Review

Is Gene-Size an Issue for the Diagnosis of Skeletal Muscle Disorders?

Marco Savarese[a,b,*], Salla Välipakka[a,b], Mridul Johari[a,b], Peter Hackman[a,b] and Bjarne Udd[a,b,c,d]
[a]*Folkhälsan Research Center, Helsinki, Finland*
[b]*Department of Medical Genetics, Medicum, University of Helsinki, Helsinki, Finland*
[c]*Neuromuscular Research Center, Tampere University and University Hospital, Tampere, Finland*
[d]*Department of Neurology, Vaasa Central Hospital, Vaasa, Finland*

Abstract. Human genes have a variable length. Those having a coding sequence of extraordinary length and a high number of exons were almost impossible to sequence using the traditional Sanger-based gene-by-gene approach. High-throughput sequencing has partly overcome the size-related technical issues, enabling a straightforward, rapid and relatively inexpensive analysis of large genes.

Several large genes (e.g. *TTN, NEB*, *RYR1*, *DMD*) are recognized as disease-causing in patients with skeletal muscle diseases. However, because of their sheer size, the clinical interpretation of variants in these genes is probably the most challenging aspect of the high-throughput genetic investigation in the field of skeletal muscle diseases.

The main aim of this review is to discuss the technical and interpretative issues related to the diagnostic investigation of large genes and to reflect upon the current state of the art and the future advancements in the field.

Keywords: Large genes, variant interpretation, genetic diagnosis, variants of uncertain significance (VUS), copy number variants (CNV)

INTRODUCTION

Human protein-coding genes have a variable length, ranging from a few hundred nucleotides up to several millions [1, 2]. Most of the large genes, e.g. *DMD*, have introns with an extraordinary length [2, 3]. Some other genes have adapted with evolution, reducing the size of their introns to have a higher transcriptional efficiency [4, 5]. *TTN* gene, for example, has evolved through many gene duplication events by reducing the size of its introns and thus optimizing its transcription [6].

Large genes have been reported to be enriched in pathways linked to cancer or other human diseases, including cardiomyopathy and skeletal muscle diseases [4]. To study the genes with a long coding sequence and a large number of exons, the traditional Sanger sequencing was an extremely expensive, time consuming and laborious approach. This technical bias has hampered a proper investigation of these genes for diagnostic purposes, reducing the number of variants identified and hampering a correct diagnosis in probably thousands of patients.

The rapid and thorough investigation of multiple genes, made possible by the introduction of high throughput sequencing (HTS), has allowed the

*Correspondence to: Marco Savarese, Ph.D., Folkhälsan Institute of Genetics, Department of Medical Genetics, University of Helsinki, Biomedicum, Haartmaninkatu 8 -Pb 63 00014 Helsinki–Finland. Tel.: +358 294125069; E-mail: marco.savarese@helsinki.fi.

Table 1
Large genes causing a skeletal muscle disease

	Coding exons	Reference transcript ID*	Size of the coding sequence#
TTN	363	NM_001267550.1	107976
NEB	183	NM_001271208.1	25683
RYR1	106	NM_000540.2	15117
PLEC	32	NM_000445.3	13725
DMD	79	NM_004006.2	11058
SPEG	41	NM_005876.4	9804
COL12A1	66	NM_004370.5	9192

*As listed in the Leiden Database. #longest transcript.

-Technical: repetitive sequences

-Biological: isoforms/splicing events

-Interpretative: variant interpretation

Fig. 1. – Gene-size related difficulties. Three are the main issues related to the diagnostic investigation of large genes: the technical issues due to the presence of repetitive sequences and the subsequent mapping difficulties; the biological issues due to alternative splicing events resulting in isoforms with a different expression; the interpretative issues related to the clinical interpretation of the high number of rare variants identified in large genes.

analysis of large genes. The use of HTS has also resulted into a growing number of variants identified in these genes and in an increased diagnostic yield [7–9], expanding the spectrum of diseases associated with large-genes [10–14].

Here, we will briefly review the findings related to the three skeletal muscle disease-genes with the longest coding sequence (Table 1) [15]. We will then discuss the technical and interpretative difficulties, related to the sheer size of these genes, met during the diagnostic workflow (Fig. 1). Finally, we will mention the advancements in the field and the possible outcome of the recently developed, cutting-edge sequencing technologies soon to be used in a diagnostic setting.

THE TITIN GENE, *TTN*

The 363-coding exon *TTN* gene encodes titin, the largest known human protein [16]. Titin plays several crucial structural and functional roles in the muscle through a wide network of interactions and interactors [17]. *TTN* mutations are responsible for a wide spectrum of skeletal muscle disorders with or without an overt cardiac involvement [17]. Skeletal muscle titinopathies are mainly recessive and include congenital myopathies and proximal or distal myopathies with a later onset [17–19]. Mutations in the last exons

can result in a dominant form, the Tibial Muscular Dystrophy, a late onset distal myopathy [20]. Missense mutations in a specific exon (exon 344) have been associated with an adult onset hereditary myopathy with early respiratory failure (HMERF) [21, 22].

THE NEBULIN GENE, *NEB*

With its 183 exons, *NEB* encodes nebulin, a big protein of 600–900 kDa [23]. Nebulin has a highly repetitive structure and can bind hundreds of actin monomers, thereby regulating the length of actin filaments and their interaction with myosin [23].

NEB mutations are the most common cause of autosomal recessive congenital nemaline myopathy, but onset may range from the severe forms with a perinatal onset to milder forms with a later onset [24].

However, recessive disease-causing variants in *NEB* have been identified also in patients with a distal myopathy [25, 26], core rod myopathy [27], and fetal akinesia/lethal multiple pterygium syndrome [12, 28]. *NEB* with its 32-kb triplicate region (eight exons repeated three times: 82–89, 90–97, 98–105) is prone to copy number variants (CNV) [29]. Moreover, recently, Kiiski and colleagues described a large in-frame deletion, dominantly inherited in a three-generation family, causing a distal nemaline/cap myopathy [30].

THE RYANODINE RECEPTOR GENE, *RYR1*

RYR1 gene encodes ryanodine receptor 1, an intracellular channel responsible for the release of Ca2+ from sarcoplasmic reticulum [31].

RYR1 mutations cause a wide spectrum of dominant and recessive myopathies [32]. *RYR1*-related myopathies are usually classified in several histological subtypes, including central core disease, multiminicore disease, core–rod myopathy, centronuclear myopathy and congenital fiber-type disproportion [33–37]. Moreover, *RYR1* mutations are a well-known cause of dominant malignant hyperthermia (MH) susceptibility [38], and of exercise-induced rhabdomyolysis [39]. Recently, a calf-predominant myopathy with core pathology was associated with dominantly inherited *RYR1* mutations [14].

ALTERNATIVE SPLICING EVENTS AND MULTIPLE ISOFORMS

Large multi-exonic genes undergo extensive alternative splicing events in different developmental and physiological states [40]. Alternative splicing is a highly regulated process by which a single gene produces multiple distinct mRNA isoforms and protein variants of different size [40].

Although *TTN* transcripts have traditionally been classified in six main isoforms [16], we and other groups have described a more complex splicing pattern with an elevated number of alternative splicing events, resulting into exon skipping events and the use of alternative 5′ and 3′ splice sites. During the prenatal development, larger and more compliant titin isoforms are expressed. A perinatal switch in titin isoforms leads to the production of shorter transcripts [41], resulting into a smaller and less compliant protein expressed in the later period of life [42]. Interestingly, *TTN* mutations located in exons with a higher fetal expression, result into a form of arthrogryposis multiplex congenita characterized by reduced fetal movements and a congenital amyoplasia and severe hypotonia [19]. A different expression of specific exons among anatomically different muscles is also expected although the experimental setting of our recent study on titin splicing in adult skeletal muscles did not allow us to observe any clear splicing difference among the anatomically different muscles [43].

A complex splicing pattern has also been reported for *NEB*. In particular, specific regions of the gene (exons 63–66; exons 143–144; exons 167–177) undergo extensive alternative splicing. A study performed by Laitila and colleagues combining expression array and RT-PCR data suggests that anatomically different muscles do not show specific *NEB* isoforms [44]. *NEB* splicing is also developmentally-regulated, resulting in different isoforms with potentially different functional roles [45].

In *RYR1*, a fetal isoform (ASI-) lacks residues 3481–3485 in exon 70 [46]. The alternative splicing of this region plays an important role in adapting to different physiological and pathological conditions [47].

A number of RNA-binding proteins act as splicing regulators, determining the isoform expression of large genes. The muscle-specific splicing factor RBM20 is responsible for *TTN* alternative splicing [41, 48] and it targets other several important genes, including cardiomyopathy and skeletal muscle disease-related genes (e.g. *CACNA1c*, *RYR2*, *LDB3*, *DAB1*, *CAMK2D* and *SPEN*) [49]. Similarly, CUG binding protein 1 (CUG-BP1) regulates the alternative splicing of RYR1 [47].

A better understanding of these events and a further characterization of the splicing regulators will probably provide new insight in the pathogenesis of human diseases and, probably, novel potential pharmaceutical and therapeutic targets.

THE INTERPRETATION OF RARE VARIANTS IN LARGE GENES

Because of their sheer size, rare variants in large genes are observed almost in any test able to investigate these genes. The evaluation of the clinical meaning of these variants is a challenging multi-step process based on specific criteria, as suggested by the ACMG/AMP guidelines [50]. These guidelines represent a general framework and their application to large genes does not allow a straightforward distinction between the few causative mutations and the large number of rare, clinically irrelevant, variants. Thereby, clinical geneticists report most of these experimentally identified rare variants as variants of uncertain significance (VUS).

When interpreting variants in large genes, a 'deep phenotyping' is crucial to identify a correlation between the observed phenotype and the known gene-associated clinical presentations [51]. The recent large HTS-based studies are further expanding the already broad range of clinical phenotypes associated to the genes discussed in this review [13, 14, 19]. Traditionally, the diagnosis of skeletal muscle disorders benefits from a careful evaluation of clinical signs and symptoms, of creatine kinase level, of histopathological findings on a muscle biopsy and of electromyography records. However, each of the aforementioned tests does not have enough specificity to discriminate among the different genetic forms. A comprehensive diagnostic approach and analysis is thereby required. On the other side, in the last few years, several international studies are successfully identifying and describing specific patterns of muscle involvement, evaluated through MRI scans, in genetically different muscle disorders [52–55].

The different forms of titinopathies show specific progression-related patterns of muscular involvement [18, 19, 21, 56]. In the *RYR1*-related dominant central core myopathies, MRI studies show a selective involvement of vasti, sartorius and adductor magnus

in the thigh and of soleus, gastrocnemii, and peroneal group in the leg with relative sparing of rectus femoris, gracilis, adductor longus and tibialis anterior [57]. A similar, although more diffuse, pattern is seen in recessive *RYR1* myopathies [58], although in the new distal calf-predominant *RYR1*-myopathy the target muscle was the medial gastrocnemius. A small series study has showed the characteristic involvement of the tibialis anterior and soleus and the sparing of the thigh muscles in *NEB*-related nemaline myopathies [59].

As we previously suggested for recessive titinopathies [13], also for other recessive diseases due to mutations in large genes, the identification of bi-allelic variants resulting into a premature stop codon (nonsense variants or small indels causing a frameshift) or the detection of previously reported mutations easily addresses the diagnosis. Novel missense and splice variants require an extensive and comprehensive characterization including in *silico*, in *vitro* and in *vivo* tests [13, 60].

Missense variants can result in a diagnosis only when sufficient evidence supporting their pathogenicity is obtained [50]. Many computational tools for predicting the pathogenicity of missense variants have been developed [61–63]. They take into account the amino acid or nucleotide conservation or the biochemical/structural/functional properties of the amino acid change [64].

Recently, a deep learning network for pathogenicity prediction, named PrimateAI, has been developed [65]. The program has been trained using hundreds of thousands of common variants from a large population sequencing data from six non-human primate species [65]. The sole analysis of the human population does not allow a correct evaluation of the frequency of a specific variant. In humans, in fact, the total number of common variants has been reduced by bottleneck events that have largely reduced the ancestral diversity. Because of these bottleneck events, only 0.1% of the missense variants have a MAF > 0.1% in the human population and, consequently, most of the human missense variants are ultra-rare or private. PrimateAI evaluates the allele frequencies of a specific variant in different primate species: if a variant, affecting a conserved amino acid, is polymorphic in these species, then most probably it will be benign also in humans [65].

Recently, Laddach and colleagues developed a web application, TITINdb, that integrates information about *TTN* structure, sequence, variant and disease in a single, user-friendly environment. TITINdb is a precious resource to map *TTN* variants to domain structures and to predict their impact using computational methods based on the protein structure and sequence [66].

A different approach for in *silico* prediction is represented by ensemble methods able to combine the results of several individual predictors to improve the predictive performance [64, 67]. Recently, a novel ensemble method, named REVEL, has been released [68]. REVEL is reported to outperform the other existing methods for distinguishing possible disease causing missense mutations from rare missense variants with an MAF below 3% [68, 69].

Still more complex is the interpretation of synonymous single nucleotide variants (sSNVs) that are often thought to be functionally irrelevant since they do not alter the protein sequences. However, sSNVs have been associated to hundreds of different human diseases since they can affect the transcription and the splicing regulation, the microRNA binding, the mRNA folding, and, finally, the translation [70–72]. Recently, Shi and colleagues have developed IDSV (Identification of Deleterious Synonymous Variants), a computational model able to predict the possible deleterious effect of sSNVs by using a wide variety of features [73].

Similarly, exonic variants causing missense changes as well as intronic variants may be cryptic splice mutations. Different bioinformatic tools have been developed to predict a possible splicing effect of an identified variant [74, 75]. SpliceAI is a deep residual neural network that uses the genomic sequence of the pre-mRNA transcript to predict whether each position in a pre-mRNA transcript acts as a splice donor, splice acceptor, or neither and, also, to estimate the splicing effects of genetic variants in each genomic position [76]. The prediction score provided by SpliceAI for each variant reflects the probability of the variant altering the splicing [76]. With an accuracy over 95%, SpliceAI is reported to outperform the other available tools [76].

In *silico* predictors may provide supporting evidence for pathogenicity. However, a more reliable evidence is provided by in *vitro* studies. Biochemical and biophysical studies, using wild-type and mutated constructs, have been used to characterize the effect of missense variants in large genes. Using thermal denaturation monitored by circular dichroism spectroscopy, Chauveau and colleagues demonstrated the reduced stability of the missense mutation within the enzymatic site of the *TTN* kinase domain (p.Trp34072Arg) [77]. Similarly, Hastings and col-

leagues proved that a *TTN* mutation (p.Ala178Asp), located in the Z-disk region, leads to partial misfolding of bacterially expressed Z1Z2 protein fragment [78].

A second possibility is to study the effect of a variant on protein–protein interactions. Using plasmid vectors for the expression of human nebulin super repeats, Marttila and colleagues demonstrated that a *NEB* missense variant (p.Ser6366Ile) causes an increased nebulin–actin affinity and a second missense variant (p.Thr7382Pro) reduces the affinity of nebulin for tropomyosin [79]. Recently, an interesting nebulin super-repeat panel has been described by Laitila and colleagues [80]. The panel allows the study of the actin binding of each single super-repeat and it is a precious and innovative tool to assess the effect of *NEB* missense changes identified in patients on nebulin-actin interaction [80]. Finally, *in vitro* studies can provide a direct evidence of an abnormal Ca2 + homeostasis, suggesting a disease-causing effect of *RYR1* variants [81–83].

Variants in canonical splice sites or predicted as being splice-disrupting also need a further cDNA validation and characterization. This is particularly important considering that large genes have multiple isoforms with a development- and tissue-specific expression [43–45, 47]. Splice variants can result in an out-of-frame deletion or insertion and, consequently, in a premature truncation; however, they can also result in a slightly longer or smaller protein (as a consequence of an in-frame deletion or insertion, in particular in presence of symmetric exons). A further characterization of the protein expression and function is strongly recommended for a proper evaluation of the mis-splicing effect.

A good example of the aforementioned issues with interpretation of splicing variants is represented by the recently identified recurrent *TTN* intronic splice-site variant (c.39974-11T>G) [84]. A large segregation in eight families where the variant, in *trans* with a second causative variant, co-segregated with the disease and a comprehensive analysis of expression data strongly suggested the pathogenic role of the identified variant [84].

A more robust proof of pathogenicity can be provided with functional genomics studies. Functional genomics approaches include a number tools, requiring for example patients' cells, micro-organism or animal models, that can be used (often in combinations among them or with *in vitro* studies) to obtain additional evidence for pathogenicity of genetic variants [85].

The availability of protocols to reprogram somatic cells into pluripotent stem cells (iPSCs) enables, for example, the study of sarcomere organization in iPSC myocytes and cardiomyocytes derived from patients' fibroblasts [86, 87]. On the other hand, RNA-guided CRISPR (clustered regularly interspaced short palindromic repeat)-associated Cas proteins can be utilized to create knock in cellular and, above all, animal models and mimic patients', and hopefully disease, states [88, 89].

So far, animal models have been mainly used to prove that a novel gene, previously not reported as disease causing, is implicated in the observed disease (gene discovery) and/or to provide information on the pathophysiological mechanisms triggered by the gene mutations [90, 91].

For pathophysiology studies and for testing potential therapeutic strategies, zebrafish models of nemaline myopathy, titinopathy and Ryr1-related myopathies have been used [92–94].

Similarly, to study the physiopathology of the dominant tibial muscular dystrophy (TMD) and the recessive limb-girdle muscular dystrophy (LGMD2J or LGMD R10 titin-related) due to heterozygosity and homozygosity for the FINmaj mutation, Charton and colleagues generated a mouse model carrying the same mutation [95]. Several *RYR1* knock in mouse models have been generated to mimic the equivalent mutations identified in humans [96–99]. Recently, Laitila and colleagues have generated and characterized a mouse model with compound heterozygous Neb mutations (a missense p.Tyr2303His and a nonsense p.Tyr935*), matching the genotype observed in patients with a nemaline myopathy [100, 101].

An interesting perspective is represented by the recently developed Gene Replacement (GR) technology that enables to replace mouse genes with their full human orthologs [102]. The new full gene-replacement model would mimic the same expression, regulation and function of the human gene, improving our understanding of the gene function, of the disease mechanisms triggered by gene mutations and, finally, providing a valuable model for possible treatment options [102]. The feasibility of this approach for large genes is still to be proven.

Finally, considering the sheer size of these genes and their complex structure with repetitive areas and GC-rich regions, sequencing parameters (e.g. depth and coverage) need to be carefully evaluated for a proper interpretation of the genetic results [11]. The sole DNA sequencing, in particular a non-custom tailored exome sequencing, and the traditional algo-

rithms in use in a diagnostic setting (mainly aiming at the detection of SNV or small indels) can still result into a number of elusive damaging variants. As discussed below, a more exhaustive workflow, including further bioinformatic analyses and second-tier tests, often results in a higher diagnostic rate, revealing variants missed by the traditional diagnostic methods.

THE IDENTIFICATION OF COPY NUMBER VARIANTS FROM HTS DATA

Copy number variants are genomic regions of loss or gain of at least 50 bp in size, which are formed by distinct mechanisms compared to SNVs and indels [103]. CNVs are estimated to cause approximately 10% of disorders, and they seem to be even more involved in neurological disorders than in many other disorder groups [104, 105]. CNVs can be detected from various types of next generation sequencing (NGS) data, and numerous CNV detection algorithms have been developed during recent years [103, 106]. Usually, different technical approaches are needed for WES and gene panel data compared to WGS data, since the former produce non-continuous sequencing data [107]. The CNV detection algorithms designed for WES and gene panel data require high average read-depth and uniform coverage to provide sensitive and reliable CNV detection results, which puts restrictions on the quality of NGS data [108]. Additionally, the whole spectrum of genomic structural variation can be detected only from WGS data, as opposed to deletions and duplication, which are detectable also from targeted sequencing data [107].

The CNV detection algorithms tend to have differing CNV detection accuracy and biases in detected CNV classes: therefore, utilizing more than one algorithm is generally recommended to achieve comprehensive CNV detection results [106, 109–111]. Kosugi and colleagues list CNV detection algorithms for WGS data with relatively best performances for each structural variation category, including CNVs [103]. For now, studies with comparably comprehensive algorithm comparisons are not available for WES and gene panel sequencing data, but numerous studies of smaller scale have been published to aid in making the choice [105, 106, 109, 110].

Large deletions or duplications in the DMD gene are a well-known cause of Duchenne and Becker muscular dystrophies (70–80% of cases) and multiplex ligation-dependent probe amplification (MLPA) is the current standard for clinical CNV analysis in

DMD [112, 113]. Two hot spots, proximal with exons 2–20 and distal with exons 45–55, contain most of the CNV. Nevertheless, the detected CNVs are highly heterogeneous. The breakpoints land mostly in the very large introns. For most of the DMD and BMD patients (>90%), the phenotype severity depends on the effect of the CNV on translation, premature termination of protein synthesis through deleterious change in read-frame being the most notable [112]. Therefore, detecting CNVs precisely on exon level is highly important in the case of dystrophinopathies. This is quite feasible from NGS data, with separate approaches developed specifically for analyzing CNVs in the gene DMD due to its high clinical significance, but more general approaches have provided detections as well [114, 115].

The large size of certain genes involved in neuromuscular disorders in itself does not pose a problem for CNV analysis from NGS data. However, nebulin and titin provide unique challenges for CNV detection due to regions of segmental duplications [16, 23, 116]. Repetitive gene sequences are challenging to sequence and align accurately, which leads to imprecise basis for accurate CNV analysis on these regions [107]. These regions are especially difficult to decipher from non-continuous short-read sequencing data from WES and gene panels. For unambiguous sequencing and mapping of these repeated regions, special approaches are probably needed, namely specific probe designs and/or long-read sequencing, even with WGS approaches [11]. These regions could be actually of special interest: changes in the number of triplicate region blocks in the gene *NEB* are potentially pathogenic [116]. CNVs from other regions of these giant genes have been detected during recent years, also with NGS approaches. In *TTN*, CNVs have been detected to cause myopathies with or without cardiac involvement in compound heterozygosity with other variant types [29, 115, 117] and in *NEB* with even more variable outcomes [24, 30, 115]. Recently, a large heterozygous deletion in *NEB* was identified as being the cause of a distal myopathy through a probable dominant negative mechanism on molecular level [30]. This is the first reported dominant disease for *NEB*, suggesting that different rules may apply for CNVs than for other types of variants in consideration of their clinical significance and effects. Generally, CNVs as disease causing discoveries seem to be rare for these genes, maybe due to difficulties in analyzing the genes with old methods.

CNVs have been discovered also in some other genes causing a skeletal muscle disorder, such as

LAMA2 [118], *MTM1* [119], *RYR1* [120], *SACS* [121], *SGCB* [122] and others [123]. Nevertheless, it is probable that large genes with few verified CNVs detected by any methods so far (e.g. *PLEC* or *LARGE*) will have ones soon, with the amount of comprehensive genetic variation studies increasing.

Following detection of potentially interesting CNVs, inferring their clinical significance and effects is even less straightforward than for SNVs or small indels: the variant databases available for CNVs are generally not as well curated as for other variant types [124]. Additionally, comparing CNVs to ones in these databases is not unequivocal, since the reported breakpoints of the CNVs could differ depending on their original detection methods [124, 125]. The American College of Medical Genetics (ACMG) has published very recently guidelines for interpreting and reporting germline CNVs [126]. The application of these guidelines will probably modify the interpretation of CNVs in a diagnostic setting. The CNV guidelines will be probably improved and customized for a more efficient diagnostic use. In the meanwhile, the CNV detection results still need to be regarded with caution and may need to be verified with complementary methods [105, 111, 123]. This increases workload and costs, thus setting back use of NGS methods as an independent first-tier diagnostic test.

RNA SEQUENCING AS SECOND TIER TEST

RNA sequencing (RNAseq) is a convenient second tier test that complements a DNA-based method and is able to identify possible elusive variants [127, 128]. The low detection rate of DNA tests is probably due to several reasons. Complex genetic mechanisms, such as a digenic or oligogenic inheritance, and the presence of causative mosaic variants, can probably explain the observed phenotype in part of the unsolved cases [129–132]. On the other hand, part of the patients with an undiagnosed disease carry variants that are not detectable (for example deep intronic variants in exome sequencing) or variants not correctly interpreted (e.g. synonymous variants) [10].

Cryptic splice mutations explain 9–11% of cases with intellectual disability or autism spectrum disorders, respectively [76]. A correct evaluation of elusive splice variants, using bioinformatic tools and RNASeq, can result into a similar increase in the diagnostic yield in most of the other rare genetic diseases. Similarly, the integration of RNAseq with genome sequencing has resulted in an improved diagnostic rate for a wide spectrum of undiagnosed Mendelian diseases [133]. Moreover, the availability of the most appropriate tissue for RNA extraction/analysis further increases the diagnostic rate [18, 128]. For skeletal muscle disorders, muscle is the most informative tissue due to the higher expression of disease genes [128].

After the first report of a novel re-occurring *COL6A2* mutation (c.930 + 189C > T) identified using a well-designed workflow for prioritizing candidate aberrant splicing events [127], a similar approach has been successfully used to screen unsolved patients with a nemaline myopathy [134]. Hamanaka and colleagues identified a novel deep-intronic *NEB* pathogenic variant (c.1569 + 339A > G) and a synonymous *NEB* pathogenic variant (c.24684G > C; p.Ser8228Ser affecting the last nucleotide of exon 175), both resulting in an aberrant splicing [134].

A different approach, described by Lee and colleagues, is of the extreme interest [133]. Instead of analyzing the entire transcriptome to search for an outlier (as described in ref. [127, 134]), they used RNAseq to evaluate the effect on the transcripts of rare, potentially causative, genetic variants using the splicing predictors as a prioritizing method [133].

A muscle biopsy is routinely collected during the diagnostic procedure for patients with skeletal muscle disorders. However, the use of transcriptome analysis for a diagnostic purpose will benefit from the development of methods to transdifferentiate *ex vivo* skin fibroblasts or blood mononuclear cells, not requiring invasive medical procedures, to specific cell types. This will virtually enable the analysis of transcripts with a development-specific expression.

It is noteworthy that the interpretation of RNAseq data is still challenging because of the presence of natural splice variants that need to be distinguished from the pathogenic splicing defects. Improved algorithms and defined guidelines will probably facilitate the interpretation of these data.

Finally, the effect of splice defects on translation needs to be carefully evaluated.

FINAL CONSIDERATIONS AND FUTURE PERSPECTIVES

The introduction of HTS has allowed us to overcome the technical difficulties related to the size of

the gene. DNA and RNAsequencing and, above all, their combined use enable an exhaustive analysis of most of the human genes and their size is not an issue anymore [133].

The use of long-reads and linked-reads sequencing technologies will probably result into several technical improvements, allowing the detection of complex structural variations, large segmental duplications and possible microsatellite expansions, such as trinucleotide repeat expansion [135–137]. These technologies will also improve the variant detection in repetitive regions where short reads do not map uniquely. Finally, an important (and often neglected) aspect is related to the phase information [138]. Short reads collapse the diploid genome in a single sequence. Phasing variants using segregation studies is sometimes time-consuming and, somehow, not cost-effective. These new technologies provide phase information over long contiguous DNA segments [138]. Large genes will probably benefit the most from the introduction of long-reads and linked-reads sequencing technologies.

The most challenging aspect for the diagnosis of skeletal muscle disorders is definitively related to the clinical interpretation of the high number of rare variants detected in large genes (Table 2). Despite the possible improvements in the variant interpretation and the definition of gene-tailored guidelines, the variant interpretation is a dynamic process. It evolves because of multiple factors including a better understanding of the disease, the availability of

more performant in *silico* predictors and of additional population data, the development of novel *in vitro* and *in vivo* functional studies and the identification of novel cases [139–141]. Previous studies have questioned, for example, the pathogenicity of mutations associated with the limb girdle muscular dystrophies (LGMDs) or with a cardiomyopathy, suggesting the need for a periodic, careful re-evaluation of the experimental findings [139, 140]. A recent study by Appelbaum and colleagues has discussed the ethical duty to reinterpret experimentally identified variants, concluding that we all need to re-evaluate periodically our findings in the light of technical and interpretative improvements [142].

A crucial aspect for a proper evaluation of the sequence variants is represented by the choice of appropriate functional studies. Although recommendations have been recently issued to provide a detailed guidance on the evaluation of functional data [143, 144], for the large genes discussed here, we do not have a general agreement on the assays providing sufficient evidence. Moreover, the large size of the coding region is a considerable issue for specific applications (e.g. cloning full-length sequence or – mainly for titin - protein expression study). It is however noteworthy that, in the context of MYH7-associated inherited cardiomyopathies, a panel of experts suggested that only functional data from mammalian knock-in models provide supporting evidence of the variant damaging effect [145].

Table 2
Challenges and possible improvements in variant interpretation

Key points for variant interpretation	Challenges	Possible improvements
Deep phenotyping	Identification of clinical gene-related hallmarks	International natural history studies on large cohorts of patients; a large consensus on the diagnostic and prognostic value of each test/hallmark
Population data: allele frequency threshold	Phenotypic divergence (1 gene = several diseases)	Large epidemiological studies
Phasing/segregation	Time-consuming and cost-ineffective PCR-based analysis	Novel sequencing technologies, TRIO or multi-sample sequencing
Elusive variants	Repetitive regions, low covered areas, CNV-prone sequences, cryptic splice-causing variants	Improved computational tools, novel sequencing technologies, second-tier tests
Variant annotation/functional validation:		
In *silico* tools	Conflicting predictions; uncertain accuracy	Improved (more accurate) computational tools
In vitro experiments	Large proteins to be dissected in more manageable fragments	Benchmark assays
In vivo or *ex vivo* experiments	High cost, non-scalability	International multidisciplinary consortia
Public disease-databases	Not standardized interpretation; limited number of shared variants	Sharing data; gene/disease-tailored guidelines for an improved variant interpretation

As discussed by Dr Rodenburg in his recent review [85], obtaining functional evidence of pathogenicity requires huge work and money. This effort brings a reward in terms of scientific impact (and of granted funds) when the novel variants are in novel disease genes. The same effort is much less rewarding in a diagnostic setting when the variants are in very well-known disease-genes. However, a correct diagnosis is important for patients and an international collaborative effort aiming at setting up and validating functional assays for the genes discussed here is strongly advisable.

Finally, HTS has contributed to the identification of digenic, or even more complex, genetic mechanisms underlying human diseases [129, 132, 146]. This should be considered when evaluating the functional and clinical impact of variants of unknown significance.

Our understanding of large genes will benefit from large, international and interdisciplinary consortia [147–150]. A larger cohort of patients, shared clinical and genetic data and shared scientific and technological resources are needed for these complex challenges [148–150]. Similarly, making available experimentally identified variants and the interpretation of their clinical significance through public databases will help to standardize the assessment of variant pathogenicity among different laboratories [151–153].

A perfect synergy among scientists and clinicians with a multidisciplinary expertise are required to set up a full translational research, going from variant identification in patients to characterization of pathophysiological mechanisms in muscular cells and animal models, from basic research to clinical developments - all for the benefit of the patients [148, 150].

ACKNOWLEDGMENTS

REFERENCES

[1] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921.

[2] Lynch M, Conery JS. The origins of genome complexity. Science. 2003;302(5649):1401-4.

[3] Koenig M, Hoffman EP, Bertelson CJ, Monaco AP, Feener C, Kunkel LM. Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell. 1987;50(3):509-17.

[4] Sahakyan AB, Balasubramanian S. Long genes and genes with multiple splice variants are enriched in pathways linked to cancer and other multigenic diseases. BMC Genomics. 2016;17:225.

[5] Grishkevich V, Yanai I. Gene length and expression level shape genomic novelties. Genome Res. 2014;24(9):1497-503.

[6] Higgins DG, Labeit S, Gautel M, Gibson TJ. The evolution of titin and related giant muscle proteins. J Mol Evol. 1994;38(4):395-404.

[7] Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. NPJ Genom Med. 2018;3:16.

[8] Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet. 2010;11(1):31-46.

[9] Gutowska-Ding MW, Deans ZC, Roos C, Matilainen J, Khawaja F, Brugger K, et al. One byte at a time: evidencing the quality of clinical service next-generation sequencing for germline and somatic variants. Eur J Hum Genet. 2019.

[10] Nigro V, Savarese M. Next-generation sequencing approaches for the diagnosis of skeletal muscle disorders. Curr Opin Neurol. 2016;29(5):621-7.

[11] Zenagui R, Lacourt D, Pegeot H, Yauy K, Juntas Morales R, Theze C, et al. A reliable targeted next-generation sequencing strategy for diagnosis of myopathies and muscular dystrophies. Especially for the Giant Titin and Nebulin Genes. J Mol Diagn. 2018;20(4):533-49.

[12] Abdalla E, Ravenscroft G, Zayed L, Beecroft SJ, Laing NG. Lethal multiple pterygium syndrome: A severe phenotype associated with a novel mutation in the nebulin gene. Neuromuscul Disord. 2017;27(6):537-41.

[13] Savarese M, Maggi L, Vihola A, Jonson PH, Tasca G, Ruggiero L, et al. Interpreting genetic variants in titin in patients with muscle disorders. JAMA Neurol. 2018;75(5):557-65.

[14] Jokela M, Tasca G, Vihola A, Mercuri E, Jonson PH, Lehtinen S, et al. An unusual ryanodine receptor 1 (RYR1) phenotype: Mild calf-predominant myopathy. Neurology. 2019;92(14):e1600-e9.

[15] Bonne G, Rivier F, Hamroun D. The 2019 version of the gene table of neuromuscular disorders (nuclear genome). Neuromuscul Disord. 2018;28(12):1031-63.

[16] Bang ML, Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, et al. The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system. Circ Res. 2001;89(11):1065-72.

[17] Savarese M, Sarparanta J, Vihola A, Udd B, Hackman P. Increasing role of titin mutations in neuromus-

cular disorders. Journal of Neuromuscular Diseases. 2016;3(3):293-308.

[18] Evila A, Palmio J, Vihola A, Savarese M, Tasca G, Penttila S, et al. Targeted next-generation sequencing reveals novel TTN mutations causing recessive distal titinopathy. Mol Neurobiol. 2017;54(9):7212-23.

[19] Oates EC, Jones KJ, Donkervoort S, Charlton A, Brammah S, Smith JE, 3rd, et al. Congenital titinopathy: Comprehensive characterization and pathogenic insights. Ann Neurol. 2018;83(6):1105-24.

[20] Hackman P, Vihola A, Haravuori H, Marchand S, Sarparanta J, De Seze J, et al. Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin. Am J Hum Genet. 2002;71(3):492-500.

[21] Palmio J, Leonard-Louis S, Sacconi S, Savarese M, Penttila S, Semmler AL, et al. Expanding the importance of HMERF titinopathy: New mutations and clinical aspects. J Neurol. 2019;266(3):680-90.

[22] Tasca G, Udd B. Hereditary myopathy with early respiratory failure (HMERF): Still rare, but common enough. Neuromuscul Disord. 2018;28(3):268-76.

[23] Bang ML, Caremani M, Brunello E, Littlefield R, Lieber RL, Chen J, et al. Nebulin plays a direct role in promoting strong actin-myosin interactions. FASEB journal : Official publication of the Federation of American Societies for Experimental Biology. 2009;23(12):4117-25.

[24] Lehtokari VL, Kiiski K, Sandaradura SA, Laporte J, Repo P, Frey JA, et al. Mutation update: the spectra of nebulin variants and associated myopathies. Hum Mutat. 2014;35(12):1418-26.

[25] Lehtokari VL, Pelin K, Herczegfalvi A, Karcagi V, Pouget J, Franques J, et al. Nemaline myopathy caused by mutations in the nebulin gene may present as a distal myopathy. Neuromuscul Disord. 2011;21(8):556-62.

[26] Wallgren-Pettersson C, Lehtokari VL, Kalimo H, Paetau A, Nuutinen E, Hackman P, et al. Distal myopathy caused by homozygous missense mutations in the nebulin gene. Brain. 2007;130(Pt 6):1465-76.

[27] Romero NB, Lehtokari VL, Quijano-Roy S, Monnier N, Claeys KG, Carlier RY, et al. Core-rod myopathy caused by mutations in the nebulin gene. Neurology. 2009;73(14):1159-61.

[28] Feingold-Zadok M, Chitayat D, Chong K, Injeyan M, Shannon P, Chapmann D, et al. Mutations in the NEB gene cause fetal akinesia/arthrogryposis multiplex congenita. Prenat Diagn. 2017;37(2):144-50.

[29] Sagath L, Lehtokari VL, Valipakka S, Udd B, Wallgren-Pettersson C, Pelin K, et al. An extended targeted copy number variation detection array including 187 genes for the diagnostics of neuromuscular disorders. J Neuromuscul Dis. 2018;5(3):307-14.

[30] Kiiski KJ, Lehtokari VL, Vihola AK, Laitila JM, Huovinen S, Sagath LJ, et al. Dominantly inherited distal nemaline/cap myopathy caused by a large deletion in the nebulin gene. Neuromuscul Disord. 2019;29(2):97-107.

[31] Lanner JT, Georgiou DK, Joshi AD, Hamilton SL. Ryanodine receptors: structure, expression, molecular details, and function in calcium release. Cold Spring Harb Perspect Biol. 2010;2(11):a003996.

[32] Lawal TA, Todd JJ, Meilleur KG. Ryanodine Receptor 1-Related Myopathies: Diagnostic and therapeutic approaches. Neurotherapeutics. 2018;15(4):885-99.

[33] Ferreiro A, Monnier N, Romero NB, Leroy JP, Bonnemann C, Haenggeli CA, et al. A recessive form of central

core disease, transiently presenting as multi-minicore disease, is associated with a homozygous mutation in the ryanodine receptor type 1 gene. Ann Neurol. 2002;51(6):750-9.

[34] Wu S, Ibarra MC, Malicdan MC, Murayama K, Ichihara Y, Kikuchi H, et al. Central core disease is due to RYR1 mutations in more than 90% of patients. Brain. 2006;129(Pt 6):1470-80.

[35] Clarke NF, Waddell LB, Cooper ST, Perry M, Smith RL, Kornberg AJ, et al. Recessive mutations in RYR1 are a common cause of congenital fiber type disproportion. Hum Mutat. 2010;31(7):E1544-50.

[36] Fattori F, Maggi L, Bruno C, Cassandrini D, Codemo V, Catteruccia M, et al. Centronuclear myopathies: genotype-phenotype correlation and frequency of defined genetic forms in an Italian cohort. J Neurol. 2015;262(7):1728-40.

[37] Cassandrini D, Trovato R, Rubegni A, Lenzi S, Fiorillo C, Baldacci J, et al. Congenital myopathies: Clinical phenotypes and new diagnostic tools. Ital J Pediatr. 2017;43(1):101.

[38] Gillard EF, Otsu K, Fujii J, Khanna VK, de Leon S, Derdemezi J, et al. A substitution of cysteine for arginine 614 in the ryanodine receptor is potentially causative of human malignant hyperthermia. Genomics. 1991;11(3):751-5.

[39] Dlamini N, Voermans NC, Lillis S, Stewart K, Kamsteeg EJ, Drost G, et al. Mutations in RYR1 are a common cause of exertional myalgia and rhabdomyolysis. Neuromuscul Disord. 2013;23(7):540-8.

[40] Park E, Pan Z, Zhang Z, Lin L, Xing Y. The expanding landscape of alternative splicing variation in human populations. Am J Hum Genet. 2018;102(1):11-26.

[41] Guo W, Bharmal SJ, Esbona K, Greaser ML. Titin diversity—alternative splicing gone wild. J Biomed Biotechnol. 2010;2010:753675.

[42] Greaser ML, Krzesinski PR, Warren CM, Kirkpatrick B, Campbell KS, Moss RL. Developmental changes in rat cardiac titin/connectin: transitions in normal animals and in mutants with a delayed pattern of isoform transition. J Muscle Res Cell Motil. 2005;26(6-8):325-32.

[43] Savarese M, Jonson PH, Huovinen S, Paulin L, Auvinen P, Udd B, et al. The complexity of titin splicing pattern in human adult skeletal muscles. Skelet Muscle. 2018;8(1):11.

[44] Laitila J, Hanif M, Paetau A, Hujanen S, Keto J, Somervuo P, et al. Expression of multiple nebulin isoforms in human skeletal muscle and brain. Muscle Nerve. 2012;46(5):730-7.

[45] Lam LT, Holt I, Laitila J, Hanif M, Pelin K, Wallgren-Pettersson C, et al. Two alternatively-spliced human nebulin isoforms with either exon 143 or exon 144 and their developmental regulation. Sci Rep. 2018;8(1):15728.

[46] Kimura T, Nakamori M, Lueck JD, Pouliquin P, Aoike F, Fujimura H, et al. Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca2+-ATPase in myotonic dystrophy type 1. Hum Mol Genet. 2005;14(15):2189-200.

[47] Tang Y, Wang H, Wei B, Guo Y, Gu L, Yang Z, et al. CUG-BP1 regulates RyR1 ASI alternative splicing in skeletal muscle atrophy. Sci Rep. 2015;5:16083.

[48] Guo W, Schafer S, Greaser ML, Radke MH, Liss M, Govindarajan T, et al. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. Nat Med. 2012;18(5):766-73.

[49] Bertero A, Fields PA, Ramani V, Bonora G, Yardimci GG, Reinecke H, et al. Dynamics of genome reorganization

during human cardiogenesis reveal an RBM20-dependent splicing factory. Nat Commun. 2019;10(1):1538.

[50] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. 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(5):405-24.

[51] Delude CM. Deep phenotyping: The details of disease. Nature. 2015;527(7576):S14-5.

[52] Barp A, Laforet P, Bello L, Tasca G, Vissing J, Monforte M, et al. European muscle MRI study in limb girdle muscular dystrophy type R1/2A (LGMDR1/LGMD2A). J Neurol. 2019.

[53] GoMez-Andres D, Diaz-Manera J, Alejaldre A, Pulido-Valdeolivas I, GonzaLez-Mera L, Olive M, et al. Muscle imaging in laminopathies: Synthesis study identifies meaningful muscles for follow-up. Muscle Nerve. 2018;58(6):812-7.

[54] Alonso-Jimenez A, Kroon R, Alejaldre-Monforte A, Nunez-Peralta C, Horlings CGC, van Engelen BGM, et al. Muscle MRI in a large cohort of patients with oculopharyngeal muscular dystrophy. J Neurol Neurosurg Psychiatry. 2019;90(5):576-85.

[55] Jungbluth H. Myopathology in times of modern imaging. Neuropathol Appl Neurobiol. 2017;43(1):24-43.

[56] Evila A, Vihola A, Sarparanta J, Raheem O, Palmio J, Sandell S, et al. Atypical phenotypes in titinopathies explained by second titin mutations. Ann Neurol. 2014;75(2):230-40.

[57] Jungbluth H, Davis MR, Muller C, Counsell S, Allsop J, Chattopadhyay A, et al. Magnetic resonance imaging of muscle in congenital myopathies associated with RYR1 mutations. Neuromuscul Disord. 2004;14(12):785-90.

[58] Jungbluth H, Muller CR, Halliger-Keller B, Brockington M, Brown SC, Feng L, et al. Autosomal recessive inheritance of RYR1 mutations in a congenital myopathy with cores. Neurology. 2002;59(2):284-7.

[59] Jungbluth H, Sewry CA, Counsell S, Allsop J, Chattopadhyay A, Mercuri E, et al. Magnetic resonance imaging of muscle in nemaline myopathy. Neuromuscul Disord. 2004;14(12):779-84.

[60] Savarese M, Johari M, Johnson K, Arumilli M, Torella A, Topf A, et al. Improved Criteria for the Classification of Titin Variants in Inherited Skeletal Myopathies. J Neuromuscul Dis. 2020.

[61] Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003;31(13):3812-4.

[62] 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(4):248-9.

[63] Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: Mutation prediction for the deep-sequencing age. Nat Methods. 2014;11(4):361-2.

[64] Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum Mol Genet. 2015;24(8):2125-37.

[65] Sundaram L, Gao H, Padigepati SR, McRae JF, Li Y, Kosmicki JA, et al. Predicting the clinical impact of human mutation with deep neural networks. Nat Genet. 2018;50(8):1161-70.

[66] Laddach A, Gautel M, Fraternali F. TITINdb-a computational tool to assess titin's role as a disease gene. Bioinformatics. 2017;33(21):3482-5.

[67] Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46(3):310-5.

[68] Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet. 2016;99(4):877-85.

[69] Tian Y, Pesaran T, Chamberlin A, Fenwick RB, Li S, Gau CL, et al. REVEL and BayesDel outperform other in silico meta-predictors for clinical variant classification. Sci Rep. 2019;9(1):12752.

[70] Hunt RC, Simhadri VL, Iandoli M, Sauna ZE, Kimchi-Sarfaty C. Exposing synonymous mutations. Trends Genet. 2014;30(7):308-21.

[71] Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: Exonic mutations that affect splicing. Nat Rev Genet. 2002;3(4):285-98.

[72] Spencer PS, Siller E, Anderson JF, Barral JM. Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. J Mol Biol. 2012;422(3):328-35.

[73] Shi F, Yao Y, Bin Y, Zheng CH, Xia J. Computational identification of deleterious synonymous variants in human genomes using a feature-based approach. BMC Med Genomics. 2019;12(Suppl 1):12.

[74] 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(9):e67.

[75] Leman R, Gaildrat P, Gac GL, Ka C, Fichou Y, Audrezet MP, et al. Novel diagnostic tool for prediction of variant spliceogenicity derived from a set of 395 combined in silico/in vitro studies: An international collaborative effort. Nucleic Acids Res. 2018;46(15):7913-23.

[76] Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting splicing from primary sequence with deep learning. Cell. 2019;176(3):535-48 e24.

[77] Chauveau C, Bonnemann CG, Julien C, Kho AL, Marks H, Talim B, et al. Recessive TTN truncating mutations define novel forms of core myopathy with heart disease. Hum Mol Genet. 2014;23(4):980-91.

[78] Hastings R, de Villiers CP, Hooper C, Ormondroyd L, Pagnamenta A, Lise S, et al. Combination of whole genome sequencing, linkage, and functional studies implicates a missense mutation in titin as a cause of autosomal dominant cardiomyopathy with features of left ventricular noncompaction. Circ Cardiovasc Genet. 2016;9(5):426-35.

[79] Marttila M, Hanif M, Lemola E, Nowak KJ, Laitila J, Gronholm M, et al. Nebulin interactions with actin and tropomyosin are altered by disease-causing mutations. Skelet Muscle. 2014;4:15.

[80] Laitila J, Lehtonen J, Lehtokari VL, Sagath L, Wallgren-Pettersson C, Gronholm M, et al. A nebulin super-repeat panel reveals stronger actin binding toward the ends of the super-repeat region. Muscle Nerve. 2019;59(1):116-21.

[81] Tong J, McCarthy TV, MacLennan DH. Measurement of resting cytosolic Ca2+concentrations and Ca2+store size in HEK-293 cells transfected with malignant hyperthermia

or central core disease mutant Ca2+release channels. J Biol Chem. 1999;274(2):693-702.

[82] Tong J, Oyamada H, Demaurex N, Grinstein S, McCarthy TV, MacLennan DH. Caffeine and halothane sensitivity of intracellular Ca2+release is altered by 15 calcium release channel (ryanodine receptor) mutations associated with malignant hyperthermia and/or central core disease. J Biol Chem. 1997;272(42):26332-9.

[83] Monnier N, Kozak-Ribbens G, Krivosic-Horber R, Nivoche Y, Qi D, Kraev N, et al. Correlations between genotype and pharmacological, histological, functional, and clinical phenotypes in malignant hyperthermia susceptibility. Hum Mutat. 2005;26(5):413-25.

[84] Bryen SJ, Ewans LJ, Pinner J, MacLennan SC, Donkervoort S, Castro D, et al. Recurrent TTN metatranscript-only c.39974-11T>G splice variant associated with autosomal recessive arthrogryposis multiplex congenita and myopathy. Hum Mutat. 2019.

[85] Rodenburg RJ. The functional genomics laboratory: Functional validation of genetic variants. J Inherit Metab Dis. 2018;41(3):297-307.

[86] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126(4):663-76.

[87] Wang G, McCain ML, Yang L, He A, Pasqualini FS, Agarwal A, et al. Modeling the mitochondrial cardiomyopathy of Barth syndrome with induced pluripotent stem cell and heart-on-chip technologies. Nat Med. 2014;20(6):616-23.

[88] Ford K, McDonald D, Mali P. Functional Genomics via CRISPR-Cas. J Mol Biol. 2019;431(1):48-65.

[89] Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. RNA-guided human genome engineering via Cas9. Science. 2013;339(6121):823-6.

[90] Lornage X, Romero NB, Grosgogeat CA, Malfatti E, Donkervoort S, Marchetti MM, et al. ACTN2 mutations cause "Multiple structured Core Disease" (MsCD). Acta Neuropathol. 2019.

[91] Sarparanta J, Jonson PH, Golzio C, Sandell S, Luque H, Screen M, et al. Mutations affecting the cytoplasmic functions of the co-chaperone DNAJB6 cause limb-girdle muscular dystrophy. Nat Genet. 2012;44(4):450-5, S1-2.

[92] Hirata H, Watanabe T, Hatakeyama J, Sprague SM, Saint-Amant L, Nagashima A, et al. Zebrafish relatively relaxed mutants have a ryanodine receptor defect, show slow swimming and provide a model of multi-minicore disease. Development. 2007;134(15):2771-81.

[93] Steffen LS, Guyon JR, Vogel ED, Howell MH, Zhou Y, Weber GJ, et al. The zebrafish runzel muscular dystrophy is linked to the titin gene. Dev Biol. 2007;309(2):180-92.

[94] Telfer WR, Nelson DD, Waugh T, Brooks SV, Dowling JJ. Neb: A zebrafish model of nemaline myopathy due to nebulin mutation. Dis Model Mech. 2012;5(3):389-96.

[95] Charton K, Daniele N, Vihola A, Roudaut C, Gicquel E, Monjaret F, et al. Removal of the calpain 3 protease reverses the myopathology in a mouse model for titinopathies. Hum Mol Genet. 2010;19(23):4608-24.

[96] Lee CS, Hanna AD, Wang H, Dagnino-Acosta A, Joshi AD, Knoblauch M, et al. A chemical chaperone improves muscle function in mice with a RyR1 mutation. Nat Commun. 2017;8:14659.

[97] Durham WJ, Aracena-Parks P, Long C, Rossi AE, Goonasekera SA, Boncompagni S, et al. RyR1 S-nitrosylation underlies environmental heat stroke and sudden death in Y522S RyR1 knockin mice. Cell. 2008;133(1):53-65.

[98] Kushnir A, Betzenhauser MJ, Marks AR. Ryanodine receptor studies using genetically engineered mice. FEBS Lett. 2010;584(10):1956-65.

[99] Yang T, Riehl J, Esteve E, Matthaei KI, Goth S, Allen PD, et al. Pharmacologic and functional characterization of malignant hyperthermia in the R163C RyR1 knock-in mouse. Anesthesiology. 2006;105(6):1164-75.

[100] Laitila J, McNamara E, Goullee H, Lawlor M, Ochala J, Griffiths L, et al. A mouse model with compound heterozygous nebulin mutations recapitulates the typical form of nemaline myopathy. Neuromuscular Disorders. 2017;27:S183.

[101] Laitila JM, McNamara EL, Wingate CD, Goullee H, Ross JA, Taylor RL, et al. Nebulin nemaline myopathy recapitulated in a compound heterozygous mouse model with both a missense and a nonsense mutation in Neb. Acta Neuropathol Commun. 2020;8(1):18.

[102] Koob M, Karanjeet K, Benzow K. Moving human genetics into the mouse: Full human gene-replacement models. Presented at the American Society of Human Genetics 2019 Annual Meeting Houston, Texas. 2019.

[103] Kosugi S, Momozawa Y, Liu X, Terao C, Kubo M, Kamatani Y. Comprehensive evaluation of structural variation detection algorithms for whole genome sequencing. Genome Biol. 2019;20(1):117.

[104] Truty R, Paul J, Kennemer M, Lincoln SE, Olivares E, Nussbaum RL, et al. Prevalence and properties of intragenic copy-number variation in Mendelian disease genes. Genet Med. 2019;21(1):114-23.

[105] Pfundt R, Del Rosario M, Vissers L, Kwint MP, Janssen IM, de Leeuw N, et al. Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. Genet Med. 2017;19(6):667-75.

[106] Roca I, Gonzalez-Castro L, Fernandez H, Couce ML, Fernandez-Marmiesse A. Free-access copy-number variant detection tools for targeted next-generation sequencing data. Mutat Res. 2019;779:114-25.

[107] Mason-Suares H, Landry L, S. Lebo MJCGMR. Detecting copy number variation via next generation technology. 2016;4(3):74-85.

[108] Kerkhof J, Schenkel LC, Reilly J, McRobbie S, Aref-Eshghi E, Stuart A, et al. Clinical validation of copy number variant detection from targeted next-generation sequencing panels. J Mol Diagn. 2017;19(6):905-20.

[109] Sadedin SP, Ellis JA, Masters SL, Oshlack A. Ximmer: a system for improving accuracy and consistency of CNV calling from exome data. Gigascience. 2018;7(10).

[110] Zhao M, Wang Q, Wang Q, Jia P, Zhao Z. Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. BMC Bioinformatics. 2013;14 Suppl 11:S1.

[111] Välipakka S, Savarese M, Sagath L, Arumilli M, Giugliano T, Udd B, et al. Improving copy number variant detection from sequencing data with a combination of programs and a predictive model. The Journal of Molecular Diagnostics. 2020;22(1):40-49.

[112] Tuffery-Giraud S, Beroud C, Leturcq F, Yaou RB, Hamroun D, Michel-Calemard L, et al. Genotype-phenotype analysis in 2,405 patients with a dystrophinopathy using the UMD-DMD database: a model of nationwide knowledgebase. Hum Mutat. 2009;30(6):934-45.

[113] Aartsma-Rus A, Ginjaar IB, Bushby K. The importance of genetic diagnosis for Duchenne muscular dystrophy. J Med Genet. 2016;53(3):145-51.

[114] Kozareva V, Stroff C, Silver M, Freidin JF, Delaney NF. Clinical analysis of germline copy number variation in DMD using a non-conjugate hierarchical Bayesian model. BMC Med Genomics. 2018;11(1):91.

[115] Välipakka S, Savarese M, Johari M, Sagath L, Arumilli M, Kiiski K, et al. Copy number variation analysis increases the diagnostic yield in muscle diseases. Neurology Genetics. 2017;3(6).

[116] Kiiski K, Lehtokari VL, Loytynoja A, Ahlsten L, Laitila J, Wallgren-Pettersson C, et al. A recurrent copy number variation of the NEB triplicate region: Only revealed by the targeted nemaline myopathy CGH array. Eur J Hum Genet. 2016;24(4):574-80.

[117] Roggenbuck J, Rich K, Morales A, Tan CA, Eck D, King W, et al. A novel TTN deletion in a family with skeletal myopathy, facial weakness, and dilated cardiomyopathy. Mol Genet Genomic Med. 2019;7(11):e924.

[118] Giugliano T, Savarese M, Garofalo A, Picillo E, Fiorillo C, D'Amico A, et al. Copy Number Variants Account for a Tiny Fraction of Undiagnosed Myopathic Patients. Genes (Basel). 2018;9(11).

[119] Savarese M, Musumeci O, Giugliano T, Rubegni A, Fiorillo C, Fattori F, et al. Novel findings associated with MTM1 suggest a higher number of female symptomatic carriers. Neuromuscul Disord. 2016;26(4-5):292-9.

[120] Tian X, Liang WC, Feng Y, Wang J, Zhang VW, Chou CH, et al. Expanding genotype/phenotype of neuromuscular diseases by comprehensive target capture/NGS. Neurology Genetics. 2015;1(2):e14.

[121] Piluso G, Dionisi M, Del Vecchio Blanco F, Torella A, Aurino S, Savarese M, et al. Motor chip: a comparative genomic hybridization microarray for copy-number mutations in 245 neuromuscular disorders. Clin Chem. 2011;57(11):1584-96.

[122] Giugliano T, Fanin M, Savarese M, Piluso G, Angelini C, Nigro V. Identification of an intragenic deletion in the SGCB gene through a re-evaluation of negative next generation sequencing results. Neuromuscul Disord. 2016;26(6):367-9.

[123] Fichna JP, Macias A, Piechota M, Korostynski M, Potulska-Chromik A, Redowicz MJ, et al. Whole-exome sequencing identifies novel pathogenic mutations and putative phenotype-influencing variants in Polish limb-girdle muscular dystrophy patients. Hum Genomics. 2018;12(1):34.

[124] Nowakowska B. Clinical interpretation of copy number variants in the human genome. J Appl Genet. 2017;58(4):449-57.

[125] Haraksingh RR, Abyzov A, Gerstein M, Urban AE, Snyder M. Genome-wide mapping of copy number variation in humans: comparative analysis of high resolution array platforms. PLoS One. 2011;6(11):e27859.

[126] Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). Genet Med. 2019.

[127] Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Sci Transl Med. 2017;9(386).

[128] Gonorazky HD, Naumenko S, Ramani AK, Nelakuditi V, Mashouri P, Wang P, et al. Expanding the Boundaries of RNA Sequencing as a Diagnostic Tool for Rare Mendelian Disease. Am J Hum Genet. 2019;104(3):466-83.

[129] Lee Y, Jonson PH, Sarparanta J, Palmio J, Sarkar M, Vihola A, et al. TIA1 variant drives myodegeneration in multisystem proteinopathy with SQSTM1 mutations. J Clin Invest. 2018;128(3):1164-77.

[130] D'Amico A, Fattori F, Tasca G, Petrini S, Gualandi F, Bruselles A, et al. Somatic mosaicism represents an underestimated event underlying collagen 6-related disorders. Eur J Paediatr Neurol. 2017;21(6):873-83.

[131] Donkervoort S, Hu Y, Stojkovic T, Voermans NC, Foley AR, Leach ME, et al. Mosaicism for dominant collagen 6 mutations as a cause for intrafamilial phenotypic variability. Hum Mutat. 2015;36(1):48-56.

[132] Pehlivan D, Bayram Y, Gunes N, Coban Akdemir Z, Shukla A, Bierhals T, et al. The Genomics of Arthrogryposis, a Complex Trait: Candidate Genes and Further Evidence for Oligogenic Inheritance. Am J Hum Genet. 2019;105(1):132-50.

[133] Lee H, Huang AY, Wang LK, Yoon AJ, Renteria G, Eskin A, et al. Diagnostic utility of transcriptome sequencing for rare Mendelian diseases. Genet Med. 2019.

[134] Hamanaka K, Miyatake S, Koshimizu E, Tsurusaki Y, Mitsuhashi S, Iwama K, et al. RNA sequencing solved the most common but unrecognized NEB pathogenic variant in Japanese nemaline myopathy. Genet Med. 2019;21(7):1629-38.

[135] Greer SU, Nadauld LD, Lau BT, Chen J, Wood-Bouwens C, Ford JM, et al. Linked read sequencing resolves complex genomic rearrangements in gastric cancer metastases. Genome Med. 2017;9(1):57.

[136] Marks P, Garcia S, Barrio AM, Belhocine K, Bernate J, Bharadwaj R, et al. Resolving the full spectrum of human genome variation using Linked-Reads. Genome Res. 2019;29(4):635-45.

[137] Pollard MO, Gurdasani D, Mentzer AJ, Porter T, Sandhu MS. Long reads: their purpose and place. Hum Mol Genet. 2018;27(R2):R234-R41.

[138] Jang SS, Lim BC, Yoo SK, Shin JY, Kim KJ, Seo JS, et al. Targeted linked-read sequencing for direct haplotype phasing of maternal DMD alleles: a practical and reliable method for noninvasive prenatal diagnosis. Sci Rep. 2018;8(1):8678.

[139] Di Fruscio G, Garofalo A, Mutarelli M, Savarese M, Nigro V. Are all the previously reported genetic variants in limb girdle muscular dystrophy genes pathogenic? Eur J Hum Genet. 2016;24(1):73-7.

[140] Andreasen C, Nielsen JB, Refsgaard L, Holst AG, Christensen AH, Andreasen L, et al. New population-based exome data are questioning the pathogenicity of previously cardiomyopathy-associated genetic variants. Eur J Hum Genet. 2013;21(9):918-28.

[141] Liu P, Meng L, Normand EA, Xia F, Song X, Ghazi A, et al. Reanalysis of clinical exome sequencing data. N Engl J Med. 2019;380(25):2478-80.

[142] Appelbaum PS, Parens E, Berger SM, Chung WK, Burke W. Is there a duty to reinterpret genetic data? The ethical dimensions. Genet Med. 2019.

[143] Brnich SE, Abou Tayoun AN, Couch FJ, Cutting GR, Greenblatt MS, Heinen CD, et al. Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. Genome Med. 2019;12(1):3.

[144] Gelman H, Dines JN, Berg J, Berger AH, Brnich S, Hisama FM, et al. Recommendations for the collection and use of

multiplexed functional data for clinical variant interpretation. Genome Med. 2019;11(1):85.

[145]   Kelly MA, Caleshu C, Morales A, Buchan J, Wolf Z, Harrison SM, et al. Adaptation and validation of the ACMG/AMP variant classification framework for MYH7-associated inherited cardiomyopathies: Recommendations by ClinGen's Inherited Cardiomyopathy Expert Panel. Genet Med. 2018.

[146]   Karaca E, Posey JE, Coban Akdemir Z, Pehlivan D, Harel T, Jhangiani SN, et al. Phenotypic expansion illuminates multilocus pathogenic variation. Genet Med. 2018;20(12):1528-37.

[147]   Hackman P, Udd B, Bonnemann CG, Ferreiro A, Titinopathy Database C. 219th ENMC International Workshop Titinopathies International database of titin mutations and phenotypes, Heemskerk, The Netherlands, 29 April-1 May 2016. Neuromuscul Disord. 2017;27(4):396-407.

[148]   Boycott KM, Rath A, Chong JX, Hartley T, Alkuraya FS, Baynam G, et al. International cooperation to enable the diagnosis of all rare genetic diseases. Am J Hum Genet. 2017;100(5):695-705.

[149]   Lochmuller H, Badowska DM, Thompson R, Knoers NV, Aartsma-Rus A, Gut I, et al. RD-Connect, NeurOmics and EURenOmics: Collaborative European initiative for rare diseases. Eur J Hum Genet. 2018;26(6):778-85.

[150]   Lochmuller H, Torrent IFJ, Le Cam Y, Jonker AH, Lau LP, Baynam G, et al. The International Rare Diseases Research Consortium: Policies and Guidelines to maximize impact. Eur J Hum Genet. 2017;25(12):1293-302.

[151]   Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: Improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 2018;46(D1):D1062-D7.

[152]   Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: The next generation in gene variant databases. Hum Mutat. 2011;32(5):557-63.

[153]   Amendola LM, Jarvik GP, Leo MC, McLaughlin HM, Akkari Y, Amaral MD, et al. Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. Am J Hum Genet. 2016;98(6):1067-76.