9

Genome-Wide Association Studies of Copy Number Variation in Autism Spectrum Disorder

Hae-Jin Hu and Yeun-Jun Chung
Integrated Research Center for Genome Polymorphism, The Catholic University of Korea, College of Medicine, Seoul
Korea

1. Introduction

Autism is a syndrome with a broad spectrum of phenotypes characterized by deficits in social interaction and communication, repetitive or stereotyped behaviors, and restricted interests (Rutter, 2005). Autism spectrum disorder (ASD) manifests mostly before 3 years of age (Klauck, 2006). ASDs include two related diagnoses; pervasive developmental disorder (PDD) including atypical autism, impairment in the same areas, but not meeting criteria for autism; Asperger's syndrome, which is milder than PDD, showing similar impairments in social interaction, behaviors and interests, but no significant delay in linguistic and cognitive development (Weiss, 2009). The prevalence rate of ASDs is ~0.6% and ASDs are approximately four times more common in males than in females (Veenstra-VanderWeele, 2004). Many studies have been performed to elucidate the pathogenesis of ASDs, but identified risk factors do not explain a significant proportion of the disease prevalence.

Genetic epidemiological data have been suggesting that ASDs are heritable both in autism families and in the general population (Freitag, 2007). The concordance rates of autism in monozygotic twins were reported to be significantly higher (~ 60–90%) than those in dizygotic twins (~ 10%) and the recurrence rates are known to be approximately 10-20 times higher in siblings than in normal population (Folstein & Rosen-Sheidley, 2001; Cohen et al., 2005; Bailey et al., 1995; Lauritsen et al., 2005). ASD is not a single-gene disorder with Mendelian inheritance but rather a component of various genetic disorders with apparent cytogenetic abnormalities (Eapen, 2011). Cytogenetic alterations were detected in 7.4% of ASD (Vorstman et al., 2006), and some of them have been suggested as causative factors of neurodevelopmental disorders (Merikangas et al., 2009). However, discrepancies in study results and diverse modes of inheritance have hindered the discovery of common genetic susceptibility factors to ASDs. For these reasons, despite the growing evidence supporting the genetic susceptibility to ASD development (Folstein & Rosen-Sheidley, 2001; Veenstra-VanderWeele & Cook, 2004), the genetic mechanisms of ASD is still largely unknown.

Recent technical advance in microarray-based whole-genome analysis has enabled identification of common and rare genetic alterations associated with ASDs. Several recent studies have suggested that ASDs are associated with genetic variations including single nucleotide polymorphisms (SNPs) and copy number variations (CNVs), and that these genetic variations may work together (Veenstra-VanderWeele & Cook, 2004). For example, de novo CNVs were found in ~7% of idiopathic ASD families via oligoarray-comparative

genomic hybridization and whole-genome SNP array analysis (Abrahams & Geschwind, 2008; Psychiatric GWAS Consortium Coordinating Committee et al., 2009). In addition to rare de novo variations, common genetic variations such as SNPs on 5p14.1 were found to be associated with ASDs and this finding was replicated in two independent studies (Wang et al., 2009). Recently introduced next-generation sequencing (NGS) will further accelerate mining of genetic variations linked with ASDs (Ropers, 2010). Graphical overview of the reported ASD-associated CNVs and SNP are illustrated in the Figure 1. In this chapter, we will review the recent results of CNV and SNP genome-wide association studies (GWAS) on ASDs and discuss the perspectives of the genetic susceptibility study of ASDs.

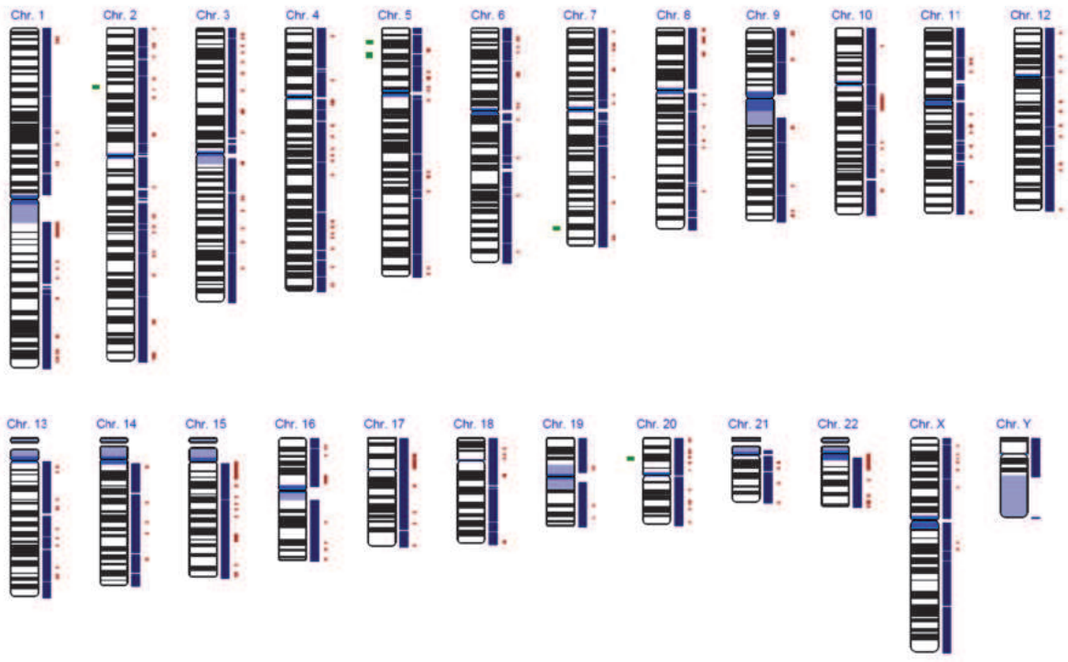


Fig. 1. Genomic map of CNVs and SNPs associated with ASDs identified by GWAS. Green and red bars on the left and right side of the karyograms indicate chromosomal locations of SNPs and CNVs, respectively. Blue bars on the right side of the karyograms present the locations of known genes. This figure was drawn by IdeogramBrowser (Müller et al, 2007).

2. CNVs associated with ASD

2.1 What is CNV?

Using array-CGH, a combination of microarray and comparative genomic hybridization (CGH) technologies, two pioneering groups of scientists have identified wide-spread CNVs in apparently healthy, normal individuals in 2004 (Iafrate et al., 2004; Sebat et al., 2004). CNV is defined as any type of genetic variant that alters the chromosomal structure, including duplications and deletions (Iafrate et al., 2004; Sebat et al., 2004; Redon et al., 2006) and now known to be one of the most prevalent types of genetic variations in the human genome (Feuk et al., 2006; Hurles et al., 2008; Carter, 2007; Estivill & Armengol, 2007). In

addition to SNPs, CNVs in normal individuals have been widening our understanding of genetic heterogeneity (Iafrate et al., 2004; Sebat et al., 2004; Redon et al., 2006). Commonly used working definition of CNV was a copy number change involving a DNA segment sized 1 kilobases (kb) or larger (Freeman et al., 2006; Feuk et al., 2006). Nowadays, definition of CNVs includes any DNA structural variants including duplications, deletions and inversions (Hurles et al., 2008). When the frequency of CNV is common (>1%) in the population, CNV is also called copy number polymorphism (CNP). However, due to lack of standardized technologies to define CNV, the size and frequency of CNV have not been well defined in human populations. Since the two pioneering studies discovered the evidence of the existence of CNVs (Iafrate et al., 2004; Sebat et al., 2004), more than 66,000 CNVs and 34,000 InDels have been identified in various populations (Redon et al., 2006; Simon-Sanchez et al., 2007; de Smith et al., 2007; Perry et al., 2008; Díaz de Ståhl et al., 2008; Yim et al., 2010; Conrad et al., 2010; Park et al., 2010) and catalogued in the public database, Database of Genomic Variants (http://projects.tcag.ca/variation/) (Feuk et al, 2006). More CNVs have been uncovered using the NGS analysis (Mills et al., 2011; Kidd et al., 2010; Kim et al., 2009).

CNVs can affect gene functions in several ways and have a potential to affect gene expression levels presumably larger than that of SNPs. Deletion or duplication may disrupt the genes located inside those regions, resulting in changes in the gene structure, which can affect the gene expression. Alternatively, disruption of the transcription regulatory regions and the enhancers can also affect the gene expression. During the recombination which is thought to be an important mechanism of CNV development, novel fusion products may be generated, which may exert positive or negative effects on gene expression and epigenetic regulations (Feuk et al, 2006; Zhang et al., 2009; Hampton et al., 2009; Przybytkowski et al., 2011; Reymond et al., 2007). Taken together, structural variations are likely to be responsible for the phenotypic variation of human beings and comprehensive mapping of CNVs can facilitate the understanding of inter-individual phenotypic differences including disease susceptibility and responsiveness to drugs (Feuk et al, 2006; Estivill & Armengol, 2007). Indeed, CNVs have been found to be associated with various types of Mendelian traits and also a substantial number of complex diseases including neurodevelopmental disorders (Buchanan & Scherer, 2008; Lee & Lupski, 2006).

2.2 CNVs in ASD

To assess the role of CNV in ASD, several different whole-genome microarray platforms based on oligonucleotides, SNPs and BAC clones have been used for ASD family studies or case-control studies (Abrahams & Geschwind, 2008; Cook & Scherer, 2008). As a result, lines of evidence have been accumulated that multiple rare de novo CNVs contribute to the susceptibility to ASD. For example, duplications and/or deletions on chromosome 15q11–q13 confer increased risk of ASD (15q11–q13 duplication syndrome, Prader-Willi syndrome and Angelman syndrome). Approximately one fourth of the individuals who have a 22q11.2 deletion and over 90% of individuals with duplication of 17p11.2 show characteristics of ASD (Cohen et al., 2005; Abrahams & Geschwind, 2008; Fernández et al., 2009). Significant associations have been reported between ASD and CNV of various genes, such as *NRXN1* (2p16.3), *NLGN3* (Xq13.1), *NLGN4* (Xp22.23) and *SHANK3* (22q13.3). There have been many reports on CNVs associated with ASDs, but, due to technical limitations and lack of standardized methods for defining the CNVs and CNV regions (CNVRs), there are inconsistencies among studies which should be removed by further GWAS. Table 1 summarizes the major CNVs identified by GWAS in ASD.

Discovery Sample		Replication Sample		Study design	Platform	CNV Detection method	Number of CNVs identified	Strong candidate loci	CN change	Reference
Case	Control	Case	Control							
1496 families with 7,917 subjects	Unaffected family members	-	-	Family-based	Affymetrix 10K	dChip	254	NRXN1 1q21 17p12 22q11.2	del	Szatmari et al., 2007
165 families	99 unaffected families	-	-	Family-based	Agilent 244K 390K ROMA	HMM	17	SLC4A10, FHIT FHIT FLJ16237 A2BP1	del del dup del del	Sebat et al., 2007
180	372	532	465	Family-based and Case-control	Array-CGH	NimbleGen	1	16p11.2	microdel	Kumar et al., 2008
751 multiplex families with 1441 cases	1420 (AGRE parents) 2814 (bipolar disporder or NIMH controls)	512 (CHB) 299 (deCODE)	434 (CHB) 18,834 (decode)	Family-based	Affymetrix 5.0 (AGRE) Affymetrix 500K (controls)	COPPER/ Birdseye (AGRE) ADM-2(CHB) HMM(deCODE)	47	16p11.2	del/dup	Weiss et al., 2008
397	372	-	-	Family-based	19K BAC Microarray	-	51	15q11-q13 22q11 16p11.2	dup dup microdel	Christian et al., 2008
427 families	500	-	1,152 matched controls	Case-control	Affymetrix 500K	dChip, CNAG, GEMCA	277	16p11.2 SHANK3-NLGN4-NRXN1-PSD DPP6-DPP10-PCDH9 ANKRD11 DPYD PTCHD1 15q24	del/dup	Marshall et al., 2008
859	1409	1,336	1,110	Case-control	Illumina HumanHap550	PennCNV	78,490	15q11-13 22q11.21 NRXN1 CNTN4 PARK2 RFWD2 AK123120 UNQ3037 GRID1 NLGN1 GYPELOC44	dup dup del del/dup del dup dup del del dup dup	Glessner et al., 2009
912 multiplex families	1,488 (CHOP) 542 (NINDS)	859	1,051	Case-control	Illumina HumanHap550	PennCNV	> 150	NRXN1 UBE3A 15q11-q13 BZRAP1 MDGA2	del dup del/dup del/dup del	Bucan et al, 2009

Discovery Sample		Replication Sample		Study design	Platform	CNV Detection method	Number of CNVs identified	Strong candidate loci	CN change	Reference
Case	Control	Case	Control							
28 children	62 Adults	-	-	Case-control	Array-CGH	Array-CyGHt	38	8p23.1 17p11.2	del del	Cho et al, 2009
996	1,287	-	3,677	Case-control and Family-based	Illumina 1M	QuantiSNP iPattern	5,478	SHANK2 SYNGAP1 DLGAP2 CSNK1D/SLC16A3 NRXN1 22q11.21 DDX53/PTCHD1	del del dup dup/del dup/del del del	Pinto rt al., 2010

ACC:Autism Case-Control cohort
ADM : aberration detection method
AGP: Autism Genome Project
AGRE: Autism Genetic Resource Exchange
CHB: Children's Hospital Boston
CHOP: Children's Hospital of Philadelphia
CNAG: Copy Number Analysis for GeneChip
COPPER: copy-number polymorphism evaluation routine
dChip: DNA Chip Analyzer
GEMCA: Genotyping Microarray based CNV Analysis
HMM: hidden Markov model
NIMH: National Institute of Mental Health
NINDS: National Institute of Neurological Disorders and Stroke

Table 1. Genome-wide CNV association studies of autism

In 2007, two pioneering studies demonstrated the association of CNVs with ASD. The Autism Genome Project Consortium performed linkage and CNV analyses using Affymetrix 10K SNP array for 1,181 ASD families with at least two affected individuals (Autism Genome Project Consortium et al., 2007). Of the 254 highly significant CNVs, the investigators emphasized four CNVs and the most interesting finding was a 300-kb sized CNV loss on chromosome 2p16 identified recurrently in two families. The deletion of this region disrupted the coding exons of the neurexin 1 gene (*NRXN1*), which interacts with neurologins and involves in synaptogenesis. Therefore, deterioration of the neurexin 1 function by deletion may affect susceptibility to ASD or its phenotypes. The structural variation in the *NRXN1* gene was reported from the previous autism studies (Chubykin et al., 2005; Feng et al., 2006). The other three interesting CNVs were 1.1-Mb sized CNV gain on chromosome 1q21, 933-kb sized de novo duplication on 17p12, and duplication on 22q11.2. The duplication on 17p12 is known to cause Charcot-Marie-Tooth 1A (CMT1A) disease (Houlden et al., 2006). In addition, other micro-duplications of the same chromosomal region have been reported in individuals with mental retardation, linguistic delay, autism and related phenotypes (Moog et al., 2004).

Sebat and his colleagues performed array-CGH analysis with 264 families and explored the association of de novo CNVs with ASD, which are not present in their respective parents (Sebat et al., 2007). The authors identified 17 de novo CNVs in 16 subjects. According to their result, the frequency of spontaneous mutation was 10% in the sporadic cases and 3% in the multiplex families, while 1% in unaffected individuals. One of the de novo CNV loci was a 4.3-Mb sized deletion at 22q13.31-q13.33, where *SHANK3* gene is located. Recurrent deletion of this

region has been previously reported in ASD (Manning et al., 2004). Durand et al. reported that mutations in *SHANK3* gene were associated with ASD and abnormal gene dosage of *SHANK3* was associated with severe cognitive deficits, linguistic delay and ASD (Durand et al., 2007). SHANK3 is a scaffolding protein found in excitatory synapses directly opposite to the presynaptic active zone. This gene has been suggested to be associated with the neurobehavioral symptoms observed in individuals with 22q13 deletions.

In 2008, four independent studies consistently reported the association of the CNV on 16p11.2 locus with autism. Weiss et al. adopted Affymetrix 5.0 SNP array to find CNVs in 751 multiplex families from the Autism Genetic Resource Exchange (AGRE) (Weiss et al., 2008). They identified 32 high- and 15 low- confidence regions. Among the candidate loci, microdeletion and microduplication on 16p11.2 were validated to be associated with ASD. This association was further confirmed in clinical testing data from Children's Hospital Boston and in a large population data from Iceland (deCODE genetics data). Kumar et al. screened 180 ASD cases and 372 controls using a 19K whole-genome tiling bacterial artificial chromosome (BAC) array to identify submicroscopic copy number changes specific to autism (Kumar et al., 2008). They observed ~500-kb sized recurrent microdeletion on 16p11.2 in two cases with autism but not in the controls. When they assessed the frequency of this putative autism-associated genomic disorder, 0.6% of the ASD cases showed the alterations while none in controls. The variation was confirmed by FISH, microsatellite analyses and array-CGH. Christian et al. also used the same 19K whole-genome tiling BAC array to identify ASD-associated CNVs in the 397 cases and 372 control set (Christian et al., 2008). Among the 51 candidate CNVs, recurrent CNVs were identified in the loci including 15q11-q13, 22q11, and 16p11.2. They were confirmed by FISH, microsatellite analysis, or quantitative polymerase chain reaction (PCR) analysis. Marshall et al. performed whole-genome screening for 427 ASD cases and 500 controls using Affymetrix 500K SNP arrays (Marshall et al., 2008). Of the 277 CNVs identified only in the cases, the CNVs on 16p11.2 locus appeared in around 1% of the ASD cases, which included both duplications and deletions. There exist SHANK3-NLGN4-NRXN1 postsynaptic density genes, DPP6-DPP10-PCDH9 (synapse complex), ANKRD11, DPYD and PTCHD1 in other associated CNVs.

New CNVs in addition to the known ones have been suggested to be associated with ASD in the subsequent studies. Glessner et al. performed a whole-genome CNV analysis with 859 cases and 1,409 controls using Illumina HumanHap550 BeadChip (Glessner et al., 2009). They generated 78,490 CNV calls and the positive findings were further evaluated in an independent cohort of 1,336 ASD cases and 1,110 controls. Through this approach, they identified several known ASD-associated genes as well as novel candidate CNVs. For example, they identified the CNVs in the loci including 15q11–q13, 22q11.21, *NRXN1* and *CNTN4*, which were previously reported to be associated with autism (Kim et al., 2009; Roohi et al., 2009; Fernandez et al., 2008). However, some of the genes or loci previously known to be associated with ASD such as A*UTS2* (Kalscheuer et al., 2007), *NLGN3* (Jamain et al., 2003), *SHANK3* (Moessner et al., 2007) and 16p11.2(Weiss et al., 2008) were not replicated in their study. Especially 16p11.2, a locus consistently reported to be associated in four previous independent studies, did not show a significant association in this study. Several new susceptibility genes such as *NLGN1* and *ASTN2* were identified in this study. Both genes encode neuronal cell-adhesion molecules. In Chubykin et al's report, mutations in neuroligin superfamily members were identified in the individuals with ASD (Chubykin et al., 2005). ASTN1 is a neuronal protein receptor integral in the process of glial-guided granule cell migration during development (Zheng et al., 1996). Furthermore, CNVs of the

genes involved in the ubiquitin pathways, such as *UBE3A*, *PARK2*, *RFWD2* and *FBXO40*, were observed in the ASD cases but not in the controls. Bucan et al. conducted high -density genotyping of 912 multiplex families from the AGRE collection and 1,488 controls using Illumina HumanHap550 BeadChip (Bucan et al., 2009). They identified more than 150 loci harboring rare variants in multiple unrelated patients and the positive findings were further validated in an independent cohort of 859 ASD cases and 1,051 controls by genomic quantitative PCR. Among the candidate loci, there are previously reported ones such as *NRXN1*(Marshall et al., 2008), *UBE3A* (Glessner et al., 2009), and 15q11-q13 (Christian et al., 2008) and novel ones such as *BZRAP1* and *MDGA2*.

In 2009, Cho et al. reported the ASD associated CNVs in east-Asians. They performed whole-genome BAC array-CGH with 28 ASD cases and with 62 controls and identified 38 CNVs including those harboring two significant loci, 8p23.1 and 17p11.2 (Cho et al., 2009). *DEFENSIN* gene family are located in the 8p23.1 CNV locus and often showed copy number polymorphisms in earlier studies (Linzmeier & Ganz, 2005). Although there have been no direct clues to connect the copy number loss of *DEFENSIN* gene and ASD, immunological dysfunction has been suggested to be associated with autism (Rutter, 2005).

Most recently, Pinto et al. analyzed the genome-wide features of rare CNVs in autism using Illumina 1M SNP arrays (Pinto et al., 2010). Based on 996 cases and 1,287 controls, they identified 5,478 rare CNVs. By examining parent-child transmission, the authors found the 226 de novo and inherited CNVs which were not present in controls. As a whole, ASD cases were found to carry a higher number of de novo CNVs than controls (1.69 fold, P=3.4X10^{-4}). A number of novel genes such as *SHANK2*, *SYNGAP1*, *DLGAP2* and the *DDX53–PTCHD1* in the CNVs were found to be associated with ASD in this study. Also, through gene set enrichment analysis, cellular proliferation, projection and motility, and GTPase/Ras signaling were found to be affected by the CNVs identified in their study. This approach demonstrated the new paradigm of autism research based on functional pathway and cross-talk.

3. SNPs in autism

Before the establishment of GWAS, the genome-wide linkage analysis has been used for the discovery of the mutations in diverse diseases (OMIM http://www.ncbi.nlm.nih.gov/omim). Location of the disease genes were successfully narrowed down by linkage disequilibrium mapping studies, but linkage approach was not always successful especially for complex diseases. In many cases, the significant linkage loci were not replicated. One potential reason is that the effect of a single risk variant on the pathogenesis of complex disease might be too small to be detected. Small genetic effects could be detected with greater power by association analyses such as GWAS with large case-control population (Risch & Merikangas, 1996). In other words, to identify common risk alleles in the common complex diseases, population-wide analysis with more common and dense variants is required. SNP-GWAS can be an ideal approach for unbiased screening and also be adopted for high-density linkage analysis. SNP-GWAS became a matured technology for exploring novel associations between genetic variants and complex diseases because over 12 million of SNAs have been catalogued and high density array fabrication/analysis technologies have been developed. In neuropsychiatric disorders with unknown etiology such as ASD, SNP-GWAS have been actively adopted to explore the genetic background of the diseases (Table 2). In this chapter, we will review the major SNP-GWAS results for ASDs.

Initial Sample Size	Replication Sample Size	Region	Reported Gene(s)	Strongest SNP	p-value of discovery set (OR*)	p-value of replication set (OR*)	Platform	Reference
72 families (148 cases)	1,295 trios	7q35	CNTNAP2	rs7794745	2.14E-05	<0.005	Affymetrix 500K	Arking et al., 2008
1,031 families (1,553 cases)	2,073 trios	5p15	SEMA5A, TAS2R1	rs10513025	1.7E-06 (0.55)	6.1E-03 (0.76)	Affymetrix 5.0 /500K	Weiss et al., 2009
438 families	487 families	5p14.1	Intergenic	rs10038113	2.75E-05 (0.67)	3.28E-03	Illumina Human 1Mv1 beadchip	Ma et al., 2009
780 families (1,299 cases)	447 families	5p14.1	CDH10,CDH9	rs4307059	1.1E-05	1.2E-2	550K/1M Illumina	Wang et al., 2009
1,204 cases 6,491 controls	108 cases 540 controls				2.2E-04 (1.19)	1.6E-2	550K/300K Illumina	
745 boys in social group 870 boys in nonsocial group	1,400 boys	2p21	Intergenic (social traits)	rs11894053 (social traits)	0.02	-	Affymetrix 500K	Ronald et al., 2010
1,558 families	2,179 families	20p12.1	MACROD2 (Str|Eur)**	rs4141463 (Str|Eur)	2.1E-08 (0.56)	4.7E-08 (0.65)	1M Illumina	Anney et al., 2010

*OR: Odds Ratio
**Str|Eur : strict diagnosis and European ancestry

Table 2. Genome-wide SNP association studies of autism

3.1 Common SNPs associated with ASD

Arking et al. performed a two-stage study on ASD using genome-wide linkage and family-based association mapping by whole-genome SNP genotyping (Arking et al., 2008). For stage I, they selected 72 multiplex ASD families and genotyped the samples using Affymetrix 500K arrays. In this approach, they could not find any significant SNPs or haplotypes. However, through the genome-wide linkage analysis, they identified 2 significant loci associated with ASD, 7q35 and 10p13-14. In the most significant locus (7q35), they identified that a polymorphism in contactin-associated protein-like 2 (*CNTNAP2*) gene, a member of the neurexin superfamily, is associated with ASD. In the second stage, they validated the significant findings of the stage I by examining 145 multiplex families and confirmed that *CNTNAP2* was an autism-susceptibility gene. This result was the first evidence that a common genetic variant in the neurexin superfamily member increases risk of autism.

In 2009, Weiss et al. explored a linkage and SNP association analysis with 1,031 multiplex autism families using Affymetrix 5.0 SNP array (Weiss et al., 2009). They found that a SNP on 5p15 locus between *SEMA5A* and *TAS2R1* gene was significantly associated with autism. In addition, the expression of SEMA5A was found to decrease in brains of autistic patients. Taken together the authors suggested a possibility of *SEMA5A* as an autism risk gene. Wang et al. used higher density SNA array and larger study populations to identify common genetic risk factors underlying ASDs (Wang et al., 2009). They used two different sets of study subjects in discovery stage using Illumina Human 1M beadchip. First set was 780 families with 1,299 affected children and the second set was 1,204 patients and 6,491

controls. They identified six significant SNPs located between cadherin 10 (*CDH10*) and cadherin 9 (*CDH9*) genes strongly associated with ASD. These two genes encode neuronal cell-adhesion molecules. Among the 6 SNPs, the most significant one was rs4307059 ($P = 3.4 \times 10^{-8}$, odds ratio = 1.19). The SNP was replicated in two independent datasets of 447 families and 108 case-540 control sets. Combined analysis using all four datasets showed that all six SNPs are associated with autism (*P* values ranging from 7.4×10^{-8} to 2.1×10^{-10}). Interestingly, 5p14.1 was consistently suggested as a novel risk locus in an independent study in the same year. Ma et al. performed GWAS with 438 Caucasian autistic families using Illumina Human 1M beadchip (Ma et al., 2009). They found that 96 SNPs were strongly associated with autism ($P < 0.0001$). They validated all 96 significant associations in independent samples of 487 families using 550K Illumina BeadChip, which was the same array platform to Wang et al's. A novel locus on 5p14.1 was found to be significantly associated with autism both in the discovery and validation dataset.

The Autism Genome Project (AGP) Consortium performed high-resolution genotyping with 1,558 families to identify significant SNPs (Anney et al., 2010). For primary analysis, they partitioned the dataset along axes of diagnosis and ancestry; spectrum versus strict; European versus all ancestries. Based on these partitioned data, they conducted four GWAS. They observed the strongest association for SNP rs4141463 in one of the four primary association analyses. Located within MACROD2, this marker crossed the GWA significance threshold of $P < 5 \times 10^{-8}$. They are performing analysis of combining data to validate the results of the primary analysis.

Despite the expected advantages of large-scale GWAS analysis, none of the candidate associations have been replicated so far, which may underscore the genetic and phenotypic heterogeneity of ASD and indicate the fact that the effect size of common alleles contributing to common disorders is much smaller than expectation (Eapen, 2011).

3.2 Rare SNPs associated with ASD

Definition of a rare variant is a variant with frequency <1%. The most deleterious variants might be naturally eliminated during evolution, but some remain as rare variants. In 'common disease–common variant' model, most of the rare SNP associations have been missed by current GWAS concept. However, rapid development of NGS will facilitate the discovery of rare variants. Rare SNP associations are more likely to be detected by re-sequencing of relevant regions in hundreds or thousands of individuals. It is anticipated that advances in re-sequencing technologies will make it feasible to search systematically for rare variant effects.

4. Conclusion

Human Genome Project has provided insight into a complete sequence of the haploid human genome and we also have got new insight of the human genetic variations. Based on this new insight, conventional target gene oriented and hypothesis-driven research design has been quickly moved to a new paradigm, hypothesis-free mining of novel disease associated genes. Indeed, over hundreds of genetic variants which may affect the susceptibility of pathogenesis of complex disease have been identified by the GWAS. The GWAS have been actively adopted in studying the causative factors of neurodevelopmental disorders including ASD. Through GWAS approach, several robust ASD-associated variants

in the genes such as *NRXN1, SHANK3, NLGN4* and *CNTNAP2* were uncovered. However, it is too early to say that GWAS have brought reliable-enough insight into ASD. Many of the significant CNVs identified in one study were not consistent or not successfully replicated in the following studies. New improved algorithms for CNV and CNVR will be needed for defining the CNVs more robustly. To sort out the platform to platform variation of CNV call, which is one of the obstacles for the meta-analysis, more reliable experimental methods should be developed. Re-sequencing large number of individuals without CNVs will help to discover the new rare variants. Considering the speed of technological innovations including algorithm, high-throughput analysis and NGS, we anticipate that current obstacles of GWAS in autism research will be removed soon. However, GWAS result itself will not be enough to get clinically applicable insight about the pathophysiology of ASD. Integration of GWAS data with other resources such as improved bio-imaging, personal whole-genome sequencing, gene-environmental interaction and metagenome analysis data about gastrointestinal commensal bacteria will enable us to get a more comprehensive insight in designing future personalized care of autism.

5. Useful website for ASD related data

- Psychiatric GWAS Consortium(PGC): https://pgc.unc.edu/index.php
- National Institute of Mental Health Center for Collaborative Genetic Studies on Mental Disorders : http://nimhgenetics.org/
- Autism Genetic Resource Exchange (AGRE): http://www.agre.org/
- Autism Genome Project (AGP): http://www.well.ox.ac.uk/monaco/autism/AGP.shtml
- The International Schizophrenia Consortium (ISC): http://pngu.mgh.harvard.edu/isc/
- Genetic Association Information Network(GAIN): http://www.genome.gov/19518664#al-4
- CNV project at the Children's Hospital of Philadelphia: http://cnv.chop.edu
- The SGENE project: http://www.sgene.eu/Summary.php
- Database of Genomic Variants (DGV): http://projects.tcag.ca/variation
- A Catalog of Published Genome-Wide Association Studies: http://www.genome.gov/gwastudies/index.cfm?#searchForm
- DECIPHER: https://decipher.sanger.ac.uk
- CNV project: http://www.sanger.ac.uk/humgen/cnv/
- GEN2PHEN: http://www.gen2phen.org/

6. Acknowledgements

Work in the authors' laboratory is supported by the Korea Health 21 R&D Project (A040002), Ministry of Health and Welfare, Republic of Korea.

7. References

[1] Abrahams BS, Geschwind DH. (2008). Advances in autism genetics: on the threshold of a new neurobiology. *Nat Rev Genet*. 9:341-355.

[2] Anney R, Klei L, Pinto D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Sykes N, Pagnamenta AT, Almeida J, Bacchelli E, Bailey AJ, Baird G, Battaglia A,

Berney T, Bolshakova N, Bölte S, Bolton PF, Bourgeron T, Brennan S, Brian J, Carson AR, Casallo G, Casey J, Chu SH, Cochrane L, Corsello C, Crawford EL, Crossett A, Dawson G, de Jonge M, Delorme R, Drmic I, Duketis E, Duque F, Estes A, Farrar P, Fernandez BA, Folstein SE, Fombonne E, Freitag CM, Gilbert J, Gillberg C, Glessner JT, Goldberg J, Green J, Guter SJ, Hakonarson H, Heron EA, Hill M, Holt R, Howe JL, Hughes G, Hus V, Igliozzi R, Kim C, Klauck SM, Kolevzon A, Korvatska O, Kustanovich V, Lajonchere CM, Lamb JA, Laskawiec M, Leboyer M, Le Couteur A, Leventhal BL, Lionel AC, Liu XQ, Lord C, Lotspeich L, Lund SC, Maestrini E, Mahoney W, Mantoulan C, Marshall CR, McConachie H, McDougle CJ, McGrath J, McMahon WM, Melhem NM, Merikangas A, Migita O, Minshew NJ, Mirza GK, Munson J, Nelson SF, Noakes C, Noor A, Nygren G, Oliveira G, Papanikolaou K, Parr JR, Parrini B, Paton T, Pickles A, Piven J, Posey DJ, Poustka A, Poustka F, Prasad A, Ragoussis J, Renshaw K, Rickaby J, Roberts W, Roeder K, Roge B, Rutter ML, Bierut LJ, Rice JP, Salt J, Sansom K, Sato D, Segurado R, Senman L, Shah N, Sheffield VC, Soorya L, Sousa I, Stoppioni V, Strawbridge C, Tancredi R, Tansey K, Thiruvahindrapduram B, Thompson AP, Thomson S, Tryfon A, Tsiantis J, Van Engeland H, Vincent JB, Volkmar F, Wallace S, Wang K, Wang Z, Wassink TH, Wing K, Wittemeyer K, Wood S, Yaspan BL, Zurawiecki D, Zwaigenbaum L, Betancur C, Buxbaum JD, Cantor RM, Cook EH, Coon H, Cuccaro ML, Gallagher L, Geschwind DH, Gill M, Haines JL, Miller J, Monaco AP, Nurnberger JI Jr, Paterson AD, Pericak-Vance MA, Schellenberg GD, Scherer SW, Sutcliffe JS, Szatmari P, Vicente AM, Vieland VJ, Wijsman EM, Devlin B, Ennis S, Hallmayer J. (2010). A genome-wide scan for common alleles affecting risk for autism. *Hum Mol Genet*. 19:4072-4082.

[3] Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, Rea A, Guy M, Lin S, Cook EH, Chakravarti A. (2008). A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *Am J Hum Genet*. 82:160-164.

[4] Autism Genome Project Consortium, Szatmari P, Paterson AD, Zwaigenbaum L, Roberts W, Brian J, Liu XQ, Vincent JB, Skaug JL, Thompson AP, Senman L, Feuk L, Qian C, Bryson SE, Jones MB, Marshall CR, Scherer SW, Vieland VJ, Bartlett C, Mangin LV, Goedken R, Segre A, Pericak-Vance MA, Cuccaro ML, Gilbert JR, Wright HH, Abramson RK, Betancur C, Bourgeron T, Gillberg C, Leboyer M, Buxbaum JD, Davis KL, Hollander E, Silverman JM, Hallmayer J, Lotspeich L, Sutcliffe JS, Haines JL, Folstein SE, Piven J, Wassink TH, Sheffield V, Geschwind DH, Bucan M, Brown WT, Cantor RM, Constantino JN, Gilliam TC, Herbert M, Lajonchere C, Ledbetter DH, Lese-Martin C, Miller J, Nelson S, Samango-Sprouse CA, Spence S, State M, Tanzi RE, Coon H, Dawson G, Devlin B, Estes A, Flodman P, Klei L, McMahon WM, Minshew N, Munson J, Korvatska E, Rodier PM, Schellenberg GD, Smith M, Spence MA, Stodgell C, Tepper PG, Wijsman EM, Yu CE, Rogé B, Mantoulan C, Wittemeyer K, Poustka A, Felder B, Klauck SM, Schuster C, Poustka F, Bölte S, Feineis-Matthews S, Herbrecht E, Schmötzer G, Tsiantis J, Papanikolaou K, Maestrini E, Bacchelli E, Blasi F, Carone S, Toma C, Van Engeland H, de Jonge M, Kemner C, Koop F, Langemeijer M, Hijmans C, Staal WG, Baird G, Bolton PF, Rutter ML, Weisblatt E, Green J, Aldred C, Wilkinson JA, Pickles A, Le Couteur A, Berney T, McConachie H, Bailey AJ, Francis K, Honeyman G, Hutchinson A, Parr

JR, Wallace S, Monaco AP, Barnby G, Kobayashi K, Lamb JA, Sousa I, Sykes N, Cook EH, Guter SJ, Leventhal BL, Salt J, Lord C, Corsello C, Hus V, Weeks DE, Volkmar F, Tauber M, Fombonne E, Shih A, Meyer KJ. (2007). Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 39:319-328.

[5] Bailey A, Le Couteur A, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol. Med* 25:63–77.

[6] Bucan M, Abrahams BS, Wang K, Glessner JT, Herman EI, Sonnenblick LI, Alvarez Retuerto AI, Imielinski M, Hadley D, Bradfield JP, Kim C, Gidaya NB, Lindquist I, Hutman T, Sigman M, Kustanovich V, Lajonchere CM, Singleton A, Kim J, Wassink TH, McMahon WM, Owley T, Sweeney JA, Coon H, Nurnberger JI, Li M, Cantor RM, Minshew NJ, Sutcliffe JS, Cook EH, Dawson G, Buxbaum JD, Grant SF, Schellenberg GD, Geschwind DH, Hakonarson H. (2009). Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genet*. 5:e1000536.

[7] Buchanan, J. A. & Scherer, S. W. (2008). Contemplating effects of genomic structural variation. *Genet. Med*. 10:639–647.

[8] Carter, N.P. (2007). Methods and strategies for analyzing copy-number variation using DNA microarrays. *Nat. Genet*. 39: S16–21

[9] Cho SC, Yim SH, Yoo HK, Kim MY, Jung GY, Shin GW, Kim BN, Hwang JW, Kang JJ, Kim TM, Chung YJ. (2009). Copy number variations associated with idiopathic autism identified by whole-genome microarray-based comparative genomic hybridization. *Psychiatr Genet*. 19:177-185.

[10] Christian SL, Brune CW, Sudi J, Kumar RA, Liu S, Karamohamed S, Badner JA, Matsui S, Conroy J, McQuaid D, Gergel J, Hatchwell E, Gilliam TC, Gershon ES, Nowak NJ, Dobyns WB, Cook EH Jr. (2008). Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol Psychiatry*. 63:1111-1117.

[11] Chubykin AA, Liu X, Comoletti D, Tsigelny I, Taylor P, Sudhof TC. (2005). Dissection of synapse induction by neuroligins: effect of a neuroligin mutation associated with autism. *J Biol Chem*. 280:22365-22374

[12] Cohen D., Pichard, N., Tordjman, S., Baumann, C., Burglen, L., Excoffier, E., Lazar, G., Mazet, P., Pinquier, C., Verloes, A., Héron, D. (2005). Specific genetic disorders and autism: clinical contribution towards their identification. *J Autism Dev Disord* 35:103–116.

[13] Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, Fitzgerald T, Hu M, Ihm CH, Kristiansson K, Macarthur DG, Macdonald JR, Onyiah I, Pang AW, Robson S, Stirrups K, Valsesia A, Walter K, Wei J; Wellcome Trust Case Control Consortium, Tyler-Smith C, Carter NP, Lee C, Scherer SW, Hurles ME. (2010). Origins and functional impact of copy number variation in the human genome. *Nature*. 464:704-712.

[14] Cook EH Jr, Scherer SW. (2008). Copy-number variations associated with neuropsychiatric conditions. *Nature*. 455:919-923

[15] de Smith AJ, Tsalenko A, Sampas N, Scheffer A, Yamada NA, Tsang P, Ben-Dor A, Yakhini Z, Ellis RJ, Bruhn L, Laderman S, Froguel P, Blakemore AI. (2007). Array CGH analysis of copy number variation identifies 1284 new genes variant in

healthy white males: implications for association studies of complex diseases. Hum. Mol. Genet. 16:2783–2794.

[16] Díaz de Ståhl T, Sandgren J, Piotrowski A, Nord H, Andersson R, Menzel U, Bogdan A, Thuresson AC, Poplawski A, von Tell D, Hansson CM, Elshafie AI, Elghazali G, Imreh S, Nordenskjöld M, Upadhyaya M, Komorowski J, Bruder CE, Dumanski JP. (2008). Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array. *Hum. Mutat*. 29:398–408.

[17] Durand CM, Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, Nygren G, Rastam M, Gillberg IC, Anckarsäter H, Sponheim E, Goubran-Botros H, Delorme R, Chabane N, Mouren-Simeoni MC, de Mas P, Bieth E, Rogé B, Héron D, Burglen L, Gillberg C, Leboyer M, Bourgeron T. (2007). Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet*. 39:25-27.

[18] Eapen, V. (2011). Genetic basis of autism: is there a way forward? *Curr Opin Psychiatry*. 24:226-236.

[19] Estivill, X. and Armengol, L. (2007). Copy number variants and common disorders: filling the gaps and exploring complexity in genome-wide association studies. *PLoS Genet*. 3:1787–1799.

[20] Feng J, Schroer R, Yan J, Song W, Yang C, Bockholt A, Cook EH Jr, Skinner C, Schwartz CE, Sommer SS. (2006). High frequency of neurexin 1 [beta] signal peptide structural variants in patients with autism. *Neuroscience letters*. 409:10-13.

[21] Fernández L, Nevado J, Santos F, Heine-Suñer D, Martinez-Glez V, García-Miñaur S, Palomo R, Delicado A, Pajares IL, Palomares M, García-Guereta L, Valverde E, Hawkins F, Lapunzina P. (2009). A deletion and a duplication in distal 22q11.2 deletion syndrome region. Clinical implications and review. *BMC Med Genet*. 10:48.

[22] Fernandez T, Morgan T, Davis N, Klin A, Morris A, Farhi A, Lifton RP, State MW. (2008). Disruption of Contactin 4 (CNTN4) results in developmental delay and other features of 3p deletion syndrome. *Am J Hum Genet*. 82:1385.

[23] Feuk, L., Carson, A.R., and Scherer, S.W. (2006). Structural variation in the human genome. *Nat. Rev. Genet*. 7: 85–97.

[24] Feuk, L., Marshall, C.R., Wintle, R.F. and Scherer, S.W. (2006). Structural variants: changing the landscape of chromosomes and design of disease studies. *Hum. Mol. Genet*. 15: R57–R66.

[25] Folstein, S.E., Rosen-Sheidley, B. (2001). Genetics of autism: complex aetiology for a heterogeneous disorder. *Nat Rev Genet* 2:943–955.

[26] Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurles ME, Carter NP, Scherer SW, Lee C. (2006). Copy number variation: new insights in genome diversity. *Genome Res*. 16:949-961.

[27] Freitag, C.M. (2007). The genetics of autistic disorders and its clinical relevance: a review of the literature. *Mol Psychiatry*. 12:2-22.

[28] Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, Zhang H, Estes A, Brune CW, Bradfield JP, Imielinski M, Frackelton EC, Reichert J, Crawford EL, Munson J, Sleiman PM, Chiavacci R, Annaiah K, Thomas K, Hou C, Glaberson W, Flory J, Otieno F, Garris M, Soorya L, Klei L, Piven J, Meyer KJ, Anagnostou E, Sakurai T, Game RM, Rudd DS, Zurawiecki D, McDougle CJ, Davis LK, Miller J, Posey DJ,

Michaels S, Kolevzon A, Silverman JM, Bernier R, Levy SE, Schultz RT, Dawson G, Owley T, McMahon WM, Wassink TH, Sweeney JA, Nurnberger JI, Coon H, Sutcliffe JS, Minshew NJ, Grant SF, Bucan M, Cook EH, Buxbaum JD, Devlin B, Schellenberg GD, Hakonarson H. (2009). Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature*. 459:569-573.

[29] Hampton OA, Den Hollander P, Miller CA, Delgado DA, Li J, Coarfa C, Harris RA, Richards S, Scherer SE, Muzny DM, Gibbs RA, Lee AV, Milosavljevic A. (2009). A sequence-level map of chromosomal breakpoints in the MCF-7 breast cancer cell line yields insights into the evolution of a cancer genome. *Genome Res*. 19:167-177.

[30] Houlden, H. & Reilly, M.M. (2006). Molecular genetics of autosomal-dominant demyelinating Charcot-Marie-Tooth disease. *Neuromolecular Med*. 8: 43–62

[31] Hurles ME, Dermitzakis ET, Tyler-Smith C. (2008). The functional impact of structural variation in humans. *Trends Genet*. 24: 238–245

[32] Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. (2004). Detection of large-scale variation in the human genome. *Nat. Genet*. 36:949–951

[33] Jamain S, Quach H, Betancur C, Råstam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B, Leboyer M, Gillberg C, Bourgeron T; Paris Autism Research International Sibpair Study. (2003). Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet*. 34:27-29.

[34] Kalscheuer VM, FitzPatrick D, Tommerup N, Bugge M, Niebuhr E, Neumann LM, Tzschach A, Shoichet SA, Menzel C, Erdogan F, Arkesteijn G, Ropers HH, Ullmann R. (2007). Mutations in autism susceptibility candidate 2 (AUTS2) in patients with mental retardation. *Hum Genet*. 121:501-509.

[35] Kidd JM, Sampas N, Antonacci F, Graves T, Fulton R, Hayden HS, Alkan C, Malig M, Ventura M, Giannuzzi G, Kallicki J, Anderson P, Tsalenko A, Yamada NA, Tsang P, Kaul R, Wilson RK, Bruhn L, Eichler EE. (2010). Characterization of missing human genome sequences and copy-number polymorphic insertions. *Nat Methods*. 7:365-371.

[36] Kim HG, Kishikawa S, Higgins AW, Seong IS, Donovan DJ, Shen Y, Lally E, Weiss LA, Najm J, Kutsche K, Descartes M, Holt L, Braddock S, Troxell R, Kaplan L, Volkmar F, Klin A, Tsatsanis K, Harris DJ, Noens I, Pauls DL, Daly MJ, MacDonald ME, Morton CC, Quade BJ, Gusella JF. (2008). Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet.* 82:199-207.

[37] Kim JI, Ju YS, Park H, Kim S, Lee S, Yi JH, Mudge J, Miller NA, Hong D, Bell CJ, Kim HS, Chung IS, Lee WC, Lee JS, Seo SH, Yun JY, Woo HN, Lee H, Suh D, Lee S, Kim HJ, Yavartanoo M, Kwak M, Zheng Y, Lee MK, Park H, Kim JY, Gokcumen O, Mills RE, Zaranek AW, Thakuria J, Wu X, Kim RW, Huntley JJ, Luo S, Schroth GP, Wu TD, Kim H, Yang KS, Park WY, Kim H, Church GM, Lee C, Kingsmore SF, Seo JS. (2009). A highly annotated whole-genome sequence of a Korean individual. *Nature.* 460:1011-1015.

[38] Klauck, S.M. (2006). Genetics of autism spectrum disorder. *Eur. J. Hum. Genet.* 14: 714-720.

[39] Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, Badner JA, Gilliam TC, Nowak NJ, Cook EH Jr, Dobyns WB, Christian SL. (2008). Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet*. 17:628-638.

[40] Lauritsen MB, Pedersen CB, Mortensen PB. (2005). Effects of familial risk factors and place of birth on the risk of autism: a nationwide register-based study. *J. Child Psychol. Psychiatry* 46:963–971.

[41] Lee, J. A. & Lupski, J. R. (2006). Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. *Neuron* 52:103–121.

[42] Linzmeier RM, Ganz T. (2005). Human defensin gene copy number polymorphisms: comprehensive analysis of independent variation in alpha- and beta-defensin regions at 8p22-p23. *Genomics* 86:423–430.

[43] Ma D, Salyakina D, Jaworski JM, Konidari I, Whitehead PL, Andersen AN, Hoffman JD, Slifer SH, Hedges DJ, Cukier HN, Griswold AJ, McCauley JL, Beecham GW, Wright HH, Abramson RK, Martin ER, Hussman JP, Gilbert JR, Cuccaro ML, Haines JL, Pericak-Vance MA. (2009). A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann Hum Genet*. 73:263-273.

[44] Manning MA, Cassidy SB, Clericuzio C, Cherry AM, Schwartz S, Hudgins L, Enns GM, Hoyme HE. (2004). Terminal 22q deletion syndrome: a newly recognized cause of speech and language disability in the autism spectrum. *Pediatrics*. 114:451-457.

[45] Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, Shago M, Moessner R, Pinto D, Ren Y, Thiruvahindrapduram B, Fiebig A, Schreiber S, Friedman J, Ketelaars CE, Vos YJ, Ficicioglu C, Kirkpatrick S, Nicolson R, Sloman L, Summers A, Gibbons CA, Teebi A, Chitayat D, Weksberg R, Thompson A, Vardy C, Crosbie V, Luscombe S, Baatjes R, Zwaigenbaum L, Roberts W, Fernandez B, Szatmari P, Scherer SW. (2008). Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet*. 82:477-488.

[46] Merikangas AK, Corvin AP, Gallagher L. (2009). Copy-number variants in neurodevelopmental disorders: promises and challenges. *Trends Genet*. 25:536-544.

[47] Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, Abyzov A, Yoon SC, Ye K, Cheetham RK, Chinwalla A, Conrad DF, Fu Y, Grubert F, Hajirasouliha I, Hormozdiari F, Iakoucheva LM, Iqbal Z, Kang S, Kidd JM, Konkel MK, Korn J, Khurana E, Kural D, Lam HY, Leng J, Li R, Li Y, Lin CY, Luo R, Mu XJ, Nemesh J, Peckham HE, Rausch T, Scally A, Shi X, Stromberg MP, Stütz AM, Urban AE, Walker JA, Wu J, Zhang Y, Zhang ZD, Batzer MA, Ding L, Marth GT, McVean G, Sebat J, Snyder M, Wang J, Ye K, Eichler EE, Gerstein MB, Hurles ME, Lee C, McCarroll SA, Korbel JO; 1000 Genomes Project. (2011). Mapping copy number variation by population-scale genome sequencing. *Nature*. 470:59-65.

[48] Moessner R, Marshall CR, Sutcliffe JS, Skaug J, Pinto D, Vincent J, Zwaigenbaum L, Fernandez B, Roberts W, Szatmari P, Scherer SW. (2007). Contribution of SHANK3 mutations to autism spectrum disorder. *Am J Hum Genet*. 81:1289-1297.

[49] Moog U, Engelen JJ, Weber BW, Van Gelderen M, Steyaert J, Baas F, Sijstermans HM, Fryns JP. (2004). Hereditary motor and sensory neuropathy (HMSN) IA, developmental delay and autism related disorder in a boy with duplication (17)(p11.2p12). *Genet. Couns*. 15: 73–80

[50] Müller A, Holzmann K, Kestler HA. Visualization of genomic aberrations using Affymetrix SNP arrays. *Bioinformatics* 23**:**496-497.

[51] Park H, Kim JI, Ju YS, Gokcumen O, Mills RE, Kim S, Lee S, Suh D, Hong D, Kang HP, Yoo YJ, Shin JY, Kim HJ, Yavartanoo M, Chang YW, Ha JS, Chong W, Hwang GR, Darvishi K, Kim H, Yang SJ, Yang KS, Kim H, Hurles ME, Scherer SW, Carter NP,

Tyler-Smith C, Lee C, Seo JS. (2010). Discovery of common Asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing. *Nat Genet*. 42:400-405.

[52] Perry GH, Ben-Dor A, Tsalenko A, Sampas N, Rodriguez-Revenga L, Tran CW, Scheffer A, Steinfeld I, Tsang P, Yamada NA, Park HS, Kim JI, Seo JS, Yakhini Z, Laderman S, Bruhn L, Lee C. (2008) The fine-scale and complex architecture of human copy-number variation. *Am. J. Hum. Genet*. 82: 685–695.

[53] Pinto D, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Almeida J, Bacchelli E, Bader GD, Bailey AJ, Baird G, Battaglia A, Berney T, Bolshakova N, Bölte S, Bolton PF, Bourgeron T, Brennan S, Brian J, Bryson SE, Carson AR, Casallo G, Casey J, Chung BH, Cochrane L, Corsello C, Crawford EL, Crossett A, Cytrynbaum C, Dawson G, de Jonge M, Delorme R, Drmic I, Duketis E, Duque F, Estes A, Farrar P, Fernandez BA, Folstein SE, Fombonne E, Freitag CM, Gilbert J, Gillberg C, Glessner JT, Goldberg J, Green A, Green J, Guter SJ, Hakonarson H, Heron EA, Hill M, Holt R, Howe JL, Hughes G, Hus V, Igliozzi R, Kim C, Klauck SM, Kolevzon A, Korvatska O, Kustanovich V, Lajonchere CM, Lamb JA, Laskawiec M, Leboyer M, Le Couteur A, Leventhal BL, Lionel AC, Liu XQ, Lord C, Lotspeich L, Lund SC, Maestrini E, Mahoney W, Mantoulan C, Marshall CR, McConachie H, McDougle CJ, McGrath J, McMahon WM, Merikangas A, Migita O, Minshew NJ, Mirza GK, Munson J, Nelson SF, Noakes C, Noor A, Nygren G, Oliveira G, Papanikolaou K, Parr JR, Parrini B, Paton T, Pickles A, Pilorge M, Piven J, Ponting CP, Posey DJ, Poustka A, Poustka F, Prasad A, Ragoussis J, Renshaw K, Rickaby J, Roberts W, Roeder K, Roge B, Rutter ML, Bierut LJ, Rice JP, Salt J, Sansom K, Sato D, Segurado R, Sequeira AF, Senman L, Shah N, Sheffield VC, Soorya L, Sousa I, Stein O, Sykes N, Stoppioni V, Strawbridge C, Tancredi R, Tansey K, Thiruvahindrapduram B, Thompson AP, Thomson S, Tryfon A, Tsiantis J, Van Engeland H, Vincent JB, Volkmar F, Wallace S, Wang K, Wang Z, Wassink TH, Webber C, Weksberg R, Wing K, Wittemeyer K, Wood S, Wu J, Yaspan BL, Zurawiecki D, Zwaigenbaum L, Buxbaum JD, Cantor RM, Cook EH, Coon H, Cuccaro ML, Devlin B, Ennis S, Gallagher L, Geschwind DH, Gill M, Haines JL, Hallmayer J, Miller J, Monaco AP, Nurnberger JI Jr, Paterson AD, Pericak-Vance MA, Schellenberg GD, Szatmari P, Vicente AM, Vieland VJ, Wijsman EM, Scherer SW, Sutcliffe JS, Betancur C. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature.* 466:368-372.

[54] Przybytkowski E, Ferrario C, Basik M. (2011). The use of ultra-dense array CGH analysis for the discovery of micro-copy number alterations and gene fusions in the cancer genome. *BMC Med Genomics*. 4:16.

[55] Psychiatric GWAS Consortium Coordinating Committee, Cichon S, Craddock N, Daly M, Faraone SV, Gejman PV, Kelsoe J, Lehner T, Levinson DF, Moran A, Sklar P, Sullivan PF. (2009). Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am J Psychiatry*. 166:540-556.

[56] Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, González JR, Gratacòs M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J,

Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. (2006). Global variation in copy number in the human genome. *Nature* 444: 444–454

[57] Reymond A, Henrichsen CN, Harewood L, Merla G. (2007). Side effects of genome structural changes. *Curr Opin Genet Dev*. 17:381-386

[58] Risch N, Merikangas K (1996). The future of genetic studies of complex human diseases. *Science* 273:1516–1517

[59] Ropers HH. (2010). Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet*. 11:161-187

[60] Roohi J, Montagna C, Tegay DH, Palmer LE, DeVincent C, Pomeroy JC, Christian SL, Nowak N, Hatchwell E. (2009). Disruption of contactin 4 in three subjects with autism spectrum disorder. *J Med Genet*. 46:176-182.

[61] Rutter, M. (2005). Autism research: lessons from the past and prospects for the future. *J Autism Dev Disord* 35:241–257.

[62] Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Månér S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M. (2004). Large-scale copy-number polymorphism in the human genome. *Science* 305:525–528.

[63] Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, Leotta A, Pai D, Zhang R, Lee YH, Hicks J, Spence SJ, Lee AT, Puura K, Lehtimäki T, Ledbetter D, Gregersen PK, Bregman J, Sutcliffe JS, Jobanputra V, Chung W, Warburton D, King MC, Skuse D, Geschwind DH, Gilliam TC, Ye K, Wigler M. (2007). Strong association of de novo copy number mutations with autism. *Science*. 316:445-449.

[64] Simon-Sanchez J, Scholz S, Fung HC, Matarin M, Hernandez D, Gibbs JR, Britton A, de Vrieze FW, Peckham E, Gwinn-Hardy K, Crawley A, Keen JC, Nash J, Borgaonkar D, Hardy J, Singleton A. (2007). Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. *Hum. Mol. Genet*. 16: 1–14.

[65] Veenstra-VanderWeele, J., Christian, S.L., Cook, E.H. Jr. (2004). Autism as a paradigmatic complex genetic disorder. *Annu Rev Genomics Hum Genet* 5: 379–405.

[66] Veenstra-VanderWeele J, Cook EH Jr. (2004). Molecular genetics of autism spectrum disorder. *Mol Psychiatry* 9:819–832.

[67] Vorstman, J.A., Staal, W.G., van Daalen, E., van Engeland, H., Hochstenbach, P.F., Franke, L. (2006). Identification of novel autism candidate regions through analysis of reported cytogenetic abnormalities associated with autism. *Mol. Psychiatry* 11:18-28.

[68] Wang K, Zhang H, Ma D, Bucan M, Glessner JT, Abrahams BS, Salyakina D, Imielinski M, Bradfield JP, Sleiman PM, Kim CE, Hou C, Frackelton E, Chiavacci R, Takahashi N, Sakurai T, Rappaport E, Lajonchere CM, Munson J, Estes A, Korvatska O, Piven J, Sonnenblick LI, Alvarez Retuerto AI, Herman EI, Dong H, Hutman T, Sigman M, Ozonoff S, Klin A, Owley T, Sweeney JA, Brune CW, Cantor RM, Bernier R, Gilbert JR, Cuccaro ML, McMahon WM, Miller J, State MW, Wassink TH, Coon H, Levy SE, Schultz RT, Nurnberger JI, Haines JL, Sutcliffe JS, Cook EH, Minshew NJ, Buxbaum JD, Dawson G, Grant SF, Geschwind DH, Pericak-Vance MA,

Schellenberg GD, Hakonarson H. (2009). Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*. 459:528-533.

[69] Weiss, L.A. (2009). Autism genetics: emerging data from genome-wide copy-number and single nucleotide polymorphism scans. *Expert Rev Mol Diagn*. 9:795-803.

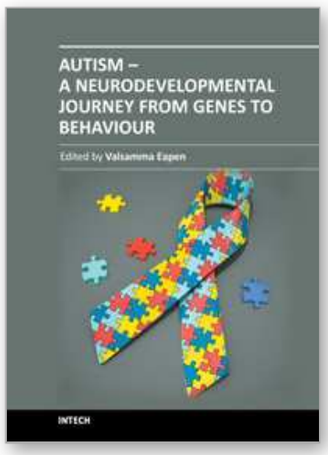
[70] Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R, Saemundsen E, Stefansson H, Ferreira MA, Green T, Platt OS, Ruderfer DM, Walsh CA, Altshuler D, Chakravarti A, Tanzi RE, Stefansson K, Santangelo SL, Gusella JF, Sklar P, Wu BL, Daly MJ; Autism Consortium. (2008). Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med*. 358:667-675.

[71] Weiss LA, Arking DE, Daly MJ, Chakravarti A. (2009). A genome-wide linkage and association scan reveals novel loci for autism. *Nature*. 461:802-808.

[72] Yim SH, Kim TM, Hu HJ, Kim JH, Kim BJ, Lee JY, Han BG, Shin SH, Jung SH, Chung YJ. (2010). Copy number variations in East-Asian population and their evolutionary and functional implications. *Hum Mol Genet*. 19:1001-1008.

[73] Zhang F, Gu W, Hurles ME, Lupski JR. (2009). Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet*. 10:451-481.

[74] Zheng, C.,Heintz, N., Hatten, M.E. (1996). CNS gene encoding astrotactin, which supports neuronal migration along glial fibers. *Science* 272: 417–419



Autism - A Neurodevelopmental Journey from Genes to Behaviour
Edited by Dr. Valsamma Eapen

ISBN 978-953-307-493-1
Hard cover, 484 pages
Publisher InTech
Published online 17, August, 2011
Published in print edition August, 2011

The book covers some of the key research developments in autism and brings together the current state of evidence on the neurobiologic understanding of this intriguing disorder. The pathogenetic mechanisms are explored by contributors from diverse perspectives including genetics, neuroimaging, neuroanatomy, neurophysiology, neurochemistry, neuroimmunology, neuroendocrinology, functional organization of the brain and clinical applications from the role of diet to vaccines. It is hoped that understanding these interconnected neurobiological systems, the programming of which is genetically modulated during neurodevelopment and mediated through a range of neuropeptides and interacting neurotransmitter systems, would no doubt assist in developing interventions that accommodate the way the brains of individuals with autism function. In keeping with the multimodal and diverse origins of the disorder, a wide range of topics is covered and these include genetic underpinnings and environmental modulation leading to epigenetic changes in the aetiology; neural substrates, potential biomarkers and endophenotypes that underlie clinical characteristics; as well as neurochemical pathways and pathophysiological mechanisms that pave the way for therapeutic interventions.

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

Hae-Jin Hu and Yeun-Jun Chung (2011). Genome-Wide Association Studies of Copy Number Variation in Autism Spectrum Disorder, Autism - A Neurodevelopmental Journey from Genes to Behaviour, Dr. Valsamma Eapen (Ed.), ISBN: 978-953-307-493-1, InTech, Available from: http://www.intechopen.com/books/autism-a-neurodevelopmental-journey-from-genes-to-behaviour/genome-wide-association-studies-of-copy-number-variation-in-autism-spectrum-disorder

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen