Next-generation sequencing still needs our generation's clinicians

OPEN

A. Reghan Foley, MD*
Sandra Donkervoort, MS, CGC*
Carsten G. Bönnemann, MD

Correspondence to Dr. Bönnemann: carsten.bonnemann@nih.gov

Neurol Genet
2015;1:e13; doi: 10.1212/NXG.0000000000000019

The full arrival and broader availability of next-generation sequencing (NGS) is transforming the practice of medicine, including neurology. Compared with the traditional one-gene-at-a-time Sanger sequencing, NGS, or massively parallel sequencing, is a radically different approach to genetic sequencing. NGS allows for a large number of genes to be captured and sequenced in parallel, creating an enormous amount of data in a relatively short period of time at much lower cost "per gene."

Traditional direct Sanger sequencing may still be a cost- and time-effective genetic diagnostic testing approach in disorders with distinctive phenotypes associated with minimal genetic heterogeneity. NGS-based approaches will be far more effective in disorders with prominent genetic heterogeneity, as is the case in the dystroglycanopathies and many forms of hereditary motor and sensory neuropathies (HMSNs). Various clinical laboratories offer panel-based NGS testing, and each laboratory has its own set of criteria for selecting the genes represented in the panel, making proper clinical categorization essential in identifying the appropriate panel for individual patients. Whole-exome sequencing (WES) offers the most comprehensive testing methodology, without the limitation of targeting only known disease candidate genes. While this is a promising genetic testing approach, particularly in patients in whom single-gene or panel approaches have not identified causative mutations, it also has some limitations as a primary testing platform, as we will discuss later. For those patients who remain undiagnosed even after WES, research-based gene discovery studies including whole-genome sequencing (WGS) or even RNA sequencing (using available tissue) may be the next step.

NGS-based technologies have already transformed diagnostic testing paradigms in the clinical neurology setting, but their impact has yet to be fully assessed. In this issue of *Neurology® Genetics*, Tian et al.[1] report on a robust and comprehensive NGS-based panel testing for 236 neuromuscular-related genes. They report that this targeted capture approach resulted in an excellent yield, identifying deleterious mutations in 29 of 35 families (83%) studied in their cohort of patients with clinically heterogeneous neuromuscular diseases, including congenital myopathies, limb-girdle muscular dystrophies (LGMDs), Charcot-Marie-Tooth (CMT) disease, metabolic myopathies, myotonic syndromes, and ion channel diseases. They considered the diagnosis to be definitive in 21 families (60%) and likely in an additional 8 families (23%). The authors correctly point out that the strength of their NGS and capture approach is the excellent coverage (with a mean depth of approximately 1,000-fold) and the considerable number of genes targeted in parallel.

The report also highlights diagnostic challenges arising from the expanding genetic heterogeneity underlying seemingly similar clinical neurologic phenotypes and the frequently striking phenotypic heterogeneity arising from mutations in single genes and how these challenges can be addressed by an NGS-based approach. In one example, Tian et al.[1] identified compound heterozygous mutations in *IGHMBP2* in a patient clinically diagnosed with infantile-onset CMT. This gene was initially found to underlie spinal muscular atrophy with respiratory distress type 1 (SMARD1), but its expanded phenotype now also includes HMSNs in the absence of respiratory symptoms. This latter presentation is clinically much less suggestive; therefore, direct genetic testing is not typically considered. *IGHMBP2*, however, may be "automatically" included in a broad NGS panel, as it was here. This can lead to a rapid diagnosis, significantly shortening the diagnostic path or "diagnostic odyssey" historically associated with consecutive targeted Sanger gene sequencing, as well as reducing the cost of the diagnostic testing.

See article

*These authors contributed equally to the manuscript.

From the Neuromuscular and Neurogenetic Disorders of Childhood Section, Neurogenetics Branch, National Institute of Neurological Disorders and Stroke, NIH, Bethesda, MD.

Funding information and disclosures are provided at the end of the editorial. Go to Neurology.org/ng for full disclosure forms. The Article Processing Charge for this editorial was waived at the discretion of the Editor.

It may seem tempting for the clinician—trying to keep up with the dizzying pace of discovery of new genes and genetic entities—to consider this technology to be the magic bullet, and in many ways it can be. With it, however, comes a number of challenges and complexities of interpretation that the clinician has to be fully aware of and that require not only genetic but also clinical expertise, perhaps more so than ever before. Hence, the proverbial ball is now back in the clinician's court. While the technicalities of the testing and variant "calling" are firmly in the genetic laboratories' court, the clinician's crucial input and interpretation is required both before and after the testing. It is essential to have a solid understanding of the extended phenotype (which may include histology, physiology, or imaging) before ordering genetic testing not only to ensure that the appropriate panel is ordered but especially to recognize those conditions for which NGS sequencing is unable to reliably identify the pathogenic genetic cause. In fact, this is the case for the most common genetic neuromuscular disorders, including trinucleotide repeat expansions (myotonic dystrophy, Friedreich ataxia), deletions, duplications, or, in particular, "copy number neutral" inversions of multiple exons—as NGS is now getting better at detecting larger deletions or duplications (Duchenne/Becker muscular dystrophy)—deletions situated in highly similar duplicated genomic regions (spinal muscular atrophy), contractions or duplications of genomic regions (facioscapulohumeral muscular dystrophy and CMT type 1A), abnormal methylation (Prader-Willi syndrome, an important differential diagnosis in the hypotonic infant), and chromosomal disorders (molecular karyotype by array comparative genomic hybridization may be better here). For this reason alone, deep phenotyping and a hierarchical differential diagnosis are imperative when selecting the genetic testing methodology appropriate for an individual patient.

NGS, like every form of genetic testing, not only requires such careful "pre-testing" phenotyping but also requires an equally careful "post-testing" analysis. The post-testing challenges for the clinician arise from interpreting the results reported by the genetic laboratory. Using this approach of "pre-test" deep phenotyping, NGS, and "post-test" clinical plausibility checking, the potential outcomes of NGS for the clinician include (1) genetic diagnosis confirmed, as suspected; (2) genetic diagnosis confirmed, but not as suspected; (3) incomplete genetic diagnosis in a highly suspected gene, such as a missing second pathogenic allele in a recessive disease; (4) genetic variants of unknown significance (VUS) in more than 1 gene, which are possibly causative; and (5) no genetic diagnosis achieved. For the purpose of this discussion, we will not consider the equally vexing problem of the identification of unexpected secondary findings on WES or WGS that are unrelated to the patient's underlying clinical diagnosis but may have significant implications for future medical management.[2]

Tian et al. provide examples of all of these potential scenarios. A good example for outcome 1 would be the identification of de novo *ACTA1* mutations in patients with histologically evident nemaline myopathy. The *IGHMBP2* mutations mentioned before illustrate scenario 2. When considering scenario 3 (potentially "missing" mutation), it is helpful to realize that in contrast to WES, NGS-based panels typically cover all genes included in the panel at superior depth. Deletion/duplication testing (to assess for copy number variations) can supplement NGS, thus decreasing the risk of "missed" mutations. Coverage for WES may be more incomplete, with only 80%–85% of the exome captured and at a lower depth.[3] A limitation of both approaches is the potential for "drop out" of critical exons, especially those with high GC-rich regions. To avoid false-negative results from NGS-based panels and WES, laboratories may offer Sanger sequencing to "fill in" regions with known reduced coverage in disease-associated candidate genes. Pseudogenes and highly repetitive regions in clinically relevant genes can also be misaligned by WES, potentially resulting in "missed" causative variants.[4] Mutations located "deep" within introns or regulatory regions may be outside of the coverage of NGS-based panels and WES and may affect splicing, pseudoexon inclusion, or expression of the gene. As briefly mentioned earlier, copy-neutral rearrangements such as inversions with deep intronic breakpoints can escape detection by even superior-depth NGS panels. In addition, capture-based NGS approaches are built on reference sequences available at the time of construction, which may change with the annotation of additional exons, as has happened with nebulin (*NEB*).

One of the main hurdles in broad panel-based testing, especially in WES and WGS, is scenario 4, the interpretation of a large number of sequence VUS, which can be incredibly time consuming and may result in incorrectly assigned pathogenicity or missed causative variants.[5] It is the laboratory's responsibility to help evaluate any variants found with regard to their predicted pathogenicity (based on allele frequency, in silico analysis tools, data in published reports and databases, and tools for predicting the functional effects of amino acid substitutions such as PolyPhen and SIFT, as detailed nicely in Tian et al.[1,6]). Given the number of sequence variants reported by WES, it is essential for the clinician to have strategies in place for sorting through the variants reported by the genetic laboratory as potentially

deleterious. This list will be narrowed by the laboratory to initially include only those with disease association; however, the likely still considerable number of variants can be more successfully analyzed with a hypothesis-driven condensed candidate gene list in mind, based on a hierarchical clinical diagnosis of a well-phenotyped patient. Familial segregation testing based on the suspected mode of inheritance and additional phenotyping may help shed light on the interpretation of these variants. In fact, this is one of the most important steps in clinical validation and also the one that can only be pursued by the clinician, frequently requiring considerable additional effort. Collaborating with geneticists and translational scientists can be an essential step for the neurologist, both in evaluating a list of potential causative variants and in working toward confirming the pathogenicity of particular variants, including laboratory-based validation of the functional impact of a variant, which is sometimes possible on a research basis. Much larger population-based data sets may be required to fully catalog all the variants in a given gene across populations and ideally define the functional impact of all of the variants observed. The public availability of very deep variant databases such as Exome Aggregation Consortium browser (http://exac.broadinstitute.org) is an important step in that direction. Thus, some variants will have to remain VUS until all of these pieces fall into place.

The challenges of determining VUS are also highlighted in Tian et al., particularly in the titin gene, which can cause dominantly and recessively inherited muscle disease. They report a patient with a phenotype of LGMD in whom 3 variants were found in the titin gene, all of which were predicted to be deleterious, raising the unresolved question of how to properly assign pathogenicity to many of the variants found in this gene. Not surprisingly, they also report 9 potentially deleterious variants in titin in patients with apparent disease-causing mutations in other genes. In fact, with the broad use of NGS, it has become apparent that pathogenicity has been assigned too liberally to variants that appear "damaging" by various algorithms but that are later found to also occur in healthy individuals.[7] Tian et al. claim that comprehensive NGS can be a cost-effective means of identifying mutations. This is true, but it is important to recognize that the "price" of this approach is time on the part of the neurologist, geneticist, bioinformatician, and translational research scientist to identify and validate variants in disease genes within the phenotypic context of the patient.

Lastly, in the case of no genetic diagnosis being identified (scenario 5), this could indicate that disease-causing mutations were missed because of the limitations of current NGS or that a novel gene may be involved. Indeed, as good as panel-based NGS is at covering sets of genes of interest, we are only looking "under the streetlight" and the truly disease-causing variants may be out there "in the dark." This moves into the field of research and gene discovery, for which NGS holds considerable power. Pursuing a genetic diagnosis in undiagnosed patients may require forming research collaborations to reanalyze or add to the already available sequence data. Such collaborations remain important, especially since they can result in the elucidation of pathophysiologic mechanisms, potentially leading to therapeutic options. For example, the discovery of the *SLC52A2* gene through WES in children with Brown-Vialetto-Van Laere syndrome—and the role of this gene in encoding the riboflavin transporter RFVT2—led to trials of high-dose oral riboflavin and subsequent preliminary evidence of clinical and biochemical improvements in children with this sensory and motor neuronopathy who had previously demonstrated relentlessly progressive symptoms.[8,9]

The identification of disease-causing mutations enables the ultimate genetic confirmation of a clinical diagnosis, ensuring appropriate management, accurate prognosis, precise recurrence risk counseling, and identification of potential clinical or trial-based therapeutic interventions. As such, a molecular genetic diagnosis is increasingly becoming the gold standard for diagnosis of neurogenetic conditions and should be attempted whenever there is a suspected monogenic etiology that could be accurately determined. It should be clear, however, that the wonders of NGS need to be firmly intertwined with the clinician's grip on the clinical phenotype. It is not the gene alone that makes a diagnosis but the entirety of the clinical phenotype and genotype. If "potentially damaging" variants are found in a gene that does not seem to fit with the patient's phenotype, then this should not be the unchallenged diagnosis, unless it constitutes a considerable expansion of the phenotypic spectrum associated with the gene, which requires substantial additional proof.

Herein lies the continued indispensable role of the clinical neurologist. In the same manner in which neurologists were called upon to study, understand, and help to interpret modalities of neuroimaging (CT and MRI) and neurophysiology (EEG and EMG) as these forms of technology became available, the neurologist of our generation is called upon to study, understand, and help to interpret the genetic information made available from NGS. Similarly, as CT, MRI, EEG, and EMG serve as complements to the clinical acumen and skills of the neurologist, so too do NGS technologies complement the clinical neurologist of today. We suggest that for our generation of neurologists, exposure to NGS technologies and an understanding of their strengths and weaknesses are the next steps toward using these technologies to improve diagnoses and clinical care for individuals affected by neurologic disease.

ACKNOWLEDGMENT

The authors thank Drs. Daniel MacArthur and Monkol Lek (MGH/Harvard, and the Broad Institute, MIT, Boston) and Leslie Hotchkiss (NINDS/NIH) for critically reading the manuscript and for their helpful suggestions.

STUDY FUNDING

A.R.F., S.D., and C.G.B. are supported by NINDS intramural funds.

DISCLOSURE

A. Reghan Foley has received travel funding from Cure CMD, the European Neuromuscular Centre and Muscular Dystrophy UK, and is employed by the NIH. Sandra Donkervoort reports no disclosures. Carsten G. Bönnemann has served on scientific advisory boards for Cure CMD, RYR1 Foundation, BioMedical Research, BioMarin Pharmaceutical, Third Rock Ventures, and Audentes Therapeutics; has received travel expenses for lectures or educational activities not funded by industry; has served as Co-Editor-in-Chief of *Journal of Neuromuscular Diseases*; has served as principal investigator without financial compensation of the CALLISTO trial sponsored by Santhera Pharmaceuticals; and is employed by and has received research support from the NIH. Go to Neurology.org/ng for full disclosure forms.

REFERENCES

1. Tian X, Liang WC, Feng Y, et al. Expanding genotype/phenotype of neuromuscular diseases by comprehensive target capture/NGS. Neurol Genet 2015;1:e14. doi: 10.1212/NXG.0000000000000015.
2. Green RC, Berg JS, Grody WW, et al; American College of Medical Genetics and Genomics. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med 2013;15:565–574.
3. Xue Y, Ankala A, Wilcox WR, Hegde MR. Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. Genet Med 2015;17:444–451.
4. Lek M, MacArthur D. The challenge of next generation sequencing in the context of neuromuscular diseases. J Neuromuscul Dis 2014;1:135–149.
5. Foley AR, Pitceathly RD, He J, et al. Whole-genome sequencing and the clinician: a tale of two cities. J Neurol Neurosurg Psychiatry 2014;85:1012–1015.
6. Richards S, Aziz N, Bale S, et al; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–423.
7. Bell CJ, Dinwiddie DL, Miller NA, et al. Carrier testing for severe childhood recessive diseases by next-generation sequencing. Sci Transl Med 2011;65:65ra4.
8. Johnson JO, Gibbs JR, Megarbane A, et al. Exome sequencing reveals riboflavin transporter mutations as a cause of motor neuron disease. Brain 2012;135:2875–2882.
9. Foley AR, Menezes MP, Pandraud A, et al. Treatable childhood neuronopathy caused by mutations in riboflavin transporter RFVT2. Brain 2014;137:44–56.

Next-generation sequencing still needs our generation's clinicians

A. Reghan Foley, Sandra Donkervoort and Carsten G. Bönnemann

Neurol Genet 2015;1;
DOI 10.1212/NXG.0000000000000019