Journal of Human Genetics (2016) 61, 497–505

www.nature.com/jhg

ORIGINAL ARTICLE

New massive parallel sequencing approach improves the genetic characterization of congenital myopathies

Jorge Oliveira[1,2], Ana Gonçalves[1,2], Ricardo Taipa[3], Manuel Melo-Pires[3], Márcia E Oliveira[1,2], José Luís Costa[4,5], José Carlos Machado[4,5], Elmira Medeiros[6], Teresa Coelho[7], Manuela Santos[8], Rosário Santos[1,2,9,12] and Mário Sousa[2,10,11,12]

Congenital myopathies (CMs) are a heterogeneous group of muscle diseases characterized by hypotonia, delayed motor skills and muscle weakness with onset during the first years of life. The diagnostic workup of CM is highly dependent on the interpretation of the muscle histology, where typical pathognomonic findings are suggestive of a CM but are not necessarily gene specific. Over 20 loci have been linked to these myopathies, including three exceptionally large genes (*TTN*, *NEB* and *RYR1*), which are a challenge for molecular diagnosis. We developed a new approach using massive parallel sequencing (MPS) technology to simultaneously analyze 20 genes linked to CMs. Assay design was based on the Ion AmpliSeq strategy and sequencing runs were performed on an Ion PGM system. A total of 12 patients were analyzed in this study. Among the 2534 variants detected, 14 pathogenic mutations were successfully identified in the *DNM2, NEB, RYR1, SEPN1* and *TTN* genes. Most of these had not been documented and/or fully characterized, hereby contributing to expand the CM mutational spectrum. The utility of this approach was demonstrated by the identification of mutations in 70% of the patients included in this study, which is relevant for CMs especially considering its wide phenotypic and genetic heterogeneity.
Journal of Human Genetics (2016) **61,** 497–505; doi:10.1038/jhg.2016.2; published online 4 February 2016

INTRODUCTION

Congenital myopathies (CMs) are a highly heterogeneous and continuously expanding, group of muscle diseases with an estimated incidence of around 6 per 100 000 live births.[1] On the severest end of the disease spectrum, CMs are characterized by muscle weakness and hypotonia with neonatal onset or during the first years of life, which often give rise to respiratory and feeding difficulties.[1,2] Typical features may include weakness of facial and bulbar muscles, and also involvement of extraocular muscles (ophthalmoparesis or ophthalmoplegia).[3] Typical forms of CM may also present during childhood as motor development delay and/or waddling gait.[1] At the other end of the scale, it is now clear that some patients with subtle congenital muscle weakness might remain undiagnosed until adulthood, when muscle strength deteriorates or respiratory insufficiency settles in.[1,4]

The classification and the diagnostic workup of CMs are highly reliant upon the identification of distinct structural abnormalities detected in muscle biopsies.[3,5] Based on pathognomonic findings CMs are subdivided into four major groups: (i) CMs with rods in nemaline myopathy (NM), (ii) CMs with cores, which includes central core disease, multiminicore disease and multicore myopathy, (iii) CMs with central nuclei (centronuclear myopathy) and (iv) congenital fiber-type size disproportion.[5] Other less common subtypes such as cap or core-rod myopathies, which might represent overlapping myopathological entities, have been also reported.[6,7] Although these findings are of extreme diagnostic value, in some patients the muscle biopsy may display only unspecific myopathic features such as type 1 fiber predominance/uniformity.[8]

More than 20 different genes are currently known to be associated with CMs (http://musclegenetable.fr/, August 2015). The majority of these loci encode basic structural components of the sarcomere or proteins involved in calcium homeostasis, both crucial for normal muscle structure and function.[1] Other genetic defects give rise to abnormal triad structure due to aberrant tubulogenesis and/or abnormal membrane recycling.[9] Among these loci there are three particularly large genes: Titin (*TTN*) with 363 exons, Nebulin (*NEB*)

[1]Unidade de Genética Molecular, Centro de Genética Médica Dr Jacinto Magalhães, Centro Hospitalar do Porto, Porto, Portugal; [2]Unidade Multidisciplinar de Investigação Biomédica (UMIB), Instituto de Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto, Porto, Portugal; [3]Unidade de Neuropatologia, Centro Hospitalar do Porto, Porto, Portugal; [4]Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Porto, Portugal; [5]Faculdade de Medicina da Universidade do Porto, Porto, Portugal; [6]Departamento de Neurologia, Hospital Egas Moniz, Centro Hospitalar de Lisboa Ocidental, Lisboa, Portugal; [7]Unidade Clínica de Paramiloidose, Centro Hospitalar do Porto, Porto, Portugal; [8]Consulta de Doenças Neuromusculares, Serviço de Neuropediatria, Centro Hospitalar do Porto, Porto, Portugal; [9]UCIBIO/REQUIMTE, Departamento de Ciências Biológicas, Laboratório de Bioquímica, Faculdade de Farmácia, Universidade do Porto, Porto, Portugal; [10]Departamento de Microscopia, Laboratório de Biologia Celular, Instituto de Ciências Biomédicas Abel Salazar (ICBAS), Universidade do Porto, Porto, Portugal and [11]Centro de Genética da Reprodução Prof. Alberto Barros, Porto, Portugal
[12]These authors contributed equally to this work.
Correspondence: Dr R Santos, Unidade de Genética Molecular, Centro de Genética Médica Dr Jacinto de Magalhães, Centro Hospitalar do Porto (CHP), Praça Pedro Nunes, 88, Porto 4099-028, Portugal.
E-mail: rosario.santos@chporto.min-saude.pt
Received 9 September 2015; revised 28 November 2015; accepted 5 January 2016; published online 4 February 2016

with 183 and the Ryanodine receptor 1 (*RYR1*) with 106. Consequently, conventional Sanger sequencing of these genes is extremely laborious and costly. These aspects might explain the lack of thorough studies for large genes such as *TTN* and the limited number of patients reported in the literature.[10] The genetic study workup is complex considering that the same clinical entity can be caused by mutations in different genes, as may the same defective gene give rise to distinct myopathies. *RYR1*-related myopathies are among the best examples. Nearly all subtypes of CMs (central core disease, centronuclear myopathy, congenital fiber-type size disproportion and core-rod myopathy) have been reported as linked to *RYR1* mutations.[3,4] In addition, although muscle histology is important for the diagnostic workup of CMs, pathognomonic findings (such as rods in NM) are not necessarily gene specific.[11] The considerable genetic and clinical heterogeneity is reflected by the large number of genetic studies reported to be inconclusive. Essentially owing to these challenging diagnostic and technical difficulties, a significant proportion of CM patients are still genetically unsolved.[2]

With the advent of the novel massive parallel sequencing (MPS) technologies, we have developed a new targeted resequencing approach, which allows the simultaneous analysis of 20 genes implicated in CMs. Besides contributing to expand the mutational spectrum of CMs, our study enabled the identification of mutations in 7 CM patients (out of 10 who were prospectively studied), thereby demonstrating its clinical utility.

MATERIALS AND METHODS

Patients

A total of 12 patients were included in this study. Two patients with mutations previously identified by conventional sequencing were used for assay validation. The first patient used as a positive control (C1) was diagnosed with a centronuclear myopathy due to a heterozygous mutation (c.1393C>T, p.Arg465Trp) in the *DNM2* gene. The second patient (C2) has a mild core myopathy with adulthood onset, caused by two heterozygous *RYR1* missense mutations (patient 1 in Duarte *et al.*[12]).

For the prospective part of this study, 10 genetically uncharacterized patients (P1–P10) were selected based on features compatible with CMs, namely (i) early disease onset (neonatal or up to early childhood) and (ii) muscle histology suggestive of CMs or with structural defects. Eight of these patients had structural changes on muscle biopsy: three with central nuclei (P1B, P2 and P8), three with rods (P3, P5 and P7) and two with minicores (P4 and P6). The remaining two patients (P9 and P10) had unspecific myopathic changes seen on muscle biopsy, mainly type 1 fiber predominance. Parents and other family members, when available, were studied to demonstrate compound heterozygosity in autosomal recessive cases and/or to ascertain the parental origin of the mutations. This research work was approved by the ethics committee of Centro Hospitalar do Porto.

Muscle histology analysis

The routine diagnostic workup consisted of an open biopsy (deltoid, quadriceps or gastrocnemius muscle) and histological (hematoxylin and eosin stain, periodic acid-Schif and gomori trichrome) and histochemical (reduced NADH, succinic dehydrogenase, cytochrome oxidase and adenosine triphosphatase) studies performed on frozen material. Semi-thin sections at 1 μm from resin embedded muscle were also routinely used.

AmpliSeq assay design

The first task consisted in the development of a targeted resequencing approach based on semiconductor technology to simultaneous analyze 20 genes linked to CMs, namely *ACTA1, BIN1, CFL2, CNTN1, DNM2, KBTBD13, KLHL40, MEGF10, MTM1, MYBPC3, MYH2, MYH7, NEB, RYR1, SEPN1, TNNT1, TPM2, TPM3, TRIM32* and *TTN* (Supplementary Data I). Assay design was performed using the Ion AmpliSeq Designer software (Life Technologies, Foster City, CA, USA) pipeline 2.2.1. Based on a list of representative transcripts for these loci (Supplementary Data I), target regions were selected from the University of California Santa Cruz, California, USA (UCSC) Genome Browser (http://genome.ucsc.edu/index.html). Targets included exonic regions (with 50 bp into flanking introns) and untranslated regions. The obtained custom Ampliseq assay covered 92.6% of these regions comprising a total of ~320 thousand base pairs (bp) and consisted of 2077 different amplicons (ranging from 64 to 239 bp, 174 bp on average), divided into two independent primer pools for multiplex PCR.

Library and template preparation

Sample quality of patient genomic DNA was evaluated by gel electrophoresis and quantified using Qubit dsDNA HS Assay Kit (Life Technologies). A total of 10 ng of genomic DNA were used in each multiplex PCR reaction. Library amplifications, digestion, bar-coding and purification was performed according to the Ion AmpliSeq Library Kit version 2.0 (Life Technologies) instructions. Library quantification was performed using the Qubit dsDNA HS Assay. All libraries were diluted to the same concentration and pooled to ensure an equal representation of the different samples.

The diluted and combined libraries were subjected to amplification by emulsion PCR using Ion Personal Genome Machine (PGM) Template OT 2 200 Kit (Life Technologies) and prepared on an Ion OneTouch 2 Instrument (Life Technologies) according to the manufacturer's protocol. Enrichment of template Ion Sphere particles was performed using the Ion OneTouch 2 enrichment system (Life Technologies).

Semiconductor sequencing

Sequencing was carried out on a PGM system platform based on semiconductor technology.[13] Two independent experiments were performed, using the Ion 316 and 318 chips, allowing four or eight bar-coded samples, respectively. The Ion Sequencing Kit v2.0 (Life Technologies) was used to perform sequencing runs, following the manufacturer's recommended protocols. Torrent Server version 4.0.2 was used to obtain the basic run metrics and to generate Binary Alignment Map and Variant Caller Format files. Torrent Variant Caller was set for germline – low stringency – calls.

Data analysis and interpretation

Variant annotation and filtering steps were performed using Ion Reporter v4.2 software (http://ionreporter.lifetechnologies.com/) (Life Technologies). Variant Caller Format files from each patient were added to Ion Reporter and were annotated using the 'annotate variants single sample' workflow incorporated in this software. A filter chain was created to restrict the number of variants, which consisted in the removal of single-nucleotide polymorphisms (SNPs) present on the UCSC common SNPs list (those with at least 1% minor allele frequency). Alamut Visual v2.4 software (Interactive Biosoftware, Rouen, France) was used to assist variant analysis and interpretation. The software also incorporates algorithms to evaluate the impact of missense mutations and the effect on splicing, and allows the simultaneous visualization of Variant Caller Format and Binary Alignment Map files. GenomeBrowse v2.0.0 (Golden Helix, Bozeman, MT, USA) was used for visual inspection of Binary Alignment Map files.

Variant validation and expression analysis

Clinically relevant variants (classified as pathogenic) were confirmed by a new PCR using in-house designed primers (sequences in Supplementary Data II) and the resulting amplicons sequenced by the Sanger method.

Variants possibly affecting splicing were evaluated by expression studies at the mRNA level. Total RNA was obtained from cryopreserved muscle samples (P4 and P7) using the PerfectPure RNA Fibrous Tissue kit (5 PRIME, Hilden, Germany) or from whole blood (patient P3) extracted with the PerfectPure RNA Blood Kit (5 PRIME). After conversion to cDNA with the High Capacity cDNA Reverse Transcription kit (Life Technologies), these samples were amplified by PCR using specific primers annealing to different regions of *TTN* and *NEB* transcripts (primers sequences in Supplementary Data II). PCR products were purified with Illustra ExoStar (GE Healthcare, Little Chalfont, UK). A new PCR was prepared with BigDye Terminator kit V3.1 chemistry

(Life Technologies). Sequencing reactions were run on an ABI 3130xl genetic analyser (Life Technologies).

RESULTS

Development of a new sequencing approach for CMs

The initial task was the development of a new gene panel based on MPS, designed using the Ion AmpliSeq software. An overall coverage of 95% was obtained for the genomic coding sequence (CDS) that was considered suitable for this research. The CDS coverage of each locus ranged between 77 and 100% with the majority of the genes ($n=15$) having a coverage above 90% (Supplementary Data I). For the remaining loci, all except the *NEB* gene could be explained by the presence of high GC content regions, having higher impact in smaller genes. These regions are amenable to filling in by Sanger sequencing. In the case of *NEB* the lower coverage (85%) is explained by the existence of a triplicated region (three almost identical sequence blocks) encompassing exons 82–89, 90–97 and 98–105. Since the majority of MPS approaches are based on the alignment of short sequences, the obtained reads and variants will not be mapped with precision.

The sequencing runs performed on the total of 12 samples generated 6.5 million sequence reads, with an average coverage depth of 257*x* (Supplementary Data III) and with 97.0% of the target regions successfully covered. A total of 2535 sequence variants were called: 2348 single-nucleotide variants and 187 insertion/deletions (INDEL), corresponding to an average of 211.4 variants per individual (Supplementary Data III).

The high number of variants obtained in this work required the development of a computational filtering strategy to facilitate the analytic process and to concentrate on the variants that most likely correlated with the patients' phenotypes. Single-nucleotide variants listed as common SNPs in the UCSC database were excluded from further analysis. As a result, 86% of the detected variants were filtered out, meaning that an average of ~29 variants per patient remained to be analyzed in detail. Data interpretation implied evaluating: (i) the variant predictable impact, (ii) the frequency in publically available genetic variant databases (dbSNP and exome variant server (EVS)) and (iii) the data available in locus specific databases (Leiden Muscular Dystrophy Pages).

A total of 14 pathogenic variants listed in Table 1 were detected in nine patients whose clinical data are shown in Table 2. The three missense mutations previously identified by conventional sequencing in the two positive controls (C1 and C2) were successfully detected. The remaining 11 mutations were identified in seven patients of the prospective study; these were located in *RYR1* ($n=4$), *NEB* ($n=4$) and *TTN* ($n=3$). One additional heterozygous mutation in the selenoprotein N 1 gene (*SEPN1*) was also identified. In this case, and so as to exclude the presence of a second pathogenic variant, this gene was completely resequenced by the Sanger method. In one of the patients (P2) with a novel heterozygous nonsense mutation in *RYR1*, gene analysis was complemented by conventional sequencing, leading to the identification of a second heterozygous mutation located in a region not covered by the assay (part of exon 101).

In this work, a total of 67 variants (the majority intronic or located in untranslated regions) were found to be unclassified variants; these included both ultra-rare sequence changes (reported in genetic databases with a frequency bellow 0.1%) and novel variants. Finally, 43 known polymorphisms were also identified, reported in the EVS database with frequencies higher than 0.1% and thus not listed in the USCS common SNP table.

Difficulties and potential pitfalls

During the validation of the MPS gene panel for CMs, the zygosity of one of the mutations (chr19:g.38986918C>G, c.6612C>G) in positive control C2 was not correctly established. After reviewing Sanger sequencing data and the MPS alignment, a rare SNP (c.6549-51C>T) was identified in the intronic region where the respective MPS forward primer anneals. This variant prevented the amplification of the normal allele and caused a PCR allele dropout, thereby explaining the apparent genotype discrepancy (Supplementary Data IV).

It is noteworthy that during the optimization of the analytical process, an intronic mutation (chr2:g.152417626_152417632del) was initially filtered out. The reference SNP (rs) cluster identification of a known variant was attributed to this 7 bp deletion during the annotation performed by Ion Reporter (Supplementary Data IV). By default the Ion Reporter algorithm associates variants to the rs number based on the genomic coordinates involved rather than the specific variant itself. As a consequence, if a deletion or duplication coincides with the genomic position of a known SNP, the software attributes the same rs number. Considering this potential pitfall, all filtered-out INDEL variants were manually rechecked.

About 9% of the total numbers of variants called were classified as false positives. These false-positive variants resulted from: (i) sequencing artifacts in homopolymeric regions (consistently found in several samples tested), (ii) variants in shorter reads due to mis-priming events (similar to those reported by McCall *et al.*[14]) and usually showing a biased proportion of mutated/normal reads) and to a lesser extent (iii) variants located in regions with low coverage depth.

RYR1-related myopathies

Three patients (P1, P1B and P2) from two families were identified has having an autosomal recessive congenital myopathy due to defects in the ryanodine receptor 1 gene (Figure 1). All three patients had marked hypotonia and respiratory distress during the neonatal period but there was some degree of clinical variability. The two brothers from the first family are both wheelchair dependent (P1 never walked) and have a severe tetraparesis, whereas patient P2 achieved independent locomotion at 3.5 years of age and currently (at 28 years of age) walks without support. The patients' muscle biopsies showed considerable fiber size variability and frequent fibers with abnormal centrally placed nuclei – pathognomonic features for a central nuclear myopathy (Figure 1a). Four novel mutations were identified in *RYR1* (Figure 1b); these comprised two nonsense mutations (p.Gln4004* and p.Arg3053*) and two missense mutations (p.Met4881Ile and p.Pro1267Arg) affecting highly conserved residues. Bioinformatic analysis resorting to two distinct algorithms corroborated the pathogenic impact of these missense variants. The genotype was concordant between all three cases, since there is compound heterozygosity of a missense and a nonsense mutation (Figure 1c).

NEB-related myopathies

Patients P3, P5 and P7 were seen to have rods in their muscle biopsy, evocative of an NM (Figure 2a). Interestingly, in the first biopsy of P5 there were features compatible with congenital fiber-type size disproportion, with a focal area showing structures resembling rods. A second biopsy performed 2 years later showed larger areas and number of the typical rods. In terms of their clinical presentation, phenotypes were rather dissimilar, varying from a very mild distal myopathy diagnosed during adulthood to a congenital form of NM with bilateral club foot and requiring non-invasive ventilation.

Table 1 Pathogenic sequence variants detected by the massive parallel sequencing gene panel

Patient Id/ gender	*Gene/accession number*	*Gene location (exon/ intron)*	*DNA change Genomic cDNA (zygosity)*	*RNA change*	*Protein*	*Mutation type*	*Mutation origin*	*Family history*	*Frequency in ESP*	*Mutation reference*
C1 M	*DNM2* NM_001005360.2	11	Chr19:g.10909219C>T c.1393C>T (heterozygous)	r.(?)	p.Arg465Trp	Missense	*De novo*? (unaffected parents were not tested)	Sporadic	N.p.	Bitoun *et al.*[32]
C2 F	*RYR1* NM_000540.2	40	Chr19:g.38986918C>G c.6612C>G (heterozygous) (*)	r.(?)	p.His2204Gln	Missense	Inherited (parent #1)	Affected brother (same genotype)	N.p.	Duarte *et al.*[12]
		98	Chr19:g.39068613G>A c.14228G>A (heterozygous)	r.(?)	p.Gly4743Asp	Missense	Inherited (parent #2)		N.p.	Duarte *et al.*[12]
P1 F	*RYR1* NM_000540.2	87	Chr19:g.39034513C>T NM_000540.2:c.12010C>T (heterozygous)	r.(?)	p.Gln4004 (*)	Nonsense	Inherited (maternal)	Affected brother (P1B, with the (same genotype)	N.p.	New
		101	c.14643G>A (**) (heterozygous)	r.(?)	p.Met4881Ile	Missense	Inherited (paternal)		N.p.	New
P2 F	*RYR1* NM_000540.2	61	Chr19:g.39002235C>T c.9157C>T (heterozygous)	r.(?)	p.Arg3053 (*)	Nonsense	Inherited? (paternal inferred)	Sporadic	N.p.	New
		28	Chr19:g.38964051C>G c.3800C>G (heterozygous)	r.(?)	p.Pro1267Arg	Missense	Inherited (maternal)		1/13006	New
P3 M	*NEB* NM_001271208.1	129	Chr2:g.152408252C>T c.19944G>A (homozygous)	r.[19944g>a; 19944_19445ins19944 +1_19944+120, =]	p.Asn6649_Ile8560delins27	Affects splicing	Inherited? (untested parents)	Affected brother (not studied)	N.p.	Lehtokari *et al.*[15] /new (splicing effect)
P4 M	*TTN* NM_001267550.2	Intron 273	Chr2:g.179473930C>T c.52102+5G>A (heterozygous)	r.52034_52102del	p.Gly17345_Leu17367del	Affects splicing	Inherited (maternal)	Affected sib (same genotype, deceased in the neonatal period)	N.p.	New
		326	Chr2:g.179434749_179434750del c.76109_76110del (heterozygous)	r.(?)	p.Ile25370Argfs*6	Frameshift deletion	Inherited (paternal)		N.p.	New
P5 M	*NEB* NM_001271208.1	171	Chr2:g.152354789_152354792dup c.24294_24297dup (homozygous)	r.(?)	p.Glu8100Serfs*5	Frameshift duplication	Inherited? (consanguineous parents)	Affected sister (not studied, deceased)	N.p.	Pelin *et al.*,[16] Lehtokari *et al.*,[17] Lehtokari *et al.*[15]
P6 M	*TTN* NM_001267550.2	304	Chr2: g.179454772_179454773insGG c.61679_61680insCC (heterozygous)	r.(?)	p.Ser20561Leufs*17	Frameshift insertion	*De novo*	Sporadic	N.p.	New
P7 M	*NEB* NM_001271208.1	Intron 122	Chr2:g.152417626_152417632del c.19102-4_19102-10del (heterozygous)	r.[19102_19206del, 19101_19102ins19101 +1_19102-1, =]	p.[Val6368_Ser6402del, Val6368_Ile8560delins25]	Affects splicing	Inherited? (maternal inferred)	Affected relative (distant cousin) (shares p.Arg7026* mutation)	N.p.	Lehtokari *et al.*[15] /new (splicing effect)
		140	Chr2:g.152394412G>A c.21076C>T (heterozygous)	r.(?)	p.Arg7026*	Nonsense	Inherited (paternal inferred)		N.p.	Lehtokari *et al.*[15]
	SEPN1 NM_020451.2	10	Chr1:g.26139280T>G c.1384T>G (heterozygous)	r.(?)	p.Scy462Gly	Selenocystein codon (TGA) loss	?		N.p.	Ferreiro *et al.*[33]

Abbreviations: C, control; ESP, NHLBI GO Exome Sequencing Project, data accessed through the exome variant server (http://evs.gs.washington.edu/EVS/); Id, identification; N.p., variant not present in the database; P, patient; (*), zygosity incorrectly determined due to allele dropout (see Results section); (**), mutation not detected by gene panel, region not included in the assay.

Table 2 Clinical data of the patients presented in this work

Patient/ gender	Age (years)	Onset	Hypotony	Muscle weakness	CK (IU/l)	Ambulant?	Facial involvement?	Cardiac involvement?	Respiratory insufficiency?	Scoliosis?
C1 M	20	Early childhood	N	Ophthalmoparesis; tetraparesis; weakness predominantly in lower limbs	47	Y	Y	N	N	N
C2 F (*)	43 (**)	Adulthood	N	Only myalgia	456	Y	N	NA	N	NA
P1 F	28	Neonatal	Y	Ophthalmoparesis; bulbar weakness; proximal tetraparesis; axial muscle weakness	NP	N (never walked)	Y	N	Y (refused NIV)	Y (corrected by surgery 13 yr)
P1B M	34	Neonatal	Y	Ophthalmoparesis; bulbar weakness; proximal tetraparesis; axial muscle weakness	18	N (never walked)	Y	N	Y (refused NIV)	Y (corrected by surgery 15 yr)
P2 F	23	Neonatal	Y	Ophthalmoparesis; bulbar weakness; proximal tetraparesis; axial muscle weakness	5	Y (3.5 yr)	Y	N	Y (NIV since 16 yr of age)	Y (corrected by surgery 12 yr)
P3 M	75	Early childhood	N	Severe and early bilateral foot drop; later with milder facial, cervical and proximal upper limb weakness	Norm.	Y	Y	N	N	N
P4 M	Deceased (at 10 yr of age)	Congenital (arthrogryposis)	Y	Facial and bulbar weakness; severe tetraparesis; axial muscle weakness	Norm.	N (never walked)	Y	Y (dilated cardi-omyopathy)	Y (IV during neonatal period)	Y (not corrected)
P5 M	16	Infancy	N	Early respiratory muscle involvement; facial and bulbar weakness; proximal tetraparesis and later distal involvement; axial muscle weakness	137	Y	Y	N	Y (NIV since 4 yr of age)	Y (corrected by surgery 15 yr)
P6 M	17	Infancy	N	Facial weakness; proximal tetraparesis	54	Y	Y	N	N	N
P7 M	12	Congenital (bilateral club foot)	Y	Facial diparesis and bulbar weakness; tetraparesis predominantly distal	NP	Y (2 yr)	Y	N	Y (NIV since 8 yr of age)	N

Abbreviations: F, female; IV, invasive ventilation; M, male; N, no; NA, not assessed; NIV, non-invasive ventilation; norm., within normal range; NP, not performed; Y, yes; yr, years; (*), reported in Duarte *et al.*[12]; (**), age at the time of publication.

The first patient (P3) is a 75-year-old man with a slowly progressive distal myopathy. His motor and cognitive development was described as normal. After his 60s there were complaints of proximal upper limb and cervical weakness. There is no respiratory or cardiac involvement. All the variants identified in the *NEB* gene of this patient were detected in homozygosity including one mutation (c.19944G>A) located in the last base of exon 129 (Figure 2b). Since no known parental consanguinity was reported, the possibility of a large deletion spanning the entire *NEB* gene was excluded by the analysis of the amplicon's normalized coverage depth (Supplementary Data V). The c.19944G>A mutation was previously reported as pathogenic and possibly affecting splicing,[15] which was corroborated by our bioinformatic analysis. However, until now, its effect had not been experimentally proven. This mutation causes the disruption of the canonical splice site (naturally occurring), and a cryptic splice site located within intron 129 is used instead (Figure 2c). The outcome of this splice-site change is a partial intronic retention (120 bp) culminating in a premature termination codon and a shortened nebulin protein (p.Asn6649_Ile8560delins27). The patient has a brother with similar muscle complaints but he was unavailable for study; no other family members were known to be affected.

The second *NEB* mutation identified in this work (c.24294_24297dup) was detected in patient P5 (Figure 2b), probably in homozygosity since his parents are first degree cousins. This frameshift duplication was previously found in heterozygosity in patients with NM.[15–17] The neonatal period of patient P5 was uneventful with only initial feeding difficulties reported. There was significant respiratory involvement starting at 2 years of age. Two years later the respiratory difficulties progressed, requiring non-invasive ventilation. Corrective surgery for scoliosis was performed at 15 years of age. He currently has a proximal tetraparesis and walks without assistance. The patient had an older brother (deceased and without genetic studies) with a similar phenotype.

Patient P7 had the severest NM presentation of the three cases included in this work. He had a congenital disease onset, presenting at birth with hypotonia and predominantly distal muscle weakness with bilateral club foot. The patient required non-invasive ventilation since the age of 8. He has two heterozygous mutations in the *NEB* gene (Figure 2b). Although his parents were not studied, these mutations are likely to have distinct origins, since only one (p.Arg7026*) is in common with a distant cousin also affected with NM. This nonsense mutation has been previously identified in a third Portuguese patient with a (*NEB*-related) NM.[15] Surprisingly, this patient is reported to carry the same nonsense mutation as found in our patient P7 and the splicing mutation (c.19944G>A) identified in our patient P3. She was described as having a typical NM, which is milder than patient P7 but more severe than P3.

The other *NEB* mutation found in patient P7, a 7 bp deletion located at the end of intron 122 (c.19102-4_19102-10del), was also previously described[15] but not evaluated at the mRNA level. We demonstrated that it affects the *NEB* splicing process (Supplementary Data VI). Two abnormal transcripts were identified in muscle-derived total RNA: one is consequential of the exon 123 skipping (r.19102_19206del) and the other is the total retention of intron 122 (r.19101_19102ins19101+1_19102-1) in mature mRNA, predictably originating a shorter polypeptide. Patient P7 is also a carrier of a pathogenic variant (c.1384T>G, p.Scy462Gly) in the *SEPN1* gene. Although it may not influence the patient's phenotype, this variant should be considered for genetic counseling purposes in this family.

TTN-related myopathies

Mutations in the *TTN* gene were identified in two unrelated cases (P4 and P6) (Figure 3). Patient P4 had a severe neonatal hypotonia

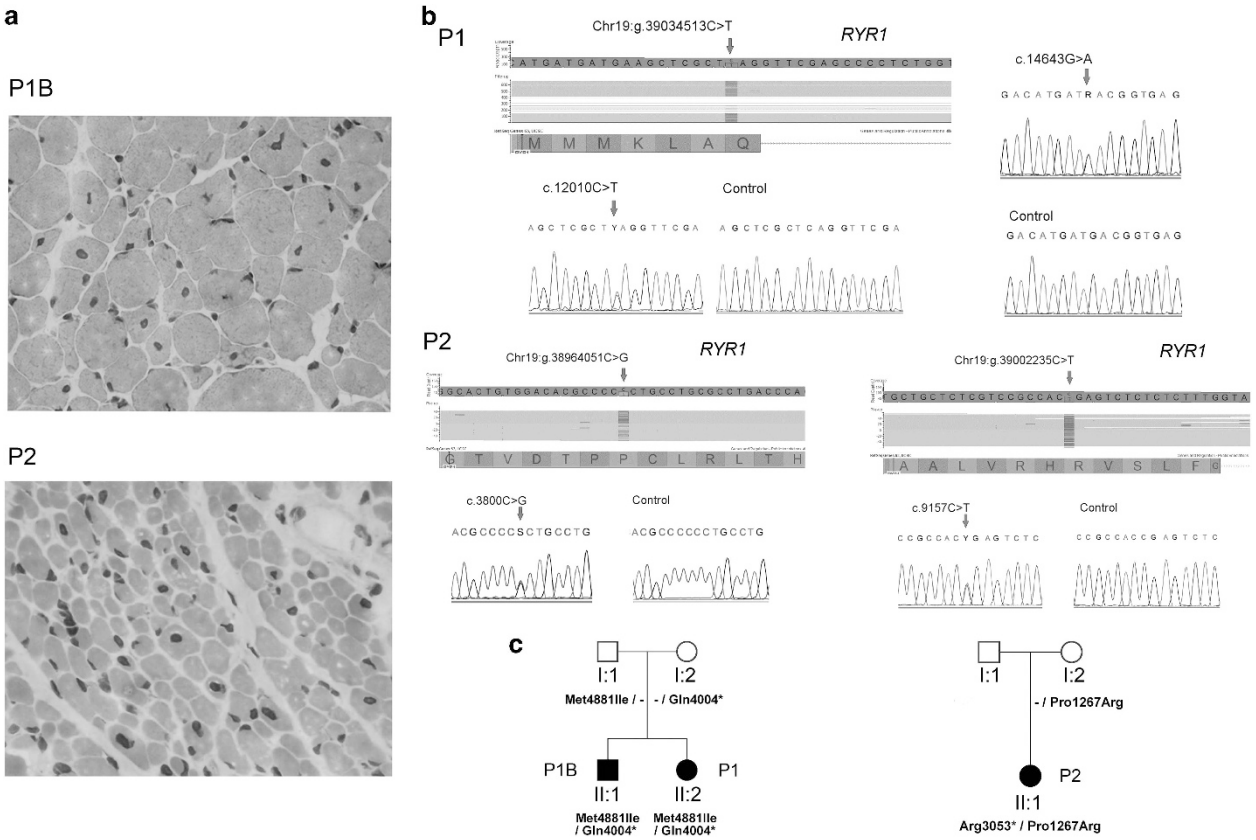

Figure 1 *RYR1*-related centronuclear myopathy. (**a**) Patients P1B and P2 histology of deltoid muscle biopsies (hematoxylin and eosin with x400 magnification). In both there is abnormal fiber size variability, with atrophic fibers and central nuclei. In patient P1B the majority of atrophic and hypotrophic fibers were type 1 fibers resembling congenital fiber-type size disproportion (not shown). (**b**) Heterozygous *RYR1* pathogenic variants identified in patient P1: (i) nonsense mutation c.12010C>T (p.Gln4004*) was identified by MPS (top part) and confirmed by Sanger sequencing (bottom); (ii) missense mutation c.14643G>A p.(Met4881Ile). Mutations identified in patient P2: (i) nonsense mutation c.9157C>T (p.Arg3053*) and (ii) missense mutation c.3800C>G (p.Pro1267Arg). (**c**) Patients' family pedigrees suggesting an autosomal inheritance pattern. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

with ophthalmoparesis, facial and bulbar weakness, dilated cardiomyopathy (cause of death at 10 years of age), requiring invasive ventilation and a feeding tube. The youngest of his two brothers had a similar phenotype and died during the neonatal period. P4's muscle biopsy showed myopathic features and multiple minicores (Figure 3a). Two novel heterozygous mutations in *TTN* were identified in patient P4. One of the mutations is a 2 bp deletion (c.76109_76110del; Figure 3b) of paternal origin, which presumably gives rise to a premature termination codon and a shorter polypeptide (p.Ile25370Argfs*6). The second mutation, maternally inherited, is an intronic single-nucleotide variant located near the splice site (c.52102+5G>A). As predicted by bioinformatic analysis, muscle RNA studies revealed that this variant is disease causing: the canonical splice site is abolished and a cryptic splice site located within exon 273 is alternatively used, originating an in-frame deletion of 69 bp (Figure 3c). The analysis of the other family members demonstrated the autosomal recessive inheritance pattern (Figure 3d).

Patient P6 had a milder phenotype, presenting with delayed walking during infancy. He had facial and limb-girdle weakness, mostly scapular weakness. Muscle biopsy showed myopathic changes and multicores (Figure 3a). The patient is currently 17 years old and shows no respiratory or cardiac involvement. The genetic study revealed a *de novo* heterozygous mutation: a 2 bp insertion in exon 304 (c.61679_61680insCC) (Figure 3b) that predictably gives rise to a premature stop codon in the A-band of Titin (p.Ser20561Leufs*17). Additionally, a novel silent variant was identified in exon 358 (c.105648G>A), but its effect at the mRNA level has yet to be determined.

DISCUSSION

We report the development of a novel and efficient method for mutation detection based on MPS and its application in the genetic characterization of patients with CMs.

Several centers reported low positivity rates in the molecular characterization of CM patients. These numbers may be seen in light of the fact that these studies were restricted to a relatively small number of candidate genes that were known, until recently.[3] Additionally, although some clinical and histological features allow some diagnostic guidance to the genetic study, in a subset of cases neither can orientate towards a (single) specific gene. Since the pathognomonic morphological abnormalities may show progressive or age-related changes,[5] the diagnostic value of the muscle biopsy depends when it is performed. As a final point, the severest end of the CM disease spectrum includes cases with prenatal onset and neonatal death, where the scarcity of clinical symptoms and the limited number of additional tests often cut short the diagnostic workup.

The introduction of new sequencing technologies (next-generation sequencing or MPS) circumvented some of these limitations, enabling

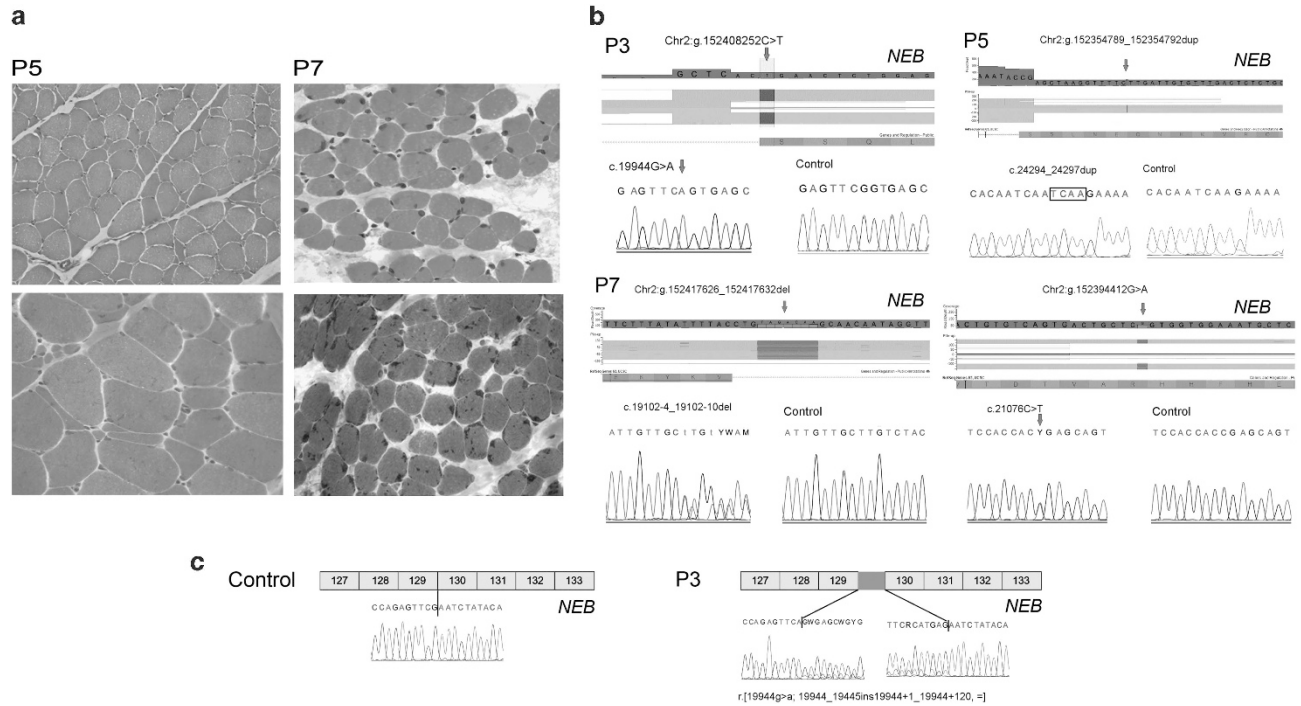

Figure 2 *NEB*-related nemaline myopathy. (**a**) Histological features of deltoid muscle biopsies of patients P5 and P7 (hematoxylin and eosin with x200 magnification in P5 and x400 in P7; GT with x400 magnification). In patient P5 there was fiber size variability, with severely atrophic fibers, corresponding to the majority to type 1 fibers. On the GT stain, there was a small area with some structures resembling rods. In P7 case, the typical rods were clearly seen on the GT stain. (**b**) Homozygous *NEB* mutations identified in patient P3 (c.19944G>A) and P5 (c.24294_24297dup), identified by MPS (top part) and confirmed by Sanger sequencing. Two heterozygous *NEB* mutations identified in patient P7: (i) an intronic deletion of 7 bp (c.19102-4_19102-10del) affecting the acceptor splice site and (ii) a nonsense mutation c.21076C>T (p.Arg7026*). (**c**) Characterization of the effect of the c.19944G>A mutation at the mRNA level. This mutation abolishes the normal donor splice site, shifting towards the use of an alternative (cryptic) splice site in intron 129, which predictably originates a premature stop codon and a shorter polypeptide. Abbreviation: GT, gomori trichrome. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

genetic diagnosis of diseases which had been considered extremely difficult to study by traditional methods. Different strategies based on MPS have been proposed for the genetic characterization of hereditary myopathies. These range from gene panels with a limited number of loci aiming to obtain high coverage depth,[18,19] to larger panels targeting several loci known to be implicated in myopathies[20] or even those with a broader scope, such as that covering all known neuromuscular disorders.[21] To address the genetic heterogeneity of these diseases, several successful applications of whole-exome or -genome sequencing were reported in the literature.[22–25] Irrespective of the approach taken, the histopathological findings should be correlated with the genetic data that are generated, as these can help narrow down the genetic studies considerably.[26]

Our strategy consisted in the development of a well delimited gene panel for CMs which in the present cohort led to an overall mutation detection rate of 70%. This is relevant considering that the assay does not include all the genomic regions, nor does it detect the presence of large genomic rearrangements. In patient P6 only one heterozygous frameshift mutation was identified. This was a *de novo* mutation (not detected in the patient's parents and paternity was confirmed). *TTN*-related myopathies are currently an expanding group of diseases, seen to have a wide clinical presentation that ranges from congenital to late adulthood onset, and with autosomal recessive or dominant inheritance.[10] The histological data of patient P6 showed the presence of multicores in NADH staining and by electron microscopy areas of sarcomeric disruption with irregularities and smearing of the Z line and disorganization of the myofibrils (data not shown). Additional experimental evidence is being gathered to clarify the impact of the apparently silent variant (c.105648G>A) and to rule out or confirm the presence of an additional *TTN* mutation, not detectable by the methods used.

About 25% of the variants detected in this group of patients required further studies (in particular, expression analysis at the mRNA level). Although bioinformatic analysis was highly suggestive of their pathogenic nature, their impact was ultimately demonstrated by these additional studies. The best example was the study of patient P3, where *NEB* gene analysis revealed the presence of an apparently silent homozygous variant (c.19944G>A), described in previous reports as likely to be deleterious.[16] Given that the patient's biopsy was performed several years ago and there were no muscle fragments to obtain RNA, the possibility of using blood-derived transcripts was evaluated. In fact, we detected suitable expression of nebulin in whole-blood nucleated cells, which enabled the study of the c.19944G>A variant and its effect on splicing. This mutation abolishes the normal splice-site consensus sequence leading to the use of an alternative (cryptic) splice in the adjacent intron. It is predictable that this partial intronic retention creates a premature termination codon and thus a shorter nebulin protein. The presence of a residual amount of normal transcripts, possibly due to a leaky mutational effect, explains the milder phenotype in this patient.

Even though the application of this gene panel was very successful, some technological limitations should be mentioned. First, one variant in a positive control was incorrectly genotyped due to an allele dropout during PCR, owing to the presence of a rare SNP in the

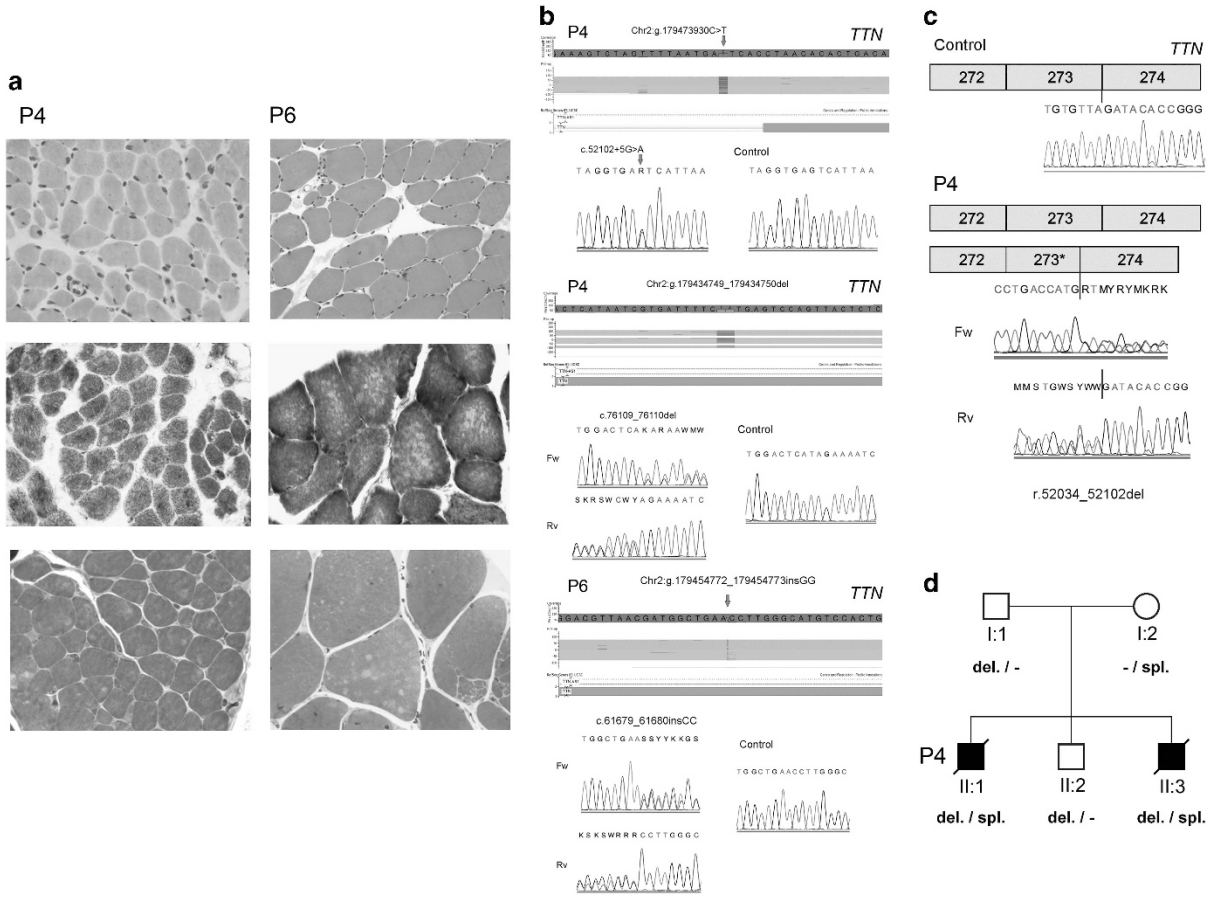

Figure 3 *TTN*-related myopathies. (**a**) Histological features of patients P4 and P6 deltoid muscle biopsies (hematoxylin and eosin with x400 magnification in P4 and x200 in P6, NADH with x400 magnification and toluidine blue semi-thin sections with x630 magnification in P4 and x400 in P6). In both cases there was fiber size variability and multiple small minicores seen on oxidative stains and semi-thin sections. (**b**) Heterozygous *TTN* mutations identified in patient P4: (i) splice-site mutation in intron 279 (c.52102+5G>A) and (ii) an out-of-frame deletion (c.76109_76110del) in exon 326. In patient P6 a heterozygous insertion of two nucleotides was detected (c.61679_61680insCC), which leads to a shift on the *TTN* reading frame (p.Ser20561Leufs*17). (**c**) The c.52102+5G>A mutation was further characterized at the mRNA level, where it was shown that it abolishes the donor splice site and an cryptic splice site in exon 273 is used in alternative, leading to an in-frame deletion (r.52034_52102del). (**d**) Pedigree of patient P4 family is compatible with an autosomal recessive inheritance pattern, where the two affected sibs share the genotype. Fw, forward; Rv, reverse. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

patient. Such events are not unusual in conventional sequencing. In an attempt to avoid this pitfall, the bioinformatic process for primer design used in the Ampliseq strategy takes into account positions of known SNPs. However, *de novo*, ultra-rare or population-specific variants will not be considered in the pipeline. This false-positive result had no practical implication in our study, because all the variants identified by MPS were confirmed by Sanger sequencing. Our main concern is the scenario where the allele dropout prevents the amplification of the mutated allele, generating a false-negative result. Data analysis is one of the biggest hurdles in this technology, given the different possibilities of error that may lead to false-positive or -negative results.[27] As shown here, besides the development of the MPS gene panel it was necessary to deploy a strategy to deal with the large number of variants by using specific variant filters. During this process, one of the mutations detected in this work was initially filtered out due to incorrect annotation. It is highly recommended, especially in large variant datasets, that different sources of variant annotation be used so as to ensure to allow accuracy and consistency. Variant annotation is one additional layer of variability in MPS[28] which very often remains disregarded.

As a future perspective, the intention is to expand this gene panel to include loci that have recently been linked to CMs, namely *KLHL41*, *LMOD3* and *SPEG*.[29–31] It would also be relevant to perform whole-exome sequencing in patients who remain genetically unsolved after screening with this gene panel.

In conclusion, the wide phenotypic heterogeneity and huge size of the candidate genes makes diagnosis of CM complex, costly and labor intensive. We have developed and demonstrated the clinical utility of a new MPS gene panel for the screening of these patients. In 7 of the 10 undiagnosed cases, pathogenic variants were identified, most of which contributed towards widening both the mutational and the phenotypic spectra of CMs.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Dr Gabriela Soares (Centro Hospitalar do Porto) and Dr Oana Moldovan (Centro Hospitalar de Lisboa Norte) for the respective referral of patients P10 and P8, and Dr Carlos Lima (Hospital Egas Moniz) for the

neuropathological characterization of patient P3. JO is a recipient of a PhD scholarship attributed by 'Fundo para a Investigação e Desenvolvimento do Centro Hospitalar do Porto'. The work was financed by the Institutions of the authors and in part by UMIB, which is funded by National Funds through FCT (Foundation for Science and Technology), under the Pest-OE/SAU/UI0215/2014.

1 Jungbluth, H. & Wallgren-Pettersson, C. in *Emery and Rimoin's Principles and Practice of Medical Genetics* 6th edn (eds Rimoin, D.L., Pyeritz, R.E. & Korf, B.) Ch. 127, 1–51 (Academic Press, 2014).
2 Oliveira, J., Oliveira, M. E., Kress, W., Taipa, R., Pires, M. M., Hilbert, P. *et al.* Expanding the MTM1 mutational spectrum: novel variants including the first multi-exonic duplication and development of a locus-specific database. *Eur. J. Hum. Genet.* **21**, 540–549 (2013).
3 North, K. N., Wang, C. H., Clarke, N., Deconinck, N., Bertini, E., Ferreiro, A. *et al.* International Standard of Care Committee for Congenital Myopathies. Approach to the diagnosis of congenital myopathies. *Neuromuscul. Disord* **24**, 97–116 (2014)
4 Snoeck, M., van Engelen, B. G., Küsters, B., Lammens, M., Meijer, R., Molenaar, J. P. *et al.* RYR1-related myopathies: a wide spectrum of phenotypes throughout life. *Eur. J. Neurol.* **22**, 1094–1112 (2015).
5 Romero, N. B & Clarke, N. F. Congenital myopathies. *Handb. Clin. Neurol.* **113**, 1321–1336 (2013).
6 Romero, N. B., Lehtokari, V. L., Quijano-Roy, S., Monnier, N., Claeys, K. G., Carlier, R. Y. *et al.* Core-rod myopathy caused by mutations in the nebulin gene. *Neurology* **73**, 1159–1161 (2009).
7 Piteau, S. J., Rossiter, J. P., Smith, R. G & MacKenzie, J. J. Congenital myopathy with cap-like structures and nemaline rods: case report and literature review. *Pediatr. Neurol.* **51**, 192–197 (2014).
8 Rocha, J., Taipa, R., Melo Pires, M., Oliveira, J., Santos, R. & Santos, M. Ryanodine myopathies without central cores—clinical, histopathologic, and genetic description of three cases. *Pediatr. Neurol.* **51**, 275–278 (2014).
9 Dowling, J. J., Lawlor, M. W. & Dirksen, R. T. Triadopathies: an emerging class of skeletal muscle diseases. *Neurotherapeutics* **11**, 773–785 (2014).
10 Chauveau, C., Rowell, J. & Ferreiro, A. A rising titan: TTN review and mutation update. *Hum. Mutat.* **35**, 1046–1059 (2014).
11 Wallgren-Pettersson, C., Sewry, C. A., Nowak, K. J. & Laing, N. G. Nemaline myopathies. *Semin. Pediatr. Neurol.* **18**, 230–238 (2011).
12 Duarte, S. T., Oliveira, J., Santos, R., Pereira, P., Barroso, C., Conceição, I. *et al.* Dominant and recessive RYR1 mutations in adults with core lesions and mild muscle symptoms. *Muscle Nerve* **44**, 102–108 (2011).
13 Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M. *et al.* An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **475**, 348–352 (2011).
14 McCall, C. M., Mosier, S., Thiess, M., Debeljak, M., Pallavajjala, A., Beierl, K. *et al.* False positives in multiplex PCR-based next-generation sequencing have unique signatures. *J. Mol. Diagn.* **16**, 541–549 (2014).
15 Lehtokari, V. L., Kiiski, K., Sandaradura, S. A., Laporte, J., Repo, P., Frey, J. A. *et al.* Mutation update: the spectra of nebulin variants and associated myopathies. *Hum. Mutat.* **35**, 1418–1426 (2014).
16 Pelin, K., Donner, K., Holmberg, M., Jungbluth, H., Muntoni, F. & Wallgren-Pettersson, C. Nebulin mutations in autosomal recessive nemaline myopathy: an update. *Neuromuscul. Disord.* **12**, 680–686 (2002).
17 Lehtokari, V. L., Pelin, K., Sandbacka, M., Ranta, S., Donner, K., Muntoni, F. *et al.* Identification of 45 novel mutations in the nebulin gene associated with autosomal recessive nemaline myopathy. *Hum. Mutat.* **27**, 946–956 (2006).
18 Kondo, E., Nishimura, T., Kosho, T., Inaba, Y., Mitsuhashi, S., Ishida, T. *et al.* Recessive RYR1 mutations in a patient with severe congenital nemaline myopathy with ophthalomoplegia identified through massively parallel sequencing. *Am. J. Med. Genet. A* **158A**, 772–778 (2012).
19 Ankala, A., da Silva, C., Gualandi, F., Ferlini, A., Bean, L. J., Collins, C. *et al.* A comprehensive genomic approach for neuromuscular diseases gives a high diagnostic yield. *Ann. Neurol.* **77**, 206–214 (2015).
20 Savarese, M., Di Fruscio, G., Mutarelli, M., Torella, A., Magri, F., Santorelli, F. M. *et al.* MotorPlex provides accurate variant detection across large muscle genes both in single myopathic patients and in pools of DNA samples. *Acta Neuropathol. Commun.* **2**, 100 (2014).
21 Vasli, N., Böhm, J., Le Gras, S., Muller, J., Pizot, C., Jost, B. *et al.* Next generation sequencing for molecular diagnosis of neuromuscular diseases. *Acta Neuropathol.* **124**, 273–283 (2012).
22 Bögershausen, N., Shahrzad, N., Chong, J. X., von Kleist-Retzow, J. C., Stanga, D., Li, Y. *et al.* Recessive TRAPPC11 mutations cause a disease spectrum of limb girdle muscular dystrophy and myopathy with movement disorder and intellectual disability. *Am. J. Hum. Genet.* **93**, 181–190 (2014).
23 Böhm, J., Vasli, N., Malfatti, E., Le Gras, S., Feger, C., Jost, B. *et al.* An integrated diagnosis strategy for congenital myopathies. *PLoS ONE* **8**, e67527 (2013).
24 Couthouis, J., Raphael, A. R., Siskind, C., Findlay, A. R., Buenrostro, J. D., Greenleaf, W. J. *et al.* Exome sequencing identifies a DNAJB6 mutation in a family with dominantly-inherited limb-girdle muscular dystrophy. *Neuromuscul. Disord.* **24**, 431–435 (2014).
25 Oliveira, J., Negrão, L., Fineza, I., Taipa, R., Melo-Pires, M., Fortuna, A. M. *et al.* New splicing mutation in the choline kinase beta (CHKB) gene causing a muscular dystrophy detected by whole-exome sequencing. *J. Hum. Genet.* **60**, 305–312 (2015).
26 Uruha, A., Hayashi, Y. K., Oya, Y., Mori-Yoshimura, M., Kanai, M., Murata, M. *et al.* Necklace cytoplasmic bodies in hereditary myopathy with early respiratory failure. *J. Neurol. Neurosurg. Psychiatry.* **86**, 483–489 (2015).
27 Robasky, K., Lewis, N. E. & Church, G. M. The role of replicates for error mitigation in next-generation sequencing. *Nat. Rev. Genet.* **15**, 56–62 (2014).
28 McCarthy, D. J., Humburg, P., Kanapin, A., Rivas, M. A., Gaulton, K., Cazier, J. B. *et al.* Choice of transcripts and software has a large effect on variant annotation. *Genome Med.* **6**, 26 (2014).
29 Gupta, V. A., Ravenscroft, G., Shaheen, R., Todd, E. J., Swanson, L. C., Shiina, M. *et al.* Identification of KLHL41 mutations implicates BTB-Kelch-mediated ubiquitination as an alternate pathway to myofibrillar disruption in nemaline myopathy. *Am. J. Hum. Genet.* **93**, 1108–1117 (2013).
30 Yuen, M., Sandaradura, S. A., Dowling, J. J., Kostyukova, A. S., Moroz, N., Quinlan, K. G. *et al.* Leiomodin-3 dysfunction results in thin filament disorganization and nemaline myopathy. *J. Clin. Invest.* **124**, 4693–4708 (2014).
31 Agrawal, P. B, Pierson, C. R., Joshi, M., Liu, X., Ravenscroft, G., Moghadaszadeh, B. *et al.* SPEG interacts with myotubularin, and its deficiency causes centronuclear myopathy with dilated cardiomyopathy. *Am. J. Hum. Genet.* **95**, 218–226 (2014).
32 Bitoun, M., Bevilacqua, J. A., Prudhon, B., Maugenre, S., Taratuto, A. L., Monges, S. *et al.* Dynamin 2 mutations cause sporadic centronuclear myopathy with neonatal onset. *Ann. Neurol.* **62**, 666–670 (2007).
33 Ferreiro, A., Quijano-Roy, S., Pichereau, C., Moghadaszadeh, B., Goemans, N., Bönnemann, C. *et al.* Mutations of the selenoprotein N gene, which is implicated in rigid spine muscular dystrophy, cause the classical phenotype of multiminicore disease: reassessing the nosology of early-onset myopathies. *Am. J. Hum. Genet.* **71**, 739–749 (2002).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)