Rapid genetic diagnosis in single-gene movement disorders

Andrew B. Singleton
Chief of the Laboratory of Neurogenetics, National Institute on Aging, Building 35 Room 1A1014, 35 Convent Drive, Bethesda, MD 20892, USA.

Andrew B. Singleton: singleta@mail.nih.gov

In this issue of the journal Ruth Walker and colleagues describe the genetic diagnosis of neuroacanthocytosis disorders using a fairly new but rapidly prevailing genetic method, exome sequencing.[1] This work highlights a standing problem in diseases with a great deal of genetic heterogeneity, how to reach a molecular diagnosis. Perhaps more importantly, it illustrates how this issue may be overcome quite routinely in the near future.

Although not unique, the marked genetic heterogeneity associated with many clinical presentations represents a diagnostic problem that is pronounced in movement disorders. Illustrative examples abound, there are at least 19 identified genetic causes of autosomal dominant ataxia,[2] primary dystonia may be caused by known mutations at 10 different loci,[3] recessive parkinsonism by mutations in at least 6 different genes,[4–9] and in the case of neuroacanthocytosis, there are four genes known to contain disease causing mutations, including a very large gene, *VPS13A*.[1]

Facing genetically heterogeneous but clearly familial disorders in the clinic, a physician faces a decision about genetic testing that must balance cost, time, likelihood of a positive result, and the impact of any result on treatment, prognosis, and the patients family. A common approach is to test the most likely candidate mutations or genes first, progressing steadily through alternates until a verdict is reached, or the most plausible and testable candidate pool is exhausted. This iterative approach can be highly inefficient; it may take a substantial amount of time and money to find a mutation and oftentimes we still fail to identify the causal genetic lesion. There are many reasons for this: while there remain a large number of single gene disorders for which we don't know the genetic cause, many gene tests are not routinely available, and variants are frequently discovered that are of unknown significance.

The quite broad phenotypic heterogeneity associated with a large number of gene mutations also presents a hurdle to identifying the specific genetic cause of clinical presentations. We know that single-gene movement disorders can present with phenotypes outside of the norm. Patients with a pathogenic *LRRK2* mutation most commonly present with typical Parkinson's disease (PD);[10] however, some of the earliest *LRRK2* families described included patients with presentations such as amyotrophy, dystonia, and dementia.[11] Likewise, *PLA2G6* mutations, which were originally associated with infantile neuroaxonal dystrophy, are known to also cause dystonia-parkinsonism.[7, 12] *VCP* mutations can lead to any one or a combination of frontotemporal dementia, inclusion body myopathy, amyotrophic lateral sclerosis, and Paget's disease of bone.[13] It is also reasonable to suspect that the same genetic lesion may result in different clinical presentations based on genetic background, and this may be evident across diverse populations. There already exists some support of this in ataxia, where *ATXN2* and *ATXN3* expansion mutations, primarily described in Caucasian spinocerebellar ataxia patients, may present as a dopa-responsive and quite pure parkinsonism in patients of Asian or African lineage.[14–16]

Given the challenges of substantial genetic and phenotypic heterogeneity, what is needed is a way to screen all gene candidates for a disorder, including those outside of immediately plausible candidate gene pool. Such a method needs to be quick and cost effective and should ideally provide unequivocal results that are easily interpretable. Exome sequencing promises to best some of these challenges. In essence this method aims to produce genetic sequence data on every protein coding DNA basepair of the human genome, representing about 1/100th of the diploid genome, or approximately 70 million base pairs per individual. There are a number of approaches to this but they each center on creating a library from a genomic DNA sample, which is enriched for the genetic regions of interest, then analyzed using a second-generation sequencer. This method generates billions of basepairs of information in quite short order. Sample preparation and sequencing takes ~ 3 weeks; but many samples can be sequenced in parallel. Following this, a computationally intensive but increasingly stable and accessible analysis generates a report of all variants detected, their population frequencies, predicted effect on the protein product, and whether the variants have been described before. Current costs for an exome sequence run at around $1000. In the paper by Ruth Walker and coworkers this is exactly the approach that was taken. In this study the investigators performed exome sequencing in two unrelated individuals with neuroacanthocytosis. In both individuals this method revealed compound heterozygous mutations in *VPS13A*, a known genetic cause of this disorder.

So is exome sequencing the panacea to current limitations faced in clinical testing? The information it provides is certainly extensive, now quite accurate, and has been used to make genetic diagnosis, including unanticipated diagnosis.[17] There are, however, some limitations to the method that bear consideration. In its current form exome sequence does not quite capture all of the protein coding genome, and is generally unable to capture or accurately sequence about 10% of this content. Exome sequencing is not particularly good at identifying repeat mutations (i.e. triplet repeats seen in spinocerebellar ataxias for instance), and although improving, methods for detecting gene deletions and duplications using exome data are not well established.

Currently, exome sequencing is not clinically validated, and will likely require additional follow up with established mutation discovery methods once a candidate variant has been identified. Despite these limitations, the comprehensive nature of exome means that this remains an attractive method for genetic testing; indeed companies are now offering this as part of a diagnostic service. Given the breadth of sequencing afforded by this method, it is also useful to consider the ramifications of unanticipated genetic findings. What are the consequences of a collateral discovery of a *BRCA1* mutation in a patient without cancer, or in identifying a carrier of a recessive mutation who is about to start a family? One of the earliest exome sequencing studies provides an example of this, finding not only the cause of Miller syndrome (*DHODH* mutation), a rare multiple malformation disorder, but also showing genetically that one of the patients had primary ciliary dyskinesia as a result of *DNAH5* mutations.[18] These are clearly complex issues, that will require a shift in our view of genetic data (and perhaps genetic determinism), most importantly this will require an informed discussion with patients seeking genetic diagnosis using this method.

One early consequence of diagnostic exome sequencing is the identification of a large number of variants of unknown significance and an opportunity exists here. It is likely that clinical diagnostic services will be the primary producer of sequences in monogenic or suspected genetic disease, eventually outstripping the throughput of research laboratories. What is called for is a centralized database of variants identified, linked with phenotypes, and this resource should be available for all of these clinical services to both data mine and contribute to. This provides an opportunity to build statistical confidence in the pathogenic role (or not) of identified variants based on data accumulating worldwide. While these are

likely to be difficult waters to navigate, particularly using clinical data for research and maintaining subject anonymity, this seems a worthy goal, likely to identify new genes, new mutations and perhaps new risk variants; ultimately such a service would greatly help genetic diagnostic accuracy and risk prediction.

Clearly exome sequencing is already making an impact in the research arena, and is now the primary method in identifying novel genetic causes of disease. This has been discussed extensively elsewhere, so I will not go into details here, but there exists a burgeoning gene discovery literature that includes examples in movement disorders (*EIF4G1* and *VPS35* mutations in PD, *TGM6* mutations in SCA).[19–22]

Lastly, it is worth noting that exome sequencing likely represents a half step toward a complete genetic solution for genetic diagnosis, in the form of whole genome sequencing. This method that has already been used to find new genetic causes of disease in the research setting.[23] One would expect this approach to rapidly find its way into the diagnostic field, probably replacing exome sequencing, because whole genome sequencing provides much more complete coverage and is better positioned to identify both structural and non-protein coding mutations. Many of the challenges that we will face with the implementation and interpretation of exome data will prepare us for whole genome sequencing data, although we will likely have to be more thoughtful about how we approach low risk variability. The use of exome sequencing, and the inevitable transition to whole genome sequencing in clinical diagnosis presents us with a host of opportunities, challenges, and food for thought.

Acknowledgments

REFERENCES

1. Walker RH, Schulz VP, Tikhonova IR, et al. Genetic diagnosis of neuroacanthocytosis disordes using exome sequencing. Movement Disorders. 2011

2. Klockgether T. Update on degenerative ataxias. Curr Opin Neurol. 2011; 24(4):339–345. [PubMed: 21734495]

3. Ozelius LJ, Lubarr N, Bressman SB. Milestones in dystonia. Mov Disord. 2011; 26(6):1106–1126. [PubMed: 21626555]

4. Bonifati V, Rizzu P, van Baren MJ, et al. Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. Science. 2003; 299(5604):256–259. [PubMed: 12446870]

5. Di Fonzo A, Dekker MC, Montagna P, et al. FBXO7 mutations cause autosomal recessive, early-onset parkinsonian-pyramidal syndrome. Neurology. 2009; 72(3):240–245. [PubMed: 19038853]

6. Kitada T, Asakawa S, Hattori N, et al. Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature. 1998; 392(6676):605–608. [PubMed: 9560156]

7. Paisan-Ruiz C, Bhatia KP, Li A, et al. Characterization of PLA2G6 as a locus for dystonia-parkinsonism. Ann Neurol. 2009; 65(1):19–23. [PubMed: 18570303]

8. Ramirez A, Heimbach A, Grundemann J, et al. Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. Nat Genet. 2006; 38(10):1184–1191. [PubMed: 16964263]

9. Valente EM, Abou-Sleiman PM, Caputo V, et al. Hereditary early-onset Parkinson's disease caused by mutations in PINK1. Science. 2004; 304(5674):1158–1160. [PubMed: 15087508]

10. Paisan-Ruiz C, Jain S, Evans EW, et al. Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. Neuron. 2004; 44(4):595–600. [PubMed: 15541308]

11. Zimprich A, Biskup S, Leitner P, et al. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. Neuron. 2004; 44(4):601–607. [PubMed: 15541309]

12. Morgan NV, Westaway SK, Morton JE, et al. PLA2G6, encoding a phospholipase A2, is mutated in neurodegenerative disorders with high brain iron. Nat Genet. 2006; 38(7):752–754. [PubMed: 16783378]

13. Johnson JO, Mandrioli J, Benatar M, et al. Exome sequencing reveals VCP mutations as a cause of familial ALS. Neuron. 2010; 68(5):857–864. [PubMed: 21145000]

14. Subramony SH, Hernandez D, Adam A, et al. Ethnic differences in the expression of neurodegenerative disease: Machado-Joseph disease in Africans and Caucasians. Mov Disord. 2002; 17(5):1068–1071. [PubMed: 12360561]

15. Gwinn-Hardy K, Singleton A, O'Suilleabhain P, et al. Spinocerebellar ataxia type 3 phenotypically resembling parkinson disease in a black family. Arch Neurol. 2001; 58(2):296–299. [PubMed: 11176969]

16. Gwinn-Hardy K, Chen JY, Liu HC, et al. Spinocerebellar ataxia type 2 with parkinsonism in ethnic Chinese. Neurology. 2000; 55(6):800–805. [PubMed: 10993999]

17. Choi M, Scholl UI, Ji W, et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A. 2009; 106(45):19096–19101. [PubMed: 19861545]

18. Ng SB, Buckingham KJ, Lee C, et al. Exome sequencing identifies the cause of a mendelian disorder. Nat Genet. 2010; 42(1):30–35. [PubMed: 19915526]

19. Chartier-Harlin MC, Dachsel JC, Vilarino-Guell C, et al. Translation initiator EIF4G1 mutations in familial Parkinson disease. Am J Hum Genet. 2011; 89(3):398–406. [PubMed: 21907011]

20. 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(1):168–175. [PubMed: 21763483]

21. Vilarino-Guell C, Wider C, Ross OA, et al. VPS35 mutations in Parkinson disease. Am J Hum Genet. 2011; 89(1):162–167. [PubMed: 21763482]

22. Wang JL, Yang X, Xia K, et al. TGM6 identified as a novel causative gene of spinocerebellar ataxias using exome sequencing. Brain. 2010; 133(Pt 12):3510–3518. [PubMed: 21106500]

23. Lupski JR, Reid JG, Gonzaga-Jauregui C, et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. N Engl J Med. 2010; 362(13):1181–1191. [PubMed: 20220177]