



**HAL**  
open science

# EXOME SEQUENCING AS A SECOND-TIER DIAGNOSTIC APPROACH FOR CLINICALLY SUSPECTED DYSFERLINOPATHY PATIENTS

Marc Bartoli, Jean-Pierre Desvignes, Nicolas Lévy, Martin Krahn

► **To cite this version:**

Marc Bartoli, Jean-Pierre Desvignes, Nicolas Lévy, Martin Krahn. EXOME SEQUENCING AS A SECOND-TIER DIAGNOSTIC APPROACH FOR CLINICALLY SUSPECTED DYSFERLINOPATHY PATIENTS. *Muscle & Nerve*, 2014, 50 (6), pp.1007-1010. 10.1002/mus.24344 . hal-01610020

**HAL Id: hal-01610020**

**<https://hal.science/hal-01610020>**

Submitted on 15 Dec 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **EXOME SEQUENCING AS A SECOND-TIER DIAGNOSTIC APPROACH FOR CLINICALLY SUSPECTED DYSFERLINOPATHY PATIENTS**

**MARC BARTOLI, PhD,<sup>1,2</sup> JEAN-PIERRE DESVIGNES, BSC,<sup>1</sup> NICOLAS LEVY MD, PhD,<sup>1,2</sup> and MARTIN KRAHN MD, PhD<sup>1,2</sup>**

<sup>1</sup>Aix Marseille Université, INSERM, GMGF, UMR\_S 910, 13385, Marseille, France

<sup>2</sup>APHM, Département de Génétique Médicale, Hôpital Timone Enfants, Marseille, 13385, France

**ABSTRACT:** *Introduction:* Autosomal recessive muscular dystrophies are heterogeneous genetic disorders, with 39 genes currently implicated. Genetic diagnosis using targeted single-gene analysis by Sanger sequencing yields negative results in 10–20% of samples, warranting clinical re-evaluation and time-consuming testing of additional genes. This applies to dysferlinopathies caused by mutations in the gene encoding dysferlin (*DYSF*), which presents mainly as limb-girdle muscular dystrophy (LGMD) or distal myopathy. *Methods:* We evaluated exome sequencing associated with data filtering for selected genes as a second-tier approach for genetic diagnosis in a cohort of 37 patients with an initial negative result on targeted *DYSF* analysis. *Results:* Exome sequencing allowed for establishing (16%) or suggesting (8%) the molecular diagnosis by implicating other known LGMD or distal myopathy genes or by revealing *DYSF* mutations previously missed using mutation-screening

techniques with incomplete detection yields. *Conclusions:* Exome sequencing associated with data filtering constitutes an efficient second-tier analysis for genes implicated in LGMD or distal myopathies.

**E**xome sequencing offers promising perspectives for molecular diagnosis of genetically heterogeneous diseases, in particular muscular dystrophies and distal myopathies, for which 39 different genes have been implicated to date (Gene Table of Neuromuscular Disorders; see [www.musclegenetable.fr](http://www.musclegenetable.fr)). In the transition of genetic diagnostic laboratories from Sanger sequencing to next generation sequencing techniques, exome sequencing can be considered either as a direct first tier of a sample submitted for diagnosis,<sup>1</sup> or after an initial negative result on targeted single-gene analysis. In autosomal recessive muscular dystrophies, the latter accounts for 10–20% of samples,<sup>2</sup> which leads to clinical re-evaluation and time-consuming testing of additional genes. This applies to dysferlinopathies caused by mutations in *DYSF* (MIM #603009), which present mainly as a limb-girdle muscular dystrophy (LGMD) or a distal myopathy (Miyoshi myopathy).<sup>3</sup>

We evaluated exome sequencing associated with data filtering for selected genes as a second-tier approach for genetic diagnosis in patients with an initial negative result on targeted *DYSF* analysis.

## METHODS

From among the largest international cohort of *DYSF* mutation analyses,<sup>4</sup> we included 37 samples that were assessed initially for diagnostic suspicion of dysferlinopathy and for which no *DYSF* disease-causing mutation had been identified. Samples were submitted to exome sequencing using SureSelect Human All Exon Kit version 4 (Agilent Technologies, Santa Clara, California) and Illumina HiSeq 2000 (Illumina, San Diego, California) systems. Sequencing data were processed using GATK<sup>5</sup> variant calling and ANNOVAR<sup>6</sup> annotation, then filtered through the list of genes in the Gene Table of Neuromuscular Disorders, specifically disease groups “muscular dystrophies” and “distal myopathies” (total of 39 genes; updated 2013). All disease-causing mutations identified using exome sequencing were confirmed using direct targeted sequencing on a genetic analyzer (Model 3500XL; Applied Biosystems—Life Technologies, Carlsbad, California). For patients P6 – P9, all coding exons and intronic boundaries of the genes *SGCA*, *GNE*, and *DYSF* were also analyzed respectively using targeted sequencing on the aforementioned genetic analyzer to identify a possible second disease-causing mutation missed by the exome sequencing.

## RESULTS

Data analysis showed a mean overall sequencing depth of 73× and a mean overall coverage of the targeted coding exons of 87%, 82%, 71%, and 63% at respective sequencing depths of 1×, 5×, 20×, and 30×. Accordingly, the 39 genes from the Gene Table of Neuromuscular Disorders, disease groups “muscular dystrophies” and “distal myopathies,” were only partially covered at a sequencing depth of 20×, which is considered adequate for detection of heterozygous germline mutations (Supplementary Table S1, available online).

In 5 samples (patients P1 – P5), molecular diagnoses could be established directly after exome sequencing by identifying compound heterozygous or homozygous disease-causing mutations in previously known LGMD or distal myopathy genes (Table 1). These corresponded to *CAPN3* mutations in 2 cases, *ANO5* mutation in 1 case, *GNE* mutation in 1 case, and *DYSF* mutation in 1 case. In the latter case, the homozygous *DYSF* mutation was most likely missed due to incomplete mutation yields using screening techniques (mainly single-strand conformational polymorphism analysis) used initially in our laboratory,<sup>7</sup> or other problems, such as allele drop-out, underlining how molecular re-evaluation using exome sequencing could be useful for reaching a definitive genetic diagnosis in such situations. The identified mutation in patient P2 has not been reported previously; its pathogenicity is implied, as it would lead to a frameshift of the reading frame with the introduction of a premature translation termination codon.

In 4 samples, exome sequencing identified only 1 disease-causing mutation in previously known LGMD or distal myopathy genes, respectively, in the following genes: *SGCA* for patient P6; *GNE* for patient P7; and *DYSF* for patients P8 and P9 (Table 1). In patients P7 and P8, the identified mutations have not been reported previously, but their pathogenicity is implied with the following rationale: In patient P7, the identified missense mutation affects a highly conserved amino acid residue, with its pathogenic effect predicted by PolyPhen-2 analysis (score = 0.996); in patient P8, the identified abnormality is a nonsense mutation that introduces a premature translation termination codon. For patients P6 – P9, all the coding exons and intronic boundaries of the respective genes were analyzed additionally using targeted direct sequencing to identify a possible second disease-causing mutation missed by the exome sequencing.

For patient P6, molecular diagnosis could be completed by identifying a second expected disease-causing mutation in *SGCA* using additional Sanger sequencing of the gene given the exome

**Table 1.** Disease-causing mutations identified in patients using exome and additional Sanger sequencing.

Patient ID	Mutated LGMD or distal myopathy gene and corresponding disease-causing mutation(s) identified using exome sequencing	Additional disease-causing mutation(s) identified in the respective genes using Sanger sequencing
P1	<i>CAPN3</i> (NM_000070) Exon 6: c.848T>C (p.Met283Thr) HTZ Previously reported* and Exon 13: c.1611C>A (p.Tyr537X) HTZ Previously reported*	Not applicable
P2	<i>CAPN3</i> (NM_000070) Exon 13: c.1642delC (p.Arg548A-lafsX47) HTZ Not previously reported* and Exon 16: c.1817C>T (p.Ser606Leu) HTZ Previously reported*	Not applicable
P3	<i>ANO5</i> (NM_213599) Exon 5: c.191dupA (p.Asn64-LysfsX15) HTZ Previously reported* and Exon 16: c.1733T>C (p.Phe578Ser) HTZ Previously reported*	Not applicable
P4	<i>GNE</i> (NM_001128227) Exon 10: c.1853T>C (p.Ile618Thr) HOZ Previously reported*	Not applicable
P5	<i>DYSF</i> (NM_003494) Exon 34: c.3708delA (p.Asp1237ThrfsX24) HOZ Previously reported*	Not applicable
P6	<i>SGCA</i> (NM_000023) Exon 6: c.739G>A (p.Val247Met) HTZ Previously reported*	<i>SGCA</i> (NM_000023) Exon 5: c.403C>T (p.Gln135X) HTZ Previously reported*
P7	<i>GNE</i> (NM_001128227) Exon 7: c.1361T>G (p.Ile454Arg) HTZ Not previously reported* PolyPhen-2 predictive analysis: probably damaging (score 0.996)	No additional disease-causing mutation identified
P8	<i>DYSF</i> (NM_003494) Exon 30: c.3207G>A (p.Trp1069X) HTZ Not previously reported*	No additional disease-causing mutation identified
P9	<i>DYSF</i> (NM_003494) Exon 47: c.5302C>T (p.Arg1768Trp) HTZ Previously reported*	No additional disease-causing mutation identified

HTZ, heterozygous; HOZ, homozygous.

\*Previously reported and not previously reported disease-causing mutations with regard to the Leiden Muscular Dystrophy Pages ([www.dmd.nl](http://www.dmd.nl)).

<sup>†</sup>PolyPhen-2 predictive analysis as reported by Adzhubei et al.<sup>9</sup>

sequencing results. This variant had not been called due to a number of reads below the 5-reads threshold. Only 3 reads have been mapped to the mutated region as visualized using the Integrated Genome Viewer<sup>8</sup> (Fig. S1 in Supplementary Material, available online).

For patients P7 – P9, no second disease-causing mutation could be identified. In these cases, mutations such as exonic deletions/amplifications or mutations in deep intronic and/or regulatory regions may have been missed. Widening of the informatics filter to the entire Gene Table of Neuromuscular Disorders list (297 genes) for patients P7 – P9 did not identify any other mutational data correlating with their clinical phenotype. At the clinical level, patients P8 and P9 presented with LGMD and Miyoshi myopathy, respectively, and dysferlin protein deficiency had been identified initially by immunohistochemical analysis on muscle biopsy samples, which correlates clinically with the mutational findings. Patient P7 presented with distal myopathy at the clinical level. Of note, patient P7 carries an additional heterozygous missense variant at exon 12 of *GNE* (c.2030G>C; p.Gly677Ala) not previously reported in the Leiden Muscular Dystrophy Pages database ([www.dmd.nl](http://www.dmd.nl)) and not reported to date in dbSNP138 ([www.ncbi.nlm.nih.gov/SNP/](http://www.ncbi.nlm.nih.gov/SNP/)).

A possibly damaging effect at the amino acid level is not likely, as PolyPhen-2 analysis<sup>9</sup> predicts a non-pathogenic effect of the induced amino acid change, yet with a possible deleterious effect on mRNA splicing predicted by activation of a cryptic splice-acceptor site (Human Splicing Finder analysis).<sup>10</sup> The latter hypothesis could not yet be confirmed due to lack of available muscle tissue.

## DISCUSSION

With the increasing accuracy and decreasing costs of next generation sequencing techniques, first-tier exome sequencing associated with data filtering will represent an efficient alternative strategy,<sup>11</sup> as compared with Sanger sequencing, in selected single genes. Although next generation sequencing techniques may soon become the standard initial analysis for samples submitted for diagnosis of myopathies, diagnostic laboratories must also consider the problem of accumulated samples initially submitted for targeted single-gene analyses that tested negative and for which additional testing may be initiated after discussion with the patient and referring physician. Moreover, a second-tier approach applied to different heterogeneous groups of myopathies after an initial negative

result of a targeted single-gene analysis should help clarify the differential diagnosis. We evaluated the utility of second-tier exome sequencing associated with data filtering after an initial negative result on targeted single-gene analysis of the *DYSF* gene, one of the most frequently implicated genes in LGMD and distal myopathies. Negative results after targeted single-gene analysis are a recurrent problem in routine genetic diagnosis practice for autosomal recessive muscular dystrophies. Altogether, 10–20% of samples remain without accurate diagnosis after targeted single-gene analyses,<sup>2</sup> due to various causes, including genetic heterogeneity and limitations in sensitivity depending on the techniques used.

In our cohort of 37 patients with an initial negative result of targeted *DYSF* analysis, exome sequencing allowed for molecular diagnosis in 6 cases (16%), with 1 case requiring additional Sanger sequencing to identify the second mutated allele. In 3 cases (8%), exome sequencing was only suggestive of a molecular diagnosis, given the absence of a second identifiable disease-causing mutation. The sequencing abnormalities identified belonged to other known LGMD or distal myopathy genes, including *DYSF* mutations previously missed in samples analyzed using mutation screening techniques with incomplete mutation detection rates.<sup>7</sup>

For the remaining 28 samples, widening of the informatics filter to the entire Gene Table of Neuro-muscular Disorders list (297 genes) identified several mutation candidates, which are under investigation.

Taken together, our results underline the utility of second-tier exome sequencing associated with data filtering for reaching an accurate genetic diagnosis in a subset of patients with an initial negative result on a targeted single-gene analysis of

the *DYSF* gene. Moreover, the incomplete diagnostic yield points to further genetic heterogeneity in patients who present clinically with dysferlinopathy. The increase of targeted sequence coverage available through updated commercial and/or customized exome capture enrichment kits should allow for progressive gains in sensitivity of this strategy for diagnostic applications in myopathies.

The authors sincerely thank Karine Bertaux, Cécile Mouradian, Véronique Blanck, Eric Salvo, Brigitte Le Goanvic, Valérie Delague, David Salgado, Christophe Bérout, Anne Boland, and Mark Lathrop for their contributions to this work. We also thank the patients and their referring physicians for their participation.

## REFERENCES

1. Dias C, Sincan M, Cherukuri PF, Rupps R, Huang Y, Briemberg H, et al. An analysis of exome sequencing for diagnostic testing of the genes associated with muscle disease and spastic paraplegia. *Hum Mutat* 2012;33:614–626.
2. Bushby K. Diagnosis and management of the limb girdle muscular dystrophies. *Pract Neurol* 2009;9:314–323.
3. Úrtizberea JA, Bassez G, Leturcq F, Nguyen K, Krahn M, Levy N. Dysferlinopathies. *Neurol India* 2008;56:289–297.
4. Krahn M, Beroud C, Labelle V, Nguyen K, Bernard R, Bassez G, et al. Analysis of the *DYSF* mutational spectrum in a large cohort of patients. *Hum Mutat* 2009;30:E345–375.
5. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491–498.
6. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
7. Nguyen K, Bassez G, Bernard R, Krahn M, Labelle V, Figarella-Branger D, et al. Dysferlin mutations in LGMD2B, Miyoshi myopathy, and atypical dysferlinopathies. *Hum Mutat* 2005;26:165.
8. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol* 2011;29:24–26.
9. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–249.
10. Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 2009;37:e67.
11. Neveling K, Feenstra I, Gilissen C, Hoefsloot LH, Kamsteeg EJ, Mensenkamp AR, et al. A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. *Hum Mutat* 2013;34:1721–1726.