

The FASEB Journal express article 10.1096/fj.05-4678fje. Published online November 23, 2005.

Common pathological mechanisms in mouse models for muscular dystrophies

R. Turk,^{*,†} E. Sterrenburg,^{*} C. G. C. van der Wees,^{*} E. J. de Meijer,^{*} R. X. de Menezes,^{*} S. Groh,[†] K. P. Campbell,[†] S. Noguchi,[‡] G. J. B. van Ommen,^{*} J. T. den Dunnen,^{*} and P. A. C. 't Hoen^{*}

^{*}Leiden University Medical Center, Center for Human and Clinical Genetics, Leiden, Nederland; [†]Howard Hughes Medical Institute, Department of Physiology and Biophysics, Iowa City, Iowa; and [‡]National Institute of Neuroscience, Department of Neuromuscular Research, Tokyo, Japan

Corresponding author: Dr. P.A.C. 't Hoen, Center for Human and Clinical Genetics, Leiden University Medical Center, Wassenaarseweg 72, 2333 AL Leiden, Nederland. E-mail: p.a.c.hoen@lumc.nl

ABSTRACT

Duchenne/Becker and limb-girdle muscular dystrophies share clinical symptoms like muscle weakness and wasting but differ in clinical presentation and severity. To get a closer view on the differentiating molecular events responsible for the muscular dystrophies, we have carried out a comparative gene expression profiling of hindlimb muscles of the following mouse models: dystrophin-deficient (*mdx*, *mdx*^{3^{cv}}), sarcoglycan-deficient (*Sgca* null, *Sgcb* null, *Sgcg* null, *Sgcd* null), dysferlin-deficient (*Dysf* null, *SJL*^{Dysf}), sarcospan-deficient (*Sspn* null), and wild-type (C57Bl/6, C57Bl/10) mice. The expression profiles clearly discriminated between severely affected (dystrophinopathies and sarcoglycanopathies) and mildly or nonaffected models (dysferlinopathies, sarcospan-deficiency, wild-type). Dystrophin-deficient and sarcoglycan-deficient profiles were remarkably similar, sharing inflammatory and structural remodeling processes. These processes were also ongoing in dysferlin-deficient animals, albeit at lower levels, in agreement with the later age of onset of this muscular dystrophy. The inflammatory proteins *Spp1* and *S100a9* were up-regulated in all models, including sarcospan-deficient mice, which points, for the first time, at a subtle phenotype for *Sspn* null mice. In conclusion, we identified biomarker genes for which expression correlates with the severity of the disease, which can be used for monitoring disease progression. This comparative study is an integrating step toward the development of an expression profiling-based diagnostic approach for muscular dystrophies in humans.

Key words: microarray • dystrophin-glycoprotein complex • inflammation • extracellular matrix • biomarker

Muscular dystrophies are a heterogeneous group of inherited neuromuscular disorders characterized by progressive muscle wasting and weakness. The genetic defects underlying many muscular dystrophies have been elucidated (1, 2). A particular subset

of muscular dystrophies is caused by mutations in genes coding for constituents of the dystrophin-associated glycoprotein complex (DGC). The DGC is a multimeric protein complex composed of integral, peripheral, and cytoplasmic proteins expressed at the sarcolemma of muscle fibers. One likely role of the DGC is to maintain the sarcolemma stability by providing a physical link between the extracellular matrix and the actin cytoskeleton. This link occurs through the trio dystrophin-dystroglycan-laminin-2. Dystrophin binds to cytoskeletal F-actin and the transmembrane β -subunit of dystroglycan. Dystroglycan binds to laminin-2 in the extracellular matrix via its α -subunit (3). The trio dystrophin-dystroglycan-laminin-2 is stabilized by other DGC components such as the sarcoglycan-sarcospan complex (SGC), which resides in the sarcolemma. In skeletal muscle, the SGC contains four glycosylated subunits (α -, β -, γ -, and δ -sarcoglycan) and sarcospan (4).

Although genetic mutations in one of the DGC components frequently result in destabilization or mislocalization of the entire complex, the clinical presentation of the gene defects is variable: dystrophin-deficiency (dystrophinopathy) results in either the lethal Duchenne or the milder Becker muscular dystrophy (DMD/BMD) (5), dystroglycan-deficiency is likely to be embryonically lethal (6), and laminin 2-deficiency results in congenital muscular dystrophy (MDC1A) (7). Mutations in sarcoglycan-proteins (sarcoglycanopathies) are responsible for several recessive forms of limb-girdle muscular dystrophy (LGMD-2C-F, reviewed in refs 8, 9), whereas sarcospan has not been associated with human disease (10).

Mutations in dysferlin, a muscle membrane protein that plays a role in membrane repair (11), cause the non-DGC related muscular dystrophies LGMD-2B and Myoshi myopathy (12). In these relatively late-onset diseases, the defective membrane repair system is ultimately unable to cope with contraction-induced injuries to the sarcolemma, explaining the observed muscle degeneration (11, 13).

The pathological manifestation of the different muscular dystrophies at the histological level is similar. Dystrophic muscle tissue is hallmarked by myofiber degeneration, infiltration of inflammatory cells, and subsequent formation of foci of fibrotic and adipose tissue. These manifestations are secondary to the primary defect. The decisive secondary factors responsible for the variability in the clinical phenotypes are still unknown.

Muscular dystrophies can be clinically grouped based on several indicators, namely age of onset, severity, presentation, affected musculature, and genetic inheritance (14). The clinical variability within each muscular dystrophy complicates the diagnosis of muscular dystrophies, particularly in young children. Consequently, a large array of molecular diagnostic methods, frequently combining analyses at the protein, DNA, and mRNA level, has to be applied to come to a specific diagnosis (15).

Gene expression profiling studies have generated more detailed insight in the molecular processes underlying DMD (16–19). However, few microarray datasets have been published on other muscular dystrophies (13, 20–23). The goal of the present study is to assess whether the muscular gene expression patterns in the individual muscular dystrophies are sufficiently different to allow for microarray-based classification, which could greatly facilitate diagnosis. In addition, comparison of the different gene expression patterns in muscular dystrophies will help to understand the molecular mechanisms underlying the specificity of the clinical symptoms.

technical replicates each). The randomized experimental design facilitates an intensity-based analysis procedure, since ratio-based analysis procedures become complicated and inefficient in multi-class comparisons (36). Spot intensities were evaluated with feature extraction software (GenePix Pro 5, Axon, Union City, CA). Local background corrected, median spot intensities were normalized simultaneously for all microarray experiments using variance stabilization and normalization (VSN) as described previously (37). This transformation coincides with the natural logarithm for the high intensities. Two filtering criteria were applied. First, genes that were flagged in at least four out of eight observations for a particular genotype were excluded to minimize the influence of unreliable spots (specks, high background). Since samples were randomly hybridized with each other, and spots were flagged only when signals in both channels were below background, it is very unlikely that genes that are expressed only in one or a few models were excluded from the analysis. Second, 88 genes previously shown to demonstrate variation in expression in the muscle due to differences in the genetic background (34) were excluded from the analysis. In the end, 2751 genes were included in the analysis.

Microarray data have been made available through the GEO data repository of the National Center for Biotechnology Information under series GSE2112.

Evaluation of differential expression between mouse models

Significance levels of differential gene expression differences between different classes of mouse models were calculated in R (38) using linear regression models that take the dye effect into account. All fixed-effects model were of this form:

$$(1) Y = \alpha + \beta \cdot \text{class} + \gamma \cdot \text{dye}$$

In all comparisons, the expression levels of Cy3-labeled targets were captured in the intercept (dye=0), whereas dye equaled 1 for Cy5-labeled targets.

In the comparison of the different strains, the expression levels of C57Bl/10 were captured in the intercept (class=0), and different β coefficients were estimated for the other strains.

In the comparison of severely affected (*mdx*, *mdx*^{3^{cv}}, *Sgca* null, *Sgcb* null, *Sgcd* null, *Sgcg* null) and non- or mildly affected (C57Bl/10, C57Bl/6, *Sspn* null, *Dysf* null, *SJL*^{Dysf}) models, the expression levels of non- or mildly affected were captured in the intercept (class=0), and class equaled 1 for severely affected mouse models.

Similarly, in the comparison of dystrophin-deficient (*mdx*, *mdx*^{3^{cv}}) and sarcoglycan-deficient (*Sgca* null, *Sgcb* null, *Sgcd* null, *Sgcg* null) mouse models, expression levels of dystrophin-deficient mice were captured in the intercept (class=0), and class equaled 1 for sarcoglycan-deficient mice.

Similarly, in the comparison of dysferlin-deficient (*Dysf* null, *SJL*^{Dysf}) vs. control (C57Bl/10, C57Bl/6) mouse models, expression levels of control mice were captured in the intercept (class=0), and class equaled 1 for dysferlin-deficient mice.

P values for class were corrected for multiple testing using the method proposed by Benjamini and Hochberg (BH-correction) (39).

To find biomarker genes for disease progression, we fitted the following models:

$$(2) Y = \alpha_2 + \beta_2 \cdot \text{class2} + \gamma_2 \cdot \text{dye}$$

$$(3) Y = \alpha_3 + \beta_3 \cdot \text{class3} + \gamma_3 \cdot \text{dye}$$

In model 2, we compared nonaffected (intercept; C57Bl/10, C57Bl/6, Sspn null) with mildly affected (class2=1; Dysf null, SJL^{Dysf}), and in model 3, we compared mildly affected (intercept; Dysf null, SJL^{Dysf}) with severely affected (class3=1; *mdx*, *mdx*^{3cv}, Sgca null, Sgcb null, Sgcd null, Sgeg null). We selected genes with significant *P* values (BH-corrected *P* value < 0.05) in both comparisons and then identified up-regulated biomarker genes as genes with both β_2 and β_3 greater than zero, whereas for down-regulated biomarker genes both β_2 and β_3 were smaller than zero.

Clustering

Unsupervised hierarchical clustering was performed using the Functional Genomics package within Spotfire DecisionSite 7.3. Expression levels were scaled such that the average expression per gene over all conditions was zero. Average linkage was the clustering method applied, and Euclidean distance was taken as similarity measure.

Functional annotation

Recent functional annotation of genes was obtained from the SOURCE database (<http://source.stanford.edu>) (40). Genes were classified based on gene ontology using the web-based tool Gene Ontology Tree Machine (GOTM; <http://genereg.ornl.gov/gotm>) (41). Using this tool, we identified the functional processes that were overrepresented in lists of up- and down-regulated genes compared with the list of all the genes on the array using a hypergeometric test. Only categories containing at least three differentially expressed genes and with a *P* value < 0.001 were considered.

Quantitative RT-PCR

Quantitative RT-PCR was performed as described previously (25). Gene expression levels were calculated using the gene expression macro provided by Bio-Rad (Bio-Rad, Hercules, CA) and normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH, stable expression in all samples) expression levels. Primer sequences are available on request.

RESULTS

We applied large-scale gene expression profiling to make a global inventory of differences in gene expression levels between mouse models for different muscular dystrophies. Nine mouse models with different genetic defects ([Table 1](#)) and two wild-type mouse strains were compared. Total RNA, isolated from hindlimb muscle of two individual mice, was amplified and hybridized to murine microarrays containing 7445 gene-specific oligonucleotides, spotted in duplicate. A dye-balanced, randomized hybridization design was chosen to avoid potential dye, sample processing, and hybridization biases (Supplemental Table 1). Combined with an intensity-based

analysis procedure, this experimental design is most efficient when comparing large numbers of groups (36).

After being filtered for unreliable spots and for genes the expression of which was known to be affected by differences in genetic backgrounds (34), 2171 genes demonstrated differential expression in the 11 models ($P < 0.05$, BH-correction). These genes, as well as their expression levels in the different models, are listed in Supplemental Table 2. We applied unsupervised hierarchical clustering on the profiles of the 300 most significantly differentially expressed genes to determine the similarity between the mouse models (Fig. 2A). According to the hierarchical clustering, two groups can be clearly discerned. The first group (I) contains the models for dystrophinopathy and sarcoglycanopathy, which have a severe dystrophic phenotype at the age of 8 wk. The second group (II) consists of models with a mild or unaffected phenotype.

The identity of the genes that contribute most to the distinction between group I and group II was determined by calculating per gene the statistical significance between the two groups with a linear regression model. After correction for multiple testing, 1028 genes appear to function as general markers for dystrophic muscles (Supplemental Table 3). In Table 2, we present the genes that are at least twofold up-regulated or down-regulated (128 and 46 genes, respectively) in the mouse models for dystrophinopathy and sarcoglycanopathy, compared with mildly affected or unaffected muscle tissue.

To determine the biological variation in gene expression between the two individual mice from each strain, we performed a hierarchical clustering on the 300 most significant genes using the average gene expression levels of all separate individuals (Fig. 2B). Within the group of mildly and nonaffected mouse models (group II), both individuals of each strain clustered together, demonstrating a high level of similarity in their gene expression profiles. In contrast, within the group of severely affected mouse models (group I), the gene expression profiles of the biological replicates were located in different clusters, which is indicative of a higher variability in gene expression profiles between individuals than between models.

To assess whether the absence of one DGC component affects the expression of other DGC components, we evaluated the expression levels of seven genes involved in the DGC (*Dmd*, *Sgca*, *Sgcb*, *Sgcg*, *Sgcd*, *Sspn*, and *Dysf*; Fig. 3). The knockout mouse models demonstrated clearly reduced mRNA levels of the targeted gene, with an exception for *Sgcd* in *Sgcd* null mice and *Dysf* in dysferlin-deficient mice, in which both only a relatively modest decrease in signal intensity is shown. The effect of a point mutation (*mdx*, *mdx*^{3cv}, and SJL^{Dysf}) seems to have a milder influence on the mRNA levels, probably due to less efficient degradation by nonsense-mediated mRNA decay (NMD), which depends on the location of the mutation (42, 43). In the mouse models for sarcoglycanopathy, a decrease in expression levels from other DGC components was observed, with largest effects in *Sgca* null mice. The gene expression of dysferlin, on the other hand, was up-regulated in severely dystrophic mouse models, with the exception of *Sgcb* null mice.

We implemented a recently published bioinformatics tool based on gene ontology (GOTM; ref 41) to investigate the functional processes active in severely affected dystrophic muscle tissue (group I). This tool associates sets of genes (e.g., up-regulated genes) with functional processes and determines whether these processes are represented at a higher frequency than can be

expected by chance, i.e., as represented in the reference gene list, consisting of all genes on the microarray. The results of this analysis are presented in [Fig. 4](#). The most strikingly overrepresented biological processes in the list of up-regulated genes are cell adhesion (*Icam1*, *P-selectin* and several integrins, laminins, and thrombospondins), inflammation (many chemokines, cytokines, cytokine receptors, lymphocyte antigens), and regulation of muscle contraction (*Atpa2*, *Calsequestrin 1* and *2*, and *Troponin C, I, T1*, and *T2*). Many of the up-regulated genes localized either in the extracellular matrix (including 12 different collagens) or in the lysosome. The most striking feature of the list of down-regulated genes is the participation of 53% of these genes (259 genes) in highly diverse metabolic processes, indicative of an overall decline in metabolic activity in dystrophic tissue. A large number of the gene products of these genes are localized in the mitochondrion (72 genes; *P* value for overrepresentation: 5.48×10^{-24}). The individual genes in the listed functional categories are given in Supplemental Table 4.

Although gene expression profiles of mouse models for dystrophinopathy show high similarity with those of mouse models for sarcoglycanopathy, some genes displayed statistically significant differential gene expression between these groups. With the linear regression model described in Materials and Methods, we found 46 differentially expressed genes ($P < 0.05$ after BH-correction for multiple testing, Supplemental Table 5). A selection of four genes with greater than twofold lower expression and five genes with greater than twofold higher expression in mouse models with sarcoglycanopathy compared with mouse models for dystrophinopathy is presented in [Table 3](#).

The clustering diagrams in [Fig. 2](#) suggest that gene expression patterns of mildly affected dysferlin-deficient animals can be discerned from those of healthy mice. With a linear regression model, 321 genes were identified that displayed statistically significant ($P < 0.05$ after BH-correction) differences in expression levels between dysferlin-deficient (SJL^{Dysf} and Dysf null mice) and the two wild-type strains (Supplemental Table 6). Interestingly, the majority of up-regulated (101/154) and down-regulated (88/167) genes were also differentially expressed between group I (severely affected) and group II (mildly or nonaffected). Since the latter group includes the dysferlin-deficient mice, these are genes with subtle changes in dysferlin-deficient mice and higher fold changes in the more severe models. We statistically evaluated which genes show a significant trend in expression level from nonaffected to mildly and to severely affected animal models. These genes can be seen as biomarker genes for disease severity. The heat map in [Fig. 5](#) illustrates the expression levels of the 33 identified biomarker genes.

Two genes up-regulated in both mildly and severely affected animal models, *S100a9*, coding for a phagocytic protein known as calgranulin B, and *Spp1*, coding for a cytokine known as osteopontin, were also up-regulated in sarcospan-deficient mice ($P < 10^{-7}$ after correction for multiple testing). This observation suggests for the first time a subtle molecular phenotype for these mice. With quantitative RT-PCR, we confirmed the much higher expression of *Spp1* and *S100a9* in sarcospan-deficient mice (29- and 14-fold increase over wild-types, respectively), as well as in the other muscular dystrophy models (Supplemental Fig. 1).

DISCUSSION

In this report, we have classified muscular dystrophies based on gene expression profiles in skeletal muscle of 8-wk-old mice. Cluster analysis readily distinguished severely affected mouse

models (dystrophinopathies and sarcoglycanopathies) from dysferlin- and sarcospan-deficient and wild-type mice. Earlier studies demonstrated a high similarity in the histology of muscles from mouse models for dystrophinopathy and sarcoglycanopathies at the age of 8 wk (27–29, 32, 44, 45). We now show that common denominators in dystrophinopathies and sarcoglycanopathies are also apparent at the molecular level. The robustness of this classification is further illustrated by the fact that one of the two mouse models for β -sarcoglycan deficiency (29, 30) that were originally included in the study, clustered with the non- or mildly affected mice. Since this was highly unexpected, a detailed analysis was performed using histological techniques and RT-PCR, showing that one of the two mice analyzed was incorrectly genotyped as a *Sgcb* null mouse instead of wild-type. As a consequence, we had to exclude the second *Sgcb* null model from further analysis.

Gene expression levels of DGC components decrease in DGC-related muscular dystrophies

Although the entire DGC is absent from the sarcolemma in dystrophinopathies, whereas only the SGC is lost in sarcoglycanopathies, we observed a general decrease in mRNA levels of the DGC components in both dystrophin- and sarcoglycan-deficient mice. The instability of the DGC apparently triggers a negative feedback to transcriptional levels of the DGC components. Notable is the up-regulation of dysferlin, which is indicative for an increased need for membrane repair systems to prevent membrane leaking or rupture, a common feature for muscular dystrophies (46, 47). Other important processes secondary to the genetic defect and the loss of a functional DGC are induction of an immune response, increased expression of cellular adhesion and extracellular matrix proteins, and changes in cytoskeletal organization. These processes were also shared between human patients with DMD and LGMD-2D (α -sarcoglycan deficiency) (16).

The inflammatory response in muscular dystrophy can be divided into multiple components

By looking at markers for different types of immune cells (48), we found evidence for the infiltration and accumulation of macrophages [*Cd68*, *Lgals3* (*Mac-2*), *Mpeg1*, *Clecsf12*], B-cells (*Cd83*, *Blnk*), T-cells (*CD8b1*, *Ly6e*), and NK-cells (*Ypel1*). The recruitment of inflammatory cells is probably mediated by CC-class chemokines that are up-regulated (*Ccl2*, *Ccl6*, *Ccl7*, *Ccl8*, and *Ccl9*), as was reported for *mdx* mice (25, 49, 50). Furthermore, an activation of components of the complement system (*C1qa*, *C1qg*, *C1qr1*, *C1qTNF3*, and *C3ar1*) was observed, which contributes to the inflammatory response in dystrophic muscle by further damaging cell membranes and releasing complement split products that attract macrophages for phagocytosis.

Genes involved in sarcomeric organization and extracellular matrix formation are up-regulated in both sarcoglycan- and dystrophin-deficient muscles

Among the up-regulated extracellular matrix proteins are 12 different collagens, laminin α 4, B1 and gamma1, and collagen binding proteins such as biglycan. The up-regulated cytoskeletal proteins include the intermediate filament vimentin, the microtubular components tubulin α -1, -2, -6, and β -5, the actin-interacting protein transgelin2 and the sarcomeric troponins I (skeletal, slow), T1 (skeletal, slow), T2 (cardiac), C (cardiac / slow skeletal), most of which having been reported before as up-regulated in *mdx* mice (25, 50–52). Such up-regulation of extracellular and

intracellular structural proteins, which is also found in human DMD patients (17–19, 53) might compensate for the loss of force-generating capacity due to the instability of the DGC-mediated link between the cytoskeleton and the extracellular matrix.

Dystrophic muscle tissue experiences a metabolic crisis

The large number of down-regulated genes functioning in diverse metabolic processes reflects probably the metabolic crisis also seen in human DMD, LGMD-2D, FSHD, and nemaline myopathy patients (16, 22, 23). From our and data of others (51, 52), we conclude that the metabolic crisis in *mdx* mice is less severe than in sarcoglycan-deficient mice. However, a reduction in the respiratory rate of mitochondria in skeletal muscle of both *mdx* mice and DMD patients was also found in previous studies using other techniques (54, 55).

Differential gene expression between dystrophinopathies and sarcoglycanopathies

Despite the high similarities between dystrophinopathies and sarcoglycanopathies at 8 wk of age, 46 genes were found significantly differentially expressed between these groups. This may indicate that there are some subtle differences between dystrophinopathies and sarcoglycanopathies. These differences may become more prominent at higher age, since the *mdx* mice, in contrast to the sarcoglycan-deficient animals, make an almost complete recovery. A more extensive temporal gene expression profiling study together with functional analysis of the identified proteins is necessary to characterize the differences (work in progress). Cysteine and glycine-rich protein 3 (*Csrp3*), a positive regulator myogenesis (56), was the most prominent, higher expressed gene in sarcoglycan-deficient compared with dystrophin-deficient animals. Interestingly, it has been found before that *Csrp3* is also up-regulated in human FSHD but not DMD patients (22). This exemplifies that muscular dystrophies share some aberrations in gene expression and differ in others, and that it would be feasible to find a specific disease signature based on the measurement of the expression of combinations of genes.

Vascular irregularities are not likely to contribute to muscular dystrophy

Mutations in smooth muscle SGC-components (*Sgcb*, *Sgcg*, *Sgcd*) lead to vascular irregularities, which are thought to exacerbate the cardiac pathology seen in mice deficient for the individual components (29, 57, 58). Alpha-sarcoglycan is not expressed in smooth muscle and therefore the *Sgca* null mouse does not have altered SGC expression in smooth muscle. We show that the gene expression profiles of *Sgca* null mice are highly similar to the gene expression profiles of the other sarcoglycan-deficient mouse models, corroborating the suggestion of Durbeej et al. that vascular irregularities are not likely to contribute to the skeletal muscle pathology (24).

Differential gene expression precedes histological changes in dysferlinopathies

The phenotype of dysferlin-deficient mice is less progressive than that of the dystrophin- and sarcoglycan-deficient mice (11, 26). HE staining shows only few necrotic fibers at the age of 8 wk, corresponding with the higher age of onset of dysferlinopathies, both in mice and humans. Still, we observe significant differences in gene expression between wild-type mice and dysferlin-deficient mice at the age of 8 wk. Among the up-regulated genes that have also been found in an earlier study in SJL mice (13) are: Troponins T1, C, and I, *Ccl2*, *Cd53*, *Cd68*, *Cd83*, P lysozyme structural, tissue inhibitor of metalloproteinase 1 and 4. Many of these genes ([Fig. 5](#))

are more severely changed in the severely affected mouse models. Thus, in dysferlin-deficient mice, similar to dystrophin and sarcoglycan-deficient mice, inflammatory and muscle remodeling processes play an important role (confirmed by gene ontology analysis with GOTM, not shown). Although differences in the expression of these genes are still limited at the age of 8 wk, they may become more prominent at higher ages. In the *mdx* mouse, however, the expression of these genes is already increased at 4 wk and peaks at 6 wk (25), in agreement with the earlier age of onset in dystrophinopathies. Consequently, these genes can be regarded as markers for disease progression.

A molecular phenotype for sarcospan-deficient mice

Genes, coding for secreted phosphoprotein 1 (*Spp1* or Osteopontin), functioning in chemotaxis, and the phagocytic S100 calcium binding protein A9 (*S100a9* or Calgranulin B), are also up-regulated in sarcospan-deficient mice. Differential expression of these genes indicates that sarcospan-deficiency may induce a subtle immune response in muscle. *Spp1* seems to be an extremely sensitive marker for muscle pathology, since it was picked up in other gene expression profiling studies in dystrophin-deficient humans and mice (18, 51).

Inter-individual variation in gene expression levels

Our study shows that the inter-individual differences in expression levels are considerable, even in genetically identical mice. Consequently, the subtle differences that discriminate between the muscular dystrophies will only become significant when large sample groups are analyzed. We suppose that the inter-individual differences in the severely affected mice are related to the physical behavior or the amount of exercise. An increase in muscle activity leads to an increase in sarcolemmal rupture, which initiates the secondary processes responsible for the severity of the disease (59, 60). Recently observed inter-individual variability in satellite cell number (61) and possibly regenerative potential in dystrophic mice, may also contribute to the observed variability. Due to the interindividual variability, at least 4 animals per group were required for meaningful comparisons. This is why classes of genotypes were combined in the performed statistical tests (e.g., the class of dysferlin-deficient mice, consisting of the 2 SJL^{Dysf} and 2 dysferlin null individuals). In humans, many more genetic and environmental factors will contribute to the inter-individual variation (62), and building of a robust classifier necessary for muscular gene expression profiling-based diagnosis of muscular dystrophies would necessitate the analysis a large number of samples per group. On the other hand, analysis of disease progression using the biomarker set described here, would be more readily achievable, although more definitive correlations between expression levels of the identified biomarkers and other pathological parameters need to be established. This type of analysis would be useful to monitor response to treatment.

CONCLUSION

In summary, we have found remarkable similarity in the expression patterns of dystrophin-, sarcoglycan-, and dysferlin-deficient mice. Genes functioning in the inflammatory response and structural organization are significantly up-regulated compared with wild-type mice, whereas metabolism genes are down-regulated. We conclude that common pathogenic mechanisms underlay the onset and progression of different forms of muscular dystrophy in mice. Given the

similarity with published human studies on specific forms of muscular dystrophy, these pathogenic mechanisms may also contribute to the different forms of muscular dystrophy in humans. Furthermore, we have identified sets of biomarker genes that can be used to monitor disease progression in muscular dystrophies, thereby taking the first step toward the development of a diagnostic approach for muscular dystrophies by expression profiling-based classification.

ACKNOWLEDGMENTS

This work was supported by the Nederlandse Stichting voor Wetenschappelijk Onderzoek (NWO), the Center for Biomedical Genetics (CBG), and the Centre for Medical Systems Biology (CMSB) established by the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research (NGI/NWO).

REFERENCES

1. Cohn, R. D., and Campbell, K. P. (2000) Molecular basis of muscular dystrophies. *Muscle Nerve* **23**, 1456–1471
2. Dalkilic, I., and Kunkel, L. M. (2003) Muscular dystrophies: genes to pathogenesis. *Curr. Opin. Genet. Dev.* **13**, 231–238
3. Ozawa, E., Yoshida, M., Suzuki, A., Mizuno, Y., Hagiwara, Y., and Noguchi, S. (1995) Dystrophin-associated proteins in muscular dystrophy. *Hum.Mol.Genet.* 4 Spec No, 1711-1716
4. Straub, V., Duclos, F., Venzke, D. P., Lee, J. C., Cutshall, S., Leveille, C. J., and Campbell, K. P. (1998) Molecular pathogenesis of muscle degeneration in the delta-sarcoglycan-deficient hamster. *Am. J. Pathol.* **153**, 1623–1630
5. Hoffman, E. P., Brown, R. H., Jr., and Kunkel, L. M. (1987) Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* **51**, 919–928
6. Williamson, R. A., Henry, M. D., Daniels, K. J., Hrstka, R. F., Lee, J. C., Sunada, Y., Ibraghimov-Beskrovnaya, O., and Campbell, K. P. (1997) Dystroglycan is essential for early embryonic development: disruption of Reichert's membrane in Dag1-null mice. *Hum. Mol. Genet.* **6**, 831–841
7. Tome, F. M. S., Evangelista, T., Leclerc, A., Sunada, Y., Manole, E., Estournet, B., Barois, A., Campbell, K. P., and Fardeau, M. (1994) Congenital muscular dystrophy with merosin deficiency. *Comp.Rend.Acad.Sci.(Paris)* **317**, 351–357
8. Ozawa, E., Noguchi, S., Mizuno, Y., Hagiwara, Y., and Yoshida, M. (1998) From dystrophinopathy to sarcoglycanopathy: evolution of a concept of muscular dystrophy. *Muscle Nerve* **21**, 421–438
9. Lim, L. E., and Campbell, K. P. (1998) The sarcoglycan complex in limb-girdle muscular dystrophy. *Curr. Opin. Neurol.* **11**, 443–452

10. Crosbie, R. H., Lim, L. E., Moore, S. A., Hirano, M., Hays, A. P., Maybaum, S. W., Collin, H., Dovico, S. A., Stolle, C. A., Fardeau, M., et al. (2000) Molecular and genetic characterization of sarcospan: insights into sarcoglycan-sarcospan interactions. *Hum. Mol. Genet.* **9**, 2019–2027
11. Bansal, D., Miyake, K., Vogel, S. S., Groh, S., Chen, C. C., Williamson, R., McNeil, P. L., and Campbell, K. P. (2003) Defective membrane repair in dysferlin-deficient muscular dystrophy. *Nature* **423**, 168–172
12. Bashir, R., Britton, S., Strachan, T., Keers, S., Vafiadaki, E., Lako, M., Richard, I., Marchand, S., Bourg, N., Argov, Z., et al. (1998) A gene related to *Caenorhabditis elegans* spermatogenesis factor *fer-1* is mutated in limb-girdle muscular dystrophy type 2B. *Nat. Genet.* **20**, 37–42
13. Lennon, N. J., Kho, A., Bacskai, B. J., Perlmutter, S. L., Hyman, B. T., and Brown, R. H., Jr. (2003) Dysferlin interacts with annexins A1 and A2 and mediates sarcolemmal wound-healing. *J. Biol. Chem.* **278**, 50,466–50,473
14. Emery, A. E. (2002) The muscular dystrophies. *Lancet* **359**, 687–695
15. Bushby, K. M. (1999) Making sense of the limb-girdle muscular dystrophies. *Brain* **122**, 1403–1420
16. Chen, Y. W., Zhao, P., Borup, R., and Hoffman, E. P. (2000) Expression profiling in the muscular dystrophies: identification of novel aspects of molecular pathophysiology. *J. Cell Biol.* **151**, 1321–1336
17. Bakay, M., Zhao, P., Chen, J., and Hoffman, E. P. (2002) A web-accessible complete transcriptome of normal human and DMD muscle. *Neuromuscul. Disord.* **12**, S125–S141
18. Haslett, J. N., Sanoudou, D., Kho, A. T., Bennett, R. R., Greenberg, S. A., Kohane, I. S., Beggs, A. H., and Kunkel, L. M. (2002) Gene expression comparison of biopsies from Duchenne muscular dystrophy (DMD) and normal skeletal muscle. *Proc. Natl. Acad. Sci. USA* **99**, 15,000–15,005
19. Noguchi, S., Tsukahara, T., Fujita, M., Kurokawa, R., Tachikawa, M., Toda, T., Tsujimoto, A., Arahata, K., and Nishino, I. (2003) cDNA microarray analysis of individual Duchenne muscular dystrophy patients. *Hum. Mol. Genet.* **12**, 595–600
20. Campanaro, S., Romualdi, C., Fanin, M., Celegato, B., Pacchioni, B., Trevisan, S., Laveder, P., De Pitta, C., Pegoraro, E., Hayashi, Y. K., et al. (2002) Gene expression profiling in dysferlinopathies using a dedicated muscle microarray. *Hum. Mol. Genet.* **11**, 3283–3298
21. Tsukahara, T., Tsujino, S., and Arahata, K. (2002) CDNA microarray analysis of gene expression in fibroblasts of patients with X-linked Emery-Dreifuss muscular dystrophy. *Muscle Nerve* **25**, 898–901

22. Winokur, S. T., Chen, Y. W., Masny, P. S., Martin, J. H., Ehmsen, J. T., Tapscott, S. J., Van Der Maarel, S. M., Hayashi, Y., and Flanigan, K. M. (2003) Expression profiling of FSHD muscle supports a defect in specific stages of myogenic differentiation. *Hum. Mol. Genet.* **12**, 2895–2907
23. Sanoudou, D., Frieden, L. A., Haslett, J. N., Kho, A. T., Greenberg, S. A., Kohane, I. S., Kunkel, L. M., and Beggs, A. H. (2004) Molecular classification of nemaline myopathies: “nontyping” specimens exhibit unique patterns of gene expression. *Neurobiol. Dis.* **15**, 590–600
24. Durbeej, M., and Campbell, K. P. (2002) Muscular dystrophies involving the dystrophin-glycoprotein complex: an overview of current mouse models. *Curr. Opin. Genet. Dev.* **12**, 349–361
25. Turk, R., Sterrenburg, E., De Meijer, E., Van Ommen, G.-J. B., Den Dunnen, J. T., and 't Hoen, P. A. (2005) Muscle regeneration in dystrophin-deficient mdx mice studied by gene expression profiling. *BMC Genomics* **6**, 98
26. Bittner, R. E., Anderson, L. V., Burkhardt, E., Bashir, R., Vafiadaki, E., Ivanova, S., Raffelsberger, T., Maerk, I., Hoger, H., Jung, M., et al. (1999) Dysferlin deletion in SJL mice (SJL-Dysf) defines a natural model for limb girdle muscular dystrophy 2B. *Nat. Genet.* **23**, 141–142
27. Cox, G. A., Phelps, S. F., Chapman, V. M., and Chamberlain, J. S. (1993) New mdx mutation disrupts expression of muscle and nonmuscle isoforms of dystrophin. *Nat. Genet.* **4**, 87–94
28. Duclos, F., Straub, V., Moore, S. A., Venzke, D. P., Hrstka, R. F., Crosbie, R. H., Durbeej, M., Lebakken, C. S., Ettinger, A. J., van der, M. J., et al. (1998) Progressive muscular dystrophy in alpha-sarcoglycan-deficient mice. *J. Cell Biol.* **142**, 1461–1471
29. Araishi, K., Sasaoka, T., Imamura, M., Noguchi, S., Hama, H., Wakabayashi, E., Yoshida, M., Hori, T., and Ozawa, E. (1999) Loss of the sarcoglycan complex and sarcospan leads to muscular dystrophy in beta-sarcoglycan-deficient mice. *Hum. Mol. Genet.* **8**, 1589–1598
30. Durbeej, M., Cohn, R. D., Hrstka, R. F., Moore, S. A., Allamand, V., Davidson, B. L., Williamson, R. A., and Campbell, K. P. (2000) Disruption of the beta-sarcoglycan gene reveals pathogenetic complexity of limb-girdle muscular dystrophy type 2E. *Mol. Cell* **5**, 141–151
31. Sasaoka, T., Imamura, M., Araishi, K., Noguchi, S., Mizuno, Y., Takagoshi, N., Hama, H., Wakabayashi-Takai, E., Yoshimoto-Matsuda, Y., Nonaka, I., et al. (2003) Pathological analysis of muscle hypertrophy and degeneration in muscular dystrophy in gamma-sarcoglycan-deficient mice. *Neuromuscul. Disord.* **13**, 193–206
32. Coral-Vazquez, R., Cohn, R. D., Moore, S. A., Hill, J. A., Weiss, R. M., Davisson, R. L., Straub, V., Barresi, R., Bansal, D., Hrstka, R. F., et al. (1999) Disruption of the sarcoglycan-

sarcospan complex in vascular smooth muscle: a novel mechanism for cardiomyopathy and muscular dystrophy. *Cell* **98**, 465–474

33. Lebakken, C. S., Venzke, D. P., Hrstka, R. F., Consolino, C. M., Faulkner, J. A., Williamson, R. A., and Campbell, K. P. (2000) Sarcospan-deficient mice maintain normal muscle function. *Mol. Cell. Biol.* **20**, 1669–1677
34. Turk, R., 't Hoen, P. A., Sterrenburg, E., de Menezes, R. X., de Meijer, E. J., Boer, J. M., van Ommen, G. J., and den Dunnen, J. T. (2004) Gene expression variation between mouse inbred strains. *BMC Genomics* **5**, 57
35. 't Hoen, P. A., de Kort, F., van Ommen, G. J., and den Dunnen, J. T. (2003) Fluorescent labelling of cRNA for microarray applications. *Nucleic Acids Res.* **31**, e20
36. 't Hoen, P. A., Turk, R., Boer, J. M., Sterrenburg, E., de Menezes, R. X., van Ommen, G. J., and den Dunnen, J. T. (2004) Intensity-based analysis of two-colour microarrays enables efficient and flexible hybridization designs. *Nucleic Acids Res.* **32**, e41
37. Huber, W., Von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18**, S96–S104
38. R: Development core team (2004) *R: A language and environment for statistical computing*. Vienna, Austria
39. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B.* **57**, 289–300
40. Diehn, M., Sherlock, G., Binkley, G., Jin, H., Matese, J. C., Hernandez-Boussard, T., Rees, C. A., Cherry, J. M., Botstein, D., Brown, P. O., et al. (2003) SOURCE: a unified genomic resource of functional annotations, ontologies, and gene expression data. *Nucleic Acids Res.* **31**, 219–223
41. Zhang, B., Schmoyer, D., Kirov, S., and Snoddy, J. (2004) GOTree Machine (GOTM): a web-based platform for interpreting sets of interesting genes using Gene Ontology hierarchies. *BMC Bioinformatics* **5**, 16
42. Nagy, E., and Maquat, L. E. (1998) A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem. Sci.* **23**, 198–199
43. Inoue, K., Khajavi, M., Ohyama, T., Hirabayashi, S., Wilson, J., Reggin, J. D., Mancias, P., Butler, I. J., Wilkinson, M. F., Wegner, M., et al. (2004) Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. *Nat. Genet.* **36**, 361–369
44. Bulfield, G., Siller, W. G., Wight, P. A., and Moore, K. J. (1984) X chromosome-linked muscular dystrophy (mdx) in the mouse. *Proc. Natl. Acad. Sci. USA* **81**, 1189–1192

45. Noguchi, S., Wakabayashi-Takai, E., Sasaoka, T., and Ozawa, E. (2001) Analysis of the spatial, temporal and tissue-specific transcription of gamma-sarcoglycan gene using a transgenic mouse. *FEBS Lett.* **495**, 77–81
46. Carpenter, S., and Karpati, G. (1979) Duchenne muscular dystrophy: plasma membrane loss initiates muscle cell necrosis unless it is repaired. *Brain* **102**, 147–161
47. Bansal, D., and Campbell, K. P. (2004) Dysferlin and the plasma membrane repair in muscular dystrophy. *Trends Cell Biol.* **14**, 206–213
48. Abbas, A. R., Baldwin, D., Ma, Y., Ouyang, W., Gurney, A., Martin, F., Fong, S., van Lookeren, C. M., Godowski, P., Williams, P. M., et al. (2005) Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun.* **6**, 319–331
49. Porter, J. D., Guo, W., Merriam, A. P., Khanna, S., Cheng, G., Zhou, X., Andrade, F. H., Richmonds, C., and Kaminski, H. J. (2003) Persistent over-expression of specific CC class chemokines correlates with macrophage and T-cell recruitment in mdx skeletal muscle. *Neuromuscul. Disord.* **13**, 223–235
50. Porter, J. D., Merriam, A. P., Leahy, P., Gong, B., and Khanna, S. (2003) Dissection of temporal gene expression signatures of affected and spared muscle groups in dystrophin-deficient (mdx) mice. *Hum. Mol. Genet.* **12**, 1813–1821
51. Porter, J. D., Khanna, S., Kaminski, H. J., Rao, J. S., Merriam, A. P., Richmonds, C. R., Leahy, P., Li, J., Guo, W., and Andrade, F. H. (2002) A chronic inflammatory response dominates the skeletal muscle molecular signature in dystrophin-deficient mdx mice. *Hum. Mol. Genet.* **11**, 263–272
52. Rouger, K., Le Cunff, M., Steenman, M., Potier, M. C., Gibelin, N., Dechesne, C. A., and Leger, J. J. (2002) Global/temporal gene expression in diaphragm and hindlimb muscles of dystrophin-deficient (mdx) mice. *Am. J. Physiol. Cell Physiol.* **283**, C773–C784
53. Haslett, J. N., and Kunkel, L. M. (2002) Microarray analysis of normal and dystrophic skeletal muscle. *Int. J. Dev. Neurosci.* **20**, 359–365
54. Kuznetsov, A. V., Winkler, K., Wiedemann, F. R., von Bossanyi, P., Dietzmann, K., and Kunz, W. S. (1998) Impaired mitochondrial oxidative phosphorylation in skeletal muscle of the dystrophin-deficient mdx mouse. *Mol. Cell. Biochem.* **183**, 87–96
55. Sharma, U., Atri, S., Sharma, M. C., Sarkar, C., and Jagannathan, N. R. (2003) Skeletal muscle metabolism in Duchenne muscular dystrophy (DMD): an in-vitro proton NMR spectroscopy study. *Magn. Reson. Imaging* **21**, 145–153
56. Arber, S., Halder, G., and Caroni, P. (1994) Muscle LIM protein, a novel essential regulator of myogenesis, promotes myogenic differentiation. *Cell* **79**, 221–231

57. Coral-Vazquez, R., Cohn, R. D., Moore, S. A., Hill, J. A., Weiss, R. M., Davisson, R. L., Straub, V., Barresi, R., Bansal, D., Hrstka, R. F., et al. (1999) Disruption of the sarcoglycan-sarcospan complex in vascular smooth muscle: a novel mechanism for cardiomyopathy and muscular dystrophy. *Cell* **98**, 465–474
58. Cohn, R. D., Durbeej, M., Moore, S. A., Coral-Vazquez, R., Prouty, S., and Campbell, K. P. (2001) Prevention of cardiomyopathy in mouse models lacking the smooth muscle sarcoglycan-sarcospan complex. *J. Clin. Invest.* **107**, R1–R7
59. McArdle, A., Edwards, R. H., and Jackson, M. J. (1995) How does dystrophin deficiency lead to muscle degeneration? – evidence from the mdx mouse. *Neuromuscul. Disord.* **5**, 445–456
60. Brussee, V., Tardif, F., and Tremblay, J. P. (1997) Muscle fibers of mdx mice are more vulnerable to exercise than those of normal mice. *Neuromuscul. Disord.* **7**, 487–492
61. Schafer, R., Zweyer, M., Knauf, U., Mundegar, R. R., and Wernig, A. (2005) The ontogeny of soleus muscles in mdx and wild-type mice. *Neuromuscul. Disord.* **15**, 57–64
62. Emery, A., and Muntoni, F. (2003) Clinical features. In: *Duchenne muscular dystrophy* pp. 26–45, Oxford University Press, Oxford.

Received July 15, 2005; accepted September 22, 2005.

Table 1**Phenotype of studied mouse models**

Disease	Model	Affected gene	Age of onset	Histopathological parameters at 8 wk					Reference
				Skeletal dystrophy	Inflammation	Central nuclei	DGC loss	SGC loss	
DMD	mdx	Dystrophin	2-3 wk	Severe	+	+	y	y	(44)
DMD	mdx ^{3cv}	Dystrophin	2-3 wk	Severe	+	+	y	y	(27)
LGMD-2D	Sgca null	α -Sarcoglycan	1 wk	Severe	+	+	n	highly reduced	(28)
LGMD-2E	Sgcb null	β -Sarcoglycan	at least 4 wk	Severe	+	+	n	y	(29,30)
LGMD-2C	Sgcg null	gamma-Sarcoglycan	2 wk	Severe	+	+	n	highly reduced	(31)
LGMD-2F	Sgcd null	delta-Sarcoglycan	2 wk	Severe	+	+	n	highly reduced	(32)
LGMD-2B	Dysf null	Dysferlin	8 wk	Mild/moderate	-	+/-	n	n	(11)
LGMD-2B	Sjl ^{Dysf}	Dysferlin	3 wk	Mild/moderate	n/a	+/-	n/a	n/a	(26)
not known	Sspn null	Sarcospan	None	None	-	-	n	n	(33)

Table 2**Differentially expressed genes in severely vs. mildly and nonaffected models****Up-regulated genes**

Accession	UGCluster	Description	Symbol	P Value*	Fold-change I vs. II
X16834	Mm.248615	Lectin, galactose binding, soluble 3	Lgals3	4.03E-14	10.29
NM_009263	Mm.288474	Secreted phosphoprotein 1	Spp1	2.17E-07	9.28
AJ251685	Mm.302602	Glycoprotein (transmembrane) nmb	Gpnmb	6.08E-12	7.63
L20315	Mm.3999	Macrophage expressed gene 1	Mpeg1	3.26E-14	6.69
NM_008590	Mm.1089	Mesoderm specific transcript	Mest	4.53E-20	6.46
NM_020008	Mm.239516	C-type lectin domain family 7, member a	Clecsf12	4.18E-12	6.40
NM_010745	Mm.2639	Lymphocyte antigen 86	Ly86	6.89E-14	6.33
NM_021443	Mm.42029	Chemokine (C-C motif) ligand 8	Ccl8	2.49E-12	5.88
NM_011593	Mm.8245	Tissue inhibitor of metalloproteinase 1	Timp1	6.97E-11	5.61
NM_011333	Mm.290320	Chemokine (C-C motif) ligand 2	Ccl2	1.06E-08	5.28
U66888	Mm.2254	EGF-like module containing, mucin-like, hormone receptor-like sequence 1	Emr1	2.80E-10	5.18
M55561	Mm.24130	CD52 antigen	Cd52	4.12E-10	4.82
X58196	Mm.14802	H19 fetal liver mRNA	H19	5.79E-16	4.68
NM_009853	Mm.15819	CD68 antigen	Cd68	5.06E-08	4.60
NM_009779	Mm.2408	Complement component 3a receptor 1	C3ar1	2.80E-10	4.58
NM_013590	Mm.177539	P lysozyme structural	Lzp-s	4.81E-09	4.22
NM_007739	Mm.371554	Procollagen, type VIII, alpha 1	Col8a1	3.76E-09	4.21
NM_008871	Mm.250422	Serine (or cysteine) proteinase inhibitor, clade E, member 1	Serpine1	6.30E-11	4.18
NM_008404	Mm.1137	Integrin beta 2	Itgb2	1.82E-08	4.18
NM_007572	Mm.370	Complement component 1, q subcomponent, alpha polypeptide	C1qa	2.55E-15	4.15
NM_013532	Mm.34408	Leukocyte immunoglobulin-like receptor, subfamily B, member 4	Lilrb4	1.48E-11	4.05
NM_011662	Mm.46301	TYRO protein tyrosine kinase binding protein	Tyrobp	2.39E-08	4.02
AF263458	Mm.34609	Placenta-specific 8	Plac8	2.54E-08	3.99
AF246265	Mm.280158	C1q and tumor necrosis factor related protein 3	C1qtnf3	9.81E-04	3.84
M33863	Mm.14301	2'-5' oligoadenylate synthetase 1G	Oas1a	5.31E-14	3.81
NM_011619	Mm.247470	Troponin T2, cardiac	Tnnt2	4.44E-15	3.73
L04694	Mm.341574	Chemokine (C-C motif) ligand 7	Ccl7	4.53E-08	3.71
NM_011175	Mm.17185	Legumain	Lgmn	2.62E-12	3.54
NM_010686	Mm.271868	Lysosomal-associated protein transmembrane 5	Laptm5	5.30E-12	3.51
NM_009814	Mm.15343	Calsequestrin 2	Casq2	8.21E-10	3.51

AF000427	Data not found	B144 mRNA, m17r splice variant		2.48E-11	3.51
NM_021467	Mm.44379	Troponin I, skeletal, slow 1	Tnni1	2.30E-06	3.38
NM_011210	Mm.130953	Protein tyrosine phosphatase, receptor type, C	Ptprc	8.19E-08	3.36
NM_008761	Mm.1870	FXFD domain-containing ion transport regulator 5	Fxyd5	3.31E-05	3.28
NM_011208	Data not found	Protein tyrosine phosphatase, nonreceptor type substrate 1	Ptpns1	6.28E-10	3.27
NM_015783	Mm.358664	Interferon, alpha-inducible protein	G1p2	2.18E-07	3.13
NM_011999	Mm.47384	C-type lectin domain family 4, member a2	Clecsf6	6.69E-14	3.10
D00208	Mm.3925	S100 calcium binding protein A4	S100a4	3.83E-11	3.08
NM_010496	Mm.34871	Inhibitor of DNA binding 2	Idb2	1.30E-09	3.05
D90156	Mm.16528	Myogenin	Myog	3.75E-10	3.03
NM_010730	Mm.248360	RIKEN cDNA C430014K04 gene	Anxa1	7.92E-07	3.03
AJ131395	Mm.297859	Procollagen, type XIV, alpha 1	Col14a1	1.86E-11	2.99
NM_009856	Mm.57175	CD83 antigen	Cd83	1.81E-07	2.97
NM_007599	Mm.18626	Capping protein (actin filament), gelsolin-like	Capg	5.57E-10	2.95
NM_009696	Mm.305152	Apolipoprotein E	Apoe	2.01E-09	2.91
NM_010253	Mm.4655	Galanin	Gal	7.07E-05	2.90
NM_008328	Mm.261270	Interferon activated gene 203	Ifi203	1.34E-09	2.89
NM_007542	Mm.2608	Biglycan	Bgn	1.74E-12	2.86
NM_009970	Data not found	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	Csf2ra	3.67E-07	2.85
NM_007574	Mm.3453	Complement component 1, q subcomponent, gamma polypeptide	C1qg	6.48E-08	2.85
NM_011607	Mm.980	Tenascin C	Tnc	4.28E-05	2.81
NM_009403	Mm.4664	Tumor necrosis factor (ligand) superfamily, member 8	Tnfsf8	9.31E-08	2.80
NM_009448	Mm.88212	Tubulin, alpha 6	Tuba6	9.17E-09	2.78
NM_011701	Mm.268000	Vimentin	Vim	4.81E-09	2.75
NM_011961	Mm.79983	Procollagen lysine, 2-oxoglutarate 5-dioxygenase 2	Plod2	6.89E-14	2.73
NM_009139	Mm.137	Chemokine (C-C motif) ligand 6	Ccl6	3.14E-07	2.69
NM_007793	Mm.6095	Cystatin B	Cstb	4.43E-07	2.61
NM_011671	Mm.171378	Uncoupling protein 2 (mitochondrial, proton carrier)	Ucp2	7.33E-08	2.61
AF149291	Mm.271711	Transgelin 2	Tagln2	3.02E-08	2.58
NM_007651	Mm.316861	CD53 antigen	Cd53	1.44E-07	2.58
M22432	Mm.138471	Similar to elongation factor 1-alpha 1 (EF-1-alpha-1)	Eef1a1	7.41E-07	2.57
NM_011035	Mm.260227	P21 (CDKN1A)-activated kinase 1	Pak1	4.62E-02	2.57
M27347	Mm.2131	Elastase 1, pancreatic	Ela1	3.38E-07	2.54

NM_009468	Mm.8180	Dihydropyrimidinase-like 3	Dpysl3	8.62E-12	2.51
X00496	Mm.276499	Ia-associated invariant chain	Ii	3.69E-10	2.44
NM_009192	Mm.7601	Src-like adaptor	Sla	2.58E-05	2.44
NM_011116	Mm.6483	Phospholipase D3	Pld3	2.75E-06	2.42
M18933	Mm.249555	Procollagen, type III, alpha 1	Col3a1	1.42E-04	2.41
NM_007802	Mm.272085	Cathepsin K	Ctsk	2.16E-07	2.36
NM_011338	Mm.2271	Chemokine (C-C motif) ligand 9	Ccl9	4.56E-04	2.36
NM_011727	Mm.347026	X-linked lymphocyte-regulated 3b	Xlr3b	6.57E-04	2.36
NM_018820	Mm.153684	SERTA domain containing 1	Sertad1	3.31E-05	2.36
NM_011206	Mm.361	Protein tyrosine phosphatase, non-receptor type 18	Ptpn18	2.33E-06	2.35
NM_020575	Mm.260635	Membrane-associated ring finger (C3HC4) 7	Axot	3.69E-04	2.33
NM_009151	Mm.332590	Selectin, platelet (p-selectin) ligand	Selpl	1.12E-05	2.33
	Data not				
X04231	found	4.5S small RNA associated with poly-(a)-containing RNAs		2.69E-05	2.32
NM_016904	Mm.3049	CDC28 protein kinase 1b	Cks1b	2.96E-04	2.31
AB031386	Mm.28385	RIKEN cDNA 1810009M01 gene	1810009M01Rik	2.25E-10	2.29
NM_019467	Mm.10747	Allograft inflammatory factor 1	Aif1	9.04E-08	2.28
NM_011653	Mm.371591	Tubulin, alpha 1	Tuba1	4.89E-07	2.28
	Data not				
AF203898	found	Nebulin		4.15E-05	2.27
AF188290	Mm.220982	Dysferlin	Dysf	7.93E-07	2.26
NM_010634	Mm.741	Fatty acid binding protein 5, epidermal	Fabp5	1.89E-05	2.26
	Data not				
NM_019389	found			6.47E-07	2.25
NM_007737	Mm.10299	Procollagen, type V, alpha 2	Col5a2	1.31E-12	2.25
NM_008597	Mm.243085	Matrix gamma-carboxyglutamate (gla) protein	Mglap	3.10E-11	2.25
NM_009515	Mm.4735	Wiskott-Aldrich syndrome homolog (human)	Was	1.06E-08	2.24
NM_009037	Mm.4876	Reticulocalbin 1	Rcn1	3.66E-09	2.24
NM_010740	Mm.681	Complement component 1, q subcomponent, receptor 1	C1qr1	6.82E-09	2.24
NM_009364	Mm.25612	Tissue factor pathway inhibitor 2	Tfpi2	2.09E-11	2.22
NM_008409	Mm.193	Integral membrane protein 2A	Itm2a	3.29E-10	2.22
AJ006837	Mm.84664	RNA, U17d small nucleolar	Rnu17d	2.99E-06	2.21
NM_008538	Mm.30059	Myristoylated alanine rich protein kinase C substrate	Marcks	3.22E-07	2.18
NM_008035	Mm.2724	Folate receptor 2 (fetal)	Folr2	1.48E-05	2.17
NM_008047	Mm.182434	Follistatin-like 1	Fstl1	1.47E-11	2.17
NM_010566	Mm.15105	Inositol polyphosphate-5-phosphatase D	Inpp5d	1.44E-07	2.15

NM_011618	Mm.358643	Troponin T1, skeletal, slow	Tnnt1	6.73E-03	2.13
NM_008516	Mm.158868	Leucine rich repeat protein 1, neuronal	Lrrn1	5.68E-05	2.13
L10905	Data not found			1.32E-04	2.13
AF068182	Mm.9749	B-cell linker	Blnk	7.74E-04	2.12
AF064749	Mm.7562	Procollagen, type VI, alpha 3	Col6a3	1.67E-08	2.12
NM_008660	Data not found				
NM_019755	Mm.18565	Myosin IF	Myo1f	1.38E-08	2.11
NM_013762	Mm.290899	Proteolipid protein 2	Plp2	1.01E-07	2.11
NM_021394	Mm.116687	Ribosomal protein L3	Rpl3	2.22E-05	2.11
X17501	Mm.1738	Z-DNA binding protein 1	Zbp1	7.99E-04	2.10
NM_011313	Mm.100144	CD48 antigen	Cd48	9.67E-06	2.09
NM_021372	Mm.339676	S100 calcium binding protein A6 (calcyclin)	S100a6	7.79E-07	2.08
NM_010276	Mm.247486	SERTA domain containing 2	Sertad2	7.39E-10	2.07
NM_011065	Mm.7373	GTP binding protein (gene overexpressed in skeletal muscle)	Gem	4.64E-08	2.06
NM_018773	Mm.221479	Period homolog 1 (Drosophila)	Per1	7.58E-05	2.06
NM_013672	Mm.4618	Src family associated phosphoprotein 2	Scap2	5.48E-08	2.05
NM_021474	Mm.276367	Trans-acting transcription factor 1	Sp1	4.90E-05	2.05
NM_016778	Mm.3295	Epidermal growth factor-containing fibulin-like extracellular matrix protein 2	Efemp2	3.34E-08	2.04
NM_015766	Mm.256798	Bcl-2-related ovarian killer protein	Bok	5.94E-06	2.03
NM_008638	Mm.443	Epstein-Barr virus induced gene 3	Ebi3	7.77E-08	2.03
NM_015807	Mm.307174	Methylenetetrahydrofolate dehydrogenase (NAD+ dependent)	Mthfd2	9.42E-06	2.03
NM_008972	Mm.19187	5',3'-nucleotidase, cytosolic	Nt5c	2.97E-05	2.03
NM_011655	Mm.273538	Prothymosin alpha	Ptma	3.66E-04	2.02
NM_010422	Mm.27816	Tubulin, beta 5	Tubb5	2.10E-03	2.02
NM_009750	Mm.90787	G elongation factor, mitochondrial 2	Hexb	1.55E-07	2.02
X62154	Mm.4502	Nerve growth factor receptor (TNFRSF16) associated protein 1	Ngfrap1	3.51E-05	2.02
AF188504	Mm.35600	Minichromosome maintenance deficient 3 (S. cerevisiae)	Mcm3	1.79E-04	2.01
U79144	Mm.250492	Chemokine-like factor super family 7	Cklfsf7	5.66E-05	2.01
NM_007570	Mm.239605	Lysyl oxidase-like 1	Loxl1	2.20E-05	2.01
NM_017380	Mm.38450	B-cell translocation gene 2, anti-proliferative	Btg2	2.28E-06	2.01
NM_019727	Mm.271891	Septin 9	Sep9	1.07E-09	2.00
AF309649	Mm.30241	Sorting nexin 1	Snx1	6.53E-03	2.00
		Interferon gamma inducible protein 30	Ifi30	1.99E-05	2.00

Down-regulated genes

Accession	UGCluster	Description	Symbol	P Value*	Fold-change I vs. II
NM_009463	Mm.4177	Uncoupling protein 1 (mitochondrial, proton carrier)	Ucp1	1.49E-11	-9.53
NM_007702	Mm.449	Cell death-inducing DNA fragmentation factor, alpha subunit-like effector A	Cidea	1.09E-14	-9.28
AF155353	Mm.27159	Ankyrin repeat and SOCS box-containing protein 2	Asb2	1.09E-14	-3.38
NM_011281	Mm.4372	RAR-related orphan receptor gamma	Rorc	1.93E-06	-3.24
NM_020589	Mm.358722	Zinc finger protein 467	Zfp467	5.31E-14	-3.03
NM_008615	Mm.148155	Malic enzyme, supernatant	Mod1	7.91E-09	-2.84
AF060220	Mm.218857	Erythroid associated factor	Eraf	9.06E-03	-2.77
NM_021391	Mm.143788	Protein phosphatase 1, regulatory (inhibitor) subunit 1A	Ppp1r1a	3.59E-11	-2.72
NM_015763	Mm.153625	Lipin 1	Lpin1	3.85E-10	-2.68
NM_008975	Mm.153891	Protein tyrosine phosphatase 4a3	Ptp4a3	2.00E-05	-2.68
AF032131	Mm.23296	E2F transcription factor 6	E2f6	3.80E-05	-2.59
NM_008768	Mm.4777	Orosomuroid 1	Orm1	4.50E-04	-2.54
AJ278735	Mm.38330	Chaperone, ABC1 activity of bc1 complex like (S. pombe)	Cabc1	3.99E-05	-2.54
NM_013849	Mm.317764	Dual specificity phosphatase 13	Dusp13	4.42E-08	-2.52
NM_010585	Mm.227912	Inositol 1,4,5-triphosphate receptor 1	Itpr1	1.45E-09	-2.48
NM_009204	Mm.10661	Solute carrier family 2 (facilitated glucose transporter), member 4	Slc2a4	5.30E-12	-2.47
NM_009161	Mm.18709	Sarcoglycan, alpha (dystrophin-associated glycoprotein)	Sgca	4.59E-03	-2.45
NM_009349	Mm.299	Indolethylamine N-methyltransferase	Inmt	2.41E-05	-2.44
NM_018819	Mm.288510	RIKEN cDNA A630084N20 gene	Brp44l	9.93E-11	-2.35
NM_007489	Mm.12177	Aryl hydrocarbon receptor nuclear translocator-like	Arntl	1.44E-07	-2.35
NM_009624	Mm.310036	Adenylate cyclase 9	Adcy9	1.18E-05	-2.34
NM_008748	Mm.39725	Dual specificity phosphatase 8	5530400B01Rik	2.80E-10	-2.33
NM_008288	Mm.28328	Hydroxysteroid 11-beta dehydrogenase 1	Hsd11b1	1.96E-08	-2.30
NM_021430	Mm.41180	RIKEN cDNA 2900002H16 gene	2900002H16Rik	1.33E-09	-2.27
AF282730	Mm.255607	Tissue inhibitor of metalloproteinase 4	Timp4	1.98E-11	-2.26
NM_009953	Mm.236081	Corticotropin releasing hormone receptor 2	Crhr2	2.98E-08	-2.21
D00926	found	Transcription factor S-II-related protein		7.92E-03	-2.18
NM_013558	Mm.14287	Heat shock protein 1-like	Hspa1l	5.86E-09	-2.18
NM_019435	Mm.30084	Nuclear protein 15.6	Np15	2.90E-03	-2.18
NM_007469	Mm.182440	Apolipoprotein C-I	Apoc1	8.17E-06	-2.17
NM_008377	Mm.245210	Leucine-rich repeats and immunoglobulin-like domains 1	Lrig1	5.14E-06	-2.17
NM_016917	Mm.28756	Solute carrier family 40 (iron-regulated transporter), member 1	Slc40a1	1.92E-07	-2.16
AF288783	Mm.256926	Liver glycogen phosphorylase	Pygl	1.56E-05	-2.16
NM_007377	Mm.6826	Apoptosis-associated tyrosine kinase	Aatk	5.86E-09	-2.15
NM_013459	Mm.4407	Adipsin	Adn	1.20E-02	-2.13
AF267660	Mm.29768	Pyruvate dehydrogenase kinase, isoenzyme 2	Pdk2	8.45E-08	-2.11
NM_009861	Mm.1022	Cell division cycle 42 homolog (S. cerevisiae)	Cdc42	1.62E-06	-2.07

NM_017479	Mm.248967	MYST histone acetyltransferase monocytic leukemia 4	Myst4	2.38E-07	-2.07
NM_011697	Mm.15607	Vascular endothelial growth factor B	Vegfb	1.18E-05	-2.05
X74504	Mm.265990	DNA segment, Chr 16, human D22S680E, expressed	D16H22S680E	1.58E-09	-2.04
X98848	Mm.249131	6-Phosphofructo-2-kinase/fructose-2,6-biphosphatase 1	Pfkfb1	1.20E-07	-2.04
NM_011428	Mm.45953	Synaptosomal-associated protein 25	Snap25	1.59E-07	-2.03
NM_010191	Mm.371560	Farnesyl diphosphate farnesyl transferase 1	Fdft1	5.75E-10	-2.03
NM_016772	Mm.291776	Enoyl coenzyme A hydratase 1, peroxisomal	Ech1	1.45E-10	-2.02
X93035	Mm.38274	Chitinase 3-like 1	Chi3l1	1.46E-06	-2.01
NM_008428	Mm.1482	Potassium inwardly rectifying channel, subfamily J, member 8	Kcnj8	1.71E-09	-2.00

Table 3**Differentially expressed genes between Mdx and Sarcoglycan-deficient mice****Lower expressed in sarcoglycan-deficient models**

Accession	UGCluster	Name	Symbol	P Value*	Fold Change SG-mdx
NM_009441	Mm.213408	Tetratricopeptide repeat domain 3	Ttc3	2.44E-02	-3.12
NM_019567	Mm.297078	Apoptotic chromatin condensation inducer 1	Acin1	8.33E-03	-3.02
NM_009484	Mm.20477	Ubiquitously transcribed tetratricopeptide repeat gene, Y chromosome	Uty	8.38E-03	-2.73
NM_007788	Mm.298893	Casein kinase II, α 1 polypeptide	Csnk2a1	1.63E-03	-2.05

Higher expressed in sarcoglycan-deficient models

Accession	UGCluster	Name	Symbol	P Value*	Fold Change SG-mdx
NM_010217	Mm.1810	Connective tissue growth factor	Ctgf	1.45E-03	2.14
NM_021503	Mm.141157	Myozenin 2	Myoz2	1.85E-03	2.15
NM_020033	Mm.143737	Ankyrin repeat domain 2 (stretch responsive muscle)	Ankrd2	1.84E-05	2.22
NM_013558	Mm.14287	Heat shock protein 1-like	Hspa11	1.71E-02	2.23
NM_013808	Mm.17235	Cysteine and glycine-rich protein 3	Csrp3	4.53E-04	3.73

Fig. 1

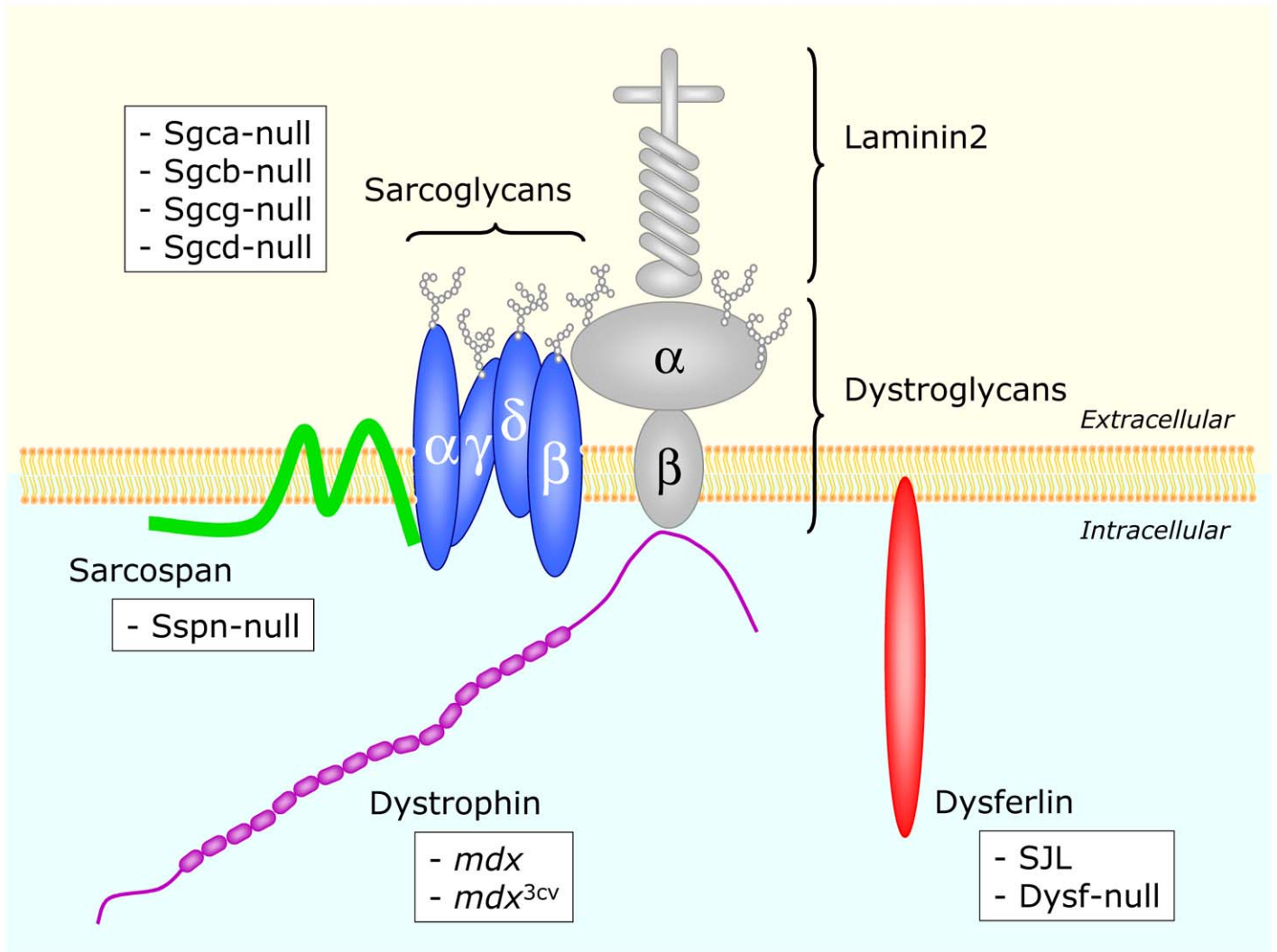


Figure 1. Mouse models for muscular dystrophy related to DGC.

Fig. 2

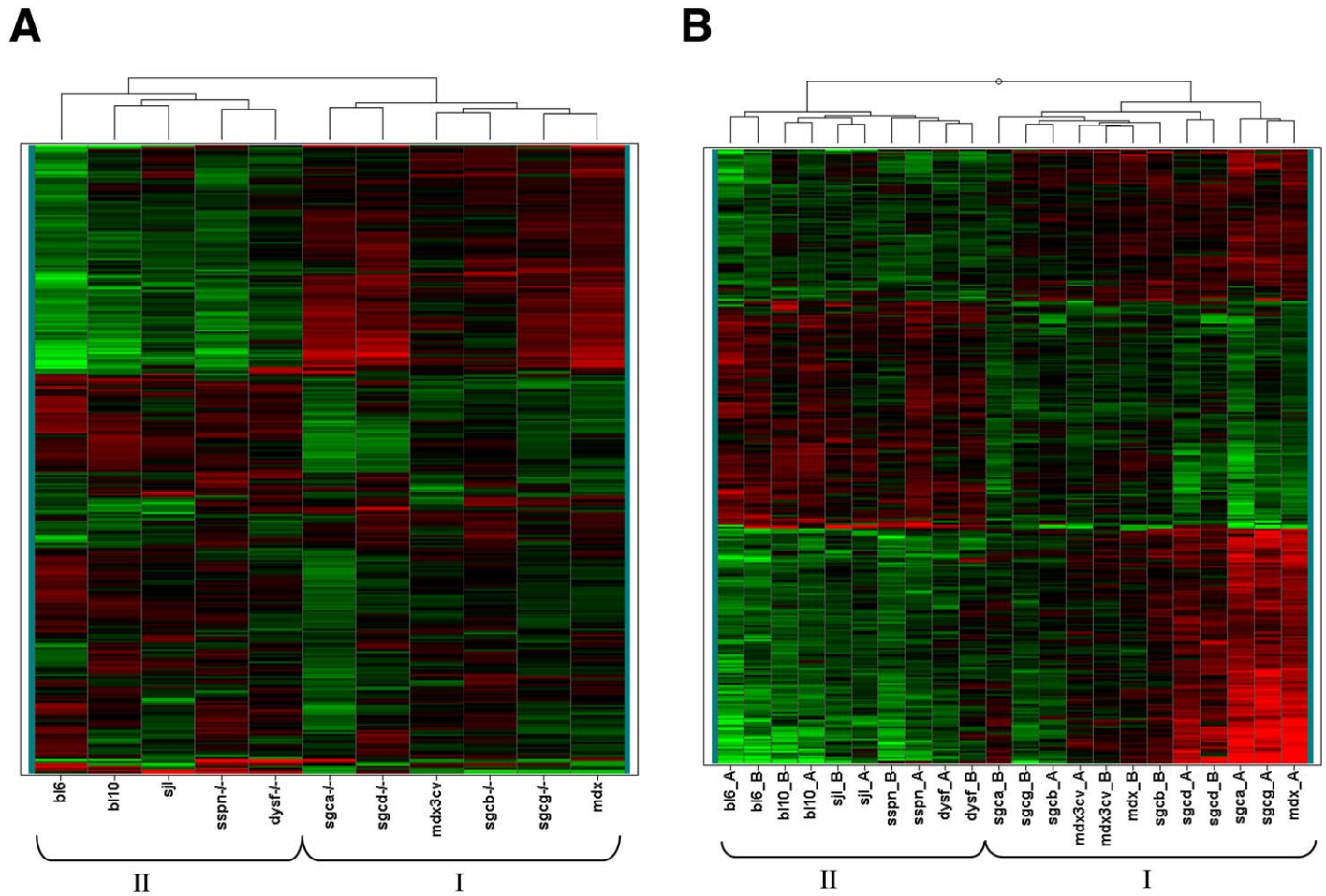


Figure 2. Unsupervised hierarchical clustering of gene expression patterns. **A)** 2-dimensional, unsupervised hierarchical clustering was performed on averaged expression levels of each strain for 300 genes that display most significant (BH-corrected P value $< 4.5 \times 10^{-8}$) differences in expression between strains. For better visualization of up- and down-regulation, gene expression levels were scaled to an average value of 0. Euclidean distance was used as a distance measure. Based on clustering, 2 major groups can be discerned (I and II). **B)** Unsupervised hierarchical clustering was performed on averaged normalized gene expression levels per individual of 300 genes that display most significant (BH-corrected P value $< 4.1 \times 10^{-5}$) differences in expression between severely affected and non- or mildly affected animal models. The clustering and visualization method applied was identical to that in **A**.

Fig. 3

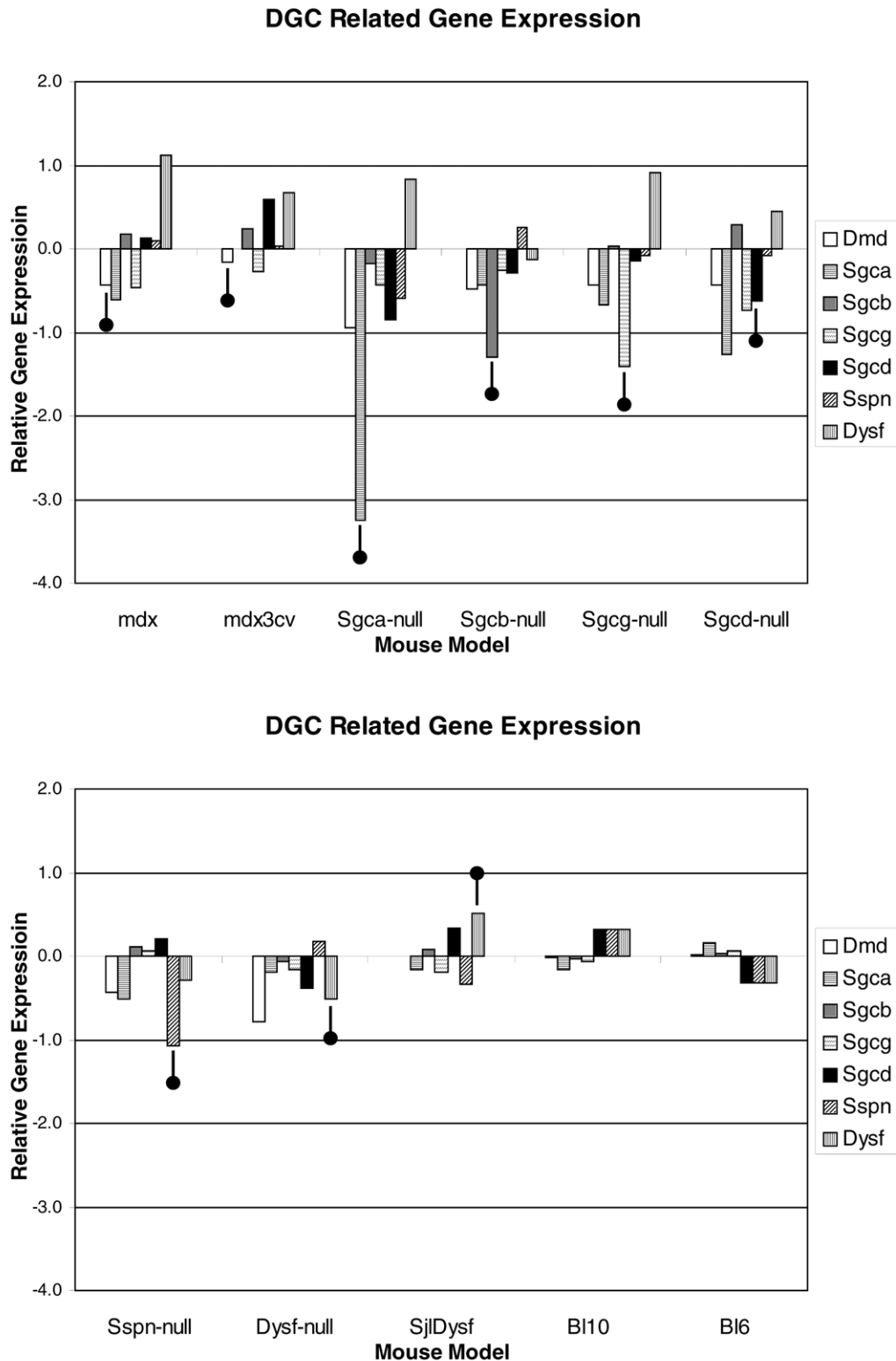


Figure 3. DGC-related genes are down-regulated in severely affected mouse models for muscular dystrophy. Differential gene expression levels between mouse models and wild-type mice were calculated by subtraction of average gene expression level of the 2 wild-type mice (B16 and B110) from gene expression levels of all models (including wild types). Differential expression levels of 7 DGC-related genes are shown (Dmd: dystrophin; Sgca: α -sarcoglycan; Sgcb: β -sarcoglycan; Sgcg: gamma-sarcoglycan; Sgcd: delta-sarcoglycan; Sspn: sarcospan; Dysf: dysferlin). Gene mutated in accompanied model is indicated with ●.

Fig. 4

Upregulated genes							
	Level 1	Level 2	Level 3	Level 4	p-value	number of genes	
Biological process	Cellular process	Cell Communication	Cell adhesion		1.58E-04	36	
	Physiological Process	Response to stimulus	Response to stress	Response to wounding	7.20E-04	21	
				Response to biotic stimulus	1.49E-07	55	
				Defense respons	4.94E-08	51	
					1.73E-04	73	
					4.25E-04	6	
	Organismal physiological process	Muscle Contraction	Regulation of muscle contraction				
	Cellular Component	Extracellular region	Extracellular matrix	Basement membrane	5.60E-04	9	
				Collagen	9.07E-08	12	
Extracellular space				1.52E-05	135		
				1.49E-05	145		
				1.09E-04	32		
Intracellular		Intracellular organelle	Intracellular membrane-bound organelle	Lysosome	4.46E-04	13	
Cytoplasm	Contractile fiber	Troponin complex		7.37E-04	4		
Molecular Function	Binding	Protein binding		2.39E-04	138		
	Signal transducer activity	Receptor activity	Transmembrane receptor activity	Hematopoietin/interferon class (D200-domain) cytokine receptor activity	8.85E-04	10	
					4.19E-04	37	
					1.71E-08	16	
	Structural molecule activity	Extracellular matrix structural constituent	Extracellular matrix structural constituent conferring tensile strength	1.91E-07	11		
Downregulated genes							
	Level 1	Level 2	Level 3	Level 4	p-value	number of genes	
Biological process	Physiological process	Cellular physiological process	Cellular metabolism		6.97E-06	328	
					1.56E-09	259	
				Cofactor metabolism	1.28E-06	19	
				Organic acid metabolism	1.04E-04	28	
				Lipid metabolism	2.71E-05	35	
				Alcohol metabolism	2.88E-07	10	
				Carbohydrate metabolism	2.48E-12	40	
	Cellular component	Intracellular	Intracellular organelle	Mitochondrion	3.07E-08	228	
				Mitochondrial membrane	5.48E-24	72	
				3.29E-11	23		
	Cytoplasm	Cytosol	Proteasome core complex	5.13E-04	6		
Molecular function	Catalytic activity	Hydrolase activity	Endopeptidase activity				
			Threonine endopeptidase activity	8.31E-04	6		
			Isomerase activity	Heme-copper terminal oxidase activity	1.35E-04	14	
					8.46E-05	6	
			Oxidoreductase activity	Oxidoreductase acting on CH-OH group of donors	Acting on CH-OH group of donors, NAD or NADP as acceptor	1.89E-10	52
						1.59E-04	12
					Oxidoreductase acting on heme group of donors	3.78E-04	11
					Acting on heme group of donors, oxygen as acceptor	8.46E-06	6
					Oxidoreductase acting on the aldehyde or oxo group of donors	8.46E-06	6
				1.49E-04	7		
Transferase activity	7.40E-06	80					

Figure 4. Functional classification of significantly up- and down-regulated genes in severely affected mouse models. Genes up-regulated in dystrophic muscle from severely affected mouse models were grouped according to biological process, cellular component, and molecular function, based on Gene Ontology classifications. Only branches of the GO-tree containing categories that were significantly overrepresented (displayed in red; $P < 0.001$) in the list of up (*upper panel*)- and down-regulated (*lower panel*) genes are shown. Listed P values are from a hypergeometric test that compares, for each category, number of genes in the set of up- or down-regulated genes with total number of genes present on the array in that category. Number of genes refers to number of genes in the list of up- or down-regulated genes in a specific category. A specification of genes present in overrepresented categories can be found in Supplemental Table 4.

Fig. 5

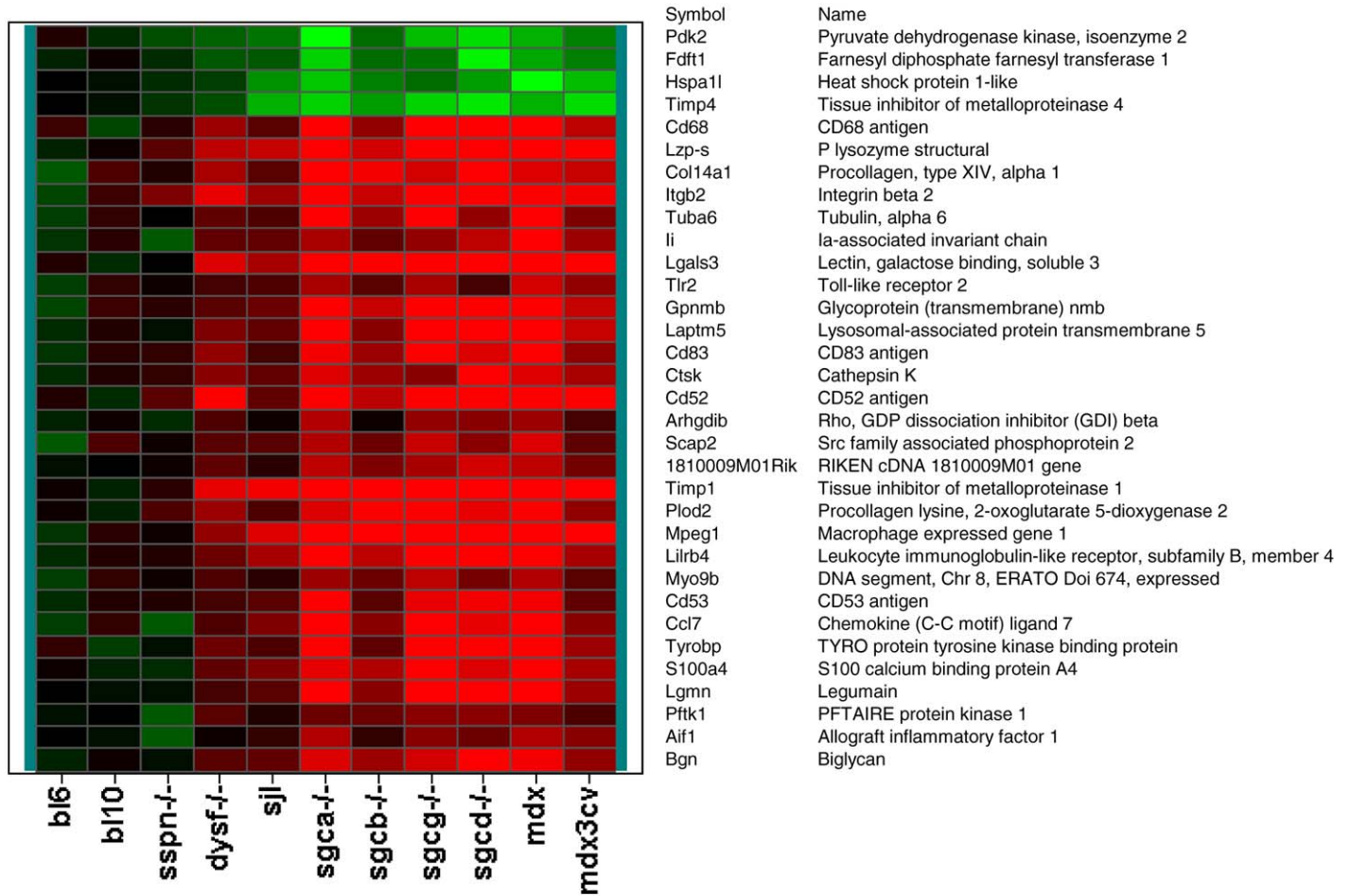


Figure 5. Biomarker genes for disease progression in muscular dystrophy heat map of averaged expression levels (relative to levels in wild-type mice) of genes that correlate with disease progression. The top 4 genes demonstrate significantly lower expression (displayed in green) in dysferlin-deficient mice compared with wild-type and sarcospan-deficient mice and even lower expression levels in the more severely affected mouse models, whereas the bottom 29 genes demonstrate significantly higher (displayed in red) expression in dysferlin-deficient mice compared with wild-type and sarcospan-deficient mice and even higher expression levels in the more severely affected mouse models.