ACTA MYOL 2023;XLII:123-128
doi: 10.36185/2532-1900-394

REVIEW

Long-read sequencing improves diagnostic rate in neuromuscular disorders

Rafaela Owusu[1], Marco Savarese[1,2]

[1] Folkhälsan Research Center, Helsinki, Finland; [2] University of Helsinki, Faculty of Medicine, Helsinki, Finland

Massive parallel sequencing methods, such as exome, genome, and targeted DNA sequencing, have aided molecular diagnosis of genetic diseases in the last 20 years. However, short-read sequencing methods still have several limitations, such inaccurate genome assembly, the inability to detect large structural variants, and variants located in hard-to-sequence regions like highly repetitive areas. The recently emerged PacBio single-molecule real-time (SMRT) and Oxford nanopore technology (ONT) long-read sequencing (LRS) methods have been shown to overcome most of these technical issues, leading to an increase in diagnostic rate.

LRS methods are contributing to the detection of repeat expansions in novel disease-causing genes (e.g., *ABCD3*, *NOTCH2NLC* and *RILPL1* causing an Oculopharyngodistal myopathy or *PLIN4* causing a Myopathy with rimmed ubiquitin-positive autophagic vacuolation), of structural variants (e.g., in *DMD*), and of single nucleotide variants in repetitive regions (*TTN* and *NEB*). Moreover, these methods have simplified the characterization of the D4Z4 repeats in *DUX4*, facilitating the diagnosis of Facioscapulohumeral muscular dystrophy (FSHD).

We review recent studies that have used either ONT or PacBio SMRT sequencing methods and discuss different types of variants that have been detected using these approaches in individuals with neuromuscular disorders.

Key words: nanopore sequencing, PacBio single-molecule real-time, neuromuscular diseases, structural variant, DNA repeat expansion

Received: November 30, 2023
Accepted: December 5, 2023
Published online: December 20, 2023

Correspondence
Marco Savarese
Folkhälsan Research Center, Biomedicum 1, Haartmaninkatu 8, 00290 Helsinki, Finland
E-mail: marco.savarese@helsinki.fi

How to cite this article: Owusu R, Savarese M. Long-read sequencing improves diagnostic rate in neuromuscular disorders. Acta Myol 2023;42:123-128. https://doi.org/10.36185/2532-1900-394

Diagnosis of neuromuscular disorders in the 21st century

Next-generation sequencing (NGS) or massive parallel sequencing (MPS) approaches were introduced in the early 21st century and their clinical use has aided molecular diagnosis of genetic diseases [1]. Interestingly, large genes associated with muscle diseases, such as *DMD* and *TTN*, were among the first targets of NGS-approaches [2,3]. Since then, NGS-based genetic tests have been developed to make the diagnosis of rare diseases more effective and accurate [4-7]. Gene panels and whole exome sequencing (WES) have become the first-tier test to identify disease-causing variants in NMDs as well as in other genetic diseases [8-17]. Although these methods have helped identify several new genes and variants causing these diseases, many patients still remain without a molecular diagnosis [8,18]. Clinically, many variants of uncertain significance (VUS) are difficult to interpret [8,10]. Reanalysis of previously identified VUS only partly increases the diagnostic rate, demonstrating that technical aspects may limit the overall detection of DNA causative variants [19]. In particular, the short-read length (usually 50-300 bp) is a clear limitation, especially in sequencing certain genomic regions such as highly repetitive areas or long homopolymers [20]. Similarly, localizing large structural variants in their entirety is extremely difficult using short reads [21-25].

The development of Long-read sequencing

Long-read sequencing (LRS), also called third-generation sequencing, generates reads with a size from 1000 bases to several kilobases (kb) [21,23]. LRS can detect both small and large structural variants, repeat expansions, and even epigenetic modifications, such as DNA methylation, with up to 99.9% accuracy [21,25,26]. Moreover, LRS eliminates bias associated with amplification since isolated DNA can be used directly for sequencing, and instead of clusters, single DNA molecules are sequenced, resulting in an improved coverage. LRS also aids variant phasing [26,27]. LRS can increase the diagnostic rate of genetic diseases (including neuromuscular disorders) and reduce the time it takes to achieve a molecular diagnosis [21,24,25,28,29]. Moreover, the combined use of LRS and of a more complete human reference genome (T2T-CHM13), also increases the detection of de novo variants [30].

In 2022, LRS was named the method of the year by Nature Methods due to the new opportunities and improvements it has given both individual labs and large-scale genomics projects. The Vertebrate Genomes Project (VGP) and the Telomere-to-Telomere Consortium (T2T) are two examples of the genomics initiatives that were made possible with the introduction of LRS [31].

Single-molecule real-time sequencing (SMRT) by Pacific Biosciences (PacBio) and Nanopore long-read sequencing by Oxford Nanopore Technologies (ONT) are the two leading LRS methods (Fig. 1) [21,32]. Both companies provide long, highly accurate reads in a short turnover time, although the processes work differently from each other. In PacBio SMRT sequencing, hairpin adapters are ligated to both ends of the target double-stranded DNA fragment, creating a template called the SMRTbell. Closed circular DNA molecules are formed, creating a library that is then loaded into flowcells containing nanoscale zero-mode waveguides (ZMWs). Fluorescently labeled polymerase at the base of the ZMWs excite fluorescent signals, which emit light that is captured in real-time, creating a circular consensus sequence (CCS) with high fidelity from multiple reads that cover the entire original DNA template. Oxford nanopore sequencing involves a nanopore embedded into a synthetic bilayer. With the help of motor proteins, DNA or RNA is unwound and translocated through the nanopore. Exonucleases, meanwhile, cut off individual nucleotides, disrupting the electric current which is captured in real time [23].

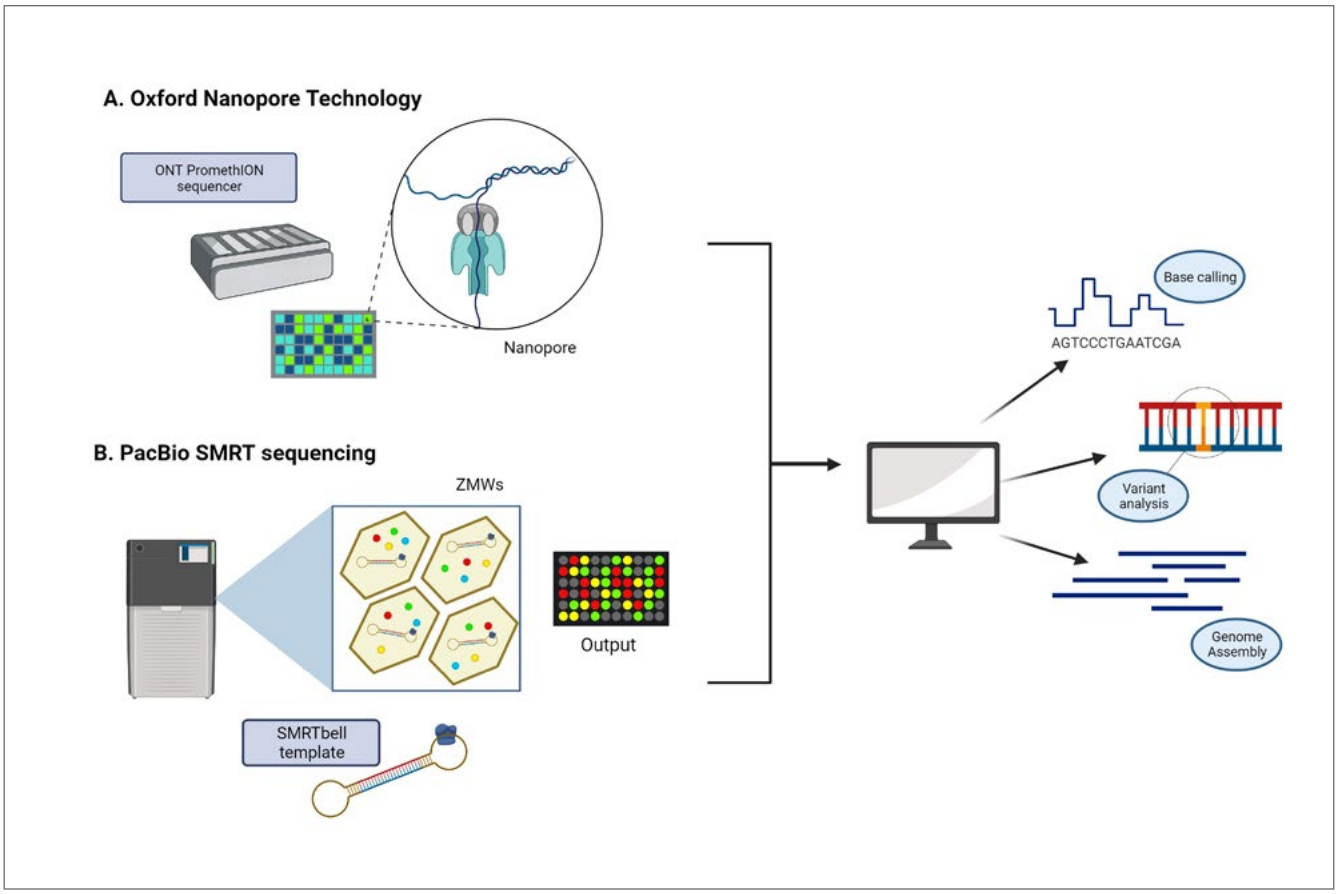

Figure 1. Long-read sequencing technologies by Oxford Nanopore Technologies and Pacific Biosciences. (A) ONT sequencing is based on a nanopore attached to a bilayer that unwinds the DNA/RNA strand. It goes through the nanopore, and individual nucleotides are then cut off with exonucleases, disrupting the electric current. This electric current is measured in real-time and captured in the output. (B) In PacBio SMRT sequencing, an SMRTbell template is formed when hairpin adapters are attached to the double-stranded DNA fragment. The SMRTbell forms a circular consensus sequence (CCS) that is loaded into a flowcell containing ZMWs. A fluorescently labeled polymerase is attached to the CCS at the base of the ZMW, which emits a fluorescent signal that can be measured in real-time. Figure created with Biorender.com

Long read sequencing for the identification and characterization of structural variants

In the field of NMDs research, LRS has been successfully used to detect novel rare structural variants, particularly in the *DMD* gene for individuals with Duchenne muscular dystrophy (DMD; MIM #310200) [33]. Bruels and colleagues used ONT sequencing to evaluate a cohort of unsolved patients with suspected pathogenic variants in muscular dystrophy genes. They collected blood and saliva samples from 12 unsolved individuals belonging to 10 different families. The average read length obtained using DNA from saliva samples was 7000 bp, while DNA from blood samples averaged a read length of 8300 bp. In an unsolved Duchenne patient, they identified one 5.9 Mb structural variant (SV); an inversion that disrupts all exons except 1 and 2 in the *DMD* gene. In addition, a *DMD* in-frame duplication was identified in an unrelated asymptomatic patient but was interpreted as benign after it was confirmed to be in tandem through nanopore. In another patient, a 3.6 Mb duplication encompassing exon 30, predicted to cause a frameshift and a premature termination, was identified in *LAMA2* in a patient presenting with sporadic congenital muscular dystrophy. The same variant was confirmed in the proband's mother. Finally, Nanopore also identified two novel splice-altering *DMD* variants in two other patients. These results were confirmed via PCR and Sanger, and in one case through a minigene splicing assay [33].

In a study by Geng and colleagues, ONT was performed secondary to RNA sequencing, cDNA capture sequencing, and optical mapping in a single family, but as the primary method in a different family. The long-read sequencing assay was customized with probes to cover the entire *DMD* gene as well as 20 kb upstream and downstream. In the first family, no pathogenic variants were found through regular genetic testing. Optical genome mapping (OGM) was later performed, and a pathogenic *DMD* variant was identified. The breakpoints of this inversion were confirmed through LRS. They revealed an 80 bp short interspersed nuclear element (SINE) and a 467 bp long terminal repeat (LTR). The inversion was predicted to affect exons 3-55. In the second family, LRS was directly performed, and an inversion was identified. In addition, several SINEs, LINEs, and LTRs were observed nearby [34].

A novel complex structural variant was identified by Xie and colleagues through long-read whole-genome sequencing (WGS) in a dystrophinopathy patient who remained without a molecular diagnosis after conventional genetic testing. LRS of the *DMD* gene identified a large-scale inversion/deletion-insertion rearrangement mediated by long interspersed nuclear element-1 (LINE-1) retrotransposons. Long-read WGS confirmed the structural variant was a 982323 bp inversion flanked by a 3719 bp deletion insertion [35]. In another recent study, LRS contributed to the identification of two Alu-mediated deletions in the *SMN1* gene and to correctly identify the breakpoints [36].

The clinical utility of LR genome sequencing in a prenatal setting for accurate and rapid characterization of structural variants was demonstrated in a complex case with a duplication involving *DMD* [37]. Chin and colleagues used nanopore sequencing to quickly sequence the whole genome of a healthy pregnant individual with a duplication of uncertain significance encompassing a portion of *DMD*. Her healthy daughter, as well as her male fetus, were found to have inherited the variant. Comparable duplications in the same location had been reported with varying clinical significance from benign to pathogenic. LRS contributed to the identification of the precise breakpoints of the duplication and proved that the duplicated region did not disrupt *DMD* and was in the same orientation as *DMD*. The variant was interpreted as likely benign and, later on, its identification in a healthy maternal uncle further supported this interpretation [37].

Repeat expansions and repeat sequences

In two studies by Yu and colleagues, DNA from patients with oculopharyngodistal myopathy (OPDM) types 3 (MIM #619473) and 4 (MIM #619790) was analyzed by long-read WGS using nanopore sequencing, after short-read WGS failed to identify a pathogenic variant. In the earlier study, LRS revealed a heterozygous GGC repeat in *NOTCH2NLC* in two patients and, after segregation studies, the gene was reported as the disease-causing gene [38,39]. In the later study, a heterozygous CCG repeat upstream of the *RILP1* gene was identified. As in the earlier study, the identified CCG repeat was reported as a novel OPDM-causing variant, and this finding was supported by segregation in the family [40].

Another recent study identified a CCG expansion, ranging from 118 to 694 repeats, in the 5′ UTR of *ABCD3* in OPDM patients from several unrelated families. The expansion results in upregulation of *ABCD3* expression [41].

Using ONT-sequencing, Yeetong and colleagues identified contracted D4Z4 repeats in seven individuals with Facioscapulohumeral muscular dystrophy (FSHD). The D4Z4 array was normal in the control groups, which included healthy individuals and unaffected parents of the FSHD patients [42].

Myotonic dystrophy type 1 (DM1; MIM: #160900) is caused by a CTG trinucleotide repeat expansion in *DMPK*. In some cases, the mutation can be up to 4000 triplets long [43,44]. Mangin and colleagues used SMRT LRS to sequence the DM1 locus in several patients, detecting de novo CCG interruptions and somatic mosaicism[44]. Similarly, Rasmussen and colleagues identified the CTG expansion in a cohort of DM1 patients using a Cas9-enrichment, combined with nanopore sequencing [45].

A study by Ruggieri and colleagues identified 99-mer repeat expansion in *PLIN4* causing myopathy. They used a multi-omics approach, combining genomic and transcriptomic data. Nanopore LRS of cDNA from RNA extracted from the muscle of affected patients revealed a 40x99-repeat sequence in exon 3 of *PLIN4*, compared to the normal 31x99-nucleotide sequence in unaffected individuals. This repeat expansion results in nearly 300 additional amino acids, which leads to an increased *PLIN4* expression in some muscle tissues [46].

Perrin and colleagues (2022) developed a strategy based on a long-range PCR combined with ONT sequencing to identify variants in the repeated sequences of *TTN*. The sequencing data helped to assign variants to specific exons and phase variants in the repeated regions [47].

Long-read sequencing for haplotyping variants

Haplotyping, using LRS, has been carried out successfully in research of many different diseases, including NMDs, in particular in Spinal muscular atrophy (SMA; MIM IDs: SMA1 #253300, SMA2 #253550, SMA3 #253400, SMA4 #271150) [48-50]. Two recent studies adopted HiFi long-reads generated with SMRT sequencing for haplotyping and phasing variants in the highly homologous *SMN1* and *SMN2* genes associated with SMA [48,50]. For this purpose, Chen and colleagues developed a new bioinformatics tool called 'Paraphase'. The pipeline identified full-length haplotypes in both genes, and the samples with more than 20x coverage were distributed into haplogroups through population-wide analysis. Co-segregation of the haplotypes was performed via pedigree-based analysis, which resulted in the identification of ten major haplogroups in *SMN1* and nine in *SMN2*[48].

By performing a comprehensive SMA trio-analysis (CASMA-trio), Li and colleagues detected silent carriers (SMA 2+0) and ascertained the inheritance patterns of *SMN1* and *SMN2* haplotypes in most of their families. The CASMA assay identified full-length *SMN* haplotypes by combining PacBio SMRT LRS with long-range PCR [50].

Conclusions

The benefits of using LRS for molecular diagnosis in rare diseases become evident by the increase in diagnostic rates observed in several studies resulting in the discovery of novel causative variants [30,33,46]. Both PacBio SMRT and ONT offer several advantages over conventional short-read sequencing methods (Fig. 2). While ONT offers the longest reads, SMRT has very high accuracy. Combining these two technologies or coupling them with other strategies, e.g., OGM, could give the most comprehensive result.

LRS is a powerful tool that can identify novel isoforms and transcript variants of very large genes [51,52]. Similarly, LRS may contribute to the characterization of splice variants in genes with complex splice patterns, such as *TTN* [53]. Finally, LRS also shows promise in the detection of structural alterations in the mitochondrial genome [54].

LRS has yet to become the standard tool for clinical applications due to its relatively high cost. However, we are observing a continued increase in quality and a decline in cost. Databases for LRS samples and standardized analytic pipelines are still needed. Several recent studies have already made efforts to facilitate the characterization of variants identified through LRS [48,55,56].

We are at the beginning of a new era of sequencing and novel exciting findings are expected to emerge through LRS.

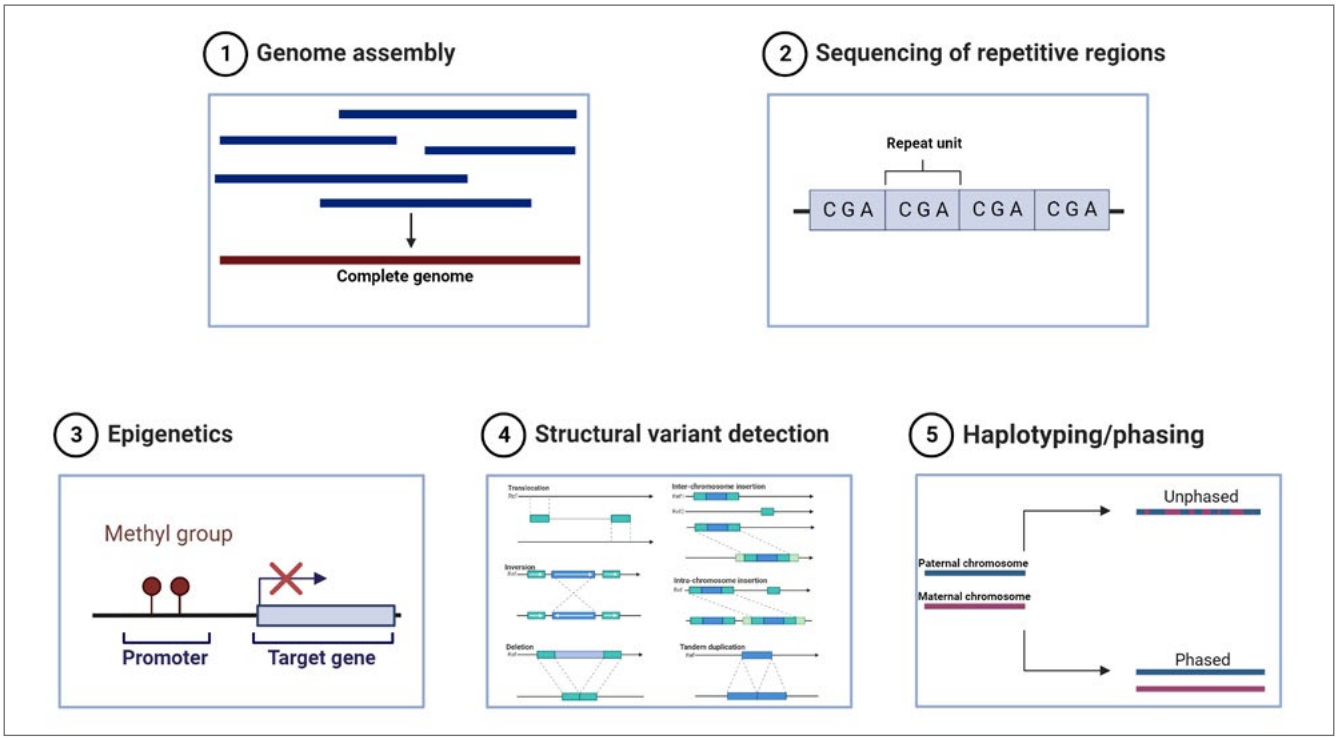

Figure 2. Advantages of long-read sequencing. (1) Comprehensive genome assembly: Longer reads overlap each other, thereby eliminating sequence gaps and covering the entire genome. Larger are also easier to piece together, which makes assembly easier. (2) Sequencing of repetitive sequences: Long-read sequencing technologies allow the sequencing of highly repetitive regions, which is an issue with SRS. (3) Epigenetics: Long-read sequencing does not require amplification of extracted DNA, which allows the detection of epigenomic modifications such as methylation. (4) Structural variant detection: Large complex structural variants have been identified using LRS. The benefit of longer reads is that the entire variant can be covered with a single read [21,25]. (5) Haplotyping/phasing: Instead of combining the maternal and paternal copy of a chromosome, tools for LRS can be used to assemble them separately; thereby, you can determine if an allele is maternally/paternally inherited and if a variant is in cis or in trans [57]. Figure created with BioRender.com

Acknowledgment

We are grateful to Johanna Ranta-aho for critically reading the manuscript.

Conflict of interest statement

Authors declare no conflict of interest.

Funding

M.S. is supported by Academy of Finland (grant #339437: "Improving the clinical interpretation of sequence variants"), and by Sigrid Jusélius Foundation.

Author contributions

RO: data collection, writing and finalizing the manuscript; MS: critical evaluation, proofreading, and supervision.

References

1 Metzker ML. Sequencing technologies the next generation. Nat Rev Genet 2010;11:31-46. https://doi.org/10.1038/nrg2626

2 Herman DS, Lam L, Taylor MR, et al. Truncations of titin causing dilated cardiomyopathy. N Engl J Med 2012;366:619-628. https://doi.org/10.1056/NEJMoa1110186

3 Bonnal RJ, Severgnini M, Castaldi A, et al. Reliable resequencing of the human dystrophin locus by universal long polymerase chain reaction and massive pyrosequencing. Anal Biochem 2010;406:176-184. https://doi.org/10.1016/j.ab.2010.07.022

4 Zatz M, Passos-Bueno MR, Vainzof M. Neuromuscular disorders: genes, genetic counseling and therapeutic trials. Genet Mol Biol 2016;39:339-438. https://doi.org/10.1590/1678-4685-GMB-2016-0019

5 Thompson R, Spendiff S, Roos A, et al. Advances in the diagnosis of inherited neuromuscular diseases and implications for therapy development. Lancet Neurol 2020;19:522-532. https://doi.org/10.1016/S1474-4422(20)30028-4

6 Nigro V, Savarese M. Next-generation sequencing approaches for the diagnosis of skeletal muscle disorders. Curr Opin Neurol 2016;29:621-627. https://doi.org/10.1097/WCO.0000000000000371

7 Biancalana V, Laporte J. Diagnostic use of Massively Parallel Sequencing in Neuromuscular Diseases: Towards an Integrated Diagnosis. J Neuromuscul Dis 2015;2:193-203. https://doi.org/10.3233/JND-150092

8 Koczwara KE, Lake NJ, DeSimone AM, Lek M. Neuromuscular disorders: finding the missing genetic diagnoses. Trends Genet 2022;38:956-971. https://doi.org/10.1016/j.tig.2022.07.001

9 Pereira R, Oliveira J, Sousa M. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. J Clin Med 2020;9:132. https://doi.org/10.3390/jcm9010132

10 Umlai UI, Bangarusamy DK, Estivill X, et al. Genome sequencing data analysis for rare disease gene discovery. Brief Bioinform 2022;23:bbab363. https://doi.org/10.1093/bib/bbab363

11 Savarese M, Di Fruscio G, Mutarelli 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 2014;2:100. https://doi.org/10.1186/s40478-014-0100-3

12 Savarese M, Di Fruscio G, Torella A, et al. The genetic basis of undiagnosed muscular dystrophies and myopathies: Results from 504 patients. Neurology 2016;87:71-76. https://doi.org/10.1212/WNL.0000000000002800

13 Evila A, Palmio J, Vihola A, et al. Targeted Next-Generation Sequencing Reveals Novel TTN Mutations Causing Recessive Distal Titinopathy. Mol Neurobiol 2017;54:7212-7223. https://doi.org/10.1007/s12035-016-0242-3

14 Nallamilli BRR, Chakravorty S, Kesari A, et al. Genetic landscape and novel disease mechanisms from a large LGMD cohort of 4656 patients. Ann Clin Transl Neurol 2018;5:1574-1587. https://doi.org/10.1002/acn3.649

15 Topf A, Johnson K, Bates A, et al. Sequential targeted exome sequencing of 1001 patients affected by unexplained limb-girdle weakness. Genet Med 2020;22:1478-1488. https://doi.org/10.1038/s41436-020-0840-3

16 Ankala A, da Silva C, Gualandi F, et al. A comprehensive genomic approach for neuromuscular diseases gives a high diagnostic yield. Ann Neurol 2015;77:206-214. https://doi.org/10.1002/ana.24303

17 Ghaoui R, Cooper ST, Lek M, et al. Use of Whole-Exome Sequencing for Diagnosis of Limb-Girdle Muscular Dystrophy: Outcomes and Lessons Learned. JAMA Neurol 2015;72:1424-1432. https://doi.org/10.1001/jamaneurol.2015.2274

18 Graessner H, Zurek B, Hoischen A, et al. Solving the unsolved rare diseases in Europe. Eur J Hum Genet 2021;29:1319-1320. https://doi.org/10.1038/s41431-021-00924-8

19 Schobers G, Schieving JH, Yntema HG, et al. Reanalysis of exome negative patients with rare disease: a pragmatic workflow for diagnostic applications. Genome Med 2022;14:66. https://doi.org/10.1186/s13073-022-01069-z

20 Wortmann SB, Oud MM, Alders M, et al. How to proceed after "negative" exome: A review on genetic diagnostics, limitations, challenges, and emerging new multiomics techniques. J Inherit Metab Dis 2022;45:663-681. https://doi.org/10.1002/jimd.12507

21 Logsdon GA, Vollger MR, Eichler EE. Long-read human genome sequencing and its applications. Nat Rev Genet 2020;21:597-614. https://doi.org/10.1038/s41576-020-0236-x

22 Matalonga L, Hernandez-Ferrer C, Piscia D, et al. Solving patients with rare diseases through programmatic reanalysis of genome-phenome data. Eur J Hum Genet 2021;29:1337-1347. https://doi.org/10.1038/s41431-021-00852-7

23 PacBio. Sequencing 101: long-read sequencing. October 17, 2023, https://www.pacb.com/blog/long-read-sequencing/

24 Mitsuhashi S, Matsumoto N. Long-read sequencing for rare human genetic diseases. J Hum Genet 2020;65:11-19. https://doi.org/10.1038/s10038-019-0671-8

25 Sanford Kobayashi E, Batalov S, Wenger AM. et al. Approaches to long-read sequencing in a clinical setting to improve diagnostic rate. Sci Rep 2022;12:16945. https://doi.org/10.1038/s41598-022-20113-x

26 Warburton PE, Sebra RP. Long-Read DNA Sequencing: Recent Advances and Remaining Challenges. Annu Rev Genomics Hum Genet 2023;24:109-132. https://doi.org/10.1146/annurev-genom-101722-103045

27 Mantere T, Kersten S, Hoischen A. Long-Read Sequencing Emerging in Medical Genetics. Front Genet 2019;10:426. https://doi.org/10.3389/fgene.2019.00426

28 Mastrorosa FK, Miller DE, Eichler EE. Applications of long-read sequencing to Mendelian genetics. Genome Med 2023;15:42. https://doi.org/10.1186/s13073-023-01194-3

29 Su Y, Fan L, Shi C, et al. Deciphering Neurodegenerative Diseases Using Long-Read Sequencing. Neurology 2021;97:423-433. https://doi.org/10.1212/WNL.0000000000012466

30 Noyes MD, Harvey WT, Porubsky D, et al. Familial long-read sequencing increases yield of de novo mutations. Am J Hum Genet 2022;109:631-646. https://doi.org/10.1016/j.ajhg.2022.02.014

31 Marx V. Method of the year: long-read sequencing. Nat Methods 2023;20:6-11. https://doi.org/10.1038/s41592-022-01730-w

32 Oehler JB, Wright H, Stark Z, et al. The application of long-read sequencing in clinical settings. Hum Genomics 2023;17:73. https://doi.org/10.1186/s40246-023-00522-3

33 Bruels CC, Littel HR, Daugherty AL, et al. Diagnostic capabilities of nanopore long-read sequencing in muscular dystrophy. Ann Clin Transl Neurol 2022;9:1302-1309. https://doi.org/10.1002/acn3.51612

34 Geng C, Zhang C, Li P, et al. Identification and characterization of two DMD pedigrees with large inversion mutations based on a long-read sequencing pipeline. Eur J Hum Genet 2023;31:504-511. https://doi.org/10.1038/s41431-022-01190-y

35 Xie Z, Sun C, Zhang S, et al. Long-read whole-genome sequencing for the genetic diagnosis of dystrophinopathies. Ann Clin Transl Neurol 2020;7:2041-2046. https://doi.org/10.1002/acn3.51201

36 Bai J, Qu Y, OuYang S, et al. Novel Alu-mediated deletions of the SMN1 gene were identified by ultra-long read sequencing technology in patients with spinal muscular atrophy. Neuromuscul Disord 2023;33:382-390. https://doi.org/10.1016/j.nmd.2023.03.001

37 Chin HL, O'Neill K, Louie K, et al. An approach to rapid characterization of DMD copy number variants for prenatal risk assessment. Am J Med Genet A 2021;185:2541-2545. https://doi.org/10.1002/ajmg.a.62349

38 Ishiura H, Shibata S, Yoshimura J, et al. Noncoding CGG repeat expansions in neuronal intranuclear inclusion disease, oculopharyngodistal myopathy and an overlapping disease. Nat Genet 2019;51:1222-1232. https://doi.org/10.1038/s41588-019-0458-z

39 Yu J, Deng, J, Guo, X, et al. The GGC repeat expansion in NOTCH2NLC is associated with oculopharyngodistal myopathy type 3. Brain. 2020;144:1819-1832. https://doi.org/https://doi.org/10.1093/brain/awab077

40 Yu J, Shan, J, Yu, M, et al. The CGG repeat expansion in RILPL1 is associated with oculopharyngodistal myopathy type 4. Report. Am J Hum Genet 2022;109:533-541. https://doi.org/10.1016/j.ajhg.2022.01.012

41 Cortese A, Beecroft, S. J., Facchini, S., et al. A CCG expansion in ABCD3 causes oculopharyngodistal myopathy in individuals of European ancestry. medRxiv. October 10 2023

42 Yeetong P, Kulsirichawaroj P, Kumutpongpanich T, et al. Long-read Nanopore sequencing identified D4Z4 contractions in patients with facioscapulohumeral muscular dystrophy. Neuromuscul Disord 2023;33:551-556. https://doi.org/10.1016/j.nmd.2023.05.004

43 Bird TD. Myotonic Dystrophy Type 1. In: Adam MP, Feldman J, Mirzaa GM, et al., eds. GeneReviews((R)) 1993.

44 Mangin A, de Pontual L, Tsai YC, et al. Robust Detection of Somatic Mosaicism and Repeat Interruptions by Long-Read Targeted Sequencing in Myotonic Dystrophy Type 1. 2021;22:2616. https://doi.org/10.3390/ijms22052616

45 Rasmussen A, Hildonen M, Vissing J, et al. High Resolution Analysis of DMPK Hypermethylation and Repeat Interruptions in Myotonic Dystrophy Type 1. Genes (Basel) 2022;13:970. https://doi.org/10.3390/genes13060970

46 Ruggieri A, Naumenko S, Smith MA, et al. Multiomic elucidation of a coding 99-mer repeat-expansion skeletal muscle disease. Acta Neuropathol 2020;140:231-235. https://doi.org/10.1007/s00401-020-02164-4

47 Perrin A, Van Goethem C, Theze C, et al. Long-Reads Sequencing Strategy to Localize Variants in TTN Repeated Domains. J Mol Diagn 2022;24:719-726. https://doi.org/10.1016/j.jmoldx.2022.04.006

48 Chen X, Harting J, Farrow E, et al. Comprehensive SMN1 and SMN2 profiling for spinal muscular atrophy analysis using long-read PacBio HiFi sequencing. Am J Hum Genet 2023;110:240-250. https://doi.org/10.1016/j.ajhg.2023.01.001

49 Kucuk E, van der Sanden B, O'Gorman L, et al. Comprehensive de novo mutation discovery with HiFi long-read sequencing. Genome Med 2023;15:34. https://doi.org/10.1186/s13073-023-01183-6

50 Li S, Han X, Zhang L, et al. An Effective and Universal Long-Read Sequencing-Based Approach for SMN1 2 + 0 Carrier Screening through Family Trio Analysis. Clin Chem 2023;69:1295-1306. https://doi.org/10.1093/clinchem/hvad152

51 Dai M, Xu Y, Sun Y, et al. Revealing diverse alternative splicing variants of the highly homologous SMN1 and SMN2 genes by targeted long-read sequencing. Mol Genet Genomics 2022;297:1039-1048. https://doi.org/10.1007/s00438-022-01874-6

52 Uapinyoying P, Goecks J, Knoblach SM, et al. A long-read RNA-seq approach to identify novel transcripts of very large genes. Genome Res 2020;30:885-897. https://doi.org/10.1101/gr.259903.119

53 Savarese M, Qureshi T, Torella A, et al. Identification and Characterization of Splicing Defects by Single-Molecule Real-Time Sequencing Technology (PacBio). J Neuromuscul Dis 2020;7:477-481. https://doi.org/10.3233/JND-200523

54 Frascarelli C, Zanetti, N., Nasca, A., et al. Nanopore long-read next-generation sequencing for detection of mitochondrial DNA large-scale deletions. Front Genet 2023;14(1089956). https://doi.org/https://doi.org/10.3389/fgene.2023.1089956

55 Lei M, Liang D, Yang Y, et al. Long-read DNA sequencing fully characterized chromothripsis in a patient with Langer-Giedion syndrome and Cornelia de Lange syndrome-4. J Hum Genet 2020;65:667-674. https://doi.org/10.1038/s10038-020-0754-6

56 Mitsuhashi S, Ohori S, Katoh K, Frith MC, Matsumoto N. A pipeline for complete characterization of complex germline rearrangements from long DNA reads. Genome Med 2020;12:67. https://doi.org/10.1186/s13073-020-00762-1

57 PacBio. Sequencing 101: ploidy, haplotypes, and phasing — how to get more from your sequencing data. 2023. https://www.pacb.com/blog/ploidy-haplotypes-and-phasing/