Distribution of Disease-Associated Copy Number Variants Across Distinct Disorders of Cognitive Development

Matthew F. Pescosolido, B.A., Ece D. Gamsiz, Ph.D., Shailender Nagpal, M.S., Eric M. Morrow, M.D., Ph.D.

Objective: The purpose of the present study was to discover the extent to which distinct *DSM* disorders share large, highly recurrent copy number variants (CNVs) as susceptibility factors. We also sought to identify gene mechanisms common to groups of diagnoses and/or specific to a given diagnosis based on associations with CNVs. **Method:** Systematic review of 820 PubMed articles on autism spectrum disorder (ASD), intellectual disability (ID), schizophrenia, and epilepsy produced 54 CNVs associated with one or several disorders. Pathway analysis on genes implicated by CNVs in different groupings was conducted. **Results:** The majority of CNVs were found in ID with the other disorders somewhat subsumed, yet certain CNVs were associated with isolated or groups of disorders. Based on genes implicated by CNVs, ID encompassed 96.8% of genes in ASD, 92.8% of genes in schizophrenia, and 100.0% of genes in epilepsy. Pathway analysis revealed that synapse processes were enriched in ASD, ID, and schizophrenia. Disease-specific processes were identified in ID (actin cytoskeleton processes), schizophrenia (ubiquitin-related processes), and ASD (synaptic vesicle transport and exocytosis). **Conclusions:** Intellectual disability may arise from the broadest range of genetic pathways, and specific subsets of these pathways appear to be relevant to other disorders or combinations of these disorders. It is clear that statistically significant CNVs across disorders of cognitive development are highly enriched for biological processes related to the synapse. There are also disorder-specific processes that may aid in understanding the distinct presentations and pathophysiology of these disorders. J. Am. Acad. Child Adolesc. Psychiatry; 2013;52(4):414-430. **Key Words:** autism, epilepsy, intellectual disability, schizophrenia, copy number variation

Genome-wide association studies (GWAS) have identified a large number of recurrent copy number variants (CNVs) that are associated with disease and are also frequently shared as susceptibility factors by several disorders of cognitive development.[1–7] However, the distribution of these CNVs across clinically distinct neurodevelopmental disorders, such as autism spectrum disorders (ASD), intellectual disability (ID), schizophrenia, or epilepsy, has not been systematically studied. For example, are there CNVs that are common to all disorders? Are there CNVs that are specific only to a subset of these disorders? Are there CNVs, either deletions or duplications, that are specific to only one disorder? Detailed and systematic examination of the distribution of disease-associated CNVs across these disorders will aid in the nosology of psychiatric disorders and will also provide insight into shared and distinct biological processes underlying the *DSM* diagnoses.

Overlapping symptoms are found across *DSM* categories of neurodevelopmental disorders. In particular, ID, ASD, and schizophrenia are all considered together here as disorders of cognitive development, as all conditions share central symptoms such as cognitive impairment and all have neurodevelopmental causes in a majority of cases.[4] Epilepsy is similarly considered among this group of clinical conditions, given the high rates of co-occurrence and many known shared etiologies with the other diagnoses.[8] Likewise,

Supplemental material cited in this article is available online.

based on a large number of studies, specific CNVs have been identified across *DSM* categories including ID, ASD, schizophrenia, and epilepsy. For example, 16p11.2 CNVs are significantly associated with disease in ASD (deletions[9–12] and duplications[9–11]), schizophrenia (duplications only[13–17]), and ID (deletions[18–21] and duplications[19–21]). 16p11.2 CNVs have also been found in patients with epilepsy.[8] In addition, 1q21.1 CNVs have been confirmed in case-control studies as significantly associated with ID (deletions and duplications[19,20,22]) and schizophrenia (deletions[14,17,23–25] and duplications[14]). By similar logic, attention-deficit/hyperactivity disorder (ADHD) may be grouped with the above conditions as a disorder of cognitive development. In one ADHD GWAS study, large rare CNVs were highest among ADHD patients with co-morbid ID and identified the locus of 16p13.11 as significantly enriched.[26] However, given the relatively low number of large, genome-wide studies in CNVs associated with ADHD, we have excluded ADHD from the current analysis.

Studies of CNVs also promise to identify potential susceptibility mechanisms associated with specific neurodevelopmental disorders. For example, post-synaptic mechanisms have been implicated in ASD as a result of genome-wide CNV studies.[27] Specific post-synaptic pathways include *SHANK3*[28,29] and neuronal cell adhesion (*NLGN3* and *NLGN4X*[30]). Some of these loci may be shared by other conditions, in particular, ID, but also schizophrenia[31] and epilepsy[32] in other cases. The present study attempts to clarify the distribution of disease-associated CNVs across four disorders of cognitive development, namely, ID, ASD, schizophrenia, and epilepsy.

First, we identify the specific CNVs (considering deletions and duplications separately) that are associated with ASD, ID, schizophrenia, and epilepsy. Then, we examine the groups of CNVs that emerge as common to all disorders, specific to subsets of disorders, or unique to a given single disorder. Next, we apply gene pathway analysis techniques to specific subsets of genes within CNVs associated with isolated disorders or given combinations of disorders. We first analyze groups of CNVs based on their association with one given *DSM* category. We then examine genetic pathways associated with CNVs from specific combinations of disorders (i.e., CNVs found in two or more *DSM* disorders) to uncover potential shared genetic mechanisms. In addition, we examine subgroups of CNVs that are associated with relatively isolated conditions (e.g., ID without autism, schizophrenia, or epilepsy) and schizophrenia (without ID, ASD, or epilepsy), with the hope of uncovering mechanisms that may be relatively specific to a given disorder. Our study provides an important and unique approach to the study of CNVs in neurodevelopmental disorders and begins to uncover some of the shared as well as distinct pathways that are at the root causes of the *DSM* diagnoses.

METHOD

Identification of Highly Recurrent CNVs

A systematic review of CNVs was pursued based on the outline in Figure S1A (available online). CNVs associated with neurodevelopmental disorders were searched in PubMed through June 2012. Specific searches were used for ASD (''copy number'' and ''autism''), ID (''copy number,'' ''intellectual disability,'' and ''mental retardation''), schizophrenia (''copy number'' and ''schizophrenia''), and epilepsy (''copy number'' and ''epilepsy''). Searches were limited to English language, humans, and publications after 2005. CNVs were considered highly recurrent and associated with disease if they were identified in a PubMed publication with the following features: genetically tested in a large disease cohort ($N > 400$); included a comparison control sample; statistically compared the CNV frequencies in cases and controls; and the CNV was significantly enriched in the case sample. CNV significance was based on criteria established by the given publication. In all cases, this represented $p < .05$, but the majority of studies used much stricter genome-wide criteria for significance. All CNV coordinates are reported in NCBI build 36, hg18. In general, we were able to consider deletions and duplications as separate CNVs.

Classification of CNV to Specific Disorders and Determination of CNV Coordinates/Genes

Highly recurrent CNVs were assigned to one or more disorder based on the following approach (Figure S1B, available online). CNVs were assigned to the phenotype for which they met rigorous case-control criteria based on criteria described above. This was ''strict'' criteria for association. Furthermore, significant CNVs were assigned to additional disorder categories based on ''broad'' criteria if a study participant with a primary disorder is reported to have symptoms meeting criteria for another disorder. An example of broad criteria is if a child with a recurrent CNV from an ASD study is described as having co-occurring ID, then the given CNV was coded for both ASD and ID disorders.