Available online at www.sciencedirect.com

ScienceDirect

Neuromuscular Disorders 25 (2015) 502–510

www.elsevier.com/locate/nmd

Targeted next-generation sequencing for the genetic diagnosis of dysferlinopathy

Ha Young Shin [a,1], Hoon Jang [b,1], Joo Hyung Han [c], Hyung Jun Park [d], Jung Hwan Lee [a], So Won Kim [e,f], Seung Min Kim [a], Young-Eun Park [g], Dae-Seong Kim [g], Duhee Bang [b], Min Goo Lee [c], Ji Hyun Lee [h,*], Young-Chul Choi [a,**]

[a] *Department of Neurology, Yonsei University College of Medicine, Seoul, Republic of Korea*
[b] *Department of Chemistry, Yonsei University, Seoul, Republic of Korea*
[c] *Department of Pharmacology, Pharmacogenomic Research Center for Membrane Transporters, Brain Korea 21 PLUS Project for Medical Sciences, Severance Biomedical Science Institute, Yonsei University College of Medicine, Seoul, Republic of Korea*
[d] *Department of Neurology, Mokdong Hospital, Ewha Womans University College of Medicine, Seoul, Republic of Korea*
[e] *Department of Pharmacology and PharmacoGenomics Research Center, Inje University College of Medicine, Busan, Republic of Korea*
[f] *Department of Clinical Pharmacology, Inje University, Busan Paik Hospital, Busan, Republic of Korea*
[g] *Department of Neurology, Pusan National University School of Medicine, Busan, Republic of Korea*
[h] *Department of Oral Biology, Yonsei University College of Dentistry, Seoul, Republic of Korea*

Received 4 December 2014; received in revised form 16 February 2015; accepted 11 March 2015

Abstract

Dysferlinopathy comprises a group of autosomal recessive muscular dystrophies caused by mutations in the *DYSF* gene. Due to the large size of the gene and its lack of mutational hot spots, analysis of the *DYSF* gene is time-consuming and laborious using conventional sequencing methods. By next-generation sequencing (NGS), *DYSF* gene analysis has previously been validated through its incorporation in multi-gene panels or exome analyses. However, individual validation of NGS approaches for *DYSF* gene has not been performed. Here, we established and validated a hybridization capture-based target-enrichment followed by next-generation sequencing to detect mutations in patients with dysferlinopathy. With this approach, mean depth of coverage was approximately 450 fold and almost all (99.3%) of the targeted region had sequence coverage greater than 20 fold. When this approach was tested on samples from patients with known *DYSF* mutations, all known mutations were correctly retrieved. Using this method on 32 consecutive patient samples with dysferlinopathy, at least two pathogenic variants were detected in 28 (87.5%) samples and at least one pathogenic variant was identified in all samples. Our results suggested that the NGS-based screening method could facilitate efficient and accurate genetic diagnosis of dysferlinopathy.

Keywords: Dysferlinopathy; *DYSF*; Next-generation sequencing; Mutation; Hybridization capture

1. Introduction

Dysferlinopathy comprises a group of autosomal recessive muscular dystrophies caused by mutations in the *DYSF* gene [dysferlin, Mendelian inheritance in man (OMIM *603009], located on chromosome 2p13 [1,2]. In dysferlinopathy, genotype–phenotype correlations are not clear. *DYSF* mutations can lead to various clinical phenotypes, such as limb-girdle muscular dystrophy (LGMD) type 2B, Miyoshi myopathy (MM), a proximodistal phenotype, distal anterior compartment myopathy (DACM), a pseudometabolic form, and asymptomatic hyper-CK-emia [3]. Furthermore, identical mutations in *DYSF* can manifest as different phenotypes, even within a single family [4,5]. A loss or decrease of dysferlin expression is a key finding of muscle immunostaining [immunohistochemical (IHC) staining] for the diagnosis of dysferlinopathy. However, a reduction in dysferlin protein on IHC staining may be seen in other myopathies, including calpainopathy, sarcoglycanopathy, and dystrophinopathy [6,7]. Thus, identification of the precise *DYSF* mutations is necessary for an accurate diagnosis of dysferlinopathy.

Sanger sequencing is considered the "gold standard" for mutation screening. However, using conventional sequencing

* Corresponding author. Department of Oral Biology, Yonsei University College of Dentistry, 50 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Korea. Tel.: +82 2 2228 3046; fax: +82 2 313 1864.
E-mail address: jihyni@yuhs.ac (J.H. Lee).
** Corresponding author. Department of Neurology, Yonsei University College of Medicine, 211 Eonju-ro, Gangnam-gu, Seoul 135-720, Korea. Tel.: +82 2 2019 3323; fax: +82 2 3462 5904.
E-mail address: ycchoi@yuhs.ac (Y.-C. Choi).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.nmd.2015.03.006
0960-8966/

methods, mutation analysis of the *DYSF* gene is a time-consuming and challenging task, because of the large size of the *DYSF* gene and the absence of mutational hot spots. The *DYSF* gene can express 14 mRNA isoforms, contains 58 coding exons, and spans over 233 kb of genomic DNA (UCSC: http://genome.ucsc.edu/, chr2:71,680,753–71,913,893). Next-generation sequencing (NGS) technologies make it possible to generate large amounts of sequence data rapidly and at a lower cost than older DNA sequencing technologies. Therefore, dysferlinopathy is a good candidate disorder for the application of NGS technologies.

High-throughput NGS-based sequencing technologies have been used to identify mutations in several genetic disorders and have shown clinical utility for genetic diagnoses [8]. Recently, a multi-gene NGS panel for congenital muscular dystrophy (CMD) was validated, demonstrating a high efficiency for detecting mutations and showing a better diagnostic yield and cost-effectiveness than conventional Sanger sequencing of associated genes [9]. However, among the 20 patients with clinically suspected CMDs, only five (25%) patients were able to get a definitive genetic diagnosis (at least one rare homozygous or two heterozygous variants in autosomal recessive disorders) by the panel. In this study, the multi-gene NGS panel was applied to CMDs patients without clinical certainty, though CMDs are genetically and phenotypically heterogeneous disorders. In addition, the panel incorporated 12 known CMD associated genes without validation of each gene in its associated disorder. Although the clinical use of NGS technologies is rapidly moving to large multi-gene targeted panels for heterogeneous group of disorders [8], there is a lack of studies that validate NGS performance for each gene in the homogeneous group of patients with its associated genetic disorder. Using NGS, the analysis of the *DYSF* gene has been applied through its incorporation in multi-gene panel or exome analyses [10–12]. However, the subjects of the previous studies were heterogeneous groups of myopathies. The validation of NGS performance has not been conducted in a homogeneous group of patients with dysferlinopathy.

In this study, we applied and validated a targeted NGS based sequencing method for mutation detection in patients with dysferlinopathy, which was confirmed by IHC staining and/or western blot analysis of skeletal muscles.

2. Patients and methods

2.1. Patients

After clinical records and muscle histopathologies were reviewed for 559 myopathy patients in our institute database, 41 patients with dysferlinopathy (24 males/17 females) were enrolled in this study. All patients showed dystrophic changes on muscle biopsy and a selective loss or lack of dysferlin on IHC staining and/or western blot analysis of skeletal muscles. The clinical features of patients and the results of IHC and western blot analysis are summarized in Supplementary Table S1. Among the 41 patients, 9 had their *DYSF* mutations confirmed by previous analysis [13]. Written informed consent was obtained from all patients. This study was approved by the Institutional Review Board of Severance Hospital, Yonsei University College of Medicine, Seoul, Korea (IRB No.: 3-2011-0270).

2.2. Muscle pathology and protein analysis

Muscle biopsies were performed on all patients. Frozen muscle sections were processed for routine histochemical staining, including hematoxylin and eosin, modified Gomori trichrome, reduced nicotinamide adenine dinucleotide-tetrazolium reductase, and adenosine triphosphatase, as well as for immunohistochemistry using a set of antibodies against skeletal muscle proteins including dysferlin (Leica Biosystems, Newcastle Upon Tyne, UK), as previously described [14]. IHC staining patterns were classified into four groups: 'Normal,' 'Negative,' 'Faint,' and 'Abnormal,' as previously described [7]. Western blot analysis was performed by standard methods using the same anti-dysferlin antibody used for IHC staining. We defined a western blot signal as 'Deficient' when the amount of dysferlin protein was less than 30% of the positive control sample [7].

2.3. Targeted sequencing

Genomic DNA from muscle or blood was extracted with Qiagen DNeasy blood & tissue kit (Qiagen, Valencia, CA). The DNA samples used in this study were divided into two groups. The first group consisted of nine patients with previously characterized mutations in the *DYSF* gene [13]. The second group comprised 32 patient samples, which were supplied for molecular diagnostic testing but had not been previously analyzed.

For mutation analysis, *DYSF* gene-containing DNA fragments were enriched by solution-based hybridization capture and followed by sequencing with an Illumina Hiseq2500 platform (Illumina, San Diego, CA, USA) with the 2×150 bp paired-end read module. The target region included all coding exons and flanking intron regions of the *DYSF* gene. Capture probes were generated by Celemics, Inc. (Seoul, Korea). The hybridization capture procedure was also performed at Celemics, Inc. according to the manufacturer's standard protocol. One μg of genomic DNA was sheared via sonication using an EpiSonic™ 1100 sonicator. Biotynilated RNA oligonucleotide baits were hybridized with sheared DNA. Captured fragments were removed from solution via streptavidin-coated magnetic beads (Dynabeads® MyOne™ Streptavidin T1, Life Technologies) and subsequently eluted. The enriched fragment library was then subjected to PCR amplification using primers specific to the linked Illumina adaptors. Resulting libraries were quantified via Agilent 2200 TapeStation before proceeding to Illumina Hiseq2500 platform. All samples were pooled into a single lane on a flow cell and sequenced together. For differentiation between samples, a six-bp index sequence (Illumina) and a six-bp in-house barcode sequence were used.

2.4. Variant analysis

From the HiSeq2500 raw data, sequencing reads from each sample were sorted by index and barcode sequences. Sorted fastq files were aligned to the hg19 reference genome using the Burrows–Wheeler Aligner (BWA; ver. 0.7.5a) algorithm [15]. Output SAM files were converted into BAM files and sorted using SAMtools (ver. 0.1.18) [16]. Duplicate removal was performed