EDITORIAL

Cognitive Phenotypes and Genomic Copy Number Variations

James R. Lupski, MD, PhD, DSc

In this issue of *JAMA*, Männik and colleagues[1] report that large (>250 kilobase [kb] pairs) structural variants in the genome, specifically deletion and duplication copy number variations (CNVs), were associated with cognitive phenotypes including intellectual disability and reduced educational achievement. These CNVs are deletions and duplications of DNA sequences in the human genome that can be considered as deviations from the normal diploid state at a given location in the genome and represent differing numbers of copies of genetic sequences. The authors also identified and examined the phenotypic consequences of genomic structural variation associated with known genomic disorders[2] among individuals in whom the genomic disorders were not initially clinically recognized.

Related article page 2044

Männik et al initially studied a large Estonian general population, using a random sample cohort of approximately 8000 from the 52 000 DNA samples available in the well-established Estonian Genome Center biobank that contains 5% of the adult Estonian population. Importantly, the study participants who were enrolled in the biobank had phenotype information available for study. They were each examined by a general practitioner who completed health- and lifestyle-related questionnaires, and reported clinical diagnoses. Public educational history also was available for analyses.

The investigators used genome-wide assays to identify CNVs to determine the medical burden of rare CNVs in the general population. Such genomic assays have been clinically available in Europe and the United States for a decade, often replacing chromosomal analysis, particularly in children manifesting developmental delay.[3] Clinical cytogenetic chromosome studies can identify changes in chromosome number (eg, trisomy 21 associated with Down syndrome) and even some chromosomal microdeletions and microduplications greater than 5 to 10 million base (Mb) pairs in size. However, the new genomic assays used in this study can detect submicroscopic genomic changes including CNVs. Thus, many more genomic changes of potential medical relevance can now be robustly identified using these genome-wide assays.

To date, the association of rare CNVs with disorders has been investigated almost exclusively using clinically ascertained cohorts. Männik et al studied the clinical features of genomic disorders[2] in adult carriers from a population cohort without selection based on clinical symptoms. The investigators identified 56 carriers of CNVs associated with known genomic disorders, previously cataloged in the Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources (DECIPHER)[4] within a combined discovery and replication sample of 7877 unrelated individuals, which represented a rate of 7.1 per 1000 personal genomes assayed, or approximately 1 in every 141 individuals studied.[1]

The presence of measurable deleterious phenotypes in the identified individuals suggests that the assumption that carriers of known syndromic CNVs identified in healthy population cohorts are asymptomatic may not be appropriate. To date, this assumption has been invoked as support for reduced penetrance of disease-associated CNV. Careful clinical reevaluation, including assessment of cognitive performance, may be required to determine if a genomic disorder trait is present. Moreover, in some instances penetrance may result from the combined effect of mutant alleles at a locus, thus variation of both alleles at the locus may be required for disease manifestation as observed for a classic recessive disease locus. This has recently been shown for between 8% and 11% of patients with the complex trait of scoliosis whereby the data are best explained by a recessive genetic model at the *TBX6* locus in the Chinese population.[5] Genes and genetic variants at other loci and environmental effects must also be considered.

Männik et al also generated a genome-wide map of rare autosomal CNVs, with a frequency of less than 0.05% and a size larger than 250 kb, and identified 10.5% of the screened general population as carriers. Carriers with deletions larger than 250 kb or duplications larger than 1 Mb compared with the Estonian population showed a statistically significant greater prevalence of intellectual disability (4.3% of deletion carriers, 5.9% of duplication carriers, and 1.7% of the Estonian population), reduced mean educational achievement, and an increased percentage of individuals not graduating from secondary school (33.5% of the deletion carriers, 39.1% of the duplication carriers, and 25.4% of the Estonian population). Evidence for an association between rare CNVs and decreased educational attainment was confirmed by analyses of UK, US, and Italian.

The genomic basis of some fraction of patients with intellectual disability, developmental disorders,[6] and neurocognitive and neurobehavioral traits[7-9] is unfolding. The role of CNVs is clearly established in the pathogenesis of a variety of conditions. New CNV mutations may be involved in the sporadic disease process.[2,10,11] The results reported by Männik et al indicate that individually rare but collectively common intermediate-size CNVs contribute to the variance in educational attainment. This phenotype is an objectively quantifiable trait that could trigger ordering of a genomic study capable of detecting CNV used in clinical care. With the recognition of a potentially causal mutation in an individual, tailored behavioral and educational interventions could be initiated with patients and family that could improve educational outcomes.[12] Although changing a person's genome is not possible, identifying those with

CNVs related to cognitive phenotypes could provide an opportunity to help them reach their fullest potential.[13]

Genomic analyses to identify disease-associated variation, including both CNV and single-nucleotide variation (SNV), expands the scope of the genetic information that can be obtained from a family history. Such genomic studies can identify new mutations that may be relevant to optimizing cognitive performance and health. Genomic studies may become a part of common medical practice, potentially assisting in the formulation of a differential diagnosis. Early studies of the incorporation of large amounts of individual genomic data into clinical practice have empowered physicians with medical knowledge including: the concept of blended clinical phenotypes, due to mutations at more than 1 mendelian disease locus; the knowledge that a new mutation plays a significant role in disease; and the documentation that rare variants are important to disease.[14,15] The latter 2 findings support the ideas formulated in the clan genomics hypothesis for disease—recent variation that arises in individuals or their nearest family members within their clan represent potential medically relevant genetic and genomic variation.[16] Whether this type of genomic knowledge improves patient outcomes is yet unknown.

Work remains. For intellectual disability, this includes more robust characterization of the specific clinical phenotypes and range of variation that can be observed in those phenotypes, with locus-specific genetic and genomic changes. It will be critically important to identify the specific genes that map within the associated CNVs that are indicated in this study, and thereby elucidate the biological roles of the proteins encoded by these genes and functional effects of variation. Only then, with full knowledge of gene mechanisms will it be possible to determine the precise factors that lead to the susceptibility to potential cognitive phenotypes.[17,18]

Understanding the specific biological systems involved and how their perturbation results in reduced intellectual abilities including cognitive performance will better characterize the molecular etiology and the potential cause for the susceptibility to intellectual disability and other neurocognitive traits observed as well as provide insights into any other possibly associated disease processes. Such understanding may eventually enable more specific intervention and clinical management for known genomic disorders that may include directed educational enhancement programs. Even though some genomic tools are now available, practitioners will need to recognize more subtle potentially clinically relevant phenotypes, such as measures of cognition, and better learn when to consider applying genomic studies for the betterment of the lives of patients, their families, and populations.

ARTICLE INFORMATION

Author Affiliations: Molecular and Human Genetics, Department of Pediatrics, and Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas; Texas Children's Hospital, Houston.

Corresponding Author: James R. Lupski, MD, PhD, DSc, Baylor College of Medicine and Texas Children's Hospital, One Baylor Plaza, Houston, TX 77030 (jlupski@bcm.edu).

Conflict of Interest Disclosures: The author has completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and reported that he owns stock in 23andMe, is a paid consultant for Regeneron Pharmaceuticals, has stock options in Lasergen Inc, is a member of the scientific advisory board of Baylor Miraca Genetics Laboratories, and is a coinventor on US and European patents related to molecular diagnostics. The Department of Molecular and Human Genetics at Baylor College of Medicine derives revenue from the chromosomal microarray analysis (CMA) and clinical exome sequencing offered in the Baylor Miraca Genetics Laboratory.

Funding/Support: Dr Lupski is supported by grants HG006542 from the US National Human Genome Research Institute/National Heart, Lung, and Blood Institute (to Baylor-Hopkins Center for Mendelian Genomics), NS058529 from the National Institute of Neurological Disorders and Stroke, GM106373 from the National Institute of General Medical Sciences, HD024064 from the National Institute of Child Health and Development, and EY021163 and EY019861 from the National Eye Institute and from the Smith-Magenis Syndrome Research Foundation.

Role of the Funder/Sponsor: The sponsors had no role in the preparation, review, or approval of the manuscript and decision to submit the manuscript for publication.

REFERENCES

1. Männik K, Mägi R, Macé A, et al. Copy number variations and cognitive phenotypes in unselected populations. *JAMA*. doi:10.1001/jama.2015.4845.

2. Lupski JR. Structural variation mutagenesis of the human genome. *Environ Mol Mutagen*. 2015. doi:10.1002/em.21943.

3. South ST, Lee C, Lamb AN, Higgins AW, Kearney HM; Working Group for the American College of Medical Genetics and Genomics Laboratory Quality Assurance Committee. ACMG Standards and Guidelines for constitutional cytogenomic microarray analysis, including postnatal and prenatal applications: revision 2013. *Genet Med*. 2013;15(11):901-909.

4. DECIPHER web page. https://decipher.sanger.ac.uk/index. Accessed May 5, 2015.

5. Wu N, Ming X, Xiao J, et al. *TBX6* null variants and a common hypomorphic allele in congenital scoliosis. *N Engl J Med*. 2015;372(4):341-350.

6. Wright CF, Fitzgerald TW, Jones WD, et al; Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet*. 2015;385(9975):1305-1314.

7. Weiss LA, Shen Y, Korn JM, et al; Autism Consortium. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med*. 2008;358(7):667-675.

8. Lupski JR. Schizophrenia: incriminating genomic evidence. *Nature*. 2008;455:178-179.

9. Stefansson H, Meyer-Lindenberg A, Steinberg S, et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. *Nature*. 2014;505(7483):361-366.

10. de Ligt J, Willemsen MH, van Bon BW, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med*. 2012;367(20):1921-1929.

11. Rauch A, Wieczorek D, Graf E, et al. Range of genetic mutations associated with severe nonsyndromic sporadic intellectual disability: an exome sequencing study. *Lancet*. 2012;380(9854):1674-1682.

12. Lacaria M, Spencer C, Gu W, Paylor R, Lupski JR. Enriched rearing improves behavioral responses of an animal model for CNV-based autistic-like traits. *Hum Mol Genet*. 2012;21(14):3083-3096.

13. Singh-Manoux A, Ferrie JE, Lynch JW, Marmot M. The role of cognitive ability (intelligence) in explaining the association between socioeconomic position and health. *Am J Epidemiol*. 2005;161(9):831-839.

14. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of Mendelian disorders. *N Engl J Med*. 2013;369(16):1502-1511.

15. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA*. 2014;312(18):1870-1879.

16. Lupski JR, Belmont JW, Boerwinkle E, Gibbs RA. Clan genomics and the complex architecture of human disease. *Cell*. 2011;147(1):32-43.

17. Lalani SR, Zhang J, Schaaf CP, et al. Mutations in *PURA* cause profound neonatal hypotonia, seizures, and encephalopathy in 5q31.3 microdeletion syndrome. *Am J Hum Genet*. 2014;95(5):579-583.

18. Coe BP, Witherspoon K, Rosenfeld JA, et al. Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet*. 2014;46(10):1063-1071.