

Research article

Open Access

Transcriptome-scale similarities between mouse and human skeletal muscles with normal and myopathic phenotypes

Alvin T Kho^{†1,2}, Peter B Kang^{†2,3}, Isaac S Kohane^{*1} and Louis M Kunkel^{*2,4}

Address: ¹Children's Hospital Informatics Program, Harvard-MIT Division of Health Sciences and Technology, Cambridge, Massachusetts, USA, ²Program in Genomics, Children's Hospital Boston and Harvard Medical School, Boston, Massachusetts, USA, ³Department of Neurology, Children's Hospital Boston and Harvard Medical School, Boston, Massachusetts, USA and ⁴Howard Hughes Medical Institute, Children's Hospital Boston, Boston, Massachusetts, USA

Email: Alvin T Kho - alvin_kho@hms.harvard.edu; Peter B Kang - peter.kang@childrens.harvard.edu; Isaac S Kohane* - isaac_kohane@harvard.edu; Louis M Kunkel* - kunkel@enders.tch.harvard.edu

* Corresponding authors †Equal contributors

Published: 07 March 2006

Received: 14 December 2005

BMC Musculoskeletal Disorders 2006, 7:23 doi:10.1186/1471-2474-7-23

Accepted: 07 March 2006

This article is available from: <http://www.biomedcentral.com/1471-2474/7/23>

© 2006 Kho et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Mouse and human skeletal muscle transcriptome profiles vary by muscle type, raising the question of which mouse muscle groups have the greatest molecular similarities to human skeletal muscle.

Methods: Orthologous (whole, sub-) transcriptome profiles were compared among four mouse-human transcriptome datasets: (M) six muscle groups obtained from three mouse strains (wildtype, *mdx*, *mdx^{5cv}*); (H1) biopsied human quadriceps from controls and Duchenne muscular dystrophy patients; (H2) four different control human muscle types obtained at autopsy; and (H3) 12 different control human tissues (ten non-muscle).

Results: Of the six mouse muscles examined, mouse soleus bore the greatest molecular similarities to human skeletal muscles, independent of the latter's anatomic location/muscle type, disease state, age and sampling method (autopsy versus biopsy). Significant similarity to any one mouse muscle group was not observed for non-muscle human tissues (dataset H3), indicating this finding to be muscle specific.

Conclusion: This observation may be partly explained by the higher type I fiber content of soleus relative to the other mouse muscles sampled.

Background

Animal models of human diseases are used extensively to study basic disease processes and test potential therapies. Even though these proxies have ethical and practical advantages, they generally do not completely recapitulate the human disease phenotype. For example, *mdx* and *mdx^{5cv}* mice have mutations in the dystrophin gene mirroring the genetic defect of human Duchenne muscular dystrophy (DMD) [1-4], yet they experience milder mus-

cle degeneration than DMD patients [3,5,6]. Consequently, extrapolating findings from a mouse model to human disease can have limitations.

Different skeletal muscle groups are dissimilarly affected in muscular dystrophies, suggesting inherent molecular and physiological differences among muscle groups. This raises the question of whether one type of mouse muscle more accurately represents particular (myopathic) charac-

teristics in a given human muscle type than another. Based on gross histology, skeletal muscles differ in at least four parameters: bulk, length, fiber architecture and fiber type proportions. It is not immediately clear how to prioritize these for cross species comparisons. Recent transcriptome studies of different skeletal muscle groups in humans [7-10] and mice [11-16] raise the related question of whether and how muscle groups are related from a molecular perspective, within and across species.

This study investigates the similarities between mouse and human skeletal muscle gene expression profiles at the level of the whole transcriptome and four distinct sub-transcriptomes. Of the six mouse muscle groups considered, soleus was found to be most similar to all human muscles tested, independent of the latter's anatomic location/muscle type, disease state, age, and sampling method (autopsy versus biopsy). These observations were consistent regardless of the sub-transcriptome used to characterize each sample profile.

Methods

Transcriptome datasets (M, H1, H2, H3)

The gene expression datasets originate from four separate, previously-published skeletal muscle studies (one on mouse and three on humans) performed on the Affymetrix GeneChip platform (Table).

Dataset M (mouse, $n = 36$) was derived from six different skeletal muscle groups: diaphragm, extensor digitorum longus, gastrocnemius, quadriceps, soleus, and tibialis anterior of eight-week-old male mice. The mice represented three genetic strains: wildtype (C57BL/10Sn), *mdx* (C57BL/10ScSn-Dmd^{mdx/J}), and *mdx*^{5cv} (B6Ros.Cg-Dmd^{mdx-5cv}), with their transcriptome profiles assayed in duplicate on the Affymetrix U74Av2 platform [17,18].

Dataset H1 (human, $n = 24$) consists of transcriptome profiles of surgically-biopsied quadriceps from 12 control and 12 DMD subjects assayed on the Affymetrix U95Av2 platform [8,19].

Dataset H2 (human, $n = 32$) contains transcriptome profiles of four different human muscle groups: deltoid, gastrocnemius, quadriceps, and tibialis anterior obtained at autopsy from each of 8 individuals – three pediatric and five geriatric – assayed on the Affymetrix U133A platform [20].

Dataset H3 (human, $n = 24$) contains transcriptome profiles of 12 distinct normal organ-tissues pooled from 10–25 individuals – bone marrow, brain, heart, liver, lung, kidney, skeletal muscle, pancreas, prostate, spinal cord, spleen, and thymus – assayed on the Affymetrix U95Av2 microarrays in duplicate [21,22].

Mouse-human gene orthologues

Curated and predicted National Center for Biotechnology Information (NCBI) Entrez-identified mouse-human orthologue gene pairs were obtained from NCBI HomoloGene, (data freeze October 07, 2005) [23]. To ensure a one-to-one and well-defined mapping, cross-species matches were restricted to those with the highest reciprocal percentage sequence homology. For instance, the human gene *ACAT2* (EntrezID 39) has two distinct mouse orthologues: *Acat2* (87.12% sequence homology, EntrezID 110460) and *Acat3* (85.10% sequence homology, EntrezID 224530). *Acat2* was considered the human *ACAT2*'s unique mouse-orthologue since it had the higher percentage homology with *ACAT2*.

The Affymetrix GeneChip platforms contain the following number of Entrez-identified genes (data freeze May 17, 2005): U74Av2, 8,894 mouse genes; U95Av2, 8,978 human genes; U133A, 12,963 human genes. Between the mouse U74Av2 and human U95Av2 platforms, there exist 5,306 mouse-human orthologue pairs. Similarly, mouse U74Av2 and human U133A platforms have 6,547 orthologue pairs.

Mathematical analyses

The primary gene set for cross-species analyses are the 5,306 orthologue gene pairs between the mouse Affymetrix U74Av2 (dataset M) and human Affymetrix U95Av2 (dataset H1 and H3) platforms. Of these 5,306 genes, 5,288 (99.7%) are shared in common with the 6,547 U74Av2/U133A orthologue pairs, and were thus used in cross-species analyses between datasets M and H2 (Affymetrix U133A). Each sample gene expression profile is represented as algebraic N -vector signifying the reported expression intensities for N (>0) distinct genes measured in that sample.

Principal component analysis (PCA) was used to obtain the subsets of genes that were dominant contributors to the global sample variation in datasets M and H1, respectively [see Additional file 1] [see Additional file 2] [see Additional file 3] [24-27]. We define the g_j 's with $|a_j| > 0.03$ in any one of principal components 1–3 (PC1-3) following PCA of datasets M (271 genes) and H1 (234 genes) to be the dominant contributors to global sample variation for datasets M and H1 respectively. Eighty-six genes are common to both sets.

Linear (Pearson) correlation was used to quantify the linear similarity between mouse and human samples [24]. Each mouse or human sample is a gene expression profile vector of length N , with the j th mouse vector component being orthologous to the j th human vector component. When calculating the mouse-human profile correlations at the whole transcriptome-scale, $N = 5,306$, the total

number of orthologue gene pairs between the mouse U74Av2-human U95Av2 platforms; and $N = 234$ (respectively, $N = 271$ – see preceding paragraph) when calculating the mouse-human profile correlations based on the subset of genes that are dominant contributors to global sample variance in dataset H1 (respectively, dataset M).

The non-parametric Wilcoxon (Mann-Whitney) rank-sum test [28] was used to assess the probability that the distribution of linear correlations of human skeletal muscle to mouse soleus is similar to the distribution of linear correlations of human skeletal muscle to mouse non-soleus muscles. The underlying expression data distributions were non-Gaussian, thus a parametric test was not used.

Results

Whole transcriptome scale correlations between mouse and human skeletal muscles

Whole transcriptome scale (5,306-gene profile) similarities between skeletal muscle datasets M (mouse) and H1 (human) were investigated. Linear correlation was used as a measure of similarity. Two additional human transcriptome datasets were considered as *in silico* positive controls: (H2) transcriptome profiles from four skeletal muscle groups of eight human subjects from autopsies measured on the Affymetrix U133A platform [20]; and (H3) transcriptome profiles from 12 distinct human tissues (two muscle, ten non-muscle) [21] measured on the same microarray platform as H1.

Linear correlations between 5,306-gene transcriptome profiles of each sample from dataset M to each sample from dataset H1 were computed. The human correlation group averages and standard deviations relative to each of the six mouse muscle groups \times three genetic strains are shown in Figure 1A–B. All 24 human quadriceps samples had a significantly higher correlation to mouse soleus than to non-soleus mouse muscles ($p < 0.02$, Wilcoxon), independent of mouse genetic background and human myopathic condition.

To determine whether the higher correlation of human quadriceps to mouse soleus could be reproduced with other human muscles, similar correlation analyses were performed with human datasets H2 and H3 relative to dataset M. Cross-species linear correlation analyses of the 5,288-gene pairs selected between datasets M and H2 showed a significantly higher correlation in 28 of the 32 human samples to mouse soleus ($p < 0.02$, Wilcoxon) than to non-soleus mouse muscles (Figure 1C–D).

To determine whether the higher correlation of human skeletal muscles to mouse soleus is specific to muscle tissue, similar correlation analyses were performed with the 5,306-gene transcriptome profiles of human tissue not

primarily composed of muscle (bone marrow, brain, liver, lung, kidney, pancreas, prostate, spinal cord, spleen and thymus) in dataset H3 relative to dataset M. The non-muscle human profiles did not return a significantly higher correlation to any mouse muscle, whereas the pooled human heart and skeletal muscle profiles in dataset H3 showed a significantly higher correlation to the mouse soleus than to non-soleus mouse muscles ($p < 0.02$, Wilcoxon) (Figures 1A, 1C). Indeed, non-muscle human tissue-to-mouse muscle correlations were significantly lower than human muscle-to-mouse muscle correlations.

Sub-transcriptome scale correlations between mouse and human skeletal muscles

We identified candidate subsets of genes (by their ontologic category or contribution to sample variation *in silico*) that might underwrite the transcriptomic resemblance of human muscle to mouse soleus. Correlation analyses of mouse-human sample profiles were performed as above, relative to four subsets of the 5,306-gene whole transcriptome separately: ST1, the 733-gene sub-transcriptome (13.8% of the whole transcriptome) recorded by the Gene Ontology Consortium [29] to be integral to cell plasma membrane – note that dystrophin and other sarcolemma-related structural proteins involved in myopathies belong to this ontologic category; ST2, the 271-gene sub-transcriptome (5.1% of the whole transcriptome) of dominant contributors to global sample variance in dataset M; ST3, the 234-gene sub-transcriptome (4.4% of the whole transcriptome) of dominant contributors to global sample variance in dataset H1; and ST4, the 4,188-gene sub-transcriptome complement to ST1-3 as a negative *in silico* control – these are genes neither integral to the cell plasma membrane nor dominant contributors to global sample variance in the muscle datasets M and H1. Twenty-one genes in ST2, and 16 in ST3, belong to the cell plasma membrane category/ST1.

Using ST1-3 to characterize sample profiles, we again observed significantly higher correlations of human muscles in datasets H1-3 to mouse soleus than to non-soleus mouse muscles ($p < 0.02$, Wilcoxon) [see Additional file 4] [see Additional file 5] [see Additional file 6]. Similarly, correlations of non-muscle human tissue in dataset H3 to mouse dataset M are lower and do not show a significantly higher correlation to any mouse muscle group – with one exception in the ST3 sample profile characterization where among non-muscle human tissue, brain and spinal cord had significantly higher correlation with mouse soleus than with non-soleus muscle profiles. In contrast, when ST4 was used to characterize sample profiles, all human muscle and non-muscle samples from dataset H1-3 have lower correlations to mouse muscle profiles in dataset M compared to their ST1-3 sample profile characterization, and human muscle profiles show no

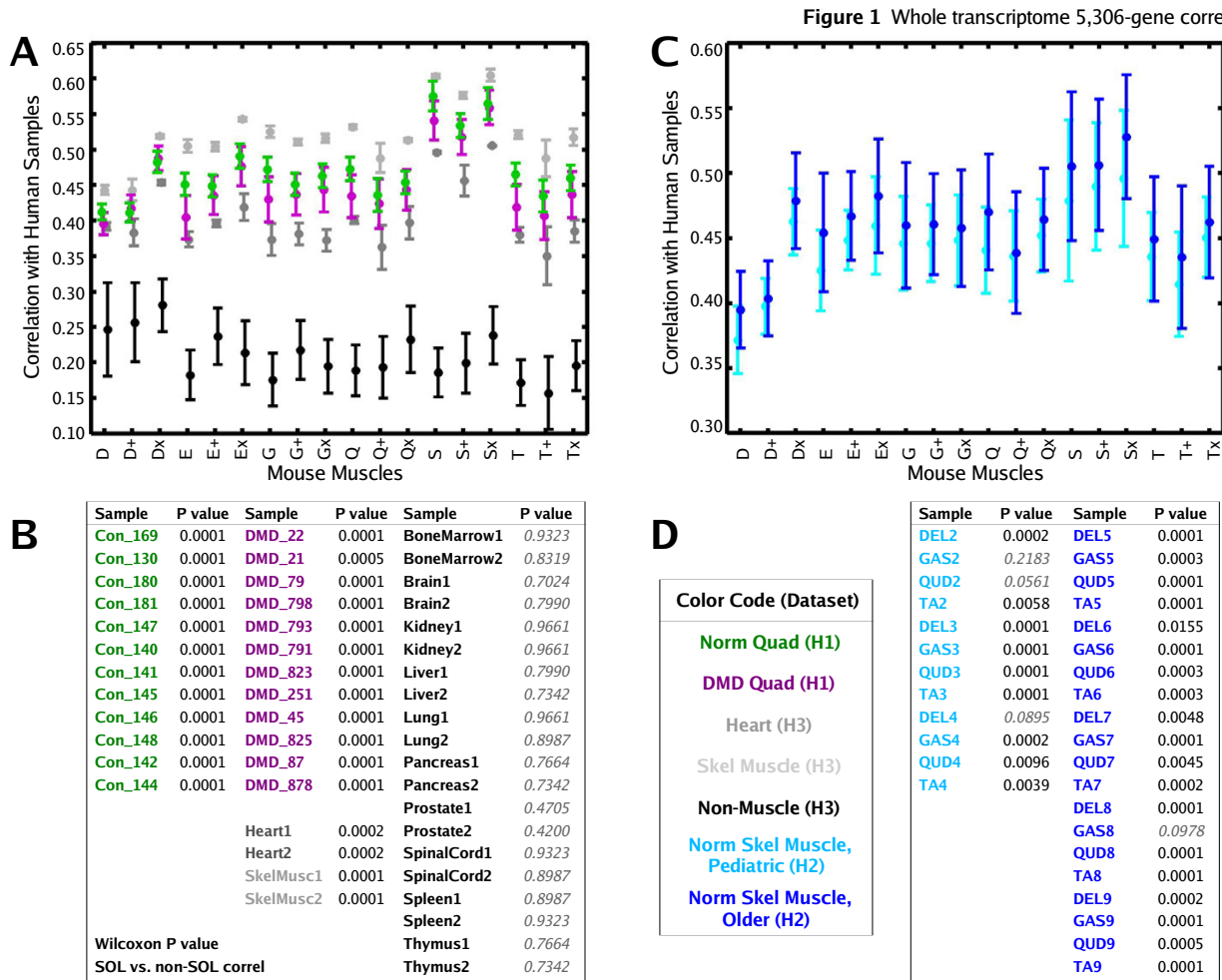


Figure 1

Correlations of human transcriptome profiles relative to homologous mouse muscle profiles. All samples are characterized by their 5,306-gene whole transcriptome profile. (A) Human linear correlation group averages \pm one standard deviation error bars relative to each of the six mouse muscle groups \times three genetic strains from dataset M. The five human groups are color coded: control (green) and DMD (magenta) quadriceps from dataset H1, pooled skeletal muscle (light gray), heart (dark gray), and tissue not primarily composed of muscle (black) from dataset H3. (B) For each human sample, the likelihood (Wilcoxon rank-sum) that there is no difference in the human sample's correlations to soleus versus to non-soleus mouse muscle profiles is shown. (C, D) Similar to (A, B) except that relative to dataset M, the human samples compared are four different skeletal muscle groups from each of eight normal subject autopsies of dataset H2, forming two color-coded groups: pediatric (light blue) and older (dark blue) samples. Among mouse muscles, human muscle sample groups were consistently most correlated with the mouse soleus.

significant correlation to any mouse muscle group (Figure 2).

Differential correlations of normal versus DMD human quadriceps to normal mouse skeletal muscle transcriptomes

We investigated whether the myopathic state (specifically DMD) of a human muscle (specifically quadriceps in

dataset H1) was reflected in its correlation against wildtype mouse skeletal muscle groups (dataset M).

Mouse-human correlation analyses were performed between datasets M and H1, relative to the above sample profile characterizations: whole transcriptome and sub-transcriptomes ST1-4. In all except the ST4 characterization, control human quadriceps to wildtype mouse mus-

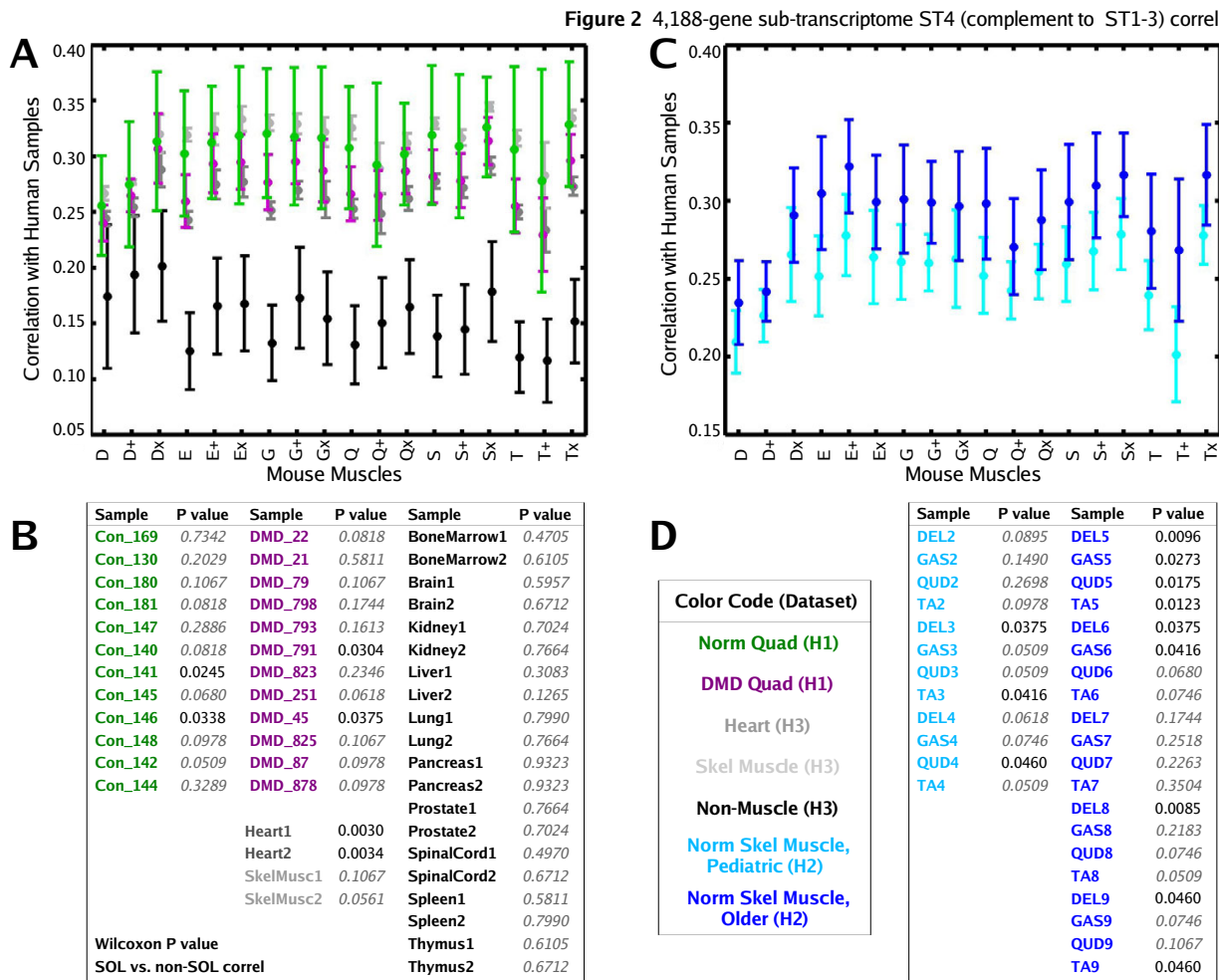


Figure 2

Analogous to Figure 1, correlations of human gene expression profiles relative to homologous mouse muscle profiles, except that samples are characterized as a 4,188-gene profile (ST4) of genes complementary to sub-transcriptomes ST1-3. This is a negative control comparison. Note that human muscle group profiles were not more correlated to any one particular mouse muscle.

cle correlations were significantly higher than DMD human quadriceps to wildtype mouse muscle correlations ($p < 0.02$, Wilcoxon) (Figure 3).

Discussion

All human skeletal muscle samples tested were found to be significantly more correlated to mouse soleus than to the other five non-soleus mouse skeletal muscle groups, independent of the human samples' anatomic location/muscle type, disease state, age and sampling method (autopsy versus biopsy). This observation was both specific to muscle tissue, and consistent even when distinct sub-transcriptomes were used to characterize the mouse-human sample profiles. These *in silico* transcriptome similarities might reflect common/analogous molecular hierarchies modulating these system biologies. From a

practical perspective, the approach taken here can be extended to assessing general animal models of human systems *in silico*.

Mouse soleus bears a greater molecular resemblance to several human skeletal muscle types under different conditions (autopsy and biopsy sampling, disease and control) than the other five mouse muscles examined. This suggests that mouse soleus may be one of the better representations of human muscle disease among mouse muscles, to study both disease processes and potential therapies. Besides having a molecular profile that is distinct from the other five mouse skeletal muscle groups [17], soleus also displays histological differences. Soleus has 58% type I fibers, whereas the other muscles have lower type I percentages: extensor digitorum longus and

Figure 3 Correlations (12 Human Control, mouse muscle) versus Correlations (12 Human DMD, mouse muscle) for each mouse muscle, Wilcoxon ranksum p-value

Mouse Muscle	5,306-gene whole transcriptome		733-gene set ST1 of cell plasma membrane genes		234-gene set ST3 of dominant contributors to sample variance in human dataset H1		271-gene set ST2 of dominant contributors to sample variance in mouse dataset M		4,188-gene set ST4 of genes complement to ST1-3	
	P value Con-DMD	Avg Con-DMD	P value Con-DMD	Avg Con-DMD	P value Con-DMD	Avg Con-DMD	P value Con-DMD	Avg Con-DMD	P value Con-DMD	Avg Con-DMD
D	0.0007	0.0163	0.0008	0.0463	0.0000	0.0632	0.0000	0.0368	0.3123	0.0157
D+	0.0949	-0.0062	0.0023	0.0426	0.9015	0.0010	0.1609	-0.0085	0.7887	0.0102
D×	0.2160	-0.0055	0.0949	0.0241	0.6207	0.0047	0.2745	-0.0066	0.7415	0.0064
E	0.0000	0.0462	0.0000	0.0818	0.0000	0.1112	0.0000	0.0862	0.0002	0.0429
E+	0.1171	0.0135	0.0010	0.0528	0.0017	0.0478	0.0696	0.0308	0.2836	0.0189
E×	0.1171	0.0145	0.0006	0.0574	0.0094	0.0459	0.0392	0.0348	0.1271	0.0239
G	0.0000	0.0421	0.0000	0.0879	0.0000	0.1002	0.0000	0.0776	0.0001	0.0439
G+	0.2317	0.0141	0.0013	0.0502	0.0083	0.0475	0.0833	0.0337	0.1768	0.0227
G×	0.0665	0.0193	0.0001	0.0617	0.0020	0.0543	0.0187	0.0396	0.0728	0.0295
Q	0.0000	0.0385	0.0000	0.0897	0.0000	0.0944	0.0000	0.0755	0.0003	0.0416
Q+	0.3123	0.0119	0.0016	0.0571	0.0392	0.0400	0.2568	0.0255	0.1939	0.0273
Q×	0.3753	0.0105	0.0017	0.0479	0.0392	0.0419	0.1432	0.0278	0.2930	0.0153
S	0.0000	0.0344	0.0000	0.0937	0.0000	0.0822	0.0000	0.0643	0.0057	0.0374
S+	0.0320	0.0165	0.0000	0.0889	0.0034	0.0488	0.0028	0.0364	0.0606	0.0308
S×	0.5094	0.0058	0.0017	0.0555	0.0478	0.0277	0.0635	0.0180	0.5777	0.0126
T	0.0000	0.0464	0.0000	0.0894	0.0000	0.1054	0.0000	0.0840	0.0001	0.0512
T+	0.0039	0.0280	0.0009	0.0628	0.0004	0.0654	0.0042	0.0485	0.0112	0.0487
T×	0.0233	0.0234	0.0001	0.0707	0.0002	0.0618	0.0083	0.0469	0.0094	0.0325

+ mdx × mdx5cv

Figure 3

Difference in correlations of control human quadriceps to each mouse skeletal muscle group, and correlations of DMD human quadriceps to each mouse skeletal muscle group – relative to five different sub-transcriptome sample profile characterizations. In other words, $\Delta(\text{Corr}(\text{control human quadriceps, mouse muscle}), \text{Corr}(\text{DMD human quadriceps, mouse muscle}))$. A Wilcoxon rank-sum test was used to assess the likelihood of non-difference in these correlations. In all but the ST4 (negative control) sample profile characterization, mouse wildtype muscles were more correlated to control (Con) than to DMD human quadriceps.

tibialis anterior 0%, gastrocnemius 1–8%, quadriceps 0–45%, diaphragm 10–12% [30-33]. Like mouse soleus, the human autopsy samples examined had a balance of fiber types I and II [20], raising the question of whether the molecular resemblances derive primarily from similarities in fiber type proportions. If this were the case, the genes that most distinguish mouse soleus from other mouse muscles, and most represent its similarities to human muscles, should include a significant number of fiber type specific genes. Consistent with this interpretation, we find that among the ten genes with the highest positive loading in PC2 (the axis of greatest variation between mouse soleus versus other muscle groups) for dataset M, six are differentially expressed between fiber types (rank in

parentheses): Myl3 (#2) [34], Tnni1 (#3) [35], Myl2 (#4) [36], Tnnc1 (#5) [37], Mb (#7) [38], and Tnnt (#9) [39]. Two of the others, Atp2a2 (#6) [40] and Idh2 (#10), [41] are involved in energy metabolism and may be expressed at different levels in different fiber types. Curiously, Fhl1 (#8) is differentially expressed between rat soleus and gastrocnemius, but this does not appear to be due to fiber type composition [42]. Myh7 (#1) is primarily expressed in cardiac muscle [43] and has been associated with cardiomyopathy, [44,45] but also appears to be expressed in skeletal muscle [46]. Two other parameters that influence muscle function – muscle length and fiber length – do not differ significantly between the five non-diaphragm skeletal muscle groups [30].

Table 1:

Dataset compositions				
Dataset	Tissue			
M (mouse)	Muscle: EDL, diaphragm, gastrocnemius, quadriceps, soleus, TA			
H1 (human)	Muscle: quadriceps			
H2 (human)	Muscle: deltoid, gastrocnemius, quadriceps, tibialis anterior			
H3 (human)	Bone marrow, brain, heart, liver, lung, kidney, skeletal muscle, pancreas, prostate, spinal cord, spleen, thymus			

Dataset compositions				
Dataset	Conditions	Source	# of Samples	Affymetrix GeneChip
M (mouse)	Control, <i>mdx</i> , <i>mdx^{5cv}</i>	Immediate autopsy	36	U74Av2
H1 (human)	Control, DMD	Surgical biopsy	24	U95Av2
H2 (human)	Control	Autopsy	32	U133A
H3 (human)	Control	Autopsy?	24	U95Av2

Functional studies have demonstrated significant differences in the mechanical properties of different mouse muscles, but the data are somewhat conflicting. One study found that the soleus muscle of dystrophin-utrophin double knockout mice had a greater reduction in twitch force than in extensor digitorum longus and diaphragm [47]. In contrast, isometric contractions followed by stretch contractions in extensor digitorum longus and soleus muscles dissected from control and *mdx* mice resulted in irreversible damage only in *mdx* extensor digitorum longus [48]. The interpretation of this data is unclear, as the findings were identified in different mouse strains. The *mdx* mouse would genetically be expected to replicate the human disease most accurately, but the divergent phenotype compared to human DMD raises doubts about this.

The use of animals to model human disease is complicated by many factors. This is especially true in myopathies where there are many different muscle types, some of which may be more representative of human disease than others. Based on phenotypic severity, *mdx* diaphragm appears to bear the closest relationship to human disease. However, the diaphragm is not an extremity muscle. Among extremity muscles in the *mdx* mouse, soleus might merit special attention in studies of the pathophysiology of the muscular dystrophies.

Conclusion

1. Mouse soleus bears a closer molecular resemblance than other mouse skeletal muscles to several different human skeletal muscles.
2. This resemblance is consistent for both control and disease human tissue, and is specific for human muscle tissue compared to non-muscle tissue.

3. These results may be due in part to differences in fiber type composition.

List of abbreviations

- DMD, Duchenne muscular dystrophy
- H1, H2, H3, human datasets
- M, mouse dataset
- NCBI, National Center for Biotechnology Information
- PC, principal component
- PCA, principal component analysis
- ST1, ST2, ST3, ST4, sub-transcriptomes

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

ATK helped conceive the study, performed the data analysis, and drafted a large portion of the manuscript. PBK participated in data interpretation, suggested the biological and therapeutic applications, and helped draft and edit the manuscript. ISK participated in data interpretation. LMK helped conceive the study, participated in data interpretation, and coordinated the overall project. All authors read and approved the final manuscript.

Additional material

Additional File 1

Adobe pdf file. Principal component analysis. A mathematical review of principal component analysis.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2474-7-23-S1.pdf>]

Additional File 2

Adobe pdf file. Principal component analysis (PCA) of mouse dataset M. (A, B) PCA of mouse muscle samples characterized as 5,306-gene whole transcriptome profiles. Sample projections into principal components PC1-2 (A) and PC1-3 (B) planes are shown. The dominant variation (along PC1) is between three skeletal muscle clusters: diaphragm, soleus, and the four remaining muscles. (C, D) Analogous to (A, B), PCA of mouse samples characterized as 271-gene profiles of dominant contributors to whole transcriptome sample variation. Dominant variance contributors are defined to be genes which have absolute loading coefficient exceeding 0.03 in PC1-3 in the PCA of the whole (5,306-gene) transcriptome case (A, B).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2474-7-23-S2.pdf>]

Additional File 3

Adobe pdf file. PCA of human dataset H1. (A, B) PCA of human muscle samples characterized as 5,306-gene whole transcriptome profiles. Sample projections into principal components PC1-2 (A) and PC1-3 (B) planes are shown. The dominant variation (along PC1) is between control and DMD samples. (C, D) Analogous to (A, B), PCA of human samples characterized as 234-gene profiles of dominant contributors to whole transcriptome sample variation. Dominant variance contributors are defined to be genes which have absolute loading coefficient exceeding 0.03 in PC1-3 in the PCA of the whole (5,306-gene) transcriptome case (A, B).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2474-7-23-S3.pdf>]

Additional File 4

Adobe pdf file. Correlations of human transcriptome profiles relative to homologous mouse muscle profiles. All samples are characterized by their 733-gene profile of cell plasma membrane genes (ST1). (A) Human linear correlation group averages \pm one standard deviation error bars relative to each of the six mouse muscle groups \times three genetic strains from dataset M. The five human groups are color coded: control (green) and DMD (magenta) quadriceps from dataset H1, pooled skeletal muscle (light gray), heart (dark gray), and tissue not primarily composed of muscle (black) from dataset H3. (B) For each human sample, the likelihood (Wilcoxon ranksum) that there is no difference in the human sample's correlations to soleus versus to non-soleus mouse muscle profiles is shown. (C, D) Similar to (A, B) except that relative to dataset M, the human samples compared are four different skeletal muscle groups from each of eight normal subject autopsies of dataset H2, forming two color-coded groups: pediatric (light blue) and older (dark blue) samples. Among mouse muscles, human muscle sample groups were consistently most correlated with the mouse soleus.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2474-7-23-S4.pdf>]

Additional File 5

Adobe pdf file. Correlations of human transcriptome profiles relative to homologous mouse muscle profiles. Here, all samples are characterized by their 271-gene profile of dominant contributors to sample variance in the mouse dataset M (i.e., genes with absolute loadings in PC1-3 exceeding 0.03 (ST2)).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2474-7-23-S5.pdf>]

Additional File 6

Adobe pdf file. Correlations of human transcriptome profiles relative to homologous mouse muscle profiles. Here, all samples are characterized by their 234-gene profile of dominant contributors to sample variance in the human dataset H1 (i.e., genes with absolute loadings in PC1-3 exceeding 0.03 (ST2)).

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2474-7-23-S6.pdf>]

Acknowledgements

The authors express their appreciation for support from the following sources: NIH P01 NS40828-01A1 (ATK, ISK, LMK), NIH R01 NS047527 (ATK), R21 NS052498 (ATK), NIH U01 HL066582-01 (ISK), NINDS K08 NS048180 (PBK), the Bernard F. and Alva B. Gimbel Foundation (LMK), and the Joshua Frase Foundation (LMK). LMK is an Investigator of the Howard Hughes Medical Institute.

References

1. Bulfield G, Siller WG, Wight PA, Moore KJ: **X chromosome-linked muscular dystrophy (mdx) in the mouse.** *Proc Natl Acad Sci U S A* 1984, **81**:1189-1192.
2. Ryder-Cook AS, Sicinski P, Thomas K, Davies KE, Worton RG, Barnard EA, Darlison MG, Barnard PJ: **Localization of the mdx mutation within the mouse dystrophin gene.** *Embo J* 1988, **7**:3017-3021.
3. Sicinski P, Geng Y, Ryder-Cook AS, Barnard EA, Darlison MG, Barnard PJ: **The molecular basis of muscular dystrophy in the mdx mouse: a point mutation.** *Science* 1989, **244**:1578-1580.
4. Chapman VM, Miller DR, Armstrong D, Caskey CT: **Recovery of induced mutations for X chromosome-linked muscular dystrophy in mice.** *Proc Natl Acad Sci U S A* 1989, **86**:1292-1296.
5. Danko I, Chapman V, Wolff JA: **The frequency of revertants in mdx mouse genetic models for Duchenne muscular dystrophy.** *Pediatr Res* 1992, **32**:128-131.
6. Durbeej M, Campbell KP: **Muscular dystrophies involving the dystrophin-glycoprotein complex: an overview of current mouse models.** *Curr Opin Genet Dev* 2002, **12**:349-361.
7. Chen YW, Zhao P, Borup R, Hoffman EP: **Expression profiling in the muscular dystrophies: identification of novel aspects of molecular pathophysiology.** *J Cell Biol* 2000, **151**:1321-1336.
8. Haslett JN, Sanoudou D, Kho AT, Bennett RR, Greenberg SA, Kohane IS, Beggs AH, Kunkel LM: **Gene expression comparison of biopsies from Duchenne muscular dystrophy (DMD) and normal skeletal muscle.** *Proc Natl Acad Sci U S A* 2002, **99**:15000-15005.
9. Haslett JN, Sanoudou D, Kho AT, Han M, Bennett RR, Kohane IS, Beggs AH, Kunkel LM: **Gene expression profiling of Duchenne muscular dystrophy skeletal muscle.** *Neurogenetics* 2003, **4**:163-171.
10. Noguchi S, Tsukahara T, Fujita M, Kurokawa R, Tachikawa M, Toda T, Tsujimoto A, Arahata K, Nishino I: **cDNA microarray analysis of individual Duchenne muscular dystrophy patients.** *Hum Mol Genet* 2003, **12**:595-600.
11. Tkatchenko AV, Le Cam G, Leger JJ, Dechesne CA: **Large-scale analysis of differential gene expression in the hindlimb mus-**

- cles and diaphragm of mdx mouse. *Biochim Biophys Acta* 2000, **1500**:17-30.
12. Porter JD, Khanna S, Kaminski HJ, Rao JS, Merriam AP, Richmonds CR, Leahy P, Li J, Guo W, Andrade FH: **A chronic inflammatory response dominates the skeletal muscle molecular signature in dystrophin-deficient mdx mice.** *Hum Mol Genet* 2002, **11**:263-272.
 13. Rouger K, Le Cunff M, Steenman M, Potier MC, Gibelin N, Dechesne CA, Leger JJ: **Global/temporal gene expression in diaphragm and hindlimb muscles of dystrophin-deficient (mdx) mice.** *Am J Physiol Cell Physiol* 2002, **283**:C773-84.
 14. Tseng BS, Zhao P, Pattison JS, Gordon SE, Granchelli JA, Madsen RW, Folk LC, Hoffman EP, Booth FW: **Regenerated mdx mouse skeletal muscle shows differential mRNA expression.** *J Appl Physiol* 2002, **93**:537-545.
 15. Porter JD, Merriam AP, Khanna S, Andrade FH, Richmonds CR, Leahy P, Cheng G, Karathanasis P, Zhou X, Kusner LL, Adams ME, Willem M, Mayer U, Kaminski HJ: **Constitutive properties, not molecular adaptations, mediate extraocular muscle sparing in dystrophic mdx mice.** *Faseb J* 2003, **17**:893-895.
 16. Porter JD, Merriam AP, Leahy P, Gong B, Feuerman J, Cheng G, Khanna S: **Temporal gene expression profiling of dystrophin-deficient (mdx) mouse diaphragm identifies conserved and muscle group-specific mechanisms in the pathogenesis of muscular dystrophy.** *Hum Mol Genet* 2004, **13**:257-269.
 17. Haslett JN, Kang PB, Han M, Kho AT, Sanoudou D, Volinski JM, Beggs AH, Kohane IS, Kunkel LM: **The influence of muscle type and dystrophin deficiency on murine expression profiles.** *Mamm Genome* 2005, **16**:739-748.
 18. **National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO), dataset 641.** http://www.ncbi.nlm.nih.gov/geo/gds/gds_browse.cgi?gds=641.
 19. **National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO), dataset 563.** http://www.ncbi.nlm.nih.gov/geo/gds/gds_browse.cgi?gds=563.
 20. Kang PB, Kho AT, Sanoudou D, Haslett JN, Dow CP, Han M, Blasko JM, Lidov HG, Beggs AH, Kunkel LM: **Variations in gene expression among different types of human skeletal muscle.** *Muscle Nerve* 2005, **32**:483-491.
 21. Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, Bar-Even A, Horn-Saban S, Safra M, Domany E, Lancet D, Shmueli O: **Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification.** *Bioinformatics* 2005, **21**:650-659.
 22. **National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO), dataset 422.** http://www.ncbi.nlm.nih.gov/geo/gds/gds_browse.cgi?gds=422.
 23. **National Center for Biotechnology Information (NCBI) HomoloGene.** <http://www.ncbi.nlm.nih.gov/HomoloGene>.
 24. Press WH: **Numerical recipes: the art of scientific computing.** New York, Cambridge University Press; 1989.
 25. Misra J, Schmitt W, Hwang D, Hsiao LL, Gullans S, Stephanopoulos G: **Interactive exploration of microarray gene expression patterns in a reduced dimensional space.** *Genome Res* 2002, **12**:1112-1120.
 26. Kho AT, Zhao Q, Cai Z, Butte AJ, Kim JY, Pomeroy SL, Rowitch DH, Kohane IS: **Conserved mechanisms across development and tumorigenesis revealed by a mouse development perspective of human cancers.** *Genes Dev* 2004, **18**:629-640.
 27. Johnson RA, Wichern DW: **Applied multivariate statistical analysis.** 5th edition. Upper Saddle River, NJ, Prentice Hall; 2002.
 28. Hollander M, Wolfe DA: **Nonparametric Statistical Methods.** New York, John Wiley and Sons; 1973.
 29. **Gene Ontology Consortium.** <http://www.geneontology.org>.
 30. Burkholder TJ, Fingado B, Baron S, Lieber RL: **Relationship between muscle fiber types and sizes and muscle architectural properties in the mouse hindlimb.** *J Morphol* 1994, **221**:177-190.
 31. Davies AS, Gunn HM: **Histochemical fibre types in the mammalian diaphragm.** *J Anat* 1972, **112**:41-60.
 32. Petrof BJ, Stedman HH, Shrager JB, Eby J, Sweeney HL, Kelly AM: **Adaptations in myosin heavy chain expression and contractile function in dystrophic mouse diaphragm.** *Am J Physiol* 1993, **265**:C834-41.
 33. Talmadge RJ, Otis JS, Rittler MR, Garcia ND, Spencer SR, Lees SJ, Naya FJ: **Calcineurin activation influences muscle phenotype in a muscle-specific fashion.** *BMC Cell Biol* 2004, **5**:28.
 34. Fodor WL, Darras B, Seharaseyon J, Falkenthal S, Francke U, Vanin EF: **Human ventricular/slow twitch myosin alkali light chain gene characterization, sequence, and chromosomal location.** *J Biol Chem* 1989, **264**:2143-2149.
 35. Wade R, Eddy R, Shows TB, Kedes L: **cDNA sequence, tissue-specific expression, and chromosomal mapping of the human slow-twitch skeletal muscle isoform of troponin I.** *Genomics* 1990, **7**:346-357.
 36. Poetter K, Jiang H, Hassanzadeh S, Master SR, Chang A, Dalakas MC, Rayment I, Sellers JR, Fananapazir L, Epstein ND: **Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle.** *Nat Genet* 1996, **13**:63-69.
 37. Schreier T, Kedes L, Gahlmann R: **Cloning, structural analysis, and expression of the human slow twitch skeletal muscle/cardiac troponin C gene.** *J Biol Chem* 1990, **265**:21247-21253.
 38. Takemasa T, Sugimoto K, Miyazaki M, Machida M, Ikeda S, Hitomi Y, Kizaki T, Ohno H, Yamashita K, Haga S: **Simple method for the identification of oxidative fibers in skeletal muscle.** *Eur J Appl Physiol* 2004, **91**:357-359.
 39. Samson F, Mesnard L, Mihovilovic M, Potter TG, Mercadier JJ, Roses AD, Gilbert JR: **A new human slow skeletal troponin T (TnTs) mRNA isoform derived from alternative splicing of a single gene.** *Biochem Biophys Res Commun* 1994, **199**:841-847.
 40. Kjellgren D, Ryan M, Ohlndieck K, Thornell LE, Pedrosa-Domellof F: **Sarco(endoplasmic reticulum Ca²⁺ ATPases (SERCA1 and -2) in human extraocular muscles.** *Invest Ophthalmol Vis Sci* 2003, **44**:5057-5062.
 41. Luo H, Shan X, Wu J: **Expression of human mitochondrial NADP-dependent isocitrate dehydrogenase during lymphocyte activation.** *J Cell Biochem* 1996, **60**:495-507.
 42. Loughna PT, Mason P, Bayol S, Brownson C: **The LIM-domain protein FHL1 (SLIM 1) exhibits functional regulation in skeletal muscle.** *Mol Cell Biol Res Commun* 2000, **3**:136-140.
 43. Kurabayashi M, Tsuchimochi H, Komuro I, Takaku F, Yazaki Y: **Molecular cloning and characterization of human cardiac alpha- and beta-form myosin heavy chain complementary DNA clones. Regulation of expression during development and pressure overload in human atrium.** *J Clin Invest* 1988, **82**:524-531.
 44. Geisterfer-Lowrance AA, Kass S, Tanigawa G, Vosberg HP, McKenna W, Seidman CE, Seidman JG: **A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation.** *Cell* 1990, **62**:999-1006.
 45. Tanigawa G, Jarcho JA, Kass S, Solomon SD, Vosberg HP, Seidman JG, Seidman CE: **A molecular basis for familial hypertrophic cardiomyopathy: an alpha/beta cardiac myosin heavy chain hybrid gene.** *Cell* 1990, **62**:991-998.
 46. Tajsharghi H, Thornell LE, Lindberg C, Lindvall B, Henriksson KG, Oldfors A: **Myosin storage myopathy associated with a heterozygous missense mutation in MYH7.** *Ann Neurol* 2003, **54**:494-500.
 47. Deconinck N, Rafael JA, Beckers-Bleux G, Kahn D, Deconinck AE, Davies KE, Gillis JM: **Consequences of the combined deficiency in dystrophin and utrophin on the mechanical properties and myosin composition of some limb and respiratory muscles of the mouse.** *Neuromuscul Disord* 1998, **8**:362-370.
 48. Moens P, Baatsen PH, Marechal G: **Increased susceptibility of EDL muscles from mdx mice to damage induced by contractions with stretch.** *J Muscle Res Cell Motil* 1993, **14**:446-451.

Pre-publication history

The pre-publication history for this paper can be accessed here:

<http://www.biomedcentral.com/1471-2474/7/23/prepub>