ARTICLE

Check for updates

Phenotype-driven genomics enhance diagnosis in children with unresolved neuromuscular diseases

Berta Estévez-Arias[1,2], Leslie Matalonga[3,4], Delia Yubero[5,6], Kiran Polavarapu[7], Anna Codina[8,9], Carlos Ortez[1,5,8], Laura Carrera-García[1,8], Jesica Expósito-Escudero[1,8], Cristina Jou[4,8,9], Stefanie Meyer[7,10], Ozge Aksel Kilicarslan[7], Alberto Aleman[7,11], Rachel Thompson[7], Rebeka Luknárová[12], Anna Esteve-Codina[3,4], Marta Gut[3,4], Steven Laurie[3,4], German Demidov[13], Vicente A. Yépez[12], Sergi Beltran[3,4], Julien Gagneur[12,14,15], Ana Topf[16], Hanns Lochmüller[3,7,11], Andres Nascimento[1,5,8], Janet Hoenicka[2,5], Francesc Palau[2,5,6,17,18]✉ and Daniel Natera-de Benito[1,5,8]✉

Establishing a molecular diagnosis remains challenging in half of individuals with childhood-onset neuromuscular diseases (NMDs) despite exome sequencing. This study evaluates the diagnostic utility of combining genomic approaches in undiagnosed NMD patients. We performed deep phenotyping of 58 individuals with unsolved childhood-onset NMDs that have previously undergone inconclusive exome studies. Genomic approaches included trio genome sequencing and RNASeq. Genetic diagnoses were reached in 23 out of 58 individuals (40%). Twenty-one individuals carried causal single nucleotide variants (SNVs) or small insertions and deletions, while 2 carried pathogenic structural variants (SVs). Genomic sequencing identified pathogenic variants in coding regions or at the splice site in 17 out of 21 resolved cases, while RNA sequencing was additionally required for the diagnosis of 4 cases. Reasons for previous diagnostic failures included low coverage in exonic regions harboring the second pathogenic variant and involvement of genes that were not yet linked to human diseases at the time of the first NGS analysis. In summary, our systematic genetic analysis, integrating deep phenotyping, trio genome sequencing and RNASeq, proved effective in diagnosing unsolved childhood-onset NMDs. This approach holds promise for similar cohorts, offering potential improvements in diagnostic rates and clinical management of individuals with NMDs.

European Journal of Human Genetics; https://doi.org/10.1038/s41431-024-01699-4

INTRODUCTION

Neuromuscular diseases (NMDs) in children constitute a diverse spectrum of conditions, including motoneuron diseases, neuropathies, myasthenias, and myopathies, all marked by muscle wasting and/or weakness, with the vast majority having a genetic basis. Despite their individual rarity, the collective global prevalence of these diseases aggregates to approximately 100–200 cases per 100,000 individuals.

Incorporating next-generation sequencing (NGS) into clinical practice has revolutionized the genetic diagnosis of inherited NMDs [1], with NGS genetic analysis now assuming a leading role over neurophysiological and histopathological tests as the primary diagnostic tool, marking a significant evolution in the diagnostic landscape [2].

NGS techniques, particularly custom panels, clinical exome sequencing (CES) or whole exome sequencing (WES), have become the primary methods for diagnosing rare diseases, especially for conditions with diverse genetic origins but similar clinical presentations. In the field of NMDs, these include hereditary neuropathies, congenital myasthenic syndromes, congenital muscular dystrophies, congenital myopathies, and arthrogryposis multiplex congenita [3–5]. Over the past 14 years, NGS has enhanced diagnostic rates and expanded our knowledge of genes associated with NMDs from 290 in 2010 to 686 in 2024 [6].

Nevertheless, even with optimal use of exome sequencing and traditional diagnostic methodologies, establishing a molecular diagnosis remains challenging in approximately half of individuals with NMDs [2, 7–10]. A combined approach, including genome

[1]Neuromuscular Unit, Department of Neurology, Hospital Sant Joan de Déu, Barcelona, Spain. [2]Laboratory of Neurogenetics and Molecular Medicine – IPER, Institut de Recerca Sant Joan de Déu, Barcelona, Spain. [3]Centro Nacional de Análisis Genómico (CNAG), Barcelona, Spain. [4]Universitat de Barcelona (UB), Barcelona, Spain. [5]Center for Biomedical Research Network on Rare Diseases (CIBERER), ISCIII, Barcelona, Spain. [6]Department of Genetic and Molecular Medicine – IPER, Hospital Sant Joan de Déu and Institut de Recerca Sant Joan de Déu, Barcelona, Spain. [7]Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada. [8]Applied Research in Neuromuscular Diseases, Institut de Recerca Sant Joan de Déu, Barcelona, Spain. [9]Department of Pathology, Hospital Sant Joan de Déu, Barcelona, Spain. [10]University Medical Center Göttingen, Department of Neurology, Göttingen, Germany. [11]Division of Neurology, Department of Medicine, The Ottawa Hospital, Ottawa, ON, Canada. [12]School of Computation, Information and Technology, Technical University of Munich, Munich, Germany. [13]Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany. [14]Computational Health Center, Helmholtz Center Munich, Neuherberg, Germany. [15]Institute of Human Genetics, School of Medicine, Technical University of Munich, Munich, Germany. [16]The John Walton Muscular Dystrophy Research Centre, Translational and Clinical Research Institute, Newcastle University and Newcastle Hospitals NHS Foundation Trust, Newcastle Upon Tyne, UK. [17]ERN ITHACA, Barcelona, Spain. [18]Division of Pediatrics, Faculty of Medicine and Health Sciences, Universitat de Barcelona (UB), Barcelona, Spain. ✉email: francesc.palau@sjd.es; daniel.natera@sjd.es

Received: 5 June 2024 Revised: 23 July 2024 Accepted: 19 September 2024
Published online: 27 September 2024

sequencing (GS) [11] and RNASeq has emerged to address this diagnostic gap and increase diagnostic rates. Additionally, the trio-based study methodology, which involves examining the individual's genome alongside those of both parents, shows promise in enhancing diagnostic efficacy [12]. Moreover, integrating deep phenotyping together with genomic approaches may serve as a critical aid for identifying causative variants [13, 14]. Use of collaborative platforms to facilitate data sharing has been shown to accelerate diagnostic advancements [15, 16].

Here, we show how the combination of standardized deep phenotyping with genomic approaches elevated the diagnostic rate for individuals with unsolved childhood-onset NMDs who had previously undergone inconclusive exome studies.

METHODS

We analyzed 58 individuals within the framework of the SolveRD Project, a program funded by the European Commission that employs, among other initiatives, a combined approach of genome sequencing and RNA sequencing for individuals with rare genetic diseases who have not received a genetic diagnosis despite prior exome sequencing [17]. Both the Neuromuscular Unit of Hospital Sant Joan de Déu, Barcelona and the Centro Nacional de Análisis Genómico (CNAG), Barcelona, are partners in the Solve-RD project (http://solve-rd.eu/).

Reporting of findings to affected individuals and their families followed the procedures of Hospital Sant Joan de Déu, Barcelona, including the production of a genetic report and appropriate genetic counseling (PIC-98-20).

Selection of individuals and phenotype data collection

We selected individuals with genetically unresolved NMDs who were followed at the Neuromuscular Unit of Hospital Sant Joan de Déu. Inclusion criteria comprised (1) being affected by a neuromuscular disease of likely genetic origin and (2) having undergone prior clinical or exome sequencing (CES or WES), without a genetic diagnosis being reached. Clinical exome sequencing had been performed using either theTruSight One Sequencing Panel (Illumina) or a Custom Comprehensive panel 17 Mb (Agilent), and whole exome sequencing using Nextera Flex for Enrichment (Illumina). The initial CES/WES had an average coverage of >100x and a minimum coverage of ≥95% in the target regions.

We performed deep phenotyping for all affected individuals, which involved comprehensive clinical examination and the collection of histopathological, neurophysiological, laboratory, and MRI muscle imaging data from electronic medical records. Phenotypic features of participants were pseudonymized and captured using standardized Human Phenotype Ontology (HPO) terms [18] in RD-Connect PhenoStore [15]. Clinical and molecular diagnoses were registered using Orphanet Rare Disease Ontology (https://www.orpha.net/consor/cgi-bin/index.php) and Online Mendelian Inheritance in Man codes (https://www.omim.org). The quality of the phenotypic data was assessed using the Monarch star rating system, which is integrated within RD-Connect PhenoStore. This system evaluates annotation sufficiency and assigns a rating ranging from bad to excellent [19].

When we refer to conducting deep phenotyping, we encompass not only the collection of the signs and symptoms of the affected individuals but also the results of complementary tests and the clinical insights derived from the detailed evaluation of these clinical symptoms and all relevant diagnostic tools from a neuromuscular perspective. These complementary tests primarily include electromyogram/electroneurogram, muscle MRI, muscle biopsy, and serum creatine kinase levels, though other tests may also have been performed depending on the case. Individuals were clinically categorized into five disease categories based on this comprehensive assessment.

Genome sequencing, data processing and variant detection

Blood samples for DNA analyses were collected from 58 families and processed using standard procedures. GS of the index case, parents and affected siblings was performed using a BGI DNBSEQ-G400 Platform. Libraries were generated using an Illumina genome capture kit and sequenced with a mean target coverage of 30x across the genome. Genomic and phenotypic data were submitted to the RD-Connect GPAP, where they can be accessed through controlled access agreements.

All samples underwent bioinformatic processing at the CNAG using the RD-Connect standard analysis pipeline [20]. This pipeline, based on GATK best practices [21], involved mapping sequencing reads to the GRCh37d5 version of the human genome, removing duplicate reads using Picard version 1.110 (http://broadinstitute.github.io/picard), insertion and deletion realignment, and recalibrating base quality scores using GATK version 3.6.

Single-nucleotide variants (SNVs), short insertions and short deletions were identified using GATK HaplotypeCaller version 3.6 [21]. SNVs, insertions and deletions with a minimum depth of coverage of 8 and a minimum genotype quality of 20 were uploaded to the RD-Connect GPAP. Structural variants (SVs) and copy number variants (CNVs) were identified using ClinCNV [22], ClinSV [23] and Manta [24].

Filtering and interpretation of single nucleotide variants and short insertions and deletions

SNV and short insertion and deletion analysis was performed using the RD-Connect GPAP. This platform integrates filters, annotations, and effect prediction tools and guides the user in performing efficient variant filtering by selecting different parameters to apply. Briefly, we selected the trio to analyze and we used a multi-step filtering strategy to prioritize candidate variants, considering criteria such as allele frequency in population databases (GnomAD, 1000 Genomes, and the RD-Connect internal database; using minor allele frequency <0.02 for autosomal recessive or X-linked in men, and minor allele frequency <0.01 for autosomal dominant or X-linked in women), functional impact predictions, and trio segregation information.

To explore candidate genes, we first focused on the analysis of variants in genes associated with NMDs using the Muscle Gene Table [25], and subsequently we analyzed all variants without gene restriction (Fig. 1). At this stage, each candidate variant was analyzed by experts in genetics and neuromuscular phenotyping. These analysts considered the phenotype of each analyzed individual and how compatible it was with the phenotype previously described for the candidate gene. This process integrated the concept of deep phenotyping, which takes into account not only the signs and symptoms of the affected individuals but also a comprehensive evaluation of all complementary tests. This thorough approach allowed analysts to start with a hypothesis about the neuromuscular disease each individual had (neuropathy, myasthenia, myopathy, muscular dystrophy, etc.), while also providing enough flexibility to avoid prematurely dismissing variants in genes that do not initially fitted the deep phenotyping-based hypothesis. For instance, in a case of distal weakness with normal creatine kinase (CK) levels and an EMG/ENG interpreted as neurogenic, in whom a variant associated with distal myopathy is identified, the analyst understands that this variant should not be dismissed due to the potential for misinterpretation of the EMG/ENG findings, which analysts familiar with neuromuscular diseases know can often be challenging to interpret. Additionally, to avoid dismissing variants that might lead to previously undescribed phenotypes, the function of the protein encoded by the candidate gene was also consulted, as well as the expression of that protein in peripheral nerve and muscle. In some cases, lists of candidate genes were created according to the associated HPO terms or OMIM codes, as well as suspected affected pathways. The use of standardized terms enhanced phenotype annotation and comparison across cases. In cases with suspected consanguinity, the run of homozygosity (RoH) filter was used to narrow down the list of candidate variants to only those within RoH of at least 500 Kb in length [26].

Variants meeting these criteria were classified following the American College of Medical Genetics and Genomics guidelines (ACMG) [27]. They underwent further pathogenicity assessment through the application of various software tools: CADD (https://cadd.gs.washington.edu/), Mutation Taster (https://www.mutationtaster.org/), FATHMM-MKL (http://fathmm.biocompute.org.uk/), and PROVEAN (https://www.jcvi.org/research/provean). The MetaDome web server (https://stuart.radboudumc.nl/metadome/) facilitated the evaluation of protein tolerance to variations. All variants were evaluated for splicing alterations using: SpliceAI (https://spliceailookup.broadinstitute.org/), SPiP [28] and MaxEntScan (http://hollywood.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html). Filtering settings were saved and applied recurrently for further analyses to speed up the prioritization process.

Variants were prioritized considering their type, suspected impact on protein, and in silico pathogenicity predictor scores, as well as previous evidence reported in the literature, if available. Genotype-phenotype correlations were analyzed based on existing data from databases such as Online Mendelian Inheritance in Man (OMIM), extensive literature searches for genes not yet associated with disease, and continuous dialogue between geneticists and clinicians. Establishing a multidisciplinary team

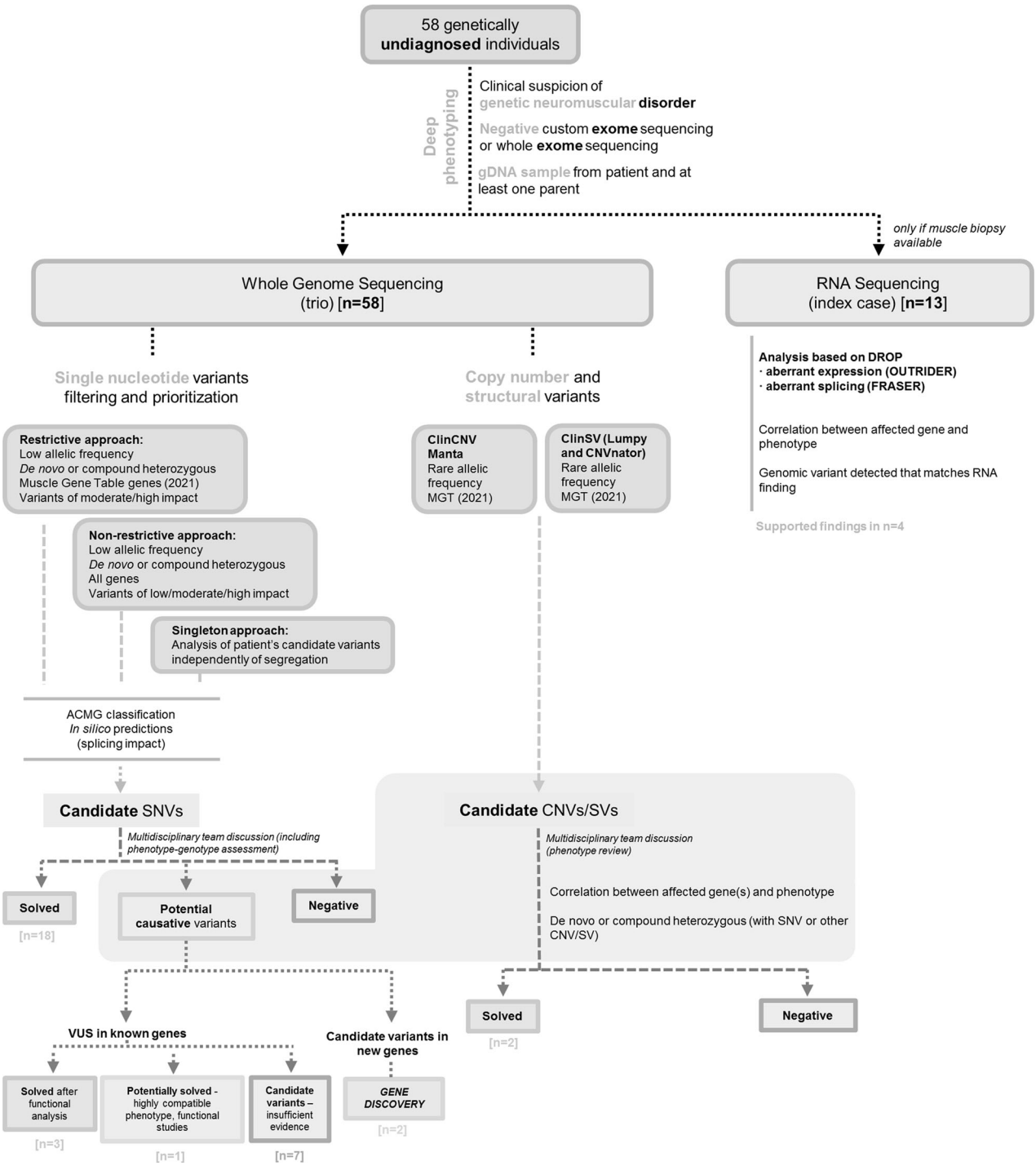

Fig. 1 Workflow used for the study of 58 individuals with childhood-onset neuromuscular diseases. Note that the schema includes both the genome analysis conducted for all 58 individuals and the RNA analysis performed in the 13 cases where muscle biopsy was available. The various analysis strategies, with different applied filters, are summarized.

specialized in neuromuscular disorders facilitated ongoing discussion of findings, enhancing variant interpretation and prioritization. The precision phenotyping previously conducted for each affected individual was continuously revisited during this prioritization process, assessing the compatibility of each variant with the phenotype of each affected individual (Fig. 1).

Functional studies were applied to facilitate the reclassification of variants, both in cases with variants of uncertain significance (VUS) within genes with a compatible reported phenotype and in genes not previously associated with any human condition but presenting a pathophysiological mechanism that hints at potential causality (Fig. 1).

Structural variants and copy number variants filtering

The analysis was performed independently using two different approaches. The first included ClinCNV as a CNV caller and Manta as an SV caller. The second used an ensemble of methods (Lumpy and CNVnator), implemented as the tool ClinSV (Fig. 1).

ClinCNV and Manta results were annotated with the allele frequencies from the whole Solve-RD GS cohort. A distribution plot of the median number of singleton CNVs and SVs per chromosome per sample was made and samples with a visually excessive number of singleton CNVs/SVs (namely, more than 30 SVs or 3 CNVs) were filtered out from the corresponding analysis results. 2954 CNV calling results and 3091 SV calling

results remained out of 3126 tested GS samples. Rare variants (observed in less than 0.002 samples per Solve-RD cohort) were intersected with the Muscle Gene Table candidate gene list and annotated with various annotations such as diseases associated with particular genes and their known modes of inheritance, de novo status, etc.

RNASeq data processing and analysis

RNA was extracted from muscle tissue in 13 affected individuals, who had undergone prior muscle biopsies for diagnostic purposes (individuals 4, 5, 26, 46, 53, 54, 55, 56, 57, 58, 59, 60, and 61). The RNeasy Mini Kit (Qiagen) was used for extractions and RNA samples were quality controlled using the Qubit RNA BR Assay kit (Thermo Fisher Scientific) to assess quantity, and the Agilent Fragment Analyzer DNF-471 RNA Kit (15nt) to evaluate integrity.

RNA-Seq libraries were generated from total RNA using the TruSeq Stranded mRNA Library Prep Kit (Illumina). In brief, mRNA was enriched with oligo-dT magnetic beads from 500 ng of total RNA. The resulting blunt-ended double-stranded cDNA was 3′ adenylated, and Illumina platform-compatible adaptors with unique dual indices and unique molecular identifiers (Integrated DNA Technologies) were ligated. The ligation product underwent enrichment with 15 PCR cycles. The final library was validated using an Agilent Bioanalyzer DNA 7500 assay.

Libraries were sequenced on a NovaSeq 6000 (Illumina) in paired-end mode with a read length of 2× 151 bp, following the manufacturer's protocol for dual indexing. Image analysis, base calling, and quality scoring were processed using the manufacturer's software Real Time Analysis (RTA 3.4.4), followed by the generation of FASTQ sequence files.

RNA-seq reads were trimmed with trim_galore version 0.6.7 [29] and mapped against the GRCh37 human reference with STAR/2.7.8a [30] using the - -twopassMode = BASIC parameter. Qualimap [31], featureCounts [32], STAR log files, and custom scripts calculated the mapping quality metrics. Further quality control of RNA-seq samples was performed via DROP v1.3.3 [33], which evaluates sex mislabeling and RNA-DNA mismatches.

Detection of aberrant expression was fully based on DROP v1.3.3 [33]. We used the GRCh37 primary assembly (GENCODE release 34) [34] as a reference genome. Read pairs that fell completely within exonic regions were counted using the summarizeOverlaps function from the GenomicAlignments [35] R package. Reads that fully overlapped more than one feature were assigned to each feature. Genes with a 95th percentile FPKM < 1 were filtered out. Expression outliers were found using OUTRIDER [36]. Significant events were defined as those with an FDR ≤ 0.1 computed either i) across all expressed genes or ii) in genes known to cause NMDs. All aberrant events were further inspected using the Integrative Genome Viewer (https://igv.org/).

Detection of aberrant splicing was also entirely based on DROP v1.3.3, which uses FRASER 2.0 [37], an annotation-free aberrant splicing detection algorithm. Exon-exon and exon-intron junctions with <20 reads in all samples and for which the total number of reads at the donor and acceptor splice site was 0 in more than 75% of the samples were filtered out. Splicing outlier genes were defined as those with Holm's adjusted *p*-value across junctions of the tested gene ≤0.1. Outlier junctions are defined as those in splicing outlier genes, with an FDR ≤ 0.1 (also computed on all genes and on genes known to cause NMDs) and effect size (absolute delta intronic Jaccard index) larger than 0.1.

To increase statistical power in both aberrant expression and splicing detection, other RNA muscle samples different than those from the Solve-RD cohort, but processed in the same manner, were added (N = 92).

RESULTS

Demographics

Our cohort of individuals with unsolved neuromuscular diseases totaled 58 index cases (29 males and 29 females) from 58 unrelated families. In two families, other affected relatives were included in addition to the index case: an affected sibling in family 3, and one sister and two affected cousins in family 7.

Forty-two of the 58 individuals were Spanish: 37 of them were of Caucasian origin and five, were of Roma origin. Seven individuals were Moroccan, two were Ecuadorian, two were Romanian, and one each of Chinese, Venezuelan, Colombian, Gambian, and Emirati origin. Fourteen individuals were born from consanguineous couples. All individuals experienced disease onset within the first 5 years of life, with the majority occurring during either the antenatal or neonatal period (34/58, 59%). Three individuals (3/58, 5%) died prior to this study at the ages of one month, 12 years, and 16 years. The mean age at the last evaluation was 12.63 years (SD: 7.07), ranging from the neonatal period to 32 years.

Classification of individuals according to phenotype

All 58 individuals had previously undergone inconclusive CES or WES. Individuals were broadly categorized into five disease categories according to the main clinical features: 23 individuals had a myopathy, 12 individuals had arthrogryposis multiplex congenita, 8 individuals had an axonal neuropathy, 5 individuals had a muscular dystrophy and 4 individuals had a congenital myasthenic syndrome. Six individuals were categorized in the group "other diseases": three had a suspicion of mitochondrial disease, two had a delayed acquisition of motor milestones, and one had a complex paraparesis (Fig. 2).

The quality of phenotypic data, annotated with HPO terms, was evaluated by the Monarch star rating system, consistently meeting a minimum of 3.5 stars for all individuals.

Diagnostic yield of the standardized deep phenotyping and trio genome sequencing approach

A genetic diagnosis was successfully established in 23 out of the 58 index individuals (40%) (Table 1 and Fig. 3). In 21 cases, SNVs or small insertions or deletions were identified as the cause, while SVs and CNVs were responsible in 2 cases (Table 1 and Fig. 1).

SNV and short insertion and deletion analysis pinpointed causal coding or essential splice-site pathogenic variants in 18 out of the 23 solved (78%). Three diagnoses (13%) necessitated further functional studies, such as RNASeq or experimental protein expression analysis (individuals 53, 54 and 58 with variants in *NEB*, *LAMA2* and *DPM2*; Table 1). Additionally, among the 35 individuals still classified as unsolved, we identified one variant of uncertain significance (VUS) that is likely responsible for the observed phenotype in one case, but requires further in silico and functional studies for confirmation of causality (individual 40 with one de novo missense variant in *DOT1L*; Table 2). This could potentially increase the overall diagnostic rate to 41%. Furthermore, we identified 13 potentially causative variants in 7 unsolved individuals meeting the following criteria: rare variants (allelic

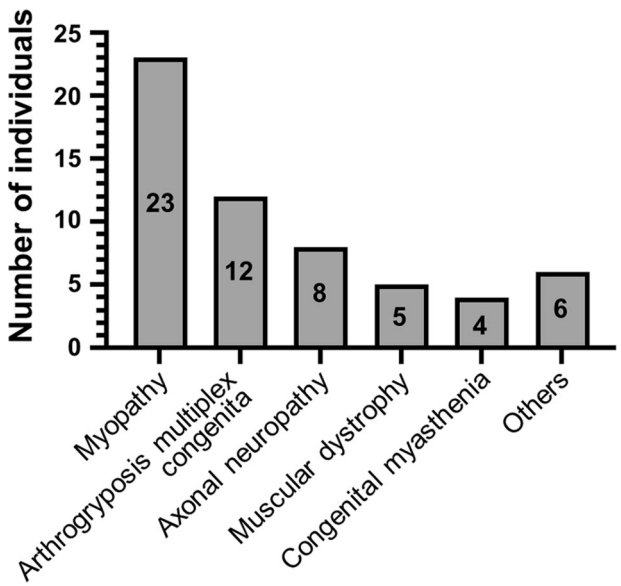

Fig. 2 Classification of individuals according to phenotype. Note that in the "others" group, three individuals with clinical features suggestive of mitochondrial disease are included, two with delayed acquisition of motor milestones, and one with complex paraparesis.

Table 1. Causative variants found in the 23 individuals that were considered solved.

Solved cases				
Individual	Gene (inheritance)	Ref transcript	Causative variant(s)	ACMG classification
Congenital myopathy				
17	***TTN*** (AR)	NM_001267550.2	c.38737G>T (p.Glu12913Ter)	Pathogenic (PS4, PVS1, PM2)
			c.32104del (p.Ser10702LeufsTer6)	Pathogenic (PVS1, PM2, PM3)
25	***TTN*** (AR)	NM_001267550.2	c.67495C>T (p.Arg22499Ter)	Pathogenic (PS4, PVS1, PM2, PM3)
			c.38737G>T (p.Glu12913Ter)	Pathogenic (PS4, PVS1, PM2, PM3)
27	***TTN*** (AR)	NM_001267550.2	c.103531A>T (p.Lys34511Ter)	Pathogenic (PVS1, PM2, PM3)
			c.38661_38665del (p.Lys12887AsnfsTer6)	Pathogenic (PM3, PVS1, PM2)
48	***TTN*** (AR)	NM_001267550.2	c.50248+1G>C	Pathogenic (PVS1, PM2, PM3)
			c.35876-2A>G	Pathogenic (PS4, PVS1, PM2)
7	***TRIP4*** (AR)	NM_016213.5	c.1611dup (p.Val538CysfsTer17)	Likely pathogenic (PVS1, PM2)
			c.1611dup (p.Val538CysfsTer17)	
20	***TRIP4*** (AR)	NM_016213.5	c.1611dup (p.Val538CysfsTer17)	Likely pathogenic (PVS1, PM2)
			c.1611dup (p.Val538CysfsTer17)	
2	***GLDN*** (AR)	NM_181789.4	c.319_325del (p.Glu107MetfsTer4)	Likely pathogenic (PVS1, PM2)
			c.1435C>T (p.Arg479Ter)	Likely pathogenic (PVS1, PM2, PP5)
53	***NEB*** (AR)	NM_001164508.2	c.2784del (p.Asp929IlefsTer28)	Pathogenic (PM3, PVS1, PM2)
			c.7432-1121G>A	Likely pathogenic (PM2, BP7, PM3) **Required functional studies** +**PS3** 108bp pseudoexon created in intron 54 detected by RNASeq
19	***RYR1*** (AD)	NM_000540.3	c.14566G>A (p.Ala4856Thr)	Likely pathogenic (PM1, PP2, PM2, PP3)
59	***RYR3*** (AR)	NM_001036.6	c.3215G>A (p.Arg1072Gln)	Likely pathogenic (PM2, PP3, PM3)
			c.3215G>A (p.Arg1072Gln)	
22	***TPM3*** (AD)	NM_152263.4	c.93_113dup (p.Ala33_Gln39dup)	Likely pathogenic (PM2, PM4, PM6)
Muscular dystrophy				
58	***DPM2*** (AR)	NM_003863.4	c.188T>A (p.Leu63Gln)	Likely pathogenic (PM2, PP3, PM3) **Required functional studies** +**PS3** reduced dystroglycan in muscle biopsy
			c.188T>A (p.Leu63Gln)	
52	***HMGCR*** (AR)	NM_000859.3	c.1522_1524del (p.Ser508del)	Likely pathogenic (PM2, PM4, PP5, PM3)
			c.1522_1524del (p.Ser508del)	
54	***LAMA2*** (AR)	NM_000426.4	c.909+7A>G	Likely pathogenic (PM2, PP3, PM3) **Required functional studies** +**PS3** exon 6 skipping detected by RNASeq and reduced merosin in muscle biopsy
			c.909+7A> G	
55	***SGCG*** (AR)	NM_000231.3	c.579-1G>A	Pathogenic (PVS1, PM2, PP5) +PS3 exon 7 truncation detected by RNASeq
			c.579-1G>A	
Axonal neuropathy				
24	***DYNC1H1*** (AD)	NM_001376.5	c.1741A>T (p.Met581Leu)	Pathogenic (PS4, PM1, PP2, PM2, PP3)
29	***UBA1*** (AD)	NM_153280.3	c.1660C>T (p.Pro554Ser)	Likely pathogenic (PM3, PM2, BP4)
30	***PLEKHG5*** (AR)	–	Homozygous complex structural variant involving an inversion and a duplication in chr1:6524620-6529620	
Congenital myasthenia				
60	***CHRNB1*** (AR)	NM_000747.3	c.114T>A (p.Tyr38Ter)	Likely pathogenic (PVS1, PM2)
			c.26T>C (p.Leu9Pro)	Likely pathogenic (PM2, PM3, PP3, PP4)
1	***SLC18A3*** (AR)	NM_003055.3	c.4G>T (p.Glu2Ter)	Likely pathogenic (PVS1, PM2)
			c.923T>G (p.Leu308Arg)	Likely pathogenic (PM2, PP3, PM3)
Arthrogryposis multiplex congenita				
12	***BICD2*** (AD)	NM_001003800.2	c.614T>C (p.Phe205Ser)	Likely pathogenic (PM2, PP3, PP2, PM6)
8	***PIEZO2*** (AR)	NM_022068.4	c.2004del (p.Glu668AspfsTer5)	Pathogenic (PM3, PVS1, PM2)
			c.2004del (p.Glu668AspfsTer5)	
Others				
28	Not applicable	–	Heterozygous deletion in chr7:5669468-6368115 involving 12 genes	

Results are indicated according to the phenotype groups. Individuals for whom RNASeq was performed are highlighted in gray.
Those variants that required functional studies are in bold.

frequency (AF) < 0.01), with segregation patterns consistent with previous reports for the affected gene, with predictions of high or moderate protein impact, and with prediction of a damaging effect by CADD score. Variant selection was restricted to genes from the MGT 2021 [25]. However, their pathogenicity remained unclear after we performed further in silico analysis and, thus, they were not considered responsible for the clinical presentation of the affected individuals. Some of these variants are summarized in Supplementary Table 1, which includes a column correlating the phenotype of the affected individual with the phenotype described for the candidate genes. The pathogenicity of these variants, as well as a reanalysis of those cases without candidate variants, will be subject to periodic reevaluation.

SNVs responsible for the phenotype observed in the affected individuals were identified in 17 genes, all included in the latest version of the Muscle Gene Table [6]. Notably, four individuals in whom the existence of a single *TTN* heterozygous pathogenic variant was known before their inclusion in this study, were subsequently confirmed to have compound heterozygous pathogenic variants in the *TTN* gene. Furthermore, two unrelated individuals shared the same homozygous variant in *TRIP4*, while 15 other genes were associated with individual cases: *BICD2*, *CHRNB1*, *DPM2*, *DYNC1H1*, *GLDN*, *HMGCR*, *LAMA2*, *NEB*, *PIEZO2*, *RYR1*, *RYR3*, *SGCG*, *SLC18A3, TPM3*, and *UBA1* (Table 1 and Fig. 3B).

In addition, analysis of copy number or structural variants revealed a complex homozygous structural variant involving an inversion and a duplication of 5 kilobases in chr1:6524620–6529620 in *PLEKHG5* in individual 30 and a 698.6 kilobases deletion in chr7:5669468–6368115, involving *PMS2*, *CYTH3*, *RSPH10B*, *RNF216-IT1*, *EIF2AK1*, *OCM*, *RNF216*, *ANKRD61*, *ZNF815P*, *AIMP2*, *USP42*, and *CCZ1* in individual 28 (Table 1 and Fig. 3C).

In a total of 6 cases, causative variants were identified de novo. Of the remaining 17 cases, 16 showed inherited homozygous or compound heterozygous variants. In one case, as we have reported previously [38], we identified a maternal mosaicism that explained the presence of the same heterozygous *RYR1* variant in the affected individual and his asymptomatic mother, who had no signs of any neuromuscular disease in the comprehensive examination conducted at 44 years old, as well as a highly preserved muscle biopsy [38]. The causal variant had been identified in previous NGS studies, but was discarded after a segregation analysis that showed its presence in the asymptomatic mother. Conducting the analysis in trio allowed us to compare allelic frequencies and determine that the maternal pattern was consistent with mosaicism, confirming after validation that the heterozygous variant identified in the affected son was causative [38].

The highest diagnostic yields were observed within groups of individuals with muscle or neuromuscular junction disorders, with diagnostic rates of 80% (4/5) in muscular dystrophies, 50% (2/4) in congenital myasthenic syndromes, and 46% (11/24) in congenital myopathies (Fig. 3D). Diagnostic rates for subgroups of individuals with axonal neuropathy and arthrogryposis multiplex congenita were comparatively lower: 38% (3/8) and 17% (2/12), respectively. The 'others' subgroup had a diagnostic rate of 67% (4/6).

Reasons for increasing the diagnostic rate

To understand the reasons that prevented a genetic diagnosis being reached in the 23 individuals prior to this study, we reviewed their previous NGS studies (Fig. 3E). The main reason that explained the lack of a previous diagnosis was that at least one of the pathogenic variants was located in an exonic region with low coverage in the previous NGS study or that the analysis pipeline failed to detect the variant (7/23; 30%). Low coverage regions were defined as those with coverage below 10x. This issue was noted for the variants in *CHRNB1*, *TPM3, TTN* and *SGCG*. Notably, the *TTN* gene has a triplicated region (exons 173–199) where ensuring read depth and correct mapping of reads can be challenging. Both *TTN* and *TPM3* variants were located in poorly covered regions. Additionally, inaccuracies in the analysis pipeline led to overlooked causative variants in *SGCG* and *CHRNB1*. The second most prevalent reason was the involvement of recently discovered genes. At the time of the first NGS analysis, five genes had not been linked to any disease in humans: *GLDN, HMGCR*, *RYR3*, *SLC18A3*, and *TRIP4*. Thirdly, causative variants were found in genes not interrogated in the previous clinical exome studies (4/23; 18%) (*BICD2*, *DYNC1H1, DPM2* and *PIEZO2)*. Reassessing our cohort, we find that reanalyzing the initial NGS data could have identified 8 cases out of 58 (14%). However, the remaining 15 cases of the 23 solved cases (15/58, 26%) could not have been resolved solely through reanalysis. The cases that could have been solved with reanalysis were those attributable to (1) new gene-disease associations that emerged after the date of the first analysis, (2) known genes that were not included in the candidate list analyzed during the first analysis, provided that the gene in which the pathogenic variants were located had been sequenced, as well as (3) those individuals unsolved due to an error in the pipeline of the initial analysis.

Auxiliary functional studies were required to diagnose 3 individuals, reclassifying the variants from VUS to likely pathogenic (3/23; 13%). The combination of GS and RNASeq analysis identified the creation of a 108 bp pseudoexon in intron 54 of *NEB*, and the causative deep intronic variant c.7432–1121G>A in individual 53, who also had the variant *NEB:* c.2784del (p.Asp929IlefsTer28) in trans. RNASeq analysis also confirmed that the candidate variant *LAMA2*: c.909+7A>G was responsible for exon 6 skipping in individual 54. In individual 58, detection of reduced levels of the alpha-dystroglycan protein in muscle was necessary to classify the variant identified in *DPM2* as likely pathogenic.

DISCUSSION

This study shows a successful approach to addressing diagnostic challenges and establishes an effective workflow to apply for individuals with NMDs in which exome testing alone yields inconclusive results. Achieving a diagnostic rate of 40% surpasses the rate reported for studies solely employing exome reanalysis, in which the average rate was 15% [39]. Our strategy for identifying causal variants involved deep phenotyping conducted by a multidisciplinary team comprising neurologists, pathologists, rehabilitation physicians, pulmonologists and cardiologists, together with genetic variant analysis and interpretation by geneticists and clinical scientists. To our knowledge, no previous study has investigated the combined impact of applying deep, standardized phenotyping with trio genome sequencing and RNASeq in unsolved NMD cases.

GS was essential in cases involving complex structural or non-coding variants and poorly covered regions in the first NGS test. The ability of GS to interrogate intronic regions and detect CNVs, as well as its enhanced coverage of GC-rich regions, offers a key diagnostic advantage over exome sequencing in unsolved NMD cases. Notably, 15 out of 23 cases would not have been resolved solely through reanalysis of previous exome data. Furthermore, RNASeq was necessary in three cases for understanding the impact of variants in intronic regions, which are often challenging to interpret, and facilitated the identification of events such as the creation of pseudoexons, exon skipping, or exon truncation.

The trio analysis approach accelerates variant filtering significantly, particularly when identifying de novo and compound heterozygous variants. Inheritance information can be directly taken into account when filtering and prioritizing variants. Furthermore, the availability of allelic frequency analysis sheds light when investigating biallelic variants: identifying a variant in one parent and searching for a variant in trans allows for identifying situations involving compound heterozygosity while ruling out parental mosaicism. We showed that a trio approach

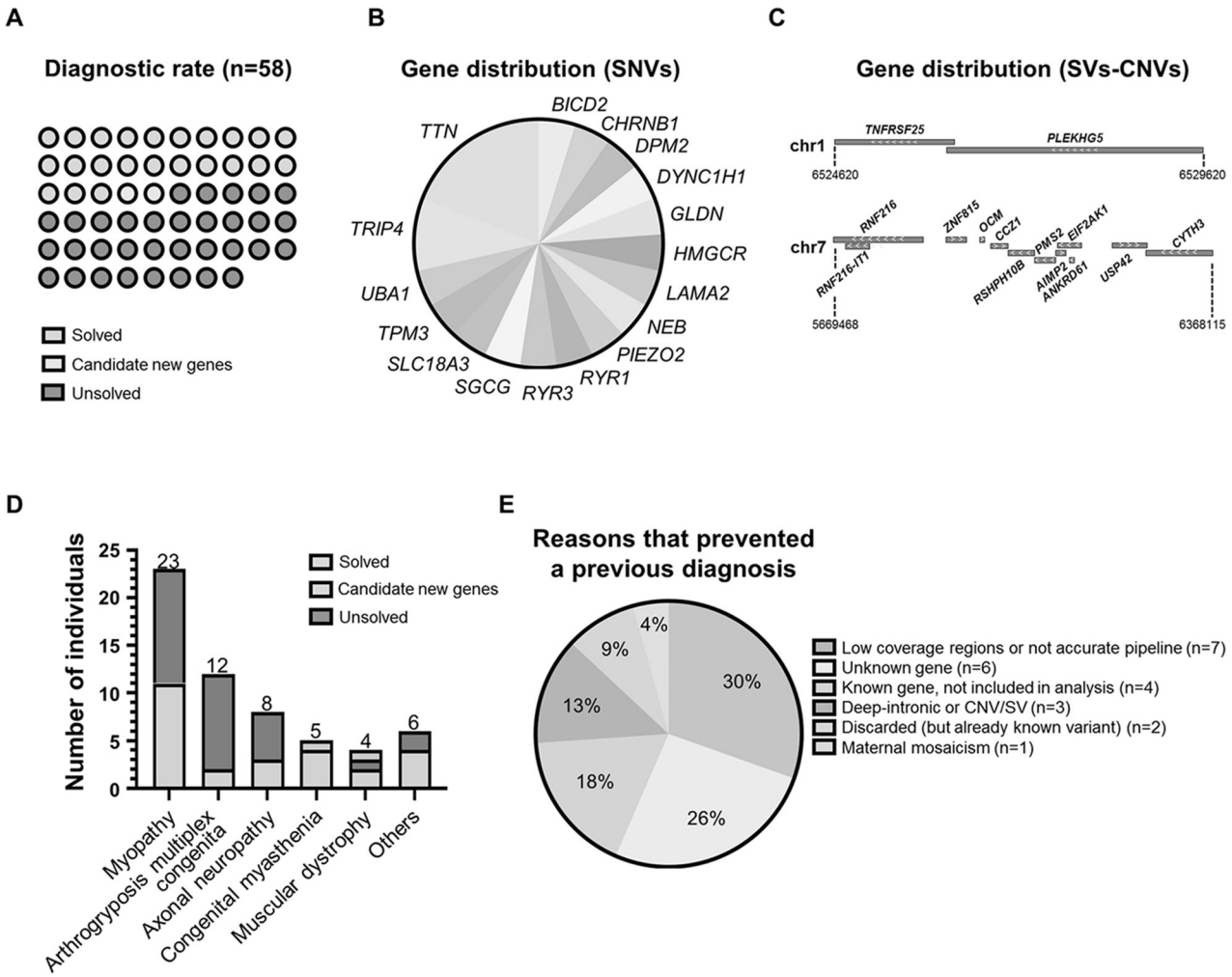

Fig. 3 Diagnostic rate overview. A Graphical representation of the achieved diagnostic rate. In green, the 23 solved cases. In 2 cases (blue), we identified candidate variants in potential new genes that are currently being studied through functional analysis. **B** The pie chart shows the gene distribution of diagnosed cases (where we found SNVs). All of them are responsible for the diagnosis of unique cases, with the exception of *TRIP4* (2 cases) and *TTN* (4 cases). The rest are represented in alphabetic order. **C** Gene distribution of CNVs and SVs. **D** Diagnostic yield according to phenotype classification: the highest diagnostic yields were observed within groups of individuals with muscular dystrophies, congenital myasthenic syndromes, and congenital myopathies. **E** Summary of reasons that prevented a previous diagnosis.

was essential in diagnosing an affected child with an *RYR1*-congenital myopathy: the presence of the variant in heterozygosity in the asymptomatic mother was justified by comparing her allelic frequency, which was much lower and compatible with mosaicism, as later reported in Estévez-Arias et al. [38]. This case would not have been solved solely by reanalyzing previous genomic data either.

The examination of factors preventing a prior diagnosis (Fig. 3E) revealed that 8 out of 23 solved cases (8 out of the total cohort of 58 individuals; 14%) could have been identified through a reanalysis of the initial NGS data. This finding aligns with other studies applying exome reanalysis [39], underscoring its utility while revealing that reanalysis enables a lower proportion of diagnosis compared to the comprehensive approach proposed here. Furthermore, despite the perceived simplicity of reanalysis, its systematic execution within hospital clinical laboratories poses significant challenges, such as data accessibility issues and time-consuming requirements for geneticists, which complicates periodic reanalysis in clinical settings. Nonetheless, structured, standardized, and easily accessible data facilitate periodic reanalysis looking for new genes or new genotype/phenotype associations, making it more feasible to implement in a research-oriented environment.

Additionally, the possibility of conducting functional studies to aid in interpreting candidate variants was essential in our workflow. It is a key aspect when considering or discarding findings in genomic analyses. Assessing the biological impact of a variant through experimental biology is not routinely applied as part of the diagnostic activity, but its efficacy is undeniable [40]. A total of 3 of the diagnosed cases would not have been resolved if several functional analyses had not been explored.

Finally, it is worth noting that among the undiagnosed cases (60%), we presume to have identified some new candidate genes, considering the continually growing catalog of approximately 15–30 novel gene-disease associations reported annually for NMDs. Variants responsible for the phenotype observed in the affected individuals were identified in 18 genes already included in the latest version of the Muscle Gene Table [6] (Fig. 3), except for a big deletion involving several genes (family 28).

We explored gene discovery in cases where the variants met the established filters and could be considered candidate findings according to the prioritization strategy. After excluding candidate genomic variants in genes associated with NMDs and abnormalities in splicing or expression in genes that could explain the patient's clinical presentation, an in silico study was undertaken, together with a literature review to determine if the implicated

Table 2. Variant found in the potentially solved case.

Potentially solved cases – candidate variants with highly compatible phenotype and functional studies in progress				
Individual	Gene (inheritance)	Ref transcript	Candidate variant(s)	ACMG classification
Others				
40	***DOT1L*** (AD)	NM_032482.3	c.766C>T (p.Arg256Trp)	Variant of uncertain significance PM2, PM6

The high compatibility between the reported phenotype and the phenotypical description guided us to further investigate this variant through functional studies.
Those variants that required functional studies are in bold.

gene could be associated with the individual's phenotype. We identified potential new genes in two cases, for which we are currently performing specific experimental biology studies to assess the impact of the identified variants on the gene products.

The main limitation of our study is that, although we have analyzed a substantial and specific series, generalizing these findings requires validation across additional series, both in the short term within research settings and potentially in clinical settings in the future. This validation is necessary to confirm the broader efficacy of trio-based genome sequencing coupled with RNASeq guided by deep phenotyping. Notably, the proposed workflow excels in identifying variants within genes already associated with diseases but poses challenges in identifying novel genotype-phenotype correlations. The discovery of new genes demands extensive efforts and functional studies for conclusive validation.

In summary, a systematic genetic analysis approach, based on standardized deep phenotyping, trio genome sequencing and RNASeq analysis, and undertaken by a multidisciplinary team in a research setting, demonstrated efficacy in our clinical series of individuals with unsolved NMDs. Additionally, this approach holds the potential to uncover new genes and genotype-phenotype associations, further enhancing our understanding of neuromuscular diseases.

DATA AVAILABILITY

Any data not published within the article will be shared by the corresponding author upon reasonable request. The genomic and phenotypic data is available to registered users of the RD-Connect GPAP. Registration is openly available to rare disease researchers across the world.

REFERENCES

1. Lek M, MacArthur D. The challenge of next generation sequencing in the context of neuromuscular diseases. J Neuromuscul Dis. 2014;1:135–49.
2. Thompson R, Spendiff S, Roos A, Bourque PR, Warman Chardon J, Kirschner J, et al. Advances in the diagnosis of inherited neuromuscular diseases and implications for therapy development. Lancet Neurol. 2020;19:522–32.
3. O'Grady GL, Lek M, Lamande SR, Waddell L, Oates EC, Punetha J, et al. Diagnosis and etiology of congenital muscular dystrophy: We are halfway there. Ann Neurol. 2016;80:101–11.
4. Harris E, Topf A, Barresi R, Hudson J, Powell H, Tellez J, et al. Exome sequences versus sequential gene testing in the UK highly specialised Service for Limb Girdle Muscular Dystrophy. Orphanet J Rare Dis. 2017;12:151.
5. Ghaoui R, Cooper ST, Lek M, Jones K, Corbett A, Reddel SW, et al. Use of Whole-Exome Sequencing for Diagnosis of Limb-Girdle Muscular Dystrophy: Outcomes and Lessons Learned. JAMA Neurol. 2015;72:1424–32.
6. Benarroch L, Bonne G, Rivier F, Hamroun D. The 2024 version of the gene table of neuromuscular disorders (nuclear genome). Neuromuscul Disord. 2024;34:126–70.
7. Natera-de Benito D, Töpf A, Vilchez JJ, González-Quereda L, Domínguez-Carral J, Díaz-Manera J, et al. Molecular characterization of congenital myasthenic syndromes in Spain. Neuromuscul Disord. 2017;27:1087–98.
8. Natera-de Benito D, Ortez C, Jou C, Jimenez-Mallebrera C, Codina A, Carrera-García L, et al. The Phenotype and Genotype of Congenital Myopathies Based on a Large Pediatric Cohort. Pediatr Neurol. 2021;115:50–65.
9. Engel AG. Genetic basis and phenotypic features of congenital myasthenic syndromes. Handb Clin Neurol. 2018;148:565–89.
10. Bansagi B, Griffin H, Whittaker RG, Antoniadi T, Evangelista T, Miller J, et al. Genetic heterogeneity of motor neuropathies. Neurology. 2017;88:1226–34.
11. Ewans LJ, Minoche AE, Schofield D, Shrestha R, Puttick C, Zhu Y, et al. Whole exome and genome sequencing in mendelian disorders: a diagnostic and health economic analysis. Eur J Hum Genet. 2022;30:1121–31.
12. Wright CF, FitzPatrick DR, Firth HV. Paediatric genomics: diagnosing rare disease in children. Nat Rev Genet. 2018;19:253–68.
13. Yuan X, Wang J, Dai B, Sun Y, Zhang K, Chen F, et al. Evaluation of phenotype-driven gene prioritization methods for Mendelian diseases. Brief Bioinform. 2022;23:bbac019.
14. Thompson R, Papakonstantinou Ntalis A, Beltran S, Töpf A, de Paula Estephan E, et al. Increasing phenotypic annotation improves the diagnostic rate of exome sequencing in a rare neuromuscular disorder. Hum Mutat. 2019;40:1797–812.
15. Laurie S, Piscia D, Matalonga L, Corvó A, Fernández-Callejo M, Garcia-Linares C, et al. The RD-Connect Genome-Phenome Analysis Platform: Accelerating diagnosis, research, and gene discovery for rare diseases. Hum Mutat. 2022;43:717–33.
16. Lochmüller H, Badowska DM, Thompson R, Knoers NV, Aartsma-Rus A, Gut I, et al. RD-Connect, NeurOmics and EURenOmics: collaborative European initiative for rare diseases. Eur J Hum Genet. 2018;26:778–85.
17. Zurek B, Ellwanger K, Vissers LELM, Schüle R, Synofzik M, Töpf A, et al. Solve-RD: systematic pan-European data sharing and collaborative analysis to solve rare diseases. Eur J Hum Genet. 2021;29:1325–31.
18. Robinson PN, Köhler S, Bauer S, Seelow D, Horn D, Mundlos S. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. Am J Hum Genet. 2008;83:610–5.
19. Mungall CJ, McMurry JA, Kohler S, Balhoff JP, Borromeo C, Brush M, et al. The Monarch Initiative: An integrative data and analytic platform connecting phenotypes to genotypes across species. Nucleic Acids Res. 2017;45:D712–22.
20. Laurie S, Fernandez-Callejo M, Marco-Sola S, Trotta JR, Camps J, Chacón A, et al. From Wet-Lab to Variations: Concordance and Speed of Bioinformatics Pipelines for Whole Genome and Whole Exome Sequencing. Hum Mutat. 2016;37:1263–71.
21. Depristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43:491–8.
22. Demidov G, Sturm M, Ossowski S. ClinCNV: multi-sample germline CNV detection in NGS data. bioRxiv. 2022. https://www.biorxiv.org/content/10.1101/2022.06.10.495642v1.
23. Minoche AE, Lundie B, Peters GB, Ohnesorg T, Pinese M, Thomas DM, et al. ClinSV: clinical grade structural and copy number variant detection from whole genome sequencing data. Genome Med. 2021;13:1–19.
24. Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics. 2016;32:1220–2.
25. Cohen E, Bonne G, Rivier F, Hamroun D. The 2022 version of the gene table of neuromuscular disorders (nuclear genome). Neuromuscul Disord. 2021;31:1313–57.
26. Matalonga L, Laurie S, Papakonstantinou A, Piscia D, Mereu E, Bullich G, et al. Improved Diagnosis of Rare Disease Patients through Systematic Detection of Runs of Homozygosity. J Mol Diagnostics. 2020;22:1205–15.
27. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405–24.
28. Leman R, Parfait B, Vidaud D, Girodon E, Pacot L, Le Gac G, et al. SPiP: Splicing Prediction Pipeline, a machine learning tool for massive detection of exonic and intronic variant effects on mRNA splicing. Hum Mutat. 2022;43:2308–23.
29. FelixKrueger/TrimGalore: v0.6.10 - add default decompression path. https://doi.org/10.5281/ZENODO.7598955 (2023).
30. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

31. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics. 2016;32:292–4.
32. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–30.
33. Yépez VA, Mertes C, Müller MF, Klaproth-Andrade D, Wachutka L, Frésard L, et al. Detection of aberrant gene expression events in RNA sequencing data. Nat Protoc. 2021;16:1276–96.
34. Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, et al. GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. 2019;47:D766–73.
35. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013;9:e1003118.
36. Brechtmann F, Mertes C, Matusevičiūtė A, Yépez VA, Avsec Ž, Herzog M, et al. OUTRIDER: A Statistical Method for Detecting Aberrantly Expressed Genes in RNA Sequencing Data. Am J Hum Genet. 2018;103:907–17.
37. Scheller IF, Lutz K, Mertes C, Yépez VA, Gagneur J. Improved detection of aberrant splicing with FRASER 2.0 and the intron Jaccard index. Am J Hum Genet. 2023;110:2056–67.
38. Estévez-Arias B, Matalonga L, Martorell L, Codina A, Ortez C, Carrera-García L, et al. Improving Diagnostic Precision: Phenotype-Driven Analysis Uncovers a Maternal Mosaicism in an Individual with RYR1-Congenital Myopathy. J Neuromuscul Dis. 2024;11:647–53.
39. Tan NB, Stapleton R, Stark Z, Delatycki MB, Yeung A, Hunter MF, et al. Evaluating systematic reanalysis of clinical genomic data in rare disease from single center experience and literature review. Mol Genet Genom Med. 2020;8:e1508.
40. Pijuan J, Rodríguez-Sanz M, Natera-de Benito D, Ortez C, Altimir A, Osuna-López M, et al. Translational Diagnostics: An In-House Pipeline to Validate Genetic Variants in Children with Undiagnosed and Rare Diseases. J Mol Diagn. 2021;23:71–90.

ACKNOWLEDGEMENTS

We are indebted to the "Biobanc de l'Hospital Infantil Sant Joan de Déu per a la Investigació" for the sample and data procurement. AN, BEA, CJ, CO, DNdB, JEE, LCG and HL are members of the European Reference Network for Neuromuscular Diseases – Project ID N° 870177. FP is member of the European Reference Network for Rare Malformation Syndromes, Intellectual and Other Neurodevelopmental Disorders.

AUTHOR CONTRIBUTIONS

Conceptualization: BEA, AN, JH, FP, DNdB; Formal analysis: BEA, LM, KP, SM, OAK, AA, RT, AEC, SL, GD, VAY, HL, DNdB; Methodology: LM, DY, AC, RL, AEC, MG, SL, GD, VAY, SB, JG; Supervision: SL, GD, VAY, SB, JG, AT, HL, AN, JH, FP, DNdB; Writing-original draft: BEA, AN, JH, FP, DNdB; Writing-review & editing: all authors.

FUNDING

BEA is supported by the predoctoral program "Joan Oró" of the Secretary of Universities and Research of the Department of Research and Universities of the Government of Catalonia (grant 2024 FI-1 00075), co-financed by the European Union. DNdB is supported by the Miguel Servet program from Instituto de Salud Carlos III, Spain (CP22/00141). FP is supported by the Spanish State Research Agency (AEI, grant PID2020-114655RB-100), CIBERER (grant ACCI2021-20), the EU Horizon Europe Health Programme, Fundación Ramón Areces and AGAUR (grant 2021-SGR-01610)). JH is supported by the Instituto de Salud Carlos III (grant PI22-00680), IRSJD Carmen de Torres 2023 grant and Torró Solidari-RAC1 i Torrons Vicens. HL receives support as principal investigator from the Canadian Institutes of Health Research (CIHR) for his Foundation Grant (FDN-167281) on Precision Health for Neuromuscular Diseases and Network Grant for NMD4C, the neuromuscular network for Canada, from the Canada Foundation for Innovation for CFI-JELF 38412, and the Canada Research Chairs program for his Canada Research Chair in Neuromuscular Genomics and Health (950-232279). He receives support as co-investigator from CIHR for ProDGNE (Transnational Team Grant), from the European Commission (Grant # 101080249) and the Canada Research Coordinating Committee New Frontiers in Research Fund (NFRFG-2022-00033) for SIMPATHIC, and from the Government of Canada Canada First Research Excellence Fund (CFREF) for the Brain-Heart Interconnectome. SM is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within the Clinician Scientist Program "Cell Dynamics in Disease and Therapy" at the University Medical Center Goettingen (project number 413501650). This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 779257 (Solve-RD). This study makes use of data and tools shared/provided through the RD-Connect GPAP, which received funding originally from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 305444. The remaining authors report no disclosures.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICAL APPROVAL

Data were collected and analyzed following the ethics guidelines of Hospital Sant Joan de Déu (PIC 98-20). The family provided written informed consent for the study.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41431-024-01699-4.

Correspondence and requests for materials should be addressed to Francesc Palau or Daniel Natera-de Benito.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.