Global Medical Genetics 12 (2025) 100038

Contents lists available at ScienceDirect

Global Medical Genetics

journal homepage: www.keaipublishing.com/en/journals/gmg/

Unlocking the genetic blueprint of duchenne muscular dystrophy: A personalized approach with MLPA and WES

Priyanshu Mathur[a], Ashmeet Kaur[b,*], Kamlesh Kumar Agarwal[a], Lokesh Kumar Agarwal[a], Avisha Mathur[c], Deepti Choudhary[d]

[a] Department of Pediatric Medicine, SMS Medical College, Jaipur, India
[b] Department of Pathology, SMS Medical College, Jaipur, India
[c] Department of Ophthalmology, SMS Medical College, Jaipur, India
[d] LifeCell International Pvt Ltd, Vijayaraghava Road, I Lane, T. Nagar, Chennai 600017, India

ARTICLE INFO

Keywords:
Duchenne muscular dystrophy
Dystrophin gene
MLPA
Whole exome sequencing
Genetic variants
Novel variants
Genotype-phenotype correlation
Exon deletions

ABSTRACT

Background: Duchenne muscular dystrophy (DMD) is a progressive X-linked disorder causing muscle degeneration and multisystem involvement, requiring precise genetic diagnosis for timely intervention and treatment.
Objective: To investigate the genetic landscape of DMD using a two-tiered diagnostic approach combining MLPA and WES, and to correlate genetic findings with clinical outcomes for improved management.
Materials and methods: A cross-sectional study of 80 male DMD patients was conducted using a sequential genetic approach, combining MLPA and WES, with bioinformatics and statistical analyses to explore genotype-phenotype correlations.
Results: Pathogenic variants were identified in 65 cases (81.2 %), with deletions (67.5 %) being the most common, followed by duplications (6.3 %), splice-site (3.8 %), and nonsense variants (3.8 %). WES identified additional pathogenic variants in MLPA-negative cases, including novel mutations, expanding the known genetic spectrum of DMD. The combined MLPA-WES approach significantly improved diagnostic yield ($\chi^2 = 12.90$, $p < 0.001$). Functional analysis revealed disruptions in glycogen metabolism (46 %), calcium transport (24 %), and mitochondrial function (12 %), with dystrophin-associated proteins (DAG1, SGCD) critically involved in muscle stability. Out-of-frame deletions were significantly associated with early disease onset ($\chi^2 = 49.03$, $p < 0.001$) and severe phenotypes ($\chi^2 = 47.04$, $p < 0.001$), supporting exon-skipping therapy. In-frame deletions correlated with milder progression, while nonsense variants posed a 2.5-fold increased risk of early cardiomyopathy ($p = 0.002$), emphasizing the need for early intervention.
Conclusion: Combining MLPA and WES enhances DMD diagnostic accuracy, enabling timely clinical interventions. Integrating functional analysis with genotype-phenotype correlations supports personalized therapeutic strategies, improving patient outcomes.

* Correspondence to: 23 Vidyut Nagar A, Prince Road, Ajmer road, Jaipur 302021, India.
E-mail address: ashmeetkochar@gmail.com (A. Kaur).

https://doi.org/10.1016/j.gmg.2025.100038
Received 5 January 2025; Received in revised form 24 January 2025; Accepted 25 January 2025
Available online 6 February 2025
2699-9404/© 2025 The Author(s). Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Duchenne muscular dystrophy (DMD) is a severe X-linked recessive disorder caused by pathogenic variants in the dystrophin gene, essential for maintaining the structural integrity of muscle fibers. Affecting approximately 1 in 5000 male births worldwide, DMD is characterized by progressive muscle weakness, loss of ambulation, and premature death due to respiratory or cardiac failure [1,2]. The dystrophin gene, the largest known human gene, spans 2.4 Mb on the X chromosome, with over 70 % of pathogenic variants resulting in deletions or duplications of one or more exons [3].

The genotype-phenotype relationship in DMD is significantly influenced by the nature and location of these pathogenic variants, particularly their effect on the reading frame of the dystrophin transcript. Out-of-frame pathogenic variants produce nonfunctional dystrophin, resulting in severe phenotypes, while in-frame pathogenic variants allow partial dystrophin production, often leading to milder phenotypes [4,5]. This understanding is vital for improving diagnostic strategies and developing optimized therapeutic approaches.

Beyond skeletal muscle instability, dystrophin loss causes systemic effects, including progressive cardiomyopathy driven by sarcolemmal instability and myocardial fibrosis, which are major contributors to morbidity and mortality in DMD [1,6]. Dystrophin isoforms in the central nervous system (CNS) are linked to cognitive and behavioral functions, with their absence associated with intellectual disabilities, attention deficits, and autism spectrum disorders in some patients [6]. The loss of dystrophin-associated proteins (DAPs) further exacerbates these effects, disrupting membrane repair, signaling pathways, and cellular stress responses. Addressing these secondary manifestations through therapies targeting oxidative stress and myocardial fibrosis is crucial for improving outcomes [6]. Multidisciplinary care, including cardiac monitoring and neurocognitive support, is equally essential.

Accurate mutation detection is pivotal for guiding diagnosis and treatment planning. Multiplex Ligation-Dependent Probe Amplification (MLPA) serves as a robust first-line tool for identifying large deletions and duplications in the dystrophin gene [7,8]. However, it cannot detect smaller pathogenic variants, such as point pathogenic variants or splice variants. Whole Exome Sequencing (WES) complements MLPA by providing a comprehensive analysis of coding regions, uncovering pathogenic variants often missed by traditional methods [9,10]. By combining MLPA for large-scale mutation detection with WES for smaller variants, this study investigates the prevalence and spectrum of pathogenic variants, their phenotypic correlations, and their implications for diagnostic and therapeutic strategies.

In the context of precision medicine, tailoring treatments to individual mutation profiles has gained prominence. Exon-skipping therapies, such as eteplirsen and golodirsen, restore the reading frame of specific pathogenic variants, enabling the production of truncated yet functional dystrophin [11]. For patients ineligible for exon-skipping therapies, gene transfer approaches like Elevidys (delandistrogene moxeparvovec) provide a significant advancement by delivering a functional dystrophin-like protein and have gained approval in several regions worldwide [12]. Additionally, gene-editing technologies like CRISPR-Cas9 hold the promise of directly correcting pathogenic variants, offering long-term solutions [11,13]. Therapies targeting secondary mechanisms, such as oxidative stress, further enhance this personalized approach, mitigating disease progression and improving muscle function [11].

This study aims to comprehensively explore the genetic landscape of Duchenne muscular dystrophy (DMD) through a two-tiered diagnostic approach that combines Multiplex Ligation-dependent Probe Amplification (MLPA) and Whole Exome Sequencing (WES). By integrating genotypic data with bioinformatics tools, such as Gene Ontology (GO) enrichment and protein interaction network analyses, alongside clinical evaluation, the study seeks to elucidate the functional impact of genetic alterations, identify disrupted biological pathways, and explore potential therapeutic targets. These insights contribute to the advancement of personalized treatment strategies tailored to individual genetic profiles.

Materials and methods

This cross-sectional study was conducted on suspected male Duchenne muscular dystrophy (DMD) patients attending the outpatient department (OPD) of J.K. Lon Hospital, S.M.S. Medical College, Jaipur. The study aimed to investigate pathogenic variants in the dystrophin gene using a sequential approach: Multiplex Ligation-dependent Probe Amplification (MLPA) to detect large deletions and duplications, followed by Whole Exome Sequencing (WES) for MLPA-negative cases to identify smaller pathogenic variants. WES was performed at LifeCell International, Chennai.

Inclusion criteria

Male children presenting with clinical indicators of DMD, such as elevated creatine phosphokinase (CPK) levels, proximal muscle weakness, calf-muscle hypertrophy, a positive Gowers sign, and early-onset symptoms (≤ 6 years) or late-onset symptoms (> 6 years). Detailed clinical histories and evaluations were conducted.

Exclusion criteria

Individuals with alternative confirmed diagnoses explaining their symptoms or incomplete clinical/laboratory data were excluded.

Ethical clearance

Approval was obtained from the Ethics Committee, S.M.S. Medical College, Jaipur (Approval number: 1025MC/EC/2021), dated 18th November 2021.

Sample collection and DNA extraction

Peripheral venous blood samples were collected from 80 participants. Genomic DNA was extracted using the QIAamp® DNA Blood Mini Kit (Qiagen Gmbh, Germany), and quality/quantity assessments were performed using NanoDrop and Qubit Fluorometer, with integrity validated by agarose gel electrophoresis.

MLPA analysis

Large deletions and duplications in the dystrophin gene were screened using the SALSA MLPA probemix (MRC Holland), with results analyzed using Coffalyser.Net software. MLPA served as the primary screening method due to its high sensitivity in detecting structural variants.

Whole Exome Sequencing (WES)

WES was conducted on MLPA-negative cases to detect smaller pathogenic variants, such as point variants and splice site variants. Exonic and flanking intronic regions were captured using hybridization-based methods and sequenced on Illumina platforms at a mean depth of 50–60X, ensuring $>$ 90 % coverage at 20X depth.

The analysis targeted the dystrophin gene along with 133 additional genes associated with dystrophinopathy phenocopies (e.g., SGCA, SGCB, SGCD, SGCG, LAMA2, ANO5), facilitating a comprehensive differential diagnosis (Table S1).

Alignment and variant annotation

Sequences were aligned to the human reference genome (GRCh37/hg19) using the BWA-MEM aligner, and variants were identified using the Sentieon HaplotypeCaller module. Annotation was performed using publicly available databases (OMIM, GNOMAD, GWAS, 1000 Genomes Project). Pathogenicity classification followed ACMG guidelines, incorporating population frequency, computational evidence, segregation analysis, and functional studies.

Functional analysis

Functional analysis was performed using the STRING database (v11.5) to analyze dystrophin-centered protein interactions [14]. High-confidence interactions (score $>$ 0.7) were selected to map the network, with experimental and predicted interactions considered. Gene Ontology (GO) enrichment analysis was conducted using the DAVID platform [15,16] to identify key biological processes, cellular components, and molecular functions associated with dystrophinopathy. The top enriched terms were selected based on statistical significance ($p < 0.05$), and data visualization was done using donut charts for interpretation.

Statistical analysis

Descriptive statistics summarized demographic and clinical data. Chi-square and Fisher's exact tests were used for categorical variables, and t-tests/ANOVA for continuous variables. Logistic regression analyzed genotype-phenotype correlations, adjusting for confounders. Statistical significance was set at $p < 0.05$.

Results

Patient cohort characteristics

The cohort comprised 80 male children suspected of having Duchenne muscular dystrophy (DMD), with a mean age of 7.92 years (SD: 3.51) and a median age of 8 years, ranging from 1 to 17 years.

Genetic variants identified

Pathogenic variants were identified in 65 cases (81.2 %), with deletions being the most common type (67.5 %), followed by duplications (6.3 %), splice site variants (3.8 %), and nonsense variants (3.8 %). The pie chart (Fig. 1) represents the distribution of different mutation types identified in the cohort. Deletions were identified as the predominant genetic alteration. The sequential use of MLPA followed by WES significantly improved the diagnostic yield (chi-square = 12.90, $p = 0.0003$), with MLPA identifying pathogenic variants in 59 cases, predominantly deletions, and WES detecting an additional 6 pathogenic variants (3 splice-site and 3 nonsense), increasing the diagnostic yield to 81.2 %. Despite these advancements, 15 cases (18.8 %) remained without identified

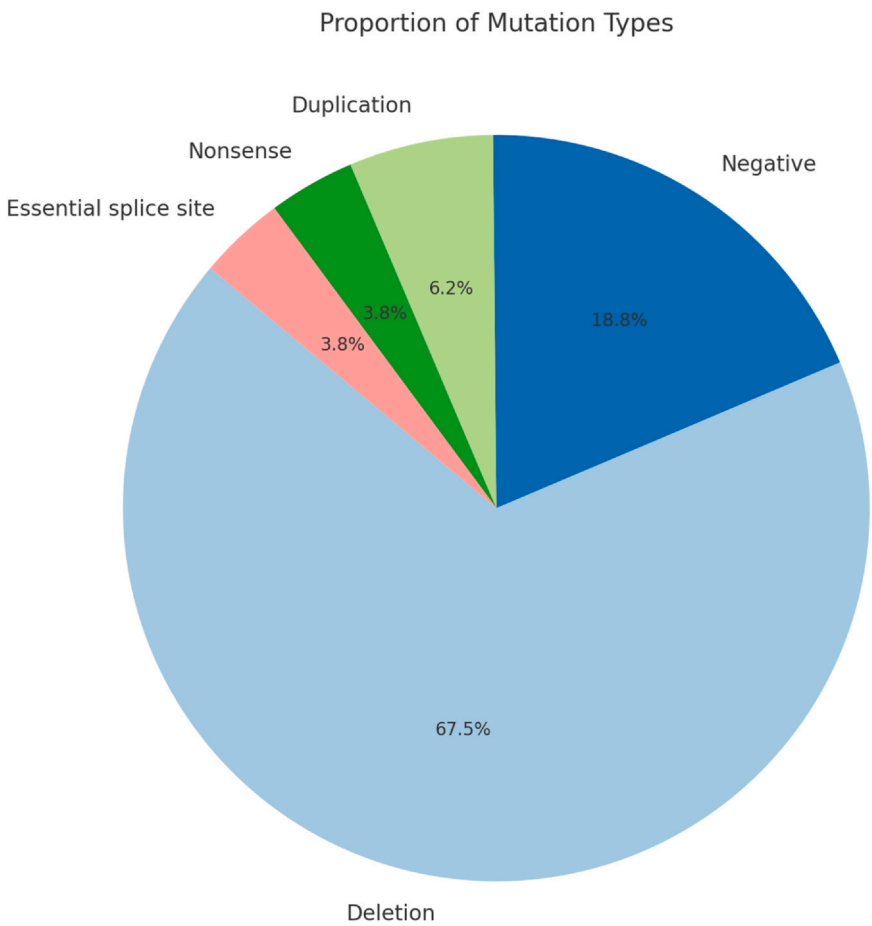

Fig. 1. Proportion of Mutation Types in Duchenne muscular dystrophy (DMD) cohort.

pathogenic variants, potentially harboring variants in non-coding regulatory or deep intronic regions or complex structural rearrangements beyond WES detection. Advanced diagnostic techniques such as RNA sequencing or long-read sequencing may help uncover such variants.

Novel pathogenic variants identified by WES

Whole exome sequencing identified four novel pathogenic variants as illustrated in Table 1, expanding the genetic spectrum of DMD in the Indian population. These included the c .5959 G > T (p.Glu1987Ter) mutation in exon 42, introducing a premature stop codon and resulting in severe disease progression; the c.2293–1 G > T splice-site variant in intron 18, leading to exon skipping and severe clinical outcomes; and the c.1149 + 2 T > C mutation in intron 10 and c.9286 + 1 G > C in intron 63, both of which disrupted splicing and contributed to varying disease severities. Table1 summarizes the variants detected in the dystrophin gene through Whole Exome Sequencing (WES) in Duchenne muscular dystrophy (DMD) patients.

Exon-specific analysis

Deletions were the most frequently observed variant, with the exon 41–50 range being the primary hotspot (47.76 %). Out-of-frame deletions in this region accounted for 52.46 % of cases and were strongly associated with severe Duchenne muscular dystrophy (DMD) phenotypes. In contrast, in-frame deletions were more common in the exon 0–10 (42.86 %) and 11–20 (28.57 %) ranges, correlating with milder Becker muscular dystrophy (BMD) phenotypes. Duplications were predominantly found in exon range 0–10 (37.50 %), while splice-site variants were mostly identified in exon ranges 0–10 (66.67 %) and 11–20 (33.33 %). Nonsense variants were exclusively located in exon range 0–10. The bar chart (Fig. 2) illustrates the distribution of various mutation types across different exon ranges in the dystrophin gene. Exon analysis demonstrated distinct deletion patterns and their correlation with clinical phenotypes, with deletions involving exons 45 and 46 strongly associated with severe early-onset DMD cases. In contrast, deletions in exons 1–10 and 31–40 were linked to later-onset, milder BMD-like phenotypes. Statistically significant associations were observed between exon deletions and symptom onset age ($p < 0.05$).

Table 1
Variants Identified in the Dystrophin Gene by Whole Exome Sequencing in DMD Patients.

SNo	Age Group	Mutation Type	Exon/Intron Location	Mutation Details	Clinical Outcome	Novelty	Additional Notes
1.	6	c .5959 G > T (p.Glu1987Ter)	Exon 42	Stop Gain	Likely Pathogenic	Novel	Not previously reported in databases; results in a truncated dystrophin protein.
2.	5	c .10171 C > T (p.Arg3391Ter)	Exon 70	Nonsens-e/stop	Pathogenic	Not Novel	Associated with a mild phenotype.
3.	10	c.2293 −1 G > T	Intron 18 (splice acceptor site of exon 19)	Essential splice site	Likely Pathogenic	Novel	Novel splice variant causing exon skipping, leading to severe clinical outcomes.
4.	11	c.1149 + 2 T > C	Intron 10	Essential splice site	Pathogenic	Novel	Not reported in 1000 Genomes or gnomAD; disrupts splicing, leading to severe phenotype.
5.	8	c.9286 + 1 G > C	Intron 63	Essential splice site	Likely Pathogenic	Novel	Affects the GT donor splice site; leads to mild phenotype but disrupts normal splicing.
6.	10	c .3259 C > T (p.Gln1087Ter)	Exon 24	Non sense	Likely Pathogenic	Not Novel	Associated with a mild phenotype.

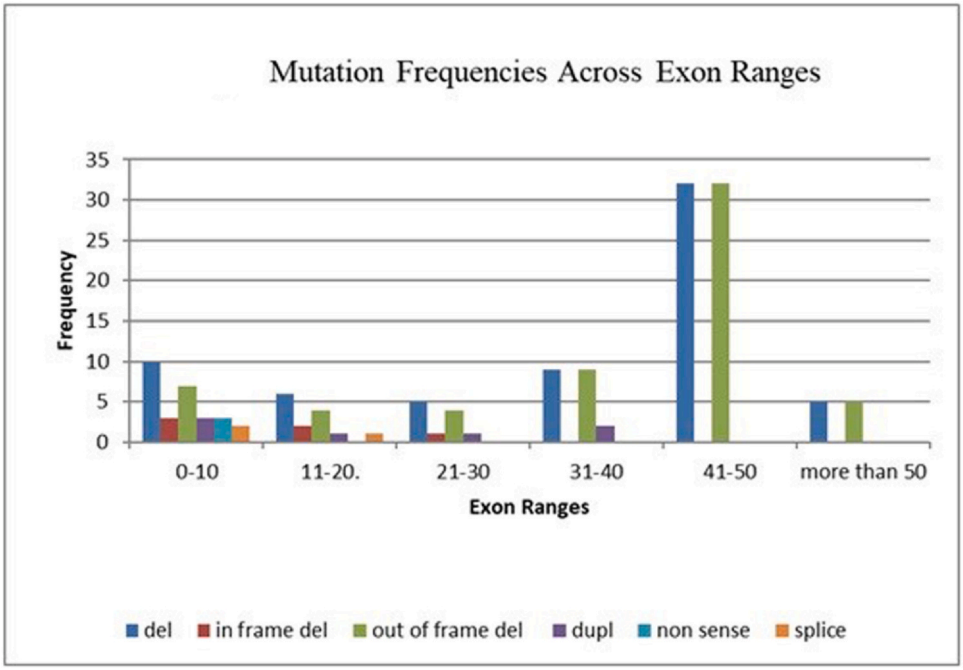

Fig. 2. Correlation Between Deletion Frequencies and Age of Symptom Onset Across exon ranges.

Age-based analysis of variants

For the purpose of analysis, the cohort was categorized into two groups based on the age of symptom onset: an early-onset group (≤6 years) and a late-onset group (> 6 years). Among the 22 early-onset cases (27.5 %), out-of-frame deletions were predominant (68.2 %, 15 cases) and were associated with severe clinical phenotypes, although statistical significance was not observed ($p = 0.234$). Other variants identified in this group included in-frame deletions (13.6 %, 3 cases), duplications (9.1 %, 2 cases), and rare pathogenic variants (9.1 %, 2 cases).

In the late-onset group (72.5 %), out-of-frame deletions were also the most common (53.4 %, 31 cases), but with slower disease progression ($p = 0.234$). Other variants in this group included in-frame deletions (8.6 %, 5 cases), duplications (5.2 %, 3 cases), and rare pathogenic variants (6.9 %, 4 cases). The difference in the frequency of rare pathogenic variants between the two groups was not statistically significant ($p = 0.739$). Fisher's Exact Test suggested a higher proportion of severe pathogenic variants in early-onset cases (Odds Ratio = 2.23, $p = 0.195$).

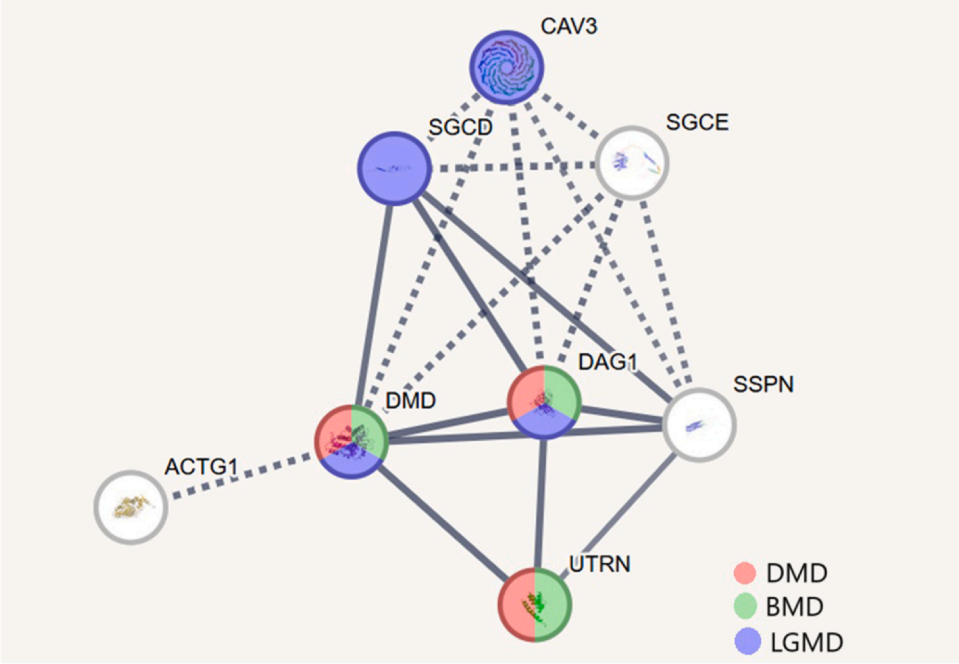

Fig. 3. Dystrophin-Centered Protein Interaction Network in Muscular Dystrophies (Alpha-Dystroglycan, Sarcospan, Utrophin, Actin Gamma, Epsilon-Sarcoglycan, Delta-Sarcoglycan, Caveolin-3).

Statistical analysis of variant distribution

Chi-square analysis was performed to assess the distribution of variant types across the two age groups and different exon ranges. Deletions were significantly associated with specific exon ranges in both early-onset and late-onset groups ($\chi 2 = 22.18$, $p < 0.001$), with the exon 41–50 region showing the highest frequency. Duplications did not show a significant association with age groups or exon ranges ($\chi 2 = 8.12$, $p = 0.15$). Nonsense variants were found exclusively in patients aged > 6 years ($\chi 2 = 10.45$, $p < 0.05$), primarily affecting exon 41–50. Splice-site variants showed a significant association with older patients and specific exon ranges ($\chi 2 = 15.32$, $p < 0.01$).

Functional analysis of protein interactions

Functional analysis using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (Fig. 3) revealed dystrophin-centered protein interaction networks, identifying key interacting partners such as dystroglycan 1 (DAG1), sarcoglycan delta (SGCD), sarcospan (SSPN), utrophin (UTRN), and actin gamma 1 (ACTG1), which play crucial roles in maintaining muscle integrity. Experimentally validated and predicted interactions (confidence scores 0.937–0.998) emphasized dystrophin's involvement in sarcolemmal stability. Disruptions in the sarcoglycan complex components SGCD and sarcoglycan epsilon (SGCE) were found to compromise membrane integrity, while utrophin (UTRN) emerged as a potential compensatory mechanism.

Gene Ontology (GO) enrichment analysis

Gene Ontology enrichment analysis via the DAVID platform identified significant impairments in key biological pathways, with glycogen metabolism (46 %), calcium transport (24 %), and mitochondrial function (12 %) being the most affected. Additional disrupted pathways included sarcoplasmic reticulum (5 %), muscle protein synthesis (9 %), actin cytoskeletal organization (9 %), and heme synthesis (2 %). The donut chart (Fig. 4) presents the results of GO enrichment analysis, showing the percentage distribution of key biological processes affected by dystrophin mutations.

Genotype-phenotype correlation by variant type

In-frame deletions within exons 1–10 and 31–40 were associated with a slower disease course. Out-of-frame deletions were significantly linked to early disease onset ($\chi^2 = 49.03$, $p < 0.001$) and severe phenotypes, including loss of ambulation ($\chi^2 = 47.04$, $p < 0.001$). Duplications spanning exons 2–20 exhibited variable severity with muscular and cardiac symptoms, with no significant association observed with cardiac complications ($\chi^2 = 4.22$, $p = 0.376$).

Splice-site variants disrupted dystrophin transcript processing, leading to combined muscular and neurocognitive impairments, with a significant association observed in older patients and specific exon ranges ($\chi^2 = 15.32$, $p < 0.01$). Nonsense variants, particularly in

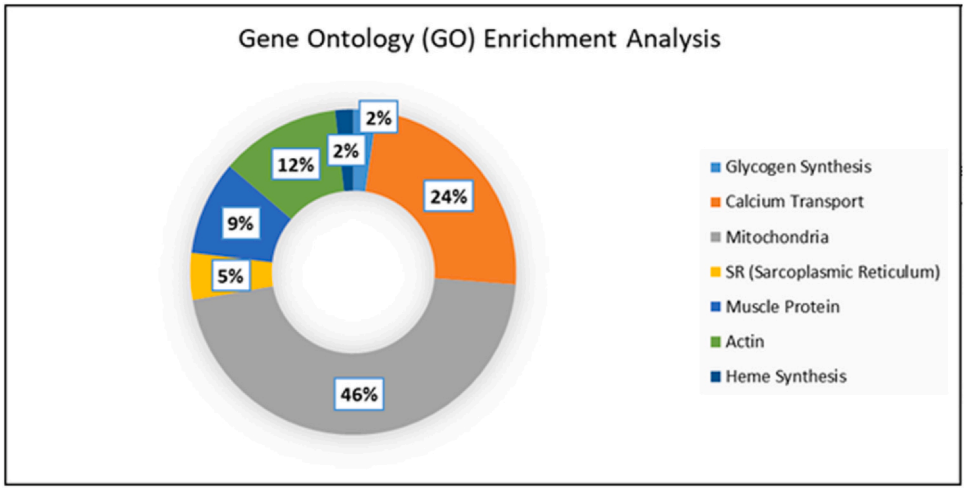

Fig. 4. GO Enrichment Analysis: Biological Processes, Molecular Functions, and Cellular Components.

exons 60–70, resulted in severe phenotypes due to premature termination of dystrophin synthesis, contributing to progressive muscle degeneration and cardiomyopathy ($\chi^2 = 10.45$, $p < 0.05$). Nonsense variants, associated with truncated dystrophin and oxidative stress, demonstrated a 2.5-fold increased risk of early cardiomyopathy ($p = 0.002$). Table 2 presents the correlation between identified genetic variant and their molecular impact, clinical phenotype, and potential therapeutic approaches in DMD patients.

Discussion

This study provides a comprehensive analysis of Duchenne muscular dystrophy (DMD) using a stepwise diagnostic approach, beginning with Multiplex Ligation-dependent Probe Amplification (MLPA) followed by Whole Exome Sequencing (WES), achieving a diagnostic yield of 81.2 %. MLPA identified pathogenic variants in 59 cases (73.8 %), with deletions being the predominant mutation type. Notably, exons 45 and 46 were the most frequently deleted, emphasizing their critical role in dystrophin functionality. These deletions were found to significantly impact protein interactions and compromise sarcolemmal stability, findings that are consistent with previous studies conducted by Goyal et al. [17] and Werneck et al. [18].

Regional variations in mutation patterns were observed in our study, particularly when compared with the findings of Orso et al. [19]. While our cohort exhibited the highest frequency of deletions in exons 45 and 46, Orso et al. [19] reported a greater prevalence of deletions in the exon 50–52 region among the Italian population. These discrepancies highlight potential population-specific genetic variations, which may be influenced by ethnic diversity, environmental factors, and regional screening practices. Furthermore, while our study predominantly identified deletions in the central hotspot region, Orso et al. observed broader variations across different exon clusters. These findings underscore the importance of region-specific genetic data in refining diagnostic approaches and enhancing our understanding of DMD's genetic heterogeneity.

WES was performed specifically for the 21 MLPA-negative cases, detecting rare pathogenic variants—3 splice site and 3 nonsense pathogenic variants—thus contributing an additional 7.6 % to the overall diagnostic yield. Statistical analysis confirmed that the combined MLPA-WES approach significantly enhanced diagnostic capability compared to MLPA alone (chi-square = 12.90,

Table 2
Genotype-Phenotype Correlation in DMD.

SNo	Genotype	Molecular Impact	Clinical Phenotype	Therapeutic Approach
1.	**Out-of-frame deletions**	Severe protein interaction loss, calcium dysregulation	Early onset, rapid progression	Exon-skipping therapies, gene therapy
2.	**In-frame deletions**	Partial dystrophin function, mild pathway disruption	Milder phenotype, slower progression	Supportive therapy, corticosteroids
3.	**Duplications**	Disrupted transcript stability	Variable severity, cardiac involvement	Gene therapy, exon-skipping
4.	**Splice site variants**	Aberrant splicing, severe protein loss	Severe phenotype, cognitive involvement	Antisense oligonucleotide-based therapies
5.	**Nonsense mutations**	Premature truncation, oxidative stress	Severe weakness, early cardiomyopathy	Nonsense read-through drugs, personalized therapy
6.	**Deep intronic mutations**	Regulatory disruption, variable protein expression	Variable severity, often milder presentation	Advanced sequencing for precision diagnosis

p = 0.0003). Rare pathogenic variants, such as splice site and nonsense variants, were closely linked to severe clinical presentations, including early symptom onset (≤6 years) and rapid progression to loss of ambulation.

The results of our study align closely with those reported by Juan-Mateu et al. [20], who utilized MLPA and sequencing techniques to establish genotype-phenotype correlations. In our study, MLPA identified pathogenic variants in 73.8 % of cases, while WES contributed an additional 7.6 %, resulting in a total diagnostic yield of 81.2 %. Juan-Mateu et al. [20], in contrast, reported a higher diagnostic yield of 100 %, with MLPA detecting variants in 81.8 % of cases and sequencing identifying all remaining variants [19]. A key distinction between the two studies is the sample size, as Juan-Mateu et al. analyzed a larger population and incorporated muscle cDNA sequencing, which provided deeper insights into transcript-level abnormalities.

Deletion hotspots in exons 45 (38 %) and 50 (24 %) were prominent in our cohort. Out-of-frame deletions accounted for 85.2 % of cases and were strongly associated with severe phenotypes, rapid disease progression, and early loss of ambulation due to the complete absence of functional dystrophin. Chi-square analysis demonstrated a significant association between genotype and early onset (≤6 years) (χ^2 = 49.03, p < 0.001) and severe phenotypes such as loss of ambulation (χ^2 = 47.04, p < 0.001). Although these deletions were more common in the early-onset group (Odds Ratio = 2.23, p = 0.195), the difference was not statistically significant. In contrast, in-frame deletions observed in 18 % of cases were linked to milder Becker muscular dystrophy (BMD)-like presentations, with slower disease progression. Duplications spanning exons 2–20 exhibited variable severity with muscular and cardiac symptoms but did not show a significant association with cardiac complications (χ^2 = 4.22, p = 0.376). These findings reinforce the well-established reading frame rule, which remains a valuable predictor of clinical severity. They also align with Falzarano et al., who identified exons 45–55 as critical for maintaining the dystrophin-glycoprotein complex (DGC) and their importance in therapeutic development.

Splice-site variants disrupted dystrophin transcript processing, leading to combined muscular and neurocognitive impairments, with a significant association observed in older patients and specific exon ranges (χ^2 = 15.32, p < 0.01). Nonsense variants, particularly in exons 60–70, resulted in severe phenotypes due to premature termination of dystrophin synthesis, contributing to progressive muscle degeneration and cardiomyopathy (χ^2 = 10.45, p < 0.05). Nonsense variants, associated with truncated dystrophin and oxidative stress, demonstrated a 2.5-fold increased risk of early cardiomyopathy (p = 0.002).

Comparison with the study by Sekar et al. [21] reveals notable differences in genetic analysis scope and clinical interpretations. While Sekar et al. [21] focused primarily on detecting deletions and duplications using MLPA, our study provided a broader perspective by integrating WES to capture smaller pathogenic variants such as splice-site and nonsense variants, which are often missed by MLPA. Both studies identified exon deletions in the 45–50 region as the most common, associated with severe early-onset phenotypes. However, our study offers a more comprehensive diagnostic approach, enhancing the detection of rare variants and their clinical implications.

In addition to known pathogenic variants, our study identified several novel variant in the dystrophin gene that have not been previously reported in global databases such as 1000 Genomes and gnomAD, highlighting population-specific genetic variations in the Indian cohort. Among these, the novel stop-gain mutation c .5959 G > T (p.Glu1987Ter) in exon 42 results in a truncated dystrophin protein, potentially contributing to disease severity. Additionally, the novel splice-site variant c.2293–1 G > T in intron 18, leading to exon skipping, was associated with severe clinical outcomes. These findings contrast with previous Indian cohort studies, such as Sekar et al. [21], which primarily focused on deletions in the exon 45–55 hotspot region. Unlike these earlier studies, our findings broaden the known mutational spectrum, capturing a diverse range of single nucleotide variants (SNVs) and splice-site variants. This expanded genetic landscape underscores the importance of utilizing advanced sequencing technologies like WES to uncover rare variants that may be overlooked by conventional screening methods.

In this study, we employed a systematic approach to establish a comprehensive understanding of Duchenne muscular dystrophy (DMD) by correlating molecular data with clinical outcomes. Our process involved a stepwise analysis, beginning with protein interaction networks, followed by gene ontology (GO) enrichment analysis, linking molecular disruptions to clinical phenotypes, and finally examining genotypic variations that drive these disturbances. Based on these insights, targeted therapeutic interventions were proposed to address the molecular and phenotypic consequences of DMD.

The protein-protein interaction (PPI) analysis using the STRING database highlighted critical dystrophin-associated partners, including DAG1, SGCD, SSPN, UTRN, and ACTG1, essential for muscle integrity. GO enrichment analysis further identified significant disruptions in pathways related to glycogen metabolism, calcium transport, mitochondrial function, and sarcoplasmic reticulum activity. Disruptions in dystrophin interactions were associated with impairments in calcium transport (24 %), glycogen synthesis (46 %), and mitochondrial function (12 %), all contributing to muscle degeneration. Utrophin (UTRN) emerged as a compensatory mechanism, while sarcoglycan disruptions (SGCD, SGCE) compromised membrane integrity, accelerating disease progression.

Mapping these molecular disruptions to clinical manifestations provided valuable insights. Progressive muscle weakness and early loss of ambulation were linked to compromised muscle protein and actin cytoskeleton integrity. Mitochondrial dysfunction contributed to fatigue and exercise intolerance due to impaired ATP production. Muscle cramps and stiffness were attributed to calcium transport abnormalities affecting contraction-relaxation cycles. Cardiac complications resulted from SGCD dysfunction, leading to progressive cardiomyopathy. Cognitive impairment was associated with disruptions in dystrophin isoforms in the central nervous system, affecting synaptic function and cognitive processing.

The genotype-phenotype correlation analysis (Table 2) revealed distinct molecular disruptions linked to specific mutation types, offering valuable insights for tailored therapeutic interventions. Out-of-frame deletions were associated with severe protein interaction loss and calcium dysregulation, leading to early onset and rapid disease progression. Exon-skipping therapies, such as eteplirsen targeting exon 51, restore the reading frame and produce a truncated but functional dystrophin protein.In-frame deletions preserved partial dystrophin functionality, resulting in milder phenotypes with slower progression. Supportive therapies, including physical rehabilitation and targeted metabolic interventions, are crucial in managing disease progression and improving quality of life.

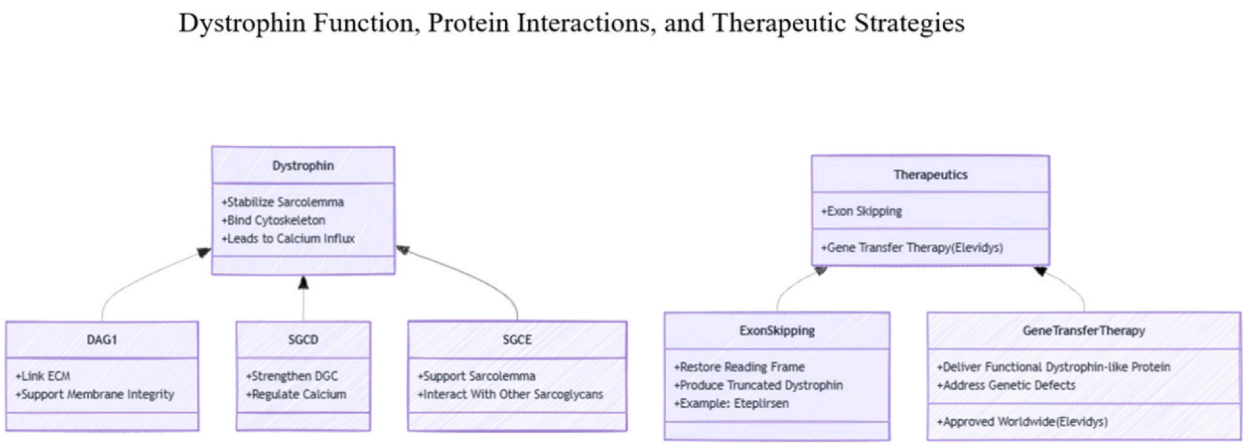

Fig. 5. Dystrophin Function, Protein Interactions, and Therapeutic Strategies.

Duplications caused transcript instability, exhibiting variable severity often accompanied by cardiac involvement. Gene therapy approaches, including Elevidys, aim to introduce functional dystrophin-like proteins to mitigate disease progression.Splice site variants caused aberrant splicing, leading to severe phenotypes with cognitive involvement. RNA-based therapies are being explored to restore proper dystrophin production. Nonsense variant resulted in premature dystrophin truncation and oxidative stress, contributing to early cardiomyopathy and severe muscle weakness. Nonsense readthrough therapies like Ataluren facilitate ribosomal bypass of premature stop codons, allowing dystrophin production [22].

In addition to mutation-specific interventions, broader therapeutic strategies are essential. Utrophin upregulation with agents such as ezutromid enhances cytoskeletal stability, compensating for dystrophin deficiency. Calcium-handling therapies, including L-type calcium channel blockers and SERCA activators, help manage muscle calcium homeostasis. Metabolic support interventions, such as Coenzyme Q10 and creatine supplements, target mitochondrial dysfunction and glycogen metabolism deficits to improve muscle endurance and reduce fatigue.

Personalized treatment strategies based on genetic profiles maximize therapeutic efficacy and patient outcomes. Future research directions include the application of CRISPR-based gene editing to achieve precise genetic corrections. This study underscores the importance of integrating molecular diagnostics with clinical management to optimize treatment approaches and advance research efforts in DMD. Fig. 5 outlines dystrophin's role in sarcolemma stabilization and interactions with key proteins, alongside potential therapeutic strategies such as exon-skipping and gene transfer therapy.

Sekar et al.'s [21] study provided valuable insights through transcript-level analysis and a structured classification system, offering a focused diagnostic framework. Our study expands on this by integrating functional enrichment analyses and protein interaction networks, linking genetic variants to molecular disruptions associated with key clinical symptoms such as muscle weakness, fatigue, and cardiac complications. Additionally, our study extends beyond diagnosis by proposing targeted therapeutic strategies, including exon-skipping therapies, metabolic support, and calcium regulation, enhancing clinical applicability. While Sekar et al. [21] provide a solid genetic framework, our study offers a broader perspective, combining genetic, molecular, and therapeutic insights for improved patient management.

Despite the comprehensive nature of this study, few limitations should be acknowledged. The use of WES, while comprehensive, may not capture certain non-coding and deep intronic variants that could contribute to disease pathogenesis. The sample size, though informative, may limit the broader applicability of the findings. The absence of functional validation for identified variants presents another area for further exploration to strengthen genotype-phenotype correlations. Future research with larger cohorts, additional sequencing methods such as RNA sequencing, and experimental validation could provide deeper insights into the complexities of dystrophinopathies.

In conclusion, this study demonstrates that the combined application of MLPA and WES provides a robust diagnostic approach for Duchenne muscular dystrophy (DMD), significantly improving the detection of both large deletions and rare variants. The identification of novel variants expands the genetic spectrum of DMD in the Indian population, emphasizing the need for region-specific screening strategies to facilitate early and accurate diagnosis.

By reinforcing the crucial role of genotype-phenotype correlations, these findings enhance clinical decision-making and prognostication, paving the way for personalized treatment approaches. Tailored therapeutic interventions, including exon-skipping therapies and emerging gene-editing techniques, can now be better aligned with individual genetic profiles.

Future research should focus on incorporating advanced technologies such as RNA sequencing and CRISPR-based gene editing to uncover regulatory variants and further optimize treatment strategies. Ultimately, this study underscores the value of integrating genetic profiling into routine clinical practice, contributing to improved patient management and long-term outcomes in DMD.

Ethical clearance

Obtained from the Office of the Ethics Committee, S.M.S. Medical College and Attached Hospitals, Jaipur, dated 18th November 2021 and number- 1025MC/EC/2021.

Declaration of Competing Interest

None.

Acknowledgement

This study utilized the DAVID Bioinformatics Resource for GO enrichment analysis. The authors acknowledge the developers and maintainers of DAVID and have cited the relevant publications as per guidelines (Sherman et al., 2022; Huang et al., 2009). Protein-protein interactions were analyzed using the STRING database (v11.5), a resource for known and predicted protein interactions, as described by Szklarczyk et al. (2021).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.gmg.2025.100038.

References

[1] S. Crisafulli, J. Sultana, A. Fontana, F. Salvo, S. Messina, et al., Global epidemiology of Duchenne muscular dystrophy: an updated systematic review and meta-analysis, Orphanet J. Rare Dis. 15 (1) (2020) 141, https://doi.org/10.1186/s13023-020-01430-8.
[2] A.L. Beaudet, A.B. Boggs, A.C. Chinault, C.M. Eng, D. Gaudio, et al., Molecular diagnosis of Duchenne/Becker muscular dystrophy: Enhanced detection of dystrophin gene rearrangements by oligonucleotide array-comparative genomic hybridization, Hum. Mutat. 29 (9) (2008) 1100–1107.
[3] E.P. Hoffman, R.H. Brown, L.M. Kunkel, Dystrophin: the protein product of the Duchenne muscular dystrophy locus, Cell 51 (6) (1987) 919–928.
[4] S. Kohli, R. Saxena, E. Thomas, J. Singh, I.C. Verma, Gene changes in Duchenne muscular dystrophy: comparison of multiplex PCR and multiplex ligation-dependent probe amplification techniques, Neurol. India 58 (6) (2010) 852–856, https://doi.org/10.4103/0028-3886.73744.
[5] A. Aartman-Rus, I. Fokkema, J. Verschuuren, I. Ginjaar, J. van Deutekom, G.J. van Ommen, et al., Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy pathogenic variants, Hum. Mutat. 30 (2) (2009) 293–299.
[6] D.G. Allen, N.P. Whitehead, S.C. Froehner, Absence of dystrophin disrupts skeletal muscle signaling: roles of $Ca2+$, reactive oxygen species, and nitric oxide in the development of muscular dystrophy, Physiol. Rev. 96 (1) (2016) 253–305, https://doi.org/10.1152/physrev.00007.2015.
[7] K. Tallapaka, P. Ranganath, A. Ramachandran, M.S. Uppin, S. Perala, S. Aggarwal, et al., Molecular and histopathological characterization of patients presenting with the Duchenne muscular dystrophy phenotype in a tertiary care center in southern India, Indian Pedia 56 (7) (2019) 556–559.
[8] M. Michalak, M. Opas, Duchenne Muscular Dystrophy, John Wiley & Sons Ltd,, 2001.
[9] C.F. Wright, D.R. FitzPatrick, H.V. Firth, Paediatric genomics: diagnosing rare disease in children, Nat. Rev. Genet 19 (1) (2018) 25–40.
[10] K. Zhang, X. Yang, G. Lin, Y. Han, J. Li, Molecular genetic testing and diagnosis strategies for dystrophinopathies in the era of next-generation sequencing, Clin. Chim. Acta 491 (2019) 66–73.
[11] C. Happi Mbakam, G. Lamothe, J.P. Tremblay, Therapeutic strategies for dystrophin replacement in Duchenne muscular dystrophy, Front Med 9 (2022) 1–19.
[12] C. Pisani, G. Strimpakos, F. Gabanella, M.G. Di Certo, A. Onori, C. Severini, et al., Utrophin up-regulation by artificial transcription factors induces muscle rescue and impacts the neuromuscular junction in mdx mice, Biochim. Biophys. Acta Mol. Basis Dis. 1864 (2018) 1172–1182.
[13] D.R. Rodríguez-Rodríguez, R. Ramírez-Solís, M.A. Garza-Elizondo, M. De Lourdes Garza-Rodríguez, H.A. Barrera-Saldaña, Genome editing: a perspective on the application of CRISPR/Cas9 to study human diseases, Int. J. Mol. Med. 43 (4) (2019) 1559–1574.
[14] D. Szklarczyk, A.L. Gable, K.C. Nastou, D. Lyon, R. Kirsch, S. Pyysalo, et al., The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets, Nucleic Acids Res. 49 (D1) (2021) D605–D612.
[15] B.T. Sherman, M. Hao, J. Qiu, X. Jiao, M.W. Baseler, H.C. Lane, et al., DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update), Nucleic Acids Res (2022).
[16] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID Bioinformatics Resources, Nat. Protoc. 4 (1) (2009) 44–57.
[17] M. Goyal, A. Gupta, K. Agarwal, S. Kapoor, S. Kumar, Duchenne muscular dystrophy: genetic and clinical profile in the population of Rajasthan, India, Ann. Indian Acad. Neurol. 24 (3) (2021) 255–261.
[18] L.C. Werneck, R.H. Scola, G.H.B. Maegawa, M.C.M. Werneck, Comparative analysis of PCR-deletion detection and immunohistochemistry in Brazilian Duchenne and Becker muscular dystrophy patients, Am. J. Med Genet 103 (2) (2001) 115–120.
[19] M. Orso, A. Migliore, B. Polistena, E. Russo, F. Gatto, M. Monterubbianesi, et al., Duchenne muscular dystrophy in Italy: A systematic review of epidemiology, quality of life, treatment adherence, and economic impact, PLoS One 18 (6) (2023) e0287774.
[20] J. Juan-Mateu, L. González-Quereda, M.J. Rodríguez, M. Baena, E. Verdura, A. Nascimento, et al., DMD variant in 576 dystrophinopathy families: A step forward in genotype-phenotype correlations, PLoS One 10 (8) (2015) e0135189, https://doi.org/10.1371/journal.pone.0135189.
[21] D. Sekar, S. Vengalil, V. Preethish-Kumar, K. Polavarapu, A. Nalini, N. Gayathri, et al., MLPA identification of dystrophin variant and in silico evaluation of the predicted protein in dystrophinopathy cases from India, BMC Med. Genet. 18 (2017) 67.
[22] M. Falzarano, C. Scotton, C. Passarelli, A. Ferlini, Duchenne muscular dystrophy: From diagnosis to therapy, Molecules 20 (10) (2015) 18168–18184.