Methods

Additional Information on Human Samples. All NM biopsies were subjected to standard neuropathological examination of cryosections stained with hematoxylin and eosin, modified Gomori trichrome, nicotinamide-adenine dinucleotide (NADH) reductase, and myosin ATPase at pH 4.3 and 9.4. Fiber-typing was further confirmed and quantified by using immunohistochemistry of fast and slow myosin heavy chains as described below.

Based on the clinical and pathological information on each patient, samples were classified into subgroups as defined by European Neuro-Muscular Centre criteria, namely severe, intermediate, or typical (1) (Table 1). Twenty-one normal control tissues came from surgical samples from unaffected individuals (n = 15) and autopsy specimens (n = 6). All 21 samples were selected and assessed as previously described (2). All samples were obtained under institutionally approved protocols, snap frozen, and stored in liquid nitrogen. Preliminary comparative analysis of 22 NM and normal samples, in relation to biopsies from patients with inflammatory myopathies, has been recently reported (2).

RNA Target Preparation and Hybridization. Approximately 6-8 μg of RNA was extracted from each specimen using Trizol (Life Technologies, Rockville, MD) and chloroform/isoamyl alcohol (49:1), and was used for target preparation. Each target was individually hybridized on both U95Av2 (12,626 probe sets) and U95B (12,620 probe sets) GeneChip oligonucleotide arrays (Affymetrix). Following hybridization, the signal amplification staining option was chosen on the Affymetrix Fluidics Station 400, and the GeneChips were scanned by the Affymetrix HP Gene Array Scanner. The expression level and a "present", "absent", or "marginal" call was assigned to each probe set by the MAS5.0 Affymetrix software. Affymetrix MAS5.0 software scored 4,537 (18%) probe sets "present" across all 21 normal samples, 4,165 (16.5%) "present" across all 13 NM samples, and 3,673 (14.5%) probe sets "present" across all 34 samples, indicating that these transcripts were present at sufficient levels for accurate quantitation. Correlation coefficients (r) between signals from replicate microarrays were typically about 0.99 when hybridized with the same targets and about 0.97 when hybridized with independently isolated and labeled RNA samples.

Quantitative Real-Time Reverse Transcriptase PCR (TaqMan). "Assays-on-Demand" gene expression kits were obtained for *PFKFB1* (Hs00159997 m1), *UCP3* (Hs00243297 m1), *BCL6* (Hs00153368 m1), *FADS1* (Hs00203685 m1), *FEZ1* (Hs00192714 m1), and *PHKA* (Hs00267497 m1) (Applied Biosystems, Foster City, CA). RNA (1 μg) from different NM and normal samples was reverse-transcribed to cDNA using reverse transcription reagent kits (Roche Molecular Systems, Branchburg, NJ). The resulting cDNA was amplified in a real-time quantitative PCR using the TaqMan Universal PCR Master Mix, the recommended TaqMan protocol and the ABI 7700 Sequence Detection System (Applied Biosystems). A standard curve for serial dilutions of 18S rRNA was obtained in a similar fashion. The outcome of each amplification run was compared to the standard curve as described by the Taqman protocol, (Applied Biosystems), enabling the calculation of the RNA expression level fold changes between NM and normal samples.

Antibodies. Primary antibodies for IF and immunoblotting (IB) were obtained from indicated sources and used at the following dilutions: 1:500-IF, 1:400-IB for CDK4 (ab6315) mouse monoclonal (Abcam, Cambridge, U.K.); 1:600-IF, 1:100-IB for TGFbeta2 (sc-90) rabbit polyclonal (Santa Cruz Biotechnology, Santa Cruz, CA); 1:500-IF, 1:400-IB for tubulin TY-06 (ab7792) mouse monoclonal (Abcam, Cambridge, U.K.); 1:100-IF CD56 (anti-NCAM1) (31660D) mouse monoclonal (PharMingen International, San Jose, CA); 10 μg/ml-IB for uncoupling protein 3 (UCP32-A) (Alpha Diagnostic International, San Antonio, TX); 1:100-IF PAX7 mouse monoclonal (Developmental Studies Hybridoma Bank, University of Iowa, Iowa City, IA); 1:30 myosin heavy chain-developmental (NCL-MHCd), 1:20 myosin heavy chain-neonatal (NCL-MHCn), 1:200 myosin heavy chain-fast (NCL-MHCf), and 1:50 myosin heavy chain-slow (NCL-MHCs) (Novocastra Laboratories, Newcastle upon Tyne, U.K.). All secondary antibodies for IF were diluted at 1:400: AlexaFluor 488 goat anti-rabbit IgG (A11008), AlexaFluor 488 (A11001), and 594 (A11005) goat anti-mouse IgG (Molecular Probes, Eugene, OR). The secondary antibodies for immunoblotting were peroxidase-conjugated AffiniPure donkey anti-mouse IgG (715-035-150) and donkey anti-rabbit IgG (711-035-152) at 1:5,000 dilution (Jackson ImmunoResearch, West Grove, PA).

Western Blotting. Protein extracts of NM and normal tissue samples for immunoblotting were prepared in lysis buffer (100 mM Tris, pH 6.8, 4% SDS, 20% glycerol, 282 mM beta-mercaptoethanol) and boiled for 10 min. Proteins were quantitated by spectrophotometry using a Bio-Rad DC Protein Assay (reagents A,B,S) (Bio-Rad Laboratories, Hercules, CA). Total protein (100ng) was separated by electrophoresis on a SDS/15% polyacrylamide gel and transferred to nitrocellulose membranes by electroblotting. The primary and secondary antibodies used (as listed above) against selected proteins were visualized by enhanced chemiluminescence ECL (Amersham Pharmacia Biotech, Buckinghamshire, UK). Equivalent loading of lanes was determined by Ponceau S staining.

Data Analysis. The data are first normalized (3-5), wherein the vector of signals in each array is linearly transformed so that their scatter plots have a linear regression of slope 1 through the origin against a reference data set. Reference data sets were composed of the average probe-by-probe signals from a normal (T146) and an NM (T80) experiment which had maximal average correlation coefficients against all other chip experiments within their respective disease classes. This normalization corrects any uniform linear aberrations of the reported signals between any two replicate measurement which may arise from idiosyncrasies in the hybridization or scanning protocol. In this context, we say that microarrays A1 and A2 are replicate measurements when both microarrays assay labeled total RNA which originated from patients (not necessarily the same patient) within a common disease class. (Note that the data are heterogeneous in the sense that each disease category consists of at least four prominent overlapping patient non-disease-related subcategories: Adult, Pediatric, Autopsy, and Surgery.) Our expectation is that the reported signals from A1 and A2 ought to be "close" to one another on average, i.e., at most an insignificant percentage of genes out of the 12,625 will have markedly different reported expression levels, and these are due largely to individual biological variation. Graphically,

this translates into the scatter plot of A1 versus A2 having data points mostly along the $x = y$ line and with minimal scatter.

Two statistical methods were applied to determine differential gene expression between the two disease classes:

(A) an unequal variance, two-tailed *t* test which looks for genes whose signal differences are significant between normal (NO) and NM; and

(B) a method in ref. 5 which assesses the differentiability of a gene by its signal fold change in the NM with respect to the NO class.

Both techniques assess changes in signal difference or fold across the two classes against changes in difference or fold within the classes, i.e., among duplicates. Unless otherwise noted, the terms data or signal will denote the *normalized* data or signal.

(A) Difference.

We applied the two-tail unequal variance *t* test following ref. 6 and selected genes whose signals had a $P < 0.001$ of a significant difference between NO and NM. 72 and 56 genes (U95Av2 and U95B, respectively) were found to be thus up-regulated, whereas 152 and 126 (U95Av2 and U95B, respectively) were down-regulated from NO to NM.

(B) Fold change.

We briefly describe this method (5): Let $a_1, a_2, \cdots, a_M$, and $b_1, b_2, \cdots, b_M$, be M replicate signals for an arbitrary gene in the NO and NM class respectively. We calculate the logarithmic geometric mean (μ) and variance (σ^2) of the fold change of the gene between the two classes as follows:

$$\mu = \left(\frac{\prod_j^M b_j}{\prod_j^M a_j}\right)^{\frac{1}{M}} = \frac{1}{M}\sum_{j=1}^{M}\log\left(\frac{b_j}{a_j}\right),$$

$$\sigma^2 = \frac{1}{M}\sum_{j=1}^{M}\left(\log\left(\frac{b_j}{a_j}\right) - \mu\right)^2 .$$

[1]

Due to non-disease-related phenotype heterogeneity and unequal number of samples in each disease class, the folds were formed from a_j and b_k that have been phenotypically matched, i.e., a_j and b_k are not signals from individual arrays but are the average signals over respective subcategories. $M = 4$ is the number of distinct phenotypic subcategories arising from the combination of Adult, Pediatric, Autopsy, and Surgery class labels.

Next, rearrange a_j and b_j so that $a_1 \geq a_2 \geq \cdots \geq a_M$, and $b_1 \geq b_2 \geq \cdots \geq b_M$. For the same gene, we also calculate the geometric average fold changes arising within each disease class denoted μ_{NO}, μ_{NM} which is attributable to measurement variability. Let m be the largest integer that is less than $M/2$. By design, μ_{NO} and μ_{NM} are nonnegative.

$$\mu_{NO} = \left(\frac{\prod_j^m a_{m+1-j}}{\prod_j^m a_j} \right)^{\frac{1}{m}} = \frac{1}{m} \sum_{j=1}^{m} \log \left(\frac{a_{m+1-j}}{a_j} \right) ,$$

[2]

$$\mu_{NM} = \left(\frac{\prod_j^m b_{m+1-j}}{\prod_j^m b_j} \right)^{\frac{1}{m}} = \frac{1}{m} \sum_{j=1}^{m} \log \left(\frac{b_{m+1-j}}{b_j} \right) .$$

[3]

We say that the gene is significantly up-regulated when

$$\mu - \sigma > 0.24, \qquad \mu > \min(\mu_{NO}, \mu_{NM}) ,$$

[4]

and down-regulated when

$$\mu + \sigma < -0.24, \qquad \mu < -\min(\mu_{NO}, \mu_{NM}) .$$

[5]

The choice of 0.24 is arbitrary. The false discovery rate (7) of the method with these parameters is found to be 0.33 following a permutation test whereby the disease class labels of the microarrays are randomly shuffled, the resulting data configuration is re-analyzed, and the procedure is repeated 10,000 times. 297 and 157 genes (U95Av2 and U95B, respectively) were found to be up-regulated, whereas 159 and 153 (U95Av2 and U95B, respectively) were down-regulated in the NM with respect to the NO class.

Hierarchical Clustering, Relevance Network, and Nearest-Neighbor Analyses. Hierarchical clustering was performed as described by Greenberg *et al* (2), using the Cluster and TreeView Software (8) with centered linear correlation as a measure of similarity using average linkage. Relevance networks, a nonhierarchical clustering method that computes

Pearson correlation coefficients between all genes and/or quantitative phenotypes (i.e., "factors") in microarray expression datasets, were computed as described (9). Correlation coefficient cutoffs were selected to allow the display of only those links between factors with similar expression patterns, allowing for identification of three-dimensional expression relationships between groups of factors.

To determine which genes had the most similar expression to each member of the set of "reliably changed" transcripts, we performed nearest-neighbor analyses (10) using the Pearson correlation coefficient as a distance measure (web supplement *Nearest Neighbors* available at http://www.tch-genomics.org/beggslab/). For each selected gene, neighbors that were at a distance of 0.6 or closer were chosen. This distance was selected based on a permutation analysis (9), to determine how often such correlations occurred by chance. If there were several such neighbors, then only the closest 50 were selected.

References

1. Wallgren-Pettersson, C. & Laing, N. G. (2001) *Neuromuscul. Disord.* **11**, 589-595.
2. Greenberg, S. A., Sanoudou, D., Haslett, J. N., Kohane, I. S., Kunkel, L. M., Beggs, A. H. & Amato, A. A. (2002) *Neurology* **59**, 1170-1182.
3. Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E. S. & Golub, T. R. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 2907-2912.
4. Butte, A. J., Ye, J., Haring, H. U., Stumvoll, M., White, M. F. & Kohane, I. S. (2001) *Pac. Symp. Biocomput.*, 6-17.
5. Zhao, Q., Kho, A., Kenney, A. M., Yuk Di, D. I., Kohane, I. & Rowitch, D. H. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 5704-5709.
6. Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992) *Numerical Recipes: The Art of Scientific Computing* (Cambridge Univ. Press, Cambridge, U.K.).
7. Tusher, V. G., Tibshirani, R. & Chu, G. (2001) *Proc Natl Acad Sci USA* **98**, 5116-5121.
8. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 14863-14868.
9. Butte, A. J., Tamayo, P., Slonim, D., Golub, T. R. & Kohane, I. S. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 12182-12186.
10. Dasarathy, B. (1991) *Nearest Neighbor Norms: NN Pattern Classification Techniques* (IEEE Computer Society Press, Los Alamos, CA).