The importance of rare DNA variation in neurologic disease

Cautionary tale

Jonathan L. Haines, PhD

Correspondence to
Dr. Haines:
Jonathan@chgr.mc.vanderbilt.edu

Neurology® 2013;80:974–975

Neurology has been one of the leaders in applying the developing tools of genetics to understand the etiology of disease. This extends as far back as the early 1980s, when the then-new methods of detecting DNA variation (restriction fragment length polymorphisms) were combined with relatively new analytical methods (linkage analysis) to identify the location of the Huntington disease gene,[1] Alzheimer disease genes,[2] and a Parkinson disease (PD) gene.[3] Extremely rare highly penetrant mutations in these genes are causative for disease. These discoveries redirected entire fields of study and greatly improved our understanding of the underlying pathophysiology. However, they directly explain just a tiny portion of the genetic causes of their disease.

More recently, genome-wide association studies have identified numerous genes with common variation that are associated with many of these diseases, including PD.[4] Unlike the rare mendelian mutations, the effect sizes (or strength of the association) are very small, so the majority of the genetic architecture of PD remains unknown. With the emergence of rapid and deep resequencing technologies, the search for genetic variation in PD has turned to identification of rare variants of strong, but not mendelian level, effect sizes. Recently, rare variation in 2 genes, *VPS35*[5] and *EIF4G1*,[6] was associated with PD. In a new study in this issue of *Neurology*®, Nuytemans et al.[7] raise doubts about the pathogenicity of some of these variants and highlight the need to be particularly cautious about interpreting the role of rare variation in disease.

In their report, Nuytemans et al.[7] used high-throughput sequencing to comprehensively examine these 2 genes in 213 PD cases and 273 controls. As in earlier reports, they identified a number of rare variants in both genes. However, careful analysis, including the valuable addition of family data, raises doubts about the pathogenicity of some variants and the penetrance of others. For example, the previously identified D620N mutation in *VPS35* appears to segregate in multiple families and is highly penetrant.[5,8] However, the potential causal effect of additional variations in *VPS35* is less clear, as they have been seen in only one or a few sporadic cases; the lack of familial segregation of the observed Y507F variant in the current study highlights this problem.

Many more individually rare variants are observed in *EIF4G1*. As in *VPS35*, 1 variation, R1205H, appears to cosegregate with PD in multiple families, strongly suggesting a causative effect. However, it also occurs in several population controls and in an unaffected member of a pedigree in the current study, raising substantial questions about the level of penetrance for this likely causal variant. Of the other rare variants, several occur in control samples, are not predicted to damage protein function, or cluster in a gene region that is likely to be tolerant of variation, bringing into question any role in PD.

Rapid advances in DNA sequencing technology have made it possible to assay comprehensively any (or every) gene for variation and thus generate massive amounts of new data to be analyzed and interpreted. In early gene discovery studies, the focus on very rare mendelian diseases made declaration of causality straightforward when cosegregation of a gene variant with disease was observed across multiple families and the mutation was not seen in a reasonable number of control samples. It is now clear that we cannot apply the same approach to the rare variants identified in non-mendelian diseases, such as PD, Alzheimer disease, multiple sclerosis, or many other neurologic diseases.

As these data are collected, it is critical that multiple lines of support for their role in disease be investigated. Just because a gene has been implicated in a disease does not mean that every rare variant in the gene plays a role in the disease. As with mendelian mutations, one of the most powerful methods is to examine cosegregation within families. There is no better way to enrich a sample for any rare variant than to examine relatives of someone with the variant. Another approach is to examine very large samples (tens of thousands) of cases and controls, an approach that is becoming more common, but only through consortia efforts. New statistical methods are being developed that test the overall "burden" of rare variants in cases and controls, compensating for their individual rarity by considering all variants in a single test. Of course, the final test is to demonstrate the biological significance of the variant. Clues to

See page 982

From the Center for Human Genetics Research, Departments of Molecular Physiology & Biophysics, Neurology, and Ophthalmology & Visual Sciences, Vanderbilt University School of Medicine, Nashville, TN.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the author, if any, are provided at the end of the editorial.

functional significance are increasingly found through bioinformatic annotation of the variation,[9] but in vitro and in vivo studies will ultimately be needed to confirm our suspicions.

In the rush to use and apply the latest technologies, we must remain cautious in our interpretations and wait until the full story can be told. Over the next few years, multitudes of rare variants will be identified and implicated in disease. However, in most cases, these variations will have unknown significance and will thus be of little use in diagnosis, prognosis, or treatment. How and when to move toward their use in clinical practice is a topic of intense conversation[10] that raises many legal, social, and ethical issues that are not yet resolved.

STUDY FUNDING

No targeted funding reported.

DISCLOSURE

The author reports no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

REFERENCES

1. Gusella JF, Wexler NS, Conneally PM, et al. A polymorphic DNA marker genetically linked to Huntington's disease. Nature 1983;306:234–238.
2. Pericak-Vance MA, Haines JL. Genetic susceptibility to Alzheimer disease. Trends Genet 1995;11:504–508.
3. Polymeropoulos MH, Lavedan C, Leroy E, et al. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. Science 1997;276:2045–2047.
4. Pankratz N, Beecham GW, DeStefano AL, et al. Meta-analysis of Parkinson's disease: identification of a novel locus, RIT2. Ann Neurol 2012;71:370–384.
5. Vilarino-Guell C, Wider C, Ross OA, et al. VPS35 mutations in Parkinson disease. Am J Hum Genet 2011;89:162–167.
6. Chartier-Harlin MC, Dachsel JC, Vilarino-Guell C, et al. Translation initiator EIF4G1 mutations in familial Parkinson disease. Am J Hum Genet 2011;89:398–406.
7. Nuytemans K, Bademci G, Inchausti V, et al. Whole exome sequencing of rare variants in *EIF4G1* and *VPS35* in Parkinson disease. Neurology 2013;80:982–989.
8. Zimprich A, Benet-Pages A, Struhal W, et al. A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. Am J Hum Genet 2011;89:168–175.
9. Moreau Y, Tranchevent LC. Computational tools for prioritizing candidate genes: boosting disease gene discovery. Nat Rev Genet 2012;13:523–536.
10. Clayton EW, McGuire AL. The legal risks of returning results of genomics research. Genet Med 2012;14:473–477.