



OPEN ACCESS

EDITED BY

Lin Ning,
University of Electronic Science and
Technology of China, China

REVIEWED BY

Atsushi Hijikata,
Tokyo University of Pharmacy and Life
Sciences, Japan
Hywel John Williams,
Cardiff University, United Kingdom

*CORRESPONDENCE

Atiyeh M. Abdallah,
✉ aabdallah@qu.edu.qa

RECEIVED 11 February 2026

REVISED 22 April 2026

ACCEPTED 24 April 2026

PUBLISHED 05 June 2026

CITATION

Ibrahim AA, Fakhro KA and Abdallah AM
(2026) Diagnostic yield of long-read
sequencing for rare diseases: a
systematic review.
Front. Genet. 17:1809097.
doi: 10.3389/fgene.2026.1809097

COPYRIGHT

© 2026 Ibrahim, Fakhro and Abdallah. This
is an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Diagnostic yield of long-read sequencing for rare diseases: a systematic review

Amal Abdulsalam Ibrahim¹, Khalid A. Fakhro^{2,3,4} and
Atiyeh M. Abdallah^{1*}

¹Department of Biomedical Sciences, College of Health Sciences, QU Health, Qatar University, Doha, Qatar, ²College of Health and Life Sciences, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar, ³Laboratory of Genomic Medicine, Sidra Medicine, Doha, Qatar, ⁴Department of Genetic Medicine, Weill Cornell Medicine, Doha, Qatar

Background: Nearly half of patients with rare genetic disorders remain undiagnosed, which may in part be due to limitations of current short-read sequencing (SRS) approaches in detecting complex genomic alterations. Long-read whole genome sequencing (lrWGS) technologies can address these limitations through enhanced detection of structural variants (SVs), repetitive regions, and epigenetic changes.

Methods: To evaluate the diagnostic yield of lrWGS in patients with rare genetic diseases receiving inconclusive or negative results from standard testing, we searched the PubMed, Science Direct, Scopus, and ProQuest databases to July 2025 for studies applying lrWGS to unresolved rare disease cases and reporting diagnostic outcomes. Risk of bias was assessed using the QUADAS-2 tool.

Results: Nine studies involving 646 previously unresolved cases that underwent lrWGS met the inclusion criteria. Of these, 29 individuals (24 unique diagnoses involving 25 genes) received a definitive diagnosis through lrWGS, a diagnostic yield of 4.5%. SVs accounted for the majority of identified variants (41.67%), followed by combined SV/single-nucleotide variants (20.83%), methylation changes (16.67%), and other variant types (copy number variations, indels, and tandem repeats). Most detected variants were in regions typically inaccessible to short-read whole-exome sequencing (WES). lrWGS also enabled phasing and methylation analysis in a single assay, which was valuable for compound-heterozygosity detection and diagnostic interpretation.

Conclusion: lrWGS shows clear potential for improving diagnostic rates in previously unresolved rare disease cases, particularly when applied after WES and combined with advanced tools such as phasing and methylation profiling. As technologies evolve and become more accessible, lrWGS may increasingly become a first-tier diagnostic approach, especially in phenotypically complex conditions.

Systematic Review Registration: <https://osf.io/y5azb/overview>, identifier 10.17605/OSF.IO/Y5AZB.

KEYWORDS

epigenesis, genetic, genetic diseases, genetic variation, high-throughput nucleotide sequencing, inborn, whole genome sequencing

1 Introduction

Rare genetic disorders, while individually uncommon, represent a significant global health burden in aggregate, affecting an estimated 400 million individuals worldwide, the majority (50%–75%) of whom are children (Rahit and Tarailo-Graovac, 2020). An estimated 7,000 rare disorders have been identified, with around 80% believed to have a genetic origin

(Wright et al., 2018). Most of these rare genetic disorders follow Mendelian patterns of inheritance, i.e., caused by mutations in a single gene (monogenic) inherited through autosomal or sex chromosomes in a dominant or recessive manner (Chakravarti, 2021; Rahit and Tarailo-Graovac, 2020). Additionally, some rare genetic disorders arise from *de novo* (newly arising) mutations, which occur spontaneously and are not inherited from parents (Mani, 2017).

Given the predominant genetic contribution to rare diseases, next-generation sequencing (NGS) has become an essential tool for establishing a molecular diagnosis. NGS approaches include targeted or candidate gene sequencing (CGS), whole-exome sequencing (WES), whole-genome sequencing (WGS), transcriptome sequencing (RNA-seq), and methylation sequencing (MeS). Among these, WES is the most widely used and cost-effective method in clinical practice, as it focuses on protein-coding regions, which represent only 1%–2% of the genome but include approximately 85% of known pathogenic variants (Caspar et al., 2018; Sullivan et al., 2023). In cases where a patient's phenotype strongly suggests a specific genetic disorder, targeted sequencing or gene panels can be used as a first-line diagnostic approach (Mitsuhashi and Matsumoto, 2020). However, WES has some limitations, particularly in detecting variants located outside coding regions or those involving structural variants (SVs). As a result, WGS is used in WES-negative cases to capture the whole genome and thus analyze regulatory elements, intergenic regions, intronic regions, and complex SVs that may be missed by WES (Yadav et al., 2023). The choice between WES and WGS depends on specific clinical and research objectives, yet, despite the adoption of NGS, diagnostic yield remains around 25%–50%, leaving under half of individuals with suspected Mendelian disorders undiagnosed (Mastrolorsa et al., 2023; Yu et al., 2023).

One possible contributor to this diagnostic gap is the reliance of NGS on short-read sequencing (SRS), which generates reads of 50–300 base pairs, making accurate mapping and assembly difficult. This limitation leads to challenges in resolving highly repetitive sequences, GC-rich regions (e.g., 5'- and 3'-untranslated regions [UTRs]), and distinguishing functional genes from their pseudogenes. Furthermore, SRS lacks sensitivity in detecting complex SVs, such as inversions, translocations or copy number changes in repetitive regions (Dierckxsens et al., 2021; Sathyanarayana et al., 2022; Yu et al., 2023). Given that SVs play a significant role in many genetic diseases, overcoming these limitations is important for improving rare disease diagnostics (Weischenfeldt et al., 2013).

The limitations of SRS highlight the need for alternative sequencing technologies that address these limitations, such as long-read sequencing (LRS), also known as third-generation sequencing. Long-read sequencing generates reads with an average read length of 10–15 kb, enabling more accurate genome assembly and variant detection (Olivucci et al., 2024). Currently, two LRS technologies are widely available: Pacific Biosciences (PacBio) single-molecule real-time (SMRT) sequencing and Oxford Nanopore Technologies (ONT) sequencing. These platforms differ in their sequencing chemistry, data production, and analysis algorithms. PacBio uses fluorescence-based detection, where DNA polymerase

incorporates labeled nucleotides, resulting in high-accuracy reads (99.9%), making it suitable for applications requiring high precision, such as rare variant detection, resolving complex genomic regions, and phasing. In contrast, ONT relies on DNA molecules passing through biological nanopores, with electrical current changes used to identify nucleotides (Zhong et al., 2021). This allows ONT to provide real-time sequencing and analysis, making it particularly useful for rapid pathogen detection, outbreak surveillance (e.g., COVID-19), and clinical microbiology where immediate results are necessary. Additionally, ONT's portable devices, such as the MinION, enable sequencing in field or point-of-care settings, expanding its use beyond traditional laboratory environments (Zheng et al., 2023). This review does not aim to compare the two existing LRS technologies. Therefore, no further distinctions are made between them, and the two are collectively referred to as LRS throughout.

Both technologies share several key advantages over SRS, making them valuable for different clinical and research applications. Among these, a primary advantage is the ability to detect large and complex SVs that are often missed by SRS. These include large deletions, such as those affecting *PRKARIA* in Carney complex (Merker et al., 2018) and *G6PC* in glycogen storage disease IA (Miao et al., 2018), as well as chromosomal translocations linked to developmental disorders (Dutta et al., 2019). Additionally, LRS excels in sequencing long tandem repeat expansions and GC-rich regions, which are implicated in diseases like Fragile X syndrome, myotonic dystrophy type 1, and Huntington's disease (Ardui et al., 2017; Cummings et al., 2017; Hoijer et al., 2018; Loomis et al., 2013). Another critical advantage is its ability to differentiate clinically relevant genes from their pseudogenes, as seen in autosomal dominant polycystic kidney disease (ADPKD), enabling the differentiation of *PKD1* from highly similar pseudogenes (Borras et al., 2017). Also, LRS enables haplotype phasing, allowing the assignment of variants to maternal or paternal chromosomes, which is particularly useful when parental testing is unavailable, as demonstrated in a case where LRS confirmed the presence of two *USH2A* variants in *trans*, confirming compound heterozygosity (Gupta et al., 2023). Beyond variant detection, LRS technologies provide additional benefits. In particular, unlike PCR-based methods, LRS enables native DNA sequencing, allowing the simultaneous detection of sequence variation and methylation modifications within a single assay, without additional processing steps (Olivucci et al., 2024). These features suggest that LRS may contribute to improved diagnostic rates in Mendelian disorders, particularly in cases involving complex variants that are difficult to resolve using short-read sequencing approaches.

This systematic review aims to provide an overview of the diagnostic yield of LRS in rare disease cohorts previously undiagnosed by SRS technologies. Many other studies have reported the clinical applicability of LRS for resolving individual rare disease cases and have highlighted its advantages; others have also compared ONT and PacBio technologies (Mantere et al., 2019; Marwaha et al., 2022; Mitsuhashi and Matsumoto, 2020; Oehler et al., 2023; van Dijk et al., 2023; Yu et al., 2023). However, to our knowledge, this is the first systematic review that comprehensively assesses the diagnostic utility of LRS across rare disease cohorts, offering a broader perspective on its clinical utility.

TABLE 1 Inclusion and exclusion criteria according to the PICOS statement.

Criteria	Included	Excluded
Population	Patients with undiagnosed rare genetic disorders who have negative, unresolved, or inconclusive findings from prior genetic testing	Patients with non-rare genetic disorders (such as cancer, infectious diseases, common disorders) or a diagnosed rare disorder with a known causative gene
Interventions	Studies utilizing whole genome long-read sequencing technologies (PacBio or Oxford Nanopore)	Studies utilizing targeted approaches (incl. long-reads) or other sequencing technologies
Comparators	Studies comparing whole genome long-read sequencing to standard diagnostic technologies	NA
Study design	Prospective and retrospective cohort studies, research articles	Reviews, case reports, books, protocols, guidelines, poster presentations, and animal studies
Primary outcome	Diagnostic yield of long-read sequencing in detecting a genetic cause (incl. single nucleotide variants (SNVs), structural variants (SVs), copy number variants (CNVs), tandem repeats, and methylation patterns)	NA
Secondary outcome	Variant types and genomic regions identified by LRS	NA

PICOS: population, Intervention, Comparator, Outcomes, and Study design. SNVs: single-nucleotide variants. SVs: structural variants. CNVs: copy number variants.

2 Methods

2.1 Search strategy

This systematic review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA) guidelines (Moher et al., 2015). A comprehensive literature search was performed using the PubMed, Science Direct, Scopus, and ProQuest databases, covering studies from their inception until July 2025. The primary outcome of this review was to assess the diagnostic application of whole-genome long-read sequencing (lrWGS) technologies for rare genetic disorders, particularly in cases with previously negative, unresolved, or inconclusive genetic testing. This included evaluating the ability of lrWGS to detect single nucleotide variants (SNVs), SVs, copy number variants (CNVs), tandem repeats, and methylation patterns. The secondary outcome was to assess the types of variants (e.g., SNVs, SVs, tandem repeats) and the genomic regions involved (e.g., exonic, intronic, regulatory) among findings identified exclusively by LRS. The inclusion and exclusion criteria are summarized in Table 1. To ensure broad inclusion, the following search strategy was used: (“long read sequencing” OR “PacBio” OR “Oxford Nanopore”) AND (“Diagnosis”). Following initial screening of titles and abstracts, eligible studies were reviewed in full to confirm their inclusion in the analysis.

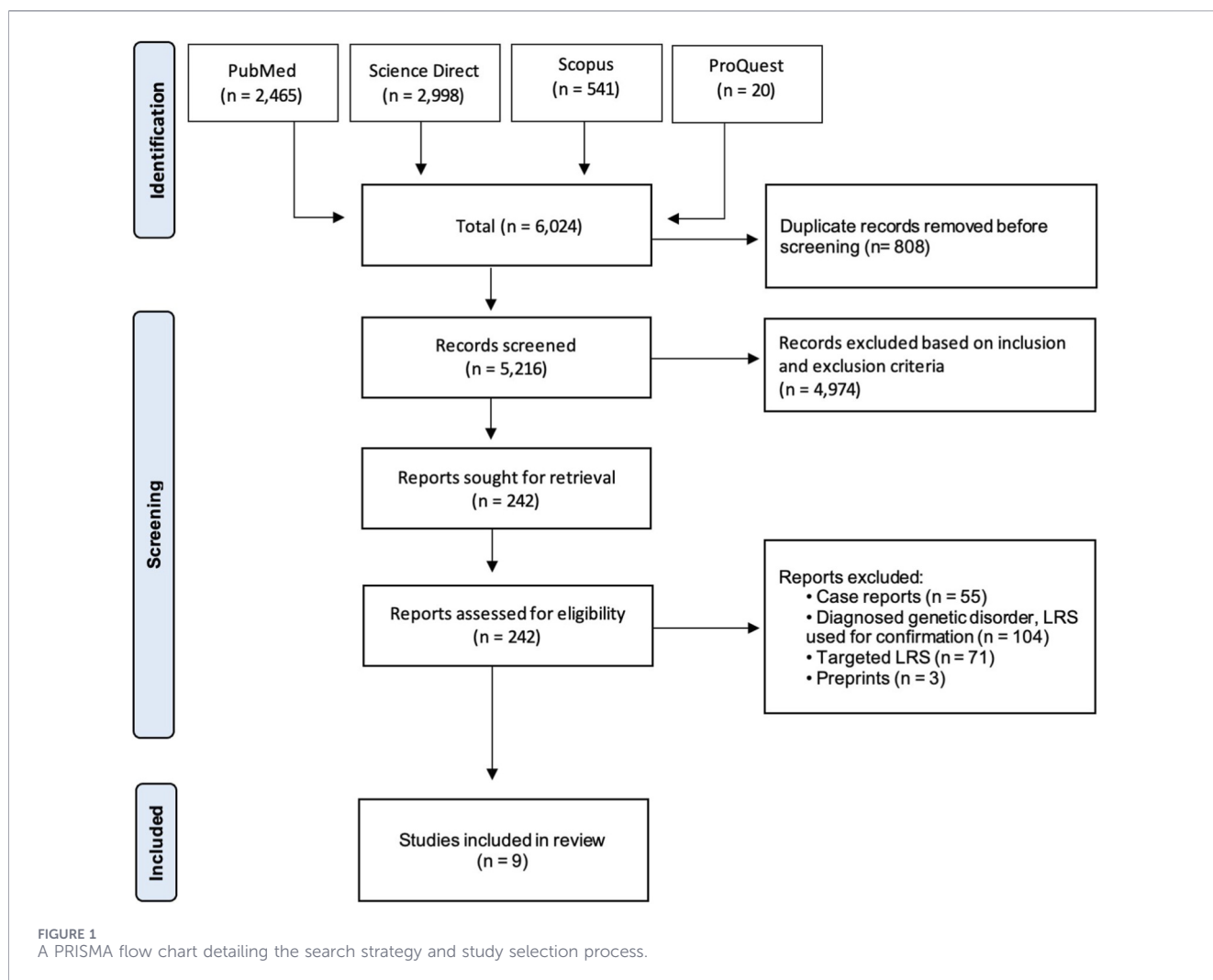
2.2 Study selection

Retrieved records were initially assessed by two authors and subsequently reviewed by the third author. Any disagreements were resolved through discussion and consensus. The total number of records retrieved from each database was documented. Studies were selected for full-text review if they met the following inclusion criteria: 1) peer-reviewed original research articles; 2) patients with undiagnosed rare genetic disorders who had previously undergone genetic testing with negative, unresolved, or

inconclusive findings; 3) use of lrWGS technologies (Oxford Nanopore or PacBio); and 4) reporting of diagnostic outcomes or variant detection using long-read approaches. Studies were excluded if they met the following criteria: 1) did not utilize lrWGS; 2) focused on non-rare genetic disorders (such as cancer, infectious diseases, or other common genetic conditions), or involved a diagnosed genetic disorder with a clear phenotypic presentation and known causative gene; and 3) were review articles, case reports, conference abstracts, guidelines, animal studies, or other non-peer-reviewed material. In addition to removing duplicates, titles and abstracts were screened to eliminate irrelevant records. The remaining articles were reviewed in full to determine final eligibility. The inclusion and exclusion criteria are detailed in Table 1, and the overall screening and selection process is presented in the PRISMA flow diagram (Figure 1).

2.3 Quality control assessment and data extraction

The quality of eligible studies was assessed by two authors using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool and reviewed by the third author (Qu et al., 2018). This tool evaluates four domains: patient selection, index test, reference standard, and flow and timing. Each domain was rated as having “low,” “high,” or “unclear” risk of bias, and any disagreements were resolved in consultation with the senior reviewers. In this review, the index test domain was frequently rated as “high” or “unclear” risk due to the lack of blinding to previous genetic test results. This limitation is expected in studies using lrWGS as a second-tier test in clinical practice. Data were extracted from the tables and full texts of all included articles and collated in a Microsoft Excel spreadsheet. The primary author performed data extraction, which was then independently reviewed by the senior authors for accuracy. Extracted variables included study design, sample size, sequencing platform (Oxford Nanopore or PacBio), diagnostic yield, and the type and complexity



of detected variants (e.g., SNVs, SVs, CNVs, tandem repeats, and methylation). Where applicable, comparisons with short-read technologies were noted, as well as clinical information such as zygosity and phenotypic findings. The identified variants were confirmed using public databases, including OMIM and the UCSC Genome Browser, to check the associated phenotype, chromosomal location, and predicted variant effect.

3 Results

3.1 Search outcome

The initial search retrieved 6,024 records across four databases. After removing 808 duplicates, 5,216 records were screened by title and abstract using predefined inclusion and exclusion criteria. Of these, 4,974 records were excluded for not meeting eligibility criteria. A total of 242 full-text articles were assessed for eligibility. As a result, nine studies were included in the final analysis. The study selection process is illustrated in the PRISMA flow diagram (Figure 1), and an overview of the included studies is provided in Table 2.

3.2 Quality of the eligible studies

The risk of bias across the nine studies was evaluated using the QUADAS-2 tool (Figure 2). Overall, most studies had a low risk of bias in the domains of patient selection, reference standard, and flow and timing. However, the index test domain was rated as high risk in all studies, primarily due to the lack of blinding to prior genetic test results. This reflects routine clinical practice where lrWGS is applied as a second-tier test rather than a standalone diagnostic tool.

3.3 Diagnostic yield of lrWGS

Among the nine included studies, a total of 646 previously unresolved or inconclusive cases underwent lrWGS. In the study by (AlAbdi et al., 2023), cases were reported by family rather than individuals; therefore, each family was counted as a single case, as only one affected individual per family was tested. Of the total cases, 29 were diagnosed exclusively through lrWGS, yielding an overall diagnostic rate of 4.5% (29/646). Diagnostic yield varied across studies depending on cohort size, selection criteria, and whether lrWGS was applied as a first-tier test or as a targeted follow-up to previous short-read testing. Most studies performed lrWGS after

TABLE 2 Characteristics of included studies.

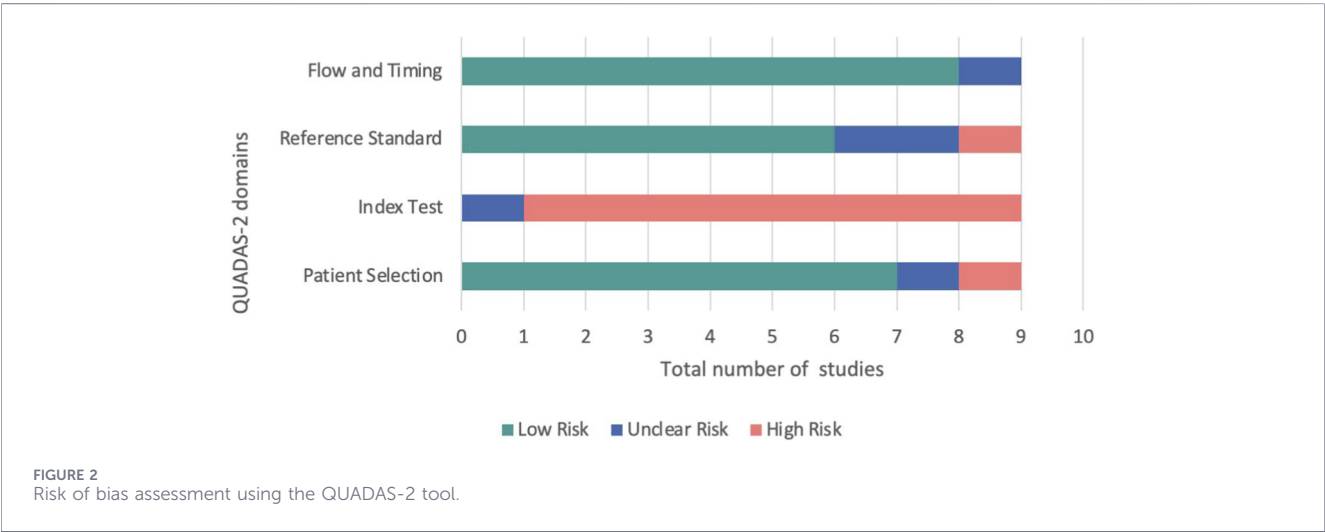
Study	Country	Hospital/ Data collection program	Total cases	Negative/ Inconclusive cases	Cases genotyped with lrWGS	Positive control cases	Clinical focus/ Disease category	Age group	Sex distribution (M:F)
Cohen et al. (2022)	United States	Children's Mercy Research Institute in Kansas City-GA4K	1,083	584	472	125	Diverse rare diseases	P	488:595
Colin et al. (2023)	France, Belgium	AnDDI-Rares network	15	15	2	NA	Mendelian disorders	P and A	3:12
De Clercq et al. (2024)	Belgium	Ghent University Hospital	6	6	6	NA	Rare diseases with unresolved SVs	P and A	2:4
LaFlamme et al. (2024)	United States of America, Canada, Australia, New Zealand	St. Jude Children's Research Hospital (United States of America), London Health Science Centre (Canada), Seattle Children's Hospital (United States of America), Austin Health (Australia), University of Otago (New Zealand)	1,194	582	3	NA	DEEs	P	556:638
Sanford Kobayashi et al. (2022)	United States of America	RCIGM	35	30	35	5	Severe, syndromic pediatric disease phenotypes	P	NA
AlAbdi et al. (2023)	Saudi Arabia	KFSHRC	34 families	34 families	34 families	NA	Autosomal recessive diseases	NA	NA
Negi et al. (2025)	United States of America	Broad Institute and CNH	98 (41 families)	42	98	5	Rare monogenic diseases	P and A	NA
Hiatt et al. (2021)	United States of America	HudsonAlpha	6 (trio)	6	6	NA	NDDs	NA	2:4

(Continued)

TABLE 2 Continued

Study	Country	Hospital/ Data collection program	Total cases	Negative/ Inconclusive cases	Cases genotyped with lrWGS	Positive control cases	Clinical focus/ Disease category	Age group	Sex distribution (M:F)
Sinha et al. (2025)	UAE	Dubai Health Genomic Medicine Center	68	51	51	17	Rare diseases (45% neurological cases)	P and A	24:22 22 NA

lrWGS, long-read whole genome sequencing; F, female; M, male; P, pediatrics; A, adults; SVs, structural variants; CNH, Children’s National Hospital; RCIGM, Rady Children’s Institute for Genomic Medicine; GA4K, genomic answers for kids; AnDDI, anomalies of developmental delay and intellectual disability; KFSHRC, king faisal specialist hospital and research center; HudsonAlpha, HudsonAlpha Institute for Biotechnology, Huntsville; DEEs: developmental and epileptic encephalopathies; NDDs, neurodevelopmental disorders; NA, not available or not reported.



prior short-read WES/WGS or targeted gene panel sequencing. The platforms used included PacBio Sequel II/IIe and Oxford Nanopore PromethION. Detailed diagnostic yield and variant characteristics of diagnosed cases are summarized in [Tables 3, 4](#).

3.4 Variant types and genomic features identified

A total of 24 unique diagnoses were reported across 29 diagnosed cases involving 25 genes. The distribution of variant types detected by lrWGS is shown in [Figure 3A](#). SVs were the most frequently identified, accounting for 41.67% of diagnostic findings (n = 10), followed by combined SV/SNV events in heterozygous inheritance (20.83%, n = 5), methylation profiling (16.67%, n = 4), SNVs (12.50%, n = 3), and other categories such as indels and tandem repeats (TRs), each contributing 4.17% (n = 1). Many of the detected variants missed using SRS were within noncoding regions (e.g., introns, promoters), or complex regions (e.g., insertions, translocations, methylation, and pseudogenes).

Chromosome-level mapping of the implicated genes is shown in [Figure 3B](#), with affected genes color-coded based on the genomic region impacted. Where a specific diagnostic gene was reported, it is shown on the map; in cases involving complex rearrangements with no single implicated gene, the label “Multi-Genes” was used. A full

list of coordinates and variant annotations is provided in the [Supplementary Table](#).

4 Discussion

Despite significant advances in NGS, up to half of individuals with rare genetic disorders remain undiagnosed. One potential causative factor is the limitation of SRS-based methods in detecting complex genomic alterations such as SVs, repetitive regions, and epigenetic changes. This systematic review aimed to evaluate whether lrWGS provides added diagnostic value in patients with previously unresolved rare diseases. To our knowledge, this is the first comprehensive review focusing on the application of lrWGS in clinical cohorts that had negative or inconclusive results from prior genetic testing.

From an initial screen of 6,024 articles, a total of nine studies were finally included, comprising 646 individuals who had previously received inconclusive or negative results from standard genetic testing. Of these, 29 were newly diagnosed through lrWGS (24 unique diagnoses involving 25 genes), resulting in an overall diagnostic yield of 4.5%. While this diagnostic yield of 4.5% is clinically meaningful, particularly given that all included cases had undergone extensive prior testing, it remains modest. Several factors may contribute to this. First, lrWGS is still an emerging technology,

TABLE 3 Primary outcome (diagnostic yield) of lrWGS in unresolved rare disorders cohorts from included studies.

Study	Prior genetic testing	LRS platform	Average coverage	Variant analysis targets	References genome	LRS-diagnosed cases/ Tested (n/N)	LRS-diagnostic yield (%)	Comments
Cohen et al. (2022)	srWES, srWGS, or panel testing	PacBio Sequel II	≥ 30	SNVs, SVs, TRs	GRCh38	6/472	1.27	Diagnostic yield recalculated based on methods section and Supplementary Table S1 , including only lrWGS-genotyped cases (n = 472) with prior negative testing and lrWGS-only solved cases. Original abstract yield (64/584) included mixed technologies
Colin et al. (2023)	Array-CGH, Sanger Sequencing, gene panel testing, srWES	ONT-PromethION	NA	SVs	GRCh37/hg19	1/2	-	Diagnostic yield was not calculated as LRS was performed on only two cases
De Clercq et al. (2024)	Karyotyping, FISH, microarray, shallow WGS, srWES, Mate-pair sequencing	ONT-MinION or PromethION	NA	SVs	GRCh38	2/6	-	Diagnostic yield not calculated as LRS was applied to a preselected cohort with unresolved but previously identified SVs
LaFlamme et al. (2024)	Gene panel, WES, or srWGS	ONT-PromethION	NA	Methylation	GRCh38	3/3	-	Diagnostic yield not calculated as LRS was selectively applied for DMR confirmation; only 3 cases underwent lrWGS.
Sanford Kobayashi et al. (2022)	srWGS	Pacbio Sequel II and IIe	25.2–38.8	SNVs, SVs, Methylation	GRCh38	1/30	3.33	Diagnostic yield recalculated as 3.33% (1/30) based on full LRS-tested cohort; original 25% (1/4) yield was subgroup-specific (immunodeficiency)
AlAbdi et al. (2023)	srWES	Pacbio Sequel IIe	NA	SNVs, SVs	GRCh37/hg19	6/34	17.65	Diagnostic yield recalculated as 17.65% (6/34) based on LRS-only diagnoses; original study reported 38% (13/34), including WES reanalysis

(Continued)

TABLE 3 Continued

Study	Prior genetic testing	LRS platform	Average coverage	Variant analysis targets	References genome	LRS-diagnosed cases/ Tested (n/N)	LRS-diagnostic yield (%)	Comments
Negi et al. (2025)	srWGS	ONT-PromethION	36	SNVs, SVs, TRs, Methylation	GRCh38	3/42	7.14	Diagnostic yield recalculated as 7.14% (3/42) based on LRS-only diagnoses; the study reported 11/42, but remaining were identifiable by SRS reanalysis or from the positive cohort
Hiatt et al. (2021)	Array, Single gene test(s) or panel(s), srWES, srWGS	PacBio Sequel II	25–44	SNVs, SVs	GRCh38	2/6	33.33	Two LRS-solved cases reported; cohort selected based on absence of causal variants
Sinha et al. (2025)	srWES, microarray	ONT - PromethION	49	SNVs, SVs, CNVs, Methylation	GRCh37/hg19	5/51	9.80	Diagnostic yield (5/51) aligns with the study's reported results

LRS, long-read sequencing; lrWGS, long-read whole genome sequencing; srWES, short-read whole exome sequencing; srWGS, short-read whole genome sequencing; SNVs, single nucleotide variants; SVs, structural variants; CNVs, copy number variants; TRs, tandem repeats; DMRs, differentially methylated regions; WES, whole exome sequencing; ONT, oxford nanopore technologies; PacBio, Pacific Biosciences; NA, not available or not reported; FISH, fluorescence *in situ* hybridization.

LRS-Diagnosed Cases/Tested (n/N), Number of cases diagnosed exclusively by lrWGS, shown as n (LRS-only diagnosed cases) out of N (total number of previously negative or unresolved cases tested by lrWGS). LRS, Diagnostic Yield (%), Percentage of cases uniquely solved by lrWGS, calculated as $(n/N \times 100\%)$.

and both laboratory workflows and bioinformatics tools are continuously evolving; for example, recent consensus-based frameworks that integrate multiple structural-variant callers have demonstrated improved accuracy and robustness, highlighting the importance of ongoing methodological development ([Pietryga et al., 2025](#)). Second, although lrWGS enables improved detection of variants in non-coding regions, including deep intronic and regulatory elements, the interpretation of such variants remains challenging due to incomplete functional genomic annotation and limited clinical frameworks for assessing pathogenicity ([Austin-Tse et al., 2022](#)). Third, using reference genomes such as GRCh38 can limit the detection of variants in highly repetitive or previously unresolved regions; new resources like the Telomere-to-Telomere (T2T-CHM13) reference genome have been shown to improve read mapping accuracy and variant detection in these regions, thereby boosting diagnostic performance ([Aganezov et al., 2022](#)). However, their clinical use remains limited because annotation tools and databases still mainly depend on GRCh38. Finally, emerging artificial intelligence-based methods for genome interpretation, such as AlphaGenome, are showing promise in enhancing the prioritization and functional prediction of non-coding variants, especially in regulatory and deep intronic regions; however, their clinical use remains in the early stages ([Avsec et al., 2026](#)). Together, these factors suggest that the current diagnostic yield likely underestimates the full potential of lrWGS as both technologies and analytical frameworks continue to develop.

The diagnostic yield varied across the included nine studies, likely due to several factors such as sample size, patient selection criteria, variability

in alignment and variant calling tools, disease focus, sequencing platform, extent of prior genetic testing, and the specific application of lrWGS (whether used as a follow-up to confirm suspected findings or as a first-tier comprehensive standalone analysis). The investigative focus (e.g., SVs, SNVs, methylation profiling) also influenced yield.

The highest diagnostic yield was reported in ([Hiatt et al., 2021](#)), at 33.3% (2 of six individuals) using the PacBio sequencing platform. However, due to the small sample size and the specific focus on neurodevelopmental disorders, the generalizability of this result is limited. Among larger cohorts, ([AlAbdi et al., 2023](#)), reported a diagnostic yield of 17.65% (6 out of 34 families), also using PacBio. All participants had undergone WES prior to lrWGS, and the higher yield can be partially attributed to the fact that applying lrWGS provided insights into intronic and regulatory regions that were previously undetectable using WES. Notably, five of the six solved cases involved non-coding variants in genes such as *TYMS*, *STK25*, *SLC4A4*, *SNAP91*, and *LEMD2* (see [Supplementary Material](#) for details). A comparable yield was reported in ([Sinha et al., 2025](#)) at 9.8% (5 out of 51) using ONT. Like ([AlAbdi et al., 2023](#)), this study applied lrWGS after WES and included methylation analysis, which contributed to the identification of disease-causing variants. The cohort was predominantly of Arab ancestry and included a high proportion of neurological cases, further supporting the value of lrWGS in diverse and clinically complex populations. In ([Negi et al., 2025](#)), ONT was used following prior WGS, yielding a 7.1% diagnostic rate (3 out of 42). This study integrated SNV, SV, and methylation analyses and employed Parakit, a pangenome assembly-based method, enabling resolution of complex regions such as *CYP21A2*.

TABLE 4 Summary of diagnosed cases by lrWGS in included studies.

Study/ Citation	Clinical phenotype	Gene (OMIM ID)	Variant type	Gene region	Zygoty	Final reported diagnosis	Comments/ LRS contribution
Cohen et al. (2022)	Septo-optic dysplasia, hypotonia, strabismus, tremor, DD	AARS2 (MIM: 612035)	SNV, SV (deletion)	SNV: Exon 4 SV: Exons 5–7	AR/ compound het	AARS2-related disease	The SNV was detected previously; however, deletion was difficult to detect using prior genetic testing
Cohen et al. (2022)	Global DD, dystonia	STARD7 (MIM: 616712)	TR (pentamer expansion)	Intron 1	AD/HET	Novel expansion disorder	
Colin et al. (2023)	Alopecia, micropenis, hypotonia, global developmental delay, autism spectrum features, attention deficit, wide- based gait	LSS (MIM: 600909), MCM3AP (MIM: 603294)	SNV, SV (deletion)	SNV: Exonic SV: Promoter, exonic	AR/HOM	Consistent with OMIM; Alopecia-intellectual disability syndrome 4 (MIM: 618840)	The SNV was detected previously; however, SV was missed by SRS.
De Clercq et al. (2024)	Severe intellectual disability and dysmorphic facial features	CELF2 (MIM: 602538)	SV (translocation)	Repeat region (chr9), coding region (chr10)	AD/HET	Consistent with OMIM; Developmental and epileptic encephalopathy (MIM: 619561)	Previously identified by karyotyping and confirmed by FISH. Breakpoints of the t (9; 10) translocation were precisely delineated using LRS, enabling a novel molecular diagnosis
De Clercq et al. (2024)	Mild intellectual disability, Down-like features, and autism spectrum disorder	MYT1L (MIM: 616521)	SV (deletion)	Exon 9	AD/HET	Consistent with OMIM; Intellectual developmental disorder, autosomal dominant (MIM: 616521)	Previously identified through WES. The exon 9 MYT1L deletion was precisely delineated using LRS.
LaFlamme et al. (2024)	Epilepsy of infancy with migrating focal seizures (EIMFS)	Multiple genes	Hypermethylation- SV (translocation)	Multiple genes	AD/HET	NA	DMR was detected; LRS was performed to confirm the underlying cause
LaFlamme et al. (2024)	Developmental and epileptic encephalopathies	CSNK1E (MIM: 600863)	Hypermethylation- CGG repeat	5'UTR and intron 1	AD/HET	NA	DMR was detected; LRS was performed to confirm the underlying cause
Sanford Kobayashi et al. (2022)	Hepatosplenomegaly; laboratory evaluation indicated abnormal immunoglobulins and imaging was remarkable for abnormal osseous mineralization	IKBKG (MIM: 300248)	SNV	Exon 10 (stop- loss variant)	X-linked	Immunodeficiency with or without anhidrotic ectodermal dysplasia	NGS dead zone due to a pseudogene with greater than 99% homology

(Continued)

TABLE 4 Continued

Study/ Citation	Clinical phenotype	Gene (OMIM ID)	Variant type	Gene region	Zygosity	Final reported diagnosis	Comments/ LRS contribution
AlAbdi et al. (2023)	Increased serum lactate, encephalopathy, abnormal breathing, metabolic acidosis, widened subarachnoid space, focal T2 hyperintense brainstem lesion	<i>TYMS</i> (MIM: 188350)	SV (insertion)	Intron 3	AR/HET	Lethal neonatal lactic acidosis	Novel phenotype
AlAbdi et al. (2023)	Hearing impairment, microcephaly, short stature, decreased body weight, Infra-orbital crease, capillary hemangiomas (flat), hypertonia, failure to thrive	<i>STK25</i> (MIM: 602255)	SV (deletion)	Deletion of regulatory region in <i>STK25</i> + deletion of intron 20 of <i>ANO7</i>	AR/HOM	<i>STK25</i> -related neurodevelopmental disorder	Novel phenotype
AlAbdi et al. (2023)	Leber congenital amaurosis	<i>RP1L1</i> (MIM: 608581)	SV	Exonic, in-frame	AR/ Compound het	Consistent with OMIM; Retinitis pigmentosa 88 (OMIM: 618826)	Known gene and phenotype
AlAbdi et al. (2023)	Band keratopathy, glaucoma, corneal calcification, choroidal detachment	<i>SLC4A4</i> (MIM: 603345)	SNV	Promoter region	AR/HOM	<i>SLC4A4</i> -related band keratopathy	Novel phenotype
AlAbdi et al. (2023)	Dwarfism, microcephaly, global developmental delay, spastic tetraparesis, gastroesophageal reflux, failure to thrive, abnormal racial shape, deeply set eyes, prominent nose, hypoplasia of dental enamel, high palate, abnormal sternum morphology, brain atrophy (mild), interhemispheric cyst (small)	<i>SNAP91</i> (MIM: 607923)	SNV	Intronic	AR/HOM	<i>SNAP91</i> -related microcephalic primordial dwarfism	Novel phenotype
AlAbdi et al. (2023)	Microcephaly, plagiocephaly, abnormal number of hair whorls, anteverted nares, microtia, deep palmar crease, slow saccadic eye movements, heart murmur, distal arthrogryposis, adductor longus contractures, Increased muscle tone,	<i>LEMD2</i> (MIM: 616312)	Indel	Intron 4	AR/HOM	<i>LEMD2</i> - related neurodevelopmental disorder	Novel phenotype

(Continued)

TABLE 4 Continued

Study/ Citation	Clinical phenotype	Gene (OMIM ID)	Variant type	Gene region	Zygoty	Final reported diagnosis	Comments/ LRS contribution
	global developmental delay, failure to thrive, diffuse white matter abnormalities, hip dysplasia dilatation of the renal pelvis, mildly echogenic liver						
Negi et al. (2025)	NA	<i>CYP21A2</i> (MIM: 613815)	SNV/SV (deletion)	Intron/Exon	AR/ Compound het	Consistent with OMIM; Adrenal hyperplasia, congenital, due to 21-hydroxylase deficiency (MIM: 201910)	Used Parakit, novel collapsed pangenome method for mapping long reads
Negi et al. (2025)	NA	<i>LHCGR</i> (MIM: 152790)	SNV/SV (deletion)	SNV: Exon 11 SV: Deletion of Exon 9	AR/ Compound het	NA	Biallelic variants in <i>LHCGR</i> were detected using LRS-only. Both variants phased by LRS in absence of parental data
Negi et al. (2025)	Complex neurodevelopmental phenotype including minimal expressive language, autistic features, and dysmorphic features	<i>ARID1B</i> (MIM: 614556)	SNV/Methylation	Intronic, splicing effect	AD/HET	Consistent with OMIM; Coffin-Siris syndrome 1 (MIM: 135900)	Methylation information provided through LRS.
Hiatt et al. (2021)	Intellectual disability, developmental delay, and seizures	<i>CDKL5</i> (MIM: 300203)	SV (insertion)	Exon 3	XLD	Consistent with OMIM; Developmental and epileptic encephalopathy 2 (MIM: 300672)	Mosaicism is suspected according to SNV analysis near the insertion region
Hiatt et al. (2021)	Intellectual disability, facial dysmorphism, hypotonia	Multiple genes <i>DGKB</i> (MIM: 604070) <i>MLLT3</i> (MIM: 159558)	SV (insertion-translocation-inversion)	<i>MLLT3</i> : Exons 3–4, exons 9–10	AR/HET		Two insertional translocations between chromosomes 7 and 9 and an inversion result in disruption of two protein-coding genes. Expression of <i>DGKB</i> was not examined, as the gene is not expressed at appreciable levels in blood

(Continued)

TABLE 4 Continued

Study/ Citation	Clinical phenotype	Gene (OMIM ID)	Variant type	Gene region	Zygoty	Final reported diagnosis	Comments/ LRS contribution
Sinha et al. (2025)	Developmental delay, ID, dysmorphic features, obesity	<i>NSMCE3</i> (MIM: 608243)	SV (deletion)	Multiple regions	AD/HET	Consistent with OMIM; Lung disease, immunodeficiency, and chromosome breakage syndrome (MIM: 617241)	
Sinha et al. (2025)	Eye disorder-foveal hypoplasia and/or anterior segment dysgenesis	<i>SLC38A8</i> (MIM: 615585)	SNV/SV (deletion)	Exons 8–3'UTR	Compound het	Consistent with OMIM; Foveal hypoplasia 2, with or without optic nerve misrouting and/or anterior segment dysgenesis (MIM: 609218)	
Sinha et al. (2025)	Learning disabilities with distinctive brittle hair, a hallmark of trichothiodystrophy non- photosensitive 1 associated with non- functional MPLKIP protein	<i>MPLKIP</i> (MIM: 609188)	SV (deletion)	3'UTR	AR/HOM	Consistent with OMIM: Trichothiodystrophy 4, nonphotosensitive (MIM: 234050)	Further investigation is required to confirm this mechanism and to understand the functional impact of the 3'UTR deletion in this gene
Sinha et al. (2025)	Craniosynostosis, intellectual deficit, short stature, and facial dysmorphism	<i>NSD1</i> (MIM: 606681)	SV (duplication) and Methylation	Multiple regions	AD/HET	Consistent with OMIM; Sotos syndrome (MIM: 117550) Methylation: Hunter McAlpine syndrome (HMA)	
Sinha et al. (2025)	Skeletal/musculoskeletal	<i>SMN1</i> (MIM: 600354)	Methylation	NA	AR/HOM	Consistent with OMIM: Spinal muscular atrophy	Patient has 4 copies of <i>SMN2</i>

AD, autosomal dominant; AR, autosomal recessive; CNV, copy number variant; DD, developmental delay; DMR, differentially methylated region; FISH, fluorescence *in situ* hybridization; HET, heterozygous; HOM, homozygous; ID, intellectual disability; Indel, insertion/deletion; LRS, long-read sequencing; OMIM, online mendelian inheritance in man; SRS, short-read sequencing; SNV, single nucleotide variant; SV, structural variant; TR, tandem repeat; UTR, untranslated region; WES, whole exome sequencing; XLD, X-linked dominant.

Additionally, phasing allowed identification of compound heterozygous variants in *LHCGR* despite the absence of parental data. Sanford Kobayashi et al. (Sanford Kobayashi et al., 2022) applied PacBio to 30 previously undiagnosed cases and reported a 3.3% diagnostic yield (1 case). Although the original study reported a higher percentage (25%), that number was based on a small disease-specific subgroup. When recalculated across the full lrWGS-tested cohort, the more conservative estimate more accurately reflects the yield attributable solely to lrWGS (Sanford Kobayashi et al., 2022).

The lowest diagnostic yield was seen in (Cohen et al., 2022), where PacBio was used on a large cohort of 472 individuals who had undergone a range of prior testing methods (WES, WGS, gene panels). Despite this, lrWGS uniquely identified diagnoses in six individuals (1.3%), which, though modest, demonstrates incremental value in highly pre-screened populations.

Several studies applied lrWGS selectively and did not permit diagnostic yield calculation. For instance, (Colin et al., 2023),

performed ONT sequencing on only two cases from a cohort of 15 unresolved Mendelian disorders, identifying a diagnosis in one. Due to the limited testing, yield was not calculated. Similarly, (De Clercq et al., 2024), applied lrWGS to a small, preselected group of six individuals with previously identified but unresolved SVs. While two were solved, the pre-enrichment bias prevents generalization. (LaFlamme et al., 2024). used ONT to investigate methylation in three individuals with suspected imprinting disorders and successfully confirmed differentially methylated regions (DMRs) in all three; however, since lrWGS was applied only to confirm prior findings, the yield was not assessed.

With respect to the genetic findings in cases exclusively solved by lrWGS, SVs represented the largest category, accounting for 41.67% (10 out of 24 cases). These were followed by combined SV/SNV findings in a compound heterozygous state (20.83%, five out of 24), methylation-related alterations (16.66%, four out of 24), SNVs (12.50%, three out of 24), TRs, and indels, each contributing 4.17% (1 out of 24).

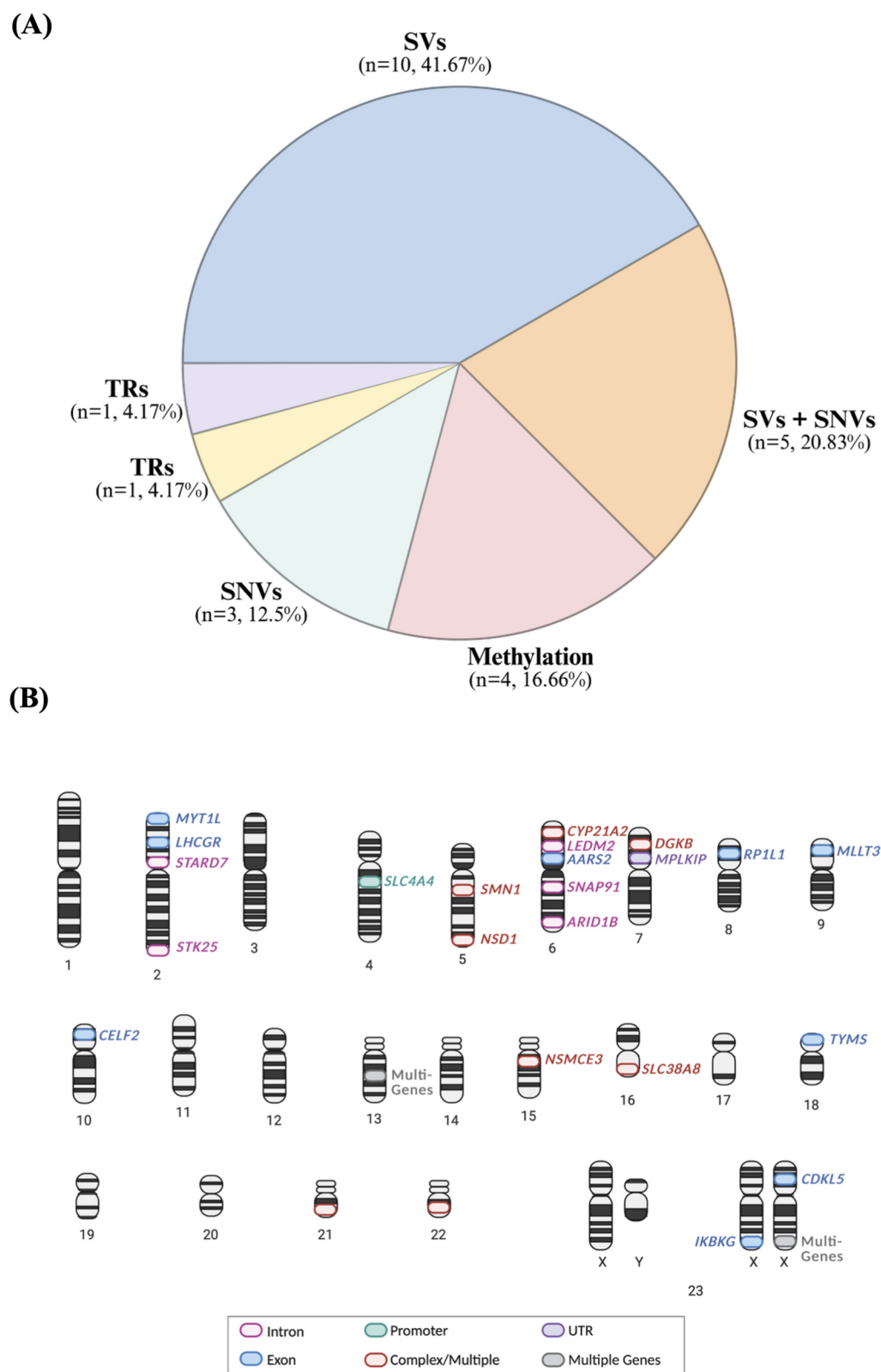


FIGURE 3 Genetic findings identified by lrWGS in cohorts with unresolved rare genetic disorders. (A) Pie chart showing the distribution of variant types identified through lrWGS. (B) Chromosomal map of affected genes based on lrWGS-detected variants. Genes are color-coded by the genomic region impacted (e.g., exon, intron, promoter, UTR, complex/multiple, or multi-gene). When a specific diagnostic gene was reported, it is shown; in cases where no single gene was specified due to complex rearrangements or broad regions, the label "Multi-Genes" is used. Detailed genomic coordinates and full gene lists are provided in the Supplementary Materials. Abbreviations: SNV, single nucleotide variant; SV, structural variant; CNV, copy number variant; TR, tandem repeat; UTR, untranslated region.

These categories highlight the well-documented limitations of SRS and the strengths of long-read platforms, which can span repetitive and GC-rich regions, resolve complex SV breakpoints, and detect non-sequence-based alterations such as differential methylation (Romagnoli et al., 2023; Wojcik et al., 2023). The affected genomic regions were diverse, with variants identified in introns, promoters, UTRs, exons, and even across multiple genes. Many of these regions are typically inaccessible or poorly covered by short-read methods, particularly WES. Notably, some diagnoses involved highly complex rearrangements such as insertion-translocation-inversion events spanning multiple genes, including *DGKB* and *MLLT3*. While PacBio HiFi reads were generally preferred for high-accuracy SNV and SV detection, ONT was frequently used in studies investigating methylation profiles or requiring rapid turnaround due to its real-time sequencing capabilities (LaFlamme et al., 2024; Negi et al., 2025; Olivucci et al., 2024; Sinha et al., 2025; Zheng et al., 2023).

Interestingly, the diagnostic gains from lrWGS were not only technical but also interpretative. In several cases, prior SRS had already flagged one pathogenic variant, but phase determination or confirmation of compound heterozygosity was not possible without parental samples. LRS enabled phasing in such situations, successfully resolving compound heterozygosity in genes like *LHCGR* and *CYP21A2* (Negi et al., 2025). Moreover, lrWGS provided novel clinical insights. Some cases involved potentially new gene-disease associations, such as *TYMS*-related lactic acidosis, while others revealed atypical presentations of known disorders, including *STK25*-associated neurodevelopmental delay. These examples highlight the ability of long-read sequencing not only to close diagnostic gaps but also to expand phenotype-genotype correlations and uncover new disease mechanisms (AlAbdi et al., 2023).

Across the studies included in our systematic review, lrWGS was most often used as a second-tier diagnostic tool, following inconclusive SRS, and when no specific genetic diagnosis was suspected. However, emerging evidence suggests that phenotype-guided LRS of known disease genes, for example, *GBA* in Parkinson's disease (Graham et al., 2020), *PKD1* in autosomal dominant polycystic kidney disease (ADPKD) (Hort et al., 2023), and *MUC1* in autosomal dominant tubulointerstitial kidney disease (Okada et al., 2022), can result in substantially higher diagnostic yields, reaching 50%–70%, particularly when guided by strong clinical suspicion.

Notably, ONT enables real-time analysis, which may support rapid sequencing in specific clinical contexts, and has been explored as a potential first-tier diagnostic tool in critically ill pediatric patients (Kamolvisit et al., 2025). Additional advantages include the simultaneous detection of methylation and sequence variants in a single assay and its capacity to phase alleles even in the absence of parental samples. However, several challenges currently limit its broader clinical implementation. These include higher base-level error rates than short-read sequencing, a lack of standardized bioinformatics pipelines, and ongoing difficulties in interpreting structural and non-coding variants. In addition, higher costs, lower throughput, and increased computational and data storage requirements pose practical barriers to routine use (Amarasinghe et al., 2020). While the overall improvement in diagnostic yield remains modest, the available evidence suggests that, when applied in targeted, phenotype-driven scenarios, long-read genome sequencing holds meaningful clinical utility and continues to contribute to the discovery of novel and complex genetic diagnoses.

Despite these promising results, our review has several limitations. Many of the included studies lacked consistency in prior genetic testing approaches, complicating comparisons of diagnostic yield. Sample sizes were generally small, and some cohorts were subject to selection bias. As reflected by the QUADAS-2 assessment, all studies had a high risk of bias in the index test domain, primarily due to a lack of blinding. Furthermore, most studies originated from North America and Europe, with only two studies from the Middle East (AlAbdi et al., 2023; Sinha et al., 2025) emphasizing the need for broader geographic and ancestral representation in future research.

5 Conclusion

Our findings highlight the emerging role of lrWGS in diagnosing previously unresolved rare genetic disorders, especially when applied after WES and combined with phasing and methylation profiling. While the overall diagnostic yield remains modest, lrWGS demonstrates clear advantages in detecting complex and previously inaccessible variants.

As sequencing technologies continue to evolve, we anticipate reductions in cost, improvements in accuracy, and the development of more targeted long-read applications, including panels or exome-level tests. With its real-time analysis, ability to phase alleles, and single-test methylation detection, lrWGS may soon find broader utility as a first-line genetic tool for both diagnostic and screening applications.

Future research could focus on expanding population diversity, standardizing prior testing protocols, and reducing bias through blinded (un-selected) assessment. With continued innovation and clinical integration, long-read genome sequencing has the potential to significantly enhance diagnostic outcomes and further our understanding of rare genetic diseases globally.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#), further inquiries can be directed to the corresponding author.

Author contributions

AI: Conceptualization, Data curation, Formal Analysis, Methodology, Writing – original draft, Writing – review and editing. KF: Conceptualization, Supervision, Validation, Writing – review and editing. AA: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing – review and editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This publication was supported by Qatar University, internal grants no QUST-1-CHS-2025-384 and QUCG-CHS-25/26-736. The findings presented herein are solely the responsibility of the authors.

Acknowledgements

We would like to acknowledge our colleagues from Sidra Medicine for their valuable discussions during the development of this systematic review, including Elbay Aliyev, Najeeb Syed, Aljazi Al-Maraghi, Waleed Aamer, Rulan Shaath, Iman Al-Azwani, Mona Abdi, Ammira S.A. Akil, and Sasirekha Palaniswamy.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

References

- Aganezov, S., Yan, S. M., Soto, D. C., Kirsche, M., Zarate, S., Avdeyev, P., et al. (2022). A complete reference genome improves analysis of human genetic variation. *Science* 376 (6588), eabl3533. doi:10.1126/science.abl3533
- AlAbdi, L., Shamseldin, H. E., Khouj, E., Helaby, R., Aljamal, B., Alqahtani, M., et al. (2023). Beyond the exome: utility of long-read whole genome sequencing in exome-negative autosomal recessive diseases. *Genome Med.* 15 (1), 114. doi:10.1186/s13073-023-01270-8
- Amarasinghe, S. L., Su, S., Dong, X., Zappia, L., Ritchie, M. E., and Gouil, Q. (2020). Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* 21 (1), 30. doi:10.1186/s13059-020-1935-5
- Ardui, S., Race, V., Zablotskaya, A., Hestand, M. S., Van Esch, H., Devriendt, K., et al. (2017). Detecting AGG interruptions in Male and female FMR1 premutation carriers by single-molecule sequencing. *Hum. Mutat.* 38 (3), 324–331. doi:10.1002/humu.23150
- Austin-Tse, C. A., Jobanputra, V., Perry, D. L., Bick, D., Taft, R. J., Venner, E., et al. (2022). Best practices for the interpretation and reporting of clinical whole genome sequencing. *NPJ Genom Med.* 7 (1), 27. doi:10.1038/s41525-022-00295-z
- Avsec, Ž., Latysheva, N., Cheng, J., Novati, G., Taylor, K. R., Ward, T., et al. (2026). Advancing regulatory variant effect prediction with AlphaGenome. *Nat.* 649 (8099), 1206–1218.
- Borras, D. M., Vossen, R., Liem, M., Buermans, H. P. J., Dauwerse, H., van Heusden, D., et al. (2017). Detecting PKD1 variants in polycystic kidney disease patients by single-molecule long-read sequencing. *Hum. Mutat.* 38 (7), 870–879. doi:10.1002/humu.23223
- Caspar, S. M., Dubacher, N., Kopps, A. M., Meienberg, J., Henggeler, C., and Matyas, G. (2018). Clinical sequencing: from raw data to diagnosis with lifetime value. *Clin. Genet.* 93 (3), 508–519. doi:10.1111/cge.13190
- Chakravarti, A. (2021). Magnitude of mendelian versus complex inheritance of rare disorders. *Am. J. Med. Genet. Part A* 185 (11), 3287–3293. doi:10.1002/ajmg.a.62463
- Cohen, A. S. A., Farrow, E. G., Abdelmoity, A. T., Alaimo, J. T., Amudhavalli, S. M., Anderson, J. T., et al. (2022). Genomic answers for children: dynamic analyses of >1000 pediatric rare disease genomes. *Genet. Med.*, 24(6), 1336–1348. doi:10.1016/j.gim.2022.02.007
- Colin, E., Duffourd, Y., Chevarin, M., Tisserant, E., Verdez, S., Paccaud, J., et al. (2023). Stepwise use of genomics and transcriptomics technologies increases diagnostic yield in Mendelian disorders. *Front. Cell Dev. Biol.* 11, 1021920. doi:10.3389/fcell.2023.1021920
- Cummings, B. B., Marshall, J. L., Tukiainen, T., Lek, M., Donkervoort, S., Foley, A. R., et al. (2017). Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci. Transl. Med.* 9 (386), eal5209. doi:10.1126/scitranslmed.aal5209
- De Clercq, G., Vantomme, L., Dewaele, B., Callewaert, B., Vanakker, O., Janssens, S., et al. (2024). Full characterization of unresolved structural variation through long-read sequencing and optical genome mapping. *Sci. Rep. Nat. Publ. Group.*, 14(1), 29142. doi:10.1038/s41598-024-80068-z
- Dierckxsens, N., Li, T., Vermeesch, J. R., and Xie, Z. (2021). A benchmark of structural variation detection by long reads through a realistic simulated model. *Genome Biol.* 22 (1), 342. doi:10.1186/s13059-021-02551-4
- Dutta, U. R., Rao, S. N., Pidugu, V. K., V, S. V., Bhattacharjee, A., Bhowmik, A. D., et al. (2019). Breakpoint mapping of a novel *de novo* translocation t(X;20)(q11.1;p13) by positional cloning and long read sequencing. *Genomics* 111 (5), 1108–1114. doi:10.1016/j.ygeno.2018.07.005
- Graham, O. E. E., Pitcher, T. L., Liau, Y., Miller, A. L., Dalrymple-Alford, J. C., Anderson, T. J., et al. (2020). Nanopore sequencing of the glucocerebrosidase (GBA) gene in a New Zealand Parkinson's disease cohort. *Park. Relat. Disord.* 70, 36–41. doi:10.1016/j.parkreldis.2019.11.022
- Gupta, S. M., Lawlor, J. M. J., Handley, L. H., Ramaker, R. C., Rogers, B. B., Partridge, E. C., et al. (2023). Familial co-segregation and the emerging role of long-read sequencing to re-classify variants of uncertain significance in inherited retinal diseases. *NPJ Genom Med.* 8 (1), 20. doi:10.1038/s41525-023-00366-9
- Hiatt, S. M., Tsai, Y. C., Clark, T. A., Kotturi, P., Dahl, N., Stattin, E. L., et al. (2018). Detailed analysis of HTT repeat elements in human blood using targeted amplification-free long-read sequencing. *Hum. Mutat.* 39 (9), 1262–1272. doi:10.1002/humu.23580
- Hort, Y., Sullivan, P., Wedd, L., Fowles, L., Stevanovski, I., Deveson, I., et al. (2023). Atypical splicing variants in PKD1 explain most undiagnosed typical familial ADPKD. *NPJ Genom Med.* 8 (1), 16. doi:10.1038/s41525-023-00362-z
- Kamolvisit, W., Cheawsamoot, C., Chetruengchai, W., Kor-Anantakul, P., Thangpong, R., Srichomthong, C., et al. (2025). Singleton rapid long-read genome sequencing as first tier genetic test for critically ill children with suspected genetic diseases. *Eur. J. Hum. Genet.* 33, 768–774. doi:10.1038/s41431-025-01818-9
- LaFlamme, C. W., Rastin, C., Sengupta, S., Pennington, H. E., Russ-Hall, S. J., Schneider, A. L., et al. (2024). Diagnostic utility of DNA methylation analysis in genetically unsolved pediatric epilepsies and CHD2 epismutation refinement. *Nat. Commun.* 15 (1), 6524. doi:10.1038/s41467-024-50159-6
- Loomis, E. W., Eid, J. S., Peluso, P., Yin, J., Hickey, L., Rank, D., et al. (2013). Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Res.* 23 (1), 121–128. doi:10.1101/gr.141705.112
- Mani, A. (2017). Pathogenicity of *de novo* Rare Variants: challenges and Opportunities. *Circ. Cardiovasc. Genet.* 10 (6), e002013. doi:10.1161/CIRCGENETICS.117.002013
- Mantere, T., Kersten, S., and Hoischen, A. (2019). Long-read sequencing emerging in medical genetics. *Front. Genet.* 10, 426. doi:10.3389/fgene.2019.00426
- Marwaha, S., Knowles, J. W., and Ashley, E. A. (2022). A guide for the diagnosis of rare and undiagnosed disease: beyond the exome. *Genome Med.* 14 (1), 23. doi:10.1186/s13073-022-01026-w

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2026.1809097/full#supplementary-material>

- Mastrorosa, F. K., Miller, D. E., and Eichler, E. E. (2023). Applications of long-read sequencing to Mendelian genetics. *Genome Med.* 15 (1), 42. doi:10.1186/s13073-023-01194-3
- Merker, J. D., Wenger, A. M., Sneddon, T., Grove, M., Zappala, Z., Fresard, L., et al. (2018). Long-read genome sequencing identifies causal structural variation in a Mendelian disease. *Genet. Med.* 20 (1), 159–163. doi:10.1038/gim.2017.86
- Miao, H., Zhou, J., Yang, Q., Liang, F., Wang, D., Ma, N., et al. (2018). Long-read sequencing identified a causal structural variant in an exome-negative case and enabled preimplantation genetic diagnosis. *Hereditas* 155, 32. doi:10.1186/s41065-018-0069-1
- Mitsuhashi, S., and Matsumoto, N. (2020). Long-read sequencing for rare human genetic diseases. *J. Hum. Genet.* 65 (1), 11–19. doi:10.1038/s10038-019-0671-8
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* 4 (1), 1. doi:10.1186/2046-4053-4-1
- Negi, S., Stenton, S. L., Berger, S. I., Canigiula, P., McNulty, B., Violich, I., et al. (2025). Advancing long-read nanopore genome assembly and accurate variant calling for rare disease detection. *Am. J. Hum. Genet.* 112 (2), 428–449. doi:10.1016/j.ajhg.2025.01.002
- Oehler, J. B., Wright, H., Stark, Z., Mallett, A. J., and Schmitz, U. (2023). The application of long-read sequencing in clinical settings. *Hum. Genomics* 17 (1), 73. doi:10.1186/s40246-023-00522-3
- Okada, E., Morisada, N., Horinouchi, T., Fujii, H., Tsuji, T., Miura, M., et al. (2022). Detecting MUC1 variants in patients clinicopathologically diagnosed with having autosomal dominant tubulointerstitial kidney disease. *Kidney Int. Rep.* 7 (4), 857–866. doi:10.1016/j.ekir.2021.12.037
- Olivucci, G., Iovino, E., Innella, G., Turchetti, D., Pippucci, T., and Magini, P. (2024). Long read sequencing on its way to the routine diagnostics of genetic diseases. *Front. Genet.* 15, 1374860. doi:10.3389/fgene.2024.1374860
- Pietryga, A., Chilinski, M., Gadakh, S., and Plewczynski, D. (2025). ConsensusSV-ONT - a modern method for accurate structural variant calling. *Sci. Rep.* 15 (1), 17144. doi:10.1038/s41598-025-01486-1
- Qu, Y. J., Yang, Z. R., Sun, F., and Zhan, S. Y. (2018). Risk on bias assessment: (6) A revised tool for the quality assessment on diagnostic accuracy studies (QUADAS-2). *Zhonghua Liu Xing Bing Xue Za Zhi* 39 (4), 524–531. doi:10.3760/cma.j.issn.0254-6450.2018.04.028
- Rahit, K. T. H., and Tarailo-Graovac, M. (2020). Genetic modifiers and rare Mendelian disease. *Genes* 11 (3), 239. doi:10.3390/genes11030239
- Romagnoli, S., Bartalucci, N., and Vannucchi, A. M. (2023). Resolving complex structural variants via nanopore sequencing. *Front. Genet.* 14, 1213917. doi:10.3389/fgene.2023.1213917
- Sanford Kobayashi, E., Batalov, S., Wenger, A. M., Lambert, C., Dhillon, H., Hall, R. J., et al. (2022). Approaches to long-read sequencing in a clinical setting to improve diagnostic rate. *Sci. Rep.* 12 (1), 16945. doi:10.1038/s41598-022-20113-x
- Sathyanarayana, S. H., Deharvengt, S. J., Huang, G., Barney, R. E., Shah, P. S., and Lefferts, J. A. (2022). Applications of long-read sequencing technology in clinical genomics. *Adv. Mol. Pathology* 5 (1), 85–108. doi:10.1016/j.yamp.2022.07.004
- Sinha, S., Rabea, F., Ramaswamy, S., Chekroun, I., El Naofal, M., Jain, R., et al. (2025). Long read sequencing enhances pathogenic and novel variation discovery in patients with rare diseases. *Nat. Commun.* 16 (1), 2500. doi:10.1038/s41467-025-57695-9
- Sullivan, J. A., Schoch, K., Spillmann, R. C., and Shashi, V. (2023). Exome/genome sequencing in undiagnosed syndromes. *Annu. Rev. Med.* 74, 489–502. doi:10.1146/annurev-med-042921-110721
- van Dijk, E. L., Naquin, D., Gorrichon, K., Jaszczyszyn, Y., Ouazahrour, R., Thermes, C., et al. (2023). Genomics in the long-read sequencing era. *Trends Genet.* 39 (9), 649–671. doi:10.1016/j.tig.2023.04.006
- Weischenfeldt, J., Symmons, O., Spitz, F., and Korbel, J. O. (2013). Phenotypic impact of genomic structural variation: insights from and for human disease. *Nat. Rev. Genet.* 14 (2), 125–138. doi:10.1038/nrg3373
- Wojcik, M. H., Reuter, C. M., Marwaha, S., Mahmoud, M., Duyzend, M. H., et al. (2023). Beyond the exome: what's next in diagnostic testing for Mendelian conditions. *Am. J. Hum. Genet.* 110 (8), 1229–1248. doi:10.1016/j.ajhg.2023.06.009
- Wright, C. F., FitzPatrick, D. R., and Firth, H. V. (2018). Paediatric genomics: diagnosing rare disease in children. *Nat. Rev. Genet.* 19 (5), 253–268. doi:10.1038/nrg.2018.12
- Yadav, D., Patil-Takbhate, B., Khandagale, A., Bhawalkar, J., Tripathy, S., and Khopkar-Kale, P. (2023). Next-generation sequencing transforming clinical practice and precision medicine. *Clin. Chim. Acta* 551, 117568. doi:10.1016/j.cca.2023.117568
- Yu, S. Y., Xi, Y. L., Xu, F. Q., Zhang, J., and Liu, Y. S. (2023). Application of long read sequencing in rare diseases: the longer, the better? *Eur. J. Med. Genet.* 66 (12), 104871. doi:10.1016/j.ejmg.2023.104871
- Zheng, P., Zhou, C., Ding, Y., Liu, B., Lu, L., Zhu, F., et al. (2023). Nanopore sequencing technology and its applications. *MedComm* (2020) 4 (4), e316. doi:10.1002/mco2.316
- Zhong, Y., Xu, F., Wu, J., Schubert, J., and Li, M. M. (2021). Application of next generation sequencing in laboratory medicine. *Ann. Lab. Med.* 41 (1), 25–43. doi:10.3343/alm.2021.41.1.25