This study
E24	C70009920T	V1376I	No	This study
E24	C70009920T	V1376I	No	This study
E24	dup(AACGGT) 70009882–70009887	dup(LP) 1387–1388	Yes (SHANK1 & SHANK3)	Berkel et al. 2010
E24	G70009529A	T1506M	Yes (SHANK1 & SHANK3)	Berkel et al. 2010
E24	C70009443T	D1535N	Yes (SHANK1 & SHANK3)	This study
E25	A69997007G	L1722P	Yes (SHANK1 % SHANK3)	This study
E25	C69996987T	A1729T	Yes (SHANK3)	Berkel et al. 2010

Nucleotide positions are according to NM 012309.3 from NCBI36/hg18 on the positive DNA strand; The patients with ASD and the controls used for this analysis came from this study of Berkel et al. 2010 comprising altogether 851 ASD cases and 1090 controls.

We observed a significant enrichment of variants affecting conserved amino acids in patients compared to controls. 80 % of variants detected only in the patient sample compared to only 35.3 % of these in the control sample affected conserved amino acids (Fisher´s exact test 1-sided, P=0.013, OR=6.83). Because several independent patients carried the

conserved variants we also tested the effect of carriers to enrichment. We noticed that 29 of 851 patients (3.4 %) and 16 of 1090 controls (1.5 %) carried variants affecting conserved amino acids (Fisher exact test 1-sided, P=0.004, OR=2.37).

All but one of the variants were predicted to be damaging with Polyphen2. The benign variant was found only in controls. To test the functional impact of *SHANK2* variations we performed expression studies in neuronal cell cultures. All mutations identified in our first screen of 230 patients with ASD and 230 controls (N=16) were tested. Mutation sites were first introduced into rat *ProSAP1A* cDNA using the site directed mutagenesis kit (Stratagene) and confirmed by sequencing. After that cDNAs were tested for expression by Western blot analysis. As a result wild type showed over-expression compared to mutant *ProSAP1A* cDNA.

The effect of the variants was then investigated by transfecting the cDNAs into primary hippocampal neurons. Cell culture experiments of rat hippocampal primary neurons were performed as described previously (Boeckers et al. 2005). For the quantification of excitatory synapse number, cells were counterstained with anti-Bassoon antibodies. From randomly chosen transfected neurons, Bassoon positive spots from primary dendrites were counted and the length of dendrites was measured. The total number of spines was expressed as density per 10 mm length of dendrite.

Variants affecting conserved amino acids were associated with a higher reduction of synapse density *in vitro*. Variant R818H was virtually absent outside Europe and had the highest allelic frequency (3 %) in Finland, but overall it´s frequency was not higher in patients with ASD (32/3250, 1 %) compared with controls (27/2030, 1.33 %). Almost 5 % of the Finnish population was heterozygous for this conservative variant which reduced the number of synapses *in vitro*. After Bonferroni correction, four of the variants could be shown to significantly affect synaptic density. Of these variants, A729T, G1170R and L1722P were observed only in patients and S557N was identified more frequently in ASD than in controls. The locations of these variants in the SHANK2 protein are presented in Figure 14. Synaptic density was also tested as a continuous trait, which showed that variants identified in patients significantly decreased synaptic density *in vitro*, compared to the variants present in controls or both. Variants affecting conserved amino acids were associated in higher reduction of synapse density (P=0.004).

We also searched if three patients with *SHANK2 de novo* deletion had additional CNVs elsewhere in the genome which could modulate the impact of *SHANK2* deletions in the development of ASD. ASD patients with *SHANK2* deletion analyzed here were AU038_3 (this study), 5237_3 (Berkel et al. 2010, Pinto et al. 2010) and 6391_3 (Pinto et al. 2010). Co-occurrence of *SHANK2 de novo* deletions and inherited CNVs are presented in Table 7.

Table 7. Additional CNVs containing ASD candidate genes detected in patients with *SHANK2* deletion.

CHR	**GENE**	**AU038_3**	**5237_3**	**6319_3**
1q32.1	*CAMSAP1L*	paternal dup	-	-
3p26.3	*CNTN4*	-	-	maternal del
3q26.3	*NLGN1*	-	maternal dup	-
5q31.3	*PCDHA1-10*	-	-	paternal del
6p25.3	*DUSP22*	-	*de novo* del	-
11q13	*SHANK2*	*de novo* del	*de novo* del	*de novo* del
15q11.2	*TUBGCP5*	-	-	paternal del
	NIPA2	-	-	paternal del
	CYFIP1	-	-	paternal del
	NIPA1	-	-	paternal del
	WHAMML1	-	-	paternal del
15q13.2	*FAM7A1*	maternal del	maternal del	-
	FAM7A2	maternal del	maternal del	-
	ARHGAP11B	maternal del	maternal del	-
15q13.3	*CHRNA7*	maternal dup	paternal dup	-

All three individuals with autism who carry *de novo SHANK2* deletions also have rare CNVs in 15q11-13, a chromosomal region long associated with Angelman syndrome, Prader-Willi syndrome, ASD and other neuropsychiatric disorders. Patient 6319_3 is missing one copy of *CYFIP1* (cytoplasmic FMR1 interacting protein), a gene located in 15q11-13. It codes for a protein that binds FMRP, the protein missing in Fragile X syndrome. The deletion also contained genes *NIPA1* (non-imprinted in Prader-Willi/Angelman syndrome 1)*, NIPA2* (non-imprinted in Prader-Willi/Angelman syndrome 2*), CYFIP1* and *TUBGCP5* (tubulin, gamma complex associated protein 5). The deletion in this locus has been associated with schizophrenia, neurodevelopmental disorder, epilepsy and autism (Cooper et al. 2011, Stefansson et al. 2008). Patients AU038_3 and 5237_3 carry an extra copy of *CHRNA7* (cholinergic receptor, nicotinic, alpha 7), a gene that encodes a receptor of the acetylcholine neurotransmitter. This locus at 15q13.3 has been tied to ID, epilepsy, schizophrenia and ASD

(International schizophrenia consortium 2008, Sharp et al. 2008, Helbig et al. 2009, Miller et al. 2009).

Other additional CNVs detected in patients with *SHANK2* deletion contained ASD candidate genes including *CAMSAP1L* which codes calmodulin regulated spectrin-associated protein, *CNTN4* (contactin 4), *NLGN1* (neuroligin 1), protocadherin gene cluster *PCDHA1-10* and tyrosine phosphatase coding gene *DUSP22* (dual specificity protein phosphatase 22).

Recent results of exome sequencing revealed one additional frame shift mutation in *SHANK2* (Sanders et al. 2012). Another study detected aberrations of *SHANK2, CHRNA7* and also *ARHGAP11B* (Rho GTPase activating protein 11B) in a patient with intellectual disability and language impairment supports (Chilian et al. 2013). Notably, the combination of the disrupted genes was the same as in this study. The identification of mutations in *NRXN1, NLGN3/4X* and *SHANK2/3* strengthens the role of synaptic gene dysfunction in ASD. The NRXN-NLGN-SHANK pathway is associated with synaptogenesis, and imbalance between excitatory and inhibitory currents is a risk factor of ASD. It is currently unknown which or how many risk factors are sufficient to contribute to ASD. Co-occurrence of *de novo* mutations, together with inherited variations might play a role in genetic susceptibility to ASD. All of the 15q CNVs in this study were inherited, whereas the *SHANK2* deletions were spontaneous. For these patients, it is likely the genome cannot cope with the extra *de novo* event and these variants together contribute to ASD.

5.4 Genome wide scan, promoter and functional analysis in ASD (IV)

In Study IV we performed a GWA scan in a novel set of Finnish ASD families, paired with functional analysis, and a candidate gene based analysis of epistasis in ASD. In order to better understand the transcriptional regulation of the currently best known autism candidate genes, we performed an *in silico* analysis of their promoter sequences obtained from Genbank. In addition, we conducted association and promoter analyses of the autism candidate gene, *AVPR1A.*

5.4.1 Genome-wide scan and epistasis

We performed a genome wide scan in 81 ASD families (N=257) with TDT. Each family had 1-3 autistic individuals. In addition to TDT, we carried out a case-control association analysis, where the control material consisted of 750 unrelated Finnish individuals. The association and epistasis analysis were carried out with PLINK. The best associated genes from the TDT and case-control analyses are shown in Table 8. Epistatic effects were considered on a significance level of p-value ≤ 10^{-6}.

Table 8. Top-results in TDT and case-control analyses.

LOCUS	SNP	P (TDT)*	P (CASE-CONTROL)**	GENE	NEAR GENE
2p23.3	rs12613835	-	8.3×10^{-5}	*DTNB*	
2q31.1	rs4668086	1.1×10^{-3}	2.0×10^{-6}	*CERS6(LASS6)*	*NOSTRIN*
3p14.1	rs4855550	2.9×10^{-5}	3.0×10^{-6}		*FOXP1, UBA3, ARL6IP5*
5q21.3	rs902505	4.0×10^{-6}	–		*PJA2, MAN2A1*
12q21.2	rs10506759	2.0×10^{-6}	6.0×10^{-5}		*NAV3*
14q32.33	rs4983511	5.1×10^{-5}	8.0×10^{-6}		*INF2, TMEM179*
16p13.2	rs11649176	5.0×10^{-6}	3.6×10^{-5}	*RBFOX1*	
	rs17665171	3.0×10^{-6}	4.5×10^{-5}	*RBFOX1*	
	rs2058528	2.0×10^{-6}	6.4×10^{-5}	*RBFOX1*	
17q21.1	rs2071427	-	1.2×10^{-4}	*NR1D1*	*THRA*
20q12	rs6093634	-	8.2×10^{-6}	*PTPRT*	
Xq21.31	rs5941380	2.5×10^{-7}	–		*PCDH11X*
	rs35429716	3.5×10^{-6}	–	*PCDH11X*	

*permuted
**genomic-control corrected p-values

The TDT and case-control analyses revealed strong association at chromosomal loci 2p23.3, 2q31.1, 3p14.1, 5q21.3, 12q21.2, 14q32.33, 16p13.2, 17q21.1, 20q12 and Xq21.31. The strongest association was detected at SNPs located in gene *RBFOX1* at 16p13.2 and at SNP rs5941380 at Xq21.31.

In the TDT test, the narrow association peak at chromosome 16p13.2 is in the *RBFOX1* (RNA binding protein, fox-1 homolog (C. elegans) 1) gene, which is one of the largest genes in the

human genome. *RBFOX1* is strongly expressed in the brain and it regulates tissue-specific splicing (Fukumura et al. 2007). Transcriptional and splicing dysregulation are the underlying mechanism of neuronal dysfunction in autism (Voineagu et al. 2011). In earlier studies *RBFOX1* has been shown to be associated with several neurodevelopmental and neuropsychiatric disorders including ASD, mental retardation, epilepsy (Bhalla et al. 2004, Martin et al. 2007, Sebat et al. 2007, Wang et al. 2009, Wintle et al. 2011, Voineagu et al. 2011), ADHD (Elia et al. 2010), bipolar disorder, and schizophrenia (Xu et al. 2008, Hamshere et al. 2009, Le-Niculescu et al. 2009). The most significant SNPs in our study reside around exon 3 (in isoform 4, UCSC). This is noteworthy, as the RNA binding domain of the protein is coded by these first exons. Thus, mutations in this area might affect the regulation of transcription of isoform 4 and exclusion of the RNA binding domain.

PCDH11X/Y is located in a human XY homology region on Xq21.31 in loci where two SNPs showed strong association with autism in this study (rs5941380 and rs35429716). This gene belongs to the protocadherin gene family, a subfamily of the cadherin superfamily. The PCDH11X/Y protein is involved in cell-cell recognition essential for the segmental development and function of the central nervous system (Kim et al. 2011). The *PCDH11X* gene has been linked as a risk factor in late onset Alzheimer's disease (Carrasquillo et al. 2009). In a re-sequencing study of the X chromosome by Piton et al. (2011) one mutation was detected in a schizophrenia patient. Earlier, two CNVs covering parts of the *PCDH11X* gene have been detected in patients with autism (Pinto et al. 2010) and three CNVs in this thesis, see chapter 4.3. Considering these findings, it is interesting that in this thesis, in Study I, the best linkage peak was detected at Xq21 near *PCDH11X/Y* with a different sample set.

Suggestive associations were also detected in *DTNB* and *NR1D1*. DTNB is a member of the dystrophin-related protein family and is required for the maturation or maintenance of a subset of inhibitory synapses in the cerebellum (Grady et al. 2006). *NR1D1* codes for nuclear receptor subfamily 1, group D, member 1, and it is one of the circadian rhythm regulating genes. Circadian rhythm dysfunctions are typical in ASD and mood disorders. The *NR1D1* gene has previously been reported to be associated with bipolar disorder (Kripke et al. 2009).

Finally, we obtained suggestive evidence for SNP rs6093634 in *PTPRT (*protein tyrosine phosphatase, receptor type, T) at 20q12 in the case-control analyses ($P=8.2 \times 10^{-6}$). *PTPRT*

expression is restricted to the central nervous system and it is involved in both signal transduction and cellular adhesion in the central nervous system (McAndrew et al. 1998).

Despite our rather restricted sample size, we were able to find statistically significant associations in analysis of epistasis. 244 ASD candidate genes were chosen to epistasis analyses (See chapter 5.4.2; selection of candidate genes). We found tentative evidence of epistasis for 11 SNP combinations with p-values at least 10^{-7}. The best two associations are between *NRXN1* and *UBA52* (ubiquitin A-52 residue ribosomal protein fusion product 1). Epistasis interactions were also observed between genes *UBA52* and *TSC1* (tuberous sclerosis 1), *RBFOX1* and *SCN1A* (sodium channel, voltage-gated, type I, alpha subunit), as well as between *RBFOX3* (RNA binding protein, fox-1 homolog (C. elegans) 3) and *DLG2* (discs, large homolog 2 (Drosophila)). Notably, many epistasis locus pairs are not between the best associating genes, but also between genes which were not markedly associated in either the TDT or case-control analyses.

5.4.2 Transcription factor binding sites

For TF site prediction, candidate genes were selected based on the results of association analyses in this study, previous studies (AutismKB database), and functions of genes related to synaptic formation or regulation. In total, 244 genes were selected (included in epistatic analyses). From these 244 genes, 72 gained a score over 8 in the AutismKB database and were chosen for more refined analysis as the most plausible genes affecting autism. Neurospecific TF sites were predicted for 0-1 kb upstream of the genes with the Match algorithm (TRANSFAC database) (Matys et al. 2003). The predicted TF sites were compared with TF sites for a random set of 210 brain expressing genes.

TF site prediction showed differential distribution for presence of EGR binding site: it was found in 11.5% of candidate genes but in only 6.2 % of control genes ($P=2.5x10^{-7}$). Transcription factors EGR1 (early growth response protein 1) and EGR3 (early growth response protein 3) can bind to this site. EGRs belong to “immediate early genes” (IEGs). In the CNS, IEG transcription factors have been shown to contribute to neurite outgrowth,

neurotransmitter fate, and synapse plasticity (Lyons and West 2011). It is noteworthy that EGR3 is functionally linked to dopamine, glutamate, and neuregulin signaling. It has been found in three CNV (Papanikolaou et al. 2006, Ozgen 2009, Berkel et al. 2010) and one linkage study of autism (Bailey et al. 1998) and is significantly associated with schizophrenia (Yamada et al. 2007, Zhang et al. 2012) and bipolar disorder (Gallitano et al. 2012). In our GWA analysis, the best associating SNP, in the promoter area of EGR3, yielded P=0.0016. We believe that defects in EGR transcription factors, EGR1 and EGR3, or in their expression might disrupt the regulation of autism-linked neuronal genes perhaps more severely than that of other neuronal genes.

5.4.3 *AVPR1A*; promoter and association analysis

Finally, we carried out a promoter analysis and an extended association analysis of the *AVPR1A* gene. *AVPR1A* is known to affect social communication (Donaldson and Young 2008), and has also been shown to associate with autism (Kim et al. 2002, Wassink et al. 2004, Yirmiya et al. 2006, Yang et al. 2010a, Yang et al. 2010b, Tansey et al. 2011). Genotyping of microsatellites of *AVPR1A* gene was performed in 205 families affected with autism, consisting altogether of 737 members. The phenotype was restricted to infantile autism. We analyzed the association of the three microsatellites AVR, RS3 and RS1 and 12 tagSNPs in the promoter and coding regions of *AVPR1A* (Figure 15). FBAT version v2.0.2c was used for the family-based genetic association analyses. We made a comparison of the allele lengths between different studies because the primers used may vary. In our study the primers used are the same as in previous studies: Yirmiya et al. (2006) and Ukkola et al. (2009). A comparison of the length and association results of the RS1 alleles between this and previous autism studies (Kim et al. 20020, Wassink et al. 2004, Yirmiya et al. 2006, Yang et al. 2010a, Tansey et al. 2011) is shown in Table 9.

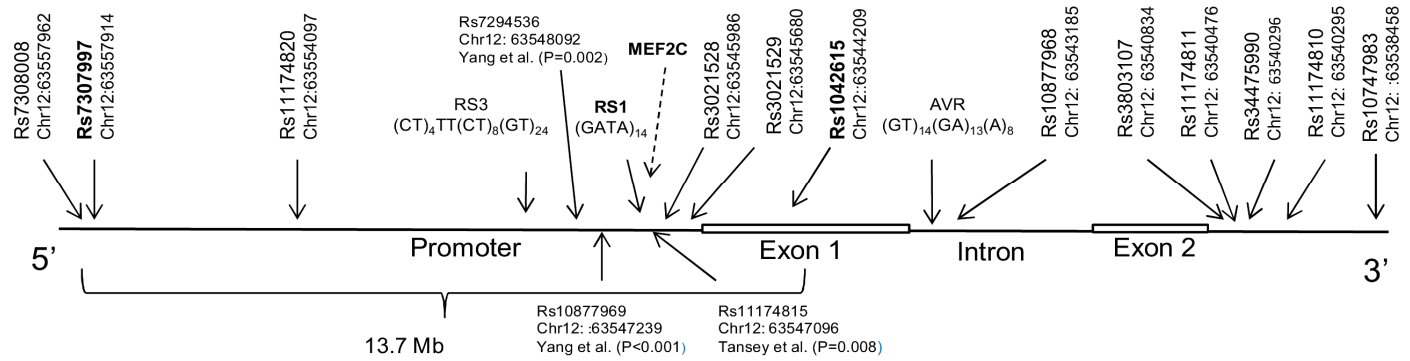


Figure 15. Schematic drawing of the human *AVPR1A* gene. Locations on markers studied here are presented. Markers associated with autism in this study are shown in bold. SNPs associated with autism by others are also presented. The binding site of MEF2C is located on the RS1 microsatellite.

Table 9. Comparison of the length of microsatellite allele RS1 of the *AVPR1A* gene between different autism studies. Alleles printed in bold were associated in autism in these studies.

Marker	Allele code in this study	P-value*	This study	Kim et al. 2002	Wassink et al. 2004	Yirmiya et al. 2006	Yang et al. 2010	Tansey et al. 2011
RS1	2		306 bp	312 bp	308 bp	306 bp	$(GATA)_9$	306 bp
	3	**0.018**	**310 bp**	316 bp	**312 bp**	310 bp	**$(GATA)_{10}$**	310 bp
	4		314 bp	320 bp	316 bp	314 bp	$(GATA)_{11}$	314 bp
	5		318 bp	324 bp	320 bp	318 bp	$(GATA)_{12}$	318 bp
	6		322 bp	328 bp	324 bp	322 bp	$(GATA)_{13}$	322 bp
	7		326 bp	332 bp	328 bp	326 bp	$(GATA)_{14}$	326 bp
	8		330 bp	336 bp	332 bp	330 bp	$(GATA)_{15}$	330 bp

*P-value of the allele associated with autism in this study in the recessive model.

The microsatellite locus RS1 gave globally significant P value with recessive model (0.038, permuted). The strongest association of RS1 alleles was observed at allele length of 310 bp (P=0.018). This allele has been shown to be associated with autism in earlier studies (Wassink et al. 2004, Yang et al. 2010a). In our study, we also noted several haplotypes which were associated with autism in our Finnish sample (Table 10). Still, the best associations were concentrated around the locus RS1.

Table 10. Haplotypes of studied alleles in *AVPR1A* gene associated with autism in this study.

Haplotype			***Additive***			***Dominant***			***Recessive***		
Markers	**Allele**	**afreq**	**Fam#***	**Z**	**P****	**Fam#***	**Z**	**P****	**Fam#***	**Z**	**P****
RS3-AVR	**4-4**	0.048	33	-2.403	**0.016**	33	-2.319	**0.021**			
RS3-RS1	**5-2**	0.046	38	2.205	**0.028**	38	2.114	**0.033**			
RS1-AVR	**2-3**	0.022	22	2.041	**0.043**	22	2.041	**0.039**			
RS1-SNP6	**3-1**	0.181							16	2.506	**0.019**
RS3-SNP6	**4-1**	0.070	41	-2.864	**0.004**	41	-2.791	**0.004**			
SNP6-SNP2	**2-2**	0.401							18	3.033	**0.002**
SNP6-SNP2	**1-1**	0.597				18	-3.033	**0.002**			

*minsize 10 **permuted ***whole marker permutation test (chisq sum). SNP2=rs7307997, SNP6=rs1042615

Notably, practically all associations found with *AVPR1A* are located in the promoter region of the gene. The promoter is polymorphic in length in all species studied so far (Hong et al. 2009). The microsatellites RS1 and RS3, located in the promoter region have shown differences in relative promoter activity by allele length: shorter alleles of RS1 decrease *AVPR1A* transcription which in turn increase the amygdala activity leading to social withdrawal (Meyer-Lindenberg et al. 2009, Tansey et al. 2011) characteristic to the autism phenotype. The mechanism of how the microsatellite allele length actually affects the transcription efficacy is not known.

Given that the microsatellite RS1 (and a few SNPs near it) has repeatedly been associated with several traits including autism (Kim et al. 2002, Ukkola et al. 2009, Skuse and Gallagher

2011), we became interested in this region, which is a part of the *AVPR1A* promoter. Somewhat surprisingly, the promoter analysis we carried out on the *AVPR1A* gene with TRANSFAC revealed a highly likely binding site of MEF2A/MEF2C and PBX1-PBX3 transcription factors residing on the 3´end of RS1 microsatellite. Of these, MEF2C has previously been implicated in autism (Novara et al. 2010). *PBX2* has been reported to be downregulated in autism (Kuwano et al. 2011).

MEF2 binding sites might be very important for the regulation of receptor transcription. Interestingly for autism, MEF2s have been shown to have direct roles in synapses: distinct forms of MEF2 TF act as positive or negative regulators of dendritic spine formation, with MEF2C playing a key regulator role in synaptic plasticity (Li et al. 2008, Barbosa et al. 2008, Dietrich et al. 2012). MEF2s interact with fragile X mental retardation protein (FMRP), to control the excitatory synapse elimination in an activity-dependent manner (Pfeiffer et al. 2010). MEF2C is also known to interact with MECP2, which is well known to be mutated in Rett syndrome (Amir et al. 1999). Moreover, *MEF2C* has been independently shown in autism and related disorders (Le Meuer et al. 2010, Zweier and Rauch 2012). Finally recent exome sequencing study of autism observed functional de novo mutation in *MEF2C* (Neale et al. 2012). We speculate that the relevance of MEF2C binding site in the RS1 microsatellite might be such that mutation(s) in the binding site sequence, or possibly, the sheer length differences of the microsatellite repeat adjacent to the binding site, could change the affinity of the transcription factor. This, in turn, could change either the magnitude of expression or the expression pattern of the AVP receptor in the brain, thereby affecting the brain´s ability to respond to AVP hormone. Obviously, functional experiments are required to confirm this hypothesis.

5.4.4 Functional analysis

We carried out pathway and function analysis with the IPA (Ingenuity Systems, www.ingenuity.com) system. The genes selected in the analyses were the same as in TF prediction analyses (N=72, see chapter 5.4.2). The nominally associated SNPs (P<0.001) from both the case-control and TDT analyses were investigated for enrichment of functions and pathways. The SNP results were assigned to genes within a 20 kb window. Results with no genes within the window were assigned to the closest genes not exceeding 1 Mb in distance. As result, there were 360 nominally associated genes in the case-control data.

Despite our rather restricted sample size, the functional analyses gave biological functional classes which fit quite well with the functions of known autism associated genes: schizophrenia, expression of DNA, transactivation, development of the brain and morphology of embryonic tissue. Previous studies (Anney et al. 2011, Hussman et al. 2011) have reported the results of pathway analyses in ASD. Both of these studies observed enrichment of genes and processes that have previously been examined in autistic disorders and offer biological plausibility to these findings.

Pathway-based methodologies are a new class of approaches to search for the underlying molecular mechanisms. Functional studies have revealed several candidate genes that likely function together, where defects in this network might cause autism. Figure 16 illustrates connections between the genes studied in this thesis (Studies I-IV).

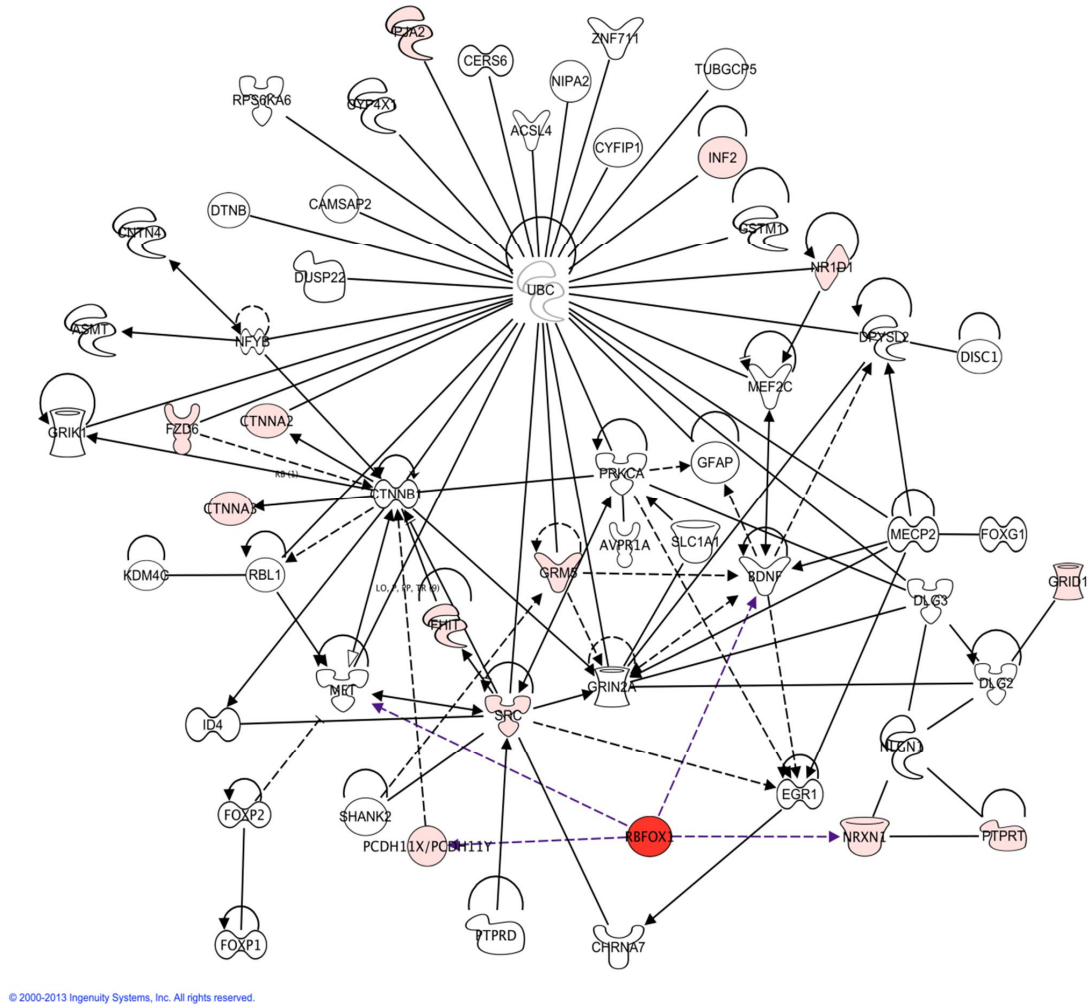


Figure 16. Connections between genes studied in this thesis. Genes that were found to associate with ASD in our GWA scan (Study IV) are colored in pink. *RBFOX1* which showed the strongest association in our dataset is colored in red. In addition, the UBC (polyubiquitin-C) protein is added as a hub. Black edges depict interactions acquired from the Ingenuity database, purple edges are based on Fogel et al. (2012). Dashed lines denote indirect proof of interaction. The map has been created with IPA. Proteins with no connections (according to IPA) were excluded.

5.5 Copy number variants (CNVs) in Finnish ASD families (unpublished data)

The results shown here are unpublished data analyzed from SNP data of the 81 ASD families analyzed in Study IV. In this chapter the large CNVs and CNVs containing ASD related genes are presented (frequency < 1% in the Finnish population). We detected an enrichment of large, rare deletions in ASD compared to the controls (P=0.02). This is in line with other published ASD studies (Pinto et al. 2010, Girirajan et al. 2011).

We identified 120 CNVs in ASD subjects which were rare (< 1 %) in the controls. Out of these, 73 % were previously reported in the AutismKB Database. The majority of them, 86 %, were inherited and 14 % were *de novo*. In most ASD families, inherited CNVs were transmitted from an apparently unaffected parent to the autistic child. Eleven of the rare CNVs were considered so called large-scale CNV's (Miller et al. 2010) as they were more than 400 kb in size (Table 11). CNVs which contain previously reported genes in ASD are presented in Table 12.

Table 11. Large CNVs in Finnish ASD samples.

Locus	ID	Start	End	Size	Dup/Del	Inheritance	Genes	References*
1q23.1	3016	158153526	158750786	597 kb	dup	paternal	*CD1A,CD1B,CD1C,CD1D,CD1E,OR10K1,OR10K2,OR10R2,OR10T2,OR10X1,OR10Z1,OR6K2,OR6K3,OR6K6,OR6N1,OR6N2,OR6P1,OR6Y1,SPTA1*	Pinto et al. 2010
3p14.2	3062	59958388	61036905	1.1 Mb	del	maternal	*FHIT*	Sebat et al. 2007, Bucan et al. 2009, Gai et al. 2011, Levy 2011, Nord 2011, Sanders 2011, Pinto 2010
10q21.1	1868	53192664	54196773	1.0 Mb	dup	maternal	*CSTF2T,DKK1,PRKG1*	Sebat et al. 2007, Sanders et al. 2011, Levy et al. 2011, Pinto et al. 2010
11p11.2-11p11.12	3017	48397233	48942781	546 kb	del	maternal	OR4A47	Pinto et al. 2010
11p11.2	2964	50095560	50697498	602 kb	dup	*de novo*	*LOC441601,LOC646813*	Szatmari et al. 2007, Christian et al. 2008, Pinto et al. 2010
15q13.2-15q13	3016	30936285	32514341	1.6 Mb	del	maternal	*CHRNA7,FAN1,KLF13,LOC100288637,MIR211,MTMR10,OTUD7A,TRPM1*	Sebat et al. 2007, Christian et al. 2008, Marshall et al. 2008, Weiss et al. 2008, Pinto et al. 2010, Bremer et al. 2011, Levy et al. 2011, Sanders et al. 2011

Locus	ID	Start	End	Size	Dup/Del	Inheritance	Genes	References*
16p11.2	1834	32137965	32941309	443 kb	dup	maternal	*LOC729264*	Marshall et al. 2008, Weiss et al. 2008, Gregory et al. 2009, Pinto et al. 2010
17q12	1831	34815551	36244358	1.4 Mb	del	*de novo*	*AATF,ACACA,C17orf78,DDX52,DHRS11,DUSP14,GGNBP2,HNF1B,LHX1,MRM1,MYO19,PIGW,SYNRG,TADA2A,ZNHIT3*	Gregory et al. 2009, Levy et al. 2011, Sanders et al. 2011
17q12	3063	34815551	36244358	1.4 Mb	del	*de novo*	*AATF,ACACA,C17orf78,DDX52,DHRS11,DUSP14,GGNBP2,HNF1B,LHX1,MRM1,MYO19,PIGW,SYNRG,TADA2A,ZNHIT3*	Gregory et al. 2009, Levy et al. 2011, Sanders et al. 2011
22q11.1	1493	16855618	17294251	439 kb	dup	paternal	*CCT8L2,XKR3*	Niklasson 2002, Levy 2011
Xq21.31	3076	91083043	91663444	580 kb	dub	paternal	*PCDH11X*	Pinto et al. 2010

*Published CNVs (AutismKB) which are overlapping the locations of our findings. CNVs in References are either deletions or duplications.

Table 12. CNVs containing ASD related genes.

Locus	ID	Start	End	Size	Dup/Del	Inheritance	Genes*	References
1q42.2	1869	231711489	231813134	102 kb	dup	maternal	***DISC1***	Pinto et al. 2010
1p13.3	1860	110133772	110348040	214 kb	dup	paternal	*AMPD2,EPS8L3,GNAI3, GNAT2,**GSTM1**,GSTM2, GSTM3,GSTM4,GSTM5*	Levy et al. 2011
14q12	3021	29006754	29309811	303 kb	dup	paternal	***FOXG1***	Pinto et al. 2010
17q21.31	3051	42880753	43202188	321 kb	dup	*de novo*	*C1QL1,CCDC103,DCAK D,EFTUD2,**GFAP**,GJC1, HIGD1B,KIF18B,NMT1, PLCD3*	Berkel et al. 2010
22q11.21	1925	18886915	18981563	95 kb	dup	paternal	*DGCR6,**PRODH***	Gai et al. 2010, Marshall et al. 2008, Nord 2011, Pinto et al. 2010, Sanders et al. 2010
22q11.21	1908	18738296	19008108	270 kb	del	maternal	*DGCR6,**PRODH***	Gai et al. 2010, Marshall et al. 2008, Nord 2011, Pinto et al. 2010, Sanders et al. 2010
Xq22.33	1908	1755236	2063606	308 kb	dup	maternal	***ASMT***	Marshall et al. 2008, Sanders et al. 2010

*ASD related genes as bold.

One of the subjects with infantile autism (3062) had a large 1.1 Mb maternally inherited deletion at the 3p14.2 region which contains the fragile histidine triad (*FHIT)* gene. *FHIT* deletions have recently been detected in Tourette syndrome (Fernandez et al. 2012). Association of 3p14.2 has been recognized in two separate studies of Asperger syndrome in Finnish families (Ylisaukko-oja et al. 2004, Rehnström et al. 2006).

A large deletion of 1.6 Mb at 15q13.2-15q13 was found in this study in patient 3016 with infantile autism. This CNV has been previously reported in several studies (Wolpert et al. 2000, Wassink et al. 2001, Silva et al. 2002, Keller et al. 2003, Sahoo et al. 2005, Sebat et al. 2007, Wassink et al. 2007, Christian et al. 2008, Marshall et al. 2008, Weiss et al. 2008, Bremer et al. 2010, Bucan et al. 2009, Gregory et al. 2009, van der Zwaag et al. 2010, Pinto et al. 2010, Bremer et al. 2011, Levy et al. 2011, Sanders et al. 2011). The deletion contains the *CHRNA7* gene. The same CNV was overlapping in ASD, schizophrenia and intellectual disability earlier but has also been detected in a control population (Guilmatre et al. 2009). Two patients with *SHANK2* deletion also had a CNV covering *CHRNA7* in Study III.

Two of the patients with autism (1831 and 3063) had a large *de novo* deletion at the 17q12 chromosomal region that has been reported in autistic subjects earlier. One of the genes in the CNV region is *HNF1B* (Homo sapiens HNF1 homeobox B). Patient 1831 also had diabetes. Interestingly, the *HNF1B* gene at this locus is associated in autism with renal disease or diabetes (Moreno-De-Luca et al. 2010, Loirat et al. 2010).

We detected CNVs in three autistic subjects at the 22q11 region, a previously known locus for 22q11 deletion/DiGeorge syndrome which alters cortical circuitry and may cause schizophrenia, autism, attention deficit/hyperactivity disorder or intellectual disability (Meechan et al. 2009, Niklasson et al. 2009). One of the patients (1493) had a large, 439 kb paternally inherited duplication in 22q11. Two of the subjects (1908, 1925) had inherited a 269 kb deletion and 95 kb duplication respectively at the 22q11.21 region containing genes *PRODH* and *DGCR6*. The 22q11 deletion spanning *PRODH* and *DGCR6* has been detected in ASD, schizophrenia and ID (Guillmatre et al. 2009). These disorders share biological pathways and they have recurrent rearrangements in synaptic and neurodevelopmental genes (Guillmatre et al. 2009).

In a recent study, adult patients with the 22q11.2 deletion were interviewed (Vorstman et al. 2013). The study reported that autism in the 22q11 deletion syndrome (22q11DS) patients does not increase the risk for psychosis in adulthood. ASD and schizophrenia should be regarded as distinct phenotypic manifestations of this deletion, hence the neuropsychiatric phenotypic expression of the 22q11.2 deletion is truly pleiotropic (Vorstman et al. 2013).

One 15 kb duplication at 1p33 was detected as a completely novel inherited variation present in the mother (1521) and her son (1518) who both have Asperger syndrome (Figure 14). This specific CNV contains a gene, *CYP4X1* (Homo sapiens cytochrome P450, family 4, subfamily X, polypeptide 1), which is expressed in several parts of the brain (Al-Anizy et al. 2006). It is mainly expressed in neurons in the brain stem, hippocampus, cortex, and cerebellum as well as in vascular endothelial cells (Bylund et al. 2002). The cytochrome P450 proteins are mono-oxygenases, which catalyze many reactions involved in drug metabolism and the synthesis of cholesterol, steroids and other lipids. The expression pattern of *CYP4X1* suggests that this protein may be involved in neurovascular functions in the brain (Bylund et al. 2002). We validated the CNV with qPCR (data not shown). Other members of this extended pedigree did not have the duplication. The grandmother as well as the aunt of the proband displayed some features of Asperger syndrome but their symptoms did not fulfill the criteria required for Asperger syndrome.

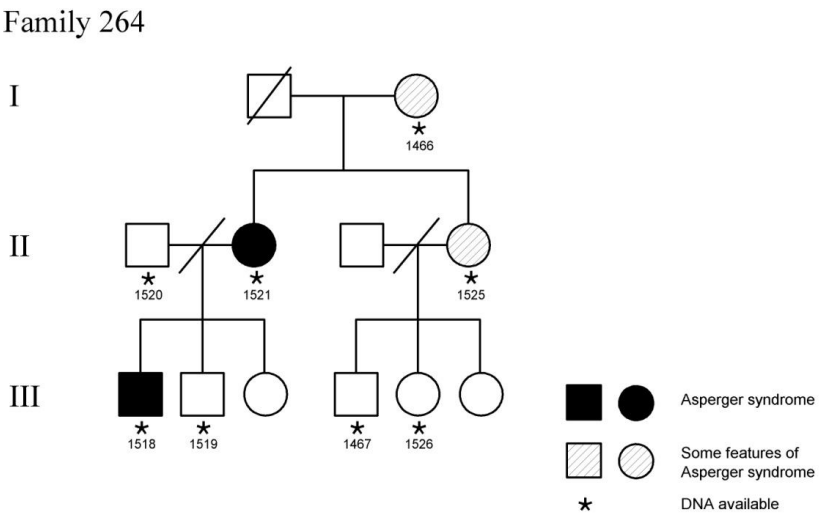


Figure 14. A novel inherited 15 kb duplication containing gene *CYP4X1* at 1p33. The duplication is present only in the mother and her son with Asperger syndrome.

Several CNVs found in this study overlapped with schizophrenia, ID, ADHD and anxiety disorders (Guilmatre et al. 2009, Fernandez et al. 2012, Elia et al. 2011). We detected several copy number variations that most likely are predisposing to autism, because they either arose *de novo* and/or overlapped with known micro-deletions and –duplications or they included genes which are important for synaptic functions and were not found in the normal population. Previously known ASD related genes which are located in the CNV regions detected in this study and whose exonic regions were covered with CNV included *DISC1, FOXG1, ASMT, PCDH11X, GSTM1, GFAP* and *PRODH* (Table 12).

Notably, a male patient (1869) had a 102 kb maternally inherited *DISC1* duplication. This patient had autism, epilepsy, and aggressive behavior. *DISC1* (disrupted in schizophrenia 1) codes the protein involved in neurite outgrowth and cortical development. *DISC1* is related to schizophrenia, bipolar disorder, and ASD (Millar et al. 2000, Ekelund et al. 2001, Hodgkinson et al. 2004, Kilpinen et al. 2008).

Finally, a girl had a paternally inherited 303 kb duplication at *FOXG1*. The phenotype was characterized by severe delay in psychomotor development and communication skills, psychotic behavior, and echolaly. The diagnosis was disintegrative disorder (F84.3).

Mutations in *FOXG1* (forkhead box G1) have previously been detected in atypical Rett syndrome (Florian et al. 2011). The *FOXG1* gene encodes a winged-helix transcriptional repressor which is essential for the development of the ventral telencephalon in the embryonic forebrain. Reported aberrations of *FOXG1* include point mutations, translocation, duplication, and a large deletion on chromosome 14q12 (Florian et al. 2011). Affected individuals with *FOXG1* mutations have shown dysmorphic features and a Rett-like clinical course, including normal perinatal period, post natal microcephaly, seizures, and severe mental retardation (Florian et al. 2011). The majority of patients with duplications on 14q12 containing *FOXG1* have developmental epilepsies, mental retardation, severe speech impairment, and infantile spasms (Brunetti-Pierri et al. 2011). Phenotypes differ presumably due to the varying size of the duplication.

6 CONCLUDING REMARKS AND FUTURE PROSPECTS

In the past decade, considerable efforts towards understanding the nature of ASD have been undertaken. Rapidly evolving genomic technologies have produced enormous advances in the genetics of ASD. Furthermore, the increasing amount of international, large study cohorts which combine ASD samples and scientific knowledge from research groups from all over the world has led highly reproducible findings.

In this thesis, the major finding was the strong association of the *RBFOX1* gene in the Finnish sample set. In future, the data will be combined with large, international data where some hints of association of *RBFOX1* have been observed. The replication of the GWA study in a bigger sample set might reveal genome-wide significant association of this gene if the associated SNPs are the same in these two different data sets. We will also perform haplotype analyses of *RBFOX1* and if the association remains, the whole gene will be sequenced. In this thesis, two SNPs in the GWA scan showed significant association with autism in the same locus Xq21, which was observed in our X-chromosomal linkage analysis earlier (Study I). Further analysis of haplotypes and homozygosity of common SNPs in this region could reveal X-linked recessive candidate variants that may be found in the isolated Finnish population.

Promoter analyses of autism candidate genes revealed the other important findings in this thesis. We observed that TF binding sites for EGR was enriched in autism candidate genes. Furthermore, a MEF2C binding site on the promoter region of the *AVPR1A* gene was located on the RS1 allele, which has shown association with autism in several studies, including ours. That might partly explain the association of the RS1 allele with autism. Cell biological studies are warranted to confirm the significance of these findings. Promoter analysis to predict transcription binding sites gives important information about the regulation of genes contributing to ASD. Next generation sequencing of the Finnish ASD samples would provide important information of the promoter regions in ASD.

In Study III of this thesis, we detected harmful mutations in the *SHANK2* gene which reduced synaptic density *in vitro*. Patients who had a deletion of *SHANK2* carried additional CNVs on chromosomal region 15q11-q13, previously associated in ASD and several other

neuropsychiatric disorders. This and previous studies suggest a multiple hit model for ASD. More studies are warranted to analyze what kind of combinations of genetic defects will be detected in the individual ASD families. In this thesis, we also performed a functional analysis of ASD related genes with IPA. The functional classes most significantly enriched in ASD were related to schizophrenia and development of the brain. We also observed connections between several genes studied in this thesis (Figure 16). It is probable that the defects on the network of ASD related genes cause the disorder, which again suggests a multiple hit model in ASD.

Rare variants have been under extensive study in the last few years in ASD. Large structural variants are present in 5-10 % of ASD cases. It is estimated that there are more than 200 CNV loci which predispose to ASD in the human genome. Each individual CNV is rare and accounts for less than 1 % of all cases of ASD. Large CNVs (>400 kb) which contain exonic regions of ASD candidate genes are likely to predispose to ASD. The same CNVs have been detected in ASD and other neuropsychiatric disorders such as ADHD, epilepsy, schizophrenia and intellectual disability. CNVs in ASD have been detected recurrently in chromosomal regions 15q13.3, 16p11.2, 17q12 and 22q11. We detected CNVs in our Finnish ASD sample in all of the aforementioned locations. CNVs detected in the ASD sample often also occur in unaffected individuals, although with lower frequency. Hence, it seems that none of the single CNVs are fully penetrant and it remains unsolved how many other predisposing genetic variants are sufficient to cause ASD.

Exome sequencing studies have revealed that non-synonymous *de novo* mutations are enriched in affected individuals. The results of exome sequencing studies together with CNV data indicate extensive locus heterogeneity but also provide a target for future diagnostics and therapeutics. Rare genetic causes for a disease could provide important knowledge for individual families. Another question is if the rare variants, even though they pinpoint to synaptic genes, are sufficient to cause ASD. It is conceivable that additional genetic variants and possibly triggering environmental factors are needed to cause autism. Much more information is coming in the next few years when next generation sequencing projects will publish more results.

Future work will have to investigate how all the rare and common predisposing variants interact together and with environmental factors. ASD comprise a broad range of variation in the severity of symptoms, from patients requiring special education and complete daily care to individuals with exceptional occupational skills. Considering the well characterized, large chromosomal aberrations known to predispose to ASD, it may be beneficial for individual families to get that information. It would decrease parental feelings of guilt and make it easier to estimate the recurrence risk for future children and other family members. There is no cure for autism yet and much more information has to be gathered from brain functions before we even get close to understanding the mechanisms of autism. Nevertheless, evidence of family-specific, rare genetic events such as mutations and *de novo* CNVs is accumulating. Exome sequencing and next qeneration sequencing studies will reveal much more evidence for this field in the next few years. Future studies will aim to translate genomic findings into molecular mechanisms. This information will be helpful for drug development which could specifically target these impaired molecular mechanisms.

7 ACKNOWLEDGEMENTS

This thesis was carried out in the Department of Medical Genetics at the University of Helsinki. I thank former and present heads of the Department of Medical Genetics, Päivi Peltomäki and Irma Järvelä for providing me with the chance to work in this department and for the excellent research facilities for this study.

This thesis was financially supported by the EC 6th FB Autism Molgen Consortium (grant number LSHM-CT-2005–512158), Sigrid Jusélius Foundation, EVO funding, the Finnish Medical Society, the Finnish Brain Foundation, Rinnekoti Foundation, University of Helsinki Funds, the Finnish Concordia Fund, Lastenlinna Foundation, Chancellor´s travel grants, the International Society of Psychiatric Genetics and the Lundbeck Foundation. I gratefully thank all the children with ASD and their families for participating in this study, without them this study would not have been possible.

I have been privileged to be supervised by two excellent mentors. I am grateful to my supervisor, Docent Irma Järvelä for running this project with enthusiasm and believing in my competence to overcome all the challenges during these years. I admire her hard working spirit and wide knowledge of the field of medical genetics. My warmest thanks go to my other supervisor, Docent Päivi Onkamo who has advised me with the statistical parts of the work and encouraged me to try several new methods which have been of great importance in this work. I admire her aim for perfection in science.

I am deeply grateful and honoured that Professor Leena Haataja accepted the role as the Opponent in the defence. MD, PhD Jacob Vorstman and Docent Liisa Myllykangas are acknowledged for reviewing this thesis and for their constructive and educative comments. MSc Minna Ahvenainen is thanked for flexible and quick language revision of this thesis.

I am extremely grateful for collaborators and co-authors of this study. Docent Raija Vanhala, Research Nurse Ulla Sarin-Seppänen, MD Reija Alen, MD, PhD, Taina Nieminen-von Wendt and the late Professor Lennart von Wendt are deeply acknowledged for their clinical effort and invaluable help with sample collection for this study. Raija Vanhala and Ulla Sarin-Seppänen are warmly thanked for helping and encouraging me many times during this

study. Huge thanks go to Raija for explaining to me the diagnostic methods in autism spectrum disorders and to Ulla for always being so cheerful and for looking for the amazing drawing on the cover of this book.

Of our international collaborators I want to especially thank Professor Thomas Bourgeron and Dr. Claire Leblond of the study of the *SHANK2* gene, and Dr. Richard Holt for the educative visit to the Wellcome Trust Center at the beginning of the project.

I have been lucky to work with a number of wonderful colleagues and good friends. I want to thank MSc Jaana Oikkonen for analyzing the GWAS data with me and for helping me with thousands of possible and impossible problems related to computers. I thank MSc Chakravarthi Kanduri for the never ending programming parts of the CNV analysis. I acknowledge MSc Ilona Kotala, PhD Karola Rehnström and PhD Tero Ylisaukko-oja for their contribution in the Xq study. I also want to acknowledge the whole research group of Irma Järvelä. Sanna Seitsonen is thanked for introducing me to the lab work in the beginning and answering my questions about thesis preparation lately. I acknowledge Minna Ahvenainen, Siiri Fuchs, Hanna Nurmi, Jenni Kallela, Ilona Kotala and Tarja Järvinen for assistance in the laboratory work of this study. Special thanks go to my good friend Liisa Ukkola-Vuoti who started and finished her thesis in the same time as I did and who has struggled with similar challenges with me during these years. We both know that *AVPR1A* is a special gene. I also thank Liisa and her husband Sauli for many nice trips, parties and other adventures in our free time. Fatma Doagu is deeply thanked for cheerful company in the office and lunch breaks. I also want to warmly thank other previous and present group members which I haven´t mention here yet, Paula Salo, Jarno Kivioja, Anju Philips, Katja Salo, Susanna Lemmelä, Suvi Torniainen and new menbers of the group Petri Peltonen and Preethy Nair. Maija Puhakka is acknowledged for the help in several practical issues at work.

I want to thank my colleague and relative Riitta Erola for inducing me to study clinical laboratory science at the University of Oulu. Warm thanks go also to my other colleagues and friends Mari Virtanen, Annariina Ahponen, Markku Yli-Pyky and Elina Malo from the University of Oulu and Inkeri Lokki who shared the office with me in Biomedicum earlier and who also shares common interests with me of human genetics and Lapland.

I want to thank my friends and colleagues Tarja Sundell and Martin Heger from our amazing conference trip to Honolulu and all the “Honolulu parties” and other activities afterwards.

I warmly acknowledge my relatives and family friends Anniina, Pikku-Aino, Paula and Heikki, Hanna Korhonen, Reetta and Petri, Anttoset, Rahkot, Lantot, Johanna and Antti, Kiiskiset, Toivat, Jantuset and Sulin-Aholat for several relaxing and fun moments during my life.

I want to thank all my friends for their presence and for not letting me forget the important things in life outside of the lab. I especially want to thank Taru-Tiina, Auli, Pia, Antti and Meri for amazing moments and adventures like “Himos-Juhannus” and cruises “abroad”.

I thank my sister Hanna and her family for all the support and being the “Voice of Reason” in my ongoing huge changes in work and life. It has been amazing to be an aunt to their 3-year old daughter Lotta. Her wonder of life has been joyful to follow. Finally I want to thank my parents for their endless help and support. Thank you for being there.

Helsinki, March 25th, 2013

Katri Kantojärvi

8 REFERENCES

1000 Genomes Project Consortium, Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., et al 2010, "A map of human genome variation from population-scale sequencing", *Nature,* vol. 467, no. 7319, pp. 1061-1073.

Abecasis, G.R., Cherny, S.S., Cookson, W.O. & Cardon, L.R. 2002, "Merlin--rapid analysis of dense genetic maps using sparse gene flow trees", *Nature genetics,* vol. 30, no. 1, pp. 97-101.

Abrahams, B.S. & Geschwind, D.H. 2008, "Advances in autism genetics: on the threshold of a new neurobiology", *Nature reviews.Genetics,* vol. 9, no. 5, pp. 341-355.

Abrams, S.A. 2002, "Nutritional rickets: an old disease returns", *Nutrition reviews,* vol. 60, no. 4, pp. 111-115.

Al-Anizy, M., Horley, N.J., Kuo, C.W., Gillett, L.C., Laughton, C.A., Kendall, D., et al 2006, "Cytochrome P450 Cyp4x1 is a major P450 protein in mouse brain", *The FEBS journal,* vol. 273, no. 5, pp. 936-947.

Alarcon, M., Cantor, R.M., Liu, J., Gilliam, T.C., Geschwind, D.H. & Autism Genetic Research Exchange Consortium 2002, "Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families", *American Journal of Human Genetics,* vol. 70, no. 1, pp. 60-71.

Alarcon, M., Yonan, A.L., Gilliam, T.C., Cantor, R.M. & Geschwind, D.H. 2005, "Quantitative genome scan and Ordered-Subsets Analysis of autism endophenotypes support language QTLs", *Molecular psychiatry,* vol. 10, no. 8, pp. 747-757.

Alkan, C., Coe, B.P. & Eichler, E.E. 2011, "Genome structural variation discovery and genotyping", *Nature reviews.Genetics,* vol. 12, no. 5, pp. 363-376.

American Psychiatric Association 1968, *DSM-II: Diagnostic and Statistical Manual of Mental Disorders,* 2th edn, American Psychiatric Association, Washington, DC, APA.

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", *Nature genetics,* vol. 23, no. 2, pp. 185-188.

Anney, R., Klei, L., Pinto, D., Almeida, J., Bacchelli, E., Baird, G., et al 2012, "Individual common variants exert weak effects on the risk for autism spectrum disorderspi", *Human molecular genetics,* vol. 21, no. 21, pp. 4781-4792.

Anney, R., Klei, L., Pinto, D., Regan, R., Conroy, J., Magalhaes, T.R., et al 2010, "A genome-wide scan for common alleles affecting risk for autism", *Human molecular genetics,* vol. 19, no. 20, pp. 4072-4082.

Anney, R.J., Kenny, E.M., O'Dushlaine, C., Yaspan, B.L., Parkhomenka, E., Buxbaum, J.D., et al 2011, "Gene-ontology enrichment analysis in two independent family-based samples highlights biologically plausible processes for autism spectrum disorders", *European journal of human genetics : EJHG,* vol. 19, no. 10, pp. 1082-1089.

Ariani, F., Hayek, G., Rondinella, D., Artuso, R., Mencarelli, M.A., Spanhol-Rosseto, A., et al 2008, "FOXG1 is responsible for the congenital variant of Rett syndrome", *American Journal of Human Genetics,* vol. 83, no. 1, pp. 89-93.

Arnold, P.D., Sicard, T., Burroughs, E., Richter, M.A. & Kennedy, J.L. 2006, "Glutamate transporter gene SLC1A1 associated with obsessive-compulsive disorder", *Archives of General Psychiatry,* vol. 63, no. 7, pp. 769-776.

Asperger, H. 1944, "Die autistischen Psychopathen im Kindersalter.", *Archiv für Psychiatrie und Nervenkrankenheiten,* vol. 1, pp. 76-136.

Auranen, M., Vanhala, R., Varilo, T., Ayers, K., Kempas, E., Ylisaukko-Oja, T., et al 2002, "A genomewide screen for autism-spectrum disorders: evidence for a major susceptibility locus on chromosome 3q25-27", *Am J Hum Genet,* vol. 71, no. 4, pp. 777-90.

Autism Genome Project Consortium, Paterson, A.D., Zwaigenbaum, L., Roberts, W., Brian, J., Liu, X.Q., et al 2007, "Mapping autism risk loci using genetic linkage and chromosomal rearrangements", *Nature genetics,* vol. 39, no. 3, pp. 319-328.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., et al 1995, "Autism as a strongly genetic disorder: evidence from a British twin study", *Psychological medicine,* vol. 25, no. 1, pp. 63-77.

Barbosa, A.C., Kim, M.S., Ertunc, M., Adachi, M., Nelson, E.D., McAnally, J., et al 2008, "MEF2C, a transcription factor that facilitates learning and memory by negative regulation of synapse numbers and function", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 105, no. 27, pp. 9391-9396.

Barnby, G., Abbott, A., Sykes, N., Morris, A., Weeks, D.E., Mott, R., et al 2005, "Candidate-gene screening and association analysis at the autism-susceptibility locus on chromosome 16p: evidence of association at GRIN2A and ABAT", *American Journal of Human Genetics,* vol. 76, no. 6, pp. 950-966.

Baron-Cohen, S. 2012, "Autism and the technical mind: children of scientists and engineers may inherit genes that not only confer intellectual talents but also predispose them to autism", *Scientific American,* vol. 307, no. 5, pp. 72-75.

Baron-Cohen, S. & Wheelwright, S. 1999, "'Obsessions' in children with autism or Asperger syndrome. Content analysis in terms of core domains of cognition", *The British journal of psychiatry : the journal of mental science,* vol. 175, pp. 484-490.

Beecham, G.W., Naj, A.C., Gilbert, J.R., Haines, J.L., Buxbaum, J.D. & Pericak-Vance, M.A. 2010, "PCDH11X variation is not associated with late-onset Alzheimer disease susceptibility", *Psychiatric genetics,* vol. 20, no. 6, pp. 321-324.

Bejerot, S. 2007, "An autistic dimension: a proposed subtype of obsessive-compulsive disorder", *Autism : the international journal of research and practice,* vol. 11, no. 2, pp. 101-110.

Bejerot, S., Nylander, L. & Lindstrom, E. 2001, "Autistic traits in obsessive-compulsive disorder", *Nordic journal of psychiatry,* vol. 55, no. 3, pp. 169-176.

Bellini, S. 2006, "The development of social anxiety in adolescents with autism spectrum disorders", *Focus on autism and other developmental disabilities,* vol. 21, no. 3, pp. 138.

Berkel, S., Marshall, C.R., Weiss, B., Howe, J., Roeth, R., Moog, U., et al 2010, "Mutations in the SHANK2 synaptic scaffolding gene in autism spectrum disorder and mental retardation", *Nature genetics,* vol. 42, no. 6, pp. 489-491.

Berkel, S., Tang, W., Trevino, M., Vogt, M., Obenhaus, H.A., Gass, P., et al 2012, "Inherited and de novo SHANK2 variants associated with autism spectrum disorder impair neuronal morphogenesis and physiology", *Human molecular genetics,* vol. 21, no. 2, pp. 344-357.

Betancur, C. 2011, "Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting", *Brain research,* vol. 1380, pp. 42-77.

Bettelheim, B. 1967, *The Empty Fortress: Infantile Autism and the Birth of the Self,* The Free Press, New York.

Bhalla, K., Phillips, H.A., Crawford, J., McKenzie, O.L., Mulley, J.C., Eyre, H., et al 2004, "The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene", *Journal of human genetics,* vol. 49, no. 6, pp. 308-311.

Bhat, S.S., Ladd, S., Grass, F., Spence, J.E., Brasington, C.K., Simensen, R.J., Schwartz, C.E., Dupont, B.R., Stevenson, R.E. & Srivastava, A.K. 2008, "Disruption of the IL1RAPL1 gene associated with a pericentromeric inversion of the X chromosome in a patient with mental retardation and autism.", *Clinical genetics,* [Online], vol. 73, no. 1, pp. 94.

Blanco, P., Sargent, C.A., Boucher, C.A., Mitchell, M. & Affara, N.A. 2000, "Conservation of PCDHX in mammals; expression of human X/Y genes predominantly in brain", *Mammalian genome : official journal of the International Mammalian Genome Society,* vol. 11, no. 10, pp. 906-914.

Bleuler, E. (ed) 1916, *Lehrbuch der Psychiatrie*, Verlag von Julius Springer, Berlin.

Bodfish, J.W., Symons, F.J., Parker, D.E. & Lewis, M.H. 2000, "Varieties of repetitive behavior in autism: comparisons to mental retardation", *Journal of Autism and Developmental Disorders,* vol. 30, no. 3, pp. 237-243.

Boeckers, T.M., Liedtke, T., Spilker, C., Dresbach, T., Bockmann, J., Kreutz, M.R., et al 2005, "C-terminal synaptic targeting elements for postsynaptic density proteins ProSAP1/Shank2 and ProSAP2/Shank3", *Journal of neurochemistry,* vol. 92, no. 3, pp. 519-524.

Bolton, P.F., Carcani-Rathwell, I., Hutton, J., Goode, S., Howlin, P. & Rutter, M. 2011, "Epilepsy in autism: features and correlates", *The British journal of psychiatry : the journal of mental science,* vol. 198, no. 4, pp. 289-294.

Bonati, M.T., Russo, S., Finelli, P., Valsecchi, M.R., Cogliati, F., Cavalleri, F., et al 2007, "Evaluation of autism traits in Angelman syndrome: a resource to unfold autism genes", *Neurogenetics,* vol. 8, no. 3, pp. 169-178.

Bourgeron, T. 2009, "A synaptic trek to autism", *Current opinion in neurobiology,* vol. 19, no. 2, pp. 231-234.

Bremer, A., Giacobini, M., Eriksson, M., Gustavsson, P., Nordin, V., Fernell, E., et al 2011, "Copy number variation characteristics in subpopulations of patients with autism spectrum disorders", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 156, no. 2, pp. 115-124.

Bruder, C.E., Piotrowski, A., Gijsbers, A.A., Andersson, R., Erickson, S., Diaz de Stahl, T., et al 2008, "Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles", *American Journal of Human Genetics,* vol. 82, no. 3, pp. 763-771.

Brune, C.W., Kim, S.J., Hanna, G.L., Courchesne, E., Lord, C., Leventhal, B.L., et al 2008, "Family-Based Association Testing of OCD-associated SNPs of SLC1A1 in an autism sample", *Autism research : official journal of the International Society for Autism Research,* vol. 1, no. 2, pp. 108-113.

Brunetti-Pierri, N., Paciorkowski, A.R., Ciccone, R., Della Mina, E., Bonaglia, M.C., Borgatti, R., et al 2011, "Duplications of FOXG1 in 14q12 are associated with developmental epilepsy, mental retardation, and severe speech impairment", *European journal of human genetics : EJHG,* vol. 19, no. 1, pp. 102-107.

Bryson, S.E., Clark, B.S. & Smith, I.M. 1988, "First report of a Canadian epidemiological study of autistic syndromes", *Journal of child psychology and psychiatry, and allied disciplines,* vol. 29, no. 4, pp. 433-445.

Bucan, M., Abrahams, B.S., Wang, K., Glessner, J.T., Herman, E.I., Sonnenblick, L.I., et al 2009, "Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes", *PLoS genetics,* vol. 5, no. 6, pp. e1000536.

Buchen, L. 2011, "Scientists and autism: When geeks meet", *Nature,* vol. 479, no. 7371, pp. 25-27.

Butler, J.M. 2006, "Genetics and genomics of core short tandem repeat loci used in human identity testing", *Journal of forensic sciences,* vol. 51, no. 2, pp. 253-265.

Buxbaum, J.D., Silverman, J., Keddache, M., Smith, C.J., Hollander, E., Ramoz, N., et al 2004, "Linkage analysis for autism in a subset families with obsessive-compulsive behaviors: evidence for an autism susceptibility gene on chromosome 1 and further support for susceptibility genes on chromosome 6 and 19", *Molecular psychiatry,* vol. 9, no. 2, pp. 144-150.

Buxbaum, J.D., Silverman, J.M., Smith, C.J., Kilifarski, M., Reichert, J., Hollander, E., et al 2001, "Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity", *American Journal of Human Genetics,* vol. 68, no. 6, pp. 1514-1520.

Bylund, J., Zhang, C. & Harder, D.R. 2002, "Identification of a novel cytochrome P450, CYP4X1, with unique localization specific to the brain", *Biochemical and biophysical research communications,* vol. 296, no. 3, pp. 677-684.

Cantor, R.M., Kono, N., Duvall, J.A., Alvarez-Retuerto, A., Stone, J.L., Alarcon, M., et al 2005, "Replication of autism linkage: fine-mapping peak at 17q21", *American Journal of Human Genetics,* vol. 76, no. 6, pp. 1050-1056.

Carrasquillo, M.M., Zou, F., Pankratz, V.S., Wilcox, S.L., Ma, L., Walker, L.P., et al 2009, "Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease", *Nature genetics,* vol. 41, no. 2, pp. 192-198.

Cashin, A., Sci, D.A. & Barker, P. 2009, "The triad of impairment in autism revisited", *Journal of child and adolescent psychiatric nursing : official publication of the Association of Child and Adolescent Psychiatric Nurses, Inc,* vol. 22, no. 4, pp. 189-193.

Chakrabarti, S. & Fombonne, E. 2001, "Pervasive developmental disorders in preschool children", *JAMA : the journal of the American Medical Association,* vol. 285, no. 24, pp. 3093-3099.

Chilian, B., Abdollahpour, H., Bierhals, T., Haltrich, I., Fekete, G., Nagel, I., et al 2013, "Dysfunction of SHANK2 and CHRNA7 in a patient with intellectual disability and language impairment supports genetic epistasis of the two loci", *Clinical genetics,* .

Ching, M.S., Shen, Y., Tan, W.H., Jeste, S.S., Morrow, E.M., Chen, X., et al 2010, "Deletions of NRXN1 (neurexin-1) predispose to a wide spectrum of developmental disorders", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 153B, no. 4, pp. 937-947.

Choy, K.W., Setlur, S.R., Lee, C. & Lau, T.K. 2010, "The impact of human copy number variation on a new era of genetic testing", *BJOG : an international journal of obstetrics and gynaecology,* vol. 117, no. 4, pp. 391-398.

Christian, S.L., Brune, C.W., Sudi, J., Kumar, R.A., Liu, S., Karamohamed, S., et al 2008, "Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder", *Biological psychiatry,* vol. 63, no. 12, pp. 1111-1117.

Clamp, M., Fry, B., Kamal, M., Xie, X., Cuff, J., Lin, M.F., et al 2007, "Distinguishing protein-coding and noncoding genes in the human genome", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 104, no. 49, pp. 19428-19433.

Cloos, P.A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., Antal, T., et al 2006, "The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3", *Nature,* vol. 442, no. 7100, pp. 307-311.

Cohen, I.L. & Tsiouris, J.A. 2006, "Maternal recurrent mood disorders and high-functioning autism", *Journal of Autism and Developmental Disorders,* vol. 36, no. 8, pp. 1077-1088.

Colella, S., Yau, C., Taylor, J.M., Mirza, G., Butler, H., Clouston, P., et al 2007, "QuantiSNP: an Objective Bayes Hidden-Markov Model to detect and accurately map copy number variation using SNP genotyping data", *Nucleic acids research,* vol. 35, no. 6, pp. 2013-2025.

Connolly, J.J., Glessner, J.T. & Hakonarson, H. 2012, "A Genome-Wide Association Study of Autism Incorporating Autism Diagnostic Interview-Revised, Autism Diagnostic Observation Schedule, and Social Responsiveness Scale", *Child development,* .

Conrad, D.F., Keebler, E.M., DePristo, M.A., Lindsay, S.J., Zhang, Y., Casals, F., et al 2011,"Variation in genome-wide mutation rates within and between human families ", *Nat Genet,* vol. 43, no. 7, pp. 712-715.

Conrad, D.F., Andrews, T.D., Carter, N.P., Hurles, M.E. & Pritchard, J.K. 2006, "A high-resolution survey of deletion polymorphism in the human genome", *Nature genetics,* vol. 38, no. 1, pp. 75-81.

Coon, H., Matsunami, N., Stevens, J., Miller, J., Pingree, C., Camp, N.J., et al 2005, "Evidence for linkage on chromosome 3q25-27 in a large autism extended pedigree", *Human heredity,* vol. 60, no. 4, pp. 220-226.

Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., et al 2011, "A copy number variation morbidity map of developmental delay", *Nature genetics,* vol. 43, no. 9, pp. 838-846.

Courchesne, E., Mouton, P.R., Calhoun, M.E., Semendeferi, K., Ahrens-Barbeau, C., Hallet, M.J., et al 2011, "Neuron number and size in prefrontal cortex of children with autism", *JAMA : the journal of the American Medical Association,* vol. 306, no. 18, pp. 2001-2010.

Crow, J.F. 2000, "The origins, patterns and implications of human spontaneous mutation", *Nature reviews.Genetics,* vol. 1, no. 1, pp. 40-47.

de Krom, M., Staal, W.G., Ophoff, R.A., Hendriks, J., Buitelaar, J., Franke, B., et al 2009, "A common variant in DRD3 receptor is associated with autism spectrum disorder", *Biological psychiatry,* vol. 65, no. 7, pp. 625-630.

de Vries, B.B., Pfundt, R., Leisink, M., Koolen, D.A., Vissers, L.E., Janssen, I.M., et al 2005, "Diagnostic genome profiling in mental retardation", *American Journal of Human Genetics,* vol. 77, no. 4, pp. 606-616.

Dealberto, M.J. 2011, "Prevalence of autism according to maternal immigrant status and ethnic origin", *Acta Psychiatrica Scandinavica,* vol. 123, no. 5, pp. 339-348.

Dellinger, A.E., Saw, S.M., Goh, L.K., Seielstad, M., Young, T.L. & Li, Y.J. 2010, "Comparative analyses of seven algorithms for copy number variant identification from single nucleotide polymorphism arrays", *Nucleic acids research,* vol. 38, no. 9, pp. e105.

DeLong, R. 2004, "Autism and familial major mood disorder: are they related?", *The Journal of neuropsychiatry and clinical neurosciences,* vol. 16, no. 2, pp. 199-213.

Dickel, D.E., Veenstra-VanderWeele, J., Cox, N.J., Wu, X., Fischer, D.J., Van Etten-Lee, M., et al 2006, "Association testing of the positional and functional candidate gene SLC1A1/EAAC1 in early-onset obsessive-compulsive disorder", *Archives of General Psychiatry,* vol. 63, no. 7, pp. 778-785.

Dietrich, J.B., Takemori, H., Grosch-Dirrig, S., Bertorello, A. & Zwiller, J. 2012, "Cocaine induces the expression of MEF2C transcription factor in rat striatum through activation of SIK1 and phosphorylation of the histone deacetylase HDAC5", *Synapse (New York, N.Y.),* vol. 66, no. 1, pp. 61-70.

Donaldson, Z.R. & Young, L.J. 2008, "Oxytocin, vasopressin, and the neurogenetics of sociality", *Science (New York, N.Y.),* vol. 322, no. 5903, pp. 900-904.

Durand, C.M., Betancur, C., Boeckers, T.M., Bockmann, J., Chaste, P., Fauchereau, F., et al 2007, "Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders", *Nature genetics,* vol. 39, no. 1, pp. 25-27.

Durand, C.M., Kappeler, C., Betancur, C., Delorme, R., Quach, H., Goubran-Botros, H., et al 2006, "Expression and genetic variability of PCDH11Y, a gene specific to Homo sapiens and candidate for susceptibility to psychiatric disorders", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 141B, no. 1, pp. 67-70.

Duvall, J.A., Lu, A., Cantor, R.M., Todd, R.D., Constantino, J.N. & Geschwind, D.H. 2007, "A quantitative trait locus analysis of social responsiveness in multiplex autism families", *The American Journal of Psychiatry,* vol. 164, no. 4, pp. 656-662.

Ehlers, S. & Gillberg, C. 1993, "The epidemiology of Asperger syndrome. A total population study", *Journal of child psychology and psychiatry, and allied disciplines,* vol. 34, no. 8, pp. 1327-1350.

Ekelund, J., Hovatta, I., Parker, A., Paunio, T., Varilo, T., Martin, R., et al 2001, "Chromosome 1 loci in Finnish schizophrenia families", *Human molecular genetics,* vol. 10, no. 15, pp. 1611-1617.

Elia, J., Gai, X., Xie, H.M., Perin, J.C., Geiger, E., Glessner, J.T., et al 2010, "Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes", *Molecular psychiatry,* vol. 15, no. 6, pp. 637-646.

Elia, J., Glessner, J.T., Wang, K., Takahashi, N., Shtir, C.J., Hadley, D., et al 2011, "Genome-wide copy number variation study associates metabotropic glutamate receptor gene networks with attention deficit hyperactivity disorder", *Nature genetics,* vol. 44, no. 1, pp. 78-84.

Elsabbagh, M., Divan, G., Koh, Y.J., Kim, Y.S., Kauchali, S., Marcin, C., et al 2012, "Global prevalence of autism and other pervasive developmental disorders", *Autism research : official journal of the International Society for Autism Research,* vol. 5, no. 3, pp. 160-179.

Endele, S., Rosenberger, G., Geider, K., Popp, B., Tamer, C., Stefanova, I., et al 2010, "Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes", *Nature genetics,* vol. 42, no. 11, pp. 1021-1026.

Eyles, D.W., Burne, T.H. & McGrath, J.J. 2012, "Vitamin D, effects on brain development, adult brain function and the links between low levels of vitamin D and neuropsychiatric disease", *Frontiers in neuroendocrinology,* .

Farrington, C.P., Miller, E. & Taylor, B. 2001, "MMR and autism: further evidence against a causal association", *Vaccine,* vol. 19, no. 27, pp. 3632-3635.

Fernandez, T.V., Sanders, S.J., Yurkiewicz, I.R., Ercan-Sencicek, A.G., Kim, Y.S., Fishman, D.O., et al 2012, "Rare copy number variants in tourette syndrome disrupt genes in histaminergic pathways and overlap with autism", *Biological psychiatry,* vol. 71, no. 5, pp. 392-402.

Fink, S., Excoffier, L. & Heckel, G. 2007, "High variability and non-neutral evolution of the mammalian avpr1a gene", *BMC evolutionary biology,* vol. 7, pp. 176.

First, M.B. 2009, "Harmonisation of ICD-11 and DSM-V: opportunities and challenges", *The British journal of psychiatry : the journal of mental science,* vol. 195, no. 5, pp. 382-390.

Florian, C., Bahi-Buisson, N. & Bienvenu, T. 2012, "FOXG1-Related Disorders: From Clinical Description to Molecular Genetics", *Molecular syndromology,* vol. 2, no. 3-5, pp. 153-163.

Fogel, B.L., Wexler, E., Wahnich, A., Friedrich, T., Vijayendran, C., Gao, F., et al 2012, "RBFOX1 regulates both splicing and transcriptional networks in human neuronal development", *Human molecular genetics,* vol. 21, no. 19, pp. 4171-4186.

Folstein, S. & Rutter, M. 1977b, "Genetic influences and infantile autism", *Nature,* vol. 265, no. 5596, pp. 726-728.

Folstein, S. & Rutter, M. 1977a, "Infantile autism: a genetic study of 21 twin pairs", *Journal of child psychology and psychiatry, and allied disciplines,* vol. 18, no. 4, pp. 297-321.

Folstein, S.E. & Rosen-Sheidley, B. 2001, "Genetics of autism: complex aetiology for a heterogeneous disorder", *Nature reviews.Genetics,* vol. 2, no. 12, pp. 943-955.

Fombonne, E. 2005, "Epidemiology of autistic disorder and other pervasive developmental disorders", *The Journal of clinical psychiatry,* vol. 66 Suppl 10, pp. 3-8.

Fombonne, E., Zakarian, R., Bennett, A., Meng, L. & McLean-Heywood, D. 2006, "Pervasive developmental disorders in Montreal, Quebec, Canada: prevalence and links with immunizations", *Pediatrics,* vol. 118, no. 1, pp. e139-50.

Freitag, C.M., Staal, W., Klauck, S.M., Duketis, E. & Waltes, R. 2010, "Genetics of autistic disorders: review and clinical implications", *European child & adolescent psychiatry,* vol. 19, no. 3, pp. 169-178.

Fukumura, K., Kato, A., Jin, Y., Ideue, T., Hirose, T., Kataoka, N., et al 2007, "Tissue-specific splicing regulator Fox-1 induces exon skipping by interfering E complex formation on the downstream intron of human F1gamma gene", *Nucleic acids research,* vol. 35, no. 16, pp. 5303-5311.

Gadow, K.D., Roohi, J., DeVincent, C.J., Kirsch, S. & Hatchwell, E. 2010, "Glutamate transporter gene (SLC1A1) single nucleotide polymorphism (rs301430) and repetitive behaviors and anxiety in children with autism spectrum disorder", *Journal of Autism and Developmental Disorders,* vol. 40, no. 9, pp. 1139-1145.

Gai, X., Xie, H.M., Perin, J.C., Takahashi, N., Murphy, K., Wenocur, A.S., et al 2012, "Rare structural variation of synapse and neurotransmission genes in autism", *Molecular psychiatry,* vol. 17, no. 4, pp. 402-411.

Gallitano, A.L., Tillman, R., Dinu, V. & Geller, B. 2012, "Family-based association study of early growth response gene 3 with child bipolar I disorder", *Journal of affective disorders,* vol. 138, no. 3, pp. 387-396.

Gardener, H., Spiegelman, D. & Buka, S.L. 2011, "Perinatal and neonatal risk factors for autism: a comprehensive meta-analysis", *Pediatrics,* vol. 128, no. 2, pp. 344-355.

Gauthier, J., Champagne, N., Lafreniere, R.G., Xiong, L., Spiegelman, D., Brustein, E., et al 2010, "De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 107, no. 17, pp. 7863-7868.

Geschwind, D.H. 2009, "Advances in autism", *Annual Review of Medicine,* vol. 60, pp. 367-380.

Gillberg, C., Gillberg, C., Rastam, M. & Wentz, E. 2001, "The Asperger Syndrome (and high-functioning autism) Diagnostic Interview (ASDI): a preliminary study of a new structured clinical interview", *Autism : the international journal of research and practice,* vol. 5, no. 1, pp. 57-66.

Gilman, S.R., Iossifov, I., Levy, D., Ronemus, M., Wigler, M. & Vitkup, D. 2011, "Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses", *Neuron,* vol. 70, no. 5, pp. 898-907.

Girirajan, S., Brkanac, Z., Coe, B.P., Baker, C., Vives, L., Vu, T.H., et al 2011, "Relative burden of large CNVs on a range of neurodevelopmental phenotypes", *PLoS genetics,* vol. 7, no. 11, pp. e1002334.

Girirajan, S., Rosenfeld, J.A., Cooper, G.M., Antonacci, F., Siswara, P., Itsara, A., et al 2010, "A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay", *Nature genetics,* vol. 42, no. 3, pp. 203-209.

Glancy, M., Barnicoat, A., Vijeratnam, R., de Souza, S., Gilmore, J., Huang, S., et al 2009, "Transmitted duplication of 8p23.1-8p23.2 associated with speech delay, autism and learning difficulties", *European journal of human genetics : EJHG,* vol. 17, no. 1, pp. 37-43.

Glessner, J.T., Connolly, J.J. & Hakonarson, H. 2012, "Rare Genomic Deletions and Duplications and their Role in Neurodevelopmental Disorders", *Current topics in behavioral neurosciences,* vol. 12, pp. 345-360.

Glessner, J.T., Wang, K., Cai, G., Korvatska, O., Kim, C.E., Wood, S., et al 2009, "Autism genome-wide copy number variation reveals ubiquitin and neuronal genes", *Nature,* vol. 459, no. 7246, pp. 569-573.

Grady, R.M., Wozniak, D.F., Ohlemiller, K.K. & Sanes, J.R. 2006, "Cerebellar synaptic defects and abnormal motor behavior in mice lacking alpha- and beta-dystrobrevin", *The Journal of neuroscience : the official journal of the Society for Neuroscience,* vol. 26, no. 11, pp. 2841-2851.

Grandin, T. & Scariano, M. 1996, *Emergence labelled autistic.* Warner Books, New York.

Gregory, S.G., Connelly, J.J., Towers, A.J., Johnson, J., Biscocho, D., Markunas, C.A., et al 2009, "Genomic and epigenetic evidence for oxytocin receptor deficiency in autism", *BMC medicine,* vol. 7, pp. 62-7015-7-62.

Guilmatre, A., Dubourg, C., Mosca, A.L., Legallic, S., Goldenberg, A., Drouin-Garraud, V., et al 2009, "Recurrent rearrangements in synaptic and neurodevelopmental genes and shared biologic pathways in schizophrenia, autism, and mental retardation", *Archives of General Psychiatry,* vol. 66, no. 9, pp. 947-956.

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., et al 2011, "Genetic heritability and shared environmental factors among twin pairs with autism", *Archives of General Psychiatry,* vol. 68, no. 11, pp. 1095-1102.

Hamdan, F.F., Gauthier, J., Araki, Y., Lin, D.T., Yoshizawa, Y., Higashi, K., et al 2011, "Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability", *American Journal of Human Genetics,* vol. 88, no. 3, pp. 306-316.

Hamshere, M.L., Green, E.K., Jones, I.R., Jones, L., Moskvina, V., Kirov, G., et al 2009, "Genetic utility of broadly defined bipolar schizoaffective disorder as a diagnostic concept", *The British journal of psychiatry : the journal of mental science,* vol. 195, no. 1, pp. 23-29.

Hanna, G.L., Veenstra-VanderWeele, J., Cox, N.J., Boehnke, M., Himle, J.A., Curtis, G.C., et al 2002, "Genome-wide linkage analysis of families with obsessive-compulsive disorder ascertained through pediatric probands", *American Journal of Medical Genetics,* vol. 114, no. 5, pp. 541-552.

Haraksingh, R.R., Abyzov, A., Gerstein, M., Urban, A.E. & Snyder, M. 2011, "Genome-wide mapping of copy number variation in humans: comparative analysis of high resolution array platforms", *PloS one,* vol. 6, no. 11, pp. e27859.

Hebebrand, J. & Buitelaar, J.K. 2011, "On the way to DSM-V", *European child & adolescent psychiatry,* vol. 20, no. 2, pp. 57-60.

Helbig, I., Mefford, H.C., Sharp, A.J., Guipponi, M., Fichera, M., Franke, A., et al 2009, "15q13.3 Microdeletions Increase Risk of Idiopathic Generalized Epilepsy", *Nature genetics,* vol. 41, no. 2, pp. 160-162.

Hirata, Y., Zai, C.C., Souza, R.P., Lieberman, J.A., Meltzer, H.Y. & Kennedy, J.L. 2012, "Association study of GRIK1 gene polymorphisms in schizophrenia: case-control and family-based studies", *Human psychopharmacology,* vol. 27, no. 4, pp. 345-351.

Hodgkinson, C.A., Goldman, D., Jaeger, J., Persaud, S., Kane, J.M., Lipsky, R.H., et al 2004, "Disrupted in schizophrenia 1 (DISC1): association with schizophrenia, schizoaffective disorder, and bipolar disorder", *American Journal of Human Genetics,* vol. 75, no. 5, pp. 862-872.

Hong, K.W., Matsukawa, R., Hirata, Y., Hayasaka, I., Murayama, Y., Ito, S., et al 2009, "Allele distribution and effect on reporter gene expression of vasopressin receptor gene (AVPR1a)-linked VNTR in primates", *Journal of neural transmission (Vienna, Austria : 1996),* vol. 116, no. 5, pp. 535-538.

Horvath, S., Xu, X. & Laird, N.M. 2001, "The family based association test method: strategies for studying general genotype--phenotype associations", *European journal of human genetics : EJHG,* vol. 9, no. 4, pp. 301-306.

Howlin, P., Goode, S., Hutton, J. & Rutter, M. 2004, "Adult outcome for children with autism", *Journal of child psychology and psychiatry, and allied disciplines,* vol. 45, no. 2, pp. 212-229.

Howlin, P., Magiati, I. & Charman, T. 2009, "Systematic review of early intensive behavioral interventions for children with autism", *American journal on intellectual and developmental disabilities,* vol. 114, no. 1, pp. 23-41.

Howlin, P., Magiati, I. & Charman, T. 2009, "Systematic review of early intensive behavioral interventions for children with autism", *American journal on intellectual and developmental disabilities,* vol. 114, no. 1, pp. 23-41.

Hu, V.W., Sarachana, T., Kim, K.S., Nguyen, A., Kulkarni, S., Steinberg, M.E., et al 2009, "Gene expression profiling differentiates autism case-controls and phenotypic variants of autism spectrum disorders: evidence for circadian rhythm dysfunction in severe autism", *Autism research : official journal of the International Society for Autism Research,* vol. 2, no. 2, pp. 78-97.

Hussman, J.P., Chung, R.H., Griswold, A.J., Jaworski, J.M., Salyakina, D., Ma, D., et al 2011, "A noise-reduction GWAS analysis implicates altered regulation of neurite outgrowth and guidance in autism", *Molecular autism,* vol. 2, no. 1, pp. 1-2392-2-1.

Icasiano, F. 2004, "Childhood autism spectrum disorder in the Barwon region: a community based study", *Journal of paediatrics and child health,* vol. 40, no. 12, pp. 696.

IMGSAC,International Molecular Genetic Study of Autism Consortium 1998, "A full genome screen for autism with evidence for linkage to a region on chromosome 7q.", *Human molecular genetics,* vol. 7, no. 3, pp. 571-578.

International HapMap Consortium, Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., et al 2007, "A second generation human haplotype map of over 3.1 million SNPs", *Nature,* vol. 449, no. 7164, pp. 851-861.

International Human Genome Sequencing Consortium 2004, "Finishing the euchromatic sequence of the human genome", *Nature,* vol. 431, no. 7011, pp. 931-945.

International Molecular Genetic Study of Autism Consortium (IMGSAC) 2001, "A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p", *American Journal of Human Genetics,* vol. 69, no. 3, pp. 570-581.

International Schizophrenia Consortium 2008, "Rare chromosomal deletions and duplications increase risk of schizophrenia", *Nature,* vol. 455, no. 7210, pp. 237-241.

Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., et al 2012, "De novo gene disruptions in children on the autistic spectrum", *Neuron,* vol. 74, no. 2, pp. 285-299.

Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., et al 2012, "De novo gene disruptions in children on the autistic spectrum", *Neuron,* vol. 74, no. 2, pp. 285-299.

Jamain, S., Quach, H., Betancur, C., Rastam, M., Colineaux, C., Gillberg, I.C., et al 2003, "Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism", *Nature genetics,* vol. 34, no. 1, pp. 27-29.

Kakinuma, H. & Sato, H. 2008, "Copy-number variations associated with autism spectrum disorder", *Pharmacogenomics,* vol. 9, no. 8, pp. 1143-1154.

Kanner, L. 1943, "Autistic disturbances of affective contact", *Nervous Child,* vol. 2, pp. 217-250.

Kanold, P.O. 2009, "Subplate neurons: crucial regulators of cortical development and plasticity", *Frontiers in neuroanatomy,* vol. 3, pp. 16.

Kaufman, L., Ayub, M. & Vincent, J.B. 2010, "The genetic basis of non-syndromic intellectual disability: a review", *Journal of neurodevelopmental disorders,* vol. 2, no. 4, pp. 182-209.

Keller, K., Williams, C., Wharton, P., Paulk, M., Bent-Williams, A., Gray, B., et al 2003, "Routine cytogenetic and FISH studies for 17p11/15q11 duplications and subtelomeric rearrangement studies in children with autism spectrum disorders", *American journal of medical genetics.Part A,* vol. 117A, no. 2, pp. 105-111.

Kerin, T., Ramanathan, A., Rivas, K., Grepo, N., Coetzee, G.A. & Campbell, D.B. 2012, "A noncoding RNA antisense to moesin at 5p14.1 in autism", *Science translational medicine,* vol. 4, no. 128, pp. 128ra40.

Khouzam, H.R., El-Gabalawi, F., Pirwani, N. & Priest, F. 2004, "Asperger's disorder: a review of its diagnosis and treatment", *Comprehensive psychiatry,* vol. 45, no. 3, pp. 184-191.

Kidd, J.M., Cooper, G.M., Donahue, W.F., Hayden, H.S., Sampas, N., Graves, T., et al 2008, "Mapping and sequencing of structural variation from eight human genomes", *Nature,* vol. 453, no. 7191, pp. 56-64.

Kilpinen, H., Ylisaukko-Oja, T., Hennah, W., Palo, O.M., Varilo, T., Vanhala, R., et al 2008, "Association of DISC1 with autism and Asperger syndrome", *Molecular psychiatry,* vol. 13, no. 2, pp. 187-196.

Kilpinen, H., Ylisaukko-oja, T., Rehnstrom, K., Gaal, E., Turunen, J.A., Kempas, E., et al 2009, "Linkage and linkage disequilibrium scan for autism loci in an extended pedigree from Finland", *Human molecular genetics,* vol. 18, no. 15, pp. 2912-2921.

Kim, S.J., Young, L.J., Gonen, D., Veenstra-VanderWeele, J., Courchesne, R., Courchesne, E., et al 2002, "Transmission disequilibrium testing of arginine vasopressin receptor 1A (AVPR1A) polymorphisms in autism", *Molecular psychiatry,* vol. 7, no. 5, pp. 503-507.

Kim, S.Y., Yasuda, S., Tanaka, H., Yamagata, K. & Kim, H. 2011, "Non-clustered protocadherin", *Cell adhesion & migration,* vol. 5, no. 2, pp. 97-105.

King, C.R. 2011, "A novel embryological theory of autism causation involving endogenous biochemicals capable of initiating cellular gene transcription: a possible link between twelve autism risk factors and the autism 'epidemic'", *Medical hypotheses,* vol. 76, no. 5, pp. 653-660.

Kinney, D.K., Barch, D.H., Chayka, B., Napoleon, S. & Munir, K.M. 2010, "Environmental risk factors for autism: do they help cause de novo genetic mutations that contribute to the disorder?", *Medical hypotheses,* vol. 74, no. 1, pp. 102-106.

Kleppe, K., Ohtsuka, E., Kleppe, R., Molineux, I. & Khorana, H.G. 1971, "Studies on polynucleotides. XCVI. Repair replications of short synthetic DNA's as catalyzed by DNA polymerases.", *J Mol Biol,* vol. 56, pp. 341-361.

Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., et al 2012, "Rate of de novo mutations and the importance of father's age to disease risk", *Nature,* vol. 488, no. 7412, pp. 471-475.

Kong, A., Gudbjartsson, D.F., Sainz, J., Jonsdottir, G.M., Gudjonsson, S.A., Richardsson, B., et al 2002, "A high-resolution recombination map of the human genome", *Nature genetics,* vol. 31, no. 3, pp. 241-247.

Kripke, D.F., Nievergelt, C.M., Joo, E., Shekhtman, T. & Kelsoe, J.R. 2009, "Circadian polymorphisms associated with affective disorders", *Journal of circadian rhythms,* vol. 7, pp. 2-3391-7-2.

Kristiansson, K., Perola, M., Tikkanen, E., Kettunen, J., Surakka, I., Havulinna, A.S., et al 2012, "Genome-wide screen for metabolic syndrome susceptibility Loci reveals strong lipid gene contribution but no evidence for common genetic basis for clustering of metabolic syndrome traits", *Circulation.Cardiovascular genetics,* vol. 5, no. 2, pp. 242-249.

Kumar, R.A., KaraMohamed, S., Sudi, J., Conrad, D.F., Brune, C., Badner, J.A., et al 2008, "Recurrent 16p11.2 microdeletions in autism", *Human molecular genetics,* vol. 17, no. 4, pp. 628-638.

Kuwano, Y., Kamio, Y., Kawai, T., Katsuura, S., Inada, N., Takaki, A., et al 2011, "Autism-associated gene expression in peripheral leucocytes commonly observed between subjects with autism and healthy women having autistic children", *PloS one,* vol. 6, no. 9, pp. e24723.

Lampi, K.M., Sourander, A., Gissler, M., Niemela, S., Rehnstrom, K., Pulkkinen, E., et al 2010, "Brief report: validity of Finnish registry-based diagnoses of autism with the ADI-R", *Acta Paediatrica (Oslo, Norway : 1992),* vol. 99, no. 9, pp. 1425-1428.

Lander, E.S. 2011, "Initial impact of the sequencing of the human genome", *Nature,* vol. 470, no. 7333, pp. 187-197.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., et al 2001, "Initial sequencing and analysis of the human genome", *Nature,* vol. 409, no. 6822, pp. 860-921.

Lathrop, G.M., Lalouel, J.M., Julier, C. & Ott, J. 1984, "Strategies for multilocus linkage analysis in humans", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 81, no. 11, pp. 3443-3446.

Lathrop, G.M., Lalouel, J.M. & White, R.L. 1986, "Construction of human linkage maps: likelihood calculations for multilocus linkage analysis", *Genetic epidemiology,* vol. 3, no. 1, pp. 39-52.

Laumonnier, F., Bonnet-Brilhault, F., Gomot, M., Blanc, R., David, A., Moizard, M.P., et al 2004, "X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family", *American Journal of Human Genetics,* vol. 74, no. 3, pp. 552-557.

Lauritsen, M.B., Pedersen, C.B. & Mortensen, P.B. 2005, "Effects of familial risk factors and place of birth on the risk of autism: a nationwide register-based study", *Journal of child psychology and psychiatry, and allied disciplines,* vol. 46, no. 9, pp. 963-971.

Le Meur, N., Holder-Espinasse, M., Jaillard, S., Goldenberg, A., Joriot, S., Amati-Bonneau, P., et al 2010, "MEF2C haploinsufficiency caused by either microdeletion of the 5q14.3 region or mutation is responsible for severe mental retardation with stereotypic movements, epilepsy and/or cerebral malformations", *Journal of medical genetics,* vol. 47, no. 1, pp. 22-29.

Lee, C. & Scherer, S.W. 2010, "The clinical context of copy number variation in the human genome", *Expert reviews in molecular medicine,* vol. 12, pp. e8.

Le-Niculescu, H., Patel, S.D., Bhat, M., Kuczenski, R., Faraone, S.V., Tsuang, M.T., et al 2009, "Convergent functional genomics of genome-wide association data for bipolar disorder: comprehensive identification of candidate genes, pathways and mechanisms", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 150B, no. 2, pp. 155-181.

Levy, D., Ronemus, M., Yamrom, B., Lee, Y.H., Leotta, A., Kendall, J., et al 2011, "Rare de novo and transmitted copy-number variation in autistic spectrum disorders", *Neuron,* vol. 70, no. 5, pp. 886-897.

Leyfer, O.T., Folstein, S.E., Bacalman, S., Davis, N.O., Dinh, E., Morgan, J., et al 2006, "Comorbid psychiatric disorders in children with autism: interview development and rates of disorders", *Journal of Autism and Developmental Disorders,* vol. 36, no. 7, pp. 849-861.

Li, H., Radford, J.C., Ragusa, M.J., Shea, K.L., McKercher, S.R., Zaremba, J.D., et al 2008, "Transcription factor MEF2C influences neural stem/progenitor cell differentiation and maturation in vivo", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 105, no. 27, pp. 9397-9402.

Li, X., Zou, H. & Brown, W.T. 2012, "Genes associated with autism spectrum disorder", *Brain research bulletin,* vol. 88, no. 6, pp. 543-552.

Liu, J., Nyholt, D.R., Magnussen, P., Parano, E., Pavone, P., Geschwind, D., et al 2001, "A genomewide screen for autism susceptibility loci", *American Journal of Human Genetics,* vol. 69, no. 2, pp. 327-340.

Loirat, C., Bellanne-Chantelot, C., Husson, I., Deschenes, G., Guigonis, V. & Chabane, N. 2010, "Autism in three patients with cystic or hyperechogenic kidneys and chromosome 17q12 deletion", *Nephrology, dialysis,*

transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association, vol. 25, no. 10, pp. 3430-3433.

Lord, C., Rutter, M. & Le Couteur, A. 1994, "Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders", *Journal of Autism and Developmental Disorders,* vol. 24, no. 5, pp. 659-685.

Lupski, J.R. 2007, "Genomic rearrangements and sporadic disease", *Nature genetics,* vol. 39, no. 7 Suppl, pp. S43-7.

Lyons, M.R. & West, A.E. 2011, "Mechanisms of specificity in neuronal activity-regulated gene transcription", *Progress in neurobiology,* vol. 94, no. 3, pp. 259-295.

Manning-Courtney, P., Murray, D., Currans, K., Johnson, H., Bing, N., Kroeger-Geoppinger, K., et al 2013, "Autism spectrum disorders", *Current problems in pediatric and adolescent health care,* vol. 43, no. 1, pp. 2-11.

Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., et al 2008, "Structural variation of chromosomes in autism spectrum disorder", *American Journal of Human Genetics,* vol. 82, no. 2, pp. 477-488.

Martin, C.L., Duvall, J.A., Ilkin, Y., Simon, J.S., Arreaza, M.G., Wilkes, K., et al 2007, "Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 144B, no. 7, pp. 869-876.

Mattila, M.L., Hurtig, T., Haapsamo, H., Jussila, K., Kuusikko-Gauffin, S., Kielinen, M., et al 2010, "Comorbid psychiatric disorders associated with Asperger syndrome/high-functioning autism: a community- and clinic-based study", *Journal of Autism and Developmental Disorders,* vol. 40, no. 9, pp. 1080-1093.

Mattila, M.L., Kielinen, M., Linna, S.L., Jussila, K., Ebeling, H., Bloigu, R., et al 2011, "Autism spectrum disorders according to DSM-IV-TR and comparison with DSM-5 draft criteria: an epidemiological study", *Journal of the American Academy of Child and Adolescent Psychiatry,* vol. 50, no. 6, pp. 583-592.e11.

Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., et al 2003, "TRANSFAC: transcriptional regulation, from patterns to profiles", *Nucleic acids research,* vol. 31, no. 1, pp. 374-378.

Mazumdar, S., Liu, K.Y., Susser, E. & Bearman, P. 2012, "The disappearing seasonality of autism conceptions in California", *PloS one,* vol. 7, no. 7, pp. e41265.

McAndrew, P.E., Frostholm, A., White, R.A., Rotter, A. & Burghes, A.H. 1998, "Identification and characterization of RPTP rho, a novel RPTP mu/kappa-like receptor protein tyrosine phosphatase whose expression is restricted to the central nervous system", *Brain research.Molecular brain research,* vol. 56, no. 1-2, pp. 9-21.

McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P., et al 2008, "Genome-wide association studies for complex traits: consensus, uncertainty and challenges", *Nature reviews.Genetics,* vol. 9, no. 5, pp. 356-369.

Mednick, S.A., Huttunen, M.O. & Machon, R.A. 1994, "Prenatal influenza infections and adult schizophrenia", *Schizophrenia bulletin,* vol. 20, no. 2, pp. 263-267.

Meechan, D.W., Tucker, E.S., Maynard, T.M. & LaMantia, A.S. 2009, "Diminished dosage of 22q11 genes disrupts neurogenesis and cortical development in a mouse model of 22q11 deletion/DiGeorge syndrome", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 106, no. 38, pp. 16434-16445.

Meyer-Lindenberg, A., Kolachana, B., Gold, B., Olsh, A., Nicodemus, K.K., Mattay, V., et al 2009, "Genetic variants in AVPR1A linked to autism predict amygdala activation and personality traits in healthy humans", *Molecular psychiatry,* vol. 14, no. 10, pp. 968-975.

Millar, J.K., Wilson-Annan, J.C., Anderson, S., Christie, S., Taylor, M.S., Semple, C.A., et al 2000, "Disruption of two novel genes by a translocation co-segregating with schizophrenia", *Human molecular genetics,* vol. 9, no. 9, pp. 1415-1423.

Miller, D.T., Adam, M.P., Aradhya, S., Biesecker, L.G., Brothman, A.R., Carter, N.P., et al 2010, "Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies", *American Journal of Human Genetics,* vol. 86, no. 5, pp. 749-764.

Miller, D.T., Shen, Y., Weiss, L.A., Korn, J., Anselm, I., Bridgemohan, C., et al 2009, "Microdeletion/duplication at 15q13.2q13.3 among individuals with features of autism and other neuropsychiatric disorders", *Journal of medical genetics,* vol. 46, no. 4, pp. 242-248.

Moessner, R., Marshall, C.R., Sutcliffe, J.S., Skaug, J., Pinto, D., Vincent, J., et al 2007, "Contribution of SHANK3 mutations to autism spectrum disorder", *American Journal of Human Genetics,* vol. 81, no. 6, pp. 1289-1297.

Moreno-De-Luca, D., SGENE Consortium, Mulle, J.G., Simons Simplex Collection Genetics Consortium, Kaminsky, E.B., Sanders, S.J., et al 2010, "Deletion 17q12 is a recurrent copy number variant that confers high risk of autism and schizophrenia", *American Journal of Human Genetics,* vol. 87, no. 5, pp. 618-630.

Nachman, M.W. & Crowell, S.L. 2000, "Estimate of the mutation rate per nucleotide in humans", *Genetics,* vol. 156, no. 1, pp. 297-304.

Nagase, T., Ishikawa, K., Suyama, M., Kikuno, R., Hirosawa, M., Miyajima, N., et al 1998, "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro", *DNA research : an international journal for rapid publication of reports on genes and genomes,* vol. 5, no. 6, pp. 355-364.

Najmabadi, H., Hu, H., Garshasbi, M., Zemojtel, T., Abedini, S.S., Chen, W., et al 2011, "Deep sequencing reveals 50 novel genes for recessive cognitive disorders", *Nature,* vol. 478, no. 7367, pp. 57-63.

Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., et al 2012, "Patterns and rates of exonic de novo mutations in autism spectrum disorders", *Nature,* vol. 485, no. 7397, pp. 242-245.

Newschaffer, C.J., Croen, L.A., Daniels, J., Giarelli, E., Grether, J.K., Levy, S.E., et al 2007, "The epidemiology of autism spectrum disorders", *Annual Review of Public Health,* vol. 28, pp. 235-258.

Newschaffer, C.J., Fallin, D. & Lee, N.L. 2002, "Heritable and nonheritable risk factors for autism spectrum disorders", *Epidemiologic reviews,* vol. 24, no. 2, pp. 137-153.

Ng, P.C. & Henikoff, S. 2003, "SIFT: Predicting amino acid changes that affect protein function", *Nucleic acids research,* vol. 31, no. 13, pp. 3812-3814.

Nguyen, D.Q., Webber, C. & Ponting, C.P. 2006, "Bias of selection on human copy-number variants", *PLoS genetics,* vol. 2, no. 2, pp. e20.

Niklasson, L., Rasmussen, P., Oskarsdottir, S. & Gillberg, C. 2009, "Autism, ADHD, mental retardation and behavior problems in 100 individuals with 22q11 deletion syndrome", *Research in developmental disabilities,* vol. 30, no. 4, pp. 763-773.

Niklasson, L., Rasmussen, P., Oskarsdottir, S. & Gillberg, C. 2002, "Chromosome 22q11 deletion syndrome (CATCH 22): neuropsychiatric and neuropsychological aspects", *Developmental medicine and child neurology,* vol. 44, no. 1, pp. 44-50.

Nord, A.S., Roeb, W., Dickel, D.E., Walsh, T., Kusenda, M., O'Connor, K.L., et al 2011, "Reduced transcript expression of genes affected by inherited and de novo CNVs in autism", *European journal of human genetics : EJHG,* vol. 19, no. 6, pp. 727-731.

Novara, F., Beri, S., Giorda, R., Ortibus, E., Nageshappa, S., Darra, F., et al 2010, "Refining the phenotype associated with MEF2C haploinsufficiency", *Clinical genetics,* vol. 78, no. 5, pp. 471-477.

O'Connell, J.R. & Weeks, D.E. 1998, "PedCheck: a program for identification of genotype incompatibilities in linkage analysis", *Am J Hum Genet,* vol. 63, no. 1, pp. 259-66.

O'Roak, B.J., Vives, L., Fu, W., Egertson, J.D., Stanaway, I.B., Phelps, I.G., et al 2012b, "Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders", *Science (New York, N.Y.),* vol. 338, no. 6114, pp. 1619-1622.

O'Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P., et al 2012a, "Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations", *Nature,* vol. 485, no. 7397, pp. 246-250.

Ott, J. 1986, "Linkage probability and its approximate confidence interval under possible heterogeneity", *Genetic epidemiology.Supplement,* vol. 1, pp. 251-257.

Ozawa, S., Kamiya, H. & Tsuzuki, K. 1998, "Glutamate receptors in the mammalian central nervous system", *Progress in neurobiology,* vol. 54, no. 5, pp. 581-618.

Ozgen, H.M., van Daalen, E., Bolton, P.F., Maloney, V.K., Huang, S., Cresswell, L., et al 2009, "Copy number changes of the microcephalin 1 gene (MCPH1) in patients with autism spectrum disorders", *Clinical genetics,* vol. 76, no. 4, pp. 348-356.

Ozonoff, S. 2011, "Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study", *Pediatrics,* vol. 128, no. 3, pp. e488.

Ozonoff, S., Heung, K., Byrd, R., Hansen, R. & Hertz-Picciotto, I. 2008, "The onset of autism: patterns of symptom emergence in the first years of life", *Autism research : official journal of the International Society for Autism Research,* vol. 1, no. 6, pp. 320-328.

Papanikolaou, K., Paliokosta, E., Gyftodimou, J., Kolaitis, G., Vgenopoulou, S., Sarri, C., et al 2006, "A case of partial trisomy of chromosome 8p associated with autism", *Journal of Autism and Developmental Disorders,* vol. 36, no. 5, pp. 705-709.

Parker, S.K. 2004, "Thimerosal-containing vaccines and autistic spectrum disorder: a critical review of published original data", *Pediatrics,* vol. 114, no. 3, pp. 793.

Perala, J., Suvisaari, J., Saarni, S.I., Kuoppasalmi, K., Isometsa, E., Pirkola, S., et al 2007, "Lifetime prevalence of psychotic and bipolar I disorders in a general population", *Archives of General Psychiatry,* vol. 64, no. 1, pp. 19-28.

Persico, A.M. & Bourgeron, T. 2006, "Searching for ways out of the autism maze: genetic, epigenetic and environmental clues", *Trends in neurosciences,* vol. 29, no. 7, pp. 349-358.

Pfeiffer, B.E., Zang, T., Wilkerson, J.R., Taniguchi, M., Maksimova, M.A., Smith, L.N., et al 2010, "Fragile X mental retardation protein is required for synapse elimination by the activity-dependent transcription factor MEF2", *Neuron,* vol. 66, no. 2, pp. 191-197.

Pinto, D., Pagnamenta, A.T., Klei, L., Anney, R., Merico, D., Regan, R., et al 2010, "Functional impact of global rare copy number variation in autism spectrum disorders", *Nature,* vol. 466, no. 7304, pp. 368-372.

Pirkola, S.P., Isometsa, E., Suvisaari, J., Aro, H., Joukamaa, M., Poikolainen, K., et al 2005, "DSM-IV mood-, anxiety- and alcohol use disorders and their comorbidity in the Finnish general population--results from the Health 2000 Study", *Social psychiatry and psychiatric epidemiology,* vol. 40, no. 1, pp. 1-10.

Piton, A., Gauthier, J., Hamdan, F.F., Lafreniere, R.G., Yang, Y., Henrion, E., et al 2011, "Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia", *Molecular psychiatry,* vol. 16, no. 8, pp. 867-880.

Piton, A., Michaud, J.L., Peng, H., Aradhya, S., Gauthier, J., Mottron, L., et al 2008, "Mutations in the calcium-related gene IL1RAPL1 are associated with autism", *Human molecular genetics,* vol. 17, no. 24, pp. 3965-3974.

Piven, J., Chase, G.A., Landa, R., Wzorek, M., Gayle, J., Cloud, D., et al 1991, "Psychiatric disorders in the parents of autistic individuals", *Journal of the American Academy of Child and Adolescent Psychiatry,* vol. 30, no. 3, pp. 471-478.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., et al 2007, "PLINK: a tool set for whole-genome association and population-based linkage analyses", *American Journal of Human Genetics,* vol. 81, no. 3, pp. 559-575.

Ramensky, V. 2002, "Human non-synonymous SNPs: server and survey", *Nucleic acids research,* vol. 30, no. 17, pp. 3894.

Redon, R. 2006, "Global variation in copy number in the human genome", *Nature,* vol. 444, no. 7118, pp. 444.

Rehnstrom, K., Ylisaukko-oja, T., Nieminen-von Wendt, T., Sarenius, S., Kallman, T., Kempas, E., et al 2006, "Independent replication and initial fine mapping of 3p21-24 in Asperger syndrome", *Journal of medical genetics,* vol. 43, no. 2, pp. e6.

Reichow, B., Barton, E.E., Boyd, B.A. & Hume, K. (eds) 2012, *Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD) (Review)*, The Cochrane Collaboration edn, JohnWiley & Sons, Ltd., The Cochrane Library.

Renieri, A., Pescucci, C., Longo, I., Ariani, F., Mari, F. & Meloni, I. 2005, "Non-syndromic X-linked mental retardation: from a molecular to a clinical point of view.", *Journal of cellular physiology,* [Online], vol. 204, no. 1, pp. 8.

Richler, J. 2006, "Is there a 'regressive phenotype'of Autism Spectrum Disorder associated with the measles-mumps-rubella vaccine? A CPEA Study", *Journal of Autism and Developmental Disorders,* vol. 36, no. 3, pp. 299.

Roelfsema, M.T., Hoekstra, R.A., Allison, C., Wheelwright, S., Brayne, C., Matthews, F.E., et al 2012, "Are autism spectrum conditions more prevalent in an information-technology region? A school-based study of three regions in the Netherlands", *Journal of Autism and Developmental Disorders,* vol. 42, no. 5, pp. 734-739.

Roelfsema, M.T., Hoekstra, R.A., Allison, C., Wheelwright, S., Brayne, C., Matthews, F.E., et al 2012, "Are autism spectrum conditions more prevalent in an information-technology region? A school-based study of three regions in the Netherlands", *Journal of Autism and Developmental Disorders,* vol. 42, no. 5, pp. 734-739.

Rosenberg, D.R., MacMaster, F.P., Keshavan, M.S., Fitzgerald, K.D., Stewart, C.M. & Moore, G.J. 2000, "Decrease in caudate glutamatergic concentrations in pediatric obsessive-compulsive disorder patients taking paroxetine", *Journal of the American Academy of Child and Adolescent Psychiatry,* vol. 39, no. 9, pp. 1096-1103.

Rosenberg, D.R., Mirza, Y., Russell, A., Tang, J., Smith, J.M., Banerjee, S.P., et al 2004, "Reduced anterior cingulate glutamatergic concentrations in childhood OCD and major depression versus healthy controls", *Journal of the American Academy of Child and Adolescent Psychiatry,* vol. 43, no. 9, pp. 1146-1153.

Rozen, S. & Skaletsky, H. 2000, "Primer3 on the WWW for general users and for biologist programmers", *Methods in molecular biology (Clifton, N.J.),* vol. 132, pp. 365-386.

Russell, A.J., Mataix-Cols, D., Anson, M. & Murphy, D.G. 2005, "Obsessions and compulsions in Asperger syndrome and high-functioning autism", *The British journal of psychiatry : the journal of mental science,* vol. 186, pp. 525-528.

Russo, S., Marchi, M., Cogliati, F., Bonati, M.T., Pintaudi, M., Veneselli, E., et al 2009, "Novel mutations in the CDKL5 gene, predicted effects and associated phenotypes", *Neurogenetics,* vol. 10, no. 3, pp. 241-250.

Sahoo, T., Shaw, C.A., Young, A.S., Whitehouse, N.L., Schroer, R.J., Stevenson, R.E., et al 2005, "Array-based comparative genomic hybridization analysis of recurrent chromosome 15q rearrangements", *American journal of medical genetics.Part A,* vol. 139A, no. 2, pp. 106-113.

Salyakina, D., Cukier, H.N., Lee, J.M., Sacharow, S., Nations, L.D., Ma, D., et al 2011, "Copy number variants in extended autism spectrum disorder families reveal candidates potentially involved in autism risk", *PloS one,* vol. 6, no. 10, pp. e26049.

Samuels, J., Wang, Y., Riddle, M.A., Greenberg, B.D., Fyer, A.J., McCracken, J.T., et al 2011, "Comprehensive family-based association study of the glutamate transporter gene SLC1A1 in obsessive-compulsive disorder", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 156B, no. 4, pp. 472-477.

Sanders, S.J., Ercan-Sencicek, A.G., Hus, V., Luo, R., Murtha, M.T., Moreno-De-Luca, D., et al 2011, "Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism", *Neuron,* vol. 70, no. 5, pp. 863-885.

Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., et al 2012, "De novo mutations revealed by whole-exome sequencing are strongly associated with autism", *Nature,* vol. 485, no. 7397, pp. 237-241.

Sanger, F. & Goulson, A.R. 1975, "A Rapid Method for Determining Sequences in DNA by Primed Synthesis with DNA Polymerase", *J Mol Biol,* vol. 94, pp. 441-448.

Scahill, L. & SCAHILL 2006, "Children's Yale-Brown Obsessive Compulsive Scale modified for pervasive developmental disorders", *Journal of the American Academy of Child Adolescent Psychiatry,* vol. 45, no. 9, pp. 1114.

Scahill, L. & SCAHILL 1997, "Children's Yale-Brown Obsessive Compulsive Scale: reliability and validity.", *Journal of the American Academy of Child Adolescent Psychiatry,* vol. 36, no. 6, pp. 844.

Schellenberg, G.D., Dawson, G., Sung, Y.J., Estes, A., Munson, J., Rosenthal, E., et al 2006, "Evidence for multiple loci from a genome scan of autism kindreds", *Molecular psychiatry,* vol. 11, no. 11, pp. 1049-60, 979.

Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium 2011, "Genome-wide association study identifies five new schizophrenia loci", *Nature genetics,* vol. 43, no. 10, pp. 969-976.

Schopler, E., Reichler, R., DeVellis, R. & Daly, K. 1980, "Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS)", *March,* vol. 10, no. 1, pp. 91-103.

Schormair, B., Kemlink, D., Roeske, D., Eckstein, G., Xiong, L., Lichtner, P., et al 2008, "PTPRD (protein tyrosine phosphatase receptor type delta) is associated with restless legs syndrome", *Nature genetics,* vol. 40, no. 8, pp. 946-948.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., et al 2007, "Strong association of de novo copy number mutations with autism", *Science (New York, N.Y.),* vol. 316, no. 5823, pp. 445-449.

Shao, Y., Wolpert, C.M., Raiford, K.L., Menold, M.M., Donnelly, S.L., Ravan, S.A., et al 2002, "Genomic screen and follow-up analysis for autistic disorder", *Am J Med Genet,* vol. 114, no. 1, pp. 99-105.

Sharp, A.J., Mefford, H.C., Li, K., Baker, C., Skinner, C., Stevenson, R.E., et al 2008, "A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures", *Nature genetics,* vol. 40, no. 3, pp. 322-328.

Shugart, Y.Y., Wang, Y., Samuels, J.F., Grados, M.A., Greenberg, B.D., Knowles, J.A., et al 2009, "A family-based association study of the glutamate transporter gene SLC1A1 in obsessive-compulsive disorder in 378 families", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 150B, no. 6, pp. 886-892.

Silva, A.E., Vayego-Lourenco, S.A., Fett-Conte, A.C., Goloni-Bertollo, E.M. & Varella-Garcia, M. 2002, "Tetrasomy 15q11-q13 identified by fluorescence in situ hybridization in a patient with autistic disorder", *Arquivos de Neuro-Psiquiatria,* vol. 60, no. 2-A, pp. 290-294.

Skuse, D.H. & Gallagher, L. 2011, "Genetic influences on social cognition", *Pediatric research,* vol. 69, no. 5 Pt 2, pp. 85R-91R.

Smalley, S.L., McCracken, J. & Tanguay, P. 1995, "Autism, affective disorders, and social phobia", *American Journal of Medical Genetics,* vol. 60, no. 1, pp. 19-26.

Speevak, M.D. & Farrell, S.A. 2011, "Non-syndromic language delay in a child with disruption in the Protocadherin11X/Y gene pair", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 156B, no. 4, pp. 484-489.

Spiker, M.A. 2012, "Restricted interests and anxiety in children with autism", *Autism,* vol. 16, no. 3, pp. 306.

Stankiewicz, P. & Lupski, J.R. 2010, "Structural variation in the human genome and its role in disease", *Annual Review of Medicine,* vol. 61, pp. 437-455.

Stefanatos, G.A. 2008, "Regression in autistic spectrum disorders", *Neuropsychology review,* vol. 18, no. 4, pp. 305-319.

Stefansson, H., Rujescu, D., Cichon, S., Pietilainen, O.P., Ingason, A., Steinberg, S., et al 2008, "Large recurrent microdeletions associated with schizophrenia", *Nature,* vol. 455, no. 7210, pp. 232-236.

Steffenburg, S., Gillberg, C., Hellgren, L., Andersson, L., Gillberg, I.C., Jakobsson, G., et al 1989, "A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden", *J Child Psychol Psychiatry,* vol. 30, no. 3, pp. 405-16.

Stewart, S.E., Fagerness, J.A., Platko, J., Smoller, J.W., Scharf, J.M., Illmann, C., et al 2007, "Association of the SLC1A1 glutamate transporter gene and obsessive-compulsive disorder", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 144B, no. 8, pp. 1027-1033.

Stewart, S.E., Yu, D., Scharf, J.M., Neale, B.M., Fagerness, J.A., Mathews, C.A., et al 2012, "Genome-wide association study of obsessive-compulsive disorder", *Molecular psychiatry,* .

Stone, J.L., Merriman, B., Cantor, R.M., Yonan, A.L., Gilliam, T.C., Geschwind, D.H., et al 2004, "Evidence for sex-specific risk alleles in autism spectrum disorder", *American Journal of Human Genetics,* vol. 75, no. 6, pp. 1117-1123.

Styczynski, M.P., Jensen, K.L., Rigoutsos, I. & Stephanopoulos, G. 2008, "BLOSUM62 miscalculations improve search performance", *Nature biotechnology,* vol. 26, no. 3, pp. 274-275.

Südhof, T.C. 2008, "Neuroligins and neurexins link synaptic function to cognitive disease", *Nature,* vol. 455, no. 7215, pp. 903.

Szatmari, P., Paterson, A.D., Zwaigenbaum, L., Roberts, W., Brian, J., Liu, X., et al 2007, "Mapping autism risk loci using genetic linkage and chromosomal rearrangements", *Nature genetics,* vol. 39, no. 3, pp. 319-328.

Tansey, K., Hill, M., Cochrane, L., Gill, M., Anney, R. & Gallagher, L. 2011, "Functionality of promoter microsatellites of arginine vasopressin receptor 1A (AVPR1A): implications for autism", *Mol Autism,* vol. 2, no. 3.

Tarpey, P., Parnau, J., Blow, M., Woffendin, H., Bignell, G., Cox, C., et al 2004, "Mutations in the DLG3 gene cause nonsyndromic X-linked mental retardation", *American Journal of Human Genetics,* vol. 75, no. 2, pp. 318-324.

Tarpey, P.S., Smith, R., Pleasance, E., Whibley, A., Edkins, S., Hardy, C., et al 2009, "A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation", *Nature genetics,* vol. 41, no. 5, pp. 535-543.

Theisen, A. 2008, "Microarray-based Comparative Genomic Hybridization (aCGH)", *Nature Education,* vol. 1, no. 1.

Toro, R., Konyukh, M., Delorme, R., Leblond, C., Chaste, P., Fauchereau, F., et al 2010, "Key role for gene dosage and synaptic homeostasis in autism spectrum disorders", *Trends in genetics : TIG,* vol. 26, no. 8, pp. 363-372.

Trikalinos, T.A., Karvouni, A., Zintzaras, E., Ylisaukko-oja, T., Peltonen, L., Jarvela, I., et al 2006, "A heterogeneity-based genome search meta-analysis for autism-spectrum disorders", *Molecular psychiatry,* vol. 11, no. 1, pp. 29-36.

Ukkola Vuoti, L. 2011, "Association of the arginine vasopressin receptor 1A (AVPR1A) haplotypes with listening to music", *Journal of human genetics,* vol. 56, no. 4, pp. 324.

Ukkola, L.T. 2009, "Musical aptitude is associated with AVPR1A-haplotypes", *PLoS ONE,* vol. 4, no. 5, pp. e5534.

van der Zwaag, B., Staal, W.G., Hochstenbach, R., Poot, M., Spierenburg, H.A., de Jonge, M.V., et al 2010, "A co-segregating microduplication of chromosome 15q11.2 pinpoints two risk genes for autism spectrum disorder", *American journal of medical genetics.Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics,* vol. 153B, no. 4, pp. 960-966.

Vandenplas, S., Wiid, I., Grobler-Rabie, A., Brebner, K., Ricketts, M., Wallis, G., et al 1984, "Blot hybridisation analysis of genomic DNA", *Journal of medical genetics,* vol. 21, no. 3, pp. 164-172.

Veltman, J.A. & Brunner, H.G. 2012, "De novo mutations in human genetic disease", *Nature reviews.Genetics,* vol. 13, no. 8, pp. 565-575.

Veltman, M.W., Thompson, R.J., Roberts, S.E., Thomas, N.S., Whittington, J. & Bolton, P.F. 2004, "Prader-Willi syndrome--a study comparing deletion and uniparental disomy cases with reference to autism spectrum disorders", *European child & adolescent psychiatry,* vol. 13, no. 1, pp. 42-50.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.S., et al 2001, "The sequence of the human genome", *Science,* vol. 291, no. 5507, pp. 1304-1351.

Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., et al 2011, "Transcriptomic analysis of autistic brain reveals convergent molecular pathology", *Nature,* vol. 474, no. 7351, pp. 380-384.

Volk, H.E., Lurmann, F., Penfold, B., Hertz-Picciotto, I. & McConnell, R. 2013, "Traffic-Related Air Pollution, Particulate Matter, and Autism ", *JAMA, Psychiatry,* vol. 70, no. 1, pp. 71-77.

Vorstman, J.A., Breetvelt, E.J., Thode, K.I., Chow, E.W. & Bassett, A.S. 2013, "Expression of autism spectrum and schizophrenia in patients with a 22q11.2 deletion", *Schizophrenia research,* vol. 143, no. 1, pp. 55-59.

Vorstman, J.A., Morcus, M.E., Duijff, S.N., Klaassen, P.W., Heineman-de Boer, J.A., Beemer, F.A., et al 2006, "The 22q11.2 deletion in children: high rate of autistic disorders and early onset of psychotic symptoms", *Journal of the American Academy of Child and Adolescent Psychiatry,* vol. 45, no. 9, pp. 1104-1113.

Vorstman, J.A. & Ophoff, R.A. 2013, "Genetic causes of developmental disorders", *Current opinion in neurology.*

Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., et al 2007, "PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data", *Genome research,* vol. 17, no. 11, pp. 1665-1674.

Wang, K., Zhang, H., Ma, D., Bucan, M., Glessner, J.T., Abrahams, B.S., et al 2009, "Common genetic variants on 5p14.1 associate with autism spectrum disorders", *Nature,* vol. 459, no. 7246, pp. 528-533.

Wassink, T.H., Losh, M., Piven, J., Sheffield, V.C., Ashley, E., Westin, E.R., et al 2007, "Systematic screening for subtelomeric anomalies in a clinical sample of autism", *Journal of Autism and Developmental Disorders,* vol. 37, no. 4, pp. 703-708.

Wassink, T.H., Piven, J. & Patil, S.R. 2001, "Chromosomal abnormalities in a clinic sample of individuals with autistic disorder", *Psychiatric genetics,* vol. 11, no. 2, pp. 57-63.

Wassink, T.H., Piven, J., Vieland, V.J., Pietila, J., Goedken, R.J., Folstein, S.E., et al 2004, "Examination of AVPR1a as an autism susceptibility gene", *Molecular psychiatry,* vol. 9, no. 10, pp. 968-972.

Weiss, L.A., Arking, D.E., Gene Discovery Project of Johns Hopkins & the Autism Consortium, Daly, M.J. & Chakravarti, A. 2009, "A genome-wide linkage and association scan reveals novel loci for autism", *Nature,* vol. 461, no. 7265, pp. 802-808.

Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., et al 2008, "Association between microdeletion and microduplication at 16p11.2 and autism", *The New England journal of medicine,* vol. 358, no. 7, pp. 667-675.

Wendland, J.R., Moya, P.R., Timpano, K.R., Anavitarte, A.P., Kruse, M.R., Wheaton, M.G., et al 2009, "A haplotype containing quantitative trait loci for SLC1A1 gene expression and its association with obsessive-compulsive disorder", *Archives of General Psychiatry,* vol. 66, no. 4, pp. 408-416.

Willour, V.L., Yao Shugart, Y., Samuels, J., Grados, M., Cullen, B., Bienvenu, O.J.,3rd, et al 2004, "Replication study supports evidence for linkage to 9p24 in obsessive-compulsive disorder", *American Journal of Human Genetics,* vol. 75, no. 3, pp. 508-513.

Wing, L. 1981, "Asperger's syndrome: a clinical account.", *Psychol Med,* vol. 11, pp. 115-129.

Wing, L., Yeates, S.R., Brierley, L.M. & Gould, J. 1976, "The prevalence of early childhood autism: comparison of administrative and epidemiological studies.", *Psychol Med,* vol. 6, no. 1, pp. 89-100.

Wintle, R.F., Lionel, A.C., Hu, P., Ginsberg, S.D., Pinto, D., Thiruvahindrapduram, B., et al 2011, "A genotype resource for postmortem brain samples from the Autism Tissue Program", *Autism research : official journal of the International Society for Autism Research,* vol. 4, no. 2, pp. 89-97.

Wolpert, C.M., Menold, M.M., Bass, M.P., Qumsiyeh, M.B., Donnelly, S.L., Ravan, S.A., et al 2000, "Three probands with autistic disorder and isodicentric chromosome 15", *American Journal of Medical Genetics,* vol. 96, no. 3, pp. 365-372.

World Health Organization 1993, *The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic Criteria for Research.* WHO, Geneva.

Xu, B., Roos, J.L., Levy, S., van Rensburg, E.J., Gogos, J.A. & Karayiorgou, M. 2008, "Strong association of de novo copy number mutations with sporadic schizophrenia", *Nature genetics,* vol. 40, no. 7, pp. 880-885.

Xu, L.M., Li, J.R., Huang, Y., Zhao, M., Tang, X. & Wei, L. 2012, "AutismKB: an evidence-based knowledgebase of autism genetics", *Nucleic acids research,* vol. 40, no. Database issue, pp. D1016-22.

Yamada, K., Gerber, D.J., Iwayama, Y., Ohnishi, T., Ohba, H., Toyota, T., et al 2007, "Genetic analysis of the calcineurin pathway identifies members of the EGR gene family, specifically EGR3, as potential susceptibility candidates in schizophrenia", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 104, no. 8, pp. 2815-2820.

Yang, S.Y., Cho, S.C., Yoo, H.J., Cho, I.H., Park, M., Kim, B.N., et al 2010b, "Association study between single nucleotide polymorphisms in promoter region of AVPR1A and Korean autism spectrum disorders", *Neuroscience letters,* vol. 479, no. 3, pp. 197-200.

Yang, S.Y., Cho, S.C., Yoo, H.J., Cho, I.H., Park, M., Yoe, J., et al 2010a, "Family-based association study of microsatellites in the 5' flanking region of AVPR1A with autism spectrum disorder in the Korean population", *Psychiatry research,* vol. 178, no. 1, pp. 199-201.

Yirmiya, N., Rosenberg, C., Levi, S., Salomon, S., Shulman, C., Nemanov, L., et al 2006, "Association between the arginine vasopressin 1a receptor (AVPR1a) gene and autism in a family-based study: mediation by socialization skills", *Molecular psychiatry,* vol. 11, no. 5, pp. 488-494.

Ylisaukko-oja, T., Nieminen-von Wendt, T., Kempas, E., Sarenius, S., Varilo, T., von Wendt, L., et al 2004, "Genome-wide scan for loci of Asperger syndrome", *Mol Psychiatry,* vol. 9, no. 2, pp. 161-8.

Ylisaukko-oja, T., Rehnstrom, K., Auranen, M., Vanhala, R., Alen, R., Kempas, E., et al 2005, "Analysis of four neuroligin genes as candidates for autism", *Eur J Hum Genet,* vol. 13, no. 12, pp. 1285-92.

Yntema, H.G., van den Helm, B., Kissing, J., van Duijnhoven, G., Poppelaars, F., Chelly, J., Moraine, C., Fryns, J.P., Hamel, B.C., Heilbronner, H., Pander, H.J., Brunner, H.G., Ropers, H.H., Cremers, F.P. & van Bokhoven, H. 1999, "A novel ribosomal S6-kinase (RSK4 RPS6KA6) is commonly deleted in patients with complex X-linked mental retardation.", *Genomics,* [Online], vol. 62, no. 3, pp. 332.

Yonan, A.L., Alarcon, M., Cheng, R., Magnusson, P.K., Spence, S.J., Palmer, A.A., et al 2003, "A genomewide screen of 345 families for autism-susceptibility loci", *Am J Hum Genet,* vol. 73, no. 4, pp. 886-97.

Yoshida, T., Yasumura, M., Uemura, T., Lee, S.J., Ra, M., Taguchi, R., et al 2011, "IL-1 receptor accessory protein-like 1 associated with mental retardation and autism mediates synapse formation by trans-synaptic interaction with protein tyrosine phosphatase delta", *The Journal of neuroscience : the official journal of the Society for Neuroscience,* vol. 31, no. 38, pp. 13485-13499.

Yu, L.M. & Goda, Y. 2009, "Dendritic signalling and homeostatic adaptation", *Current opinion in neurobiology,* vol. 19, no. 3, pp. 327-335.

Zandt, F. 2007, "Repetitive behaviour in children with high functioning autism and obsessive compulsive disorder", *Journal of Autism and Developmental Disorders,* vol. 37, no. 2, pp. 251.

Zanni, G., van Esch, H., Bensalem, A., Saillour, Y., Poirier, K., Castelnau, L., et al 2010, "A novel mutation in the DLG3 gene encoding the synapse-associated protein 102 (SAP102) causes non-syndromic mental retardation", *Neurogenetics,* vol. 11, no. 2, pp. 251-255.

Zerbo, O. 2011, "Month of conception and risk of autism", *Epidemiology,* vol. 22, no. 4, pp. 469.

Zhang, R., Lu, S., Meng, L., Min, Z., Tian, J., Valenzuela, R.K., et al 2012, "Genetic evidence for the association between the early growth response 3 (EGR3) gene and schizophrenia", *PloS one,* vol. 7, no. 1, pp. e30237.

Zweier, M. & Rauch, A. 2012, "The MEF2C-Related and 5q14.3q15 Microdeletion Syndrome", *Molecular syndromology,* vol. 2, no. 3-5, pp. 164-170.