The FASEB Journal • Research Communication

The human skeletal muscle transcriptome: sex differences, alternative splicing, and tissue homogeneity assessed with RNA sequencing

Malene E. Lindholm,*[,1] Mikael Huss,[†] Beata W. Solnestam,[‡] Sanela Kjellqvist,[†] Joakim Lundeberg,[‡] and Carl J. Sundberg*

*Department of Physiology and Pharmacology, Karolinska Institutet, Solna, Sweden; [†]Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden; and [‡]Science for Life Laboratory, School of Biotechnology, Royal Institute of Technology, Solna, Sweden

ABSTRACT **Human skeletal muscle health is important for quality of life and several chronic diseases, including type II diabetes, heart disease, and cancer. Skeletal muscle is a tissue widely used to study mechanisms behind different diseases and adaptive effects of controlled interventions. For such mechanistic studies, knowledge about the gene expression profiles in different states is essential. Since the baseline transcriptome has not been analyzed systematically, the purpose of this study was to provide a deep reference profile of female and male skeletal muscle. RNA sequencing data were analyzed from a large set of 45 resting human muscle biopsies. We provide extensive information on the skeletal muscle transcriptome, including 5 previously unannotated protein-coding transcripts. Global transcriptional tissue homogeneity was strikingly high, within both a specific muscle and the contralateral leg. We identified >23,000 known isoforms and found >5000 isoforms that differ between the sexes. The female and male transcriptome was enriched for genes associated with oxidative metabolism and protein catabolic processes, respectively. The data demonstrate remarkably high tissue homogeneity and provide a deep and extensive baseline reference for the human skeletal muscle transcriptome, with regard to alternative splicing, novel transcripts, and sex differences in functional ontology.—Lindholm, M. E., Huss, M., Solnestam, B. W., Kjellqvist, S., Lundeberg, J., Sundberg, C. J. The human skeletal muscle transcriptome: sex differences, alternative splicing, and tissue homogeneity assessed with RNA sequencing. *FASEB J.* 28, 4571–4581 (2014). www.fasebj.org**

Key Words: gene expression profiling · biopsy · isoform · splice variant · novel transcript

Gene expression analysis has become a central element in the study of mechanisms involved in tissue and cell regulation. In humans, skeletal muscle is a relatively unique tissue for this purpose, as it is accessible for sampling in meaningful quantities; easily subjected to physiologically relevant, controlled interventions; highly plastic; and functionally testable in a reliable and relevant manner. It constitutes the largest tissue in the human body and is metabolically very active. Over the past 15 yr, skeletal muscle has been widely used to study the transcriptional effects of many kinds of exercise training (1–6) and unloading (7, 8), as well as diseases, including type II diabetes (9–11), muscular dystrophies (12–14), atrophy (15–17), and others (18, 19). The global, baseline transcriptome of human skeletal muscle, however, has not been analyzed systematically.

Alternative splicing, which greatly extends the transcriptomic, and thus the proteomic, complexity (20), has become an intense area of research (21). Almost 95% of all human genes have been shown to express different isoforms, and specific alternative splicing events are probable key contributors to tissue-specific phenotype, also for skeletal muscle (22). One factor that may influence isoform expression is the sex of the subject, which in relation to alternative splicing has only been investigated in, for example, human liver tissue (23), mouse liver tissue (24), and *Drosophila melanogaster* (25).

General sex-specific transcriptomic differences have been addressed previously in skeletal muscle using microarray technology. For example, Maher *et al.* (26) showed that 66 genes differ between males and females, while Roth *et al.* (6) and Welle *et al.* (27) found ~200 and 3000 differentially expressed transcripts, respec-

Abbreviations: CS, citrate synthase; FPKM, fragments per kilobase per million mapped; GEO, Gene Expression Omnibus; GO, gene ontology; hg, human genome; lncRNA, long non-coding RNA; OPLS, orthogonal projections to latent structures; OPLS-DA, orthogonal projections to latent structures discriminant analysis; PCA, principal component analysis; PGC1α, peroxisome proliferator-activated receptor γ coactivator 1 α; PPAR, peroxisome proliferator-activated receptor; RNA-seq, RNA sequencing

[1] Correspondence: Department of Physiology and Pharmacology, Karolinska Institutet, von Eulers Väg 8, 171 77 Stockholm, Sweden. malene.lindholm@ki.se

doi: 10.1096/fj.14-255000

This article includes supplemental data. Please visit *http://www.fasebj.org* to obtain this information.

tively. Baseline sex differences have also been demonstrated in biceps muscle by Liu *et al.* (28). However, the current investigation is the first to study the global transcriptome differences between males and females in skeletal muscle by using massive parallel sequencing.

In many human studies, several samples are obtained from different sites within a specific muscle belly and occasionally also in the corresponding contralateral muscle. The underlying assumption is that the sampling sites have virtually similar histological, physiological, biochemical, and transcriptional properties. Any differences in gene expression over a time period are presumed to stem from the intervention itself and not from heterogeneity in the tissue (1, 29–33). Several studies have demonstrated that this assumption may be problematic. A single skeletal muscle has been shown to be somewhat heterogeneous at different depths, especially with regard to fiber type distribution and fiber area (34–36) and also that histo- and biochemical characteristics may differ throughout the length of a specific muscle, as indicated from previous rat studies (37, 38). This difference may also hold true for the skeletal muscle transcriptome, which hitherto has not been investigated at a global level in humans. Furthermore, differences between the 2 legs of an individual, as well as the magnitude of interindividual variability are fundamental questions that remain to be addressed in a systematic manner.

The purpose of this extensive human experimental study was to thoroughly investigate the global baseline transcriptome of human skeletal muscle with respect to spatial tissue homogeneity, sex differences, alternative splicing, and novel transcript discovery. Deep RNA sequencing (RNA-seq) was used, comparing biopsies from 18 volunteers, within the same leg of an individual, between the 2 legs, and between the males and females. Our results show the quantitative baseline variability in RNA expression for protein-coding genes and long, noncoding (lnc)RNAs. This baseline can be used as a reference when investigating changes in gene expression following an intervention or when comparing to a disease state. We found that skeletal muscle tissue is remarkably homogeneous, identified many novel transcripts, and showed that ontologies for aerobic metabolism were enriched in the females, whereas protein turnover was enriched in the males.

MATERIALS AND METHODS

Experimental protocol

Before the study, the experimental protocol was explained to all subjects, and informed consent was obtained. The study was approved by the Ethics Committee of Karolinska Institutet and conformed to the Declaration of Helsinki. Eighteen sedentary subjects, 9 males and 9 females, were included (baseline characteristics are presented in **Table 1**). A maximal incremental oxygen uptake test was performed to assess the baseline physical fitness. From 6 of the subjects, 3 males and 3 females (mean±SD age 26.7±3.9 yr, BMI 22.3±2.7 cm/kg^2, and VO_2 42.7±3.3 ml·min^{-1}·kg^{-1}), 2 skeletal muscle biopsies from the vastus lateralis muscle were taken from each leg, all from different incisions (Supplemental Fig. S1*A*). The within-leg samples were taken ~3 cm apart, and all samples were taken within 15 min. From the other 12 subjects, 6 males and 6 females, 1 biopsy was obtained from each leg. All 48 biopsies were obtained under local anesthesia by the percutaneous needle technique (39) and immediately frozen in liquid nitrogen.

TABLE 1. *Baseline characteristics*

Parameter	All subjects	Males	Females
n	18	9	9
Age (yr)	27 ± 3	29 ± 3	25 ± 3
Height (cm)	177 ± 9	183 ± 7	171 ± 5
BMI (kg·m^{-2})	24 ± 4	25 ± 4	23 ± 3
Peak VO_2 (ml·kg^{-1}·min^{-1})	41 ± 4	41 ± 5	41 ± 4

Data are presented as means ± SD.

RNA isolation, library construction, and sequencing

Total RNA was extracted by the TRIzol method (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's specifications. The concentration and quality of the RNA was determined with the RNA 6000 Nanochip on the 2100 Bioanalyzer automated electrophoresis system (Agilent Technologies Inc., Santa Clara, CA, USA). Two micrograms of total RNA was used for the library preparation of each sample, which was subsequently bar coded and prepared with the automated platform (Illumina, San Diego, CA, USA; ref. 40). The libraries were clustered on a cBot cluster-generation system with an Illumina HiSeq paired-end, cluster-generation kit and subsequently sequenced as paired-end, 2 × 100 bp on an Illumina HiSeq 2000. All lanes were spiked with a 1% phiX control library. The sequencing run generated 833 million paired-end reads (after excluding 3 samples due to technical reasons). Raw data are available at Gene Expression Omnibus (GEO) under GEO accession numbers GSE58387 and GSE58608. A schematic summary of the methodology is illustrated in Supplemental Fig. S1*B*.

Sequence read alignment and analysis

The paired-end reads were trimmed with respect to quality and remaining adapter sequences using Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and subsequently aligned to the human genome reference hg19 with TopHat version 2.0.4 (41), using standard parameters except for the parameters solexa1.3-quals, p (which was set to 8) and transcriptomeindex (used with an index built from ENSEMBL v71 transcript annotation). The aligned reads were assembled into transcripts using Cufflinks version 2.0.2 (42) and exon fragments per kilobase per million mapped (FPKM) values were calculated. The aligned reads were also used to count the number of reads per gene with HTSeq, version 0.5.1. The fraction of the total sequences that were accounted for by the top transcribed genes, in FPKM values and raw counts, as well as the distribution of FPKM values for all genes, are shown in Supplemental Fig. S1*C–E*. The R/Bioconductor package DESeq2 (43) was used to call differential gene expression based on the gene counts generated by HTSeq. Principal component analysis was based on normalized gene-level counts [adjustment of raw counts by TMM scaling implemented in the edgeR package followed by transformation into $\log_2$ counts per million (log-cpm)].

Gene ontology (GO) analysis

For functional annotation analysis of the whole gene set (in total, 10,268 protein-coding genes) the Panther database (http://www.pantherdb.org) was used, whereas the Data for Annotation, Visualization and Integrated Discovery (DAVID) GO annotation tool (http://david.abcc.ncifcrf.gov) was used for genes expressed differentially between the males and females. The presented *P* values represent the probability of finding that subset of genes from a certain GO term in a random set of genes drawn from the complete gene list.

Comparison to skeletal muscle microarray data

Correlation between sequencing- and array-based expression values was performed to compare our transcriptome data to those of a completely different set of subjects. Raw and processed microarray data from Timmons *et al.* (44) were downloaded from GEO (accession GSE18583). For the transcript-level comparison, FPKM estimates for isoforms were compared to the processed microarray data (GEO file: baseline_array_ENST_from_GSE18583_series_matrix.txt). For the gene-level comparison, FPKM estimates for genes were correlated to expression values obtained by applying robust multichip averaging (RMA; implemented in the Affy package in R/BioConductor) to the raw CEL files (file GSE18583_RAW.tar). For transcripts and genes, the Spearman correlation coefficient was calculated for the set of ENST or ENSG identifiers, respectively, for which there was an expression estimate from both the array and RNA-seq.

Novel transcript discovery and characterization

Reference-based assembly was performed on the alignment (BAM) file for each sample by Cufflinks v. 2.1.1 (42) without any transcript reference. The 45 assemblies obtained were merged with the Cuffmerge tool (Cufflinks) with transcript annotation (ENSEMBL v. 71) supplied at this step. Reconstructed transcripts that were classified as "unknown intergenic" (n=3152, representing 0.1% of the total number of assembled transcripts) by Cuffmerge were extracted by selecting genomic regions assigned the "u" class code by the software. As this set still contained overlapping predicted transcripts (because slightly different versions of transcripts had been assembled from different samples), we consolidated it into a set of nonoverlapping genomic intervals (n=2430) representing putative novel intergenic transcripts. This set was compared with other sets of transcripts by using the BEDTools package (45). For comparison with BodyMap transcripts, raw FASTQ files were downloaded (ArrayExpress accession E-MTAB-513), mapped, and assembled in the same way as in the present experiment. For comparison to lncRNAs, data were downloaded from the supplementary material in Hangauer *et al.* (46), after which the genomic coordinates were converted from the human genome 18 (hg18) to the hg19 assembly using the liftOver tool in the University of California–Santa Cruz (UCSC) Genome Browser (47). Mammalian and primate conservation tracks were downloaded from the UCSC Table Browser (47). We defined as protein coding those genes where the ENSEMBL GTF file entry had a gene_biotype feature with the value "protein_coding," with data from a recent publication by Branca *et al.* (48).

Multivariate data analysis

Differences between the right (R) and left (L) legs, within-leg samples (A and B), and the males and females were analyzed with a series of different multivariate statistical tools, including principal component analysis (PCA) and orthogonal projections to latent structures (OPLS) by means of partial least squares. In OPLS, the systematic variation in the **X** matrix is divided into 2 parts, one that is linearly related to **Y** (and therefore predictive) and one that is unrelated (orthogonal) to **Y** (49, 50). Data were mean-centered and unit variance scaled before the analysis. The model complexity was estimated according to cross-validation (51). The software used for multivariate models was Simca *P*+13.0.3×64 (Umetrics, Umeå, Sweden). Multivariate data analysis techniques applied to gene and isoform expression analyses are described elsewhere (52, 53).

Supervised OPLS discriminant analysis (OPLS-DA) was applied to expression levels of 22,911 isoforms and 12,557 genes (mean FPKM >1), as **X**, and the male or female sample, as **Y,** analyzed in 34 samples (**Table 2**). The significance of differentially expressed isoforms and genes related to differences between the females and males was analyzed by means of loadings into the OPLS models, including the jackknife confidence levels of the loadings (50, 51). The significance was calculated as absolute value, ABS (loading) minus ABS (jackknife confidence interval), where a positive value for a gene indicated significance. Similarly, comparisons were also performed for left *vs.* right leg and A *vs.* B biopsies, in OPLS-DA models (for details regarding model

TABLE 2. *Model quality parameters of OPLS-DA*

Model description	PCs	Samples[a]	Isoforms/genes[b]	R_y^2	Q_y^2	R_X^2
Isoform level: male and female comparison (Fig 4*A*)	1 + 2	34	22,911	0.97	0.72	0.35
Gene level: male and female comparison (Fig. 4*B*)	1 + 2	34	12,557	0.98	0.79	0.29
Isoform level: male and female comparison, *X* and *Y* excluded (Supplemental Fig. S5*A*)	1 + 2	34	22,201	0.97	0.71	0.35
Gene level: male and female comparison, *X* and *Y* excluded (Supplemental Fig. S5*B*)	1 + 2	34	12,124	0.97	0.78	0.29
Isoform level: left and right leg comparison	1 + 0	34	22,911	0.81	−0.70	0.20
Gene level: left and right leg comparison	1 + 0	34	12,557	0.73	−1.16	0.19
Isoform level: *A* and *B* sample comparison	1 + 0	18	22,911	0.90	−0.40	0.25
Gene level: *A* and *B* sample comparison	1 + 0	28	12,369	0.88	−0.55	0.28

PCs refers to the optimal number of PCs for a particular multivariate data model, according to cross-validation. R_X^2 accounts for the explained variance. R_y^2 accounts for the explained variance for the *Y* component, and Q_Y^2 accounts for the cumulative fraction of the total variation of *Y* that can be predicted by the model. [a]For the left *vs.* right leg and male *vs.* female comparisons, the *B* biopsy is excluded; *A vs. B* comparisons include only legs with both *A* and *B* present. [b]If a variable contains only 1 or 2 variables that differ from the median in any comparison (model), that variable is excluded from subsequent analyses.

parameters, see Table 2). Furthermore, OPLS-DA was performed on the males and females when X and Y chromosome genes were excluded.

Nonsupervised PCA was applied to mean-centered and unit variance–scaled gene and isoform expression data to investigate differences between L and R legs and A and B biopsies. Included in the L and R leg comparison were 22,662 isoforms and 12,557 genes (34 samples), whereas the corresponding numbers were 22,662 isoforms and 12,369 genes for the A and B comparison (18 samples). All model parameters for PCA are shown in **Table 3**.

Pathway analysis

Genes expressed differentially between the males and females were analyzed with the Ingenuity pathway analysis software (Ingenuity Systems, Inc., Redwood, CA, USA; http://www.ingenuity.com). The analysis was performed on male and female enriched genes separately and based on the loading values for each gene from the OPLS-DA analysis (Ancillary Table A1; http://figshare.com/articles/Ancillary_Table_A1/1098772).

Citrate synthase (CS) activity

The experimental procedure was performed according to the fluorometric principles of Lowry and Passonneau (54). In brief, freeze-dried skeletal muscle samples were homogenized in 0.1 M phosphate buffer (pH 7.7) with 0.5% BSA. The tissue lysates were subsequently added to a reagent solution (0.1 M Tris-HCl, 2.5 mM EDTA, 0.5 mM L-malate, 512.5 nM NAD^+, and 399 μg MDH). Acetyl-CoA (50 μg) started the reaction, and the velocity was registered with a fluorometer (reduction of NAD^+ to NADH). An unpaired Student's *t* test was used to evaluate the differences between the males and females, which were considered significant at $P < 0.05$.

Real-time PCR

Five transcripts that differed significantly between the males and females according to the RNA-seq analysis were analyzed by using real-time PCR. Primers were supplied as TaqMan reagents (*ACTB* Hs99999903_m1, *ACTN3* Hs00153812_m1, *COL4A1* Hs00266237_m1, *EGLN3* Hs00222966_m1, *NUPR1* Hs01044304_g1, and *VEGFA* Hs00900055_m1; Life Technologies, Grand Island, NY, USA) and analyzed according to the manufacturer's specifications. β-Actin (*ACTB*) was used as an endogenous control, and expression levels were calculated by using $2^{-\Delta CT}$. An unpaired Student's *t* test was used to compare the means (significant at $P<0.05$).

RESULTS

General expression profile of human skeletal muscle

Skeletal muscle transcriptome analysis was performed through RNA-seq on skeletal muscle biopsies from 18 human volunteers, 9 males and 9 females. Six subjects had 2 biopsies taken from each leg, and the other 12 had 1 taken from each leg—in total, 48 samples. The paired-end sequencing reads were aligned to the human genome, and the normalized gene expression level was calculated as FPKM. As expected, most of the samples clustered by individual (Supplemental Fig. S2*A*). We found 12,659 unique known transcripts expressed in human skeletal muscle by using an average cutoff of FPKM >1 across 2 replicates from each individual. Of these, 10,419 represent protein-coding genes, the rest represent, for example, lncRNAs, intronic sequences, and pseudogenes. **Table 4** lists the 25 most transcribed mRNAs in skeletal muscle based on the FPKM value. Eight of the top 10 transcribed genes were mitochondrial: ATP synthase–associated genes (*MT-ATP8*, and *MT-ATP6*) and components of the electron transport chain (*MT-CO1*, *MT-CO2*, *MT-CO3*, *MT-ND1*, *MT-ND3*, and *MT-ND4*). Other common transcripts were related to structural components: α-actin 1 (*ACTA1*), myosin light chain 2 (*MYL2*), troponins (*TNNC2*, and *TNNT1*), and tropomyosin (*TPM2*). Skeletal muscle metabolic enzymes were also highly abundant; aldolase A (*ALDOA*), glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), and creatine kinase (*CKM*). For a full transcript list, see Supplemental Material S1.

GO analysis of the protein-coding transcripts showed that the dominating molecular functions were binding and catalytic activity (**Fig. 1*A***). The most common biological process was metabolism, and most of the categorized transcripts were intracellular (Supplemental Fig. S2*B*, *C*). Comparing our skeletal muscle expression profiles with 7 tissues obtained from another study (Illumina BodyMap), we noted that although laboratory protocol differences and batch effects inevitably introduce some differences, skeletal and heart muscle were the tissues that mostly resembled ours (Fig. 1*B*).

The RNA-seq data were also correlated to microarray expression data from 24 skeletal muscle samples from another published study (44). The Spearman's correlation between the expression levels was significant at both the gene and transcript levels ($\rho=0.65$ and 0.63, respectively, $P<0.0001$; Supplemental Fig. S2*D*, *E*).

TABLE 3. *Model quality parameters of PCA*

Model description	PCs	Samples[a]	Isoforms/genes[b]	R_X^2	Q_X^2
Isoform level: *A* and *B* sample comparison (Fig. 3*A*)	3	18	22,662	0.40	0.06
Isoform level: left and right leg comparison (Fig. 3*C*)	4	34	22,662	0.39	0.24
Gene level: *A* and *B* sample comparison (Fig. 3*B*)	4	18	12,369	0.46	0.12
Gene level: left and right leg comparison (Fig. 3*D*)	5	34	12,557	0.40	0.24

PCs refers to the optimal number of PCs for a particular multivariate data model, according to cross-validation. R_X^2 accounts for the explained variance. Q_X^2 accounts for the cumulative fraction of the total variation of X that can be predicted by the model. [a]For the left *vs.* right leg comparisons, the *B* biopsy is excluded; *A vs. B* comparisons include only legs with both *A* and *B* present. [b]If a variable contains only 1 or 2 variables that differ from the median in any comparison (model), that variable is excluded from subsequent analyses.

TABLE 4. *Top 25 transcribed mRNAs in human skeletal muscle*

Gene symbol	FPKM	Gene name
MT-ATP8	52,175	Mitochondrially encoded ATP synthase 8
ACTA1	36,127	Actin, α1, skeletal muscle
MT-CO1	33,148	Mitochondrially encoded cytochrome C oxidase I
MT-ATP6	29,400	Mitochondrially encoded ATP synthase 6
MT-CO3	26,881	Mitochondrially encoded cytochrome C oxidase III
MT-CO$_2$	23,721	Mitochondrially encoded cytochrome C oxidase II
MT-ND3	22,489	Mitochondrially encoded NADH dehydrogenase 3
MT-ND1	18,009	Mitochondrially encoded NADH dehydrogenase 1
MT-ND4	17,971	Mitochondrially encoded NADH dehydrogenase 4
CKM	16,918	Creatine kinase, muscle
MT-ND4L	16,439	Mitochondrially encoded NADH dehydrogenase 4L
MT-CYB	14,896	Mitochondrially encoded cytochrome B
MT-ND6	11,548	Mitochondrially encoded NADH dehydrogenase 6
MT-ND2	10,314	Mitochondrially encoded NADH dehydrogenase 2
MB	9,725	Myoglobin
TPM2	9,313	Tropomyosin 2 (β)
ALDOA	8,319	Aldolase A, fructose-bisphosphate
AC018512.1	8,238	—
TNNC2	7,797	Troponin C type 2 (fast)
MYL2	7,558	Myosin, light chain 2, regulatory, cardiac, slow
GAPDH	7,164	Glyceraldehyde-3-phosphate dehydrogenase
MT-ND5	6,232	Mitochondrially encoded NADH dehydrogenase 5
TNNT1	6,130	Troponin T type 1 (skeletal, slow)
TNNT3	6,071	Troponin T type 3 (skeletal, fast)
MTATP6P1	5,963	Mitochondrially encoded ATP synthase 6 pseudogene 1

Average of 2 isoforms/gene expressed in skeletal muscle

Analysis of the skeletal muscle isoform expression identified 23,002 known isoforms (FPKM>1; Supplementary Material S1), with 15,455 being protein coding. The average number of isoforms per gene was 2.12. However, there were 113 genes that expressed 10 or more ENSEMBL annotated isoforms in skeletal muscle tissue at baseline according to Cufflinks analysis. A histogram of the isoform expression is shown in Fig. 1*C*. We investigated 4 of the genes with 10 or more isoforms more closely; *ALDOA* (19 isoforms), *CS* (12 isoforms), *TTN* (10 isoforms), and *VEGFA* (10 isoforms). *CS* and *TTN* both expressed 1 major isoform (ENST00000351328 and ENST00000460472, respectively), whereas the remainder had comparably low expression (FPKM=1–15). *ALDOA* had 11 isoforms with an average FPKM >100, and only 5 isoforms were expressed at FPKM ≤10. *VEGFA* expressed most isoforms at similar levels (7 isoforms with FPKM=4–12). For all isoforms investigated more closely, the abundant ones were high throughout all subjects, whereas the low-expression ones were more variable. None showed an individual-specific expression pattern, indicating that the number of isoforms was not due to the number of subjects included in the analysis. A library preparation batch effect, however, was more apparent in the isoform *vs.* the gene level.

Transcriptome assembly identified 5 new protein-coding transcripts

Reference-based assembly from all sequencing data, in total more than 830 million paired-end reads, identified 2430 novel (*i.e.*, not annotated in ENSEMBL v.71) transcribed loci in human skeletal muscle. Approximately 20% overlapped lncRNAs recently identified by Hangauer *et al.* (46). Most of the novel transcripts were also expressed in at least one BodyMap tissue, and nearly all of the regions contained highly conserved regions in mammals (**Fig. 2*A***), indicating that the transcripts identified are of functional importance. The level of transcription was low in general, with 373 transcripts expressed at an average FPKM >1 across all samples. Five of the novel transcripts colocalized in the genome with newly identified peptides from a recent high-throughput human proteomics study by Branca *et al.* (48), one example is shown in Fig. 2*B*. One of these was overlapped with an lncRNA, and the remaining 4 were found in BodyMap data.

Extremely high spatial expression similarity within and between legs

To analyze spatial expression within the vastus lateralis muscle, we examined 2 biopsies from each leg of the first 6 individuals. PCA and OPLS-DA were applied to investigate whether there were any significant differences in isoform or gene expression levels. A nonsupervised PCA showed that the within-leg samples did not form 2 clusters in a PC score plot (**Fig. 3*A, B***). When we applied a supervised OPLS-DA analysis on 2 predefined clusters (the 2 biopsies from the same leg), the model parameters clearly suggested that there were no statistically significant differences (Table 2). In this case, the predictive power of the model (Q_y^2), derived from cross-validation, was negative for both isoform and gene level data (−0.40 and −0.55, respectively), clearly showing that the model was not satisfactory (Table 2). Thus,

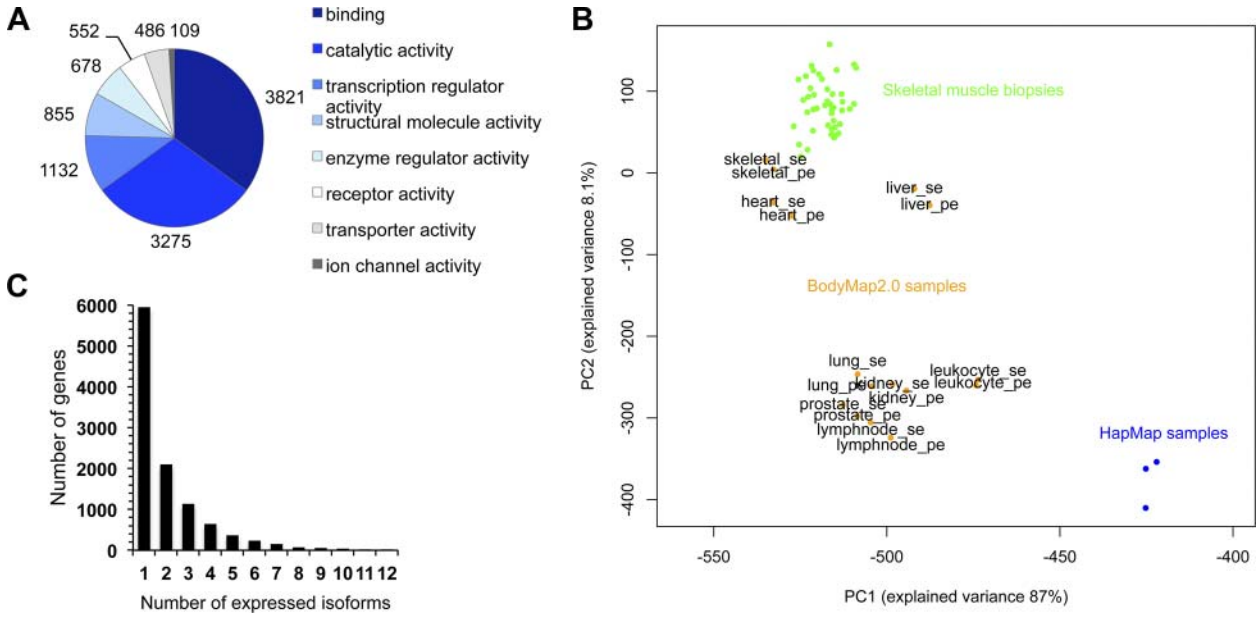

Figure 1. Transcriptional profile of human skeletal muscle. *A*) Skeletal muscle GO for molecular function, derived from the Panther database (http://www.pantherdb.org, n=10,268 protein-coding genes). Categories representing ≥1% of the total number of functions are included; numbers indicate number of genes expressed in skeletal muscle from each category. *B*) PCA plot of skeletal muscle RNA-seq samples in the present study in the context of other public human RNA-seq tissue data. Green circles represent the biopsies in the present study. Orange circles represent samples from Illumina's BodyMap project (GEO accession GSE30611); these circles are labeled with tissue followed by "_se" (for single-end sequencing) or "_pe" (for paired-end sequencing). Blue circles represent 3 randomly selected lymphoblastoid cell lines from the individuals (one Yoruban and 2 central European) from the HapMap project (ArrayExpression accession E-MTAB-197; individuals NA19238, NA06985, and NA06986). *C*) Histogram showing the number of genes expressing different numbers of alternatively spliced variants in human skeletal muscle at baseline (n=36, average FPKM>1).

the multivariate analyses clearly showed that there was a very high degree of similarity between the 2 muscle biopsies from the same leg.

For analysis of expression differences between biopsies from the 2 legs of an individual, data from 34 samples were used. PCA of both isoform and gene expression data showed that it was not possible to distinguish between the 2 legs (*i.e.*, no distinct clusters were visible among the samples in the analysis; Fig. 3*C*, *D*). OPLS-DA did not lead to a satisfactory model according to cross-validation (Q_y^2 of −0.70 and −1.16 for isoform- and gene-level analyses, respectively; Table 2). Thus, there were no statistically significant differences identified for the comparison of samples from 2 different legs. This result was also sup-

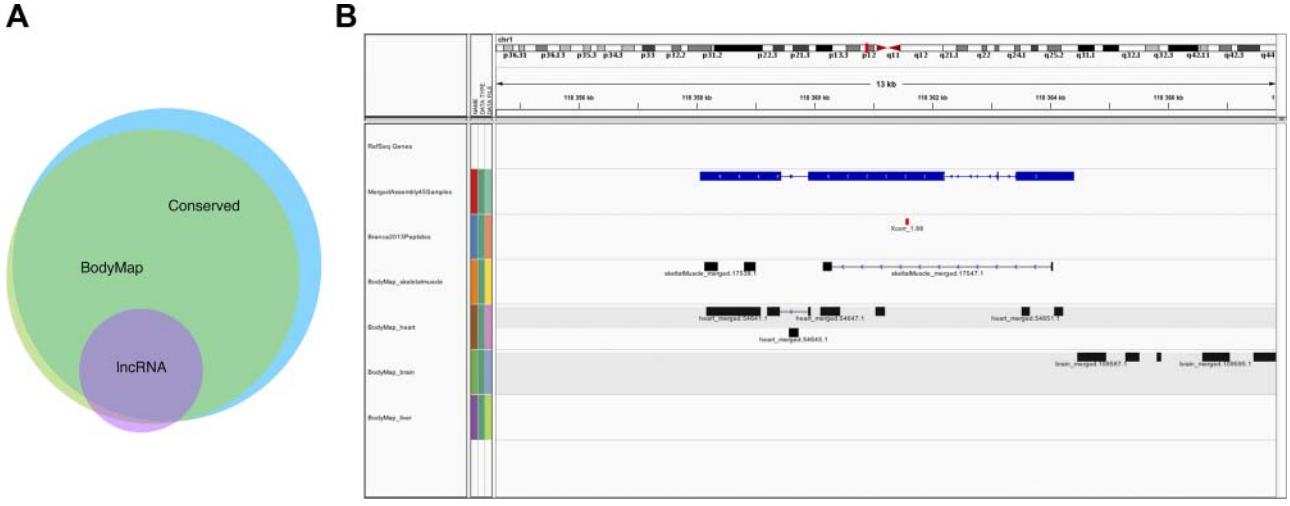

Figure 2. Identification of novel transcripts. *A*) Venn diagram showing the distribution of novel transcripts. Of the 2430 transcripts, 405 were overlapping lncRNAs, 1891 were found through reassembly of BodyMap data, and 2400 contained highly conserved regions in mammals. Eleven could not be further categorized (not shown). *B*) Example of one of the novel transcripts that overlapped with a newly identified peptide from Branca *et al.* (48). Our predicted transcript is shown in blue and the peptide in red. Reassembled BodyMap data from skeletal muscle and heart, brain, and liver tissue are also displayed in black.

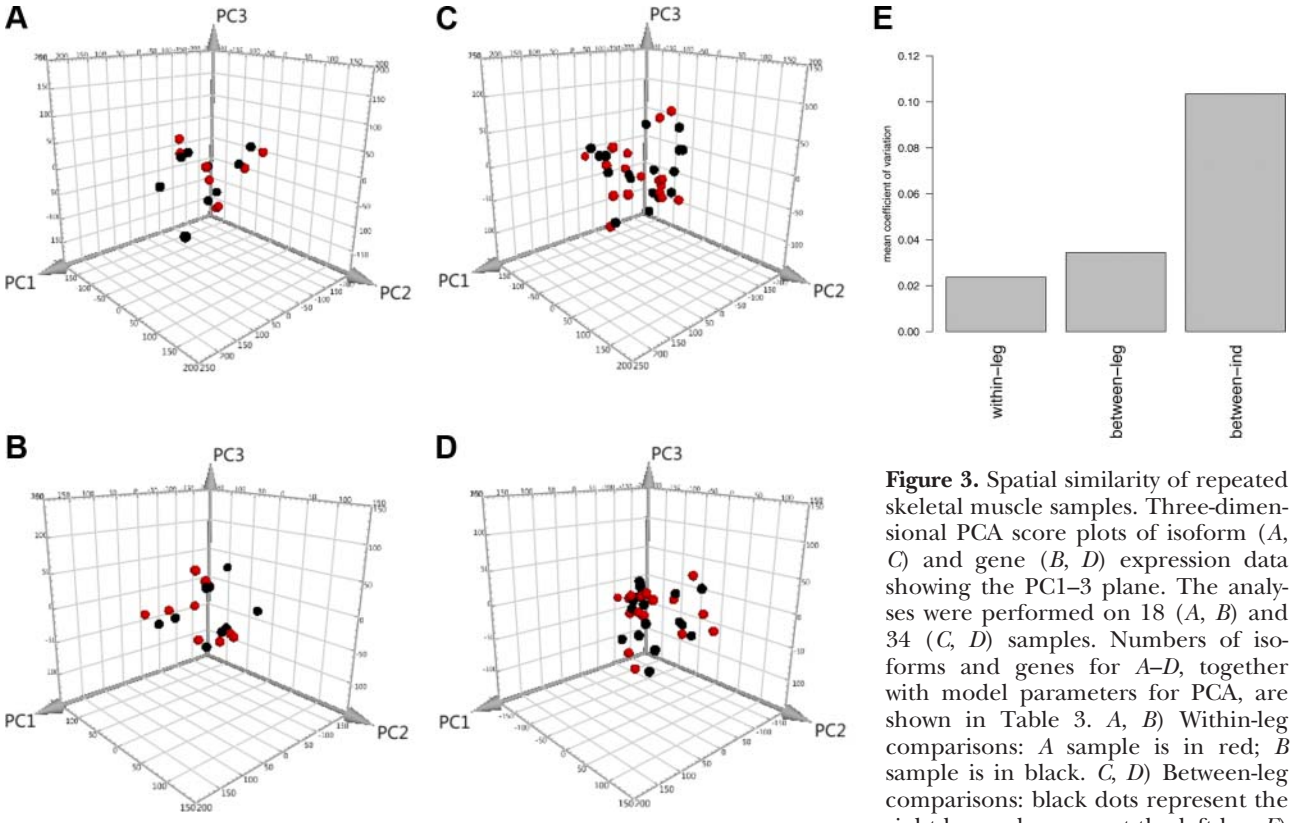

Figure 3. Spatial similarity of repeated skeletal muscle samples. Three-dimensional PCA score plots of isoform (*A*, *C*) and gene (*B*, *D*) expression data showing the PC1–3 plane. The analyses were performed on 18 (*A*, *B*) and 34 (*C*, *D*) samples. Numbers of isoforms and genes for *A*–*D*, together with model parameters for PCA, are shown in Table 3. *A*, *B*) Within-leg comparisons: *A* sample is in red; *B* sample is in black. *C*, *D*) Between-leg comparisons: black dots represent the right leg; red represent the left leg. *E*) Comparison of the mean biological coefficient of variation derived from EdgeR for within-leg samples, between-leg samples, and between individuals.

ported by the finding of no differentially expressed genes at an individual level, by DESeq2 at FDR < 0.05. However, the finding of no statistically significant differentially expressed genes across the samples does not mean that the samples were identical. To further address any transcriptional variability between samples, we continued by analyzing the biological coefficient of variation, comparing samples within the same leg, between legs, and between individuals (Fig. 3*E*). As expected, we found that samples taken from within the same leg were slightly more similar than samples from the different legs of an individual. Between individuals, the variability was 5 times greater than in samples taken within the same muscle.

Sex-specific differences in the global transcriptome

To find out whether there were differences between the males and females, a supervised OPLS-DA model was adopted (**Fig. 4*A*, *B*** and Table 2). Both isoform- and gene-level analyses showed clear differences (Fig. 4*A*, *B*). The model parameters suggested 1 predictive and 2 orthogonal components, both regarding goodness of fit of the model as well as cross-validation (R_X^2, R_Y^2, and Q_Y^2 respectively, Table 2), at both isoform and gene level (Table 2). Similar results were found when the X and Y chromosomes were excluded (Table 2 and Supplemental Fig. S3*A*, *B*). The significance of differentially expressed isoforms and genes between the males and females was analyzed by means of loadings in the OPLS-DA model. An isoform or gene was considered significant if the confidence interval of the loading value (estimated by jackknifing) did not contain the value 0. The number of statistically significant, differentially expressed isoforms and genes was 5434 and 3038, respectively (Supplemental Material S1). In the females, 1923 isoforms and 1311 genes were more highly expressed than in the males, where the corresponding numbers were 3511 and 1727, respectively (Supplemental Material S1). Supplemental Figure S3*C* shows the chromosome distribution of the genes enriched in the males and females. The number of isoforms and genes with a fold difference ≥2 was 169 and 56 for the males and 139 and 42 for the females. A technical validation using real-time PCR was performed on 5 gene transcripts (2 structural and 3 regulatory) that were found to be significantly different between the males and females. All transcripts showed enrichment in the expected direction, but in 1 case, the enrichment was not statistically significant (*ACTN3*: males, 1.49±0.97; females, 0.42±0.31, P=0.006; *COL4a1*: males, 0.13±0.06; females, 0.22±0.11, P=0.047; *EGLN3*: males, 0.05±0.02; females, 0.13±0.10, P=0.02; *NUPR1*: males, 0.98±0.55; females, 0.78±0.53, P=0.46; and *VEGFA*: males 1.02±0.54, females, 2.49±1.76, P=0.03). The data are related to β-actin (*ACTB*) and expressed as means ± SD.

There was a clear difference in functional GO between the males and females (Fig. 4*C*). In the females, the enriched genes mainly belonged to the mitochondria and were involved in biological processes related to metabolism of acids and ketones, oxidation and reduction, cellular respiration, and fatty acid metabolism. In

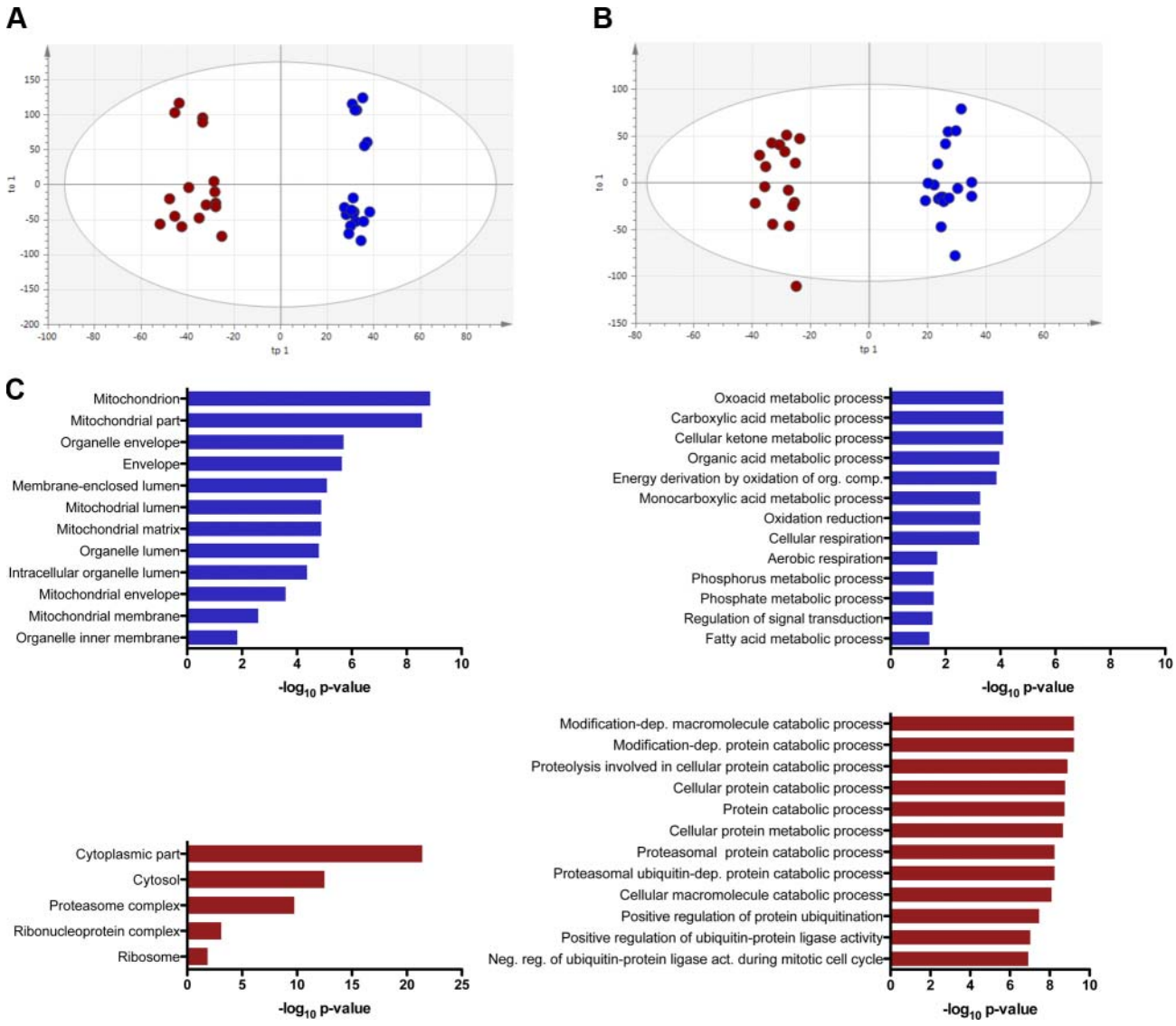

Figure 4. Sex differences in the human skeletal muscle transcriptome. *A, B*) Two-dimensional score plots of OPLS-DA showing the first predictive component (tp1) and orthogonal component (to1) plane of female (blue) and male (red) samples with respect to isoform (*A*) and gene expression data (*B*). Numbers of genes/isoforms included, as well as model parameters for the OPLS-DA, are shown in Table 2. *C*) Ontology distribution of the differential genes in the males and females, respectively, analyzed using the Data for Annotation, Visualization and Integrated Discovery (DAVID) tool (http://david.abcc.ncifcrf.gov). Left: cellular component ontology enriched in the females (top left) and the males (bottom left). Right: biological process ontology for genes enriched in the females (top right) and the males (bottom right, note that there were in total 35 that were significantly enriched in the males, and only the top 12 are shown). Ontologies included had a fold enrichment ≥1.5 and FDR < 0.05.

the males, many of the enriched genes were found in the cytoplasm and proteasome, and enriched biological processes were almost all related to protein catabolism (Fig. 4*C*). The molecular function ontology is shown in Supplemental Fig. S3*D*. The GO for isoform expression was very similar to that of the genes. Despite the higher mitochondrial transcript levels in the females, there was no significant difference in citrate synthase activity (a marker of mitochondrial density) between the male and female skeletal muscle samples (males, 16.50± 6.91; females, 21.15±5.85; P=0.18; mean ± SD).

Using Ingenuity pathway analysis we also identified potential upstream regulators controlling expression of the differential genes. In the males, 2 factors involved in translation were predicted to be activated: eukaryotic translation initiation factors 2-α kinase 3 (EIF2AK3) and 4E family member 2 (EIF4E). The most significant factor predicted to be activated in females was P53 (TP53). A mechanistic network of regulation for the females, involving P53, is illustrated in Supplemental Fig. S3*E*. Several of the peroxisome proliferator-activated receptors (PPARs; *e.g.*, PPARG and PPARA) were also predicted to be activated, as well as β-estradiol.

DISCUSSION

The current investigation provided an extensive transcriptional map of human skeletal muscle tissue, in-

cluding previously uncharacterized global isoform expression and several novel transcripts. It also showed that muscle tissue obtained from similar sites is remarkably homogeneous with regard to global gene expression profile. Finally, the results demonstrated that there are differences between male and female skeletal muscle transcriptomes, with ~3000 genes being differentially expressed.

RNA sequencing data of human skeletal muscle tissue is very limited. Studies of the global transcriptome have previously largely used microarray technology and have been focused on differential expression after an intervention or treatment, or comparing different populations (1, 4–6, 9, 12, 13, 15, 18, 19, 55). Our data identified over 10,000 transcribed protein-coding genes in skeletal muscle, in accordance with what is "normally" expressed in most other tissues (56). One important novel finding was that >23,000 different isoforms were identified in human skeletal muscle. Isoform expression detection is, however, methodologically challenging. Even with 2 × 100-bp paired-end sequencing, in many cases (particularly for annotated isoforms that differ only slightly), it is difficult to assign transcripts to different isoforms. More detailed analysis is therefore needed of specific isoforms of interest for different contexts.

From the RNA-seq data, we also performed a reference-based *de novo* assembly in order to identify new, previously unannotated transcripts. We found almost 2500 novel transcribed loci expressed in skeletal muscle, of which the majority (99.5%) are associated with lncRNAs (46), found in other RNA-seq data sets from BodyMap, or conserved in mammals, indicating that they are transcripts of functional importance. Interestingly, we identified 5 completely novel transcripts that colocalize in the genome with newly identified peptides (48), which strongly indicates that they are protein-coding genes.

In this study, we obtained multiple muscle samples at different sites and from a similar spatial volume of the muscle. Despite the samples being from different sites and the fact that the fitness of the 2 legs might differ somewhat, the spatial transcriptional homogeneity in human skeletal muscle, both within a muscle and between 2 different legs of an individual, turned out to be very high. To our knowledge, this is the first study to investigate this at a global level. However, there is a study that has analyzed a few selected genes after multiple spatial skeletal muscle sampling (57). Knowledge about baseline variability within a tissue is important for all studies involving repeated sampling. Several studies have also analyzed the effect of repeated temporal sampling for expression of a few selected genes (2, 57–59). At a global level, this remains to be investigated. Our results showed that, as long as samples are taken within 15 min from the same part and similar depth of the muscle, the transcriptional profile is not significantly different from that of repeated biopsies taken from either the same or both legs of an individual.

We found a profound transcriptomic difference between male and female skeletal muscle tissue, with >3000 genes and >5000 isoforms being significant. There is a known sex difference in fiber type composition, with males having a larger type II area, thus a more glycolytic phenotype, whereas females show a higher percentage area of type I fibers, with a more oxidative phenotype (60). These differences were reflected in our transcriptome analysis, where the males were enriched for many glycolytic mRNAs, including 5 isoforms of *LDHA*, 12 isoforms of *PFKM*, 3 isoforms of *GAPDH*, 1 *GPI*, 1 *ALDOA*, and the type II-specific marker *ACTN3*. However, the hexokinase 2 enzyme (*HK2*) was higher in the females. In contrast, the females were enriched for oxidative markers, with robust enrichment in mitochondrial function, including the cofactor PGC1α (*PPARGC1a*)—which, interestingly, is lower in skeletal muscle of type II diabetics (61)—and the important enzyme *CS*. However, at the protein level, there was no significant difference in the activity of CS, although the average was higher in the females, similar to that observed for mitochondrial enzymes in other studies (62, 63). *NRIP1*, an inhibitor of PGC1α more abundant in glycolytic fibers (64), was instead enriched in the males.

Oxidative type I fibers have a higher capillary density, and in accordance with that, we found several endothelial markers—*FLT1* and *-4*, *KDR*, and *TEK* (65, 66)—to be enriched in the females. Other angiogenic factors were also higher in the females: *ANGPT1*, 10 isoforms of *VEGFA*, and the growth factor receptor *FGFR1*. For the interpretation of expression differences between the sexes it is also important to consider that, in some cases, there are discrepancies between mRNA and the corresponding protein levels, activity, or both (67, 68). A global microarray approach to investigating baseline sex differences in expression was adopted by Maher *et al.* (26), where 66 differentially expressed genes with a fold difference ≥1.2 were identified. Half of these were concordant with our results. Our systematic GO analysis also confirmed part of their findings that a few mitochondrial and metabolic genes are higher in females than in males, whereas we also provided a clear ontology for genes enriched in males. Using oligonucleotide microarrays, Welle *et al.* (27) found ~3000 genes to be differentially expressed between males and females, a number similar to that in our study. However, out of the 2400 probes with $P < 0.01$, only 25% of the genes were concordant with our results. Older individuals were included by Welle *et al.*, thus subject variability, as well as methodological and statistical differences, can explain part of the discrepancy.

In summary, our data provide a deep and extensive baseline reference for the human skeletal muscle transcriptome, with regard to alternative splicing events, novel transcript expression, and sex differences in functional ontology. We also demonstrated a remarkably high spatial skeletal muscle tissue homogeneity, both within and between the 2 legs of an individual. That >5000 isoforms differed in expression between the males and females may explain differences in physiological phenotype and disease susceptibility between the sexes. The results could also contribute to a better understanding of the molecular pathways involved in physiological adaptations, disease development, and different dysfunctions. FJ

The authors thank Eva-Karin Gidlund and Susanna Appel (Karolinska Institutet) for assistance with physiological testing and biopsy handling and Science for Life Laboratory (SciLifeLab, Stockholm, Sweden), National Genomics Infrastructure (SciLifeLab), and Uppmax (Uppsala University, Uppsala, Sweden) for providing massive parallel sequencing and computational infrastructure. This work was supported by the Swedish National Centre for Research in Sports and the Wallenberg Advanced Bioinformatics Infrastructure. M.E.L has a PhD training grant from Karolinska Institutet.

REFERENCES

1. Timmons, J. A., Jansson, E., Fischer, H., Gustafsson, T., Greenhaff, P. L., Ridden, J., Rachman, J., and Sundberg, C. J. (2005) Modulation of extracellular matrix genes reflects the magnitude of physiological adaptation to aerobic exercise training in humans. *BMC Biol.* **3,** 19
2. Malm, C., Nyberg, P., Engstrom, M., Sjodin, B., Lenkei, R., Ekblom, B., and Lundberg, I. (2000) Immunological changes in human skeletal muscle and blood after eccentric exercise and multiple biopsies. *J. Physiol.* **529,** 243–262
3. Little, J. P., Safdar, A., Cermak, N., Tarnopolsky, M. A., and Gibala, M. J. (2010) Acute endurance exercise increases the nuclear abundance of PGC-1alpha in trained human skeletal muscle. *Am J. Physiol. Regul. Integr. Comp. Physiol.* **298,** R912–R917
4. Mahoney, D. J., Parise, G., Melov, S., Safdar, A., and Tarnopolsky, M. A. (2005) Analysis of global mRNA expression in human skeletal muscle during recovery from endurance exercise. *FASEB J.* **19,** 1498–1500
5. Gordon, P. M., Liu, D., Sartor, M. A., IglayReger, H. B., Pistilli, E. E., Gutmann, L., Nader, G. A., and Hoffman, E. P. (2012) Resistance exercise training influences skeletal muscle immune activation: a microarray analysis. *J. Appl. Physiol.* **112,** 443–453
6. Roth, S. M., Ferrell, R. E., Peters, D. G., Metter, E. J., Hurley, B. F., and Rogers, M. A. (2002) Influence of age, sex, and strength training on human muscle gene expression determined by microarray. *Physiol. Genomics* **10,** 181–190
7. Gustafsson, T., Osterlund, T., Flanagan, J. N., von Walden, F., Trappe, T. A., Linnehan, R. M., and Tesch, P. A. (2010) Effects of 3 days unloading on molecular regulators of muscle size in humans. *J. Appl. Physiol. (1985)* **109,** 721–727
8. Dalla Libera, L., Ravara, B., Gobbo, V., Tarricone, E., Vitadello, M., Biolo, G., Vescovo, G., and Gorza, L. (2009) A transient antioxidant stress response accompanies the onset of disuse atrophy in human skeletal muscle. *J. Appl. Physiol.* **107,** 549–557
9. Sreekumar, R., Halvatsiotis, P., Schimke, J. C., and Nair, K. S. (2002) Gene expression profile in skeletal muscle of type 2 diabetes and the effect of insulin treatment. *Diabetes* **51,** 1913–1920
10. Olsson, A. H., Ronn, T., Elgzyri, T., Hansson, O., Eriksson, K. F., Groop, L., Vaag, A., Poulsen, P., and Ling, C. (2011) The expression of myosin heavy chain (MHC) genes in human skeletal muscle is related to metabolic characteristics involved in the pathogenesis of type 2 diabetes. *Mol. Genet. Metab.* **103,** 275–281
11. Kulkarni, S. S., Salehzadeh, F., Fritz, T., Zierath, J. R., Krook, A., and Osler, M. E. (2012) Mitochondrial regulators of fatty acid metabolism reflect metabolic dysfunction in type 2 diabetes mellitus. *Metabolism* **61,** 175–185
12. Baron, D., Magot, A., Ramstein, G., Steenman, M., Fayet, G., Chevalier, C., Jourdon, P., Houlgatte, R., Savagner, F., and Pereon, Y. (2011) Immune response and mitochondrial metabolism are commonly deregulated in DMD and aging skeletal muscle. *PLoS One* **6,** e26952
13. 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
14. Melia, M. J., Kubota, A., Ortolano, S., Vilchez, J. J., Gamez, J., Tanji, K., Bonilla, E., Palenzuela, L., Fernandez-Cadenas, I., Pristoupilova, A., Garcia-Arumi, E., Andreu, A. L., Navarro, C., Hirano, M., and Marti, R. (2013) Limb-girdle muscular dystrophy 1F is caused by a microdeletion in the transportin 3 gene. *Brain* **136,** 1508–1517
15. Millino, C., Fanin, M., Vettori, A., Laveder, P., Mostacciuolo, M. L., Angelini, C., and Lanfranchi, G. (2009) Different atrophy-hypertrophy transcription pathways in muscles affected by severe and mild spinal muscular atrophy. *BMC Med* **7,** 14
16. Reardon, K. A., Davis, J., Kapsa, R. M., Choong, P., and Byrne, E. (2001) Myostatin, insulin-like growth factor-1, and leukemia inhibitory factor mRNAs are upregulated in chronic human disuse muscle atrophy. *Muscle Nerve* **24,** 893–899
17. Khal, J., Hine, A. V., Fearon, K. C., Dejong, C. H., and Tisdale, M. J. (2005) Increased expression of proteasome subunits in skeletal muscle of cancer patients with weight loss. *Int. J. Biochem. Cell Biol.* **37,** 2196–2206
18. Crimi, M., Bordoni, A., Menozzi, G., Riva, L., Fortunato, F., Galbiati, S., Del Bo, R., Pozzoli, U., Bresolin, N., and Comi, G. P. (2005) Skeletal muscle gene expression profiling in mitochondrial disorders. *FASEB J.* **19,** 866–868
19. Larrouy, D., Barbe, P., Valle, C., Dejean, S., Pelloux, V., Thalamas, C., Bastard, J. P., Le Bouil, A., Diquet, B., Clement, K., Langin, D., and Viguerie, N. (2008) Gene expression profiling of human skeletal muscle in response to stabilized weight loss. *Am. J. Clin. Nutr.* **88,** 125–132
20. Black, D. L. (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.* **72,** 291–336
21. Kornblihtt, A. R., Schor, I. E., Allo, M., Dujardin, G., Petrillo, E., and Munoz, M. J. (2013) Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat. Rev. Mol. Cell Biol.* **14,** 153–165
22. Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* **456,** 470–476
23. Blekhman, R., Marioni, J. C., Zumbo, P., Stephens, M., and Gilad, Y. (2010) Sex-specific and lineage-specific alternative splicing in primates. *Genome Res.* **20,** 180–189
24. Su, W. L., Modrek, B., GuhaThakurta, D., Edwards, S., Shah, J. K., Kulkarni, A. V., Russell, A., Schadt, E. E., Johnson, J. M., and Castle, J. C. (2008) Exon and junction microarrays detect widespread mouse strain- and sex-bias expression differences. *BMC Genomics* **9,** 273
25. Hartmann, B., Castelo, R., Minana, B., Peden, E., Blanchette, M., Rio, D. C., Singh, R., and Valcarcel, J. (2011) Distinct regulatory programs establish widespread sex-specific alternative splicing in Drosophila melanogaster. *RNA* **17,** 453–468
26. Maher, A. C., Fu, M. H., Isfort, R. J., Varbanov, A. R., Qu, X. A., and Tarnopolsky, M. A. (2009) Sex differences in global mRNA content of human skeletal muscle. *PLoS One* **4,** e6335
27. Welle, S., Tawil, R., and Thornton, C. A. (2008) Sex-related differences in gene expression in human skeletal muscle. *PLoS One* **3,** e1385
28. Liu, D., Sartor, M. A., Nader, G. A., Gutmann, L., Treutelaar, M. K., Pistilli, E. E., Iglayreger, H. B., Burant, C. F., Hoffman, E. P., and Gordon, P. M. (2010) Skeletal muscle gene expression in response to resistance exercise: sex specific regulation. *BMC Genomics* **11,** 659
29. Seip, R. L., and Semenkovich, C. F. (1998) Skeletal muscle lipoprotein lipase: molecular regulation and physiological effects in relation to exercise. *Exerc. Sport Sci. Rev.* **26,** 191–218
30. Koval, J. A., DeFronzo, R. A., O'Doherty, R. M., Printz, R., Ardehali, H., Granner, D. K., and Mandarino, L. J. (1998) Regulation of hexokinase II activity and expression in human muscle by moderate exercise. *Am. J. Physiol.* **274,** E304–E308
31. Kraniou, Y., Cameron-Smith, D., Misso, M., Collier, G., and Hargreaves, M. (2000) Effects of exercise on GLUT-4 and glycogenin gene expression in human skeletal muscle. *J. Appl. Physiol. (1985)* **88,** 794–796
32. Mahoney, D. J., Carey, K., Fu, M. H., Snow, R., Cameron-Smith, D., Parise, G., and Tarnopolsky, M. A. (2004) Real-time RT-PCR analysis of housekeeping genes in human skeletal muscle following acute exercise. *Physiol. Genomics* **18,** 226–231
33. Norrbom, J., Sallstedt, E. K., Fischer, H., Sundberg, C. J., Rundqvist, H., and Gustafsson, T. (2011) Alternative splice variant PGC-1alpha-b is strongly induced by exercise in human skeletal muscle. *Am. J. Physiol. Endocrinol. Metab.* **301,** E1092–E1098

34. Lexell, J., Taylor, C., and Sjostrom, M. (1985) Analysis of sampling errors in biopsy techniques using data from whole muscle cross sections. *J. Appl. Physiol.* **59,** 1228–1235
35. Lexell, J., and Taylor, C. C. (1989) Variability in muscle fibre areas in whole human quadriceps muscle: how much and why? *Acta Physiol. Scand.* **136,** 561–568
36. Polgar, J., Johnson, M. A., Weightman, D., and Appleton, D. (1973) Data on fibre size in thirty-six human muscles: an autopsy study. *J. Neurol. Sci.* **19,** 307–318
37. Punkt, K., Mehlhorn, H., and Hilbig, H. (1998) Region- and age-dependent variations of muscle fibre properties. *Acta Histochem.* **100,** 37–58
38. Kohn, T. A., and Myburgh, K. H. (2007) Regional specialization of rat quadriceps myosin heavy chain isoforms occurring in distal to proximal parts of middle and deep regions is not mirrored by citrate synthase activity. *J. Anat.* **210,** 8–18
39. Bergström, J. (1962) Muscle electrolytes in man. *Scand. J. Clin. Lab. Invest.* **14,** 1–110
40. Stranneheim, H., Werne, B., Sherwood, E., and Lundeberg, J. (2011) Scalable transcriptome preparation for massive parallel sequencing. *PLoS One* **6,** e21910
41. Trapnell, C., Pachter, L., and Salzberg, S. L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25,** 1105–1111
42. Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., Salzberg, S. L., Wold, B. J., and Pachter, L. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28,** 511–515
43. Anders, S., and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome. Biol.* **11,** R106
44. Timmons, J. A., Knudsen, S., Rankinen, T., Koch, L. G., Sarzynski, M. A., Jensen, T., Keller, P., Scheele, C., Vollaard, N. B., Nielsen, S., Akerstrom, T., Macdougald, O. A., Jansson, E., Greenhaff, P. L., Tarnopolsky, M. A., van Loon, L. J., Pedersen, B. K., Sundberg, C. J., Wahlestedt, C., Britton, S. L., and Bouchard, C. (2010) Using molecular classification to predict gains in maximal aerobic capacity following endurance exercise training in humans. *J. Appl. Physiol.* **108,** 1487–1496
45. Quinlan, A. R., and Hall, I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842
46. Hangauer, M. J., Vaughn, I. W., and McManus, M. T. (2013) Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet.* **9,** e1003569
47. Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., and Haussler, D. (2002) The human genome browser at UCSC. *Genome Res.* **12,** 996–1006
48. Branca, R. M., Orre, L. M., Johansson, H. J., Granholm, V., Huss, M., Perez-Bercoff, A., Forshed, J., Kall, L., and Lehtio, J. (2014) HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics. *Nat. Methods* **11,** 69–52
49. Trygg, J., and Wold, S. (2002) Orthogonal projections to latent structures (O-PLS). *J. Chemometr.* **16,** 119–128
50. Martens, H., Hoy, M., Westad, F., Folkenberg, B., and Martens, M. (2001) Analysis of designed experiments by stabilised PLS Regression and jack-knifing. *Chemometr. Intell. Lab. Syst.* **58,** 151–170
51. Wold, S. (1978) Cross-validatory estimation of the number of components in factor and principal components models. *Technometrics* **20,** 397–405
52. Kurtovic, S., Paloschi, V., Folkersen, L., Gottfries, J., Franco-Cereceda, A., and Eriksson, P. (2011) Diverging alternative splicing fingerprints in the transforming growth factor-beta signaling pathway identified in thoracic aortic aneurysms. *Mol. Med.* **17,** 665–675
53. Kjellqvist, S., Maleki, S., Olsson, T., Chwastyniak, M., Branca, R. M., Lehtio, J., Pinet, F., Franco-Cereceda, A., and Eriksson, P. (2013) A combined proteomic and transcriptomic approach shows diverging molecular mechanisms in thoracic aortic aneurysm development in patients with tricuspid- and bicuspid aortic valve. *Mol. Cell. Proteomics* **12,** 407–425
54. Lin, L., Sotonyi, P., Somogyi, E., Karlsson, J., Folkers, K., Nara, Y., Sylven, C., Kaijser, L., and Jansson, E. (1988) Coenzyme Q10 content in different parts of the normal human heart. *Clin. Physiol.* **8,** 391–398
55. Chopard, A., Lecunff, M., Danger, R., Lamirault, G., Bihouee, A., Teusan, R., Jasmin, B. J., Marini, J. F., and Leger, J. J. (2009) Large-scale mRNA analysis of female skeletal muscles during 60 days of bed rest with and without exercise or dietary protein supplementation as countermeasures. *Physiol. Genomics* **38,** 291–302
56. Ramskold, D., Wang, E. T., Burge, C. B., and Sandberg, R. (2009) An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. *PLoS Comput. Biol.* **5,** e1000598
57. Lundby, C., Nordsborg, N., Kusuhara, K., Kristensen, K. M., Neufer, P. D., and Pilegaard, H. (2005) Gene expression in human skeletal muscle: alternative normalization method and effect of repeated biopsies. *Eur. J. Appl. Physiol.* **95,** 351–360
58. Vissing, K., Andersen, J. L., and Schjerling, P. (2005) Are exercise-induced genes induced by exercise? *FASEB J.* **19,** 94–96
59. Guerra, B., Gomez-Cabrera, M. C., Ponce-Gonzalez, J. G., Martinez-Bello, V. E., Guadalupe-Grau, A., Santana, A., Sebastia, V., Vina, J., and Calbet, J. A. (2011) Repeated muscle biopsies through a single skin incision do not elicit muscle signaling, but IL-6 mRNA and STAT3 phosphorylation increase in injured muscle. *J. Appl. Physiol. (1985)* **110,** 1708–1715
60. Staron, R. S., Hagerman, F. C., Hikida, R. S., Murray, T. F., Hostler, D. P., Crill, M. T., Ragg, K. E., and Toma, K. (2000) Fiber type composition of the vastus lateralis muscle of young men and women. *J. Histochem. Cytochem.* **48,** 623–629
61. Wang, M., Wang, X. C., Zhang, Z. Y., Mou, B., and Hu, R. M. (2010) Impaired mitochondrial oxidative phosphorylation in multiple insulin-sensitive tissues of humans with type 2 diabetes mellitus. *J. Int. Med. Res.* **38,** 769–781
62. Thompson, J. R., Swanson, S. A., Casale, G. P., Johanning, J. M., Papoutsi, E., Koutakis, P., Miserlis, D., Zhu, Z., and Pipinos, I. I. (2013) Gastrocnemius mitochondrial respiration: are there any differences between men and women? *J. Surg. Res.* **185,** 206–211
63. Costill, D. L., Daniels, J., Evans, W., Fink, W., Krahenbuhl, G., and Saltin, B. (1976) Skeletal muscle enzymes and fiber composition in male and female track athletes. *J. Appl. Physiol.* **40,** 149–154
64. Seth, A., Steel, J. H., Nichol, D., Pocock, V., Kumaran, M. K., Fritah, A., Mobberley, M., Ryder, T. A., Rowlerson, A., Scott, J., Poutanen, M., White, R., and Parker, M. (2007) The transcriptional corepressor RIP140 regulates oxidative metabolism in skeletal muscle. *Cell Metab.* **6,** 236–245
65. Mustonen, T., and Alitalo, K. (1995) Endothelial receptor tyrosine kinases involved in angiogenesis. *J. Cell Biol.* **129,** 895–898
66. Kaipainen, A., Korhonen, J., Pajusola, K., Aprelikova, O., Persico, M. G., Terman, B. I., and Alitalo, K. (1993) The related FLT4, FLT1, and KDR receptor tyrosine kinases show distinct expression patterns in human fetal endothelial cells. *J. Exp. Med.* **178,** 2077–2088
67. Kiens, B., Roepstorff, C., Glatz, J. F., Bonen, A., Schjerling, P., Knudsen, J., and Nielsen, J. N. (2004) Lipid-binding proteins and lipoprotein lipase activity in human skeletal muscle: influence of physical activity and gender. *J. Appl. Physiol. (1985)* **97,** 1209–1218
68. Roepstorff, C., Donsmark, M., Thiele, M., Vistisen, B., Stewart, G., Vissing, K., Schjerling, P., Hardie, D. G., Galbo, H., and Kiens, B. (2006) Sex differences in hormone-sensitive lipase expression, activity, and phosphorylation in skeletal muscle at rest and during exercise. *Am. J. Physiol. Endocrinol. Metab.* **291,** E1106–E1114

Received for publication April 11, 2014.
Accepted for publication June 23, 2014.