The Journal of Molecular Diagnostics, Vol. ■, No. ■, ■ 2018

ELSEVIER

the Journal of Molecular Diagnostics
jmd.amjpathol.org

A Reliable Targeted Next-Generation Sequencing Strategy for Diagnosis of Myopathies and Muscular Dystrophies, Especially for the Giant Titin and Nebulin Genes

Reda Zenagui,* Delphine Lacourt,* Henri Pegeot,* Kevin Yauy,* Raul J. Morales,†‡ Corine Theze,* François Rivier,‡§ Claude Cances,‡¶ Guilhem Sole,‡‖ Dimitri Renard,‡** Ulrike Walther-Louvier,‡§ Xavier Ferrer-Monasterio,‡‖ Caroline Espil,‡†† Marie-Christine Arné-Bes,‡,‡‡ Pascal Cintas,‡,‡‡ Emmanuelle Uro-Coste,‡§§ Marie-Laure Martin Negrier,‡¶¶ Valérie Rigau,‡‖‖ Eric Bieth,‡*** Cyril Goizet,‡††† Mireille Claustres,††† Michel Koenig,*††† and Mireille Cossée*‡,‡‡‡

From the Laboratoire de Génétique Moléculaire and the Départements de Neurologie,† Biopathologie Cellulaire et Tissulaire des Tumeurs,‖‖ and Département de Neurologie Pédiatrique,§ CHU Montpellier, PhyMedExp, INSERM, CNRS, Universite de Montpellier, Montpellier; the Centre de Référence Maladies Neuromusculaires AOC‡; the Départements de Neurologie Pédiatrique,¶ Neurologie,‡‡ Anatomie et Cytologie Pathologiques,§§ and Génétique,*** CHU, Toulouse; the Services de Neurologie‖ and Pathologie¶¶ and the Département de Neurologie Pédiatrique,†† CHU, Bordeaux; the Département de Neurologie,** CHU, Nîmes; the Département de Génétique,††† Bordeaux; and the Laboratoire de Génétique de Maladies Rares,‡‡‡ Université Montpellier, Montpellier, France*

Accepted for publication April 5, 2018.

Address correspondence to Mireille Cossée, Laboratoire de Génétique Moléculaire/EA 7402, IURC, Institut Universitaire de Recherche Clinique, 641 Ave du Doyen Gaston Giraud, Montpellier Cedex 5, 34093, France. E-mail: mireille.cossee@inserm.fr.

Myopathies and muscular dystrophies (M-MDs) are genetically heterogeneous diseases, with >100 identified genes, including the giant and complex titin (*TTN*) and nebulin (*NEB*) genes. Next-generation sequencing technology revolutionized M-MD diagnosis and revealed high frequency of *TTN* and *NEB* variants. We developed a next-generation sequencing diagnostic strategy targeted to the coding sequences of 135 M-MD genes. Comparison of two targeted capture technologies (SeqCap EZ Choice library capture kit and Nextera Rapid Capture Custom Enrichment kit) and of two whole-exome sequencing kits (SureSelect V5 and TruSeq RapidExome capture) revealed best coverage with the SeqCap EZ Choice protocol. A marked decrease in coverage was observed with the other kits, affecting mostly the first exons of genes and the repeated regions of *TTN* and *NEB*. Bioinformatics analysis strategy was fine-tuned to achieve optimal detection of variants, including small insertions/deletions (INDELs) and copy number variants (CNVs). Analysis of a cohort of 128 patients allowed the detection of 52 substitutions, 13 INDELs (including a trinucleotide repeat expansion), and 3 CNVs. Two INDELs were localized in the repeated regions of *NEB*, suggesting that these mutations may be frequent but underestimated. A large deletion was also identified in *TTN* that is, to our knowledge, the first published CNV in this gene. *(J Mol Diagn 2018, ■: 1–17; https://doi.org/10.1016/j.jmoldx.2018.04.001)*

Myopathies and muscular dystrophies (M-MDs) are a set of phenotypically and genetically heterogeneous diseases, with >100 genes identified so far.[1] Massively parallel sequencing or next-generation sequencing (NGS) technologies revolutionized the field of molecular characterization of these heterogeneous disorders, because they allowed sequencing, in a single run, of all currently known genes.[2] In particular, several genes that were not currently analyzed because of their large size, such as the titin gene *TTN* (Online Mendelian Inheritance in Man no. 188840) with 363 exons and the nebulin gene *NEB* (Online Mendelian Inheritance in Man no. 161650) with 183 exons, can now be analyzed exhaustively by NGS.

Supported by Agence de la Biomédecine Appel d'Offre Recherche AMP, Diagnostic Prénatal et Diagnostic Génétique (M.C.).

Disclosures: None declared.

https://doi.org/10.1016/j.jmoldx.2018.04.001

Compared with whole-exome sequencing (WES), NGS strategies targeted to specific genes (exons and exon-intron junctions) require sequencers and computing resources of lower throughput and generate fewer variants that are restricted to phenotype-related genes and for which laboratories have expertise. Moreover, targeted strategies demonstrate better performance in terms of coverage (depth and uniformity), highlighted by the discovery of 20% to 30% additional variants in comparison with WES in a study of 177 unsolved cases of myopathies.[3] For these reasons, although targeted NGS does not allow detection of mutations in genes not previously implicated in M-MDs, this approach represents a first choice strategy in diagnostic practice.[4] Because there is no commercial kit dedicated for each specific syndrome, implementation of targeted NGS strategy in a diagnostic context requires an in-depth development step to ensure capture of regions of interest in an exhaustive manner. This can be particularly tricky in GC-rich sequences located in first exons of most genes and in repeated regions, such as the tandem repeat regions of *TTN* and *NEB*.[5,6] In *NEB*, the triplicated region of 24 exons is arranged as three sets of eight nearly perfectly repeated exons (82 to 89, 90 to 97, and 98 to 105). The *TTN* gene has three tandem repeats (exons 172 to 180, 181 to 189, and 190 to 198) sharing 99% DNA sequence homology among them.

In contrast to their high sensitivity for detection of substitutions, published bioinformatics pipelines are not fully reliable for insertion/deletion (INDEL) detection.[7] This is detrimental because INDELs are the second most common type of genomic variants after single-nucleotide substitutions,[8] in particular in nebulinopathies.[9] This warrants the use of customized tools associating several alignment and variant calling algorithms to study this class of variation. However, guidelines for optimal detection of biologically significant INDELs are limited.[7,9]

In a previous report of amplicon-based targeted NGS of the *DMD* gene (Multiplicom, Niel, Belgium) on MiSeq (Illumina, San Diego, CA), we reported that the MiSeq Reporter (MSR) software provided by the manufacturer can fail for INDEL calling.[10] SeqNext software (JSI Medical Systems, Ettenheim, Germany) is a private software that performs alignment, variant calling, and variant annotation. The combination of MSR/Variant Studio (an Illumina variant annotation software) with the SeqNext software was a reliable diagnostic strategy for the identification of single-nucleotide variations (SNVs), including INDELs, in the *DMD* gene.[10] With respect to copy number variants (CNVs), most NGS studies do not include CNV detection in their analyses, because of technological limitations, such as capture and PCR bias and reads mappability, which together drastically influence final read coverage for a particular region.[5]

The aim of the present study was to implement a comprehensive and reliable targeted NGS strategy for M-MD diagnosis, with a specific focus on the complex *NEB* and *TTN* genes. Because the clinical presentation does not always point to a specific entity (eg, congenital nemaline myopathy or congenital muscle dystrophy), an NGS strategy analyzing a large panel of genes, including not only the genes reported in the Gene Table of Neuromuscular Disorders (*http://www.musclegenetable.fr*), but also genes that can lead to atypical phenotypes mimicking M-MDs, was implemented. Two technologies that require low amount of template DNA and that present innovative aspects were chosen to compare: the capture kit SeqCap EZ Choice library, provided by Roche-NimbleGen (Madison, WI), which already demonstrated effectiveness in producing capture probes for comparative genomic hybridization[11]; and the Nextera kit Nextera Rapid Capture Custom Enrichment (NRCCE), from Illumina. An attractive aspect of the Illumina NRCCE kit is the use of a transposase catalyzing DNA fragmentation and ligation of adapter sequences in a single reaction step, referred to as tagmentation.[12,13] This technology allows a technical simplification for library preparation compared with Roche-NimbleGen technology, in which an initial step of DNA sonication is required. Because of the decrease in costs, development of new versions of WES capture kit providing better sequencing coverage,[14,15] acquisition by most molecular diagnostic laboratories of higher-throughput sequencers,[15] and publications of recommendations from international expert committees on the return to patients of incidental finding,[16,17] the relevance of WES compared with custom capture for diagnosis is still an issue. It was evaluated whether two recent WES technologies, SureSelect V5 (Agilent Technologies, Santa Clara, CA) and Illumina TruSeq RapidExome capture, could now be implemented for M-MD diagnosis.

On the basis of our experience and published data, in addition to the MSR/Variant Studio pipeline provided by Illumina, we decided to use the SeqNext software for independent alignment and calling of SNVs, to which we associated the SOPHIA Data Driven Medicine (DDM) software recently developed by Sophia Genetics (Ecublens, Switzerland). This software, already reported for molecular diagnosis,[18] has the advantage of detecting and annotating not only SNVs (including INDELs) but also CNVs. It also allows customizable filtering, the generation of a private database of variants, and comparison with variants generated by all SOPHIA DDM users. In addition to the CNV calling algorithm Muskat of the private SOPHIA DDM pipeline, an in-house read depth–based approach for CNV detection[19] previously validated in our laboratory on another NGS panel was used.[20] This approach allowed detection of 68 pathogenic or likely pathogenic mutations, including 13 INDELs and 3 CNVs.

Materials and Methods

Study Design and Patients

A specific customized panel of 135 genes (Supplemental Table S1) was designed, including 76 genes associated

FLA 5.5.0 DTD ■ JMDI694_proof ■ 23 May 2018 ■ 8:39 pm ■ EO: 2017_207