Metabolic, Fibrotic, and Splicing Pathways Are All Altered in Emery-Dreifuss Muscular Dystrophy Spectrum Patients to Differing Degrees

Jose I. de las Heras[1], Vanessa Todorow[2], Lejla Krečinić-Balić[2], Stefan Hintze[2], Rafal Czapiewski[1], Shaun Webb[3], Benedikt Schoser[2], Peter Meinke[2,*], Eric C. Schirmer[1,*]

[1]Institute of Cell Biology, University of Edinburgh, Edinburgh, UK

[2]Friedrich-Baur-Institute, Department of Neurology, LMU Clinic, Ludwig-Maximillians-University, Munich, Germany

[3]Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh, UK

*Co-corresponding authors:

Eric C. Schirmer

Institute of Cell Biology; University of Edinburgh, Kings Buildings; Michael Swann Building, Room 5.22; Max Born Crescent; Edinburgh, EH9 3BF, UK; Phone: +441316507075; E-Mail: e.schirmer@ed.ac.uk

Peter Meinke

Friedrich-Baur-Institute, Department of Neurology, LMU Clinics, Ludwig-Maximillians-University, Ziemssenstrasse 1, 80336 Munich, Germany; Phone +4989218078279; E-Mail: Peter.Meinke@med.uni-muenchen.de

Abstract

Emery-Dreifuss muscular dystrophy (EDMD) is a genetically and clinically variable disorder. Previous attempts to use gene expression changes find its pathomechanism were unavailing, so we here engaged a functional pathway analysis. RNA-Seq was performed on cells from 10 patients diagnosed with an EDMD spectrum disease with different mutations in 7 genes. Upon comparing to controls, the pathway analysis revealed that multiple genes involved in fibrosis, metabolism, myogenic signaling, and splicing were affected in all patients. Splice variant analysis revealed alterations of muscle-specific variants for several important muscle genes. Deeper analysis of metabolic pathways revealed a reduction in glycolytic and oxidative metabolism and reduced numbers of mitochondria across a larger set of 14 EDMD patients and 7 controls. Intriguingly, the gene expression signatures segregated the patients into three subgroups whose distinctions could potentially relate to differences in clinical presentation. Finally, differential expression analysis of miRNAs changing in the patients similarly highlighted fibrosis, metabolism, and myogenic signaling pathways. This pathway approach revealed a clear EDMD signature that can both be used as the basis for establishing a biomarker panel specific to EDMD and direct further investigation into its pathomechanism. Furthermore, the segregation of specific gene changes into three distinct categories that appear to correlate with clinical presentation may be developed into prognostic biomarkers, though this will first require their testing in a wider set of patients with more clinical information.

Keywords

Emery-Dreifuss, muscular dystrophy, EDMD, skeletal muscle

Introduction

Emery-Dreifuss muscular dystrophy (EDMD) is a genetically heterogeneous neuromuscular orphan spectrum disease affecting ~0.3-0.4 in 100,000 people [1], with clinical variability presenting even in family members carrying the same mutation [2-5]. EDMD patients present typically in mid to late childhood with early contractures of elbows and Achilles' tendons and progressive wasting of the lower leg and upper arm muscles. Cardiac involvement is also highly characteristic but tends to appear later in development and quite variably in time, though it tends to be reasonably uniform in the form it takes of cardiac conduction defects and dilated cardiomyopathy [6]. Other features vary considerably in clinical presentation, leading to the usage of 'Emery-Dreifuss-like syndromes' [7; 8]: patients from the same pedigree can show remarkable phenotypic variation [2]. The genetic variability is underscored by several confirmed linked genes and several additional candidate genes, although there are still some cases where no confirmed or candidate disease allele has been identified [9-11]. The lack of large pedigrees in combination with its genetic heterogeneity, clinical variability, already some known modifier genes, and limited patient numbers makes solving its pathomechanism difficult.

The original genes linked to EDMD, *EMD* encoding emerin and *LMNA* encoding lamin A, have both cytoskeletal and gene regulation roles leading to strong arguments for either function being responsible for the EDMD pathomechanism [12; 13]. The subsequent linking of nesprin and Sun proteins to EDMD [14; 15] failed to lend clarity since they function in mechanosignal transduction [16]. However, several recently linked genes have clear roles in genome organization and regulation [10], suggesting this is the pathomechanism. These genes encode proteins that, like emerin, are nuclear envelope transmembrane proteins (NETs) and seem to function by fine-tuning muscle gene expression by promoting the release of pro-myogenic genes from the nuclear periphery to enhance their activation while concomitantly recruiting metabolism genes (many from the alternative differentiation pathway of adipogenesis) to the nuclear envelope to better repress them [17-19]. EDMD mutations were found in 5 muscle-specific NETs with this genome organization function,

PLPP7 (also known as NET39), WFS1, TMEM38A, TMEM201, and Tmem214, and each tested had some specificity in the sets of genes that they target, though there was also some overlap [17]. These studies together with the wide range of lamin gene regulatory activities led us to the distinct and non-traditional hypothesis for the EDMD pathomechanism where moderate reductions in many genes could have the same phenotype as a shutdown of a single gene on a particular pathway. Accordingly, we considered that searching for uniformity with a pathway analysis might be more revealing than searching for uniformity in the regulation of particular genes.

The only previous study, to our knowledge, using gene expression changes to identify critical misregulated genes underlying EDMD pathophysiology, focused on the identification of genes altered specifically in EDMD compared to a set of 10 other muscular dystrophies [20]. This study only considered 8 total *LMNA*- or *EMD*-linked cases of EDMD, but EDMD now has many more genes and modifiers linked to it and, moreover, there is a wider clinical spectrum of EDMD-like phenotypes [11]. Their analysis indicated potential abnormalities in the regulation of cell cycle and myogenic differentiation, associated with perturbations in the pRB/MYOD/LMNA hub, which were consistent with changes in an *Emd*$^{-/-}$ mouse model [21]. Roughly a fifth each of EDMD mutations occurs in *LMNA* and *EMD* while another 5-6% are collectively caused by four other widely expressed nuclear envelope proteins nesprin 1 (encoded by *SYNE1*), nesprin 2 (encoded by *SYNE2*), Sun1 (encoded by *SUN1*) and FHL1 (encoded by *FHL1*) [14; 15; 22-24]. Another approximately 20% of EDMD mutations were accounted for by muscle-specific NETs that regulate muscle-specific genome organization [10]. These include NET39 (encoded by *PLPP7*), TMEM38A (encoded by *TMEM38A*), WFS1 (encoded by *WFS1*), NET5 (encoded by *TMEM201*), and TMEM214 (encoded by *TMEM214*) that affect 3D gene positioning with corresponding effects on expression [17; 19]. Accordingly, we sought to search for commonly affected pathways from a much wider range of EDMD-linked genes including *LMNA*, *EMD*, *FHL1*, *SUN1*, *SYNE1*, *PLPP7*, and *TMEM214* alleles on the expectation that the most important pathways for EDMD pathophysiology would be highlighted.

Material and methods

Patient materials

The sources of patient samples were the Muscle Tissue Culture Collection (MTCC) at the Friedrich-Baur-Institute (Department of Neurology, Ludwig-Maximilians-University, Munich, Germany) and the MRC Centre for Neuromuscular Disorders Biobank (CNDB) in London.

Ethical approval and consent to participate

All materials were obtained with written informed consent of the donor at the CNDB or the MTCC. Ethical approval of the Rare Diseases biological samples biobank for research to facilitate pharmacological, gene and cell therapy trials in neuromuscular disorders is covered by REC reference 06/Q0406/33 with MTA reference CNMDBBL63 CT-2925/CT-1402, and for this particular study was obtained from the West of Scotland Research Ethics Service (WoSRES) with REC reference 15/WS/0069 and IRAS project ID 177946. The study conduct and design complied with the criteria set by the Declaration of Helsinki.

Myoblast culture and in vitro differentiation into myotubes

Myoblasts were grown in culture at 37 ^{O}C and 5% CO_2 using a ready to use formulation for skeletal muscle (PELOBiotech #PB-MH-272-0090) and maintained in subconfluent conditions. In order to induce differentiation, the cells were grown to confluency and 24 h later the growth medium replaced with skeletal muscle differentiation medium (Cell Applications #151D-250). The differentiation medium was replaced every other day. Myotubes were selectively harvested after 6 days by partial trypsinization followed by gentle centrifugation (Figure S1). Each differentiation experiment was performed in triplicate on different days.

RNA extraction

Total RNA was extracted from each sample and separated into a high molecular weight fraction (> 200 nt, for mRNA-Seq) and a low molecular weight fraction (< 200 nt, for miRNA-

Seq) with the Qiagen RNeasy (#74134) and miRNeasy (#1038703) kits, according to the manufacturer's instructions. RNA quality was assessed with a Bioanalyzer, and all samples had a RIN > 7, with an average of 9.4 (Table S1).

mRNA-Seq analysis

Between 3-5ug of total RNA were sent to Admera Health LLC. (NJ, USA) for sequencing in paired-end mode, 2x 150 nucleotides, using an Illumina HiSeq 2500 sequencer. The sequencing library was prepared with the NEBNext Ultra II kit, with RiboZero rRNA depletion (NEB #E7103). Between 60-90 million paired end reads were obtained from each sample and mapped to the human genome (Hg38) with STAR v2.7.5a [96] using default parameters. Mapping quality was assessed with FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). Sequencing adaptors were removed with trimmomatic v0.35 [97]. Low quality reads and mitochondrial contaminants were removed, leaving on average 70 million useful reads per sample (Table S1). Differential expression analysis was performed in R with DESeq2 v1.32.0 [98] after transcript quantitation with Salmon v1.4.0 [99]. We used an FDR threshold of 5% for differential expression.

miRNA-Seq analysis

miRNA was sent to RealSeq Biosciences Inc. (SC, USA) for sequencing using an Illumina NextSeq 500 v2 sequencer in single end mode, 1x 75 nucleotides. The sequencing library was prepared with Somagenics' Low-bias RealSeq-AC miRNA library kit (#500-00012) and quality assessed by Tapestation (Lab901/Agilent). On average 5 million good quality reads were obtained per sample. Mapping and quality trimming was performed using the NextFlow nf-core/smrnaseq pipeline (web resources) with default parameters, which summarizes the reads per miRNA using the annotations from mirTop (web resources). Differential expression analysis was performed in R with DESeq2 v1.32.0. We used an FDR threshold of 0.2 for differential expression. Putative miRNA targets were extracted from miRDB (web resources) for each differentially expressed miRNA and their expression compared against

the miRNA. We kept as potential targets those genes whose expression changed in the opposite direction of the miRNA.

Bakay muscular dystrophy dataset analysis

Normalized (MAS5.0) microarray transcriptome data for a panel of 11 muscular dystrophies and healthy controls were downloaded from the Gene Expression Omnibus database (GEO), accession GSE3307 (web resources). Differential expression analysis comparing each disease to the controls was performed using Limma 3.48.1 [100]. We used an FDR threshold of 5% for differential expression.

Functional analyses

Functional analyses were performed with g:Profiler [101] and Gene Set Enrichment Analysis (GSEA v4.1.0) [35] tools. g:Profiler was used to determine enriched categories within a set of DE genes, with an FDR of 5% as threshold. GSEA was performed with default parameters, in particular using 'Signal2Noise' as ranking metric and 'meandiv' normalization mode. Redundancy in category lists was reduced by comparing the similarity between each pair of enriched categories using Jaccard similarity coefficients. Hierarchical clustering (k-means) was then applied to the resulting matrix in order to identify groups of similar functional categories, and a representative from each group chosen. Full unfiltered results are shown in Supplemental Table S3. Tissue-specific gene enrichment analysis was evaluated with TissueEnrich [102].

For the miRNA-Seq experiments, functional analysis was first performed using g:profiler on the set of DE miRNA genes. Then, putative targets for each miRNA were extracted and their expression compared to the relevant miRNA. Putative targets whose expression was not altered in the opposite direction as the miRNA were removed from the list. Significant functions were displayed using Cytoscape v3.8.2 [103], with the size of the functional labels proportional to the number of miRNAs assigned to each function.

Real time metabolic measurements

Metabolic measurements on primary human myoblast cultures were performed using a Seahorse XFp Extracellular Flux Analyzer (*Agilent Technologies*). For this, myoblasts of matched passage number were seeded in XFp Cell Culture Miniplates (103025-100, *Agilent Technologies*) at a density of 1.5×10^4 cells per well. Cell density was assessed using an automated cell counter (TC20, BioRad). Oxygen consumption rates (OCR) respectively extracellular acidification rates (ECAR) were measured using the Mito Stress Test Kit and the Glycolysis Stress Test Kit (*Agilent Technologies*) according to the manufacturer's instructions. Samples were measured in triplicates and each measurement was repeated between two and four times. Data were normalized to the number of cells and analyzed for each well.

Fuel dependency tests

Glucose dependency and fatty acid dependency were determined according the instruction of Agilent Seahorse XF Mito Fuel Flex Test kit (Agilent). The glutamine dependency was determined from the glucose and fatty acid measurements.

Mitochondrial gene quantification

Reverse transcription of RNA was performed using the QuantiTecT Reverse Transcription Kit (Qiagen) following the manufacturer's instructions. For the reaction we used the SYBR® Green Master Mix (Bio-Rad) and samples were run and measured on CFX Connect™ (Bio-Rad). As genome reference gene *B2M* (FP: 5'-TGCTGTCTCCATGTTTGATGTATCT-3'; RP: 5'-TCTCTGCTCCCCACCTCTAAGT-3') [104] Primer sequences for the mitochondrial genome were: FP: 5′-TTAACTCCACCATTAGCACC-3′; RP: 5′-GAGGATGGTGGTCAAGGGA-3′. [105] Samples were analyzed using the delta Ct method.

Splice site prediction analysis

Raw data was mapped to the human genome assembly GRCh38 (hg38) and sorted by coordinate using STAR 2.7.9a [96] for analysis in DESeq2 [98] and DEXSeq [106], trimmed using an in-built trimming function for rMATS [107] or counted using Kallisto 0.48.0 [108] for isoformSwitchAnalyzer (ISA) [109]. All analyses were performed for G1, gp1, gp2, and gp3 separately. Visualizations were conducted in R version 4.1.2.

DESeq2: Mapped reads were counted using FeatureCounts, then analysed using DESeq2 [98]. The R package fgsea was used for gene set enrichment analysis, genes were assigned to biological pathways retrieved from MSigDB v7.5.1 (c5.go.bp.v7.5.1.symbols.gmt, 7658 gene sets, [35]). Splicing pathways and their genes were plotted using GOPlot [110].

DEXseq: Mapped reads were counted using the in-built DEXseq counting function in python 3.9. Standard DEXseq workflow was followed. Exons with logFCs > |1| and p-values < 0.05 were set to be significantly different.

rMATS: Standard workflow was followed, code was executed in python 2.7. Results were analysed in R and set to be significantly differentially spliced with psi-values > |0.1| and p-values < 0.05. Pie charts displaying the distibution of event usage were generated for all groups (Supps. something) GO term enrichment analyses was performed using g:profiler2 [111]. All events were searched for muscle specific genes using a set of 867 genes relevant for muscle system process, development, structure and contraction, combined from GO terms (GO:0003012, GO:0006936, GO:0055001 and GO:0061061).

ISA: Kallisto counts were read into R and standard ISA procedure was followed, including Splicing analysis using DEXseq, coding potential using CPC 2.0 [112], domain annotation using HmmerWeb Pfam 35.0 [113], signal peptides using SignalP 5.0 [114] and predicition of intrisically unstructured proteins using IUPred2A [115]. In-built visualization tools were used for splicing maps.

Results

RNA-Seq Analysis of EDMD Patient Cells

We performed RNA-Seq on myotubes differentiated *in vitro* from myoblasts isolated from 10 unrelated EDMD patients with distinct mutations in 7 different genes to sample the genetic diversity of EDMD (**Figure 1**). Patient mutations were TMEM214 p.R179H, PLPP7/NET39 p.M92K, Sun1 p.G68D/G388S, Nesprin 1 p.S6869*, Emerin p.S58Sfs*1, FHL1 mutations c.688+1G>A, p.C224W, and p.V280M, and Lamin A mutations p.T528K and p.R571S. These mutations covered a wide range of clinical phenotypes with the age of onset ranging from early childhood to adult life and associated pathology ranging from no reported contractures to rigid spine (**Table 1**). Myoblast isolation followed by *in vitro* differentiation was chosen over directly isolating mRNA from the tissue samples in order to try to capture the earliest changes in gene expression due to the disease mutations and to reduce tertiary effects and variation from the age of the patients, range in time from onset to when biopsies were taken, and differences in biopsy site (**Table 1**). These patient variables could also affect the efficiency of myotube differentiation; so, to ensure that different percentages of undifferentiated cells in the population did not impact measuring gene expression changes, the myotubes were specifically isolated by short trypsinization thus removing all myoblast contamination (**Figures 1 and S1**). To define the baseline for comparison, 2 age-matched healthy controls were similarly analyzed. Samples from all patients and controls all yielded high-quality reads ranging between 28 and 47 million paired-end reads (**Table S1**).

First, we compared each individual patient against the controls. Compared to the controls, each individual patient had between 310 and 2,651 upregulated genes and between 429 and 2,384 downregulated genes with an FDR of 5% (**Figure 1**). The large difference in the number of differentially expressed genes between patients suggested large heterogeneity. When we calculated the intersection of DE genes in all patients, only three genes were similarly downregulated (*MTCO1P12*, *HLA-H*, *HLA-C)*, and one upregulated (*MYH14)* at 5% FDR, indicating a high degree of variation between patients (**Figure S2**). *MTCO1P12* is a mitochondrially-encoded pseudogene that has been reported to be severely downregulated in inflammatory bowel disease, associated with reduced mitochondrial energy production [25]. HLA-C is a member of the MHC class I and is involved in interferon

gamma signaling while HLA-H is a pseudogene derived from HLA-A which may function in autophagy [26-28]. Mutations in non-muscle myosin gene *MYH14* appear to be associated with hearing loss rather than muscle defects [29; 30], although it has also been recently linked to mitochondrial fission defects [31].

In order to identify common features despite the above-mentioned heterogeneity, we next compared the patients as a single group against the controls instead of one-by-one and we denoted this comparison as G1 (one single group) throughout the text (**Figure 1**). The G1 comparison revealed a signature of 1127 DE genes (894 upregulated and 233 downregulated). Approximately 60% of these genes identified were altered in the same manner in all 10 patients, although only 4 of them had an associated FDR value below 0.05 (**Figure S2**). The remaining 40% appear to be quite variable between patients in level of expression and/or direction of change. Such distinction in gene expression signatures might also contribute to clinical variation in EDMD.

Hierarchical clustering identified three broad patient subgroups based on gene expression patterns (**Figure 1)**. These groupings were: Group 1 - Emerin p.S58Sfs*1, Tmem214 p.R179H, NET39 p.M92K, Lamin A p.R571S, FHL1 p.C224W; Group 2 - Sun1 p.G68D/G388S, Fhl1 c.688+1G>A, Fhl1 p.V280M; Group 3 - Nesprin 1 p.S6869*, Lamin A p.T528K (**Fig. 1D**). The same groups were independently identified using a principal component analysis (PCA) (**Figure S2**). Of particular note, PCA and t-distributed stochastic neighbor embedding (t-SNE) analyses both revealed that clustering was independent of parameters such as patient gender or age or myotube enrichment differences (**Figure S3**). Moreover, the several FHL1 and lamin A mutations tested segregated into different expression subgroups. At the same time, the more recently identified EDMD mutations in TMEM214 and NET39 segregated with more classic emerin, FHL1, and lamin A mutations, further indicating the likelihood that their genome organizing functions could mediate core EDMD pathophysiology.

The hypothesis that EDMD is a disease of genome organization misregulation is underscored by the fact that 15% of the genes changing expression in EDMD patient cells

were altered by knockdown of at least 1 of the 4 muscle-specific genome-organizing NETs that we previously tested [17] (**Figure 1**). Interestingly, in that study most of the genes altered by knockdown of NET39, TMEM38A, and WFS1 were non-overlapping, while those altered by knockdown of TMEM214 exhibited considerable overlap with the sets altered by each of the other NETs [17]. However, here there were roughly 70 DE genes overlapping with the sets of genes altered by knockdown of each individual NET while the total number of DE genes under the regulation of any of the four NETs was 165, indicating an enrichment in the EDMD DE set for genes influenced by multiple NETs (**Figure 1**). Thus, it is not surprising that NET39 and TMEM214 were both segregated together. Another interesting observation is that the number of NET-regulated genes overlapping with Group 1 and Group 3 was similar, but much fewer were overlapping for Group 2. This suggests that gene misregulation in Groups 1 and 3 might be more strongly mediated by the muscle-specific NET-gene tethering complexes than in Group 2.

Functional Pathway Analysis of Gene Expression Changes in EDMD Patient Cells

The primary aims of this study were to determine whether a functional pathway analysis would be more effective at revealing the underlying EDMD pathomechanism than just looking for uniformly altered genes and to identify possible biomarkers in the gene expression signatures. Before using this approach with the wider set of EDMD alleles, we applied a pathway analysis to the data from the previous microarray study by Bakay and colleagues where just *LMNA* and *EMD* mutations were considered [20]. We reanalyzed Bakay's EDMD data and extracted the subset of DE genes with FDR of 5% (1349 and 1452 up- and down-regulated genes, respectively). In order to identify enriched functional categories within each set of DE genes, we used g:Profiler [32]. This tool calculates the expected number of genes to be identified for any given functional category by chance and compares it to the number of genes observed. We selected categories that were significantly enriched with an FDR of 5%. The resulting list was then summarized by selecting representative classes using a similar approach to Revigo [33], but extended to other

functional category databases in addition to gene ontology (GO) terms. Briefly, similarity matrices were generated by calculating pairwise Jaccard similarity indices between categories and used this information to group together similar functional categories based on the genes identified. Redundancy was then reduced by choosing a representative category from each group.

The functional categories enriched in the set of EDMD-upregulated genes revealed defects in cytokine signaling, organization of the extracellular matrix (ECM), and various signaling pathways important for muscle differentiation and function (*e.g*: PI3K-Akt, TGF-beta, SMADs). In addition, there was an aberrant upregulation of alternative differentiation pathways, notably adipogenesis but also angiogenesis and osteogenesis. The functions highlighted among the downregulated genes were largely related to metabolism, mitochondrial especially, as well as ribosome biogenesis, muscle contraction, and myofibril assembly (**Figure 2**). Applying the same methodology to our wider set of patient alleles highlighted fewer pathways than what we observed in the Bakay EDMD data. Among the upregulated categories, neurogenesis and ECM-related functions stood out, as well as MAPK signaling, lipid transport, and TAP binding which is linked to interferon-gamma signaling [34]. One category stands out among the downregulated genes: RNA splicing (**Figure 2**). The data above used g:Profiler which is very sensitive to the number of DE genes identified because it looks for statistical overrepresentation of genes belonging to specific functional categories among a set of previously identified DE genes. By contrast, gene set enrichment analysis (GSEA) [35] does not prefilter the data and instead ranks all genes according to the difference in expression between the two conditions tested: controls and EDMD. Next, it determines whether the distribution in the ranked list for any given functional category is random or significantly enriched statistically at either end of the ranked list. This method is especially sensitive for detecting functional categories where many genes are altered by a small amount and does not consider individual gene p-values. Therefore, we also applied GSEA to our data, querying several functional genesets within the Reactome, KEGG and WikiPathways databases, found in the Molecular Signatures

Database (MSigDB) [36]. This approach identified a larger set of functional categories that generally expanded on those identified by g:Profiler and matched better what we observed from the Bakay EDMD geneset, with strong links to ECM organization that may be relevant to fibrosis, cytokine signaling, metabolism, differentiation, and splicing (**Figure 2**).

An expansion of categories for metabolic functions included specific categories for diabetes mellitus, adipogenesis, white and brown fat differentiation, nitrogen metabolism fatty acid metabolism, retinol metabolism, and many others. Similarly, there was an expansion of cytokines supporting inflammation for the fibrotic pathways and proteoglycans and elastin adding to the previous emphasis on collagens for ECM defects. Among the differences between our data and Bakay's EDMD data, two categories stand out: RNA splicing and calcium signaling, which were only observed in our data. It is unclear how much this reflects using terminally differentiated muscle material versus early stages of differentiation *in vitro*, or a factor of microarray versus RNAseq analysis. In some cases, this is most likely due to the different transcriptome platform used. For example, applying GSEA to genomic positional genesets revealed near uniform upregulation of all mitochondrially encoded genes (**Figures 2** and **3**). This could not be observed on Bakay's data because the microarrays did not contain probes for mitochondrially encoded genes. The upregulation of mitochondrial transcripts could lead to increased oxidative stress [37]. This finding provides yet another mechanism that could lead to metabolic dysregulation on top of the alterations already indicated by the nuclear genome transcript changes. This further underscored the need to test for actual metabolic deficits in the patient cells themselves as well as to further investigate the other functional pathways highlighted by this analysis.

Detailed Analysis of Metabolic Pathways Uniformly Altered in EDMD Patients

Since metabolic disruption has been previously reported to affect muscle differentiation/myoblast fusion [38], we decided to investigate this further. While we identified a general upregulation of mitochondrially-encoded genes, Bakay's data showed a downregulation of several classes related to mitochondrial function (**Figure 2**) which was

388 due entirely to nuclear-encoded genes, as there were no mitochondrial genes represented in
389 the microarrays (**Figure 3**). When we checked the behavior of those genes in our data, we
390 did not observe the same downregulation when considering all 10 patients as a single group
391 **(Figure 3**). However, this is largely due to variability among the patient subgroups identified
392 earlier, suggesting a mechanistic breakdown between them. Group 1 which contained half of
393 our patients, including emerin and lamin A mutations, exhibited the same general
394 downregulation of the nuclear-encoded mitochondrial genes. In contrast, Group 2 displayed
395 no alteration in gene expression, while Group 3 showed upregulation although this was
396 driven mostly by patient 9 (*LMNA*) with the other patient in the group, patient 4 (*SYNE1*),
397 displaying very few changes. While no single gene was uniformly altered in the same
398 direction for all patients, several genes from glycolytic and oxidative metabolism pathways,
399 typically encoding components of mitochondrial complexes, were altered in all tested
400 patients. Other non-mitochondrial metabolic pathways were also altered such as lipid
401 translocation (**Figure 2 and Table S3**). Interestingly, downregulation of nuclear-encoded
402 mitochondrial genes was also generally observed in other muscular dystrophies included in
403 the study by Bakay and colleagues (**Figure S4**)

404 To investigate the relevance of these gene changes to cellular metabolism, we
405 performed real-time metabolic analysis using the Seahorse XFp Extracellular Flux Analyzer.
406 Myoblasts isolated from the above patients plus several additional EDMD patients and
407 controls were tested, so that we had a total of 14 EDMD patients and 8 controls for this
408 analysis (**Table 1**). Probing for glycolysis, a significant reduction of the extracellular
409 acidification rate in the EDMD samples was observed (**Figure 3**). Next, we investigated
410 mitochondrial function. When testing for basal respiration there was also a significant
411 reduction of the oxygen consumption rate in the EDMD samples (**Figure 3**). There were no
412 significant differences in fuel dependency, but ATP production was considerably reduced in
413 the EDMD samples (**Figure 3**). The significant reduction in mitochondrial respiration raised
414 another possibility to investigate, that the absolute number of mitochondria might also be
415 down due to problems in mitochondria biogenesis. Therefore, we quantified relative

mitochondria numbers using by qPCR. This revealed a clear reduction in mitochondria numbers (**Figure 3**), which with the generally elevated mitochondrial genome transcripts would suggest that a reduction in mitochondria numbers resulted in an over-compensation of expression which in turn could have resulted in inhibiting mitochondrial fission and repair. Thus, we also investigated whether genes in pathways associated with mitophagy were altered in the patients. Indeed, multiple mitophagy pathway genes were altered in all patients (**Figure 3**). Although no one individual gene was altered in all the patients, it is worth noting *CISD2* is significantly downregulated in most patients. Reduction of CISD2 has been linked to degeneration of skeletal muscles, misregulated Ca^{2+} homeostasis and abnormalities in mitochondrial morphology in mouse [39], as well as cardiac dysfunction in humans [40].

Detailed Analysis of Other Pathways Uniformly Altered in EDMD Patients

Several studies suggest that the timing of several aspects of myotube fusion could underlie some of the aberrancies observed in patient muscle [41] and, though it is unclear whether fibrosis drives the pathology or is a consequence of the pathology, fibrosis has been generally observed in EDMD patient biopsies. Contributing to these processes could be several subpathways that fall variously under the larger pathways for ECM/ fibrosis, cell cycle regulation, and signaling/ differentiation (**Figure 4**). As for the metabolic analysis, no individual genes were altered in cells from all patients, but every patient had some genes altered that could affect ECM through changes in collagen deposition (**Figure 4**). For example, 35 out of 46 collagen genes exhibited changes in at least one comparison (Bakay EDMD, G1 or one of the subgroups gp1, gp2 and gp3) and all patients had multiple of these genes altered (**Figure S5**). Note that it often appears visually that the Bakay data in the first column has little change when viewing the cluster analysis, but when looking at the full set of genes listed in the matching supplemental figures there are definitely some genes strongly changing, just not necessarily the same ones. This may be due to differences in the myogenic state of the material studied: while Bakay and colleagues used muscle biopsies containing terminally differentiated muscle fibres, we focused on the earlier stages of

444 myogenesis by *in vitro* differentiating cultured myoblasts obtained from muscle biopsies.
445 Despite this, it is important to note that while different genes may be affected, most of the
446 same pathways were highlighted in both Bakay's and our study. Collagens COL6A1,
447 COL6A2, COL6A3, COL12A1 are linked to Bethlem Muscular Dystrophy [42-45] and,
448 interestingly, all these collagens were upregulated in Group 1 patient cells and
449 downregulated in Group 3 patient cells (**Figure S5**). Matrix metalloproteinases were also
450 altered with 13 out of 28 matrix metalloproteinases exhibiting changes in at least one of the
451 comparisons and all patients had multiple of these genes altered (**Figures 4 and S5**).
452 Notably the metalloproteinase MMP1 (collagenase I), which has been proposed to resolve
453 fibrotic tissue [46], was downregulated in all but one patient, as well as in Bakay EDMD
454 samples. Likewise, multiple genes associated with fibrosis from FibroAtlas (**Figures 4 and**
455 **S6**) and with inflammation that would support fibrosis such as cytokine (**Figures 4 and S7**)
456 and INF-gamma signaling (**Figures 4 and S8**) were affected in all patients. In fact, out of
457 941 genes in FibroAtlas there were 542 altered between all the patients. Heatmaps of gene
458 clusters with similar expression patterns are shown in Figure 4, but more detailed individual
459 panels with all gene names listed are shown in **Figures S5-S11**. A few genes that stand out
460 for their functions within the INF-gamma signaling pathway include IRF4 that is a regulator of
461 exercise capacity through the PTG/glycogen pathway [47] and ILB1 that helps maintain
462 muscle glucose homeostasis [48] such that both could also feed into the metabolic pathways
463 altered.

464 Another subpathway critical for myogenesis and the timing and integrity of myotube
465 fusion is cell cycle regulation. Cell cycle defects could lead to spontaneous differentiation
466 and were previously reported in myoblasts from EDMD patients and in tissue culture cell
467 lines expressing emerin carrying EDMD mutations which could lead to depletion of the stem
468 cell population [41; 49]. All tested EDMD patients exhibited downregulation of multiple genes
469 involved in the degradation of cell cycle proteins (**Figures 2**, **4**, and **S8**) which could indicate
470 an uncoupling of the joint regulation of cell cycle and myogenesis program [50], for example
471 cells starting to fuse when they should still be dividing or *vice versa*.

Other pathways in addition to ECM deposition directly associated with myogenic differentiation, myoblast fusion, and muscle regeneration were also altered in all patients (**Figures 4 and S9-S10**), though, again, no single gene in these pathways was altered in the same way in all patients' cells. Poor differentiation and myotubes with nuclear clustering were observed in differentiated EDMD myoblast cultures [14] and in the mouse C2C12 differentiation system when EDMD-linked NETs were knocked down [17].

Previous work using C2C12 cells identified six genes whose products are required in the early differentiation stages and were under the regulation of muscle-specific genome-organizing NETs [17]. These genes (*NID1*, *VCAM1*, *PTN*, *HGF*, *EFNA5*, and *BDNF)* are critical for the timing and integrity of myotube fusion and need to be expressed early in myoblast differentiation but shut down later or they inhibit myogenesis [51-55]. All six genes were misregulated in at least 5 but none were affected in all patients (**Figure 4**). In general terms, these genes were upregulated in Group 1, downregulated in Group 3, and mixed in Group 2. All six genes were upregulated in Bakay's EDMD data, although only *NID1* and *HGF* were statistically significant at 5% FDR. Both were upregulated only in Group 1 and downregulated in Group 3. *PTN* showed a similar pattern of expression as *HGF* although the only statistically significant changes were for upregulation in Group 1.

Several myogenic signaling pathways were altered such as MAPK, PI3K, BMP and Notch signaling and several alternate differentiation pathways were de-repressed such as adipogenesis that could disrupt myotube formation and function (**Figure 2 and Table S3**). Myogenesis and adipogenesis are two distinct differentiation routes from the same progenitor cells and whichever route is taken the other becomes repressed during normal differentiation [56; 57]. We previously showed that knockout of fat- or muscle-specific genome organizing NETs yield de-repression of the alternate differentiation pathway [17; 58] and the Collas lab showed that Lamin A lipodystrophy point mutations yield de-repression of muscle differentiation genes in adipocytes [59]. We now find here that adipogenesis genes are upregulated in both Bakay's EDMD data and our data (**Figure 2**). This is especially prominent for the five patients in Group 1 while Group 3 showing strong downregulation of a

subset of the same genes and Group 2 broadly looking like an intermediate of the other two groups (**Figures 4 and S11-S12**), and thus could also contribute to the metabolic defect differences between patients.

Splicing Pathways Uniformly Altered in EDMD Patients Yield Loss of Muscle-Specific Splice Variants

Among the downregulated functional categories, mRNA splicing stood out with many genes uniformly downregulated in all patient samples (**Figures 2 and S13**). Because of that we decided to investigate various subcategories and we found that there was a striking and uniform upregulation of factors supporting alternative splicing while constitutive splicing factors involved in spliceosome assembly and cis splicing are downregulated (**Figures 5 and S14**). Expression changes of as little as 10% (log2FC > 0.1) have been shown to result in biologically relevant changes for vital proteins like kinases and splicing factors [60]. We thus assume that up- and down-regulation of whole spliceosome subcomplexes even in low log2FC ranges lead to significant splicing misregulation. Notably, snRNAs of the U1 spliceosomal subcomplex, responsible for 5' SS recognition, constitute as much as 20% of all downregulated splicing factors (log2FC > |0.1|, RNU1s and RNVU1s). Interestingly, a similar sharp cut-off between alternative and constitutive splicing has been reported in myotonic dystrophy (DM1/DM2) with similar genes being affected, namely *CELF*s, *MBNL*s, *NOVA*, *SMN1/2* and *SF3A1*, among others [61]. DM1 is one of the best studied splicing diseases and shares typical muscular dystrophy symptomology with EDMD. Moreover, a number of splicing changes in DM1 and DM2 also occur in other muscular dystrophies [62]. Of note, *MBNL3* is 4-fold transcriptionally upregulated in EDMD compared to controls. Its protein product impairs muscle cell differentiation in healthy muscle and thus needs to be downregulated upon differentiation onset [63].

Next, we performed splicing analysis to determine whether mis-splicing could drive some of the pathway alterations observed in the EDMD samples. For this purpose, we used three different methods: DEXSeq analyses exon usage, rMATS provides information about

the five most common alternative splicing (AS) events and isoform Switch Analyzer (ISA) indicates which splicing events lead to annotated isoform switches. This revealed varying amounts of alternatively spliced genes in all samples and the three subgroups (**Figure 5**). Since every method focuses on a different event type/ aspect of splicing, a higher amount of unique than overlapping genes is to be expected. Accordingly, an overlap of all three methods indicates genes with exon skipping events that lead to annotated isoform switches. The numbers of mis-spliced genes overlapping between the three algorithms was only 1 gene, *ZNF880*, in G1. In contrast, when analyzing Group 1, 2, and 3 separately, each patient grouping had many mis-spliced genes identified by all three algorithms with 18 mis-spliced genes in the intersect for Group 1, the group including half of the patients, and as much as 95 in Group 3 (**Table S4**). These genes include Nesprin 3 (*SYNE3*), the splicing factor kinase *CLK1* and the chromatin regulator *HMGN3*, all of which are potentially contributing to EDMD, given their functions. All results can be found in **Table S5**. The rMATS analysis includes five AS events: exon skipping (SE), intron retention (RI), mutually exclusive exons (MXE), alternative 3' splice site (A3SS) and alternative 5' splice site (A5SS) usage. Using this comprehensive dataset, we searched for AS events that are significantly differentially used (percent-spliced-in-(psi)-value > |0.1| and p-value < 0.05, **Table S5**). Comparing AS event inclusion between control and EDMD samples, we find thrice as much intron retention in EDMD, while all other events are similarly included as excluded. We hypothesize that this could be a result of downregulated U1 snRNAs which are necessary for proper spliceosome assembly. Supporting the likely importance of the splicing pathway to the EDMD pathomechanism, pathway analysis on these genes revealed a strong enrichment for pathways associated with metabolism, gene expression and the cytoskeleton (**Figure 5**). Moreover, Group 1 and Group 3 display an enrichment for myogenesis and muscle contraction. Using a custom-made set of genes either specific or relevant for muscle development and structure (see methods), we then scanned all significant and differential AS events (**Figures 5 and S15**). Notably, mis-splicing led to the absence of many muscle-specific splice variants (**Figure 5**), among them vital muscle structural genes like *TTN*,

556 *TNNT3*, *NEB*, *ACTA4* and *OBSCN* as well as developmental regulators of the MEF2 family.
557 Importantly, many of them are linked to a variety of muscular dystrophies. For example,
558 *TTN*, *CAPN3*, *PLEC*, and *SGCA* are linked to Limb-Girdle muscular dystrophy [64-69], *DMD* is
559 linked to Duchenne Muscular Dystrophy and Becker muscular dystrophy [70; 71], and BIN1,
560 TNNT2/3 and MBNL1 are mis-spliced in myotonic dystrophy [72].

561 Intriguingly, one of the mis-spliced genes which also displays an isoform switch in
562 Group 3 is *TMEM38A* that has been linked to EDMD [10]. The altered splicing map for
563 *TMEM38A* reveals that not only is its expression highly elevated in EDMD patients (log2FC
564 = 2.9), but also that the protein-coding isoform displays a higher usage relative to
565 abundance compared to the non-coding isoform (**Figure 5**) Many other notable mis-spliced
566 genes are involved in myotube fusion such as the previously mentioned *NID1* that is under
567 spatial genome positioning control of NET39, another of the genome organizing NETs
568 causative of EDMD. Most compellingly, three mis-spliced genes having to do with
569 myogenesis/ myotube fusion had muscle-specific splice variants absent in all patients (G1
570 rMATS). These were *CLCC1* whose loss yields muscle myotonia [73], *HLA-A/B* that
571 disappears during myogenesis and is linked as a risk factor for idiopathic inflammatory
572 myopathies [74; 75], and *SMAD2* that shuts down myoblast fusion [76]. The above examples were
573 found in all patients within a particular Group, but not always amongst all patients from the
574 study or determined by all algorithms; however, there were also some mis-spliced genes
575 that are potentially even more interesting because they were mis-spliced in all patients and
576 with all 3 algorithms yielding the same results. One of these was *ZNF880*. While overall
577 transcript numbers remained similar, the isoform predominantly expressed in control cells,
578 ENST00000422689, is strongly downregulated in Group 1 while the shorter isoform,
579 ENST00000600321, is strongly upregulated (**Figure 5**). Interestingly, the dominant isoform
580 in EDMD loses the zinc finger domain (light and dark blue) and is left with the repressive
581 KRAB domain (black). Little is known about ZNF880 except that it has an unclear role in
582 breast and rectal cancer [77; 78] and additional experiments are necessary to elucidate its role
583 in EDMD.

584

585 *miRNA-Seq Analysis of EDMD Patient Cells*

586 Changes in miRNA levels have been observed in a number of muscular dystrophies and are
587 often used as biomarkers [79; 80], but a comprehensive investigation of miRNA levels in EDMD
588 has thus far not been engaged. Thus, the *in vitro* differentiated EDMD patient cells used for
589 the preceding analysis were also analyzed by miRNA-Seq. We identified 28 differentially
590 expressed miRNAs with some variation among patients (**Figure 6**). We extracted their
591 putative targets from the miRDB database (web resources) and selected those targets
592 whose expression changed in the opposite direction of the miRNAs. Pathway analysis
593 revealed misregulation of miRNAs largely associated with the same pathways that were
594 misregulated from the RNA-Seq data e.g. metabolism, ECM/ fibrosis, and signaling/
595 differentiation (**Figure 6 and Table S6**). More specifically for metabolism 9 of the
596 misregulated miRNA were linked to metabolic functions and with only partial overlap another
597 9 linked to mitochondria function, for ECM/ fibrosis 19 of the misregulated miRNAs were
598 linked to ECM and again with only partial overlap 10 to fibrosis and 13 to cytokines and
599 inflammation. As noted before the ECM category in addition to potentially contributing to
600 fibrosis is also relevant for myotube fusion along with cell cycle regulation that was targeted
601 by 13 misregulated miRNAs and myogenesis that was targeted by 5 miRNAs. Several
602 misregulated miRNAs had functions relating to alternative differentiation pathways with 4
603 relating to adipogenesis, 12 to neurogenesis, 17 to angiogenesis and for signaling there
604 were 9 misregulated miRNAs affecting MAPK pathways, 8 for Akt signaling, 1 for JAK-STAT
605 signaling, 5 for TGF-beta signaling, 2 for Notch signaling, and 3 for TLR signaling.
606 Interestingly, some misregulated miRNAs were also reported as being linked to disease
607 states such as miR-140-3p to dilated cardiomyopathy through its repressive effect on the
608 integrin metalloproteinase gene *ADAM17* [81]. As well as working within cells, miRNAs are
609 often detected within a circulating exosomal microvesicle population that can be harvested
610 from blood serum. This makes them especially attractive as potential biomarkers when

compared to more invasive biopsies, but a much larger sample size together with more clinical informaiton will be required to clarify these as biomarkers.

EDMD Gene Expression Signature Suggests Relationships to Other Muscular Dystrophies

The earlier Bakay study analyzed patient samples from other muscular dystrophies for comparison to EDMD. We re-analyzed their data and used GSEA to determine how related the different muscular dystrophies are to our data. In order to visualize the results, for each disease we plotted the GSEA normalised enrichment score (NES) against the -log10(p-value). The higher on the y-axis the higher the confidence, while positive and negative correlations appearing to the right or to the left of the vertical, respectively. Thus, the further to the upper right the various Bakay disease sets are, the closer it is to the data derived from our set of 10 patients. When we looked at all patients as a single group, EDMD was the best match (**Figure 7**), with the highest score and lowest p-value (FDR 0.001). This indicates that despite the differences in the individual DE genes between the mature muscle data from Bakay's EDMD geneset and our early *in vitro* differentiation geneset, a clear EDMD gene expression signature was displayed in our 10 patients. The next best match is Limb-Girdle muscular dystrophy 2A (LGMD2A), which is particularly interesting because *LMNA* mutations also cause Limb-Girdle muscular dystrophy 2B (LGMD2B) and this was even further distal in pathway analysis signatures from both our EDMD and the Bakay EDMD patients than Fascioscapulohumeral muscular dystrophy (FSHD) and Duchenne Muscular Dystrophy (DMD) patients. The differences in gene signatures that broke down the 10 patients into three EDMD patient subgroups could reflect an underlying cause of clinical disease spectrum or indicate that a group may not be adequately classified as EDMD. Therefore, we performed the same GSEA analysis on each subgroup separately. Group 1, which had both classic emerin and lamin A EDMD mutations showed an even better match with the Bakay EDMD group which was again very close to LGMD2A but also to DMD, Becker Muscular Dystrophy (BMD), FSHD and Limb-Girdle muscular dystrophy 2I (LGMD2I) (**Figure 7**). LGMD2B was still separate and closer to Juvenile Dermatomyositis (JDM). For

Group 2, none of the diseases matched at FDR 5%, although the Bakay EDMD set remained the most like our set. Interestingly, two diseases exhibited an anti-correlation: DMD and BMD, which are both caused by mutations in the dystrophin gene *DMD*. In contrast, Group 3 appeared to be the most distinct and in many ways opposite to Group 1, which is a pattern that was often observed in the functional gene subsets analyzed (**Figures S5-S11**). Group 3 was anticorrelated with EDMD and most of the other muscular dystrophies, while the neurogenic Amyotrophic Lateral Sclerosis (ALS) appeared as the best match, possibly suggesting a neuronal bias in this group (**Figure 7**). This is further supported by the appearance of axonal neuropathy, ataxia, undergrowth, and speech problems in one of the two patients from this group (patient 4, *SYNE1*; **Table 1**), while none of the others exhibited any signs of neuropathy.

The relationship of the patient Groups segregated by gene signatures to potential differences in clinical presentation is underscored by the functional pathways enriched in each group over the others (**Figure 7**). Group 1 showed a strong enrichment of pathways associated with ECM and fibrosis, such as interferon signaling, TNF signaling, ECM organization, ECM proteoglycans, integrin cell surface interactions, collagen formation, and signaling by PDGF all upregulated. Adipogenesis was also particularly promoted in Group 1 compared to the others, and cardiac conduction defects were also highlighted. Group 2 was more uniquely associated with Hippo signaling and BMP2-WNT4-FOXO1 pathway and had fewer links to ECM and fibrosis. Group 3 was more uniquely associated with metabolism, particularly upregulation of oxidative phosphorylation, mitochondrial biogenesis, glucagon signaling pathway, gluconeogenesis, glycolysis and gluconeogenesis, metabolism, TP53 regulates metabolic genes, and thermogenesis pathways. This would suggest that Group 1 pathophysiology may have more characteristics of fibrosis and altered myofibers while Group 2 may have more differentiation or mechanosignaling defects and Group 3 more metabolic defects (**Figure 7**).

Discussion

Attempts to identify the EDMD pathomechanism or clinical biomarkers purely through gene expression signatures are limited because there is too little uniformity in differential gene expression between all patients. We therefore engaged a functional pathway analysis using *in vitro* differentiated myotubes derived from 10 unrelated EDMD patients with known mutations in 7 EDMD-linked genes. While it is difficult to detect many individual genes that were uniformly changed in all patients, we found many pathways that were affected in all patients. Thus, although different genes may have been targeted in different patients, the same functional pathway would be disrupted and thus yield a pathology with similar clinical features. Many pathways were disrupted when we re-analyzed data from the previously published Bakay study and we postulated that, as they just analyzed mutations in two of the over two dozen genes linked to EDMD, analyzing a larger set of linked genes might narrow down the number of pathways to highlight those most relevant to EDMD pathophysiology. Indeed, when we considered a wider set of patients with mutations in 7 different genes the set of affected pathways narrowed to the point that we could identify four likely umbrella pathways.

These four umbrella pathways all make sense for contributing to or even driving the EDMD pathomechanism [82]. Disruption of metabolism pathways from the gene expression analysis was consistent with the significantly reduced glycolysis and mitochondrial respiration output we showed in patient myoblasts compared to controls and it makes sense that this could lead to fatigue, weakness, and muscle atrophy. ECM changes and fibrosis pathways are consistent with pathology observed in EDMD and similarly could drive some of the initial pathology and, as fibrosis accumulates, contribute to disease progression. De-repression of genes from alternate differentiation pathways and defects in myogenesis through disrupted signaling pathways and cell cycle regulation could generate aberrant myotubes to yield pathology. Finally, the last disrupted pathway of splicing yields a loss of muscle-specific splice variants that could impact on all three preceding pathways.

There is much scope for intersection between the four highlighted pathways altered in all sampled EDMD patient cells. For example, amongst the de-repressed differentiation pathways was adipogenesis that could also impact on the metabolism pathway. Even amongst the few genes that were uniformly altered in all patients sampled, though not originally obvious, a more detailed reading of the literature leads to intersections with these pathways. For example, while the *MYH14* general upregulation did not make obvious sense for muscle defects since it is not part of the contractile machinery, it has been shown that a mutation in *MYH14* disrupts mitochondrial fission in peripheral neuropathy [83]. Thus, *MYH14* could potentially feed into the mitochondrial deficits noted in the patient cells. Many of the miRNAs found to be altered in the patients feed into several of these pathways. For example, miR-2392 that is increased in all patients downregulates oxidative phosphorylation in mitochondria [84] but at the same time also is reported to promote inflammation [85]. miR-140 that is up in all groups has roles in fibrosis through collagen regulation [86], is pro-adipogenic [87], and inhibits skeletal muscle glycolysis [88]. miRNAs can also be used potentially prognostically between the different groups as for example miR-146a is upregulated in Group 1, unchanged in Group 2, and downregulated in Group 3. This miRNA has a strong effect on inflammation and has been implicated in fibrosis in the heart [89]. Because there is so much functional overlap between miRNA targets and the pathways noted from the RNA-Seq analysis, it is unclear to what extent the gene expression changes observed could be indirect from the misregulated miRNAs. Nonetheless, there are 4 core functions targeted by multiple mechanisms that we argue are likely to be central to the core EDMD pathomechanism. Interestingly, the literature is filled with many examples of mutation or loss of different splicing factors causing muscle defects though no individual misspliced gene was identified as mediating these effects. Similarly, in myotonic dystrophy type 1 (DM1) there are many mis-spliced genes thought to contribute to the disease pathology [72]. For example, the splicing factor SRSF1 that is down in most patients is important for neuromuscular junction formation in mice [90].

How so many genes become misregulated has not been experimentally proven, but for lamin, emerin, Sun1, nesprin, TMEM214 and PLPP7/NET39, the fact that mutations to all individually yield many hundreds of gene expression changes with considerable overlap strongly suggests that they function in a complex at the nuclear envelope to direct genome organization. It has already been shown that knockdown of Tmem214 and NET39 as well as several other muscle-specific NETs each alters the position and expression of hundreds of genes [17]. Separately it was found that lamin B1 and the NET LAP2beta function together with two other proteins in a complex involved in tethering genes to the nuclear envelope [91] in fibroblasts and that emerin and lamins similarly function together with other proteins to tether genes in muscle cells [92]. Thus, disruption of emerin, lamin A or any other component of these tethering complexes could yield sufficiently similar gene/ pathway expression changes to yield the core characteristic clinical features of EDMD. We propose that the different muscle-specific NETs give specificity to such a complex containing lamin A and emerin and that Sun1 and nesprin proteins can impact on these complexes through mediating mechanosignal transduction and FHL1 in interpreting such signals. Since 15% of all genes changing here were affected by at least one of the muscle-specific genome-organizing NETs that were tested by knockdown, this would provide a core set of genome organization and expression changes to cause the core EDMD pathology. Since the majority of genes affected by each NET tested was unique to that NET with the exception of Tmem214, this could account for other gene expression changes that drive the segregation into subgroups which could contribute to clinical variation. This interpretation is consistent with the numbers of genes changing for mutations in different nuclear envelope proteins (**Figure 1**). The patient cells with mutations in genome organizing NETs had much fewer genes changing than the patient cells with lamin A mutations. This could be because the lamin A mutations could disrupt multiple genome tethering complexes for different genome organizing NETs and thus alter expression of more genes. Alternatively, different lamin A mutations could preferentially yield disruption of complexes with particular genome-organizing NETs and thus patients with the same complex disrupted might share gene expression signature

747 changes while those with different complexes disrupted might have less overlap so that this
748 could account for different lamin A mutations segregating into different gene expression
749 subgroups. In either case, the extreme differences in lamin A mutations gene expression
750 profiles is not entirely surprising as different lamin mutations also exhibited large differences
751 in studies of nuclear mechanics [93]; so this could also impact on mechanosignal transduction.
752 The Sun1 mutation may have affected fewer genes because its function in mechanosignal
753 transduction is redundant with Sun2 while the Nesprin 1 mutation had more genes changing
754 because it is more central to both mechanosignal transduction and to cell and nuclear
755 mechanical stability.

756 The FHL1 mutations add another level of complexity to EDMD as there are several
757 splice variants of FHL1 and only the B variant (ENST00000394155) targets to the nuclear
758 envelope [94]. That EDMD is a nuclear envelope disorder is underscored by the fact that none
759 of the FHL1 mutations occur in exons found in the much shorter C variant
760 (ENST00000618438) and the patient 8 mutation p.V280M is in an exon unique to FHL1B.
761 Thus, the nuclear envelope splice variant is the only one that could yield pathology in all
762 patients, though some of the variation could come from one of the patients also expressing
763 the mutant A splice variant (ENST00000543669).

764

765 **Conclusions**

766 While further work is needed to validate the correlations between the gene
767 expression profile subgroupings and their clinical presentation and disease progression, our
768 finding of such distinct gene expression signatures amongst clinically diagnosed EDMD
769 patients argues that the currently used clinical phenotype spectrum umbrella of the EDMD
770 classification may be too broad and it might be reclassified in more precise subtypes. What
771 is clear is that the original classifications of EDMD subtypes based just on the mutated gene
772 often allows for cases with very dissimilar gene signatures to be classified together while
773 similar gene signature cases are classified as separate classes. The two lamin A mutations
774 yielded changes in gene expression signatures that were far more different from one another

than the group 1 lamin A mutation gene signature was from the TMEM214, NET39, emerin, and FHL1 mutation gene signatures. Similarly, the three FHL1 mutations yielded greater differences between them than these many genes in Groups 1 and 2. Thus, EDMD might be better classified by similarities in gene expression signatures than by the particular gene mutated. Regardless, these subtypes or separate disorders have distinctive gene expression signatures and miRNA signatures that could be used as biomarkers both diagnostically and perhaps prognostically. To get to that point will require a more comprehensive modern description of clinical Gestalt phenotypes including *e.g.:* imaging datasets and disease progression timelines deciphering unique groups. Importantly, the different pathways we found enriched for in each subgroup could be converted to clinical recommendations based on the much more conserved individual gene expression changes for each subgroup (**Figure 7**). For example, EDMD patients have been considered by some clinicians to be at risk for malignant hyperthermia [95], though a consensus was never achieved. Our data show that the 3 genes currently associated with malignant hyperthermia (*RYR1*, *CACNA1S*, and *STAC3*) are all misregulated in Groups 1 and 3, but not Group 2 (**Figure S16**). Thus, checking expression of these genes might indicate whether a patient is likely to be at risk or not. Finally, an additional new aspect coming up from our datasets is that it might be worth further investigating the role of splicing in muscle differentiation because it might be of wider relevance to muscular dystrophy beyond DM and EDMD.

Data availability

Bakay *et al.* muscular dystrophy dataset is available at NCBI GEO with accession GSE3307. RNA-Seq and miRNA-Seq datasets will be deposited at NCBI GEO and made publicly available prior to publication.

Acknowledgements

This work was funded by Muscular Dystrophy UK 18GRO-PG24-0248 and Medical Research Council MR/R018073 to ECS and Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 470092532 to PM.

Web resources

g:Profiler, https://biit.cs.ut.ee/gprofiler/gost

TissueEnrich, https://tissueenrich.gdcb.iastate.edu/

miRDB, http://mirdb.org

'smrnaseq' miRNA-Seq analysis pipeline, https://nf-co.re/smrnaseq/1.1.0

miRTop, https://github.com/miRTop

NCBI GEO, https://www.ncbi.nlm.nih.gov/geo/

Author Contributions

JIH processed patient samples for RNA- and miRNA-Seq and analyzed the data with assistance from SW. VT performed splicing analysis. LK-B, SH, and PM performed various metabolic analyses. RC helped with generation of figures and critical discussion. BS provided patient samples and inspiration. ECS designed the study and wrote the manuscript.

References

1. Norwood, F.L.M., Harling, C., Chinnery, P.F., Eagle, M., Bushby, K., and Straub, V. (2009). Prevalence of genetic muscle disease in Northern England: in-depth analysis of a muscle clinic population Brain 132, 3175-3186.

2. Bonne, G., Mercuri, E., Muchir, A., Urtizberea, A., Becane, H.M., Recan, D., Merlini, L., Wehnert, M., Boor, R., Reuner, U., et al. (2000). Clinical and molecular genetic spectrum of autosomal dominant Emery-Dreifuss muscular dystrophy due to mutations of the lamin A/C gene. Ann Neurol 48, 170-180.

3. Mercuri, E., Poppe, M., Quinlivan, R., Messina, S., Kinali, M., Demay, L., Bourke, J., Richard, P., Sewry, C., Pike, M., et al. (2004). Extreme variability of phenotype in patients with an identical missense mutation in the lamin A/C gene: from congenital onset with severe phenotype to milder classic Emery-Dreifuss variant. Arch Neurol 61, 690-694.

4. Rankin, J., Auer-Grumbach, M., Bagg, W., Colclough, K., Nguyen, T.D., Fenton-May, J., Hattersley, A., Hudson, J., Jardine, P., Josifova, D., et al. (2008). Extreme phenotypic diversity and nonpenetrance in families with the LMNA gene mutation R644C. Am J Med Genet A 146A, 1530-1542.

5. Hoeltzenbein, M., Karow, T., Zeller, J.A., Warzok, R., Wulff, K., Zschiesche, M., Herrmann, F.H., Grosse-Heitmeyer, W., and Wehnert, M.S. (1999). Severe clinical expression in X-linked Emery-Dreifuss muscular dystrophy. Neuromuscul Disord 9, 166-170.

6. Emery, A.E., and Dreifuss, F.E. (1966). Unusual type of benign x-linked muscular dystrophy. J Neurol Neurosurg Psychiatry 29, 338-342.

7. Emery, A.E. (1989). Emery-Dreifuss syndrome. J Med Genet 26, 637-641.

8. Knoblauch, H., Geier, C., Adams, S., Budde, B., Rudolph, A., Zacharias, U., Schulz-Menger, J., Spuler, A., Yaou, R.B., Nurnberg, P., et al. (2010). Contractures and hypertrophic cardiomyopathy in a novel FHL1 mutation. Ann Neurol 67, 136-140.

9. Bonne, G., and Quijano-Roy, S. (2013). Emery-Dreifuss muscular dystrophy, laminopathies, and other nuclear envelopathies. Handb Clin Neurol 113, 1367-1376.

10. Meinke, P., Kerr, A.R.W., Czapiewski, R., de Las Heras, J.I., Dixon, C.R., Harris, E., Kolbel, H., Muntoni, F., Schara, U., Straub, V., et al. (2020). A multistage sequencing strategy pinpoints novel candidate alleles for Emery-Dreifuss muscular dystrophy and supports gene misregulation as its pathomechanism. EBioMedicine 102587, doi10.1016.

11. Meinke, P. (2018). Molecular Genetics of Emery–Dreifuss Muscular Dystrophy. eLS https://doi.org/10.1002/9780470015902.a0022438.pub2.

12. Bridger, J.M., Foeger, N., Kill, I.R., and Herrmann, H. (2007). The nuclear lamina. Both a structural framework and a platform for genome organization. FEBS J 274, 1354-1361.

13. Worman, H.J., and Bonne, G. (2007). "Laminopathies": a wide spectrum of human diseases. Exp Cell Res 313, 2121-2133.

14. Meinke, P., Mattioli, E., Haque, F., Antoku, S., Columbaro, M., Straatman, K.R., Worman, H.J., Gundersen, G.G., Lattanzi, G., Wehnert, M., et al. (2014). Muscular dystrophy-associated SUN1 and SUN2 variants disrupt nuclear-cytoskeletal connections and myonuclear organization. PLoS Genet 10, e1004605.

15. Zhang, Q., Bethmann, C., Worth, N.F., Davies, J.D., Wasner, C., Feuer, A., Ragnauth, C.D., Yi, Q., Mellad, J.A., Warren, D.T., et al. (2007). Nesprin-1 and -2 are involved in the pathogenesis of Emery-Dreifuss Muscular Dystrophy and are critical for nuclear envelope integrity. Hum Mol Genet. 16, 2816-2833.

16. Piccus, R., and Brayson, D. (2020). The nuclear envelope: LINCing tissue mechanics to genome regulation in cardiac and skeletal muscle. Biol Lett 16, 20200302.

17. Robson, M.I., de Las Heras, J.I., Czapiewski, R., Le Thanh, P., Booth, D.G., Kelly, D.A., Webb, S., Kerr, A.R., and Schirmer, E.C. (2016). Tissue-Specific Gene Repositioning by Muscle Nuclear Membrane Proteins Enhances Repression of Critical Developmental Genes during Myogenesis. Mol Cell 62, 834-847.

18. Wilkie, G.S., Korfali, N., Swanson, S.K., Malik, P., Srsen, V., Batrakou, D.G., de las Heras, J., Zuleger, N., Kerr, A.R., Florens, L., et al. (2011). Several novel nuclear envelope transmembrane proteins identified in skeletal muscle have cytoskeletal associations. Mol Cell Proteomics 10, M110 003129.

19. Zuleger, N., Boyle, S., Kelly, D.A., de Las Heras, J.I., Lazou, V., Korfali, N., Batrakou, D.G., Randles, K.N., Morris, G.E., Harrison, D.J., et al. (2013). Specific nuclear envelope transmembrane proteins can promote the location of chromosomes to and from the nuclear periphery. Genome Biol 14, R14.

20. Bakay, M., Wang, Z., Melcon, G., Schiltz, L., Xuan, J., Zhao, P., Sartorelli, V., Seo, J., Pegoraro, E., Angelini, C., et al. (2006). Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb-MyoD pathways in muscle regeneration. Brain 129, 996-1013.

21. Melcon, G., Kozlov, S., Cutler, D.A., Sullivan, T., Hernandez, L., Zhao, P., Mitchell, S., Nader, G., Bakay, M., Rottman, J.N., et al. (2006). Loss of emerin at the nuclear envelope disrupts the Rb1/E2F and MyoD pathways during muscle regeneration. Hum Mol Genet 15, 637-651.

22. Bione, S., Maestrini, E., Rivella, S., Mancini, M., Regis, S., Romeo, G., and Toniolo, D. (1994). Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy. Nat Genet 8, 323-327.

23. Bonne, G., Di Barletta, M.R., Varnous, S., Becane, H.M., Hammouda, E.H., Merlini, L., Muntoni, F., Greenberg, C.R., Gary, F., Urtizberea, J.A., et al. (1999). Mutations in the gene encoding lamin A/C cause autosomal dominant Emery- Dreifuss muscular dystrophy. Nat Genet 21, 285-288.

24. Gueneau, L., Bertrand, A.T., Jais, J.P., Salih, M.A., Stojkovic, T., Wehnert, M., Hoeltzenbein, M., Spuler, S., Saitoh, S., Verschueren, A., et al. (2009). Mutations of the FHL1 gene cause Emery-Dreifuss muscular dystrophy. Am J Hum Genet 85, 338-353.

25. Ruiz, E., Penrose, H.M., Heller, S., Nakhoul, H., Baddoo, M., Flemington, E.F., Kandil, E., and Savkovic, S.D. (2020). Bacterial TLR4 and NOD2 signaling linked to reduced mitochondrial energy function in active inflammatory bowel disease. Gut Microbes 11, 350-363.

26. Campbell, I.L., Bizilj, K., Colman, P.G., Tuck, B.E., and Harrison, L.C. (1986). Interferon-gamma induces the expression of HLA-A,B,C but not HLA-DR on human pancreatic beta-cells. J Clin Endocrinol Metab 62, 1101-1109.

27. Browne, S.K., Roesser, J.R., Zhu, S.Z., and Ginder, G.D. (2006). Differential IFN-gamma stimulation of HLA-A gene expression through CRM-1-dependent nuclear RNA export. J Immunol 177, 8612-8619.

28. Behrends, C., Sowa, M.E., Gygi, S.P., and Harper, J.W. (2010). Network organization of the human autophagy system. Nature 466, 68-76.

29. Choi, B.-O., Kang, S.H., Hyun, Y.S., Kanwal, S., Park, S.W., Koo, H., Kim, S.-B., Choi, Y.-C., Yoo, J.H., Kim, J.-W., et al. (2011). A complex phenotype of peripheral neuropathy, myopathy, hoarseness, and hearing loss is linked to an autosomal dominant mutation in MYH14. Hum Mutat 32, 669-677.

30. Iyadurai, S., Arnold, D.A., Kissel, J.T., Ruhno, C., McGovern, V.L., Snyder, P.J., Prior, T.W., Roggenbuck, J., Burghes, A.H., and Kolb, S.J. (2017). Variable phenotypic expression and onset in MYH14 distal hereditary motor neuropathy phenotype in a large, multigenerational North American family. Muscle Nerve 56, 341-345.

31. Almutawa, W., Smith, C., Sabouny, R., Smit, R.B., Zhao, T., Wong, R., Lee-Glover, L., Desrochers-Goyette, J., Ilamathi, H.S., Consortium, C.R.C., et al. (2019). The R941L mutation in MYH14 disrupts mitochondrial fission and associates with peripheral neuropathy. EBioMedicine 45, 379-392.

32. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res 47, W191-W198.

928 33. Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and
929 visualizes long lists of gene ontology terms. PLoS One 6, e21800.
930 34. Fisk, B., Ioannides, C.G., Aggarwal, S., Wharton, J.T., O'Brian, C.A., Restifo, N., and
931 Glisson, B.S. (1994). Enhanced expression of HLA-A,B,C and inducibility of TAP-1,
932 TAP-2, and HLA-A,B,C by interferon-gamma in a multidrug-resistant small cell lung
933 cancer line. Lymphokine Cytokine Res 13, 125-131.
934 35. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A.,
935 Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set
936 enrichment analysis: a knowledge-based approach for interpreting genome-wide
937 expression profiles. Proc Natl Acad Sci U S A 102, 15545-15550.
938 36. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and
939 Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics
940 27, 1739-1740.
941 37. Miranda, S., Foncea, R., Guerrero, J., and Leighton, F. (1999). Oxidative Stress and
942 Upregulation of Mitochondrial Biogenesis Genes in Mitochondrial DNA-Depleted
943 HeLa Cells. Biochem Biophys Res Commun 258, 44-49.
944 38. L'honore, A., Commere, P.-H., Negroni, E., Pallafacchina, G., Friguet, B., Drouin, J.,
945 Buckingham, M., and Montarras, D. (2018). The role of Pitx2 and Pitx3 in muscle
946 stem cells gives new insights into P38a MAP kinase and redox regulation of muscle
947 regeneration. Elife 7, e32991.
948 39. Chang, N.C., Nguyen, M., and Shore, G.C. (2012). BCL2-CISD2: An ER complex at the
949 nexus of autophagy and calcium homeostasis? Autophagy 8, 856-857.
950 40. Yeh, C.-H., Chou, Y.-J., Kao, C.-H., and Tsai, T.-F. (2020). Mitochondria and Calcium
951 Homeostasis: Cisd2 as a Big Player in Cardiac Ageing. Int J Mol Sci 21, 9238.
952 41. Meinke, P., Schneiderat, P., Srsen, V., Korfali, N., Le Thanh, P., Cowan, G.J.M.,
953 Cavanagh, D.R., Wehnert, M., Schirmer, E.C., and Walter, M.C. (2015). Abnormal
954 proliferation and spontaneous differentiation of myoblasts from a symptomatic female

carrier of X-linked Emery-Dreifuss muscular dystrophy. Neuromuscul Disord 25, 127-136.

42. Jöbsis, G.J., Keizers, H., Vreijling, J.P., de Visser, M., Speer, M.C., Wolterman, R.A., Baas, F., and Bolhuis, P.A. (1996). Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. Nat Genet 14, 113-115.

43. Lampe, A.K., Dunn, D.M., von Niederhausern, A.C., Hamil, C., Aoyagi, A., Laval, S.H., Marie, S.K., Chu, M.-L., Swoboda, K., Muntoni, F., et al. (2005). Automated genomic sequence analysis of the three collagen VI genes: applications to Ullrich congenital muscular dystrophy and Bethlem myopathy. J Med Genet 42, 108-120.

44. Baker, N.L., Morgelin, M., Pace, R.A., Peat, R.A., Adams, N.E., McKinlay Gardner, R.J., Rowland, L.P., Miller, G., De Jonghe, P., Ceulemans, B., et al. (2007). Molecular consequences of dominant Bethlem myopathy collagen VI mutations. Ann Neurol 62, 390-405.

45. Zou, Y., Zwolanek, D., Izu, Y., Gandhy, S., Schreiber, G., Brockmann, K., Devoto, M., Tian, Z., Hu, Y., Veit, G., et al. (2014). Recessive and dominant mutations in COL12A1 cause a novel EDS/myopathy overlap syndrome in humans and mice. Hum Mol Genet 23, 2339-2352.

46. Kaar, J.L., Li, Y., Blair, H.C., Asche, G., Koepsel, R.R., Huard, J., and Russell, A.J. (2008). Matrix metalloproteinase-1 treatment of muscle fibrosis. Acta Biomater 4, 1141-1120.

47. Zhu, X., Yao, T., Wang, R., Guo, S., Wang, X., Zhou, Z., Zhang, Y., Zhuo, X., Wang, R., Li, J.Z., et al. (2020). IRF4 in Skeletal Muscle Regulates Exercise Capacity via PTG/Glycogen Pathway. Adv Sci (Weinh) 7, 2001502.

48. Chiba, K., Tsuchiya, M., Koide, M., Hagiwara, Y., Sasaki, K., Hattori, Y., Watanabe, M., Sugawara, S., Kanzaki, M., and Endo, Y. (2015). Involvement of IL-1 in the Maintenance of Masseter Muscle Activity and Glucose Homeostasis PLoS One 10, e0143635.

982 49. Fairley, E., Riddell, A., Ellis, J., and Kendrick-Jones, J. (2002). The cell cycle dependent
983 mislocalisation of emerin may contribute to the Emery-Dreifuss muscular dystrophy
984 phenotype. J Cell Sci 115, 341-354.
985 50. Mademtzoglou, D., Asakura, Y., Borok, M.J., Alonso-Martin, S., Mourikis, P., Kodaka, Y.,
986 Mohan, A., Asakura, A., and Relaix, F. (2018). Cellular localization of the cell cycle
987 inhibitor Cdkn1c controls growth arrest of adult skeletal muscle stem cells. eLife 7,
988 e33337.
989 51. Caruelle, D., Mazouzi, Z., Husmann, I., Delbe, J., Duchesnay, A., Gautron, J., Martelly,
990 I., and Courty, J. (2004). Upregulation of HARP during in vitro myogenesis and rat
991 soleus muscle regeneration. J Muscle Res Cell Motility 25, 45-53.
992 52. Dietrich, S., Abou-Rebyeh, F., Brohmann, H., Bladt, F., Sonnenberg-Riethmacher, E.,
993 Yamaai, T., Lumsden, A., Brand-Saberi, B., and Birchmeier, C. (1999). The role of
994 SF/HGF and c-Met in the development of skeletal muscle. Development 126, 1621-
995 1629.
996 53. Mousavi, K., and Jasmin, B.J. (2006). BDNF is expressed in skeletal muscle satellite
997 cells and inhibits myogenic differentiation. J Neurosci 26, 5739-5749.
998 54. Peng, H.B., Ali, A.A., Dai, Z., Daggett, D.F., Raulo, E., and Rauvala, H. (1995). The role
999 of heparin-binding growth-associated molecule (HB-GAM) in the postsynaptic
1000 induction in cultured muscle cells. J Neurosci 15, 3027-3038.
1001 55. Stark, D.A., Karvas, R.M., Siegel, A.L., and Cornelison, D.D. (2011). Eph/ephrin
1002 interactions modulate muscle satellite cell motility and patterning. Development 138,
1003 5279-5289.
1004 56. Singh, R., Artaza, J.N., Taylor, W.E., Gonzalez-Cadavid, N.F., and Bhasin, S. (2003).
1005 Androgens stimulate myogenic differentiation and inhibit adipogenesis in C3H 10T1/2
1006 pluripotent cells through an androgen receptor-mediated pathway. Endocrinology
1007 144, 5081-5088.
1008 57. Bryan, B.A., Mitchell, D.C., Zhao, L., Ma, W., Stafford, L.J., Teng, B.-B., and Liu, M.
1009 (2005). Modulation of muscle regeneration, myogenesis, and adipogenesis by the

Rho family guanine nucleotide exchange factor GEFT. Mol Cell Biol 25, 11089-11101.

58. Czapiewski, R., Batrakou, D.G., de Las Heras, J.I., Carter, R., Sivakumar, A., Sliwinska, M., Dixon, C.R., Webb, S., Lattanzi, G., Morton, N., et al. (2022). Genomic loci mispositioning in Tmem120a knockout mice yields latent lipodystrophy. Nat Commun 13, 321 doi: 310.1038/s41467-41021-27869-41462.

59. Oldenburg, A.R., Delbarre, E., Thiede, B., Vigouroux, C., and Collas, P. (2014). Deregulation of Fragile X-related protein 1 by the lipodystrophic lamin A p.R482W mutation elicits a myogenic gene expression program in preadipocytes. Hum Mol Genet 23, 1151-1162.

60. Anczuków, O., Rosenberg, A.Z., Akerman, M., Das, S., Zhan, L., Karni, R., K., M.S., and Krainer, A.R. (2012). The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. Nat Struct Mol Biol 19, 220-228.

61. Todorow, V., Hintze, S., Kerr, A.R.W., Hehr, A., Schoser, B., and Meinke, P. (2021). Transcriptome Analysis in a Primary Human Muscle Cell Differentiation Model for Myotonic Dystrophy Type 1. Int J Mol Sci 22, 8607.

62. Bachinski, L.L., Baggerly, K.A., Neubauer, V.L., Nixon, T.J., Raheem, O., Sirito, M., Unruh, A.K., Zhang, J., Nagarajan, L., Timchenko, L.T., et al. (2014). Most expression and splicing changes in myotonic dystrophy type 1 and type 2 skeletal muscle are shared with other muscular dystrophies. Neuromuscul Disord 24, 227-240.

63. Lee, K.-S., Cao, Y., Witwicka, H.E., Tom, S., Tapscott, S.J., and Wang, E.H. (2010). RNA-binding protein Muscleblind-like 3 (MBNL3) disrupts myocyte enhancer factor 2 (Mef2) {beta}-exon splicing. J Biol Chem 285, 33779-33787.

64. Hackman, P., Vihola, A., Haravuori, H., Marchand, S., Sarparanta, J., De Seze, J., Labeit, S., Witt, C., Peltonen, L., Richard, I., et al. (2002). Tibial muscular dystrophy is a titinopathy caused by mutations in TTN, the gene encoding the giant skeletal-muscle protein titin. Am J Hum Genet 71, 492-500.

65. Penisson-Besnier, I., Hackman, P., Suominen, T., Sarparanta, J., Huovinen, S., Richard-Cremieux, I., and Udd, B. (2010). Myopathies caused by homozygous titin mutations: Limb-girdle muscular dystrophy 2J and variations of phenotype. J Neurol Neurosurg Psychiatry 81, 1200-1202.

66. Gundesli, H., Talim, B., Korkusuz, P., Balci-Hayta, B., Cirak, S., Akarsu, N.A., Topaloglu, H., and Dincer, P. (2010). Mutation in exon 1f of PLEC, leading to disruption of plectin isoform 1f, causes autosomal-recessive limb-girdle muscular dystrophy. Am J Hum Genet 87, 834-841.

67. Richard, I., Broux, O., Allamand, V., Fougerousse, F., Chiannilkulchai, N., Bourg, N., Brenguier, L., Devaud, C., Pasturaud, P., Roudaut, C., et al. (1995). Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell 81, 27-40.

68. Vissing, J., Barresi, R., Witting, N., Van Ghelue, M., Gammelgaard, L., Bindoff, L.A., Straub, V., Lochmuller, H., Hudson, J., Wahl, C.M., et al. (2016). A heterozygous 21-bp deletion in CAPN3 causes dominantly inherited limb girdle muscular dystrophy Brain 139, 2154-2163.

69. Roberds, S.L., Leturcq, F., Allamand, V., Piccolo, F., Jeanpierre, M., Anderson, R.D., Lim, L.E., Lee, J.C., Tome, F.M., Romero, N.B., et al. (1994). Missense mutations in the adhalin gene linked to autosomal recessive muscular dystrophy. Cell 78, 625-633.

70. Hoffman, E.P., Brown Jr, R.H., and Kunkel, L.M. (1987). Dystrophin: the protein product of the Duchenne muscular dystrophy locus. Cell 51, 919-928.

71. Kunkel, L.M., Hejtmancik, J.F., Caskey, C.T., Speer, A., Monaco, A.P., Middlesworth, W., Colletti, C.A., Bertelson, C., Muller, U., Bresnan, M., et al. (1986). Analysis of deletions in DNA from patients with Becker and Duchenne muscular dystrophy. Nature 322, 73-77.

1064 72. López-Martínez, A., Soblechero-Martín, P., de-la-Puente-Ovejero, L., Nogales-Gadea,
1065 G., and Arechavala-Gomeza, V. (2020). An Overview of Alternative Splicing Defects
1066 Implicated in Myotonic Dystrophy Type I. Genes 11, 1109.
1067 73. Jia, Y., Jucius, T.J., Cook, S.A., and Ackerman, S.L. (2015). Loss of Clcc1 results in ER
1068 stress, misfolded protein accumulation, and neurodegeneration. J Neurosci 35, 3001-
1069 3009.
1070 74. Isa, A., Nehlin, J.O., Sabir, H.J., Andersen, T.E., Gaster, M., Kassem, M., and Barington,
1071 T. (2010). Impaired cell surface expression of HLA-B antigens on mesenchymal stem
1072 cells and muscle cell progenitors. PLoS One 5, e10900.
1073 75. O'Hanlon, T.P., Carrick, D.M., Targoff, I.N., Arnett, F.C., Reveille, J.D., Carrington, M.,
1074 Gao, X., Oddis, C.V., Morel, P.A., Malley, J.D., et al. (2006). Immunogenetic risk and
1075 protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw,
1076 -DRB1, and -DQA1 allelic profiles distinguish European American patients with
1077 different myositis autoantibodies. Medicine 85, 111-127.
1078 76. Melendez, J., Sieiro, D., Salgado, D., Morin, V., Dejardin, M.-J., Zhou, C., Mullen, A.C.,
1079 and Marcelle, C. (2021). TGFβ signalling acts as a molecular brake of myoblast
1080 fusion. Nat Commun 12, 749.
1081 77. Yue, L., Wentao, L., Xin, Z., Jingjing, H., Xiaoyan, Z., Na, F., Tonghui, M., and Dalin, L.
1082 (2020). Human epidermal growth factor receptor 2-positive metastatic breast cancer
1083 with novel epidermal growth factor receptor -ZNF880 fusion and epidermal growth
1084 factor receptor E114K mutations effectively treated with pyrotinib: A case report.
1085 Medicine (Baltimore) 99, e23406.
1086 78. Bala, P., Singh, A.K., Kavadipula, P., Kotapalli, V., Sabarinathan, R., and Bashyam,
1087 M.D. (2021). Exome sequencing identifies ARID2 as a novel tumor suppressor in
1088 early-onset sporadic rectal cancer. Oncogene 40, 863-874.
1089 79. Cacchiarelli, D., Legnini, I., Martone, J., Cazzella, V., D'Amico, A., Bertini, E., and
1090 Bozzoni, I. (2011). miRNAs as serum biomarkers for Duchenne muscular dystrophy.
1091 EMBO Mol Med 3, 258-265.

80. Israeli, D., Poupiot, J., Amor, F., Charton, K., Lostal, W., Jeanson-Leh, L., and Richard, I. (2016). Circulating miRNAs are generic and versatile therapeutic monitoring biomarkers in muscular dystrophies. Sci Rep 6, 28097.

81. Fedak, P.W.M., Moravee, C.S., McCarthy, P.M., Altamentova, S.M., Wong, A.P., Skrtic, M., Verma, S., Weisel, R.D., and Li, R.-K. (2006). Altered expression of disintegrin metalloproteinases and their inhibitor in human dilated cardiomyopathy. Circulation 113, 238-245.

82. Lehka, L., and Redowicz, M.J. (2020). Mechanisms regulating myoblast fusion: a multilevel interplay. Semin Cell Dev Biol 104, 81-92.

83. Altmutawa, W., Smith, C., Sabouny, R., Smit, R.B., Zhao, T., Wong, R., Lee-Glover, L., Desrochers-Goyette, J., Ilamathi, H.S., Consortium, C.R.C., et al. (2019). The R941L mutation in MYH14 disrupts mitochondrial fission and associates with peripheral neuropathy EBioMedicine 45, 379-392.

84. Fan, S., Tian, T., Chen, W., Lv, X., Lei, X., Zhang, H., Sun, S., Cai, L., Pan, G., He, L., et al. (2019). Mitochondrial miRNA Determines Chemoresistance by Reprogramming Metabolism and Regulating Mitochondrial Transcription. Cancer Res 79, 1069-1084.

85. Narozna, M., and Rubis, B. (2021). Anti-SARS-CoV-2 Strategies and the Potential Role of miRNA in the Assessment of COVID-19 Morbidity, Recurrence, and Therapy Int J Mol Sci 22, 8663.

86. Liao, W., Liang, P., Liu, B., Xu, Z., Zhang, L., Feng, M., Tang, Y., and Xu, A. (2020). MicroRNA-140-5p Mediates Renal Fibrosis Through TGF-β1/Smad Signaling Pathway by Directly Targeting TGFBR1. Front Physiol 11, 1093.

87. Zhang, X., Chang, A., Li, Y., Gao, Y., Wang, H., Ma, Z., Li, X., and Wang, B. (2015). miR-140-5p regulates adipocyte differentiation by targeting transforming growth factor-β signaling. Sci Rep 5, 18118.

88. Liu, L., Li, T.-M., Liu, X.-R., Bai, Y.-P., Li, J., Tang, N., and Wang, X.-B. (2019). MicroRNA-140 inhibits skeletal muscle glycolysis and atrophy in endotoxin-induced

sepsis in mice via the WNT signaling pathway Am J Physiol Cell Physiol 317, C189-C199.

89. Feng, B., Chen, S., Gordon, A.D., and Chakrabarti, S. (2017). miR-146a mediates inflammatory changes and fibrosis in the heart in diabetes. J Mol Cell Cardiol 105, 70-76.

90. Liu, Y., Luo, Y., Shen, L., Guo, R., Zhan, Z., Yuan, N., Sha, R., Qian, W., Wang, Z., Xie, Z., et al. (2020). Splicing factor SRSF1 is essential for satellite cell proliferation and postnatal maturation of neuromuscular junctions in mice. Stem Cell Rep 15, 941-954.

91. Zullo, J.M., Demarco, I.A., Pique-Regi, R., Gaffney, D.J., Epstein, C.B., Spooner, C.J., Luperchio, T.R., Bernstein, B.E., Pritchard, J.K., Reddy, K.L., et al. (2012). DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. Cell 149, 1474-1487.

92. Demmerle, J., Koch, A.J., and Holaska, J.M. (2013). Emerin and histone deacetylase 3 (HDAC3) cooperatively regulate expression and nuclear positions of MyoD, Myf5, and Pax7 genes during myogenesis. Chromosome Res 21, 765-779.

93. Zwerger, M., Jaalouk, D.E., Lombardi, M.L., Isermann, P., Mauermann, M., Dialynas, G., Herrmann, H., Wallrath, L.L., and Lammerding, J. (2013). Myopathic lamin mutations impair nuclear stability in cells and tissue and disrupt nucleo-cytoskeletal coupling. Hum Mol Genet 22, 2335-2349.

94. Ziat, E., Mamchaoui, K., Beuvin, M., Nelson, I., Azibani, F., Spuler, S., Bonne, G., and Bertrand, A.T. (2016). FHL1B Interacts with Lamin A/C and Emerin at the Nuclear Lamina and is Misregulated in Emery-Dreifuss Muscular Dystrophy. Neuromuscul Disord 3, 497-510.

95. Bonne, G., Leturcq, F., and Yaou, R.B. (2004). Emery-Dreifuss muscular dystrophy. GeneReviews PMID:20301609.

96. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21.

97. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120.

98. Love, M.I., HUber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550.

99. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. Nat Methods 14, 417-419.

100. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43, e47.

101. Reimand, J., Arak, T., Adler, P., Kolberg, L., Reisberg, S., Peterson, H., and Vilo, J. (2016). g:Profiler-a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res 44, W83-89.

102. Jain, A., and Tuteja, G. (2019). TissueEnrich: Tissue-specific gene enrichment analysis. Bioinformatics 35, 1966-1967.

103. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498-2504.

104. Rooney, J.P., Ryde, I.T., Sanders, L.H., Howlett, E.H., Colton, M.D., Germ, K.E., Mayer, G.D., Greenamyre, J.T., and Meyer, J.N. (2015). PCR based determination of mitochondrial DNA copy number in multiple species. Methods Mol Biol 1241, 23-38.

105. Nakhle, J., Ozkan, T., Lněničková, K., Briolotti, P., and Vignais, M.-L. (2020). Methods for simultaneous and quantitative isolation of mitochondrial DNA, nuclear DNA and RNA from mammalian cells. Biotechniques 69, 436-442.

106. Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. Genome Res 22, 2008-2017.

1173 107. Shen, S., Park, J.W., Lu, Z.-x., and Xing, Y. (2014). rMATS: Robust and flexible
1174 detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl
1175 Acad Sci U S A 111, E5593-E5601.
1176 108. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic
1177 RNA-seq quantification. Nat Biotechnol 34, 525-527.
1178 109. Vitting-Seerup, K., and Sandelin, A. (2019). IsoformSwitchAnalyzeR: analysis of
1179 changes in genome-wide patterns of alternative splicing and its functional
1180 consequences. Bioinformatics 35, 4469-4471.
1181 110. Walter, W., Sánchez-Cabo, F., and Ricote, M. (2015). GOplot: an R package for
1182 visually combining expression data with functional analysis. Bioinformatics 31, 2912-
1183 2914.
1184 111. Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J., and Peterson, H. (2020). gprofiler2 -- an
1185 R package for gene list functional enrichment analysis and namespace conversion
1186 toolset g:Profiler. F1000Res 9, ELIXIR-709.
1187 112. Kang, Y.-J., Yang, D.-C., Kong, L., Hou, M., Meng, Y.-Q., Wei, L., and Gao, G. (2017).
1188 CPC2: a fast and accurate coding potential calculator based on sequence intrinsic
1189 features. Nucleic Acids Res 45, W12-W16.
1190 113. Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive
1191 sequence similarity searching. Nucleic Acids Res 39, W29-37.
1192 114. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). A neural network
1193 method for identification of prokaryotic and eukaryotic signal peptides and prediction
1194 of their cleavage sites. Int J Neural Syst 8, 581-599.
1195 115. Meszaros, B., Erdos, G., and Dosztányi, Z. (2018). IUPred2A: context-dependent
1196 prediction of protein disorder as a function of redox state and protein binding. Nucleic
1197 Acids Res 46, W329-W337.
1198
1199
1200

Legends to Figures and Tables

Table 1. Patients and Controls used in this study.

List of patients and controls including gene, mutation, age at biopsy, age of onset, sex, creatine kinase (CK) levels, clinical features, and the muscle groups from whence biopsies were obtained. Note that the normal range for CK levels is 25-200 U/l.

Figure 1. RNA-Seq for EDMD A. Workflow. Muscle biopsies were taken from the regions described in Table 1 and myoblasts recovered. These were then differentiated *in vitro* into myotubes, and myotubes selectively recovered by partial trypsinization. Myotube RNA was extracted and used for sequencing. Scale bar 20µm. B. Number of genes differentially expressed (FDR 5%) for each individual patient and for all patients considered together as a single group (denoted as "G1"). C. Heatmap of log2FC values for genes changing expression in EDMD patients compared to healthy controls (G1). Red is upregulated and blue is downregulated. Black indicates no change. The G1 lane is the averaged data across all patients. D. Dendrogram showing the relationships between patients, which fall into three broad groups. E. Overlaps with genome-organizing NETs gene targets. The numbers of genes altered by knockdown of muscle-specific genome-organizing NETs Wfs1, Tmem38a, NET39, or Tmem214 that were also altered in the EDMD patient cells is given. G1 refers to the analysis of all 10 patients as a single group against the healthy controls, while gp1-3 refers to the three subgroups identified. Total number of differentially expressed (DE) genes are given for each.

Figure 2. Search for functional categories changing in EDMD patients A. Box plot of log2FC values for differentially expressed genes within significantly enriched functional categories in EDMD patients analyzed in the Bakay study using g:profiler. B. Box plot of log2FC values for differentially expressed genes within significantly enriched functional categories in EDMD patients analyzed in this study using g:profiler. C. Box plot of log2FC values for leading edge genes within significantly enriched functional categories in EDMD

patients analyzed in this study using GSEA pathway (C2 geneset collection: canonical pathways). D. GSEA enrichment plot for mitochondrially encoded genes. GSEA analysis using the C1 geneset (positional) revealed an upregulation of mitochondrially encoded genes. See Supplemental Table S3 for further details.

Figure 3. Metabolism changes confirmed in patients. A. Heatmap of all mitochondria-encoded genes showing almost uniform upregulation in EDMD patients. B. Heatmap of nuclear-encoded downregulated genes in Bakay's study that supported mitochondria function. Downregulation is also generally observed in patients belonging to Group 1, but not Groups 2 or 3. Functional analyses performed in primary patient (n=12) and control (n=7) myoblast cultures for C. glycolysis and basal respiration show a significant reduction of both in EDMD myoblast cultures. D. Testing for fuel dependency we could not observe any changes (* glucose and fatty acids have been measured; glutamine calculated). E. ATP production is reduced in EDMD samples. All functional experiments have been repeated in at least three independent experiments. F. qPCR of mitochondria shows a reduction of mitochondria in EDMD myoblasts. G. Heatmap of mitophagy genes (GO-BP, GO:0000422, Autophagy of the mitochondrion) that are differentially expressed in at least one of Bakay's EDMD, G1 analysis (all 10 patients analyzed as a single group), or G3 analysis (each of the three subgroups analyzed separately).

Figure 4. Genes changing expression within functional groups supporting ECM/ fibrosis/ myotube fusion. A. Summary of altered functions in EDMD. B. Heatmap showing gene expression changes for collagens. C. Heatmap showing gene expression changes for matrix metalloproteinases. D. Heatmap showing gene expression changes for general proteins involved in fibrosis according to FibroAtlas (http://biokb.ncpsb.org.cn/fibroatlas). E. Heatmap showing gene expression changes for cytokine signaling proteins (GO-BP, GO:0019221). F. Heatmap showing gene expression changes for interferon-gamma signaling proteins (WikiPathways, WP619). G. Heatmap showing downregulation of genes

involved in the degradation of cell cycle proteins. GSEA identified downregulation in EDMD patients of REACTOME hsa-1741423 (APC-C mediated degradation of cell cycle proteins) and hsa-187577 (SCF/SKP2 mediated degradation of p27/p21) (Fig. 2C). Both genesets were combined. H. Heatmap showing gene expression changes for skeletal muscle regeneration proteins (GO-BP, GO:0043403). I. Heatmap showing gene expression changes for myoblast fusion (GO-BP, GO:0007520). J. Heatmap showing gene expression changes for myoblast differentiation (GO-BP, GO:0045445). K. Heatmap showing gene expression changes for specific genes functioning in myotube fusion that are under genome-organizing NET regulation. L. Heatmap showing the upregulation of many pro-adipogenic genes (Wikipathways, WP236). All heatmaps were generated from differentially expressed genes in at least one of Bakay's EDMD, G1 analysis (all 10 patients analyzed as a single group), or G3 analysis (each of the three subgroups analyzed separately). K-means unsupervised hierarchical clustering was used to summarize the heatmap into 8 clusters of roughly coexpressing genes. Full heatmaps displaying gene names are provided in **Figures S5-S11**. The number of genes per cluster is indicated on the right of each heatmap. Red and blue indicate upregulation and downregulation respectively.

Figure 5. Analysis of splicing defects reveals loss of muscle-specific splice variants. A. GOchord plot for misregulated splicing factors indicates the primary change is upregulation of alternative splicing and downregulation of constitutive splicing. Plot for gp1 is shown as a representative. DE genes log2FC > |0.15| and $p < 0.05$. B. Venn diagrams of alternatively spliced genes predicted by rMATS, DEXSeq and ISA for all samples (G1) and the three subgroups (gp1, gp2 and gp3). Overlap between all three methods indicates genes with exon skipping events that lead to annotated isoform switches. C. Bar chart for functional pathways reveal an enrichment in altered splice variants for pathways associated with metabolism, gene expression, cytoskeleton organization, DNA repair, proliferation/ differentiation, stress, ECM/ fibrosis/ and sarcomere structure. Number of genes detected displayed as log2. E. Pie Chart of significant AS events with psi > |0.1| in gp1 detected with

rMATS, which were used to scan for muscle specific splice variants, shown as Heatmap. AS events SE = exon skipping, MXE= mutually exclusive exons, RI = intron retention, A3SS = alternative 3' splice site, A5SS = alternative 5' splice site. F. Isoform switches as analyzed using isoformSwitchAnalyzer for ZNF880 and TMEM38A. ZNF880 shows the same isoform switch in all groups with preferential use of the shorter isoform, that contains the KRAB domain (black), but not the zinc finger domain (blue). TMEM38A shows a clear switch from the muscle isoform to a shorter isoform.

Figure 6. miRNA analysis. A. Heatmap of miRNAs that had altered levels in EDMD patient cells compared to controls. Red indicates upregulated and blue downregulated with intensity according to the log2FC values. B. Overview of functional categories linked to differentially expressed miRNAs in EDMD. The label size is proportional to the number of DE miRNAs associated with each category.

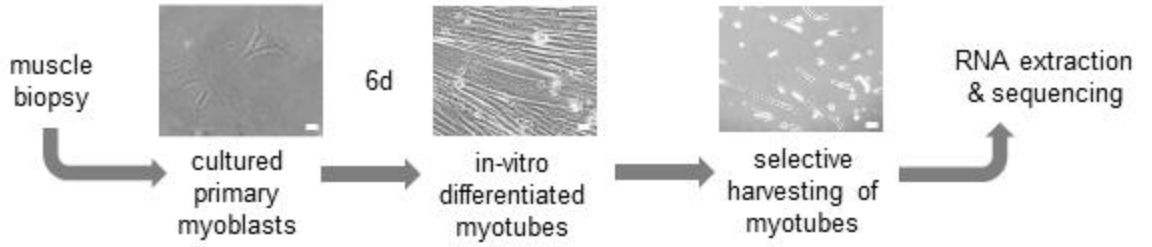
Figure 7. EDMD is distinct from other MDs. A. Scatterplots from GSEA analysis comparing Bakay data for several muscular dystrophies to the new patient data from this study (all 10 patients as a single group = G1). On the x axis, the Normalized Enriched Score (NES) is a measurement of the enrichment of the DE geneset identified in our study compared to each of the diseases in the Bakay study. The y axis shows the -log10(FDR), which is a measurement of statistical confidence. The grey horizontal line marks the 5% FDR threshold. Muscular dystrophies from the Bakay study are EDMD, Limb-Girdle muscular dystrophy 2A (LGMD2A), Limb-Girdle muscular dystrophy 2B (LGMD2B), Limb-Girdle muscular dystrophy 2I (LGMD2I), FascioScapuloHumeral muscular dystrophy (FSHD), Duchenne Muscular Dystrophy (DMD), Becker Muscular Dystrophy (BMD), Juvenile Dermatomyositis (JDM), Acute Quadriplegic Myopathy (AQM), Amyotrophic Lateral Sclerosis (ALS) and Hereditary Spastic Paraplegia (HSP). B. Same as A, for each individual patient subgroup. C. Box plot of log2FC values for differentially expressed genes within significantly enriched functional categories for each patient subgroup compared to the rest,

using g:Profiler. D. Each patient subgroup was relatively more enriched for certain functional pathways than other subgroups, suggesting that treatments for example targeting different metabolic pathways for groups 1 and 3 might partially ameliorate some patient difficulties.

1317 Table 1

ID	gene	mutation	RNAseq	metabolic analysis	age	age of onset	sex	CK(U/l)	muscle weakness	contractures	cardiac defects	neurological defects	biopsy location	other clinical information
C1	control		yes	yes	35		M		no	no	no		Gastrocnemius caput laterale	
C2	control		yes	yes	13		M		no	no	no			
C3	control			yes	43		M		no	no	no		Biceps brachii	
C4	control			yes	36		F		no	no	no		Biceps brachii	
C5	control			yes	49		F		no	no	no		Vastus lateralis	
C6	control			yes	49		F		no	no	no			
C7	control			yes	53		M		no	no	no		Vastus lateralis	
P1	TMEM214	p.179H	yes	yes	48	28	F	350	yes		no		Rectus femoris	LGMD, ptosis, type 2 fiber atrophy
P2	PLPP7	p.M92K	yes	yes	34	16	F	400	yes		no		Vastus lateralis	distal myopathy
P3	SUN1	p.G68D p.G388S	yes	yes	9	10	M	>3000	yes	yes	yes		Quadriceps femoris	
P4	SYNE1	p.6869* p.6869*	yes	yes	15		F	347	yes			yes	Quadriceps femoris	growth deficit, ataxia
P5	EMD	p.S58Sfs*1	yes	yes	17	12	F	278	yes	yes	yes		Biceps brachii	
P6	FHL1	c.688+1G>A	yes		28		M	1000	yes		yes		Biceps brachii	
P7	FHL1	p.C224W	yes	yes	46		M	300-1000	yes				Tibialis anterior	
P8	FHL1	p.V280M	yes	yes	21		M	231	yes	yes			Gastrocnemius caput laterale	rigid spine
P9	LMNA	p.T528K	yes		2		M		yes					
P10	LMNA	p.R571S	yes	yes	24	9	F	up to 12000	yes					degenerative myopathy
P11	LMNA	p.R545C		yes	18		M		yes	yes	yes			
P12	LMNA	p.R543W		yes	12		F		yes				Vastus lateralis	acrogeria
P13	unknown	unknown		yes	14		M	2800	yes	yes	yes		Tibialis anterior	degenerative myopathy
P14	FHL1	p.C224W		yes	42		M	1700	yes	yes	yes		Vastus lateralis	myofibrillar myopathy

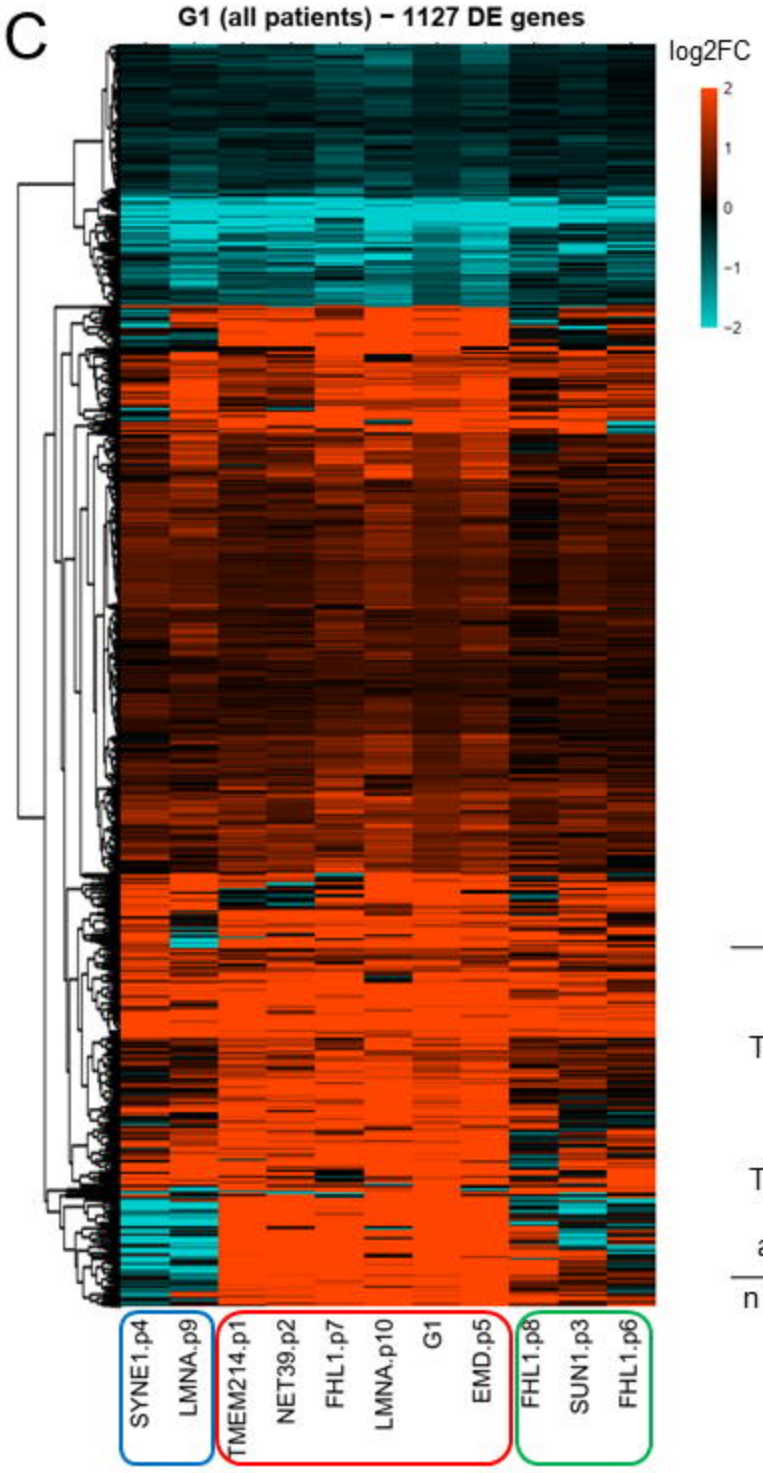
A



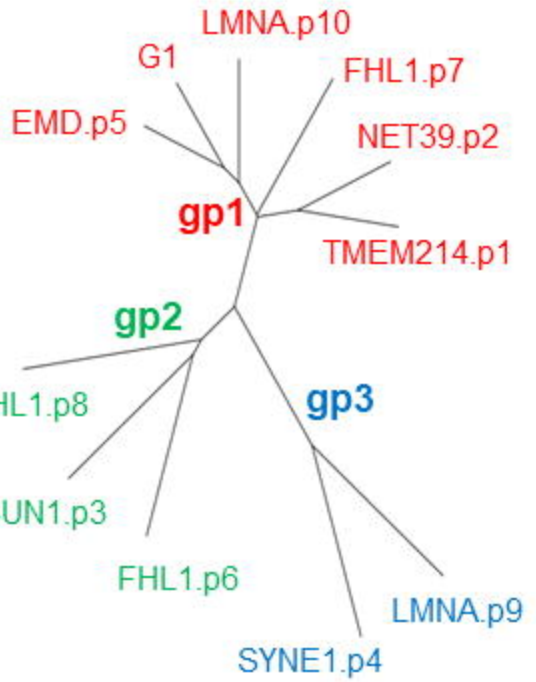
B

patient	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10	G1 (all)
gene	*TMEM214*	*PLPP7*	*SUN1*	*SYNE1*	*EMD*	*FHL1*	*FHL1*	*FHL1*	*LMNA*	*LMNA*	EDMD
up	648	1057	403	869	2062	310	2417	512	1652	2651	894
down	1087	726	498	1568	1109	429	1660	442	2384	2127	233

C



D



E

	G1(all)	gp1	gp2	gp3
PLPP7	71	351	18	330
TMEM214	70	377	19	381
WFS1	68	333	18	281
TMEM38A	70	334	18	312
any NET	166	824	38	734
n DE genes	1127	4738	265	3298