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.
<table>
<thead>
<tr>
<th>Discovery Sample</th>
<th>Replication Sample</th>
<th>Study design</th>
<th>Platform</th>
<th>CNV Detection method</th>
<th>Number of CNVs identified</th>
<th>Strong candidate loci</th>
<th>CN change</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1496 families with 7,917 subjects</td>
<td>-</td>
<td>Family-based</td>
<td>Affymetrix 10K</td>
<td>dChip</td>
<td>254</td>
<td>NRXN1 1q21 17p12 22q11.2</td>
<td>del</td>
<td>Szatmari et al., 2007</td>
</tr>
<tr>
<td>165 families</td>
<td>-</td>
<td>Family-based</td>
<td>Agilent 244K 390K ROMA</td>
<td>HMM</td>
<td>17</td>
<td>SLC4A10, FHIT FLJ16237 A2BP1</td>
<td>del del del del</td>
<td>Sebat et al., 2007</td>
</tr>
<tr>
<td>180</td>
<td>372</td>
<td>Family-based and Case-control</td>
<td>NimbleGen</td>
<td>1</td>
<td>16p11.2</td>
<td>microdel</td>
<td>Kumar et al., 2008</td>
<td></td>
</tr>
<tr>
<td>751 multiplex families with 1441 cases</td>
<td>1420 (AGRE parents) 2814 (bipolar disorder or NIMH controls)</td>
<td>Family-based</td>
<td>Affymetrix 5.0 (AGRE) Affymetrix 500K (controls)</td>
<td>COPPER/ Birdseye (AGRE) ADM-2(CHB) HMM(deC ODE)</td>
<td>47</td>
<td>16p11.2</td>
<td>del/dup</td>
<td>Weiss et al., 2008</td>
</tr>
<tr>
<td>397</td>
<td>372</td>
<td>Family-based</td>
<td>19K BAC Microarray</td>
<td>-</td>
<td>51</td>
<td>15q11-q13 22q11 16p11.2</td>
<td>dup dup microdel</td>
<td>Christian et al., 2008</td>
</tr>
<tr>
<td>427 families</td>
<td>500</td>
<td>Case-control</td>
<td>Affymetrix 500K</td>
<td>dChip, CNAG, GEMCA</td>
<td>277</td>
<td>16p11.2 SHANK3- NLGN4- NRXN1-PSD DPP6- DPP10- PCDH9 ANKRD11 DPYD PTCHD1 15q24</td>
<td>del/dup</td>
<td>Marshall et al., 2008</td>
</tr>
<tr>
<td>859</td>
<td>1409</td>
<td>Case-control</td>
<td>Illumina HumanHapl550</td>
<td>PennCNV</td>
<td>78,490</td>
<td>15q11-13 22q11.21 NRXN1 CNTN4 PARK2 RFWD2 AK125120 UNQ9037 GRID1 NLGN1 GYPELOC44</td>
<td>dup dup del/dup del dup del del del dup</td>
<td>Glessner et al., 2009</td>
</tr>
<tr>
<td>912 multiplex families</td>
<td>1,488 (CHOP) 542 (NINDS)</td>
<td>Case-control</td>
<td>Illumina HumanHapl550</td>
<td>PennCNV</td>
<td>&gt; 150</td>
<td>NRXN1 UBE3A 15q11-q13 BZRAP1 MDGA2</td>
<td>del del/dup del/dup del</td>
<td>Bucan et al., 2009</td>
</tr>
</tbody>
</table>
Table 1. Genome-wide CNV association studies of autism

<table>
<thead>
<tr>
<th>Discovery Sample</th>
<th>Replication Sample</th>
<th>Study design</th>
<th>Platform</th>
<th>CNV Detection method</th>
<th>Number of CNVs identified</th>
<th>Strong candidate loci</th>
<th>CN change</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 28 children</td>
<td>Control 62 Adults</td>
<td>-</td>
<td>Array-CGH</td>
<td>Array-CyGht</td>
<td>38</td>
<td>8p23.1 17p11.2</td>
<td>del del</td>
<td>Cho et al., 2009</td>
</tr>
<tr>
<td>996</td>
<td>1,287</td>
<td>-</td>
<td>Illumina 1M</td>
<td>QuantiSNP iPattern</td>
<td>5,478</td>
<td>SHANK2 SYNGAP1 DLGAP2 CSNK1D/S LC16A3 NRXN1 22q11.21 DDX53/PTC HD1</td>
<td>del del dup/del dup/del del</td>
<td>Pinto et al., 2010</td>
</tr>
</tbody>
</table>

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

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 AUTS2 (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 neurexin 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, \( P=3.4\times10^{-4} \)). 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.
### 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.
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 $7.4 \times 10^{-8}$ to $2.1 \times 10^{-10}$). 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

- National Institute of Mental Health Center for Collaborative Genetic Studies on Mental Disorders: http://nimhgenetics.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

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


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


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


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: