Distinctive morphological and gene/protein expression signatures during myogenesis in novel cell lines from extraocular and hindlimb muscle

John D. Porter, Sheri Israel, Bendi Gong, Anita P. Merriam, Jason Feuerman, Sangeeta Khanna, and Henry J. Kaminski

Departments of Neurology and Neurosciences, Case Western Reserve University and University Hospitals of Cleveland, Cleveland, Ohio

Submitted 7 October 2004; accepted in final form 10 November 2005

Myofibers, in turn, are the functional units of individual skeletal muscles. Myofiber type content directly translates into the contraction speed and fatigue resistance of a given skeletal muscle. However, the properties of some muscle groups are not accounted for by simple variations in content of the four conserved myofiber types. Instead, a concept of distinct skeletal muscle classes, or allotypes, has emerged, initially based on expression of novel myosin heavy chain isoforms; masticatory and extraocular muscle (EOM) allotypes express tissue-specific myosins not seen in the limb allotype (36, 37). While the potential for an allotype to express novel myosins is intrinsic to its myoblast lineage, the actual expression of tissue-specific myosin isoforms is dependent on extrinsic factors such as innervation/activation patterns (9, 11, 36). Recent studies suggest that intra- and interallotype differences are not limited to myosin, nor are they a simple function of the proportional representation of stereotypic fiber types (14, 18, 29, 44, 63, 66, 71). To date, the overall breadth, regulatory mechanisms, and significance of muscle allotypes are poorly understood.

The oversimplification that skeletal muscle is generic is an impediment to recognition that muscle group-specific traits may determine outcomes in inherited or acquired disease. Patterned muscle group involvement, usually not accounted for by existing disease models, is found in a variety of neuromuscular diseases (e.g., limb girdle muscular dystrophies, facioscapulohumeral muscular dystrophy, and distal, desmin-related, metabolic, and hereditary inclusion body myopathies) (reviewed in Refs. 27, 93). The novel properties and disease responsiveness of craniofacial muscles provide insights into the breadth and significance of skeletal muscle diversity. EOM, in particular, does not fit the established myofiber classification schemes and other “rules” that govern most skeletal muscles (62, 63, 66, 68). EOM expression profiles during development and in the adult are highly divergent from those of hindlimb (17, 18, 29, 66). Such novel baseline properties of EOM may, in turn, precondition its sparing in the dystrophin-glycoprotein complex-based muscular dystrophies (40, 42, 43, 45, 61, 65, 69, 72–74, 94) and heightened sensitivity to disorders such as myasthenia gravis (41, 66). Because neuromuscular diseases are often not fully penetrant, and targeted muscle groups can vary, it may be impossible to understand disease mechanisms without an in-depth knowledge of muscle allotypes.

Mechanisms behind muscle fiber, group, and allotype diversity are complex and incompletely understood (see Ref. 82). While immortalized cell lines have proven invaluable in dissecting myogenic mechanisms, the existing lines (e.g., C2C12 and L6) were derived from the limb allotype and cannot address the concept of allotype diversity. Here, we derived cell lines...
lines from craniofacial and hindlimb muscles as novel tools for studies of myoblast-autonomous mechanisms underlying allotype diversity. Our morphological, genomic, and proteomic data support the concept that myoblast lineages are fundamental determinants of the differentiated skeletal muscle allotypes.

MATERIALS AND METHODS

Cell line derivation. EOM- and hindlimb muscle-specific cell lines were derived from neonatal (4–6 days old) C57BL/6J mice. Gastrocnemius was chosen as representative of the limb allotype, since it has been well studied, and, like EOM, it is a predominately fast-twitch muscle in vivo. Pups were killed by decapitation, and EOM and gastrocnemius muscles were isolated, digested with collagenase, filtered, and centrifuged; cells were plated on 0.5% gelatin in F10C media (F10 plus 1% penicillin-streptomycin, 1 mM l-glutamine, and 1.2 mM CaCl2) supplemented with 15% horse serum- and fungizone-6 ng/ml fibroblast growth factor (FGF; Sigma, St. Louis, MO) at 5% CO2 and 37°C. Fibroblasts were eliminated through use of high FGF and repeated, selective absorption by plating to noncoated plastic. Cells were expanded and passed until spontaneous transformation occurred, ~30 days later (7–12 passages) and were frozen in F10C containing 15% horse serum and 7.5% DMSO and stored on liquid nitrogen. EOM and hindlimb cell lines were designated as mEOM and mLM, respectively. Animal use was approved by the Institutional Animal Care and Use Committee at Case Western Reserve University.

Myoblast propagation and fusion. Frozen cells were thawed and expanded using growth media (F10C plus 15% horse serum and 2% FGF). Cells then were plated onto 0.5% gelatin-coated plates at a density of 1.3 × 106 cells/150-mm plate in growth media without FGF. Myotube formation was induced 24 h later (0 h) by switching to differentiation media (F10C plus 2% horse serum) for 24 h. Cells were lysed between 0 and 48 h after induction, and RNA was harvested for DNA microarray and real-time quantitative PCR (qPCR) or protein isolation for immunoblotting.

Histology and immunocytochemistry. Myogenesis in the mEOM and mLM lines was documented by phase contrast microscopy. For immunocytochemistry, cells were cultured in growth and then differentiation media in chamber slides at a density of 5,700 cells/cm2, following conditions described above. Cells were permeabilized and immunostained using desmin (1:100; Vector Labs, Burlingame, CA), myogenin (1:50; Santa Cruz Biotechnology, Santa Cruz, CA), or embryonic myosin (1:25; Vector Labs) primary antibodies, and immunoreactivity was visualized with Alexa 488-conjugated goat anti-mouse secondary antibody (1:400; Molecular Probes, Eugene, OR).

DNA microarray. DNA microarray analysis was run as three independent triplicates/time point/cell line, combining multiple plates (n = 2–4) for each replicate at 0, 4, 12, 24, and 48 h after induction of myoblast fusion. Total RNA was extracted from lysates using TRIzol reagent (Gibco BRL, Rockville, MD). RNA pellets were resuspended at 1 μg RNA/μl DEPC-treated water, and 8 μg were used in a reverse transcription reaction (SuperScript II; Life Technologies, Rockville, MD) to generate first-strand cDNA. Double-strand cDNA was synthesized and used in an in vitro transcription (IVT) reaction to generate biotinylated cRNA. Fragmented cRNA (15 μg) was used in a 300-μl hybridization cocktail containing herring sperm DNA and BSA as carrier molecules, spiked IVT controls, and buffering agents. A 200-μl aliquot of this cocktail was used for hybridization to Affymetrix mouse 430A and -B (Santa Clara, CA) microarrays for 16 h at 45°C. The Affymetrix 430A/B array set includes 45,137 probe sets representing >39,000 unique transcripts. The manufacturer’s standard posthybridization wash, double-stain, and scanning protocols used an Affymetrix GeneChip Fluidics Station 400 and a Hewlett Packard Gene Array scanner.

Microarray data analysis. Raw data from microarray scans were analyzed with Affymetrix Microarray Suite (MAS) 5.0. MAS evaluates sets of perfect-match (PM) and mismatch (MM) probe sequences to obtain both hybridization signal values and present/absent calls for each transcript. The raw data series and .CEL files were posted on the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) under series record accession numbers GSE1049 (mEOM) and GSE1052 (mLM). Transcripts absent from all samples were excluded from further analysis. Microarrays were scaled to the same target intensity (MAS 5.0 scaling factor = 500), and pairwise comparisons were made between mEOM and mLM samples. Transcripts defined as differentially regulated met the criteria of 1) consistent increase/decrease call across all replicate comparisons at a given time point, based on Wilcoxon’s signed rank test (algorithm assesses probe pair saturation, calculates a P value, and determines increase, decrease, or no change calls) and 2) absolute value of the average fold difference ≥2.0.

In addition, microarray data from mEOM and mLM were normalized and analyzed using the Robust Multichip Average (RMA) algorithm (39) in ArrayAssist 2.0 (Iobion Informatics, La Jolla, CA). RMA executes a background adjustment, a quantile normalization, and a summation of individual probe set intensities using a log-scale linear additive model for log-transformed, background-corrected/normalized PM probe intensities. For each cell line, we set the 0-h time point as baseline, compared all other time points (n = 3/time point) to the 0-h data, and then averaged to generate fold difference values; the fold difference cutoff for differentially expressed transcripts was set at 2.0.

The Cluster Analysis of Gene Expression Dynamics (CAGED) 1.1 algorithm (http://genomethods.org/caged/) (75) was used to compare temporal series data from the two cell lines. CAGED uses Bayesian Clustering by Dynamics to identify a statistical model of the most probable set of clusters of transcripts in a time series, without relying on any predefined similarity threshold. After elimination of probes found to be absent from all arrays of a cell line, background-corrected/normalized probe signal output from RMA was averaged for each transcript/postinduction stage/cell line and then used as input to CAGED. To plot the mEOM and mLM data together (which allowed comparison of similarities or differences in temporal pattern for any given gene), data from each line were identity coded and then combined before CAGED analysis. To filter for differentially expressed transcripts, we required a fold difference of greater than or equal to two between any two points in the temporal series. The following parameters were used with CAGED: autoregressive model with a model order of 3, prior precision equal to 1, gamma value of 0, Bayes factor of 1, distance set to Euclidean, and logarithmic transformation.

Significance analysis of microarray (SAM) (90) was used to rank order transcripts by importance in distinguishing myogenesis in the mEOM and mLM lines. SAM provides a false discovery rate analysis; a fourfold difference cutoff was used for SAM results to identify genes that are “signatures” of the two lines. For SAM, data from all time points were combined to yield greater power to analysis of differential expression in the two cell lines.

Affymetrix transcript annotations were replaced with official gene nomenclature, and functions were assigned using information in NCBI Entrez Gene, UniGene, and PubMed and NetAffx and Weizmann Institute of Science GeneCards databases.

qPCR. Select transcripts were reanalyzed by qPCR, using the same samples as in the microarray studies. Transcript-specific primers (Supplemental Table S1; available at the Physiological Genomics web site)1 were designed using Primer Express software [Applied Biosystems (ABI), Foster City, CA], and specificity was confirmed by NCBI Basic Local Alignment Search Tool (BLAST). Reverse transcription

1 The Supplemental Material for this article (Supplemental Tables S1–S6) is available online at http://physiolgenomics.physiology.org/cgi/content/full/00234.2004/DC1.
was carried out on 2 μg of total RNA with ABI TaqMan reverse transcription reagent, qPCR used SYBR Green PCR core reagent in a 25-μl volume, with an ABI PRISM 7000 Sequence Detection System. Mouse GAPDH (forward primer: 5-AACGACCCCTTCATTGAC-3; reverse: 5-TCCACGACATACTCAGCAC-3) was used as an internal positive loading control. Fold change values represent averages from triplicate measurements, using the \(2^{-\Delta\Delta CT}\) method (35).

Proteomics—immunoblot analyses. Comparison of mEOM and mLM signaling protein profiles at the 0- and 48-h postinduction time points used the BD Biosciences Pharmingen PowerBlot system (San Diego, CA), which screens \~1,000 proteins in triplicate by Western blot. Tissue preparation followed the manufacturer’s protocol. Cultures were washed and then homogenized in RIPA lysis buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 0.5% deoxycholic acid, and 1% NP-40 with a cocktail of protease inhibitors: 20 μg/ml aprotinin, 20 μg/ml leupeptin, 10 μg/ml PMSF, 1 mM sodium orthovanadate, 10 mM sodium pyrophosphate, and 10 mM sodium fluoride), using a polytron. Samples were frozen and shipped to BD Biosciences for large-scale Western analysis, imaging, and densitometry. Only protein differences that fell within stringent acceptance criteria (BD Biosciences levels 8–10: fold difference \(\geq 1.5\) in 9 of 9 comparisons) were identified as differentially expressed.

RESULTS

 Fundamental differences exist between skeletal muscle allotypes and possibly among the muscles comprising a single allotype. There is, however, little understanding of cell and allotypes and possibly among the muscles comprising a single allotype. Here, we sought to non-cell-autonomous mechanisms that regulate the development of muscle allotypes. There is, however, little understanding of cell and allotypes and possibly among the muscles comprising a single allotype.

Myogenesis in mEOM and mLM cell lines. Cell lines were derived from primary cultures of neonatal EOM (mEOM) and gastrocnemius (mLM). Myoblasts of the mEOM and mLM lines were round, but mEOM myoblasts were smaller than those of the mLM line (mean diameter of 0-h mEOM myoblasts was \~20% less than that of comparable-stage mLM myoblasts) (Fig. 1, A and B). Myoblast elongation occurred within 4 h of induction by switching to differentiation media (Fig. 1, C and D). Myoblast fusion, to form multinucleate myotubes, followed a very similar time course in the two cell lines during the 48-h postinduction period studied here (Fig. 1, E–J). Myotubes from the mEOM line were consistently longer and thinner than those from mLM (mean diameter of 48-h mEOM myotubes was \~40% less than that of mLM myotubes of the same stage). Some myotubes of both lines became spontaneously contractile within 48 h of induction of fusion.

Cell line myogenic potential was evaluated using several established markers. Myonuclei of both lines were immunoreactive for the myogenic regulatory factor myogenin (Fig. 2). Myoblast and myotube cytoplasm was immunoreactive for the muscle-specific marker desmin (>90% of cells at 4 h) (Fig. 2). Likewise, embryonic myosin heavy chain (Myh3) protein was detected in both lines by immunoblotting and immunocytochemistry (data not shown).

DNA microarray analysis of myogenesis in the mEOM and mLM lines: quality control. Gene expression patterns were evaluated for 15 independent samples per cell line (3 replicates \times 5 time points/line), using Affymetrix 430A and -B arrays (30 arrays/cell line \times 2 lines = 60 total arrays). Data from the mEOM and mLM lines showed expression (i.e., a present or marginal call from the MAS algorithm) of 57.7 ± 1.7% (mean ± SD, \(n = 15\)) and 58.4 ± 2.0%, respectively, of the probes represented on the Affymetrix 430A arrays and
25.7 ± 1.4% and 30.8 ± 1.8%, respectively, of probes represented on the 430B arrays. This high level of consistency in percent present per array is a positive quality control measure (33, 67). Detection of GAPDH was also consistent with integrity of biotin-labeled cRNA target; ratios of 3’ to 5’ GAPDH probe signal ranged from 0.75 to 0.89 for the 30 mEOM arrays and from 0.72 to 0.86 for the 30 mL M arrays (3’-to-5’ ratios below 1 are consistent with unfragmented, fully in vitro transcribed cRNA). Collectively, the percent present and 3’-to-5’ ratio controls are excellent cumulative error measures, reflecting on all phases of microarray study conduct, and the values obtained here are positive quality control indicators.

**Overview of gene and protein expression signatures of extraocular- and hindlimb muscle-derived myoblasts.** We first used DNA microarray to compare transcripts expressed by the mEOM and mL M lines at the 0-h time point (preinduction myoblasts). An “expressed” designation required a present/marginal call from three of three replicate arrays per cell line, resulting in 16,199 expressed Affymetrix probes for mEOM and 17,817 expressed probes for mL M. The two lines showed 82.4% overlap (Fig. 3A). At the differentiating myotube stage, 48 h postinduction, there was 82.3% overlap in expressed probes (Fig. 3A). While this initial analysis does not address expression levels, these data suggest that expressed gene identities are highly conserved among myoblasts and myotubes of the two allotypes.

Transcripts that were present only in one myoblast line (i.e., in the nonoverlap regions of Fig. 3A), regardless of expression level, represent a first approximation of allotype-specific patterns. Transcripts meeting this criterion were further filtered for expression level differences using the SAM algorithm (all time points were combined for mEOM and mL M and then compared by SAM to increase the statistical power of the analysis; 4-fold cutoff). This approach identified those genes that best distinguished myoblasts of the two lines, defined here as signature genes (Table 1 and Supplemental Table S2). Twenty-two probes [19 unique genes and expressed sequence tags (ESTs)] passed the SAM filter as mEOM signature genes, including genes implicated in transcription/signaling (Dlx1, Ere, Exs1, Prkcb, Shox2, and Twist1). Ninety-eight probes (80 unique genes and ESTs) met criteria as signature genes for mL M, also including multiple transcriptional/signaling regulators (Fzd6, Hoxa1, Hoxa11, Hoxc10, Lbx1, Sfrp2, Sim2, and Wnt5a). Genes and proteins designated as mEOM or mL M signatures are underlined throughout the remainder of the results section.

**Table 1. Signature genes for mEOM and mL M myoblasts**

<table>
<thead>
<tr>
<th>Exclusive to mEOM</th>
<th>Exclusive to mL M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nid1, Ef2a3y, Il12rb1, Twist1, Serpinb2, Ere, Exs1, Kcnq4, Gsta2, Mrc2, 4631426J05Rik, Myh8, Sla38a4, Shox2, BG228852, Dlx1, Blrvb, Ddx3y, Prkcb</td>
<td>Cd24a, Ccnf2, Fzd6, Fgf13, Cepba, Timp3, Sim2, Decr, Hoxa11, Hoxa1, Tmlhe, Myf5, Rgs5, Ptgser4, Gpc4, Gzm, Elav12, Sh3bgr, Rims2, Nsg1, Pbx1, Asa4, Zfip503, Cdh10, Kir3, Hoxc13, Pdgfd, Dnr2, Xst, Lhx5, Cal5, B021831, 2310040A07Rik, Pdgfrl, 261002901Rik, Rnf130, 2210407G14Rik, Hoxa10, 8430414N03Rik, Dusp9, Echl, Aoi, Rin3, 290006419Rik, A1542195, 2810401A20Rik, BC036718, D10Ertd802, 5530060H04Rik, Slc35f1, Wnt5a, Fox, A930023L17Rik, BB600804, Pps11, A1647591, Hoxc10, Selp, BB660130, Sympo2, 9130013K24Rik, Mtap4, AW214292, Sfrp2, Elf1, Cdkn2b, Fads3, Act5, Car3, Cdkn2a, Rat2, 2310040A04Rik, 4631426J19Rik, 2700006G24Rik, Kif2, Hoxa9, A030013D21, BG068705, Pld2, 6030410K14Rik</td>
</tr>
</tbody>
</table>

All transcripts listed here were called present by the MAS algorithm in 3 of 3 independent replicate analyses of 0-h myoblasts for only 1 of the 2 cell lines and 2) passed a false discovery rate filter and exhibited ≥4-fold difference in expression level the SAM algorithm. Transcripts are listed in order from highest to lowest fold difference value. We define transcripts meeting these highly stringent selection criteria as signature genes for the 2 myoblast populations. Those transcripts shown in bold also passed filtering by both the MAS and RMA algorithms for consistent expression level differences (see MATERIALS AND METHODS for criteria). The full gene list from MAS/RMA analysis is in Supplemental Table S4. MAS, Microarray Suite; RMA, Robust Multichip Average; SAM, Significance Analysis of Microarray.
The 0-h microarray data also were evaluated using the MAS and RMA algorithms to identify all genes with consistent differences in expression level. Transcripts were defined as differentially expressed only if they passed filtering by both algorithms, yielding 466 probes, 86 with higher expression in mEOM and 380 in mLM (Supplemental Table S3). These represent 393 unique transcripts (74 for mEOM and 319 for mLM), including 234 known genes (66 for mEOM and 168 for mLM) and 159 ESTs. Seventy-six percent of the signature genes identified using the present call/SAM analysis were mLM and 159 ESTs. Seventy-six percent of the signature genes were identified as differentially expressed in mEOM and mLM (Supplemental Table S3). These data were supported by the identification of several differentially expressed proteins functional in hypertrophy, cell adhesion, and cytoskeletal or sarcomeric structure and regulation (mLM > mEOM: Actn3, Ctgf, Gmnn, Gsn, Nrap, Pak3, Pdlim, Tub3, and Vav3; mEOM > mLM: Cdt3 and Emb), and cytoskeletal or sarcomeric structure and regulation (mLM > mEOM: Actn3, Ctgf, Gmnn, Gsn, Nrap, Pak3, Pdlim, Tub3, and Vav3; mEOM > mLM: Cdt3 and Emb) (Supplemental Table S3), all of which can influence cell morphology. These data were supported by the identification of several differentially expressed proteins functional in hypertrophy, cell adhesion, and cytoskeletal organization (mLM > mEOM: ARP3, caveolin 2, dystrobrevin, GSK-3β, integrin β1, PKB/ AKT, and RAP2; mEOM > mLM: p120 catenin, MAP2B, mena, and ninjurin) (Supplemental Table S4). Compared with mEOM, mLM myoblasts were enriched in several genes known to function in limb development (Efh1, Hoxc10, Igfbp2, Kremen, Lhx1h, Popdc3, Wnt5a, and Wnt9a).

Because myoblast lineage-specific transcriptional/signaling regulators represent a putative mechanism behind muscle allop type divergence, we also focused on genes in these categories. mLM myoblasts preferentially expressed numerous genes that previously were implicated in skeletal muscle development (Dill1, Dmnt2, Dscrl11, Fhl1, Lhx1h, Musk, Myf5, Ndn, Notch3, Pterd, Sin2, Smad1, Snai2, Wnt5a, and Xdcar). By contrast, mEOM myoblasts were enriched in other transcriptional/signaling genes linked to myogenesis (Ahr, Arx, Dusp6, Igfbp7, Nid1, Pdgfrb, Prkcb, Sec8a33, Shox2, and Twist1) and some transcription factors with no known role in muscle development (Esxl and Eg2).

Myogenesis is regulated by a variety of interactive signal transduction pathways, some of which exhibited differential gene/protein regulation in mEOM and mLM myoblasts. The collective microarray and PowerBlot data established a pattern of enhanced expression of canonical MAP kinase (mRNA: Dusp1, Dusp9, Eqr13, Pdgfd, Pdgfrb, Prkcm, Rapgef2, and Rgs4; protein: MEK1, MEK2, FGFR1, and SOS) and JNK/p28 MAP kinase (mRNA: Jun, Dusp1, and Dusp9; protein: Pdgfrb, Pdgfrb, and Pdgfrb).

Table 2. Signature proteins for mEOM and mLM

<table>
<thead>
<tr>
<th>Exclusive to mEOM myoblasts</th>
<th>Exclusive to mLM myoblasts</th>
<th>Exclusive to mEOM myotubes</th>
<th>Exclusive to mLM myotubes</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-Chl (Cblc), clathrin heavy chain (Clec), GS27 (Gosr2), headpin (Serpinb13), p120 catenin (Catsn), Sam68 (Khdrbs1), synapsin I (Syn1)</td>
<td>Bp/GpR78 (Hspa5), connexin-43 (Gja1), Tat-SF1, TF2 (Neod2), TLS (Fus), CR1K (Citi), doublecortin (Dxc), G3BP (G3bp), Nogo-A (Rn4), Npat (Pnp), P53 (Mapk14), P55Cdc (Cdc20), Sec8 (Sec811), SOS (Sos2)</td>
<td>GS27 (Gosr2), headpin (Serpinb13), KIF3B (Kif3b), p43/EMAP II precursor (Scey), p120 catenin (Catsn), synapsin I (Syn1), c-Chl (Cblc), clathrin heavy chain (Clec)</td>
<td>B56a (Ppp2r5a), connexin-43 (Gja1), ERK1/2 (Mapk1/Mapk2), Tat-SF1, TF2 (Neod2), TLS (Fus), CR1K (Citi), doublecortin (Dxc), G3BP (G3bp), mEPHX (Ephp1), Npat (Pnp), P53 (Mapk14), P55Cdc (Cdc20), pClan (Clnsla), Sec8 (Sec811), SOS (Sos2)</td>
</tr>
</tbody>
</table>

All proteins listed here were highly enriched in the 0-h data in head-to-head comparisons (band was either absent for 1 line or there was a >10-fold difference in densitometric signal) using PowerBlot. Proteins listed under myoblasts were detected in comparisons of 0-h myoblasts; proteins listed under myotubes were detected in 48-h myotube comparisons. Corresponding gene symbol, when known, is indicated parenthetically. Proteins indicated in bold were differentially expressed in both the myoblast and the myotube comparisons. The full gene list of differentially expressed proteins is in Supplemental Table S5.
Overview of gene and protein expression signatures during myogenesis of the mEOM and mLM lines.

<table>
<thead>
<tr>
<th>Exclusive to mEOM</th>
<th>Exclusive to mLM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tslx, Hoxa9, Cd24a, Wnt5a, Ndn, Cdkn2b, Hoxc13</td>
<td>Sks2a, Atp5a1, Ndn, Bmp15, Nid1, Slc35fl, Gria3, Pparg</td>
</tr>
</tbody>
</table>

The cumulative expression patterns of mEOM and mLM lines across the 5 time points studied were compared here. All transcripts listed passed a false discovery rate filter and exhibited ≥4-fold difference when all time points were combined and data subjected to the SAM algorithm. Transcripts are listed in order from highest to lowest fold difference value. We define transcripts meeting these highly stringent selection criteria as signature genes for early myogenesis in the 2 cell lines. Transcripts also meeting signature gene criteria for the 0-h myoblast populations (Table 1) are indicated in bold.

Table 3. Signature genes for myogenesis in mEOM and mLM

<table>
<thead>
<tr>
<th>Exclusion Criteria</th>
<th>Gene Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exclusive to mEOM</td>
<td>Sks2a, Atp5a1, Ndn, Bmp15, Nid1, Slc35fl, Gria3, Pparg</td>
</tr>
<tr>
<td>Exclusive to mLM</td>
<td>Sks2a, Atp5a1, Ndn, Bmp15, Nid1, Slc35fl, Gria3, Pparg</td>
</tr>
</tbody>
</table>

Table 3. Signature genes for myogenesis in mEOM and mLM

- **Table 3. Signature genes for myogenesis in mEOM and mLM**

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sks2a, Atp5a1, Ndn, Bmp15, Nid1, Slc35fl, Gria3, Pparg</td>
<td>Exclusive to mEOM</td>
</tr>
<tr>
<td>Sks2a, Atp5a1, Ndn, Bmp15, Nid1, Slc35fl, Gria3, Pparg</td>
<td>Exclusive to mLM</td>
</tr>
</tbody>
</table>

Table 3. Signature genes for myogenesis in mEOM and mLM

Differentially expressed Hox family members (Hoxa1 in mLM myoblasts. Integrin-mediated adhesion and signaling differed in constitutive expression level, temporal pattern, or time course series. This approach identified transcripts that differed in constitutive expression level, temporal pattern, or both during the first 48 h of myogenesis (26 unique transcripts with higher expression in mEOM and 145 with higher expression in mLM; Supplemental Table S5). Genes identified in this analysis are defined as signature genes for myogenesis in one of the two lines (Table 3; signature genes are underlined throughout results).

The CAGED algorithm was used to identify all transcripts that met criteria for up- or downregulation during myogenesis (≥2-fold change between any two time points) and cluster them by temporal pattern similarity (Fig. 4 and Supplemental Table S6). To ensure the goodness of fit of the temporal gene expression clustering model generated by CAGED, we checked the autoregressive score and marginal log likelihood values and ensured that these values were minimized by the selected model parameters. We consider such genes to be dynamic participants in myogenesis, making any lineage differences significant for further analysis. By coding and combining the mEOM and mLM data before CAGED analysis, we could directly correlate individual transcript expression patterns across the two lines.

Six hundred fifty-eight Affymetrix probe sets (548 unique transcripts) met CAGED selection criteria and were characterized as mEOM specific (n = 152 probe sets; representing 121 unique genes and ESTs), mLM specific (n = 381; 317 unique genes and ESTs), differentially regulated by both lines with similar temporal patterns (n = 56; 53 unique genes and ESTs), or differentially regulated by both lines but with statistically different temporal patterns (n = 69; 57 unique genes and ESTs) (Supplemental Table S6). Only a minority of transcripts meeting the CAGED algorithm criteria for greater than or equal to twofold change during the first 48 h of myogenesis were shared by the mEOM and mLM lines (20%), and only 10% shared the identical temporal patterns (i.e., were assigned to the same CAGED cluster). Cell signaling and transcriptional regulation comprised the highest percentage of known genes (Fig. 5).

A total of 88 of the 548 differentially regulated transcripts identified by CAGED were validated by similar results from two or more Affymetrix probes for the same transcript. Selected genes differentially regulated during myogenesis in mEOM or mLM were further evaluated by comparing 48- with...
0-h expression levels using qPCR. Of these, 22 of 24 showed correspondence between microarray and qPCR data (Supplementary Table S6).

To assess the differential regulation of a broad group of proteins functioning as transcriptional/signaling regulators in the mEOM and mLM lines, we used PowerBlot to directly compare protein levels in 48-h myotubes from the two lines. Differential expression was identified for 72 unique proteins (Supplemental Table S4); proteins that were highly enriched (either the band was absent for one line or there was a >10-fold difference in band density) in one of the two lines were designated as signature proteins (Table 2 and underlined throughout results).

Patterns among genes and proteins differentially expressed during myogenesis in the mEOM and mLM lines. Several genes previously implicated in limb development met criteria for differential regulation during myogenesis in the mLM line (Igfbp2, Pbx3, Robo1, Sfrp2, Six1, and Wnt9a). Some transcripts related to myofiber morphology showed modulation of expression level during myogenesis in only one of the two cell lines. Hspb8, Klf5, and Prkaca2, genes related to muscle hypertrophy, and Fgf6, linked to myofiber hyperplasia, were dynamically regulated during the first 48 h of myogenesis in the mLM line only. Smox is thought to play a role in skeletal muscle hypertrophy and showed greater induction in mLM than in mEOM; Dsip1 and Igfbp3 function in myofiber size regulation and were differentially regulated in mEOM only. Likewise, higher expression of proteins that are negative regulators of muscle hypertrophy/cell spreading (melusin/Itgb1bp2 and TNIK) in mEOM, and a positive regulator of hypertrophy in mLM (CBFb), correlated with the observed myotube size differences. Cytoskeletal and sarcomemmal structural/regulatory proteins also influence myotube morphology, and several genes (differentially regulated in mEOM only: Myl1, Nef3, and Tagln; in mLM only: Dcx, Gsn, My44, Mtap4, Mybph, and Tn) were identified (Table S6). These regulators of hypertrophy/cell spreading show different expression patterns. The delta/notch (mRNA: Notch3, Ifgfbp3, and Snai2) and Wnt (mRNA: Dkk1, Dll1, Ndr1, Foxc2, Ifgfbp2, Notch3, and Snai1) signaling pathways were differentially regulated between the two lines. Analysis of transcription factors and signaling pathways that are up- or downregulated during in vitro myogenesis may provide important clues as to mechanisms responsible for the divergent fates of EOM and limb myoblasts. Several genes with established roles in early skeletal muscle development (Ankrd2, Arf6, Erbb3, Hist1h1c, Hist1h2bc, Nrk, and Six4) exhibited dynamic regulation during myogenesis in the mLM line only. Other known myogenic regulators were dynamically regulated only in mEOM (Bmp4, Egr1, Igfbp3, Peg3, and Snai2). Some pathways exhibited lineage-specific regulatory patterns. The delta/notch (mRNA: Dikk, Dll1, Dnr1, Foxc2, Ifgfbp, Notch3, Pbx3, and Snai2), and Wnt (mRNA: Camk2d, Ccnd2, Spry2, Wif1, and Wnt9a; protein: ROCK2) signaling cascades, all pathways previously implicated in limb morphogenesis (see Refs. 19, 21, 22, 31), exhibited multiple components with expression levels that were modulated only for mLM myogenesis. By contrast, few to no genes of the retinoic acid (none), Wnt (Wisp1), and delta/notch pathways (none) were modulated in patterns specific to mEOM. The trend toward more active
MAPK signaling that we identified in mLM myoblasts was maintained in differentiating myoblasts (mRNA differentially regulated in mLM only: Dusp6, Fgfl6, Mef2c, Rasal1, and Spryl1; protein, mLM > mEOM: Bnip3, GRP78, ERK1/2, FGF1, Ikkγ/NEMO, JUN, MEK1, p38, and SOS; mRNA differentially regulated in mEOM only: Ddit3, Gadd45a, Hrasls3, Rasff5, and Rad51; protein, mEOM > mLM: PKBa/AKT). TGFβ pathway signaling components showed a heterogeneous response during the first 48 h of myogenesis in the two cell lines (differentially regulated in mLM only: Inhibb, Junb, and Lhbp2; in mEOM only: Bmp4, Fos, Gdf15, Igfbp3, Prss11, and Tieg1; differentially regulated in both: Id1b, Id2b, Id3b, and Tgfβ3; at the protein level, Jun expression was higher in mLM and SMAD4 higher in mEOM).

On the basis of analysis of microarray data with CAGED, the temporal patterns of many muscle-specific transcripts were shared by the mEOM and mLM lines, including muscle development (Id1b, Id3b, and Ncam1), sarcomeric (Actn2, Actn3, Myoz2, Tcap, and Tnni2), energy metabolism (Ckm and Pygm), and ion channel/homeostasis (Atp2a1 and Sln) genes. Other genes characteristic of differentiated skeletal muscle were differentially regulated during the period studied here (mLM only: Atp2a2, Dmd, Dma, Myh1, Ryr1, Sgcg, Tnni1, and Trdn; both mEOM and mLM but with different temporal patterns: Capn3, Cav3, Mb, Myh7, Myh8, Myom1, Myom2, and Tio1). Differentiation in mEOM showed dynamic regulation of smooth muscle- (upregulated: Klf4, Nfil3, Tagln, and Pde2a; down-regulated: Areg) and bone matrix-related (upregulated: Bmp4, Cdh11, Ogn, and Postn; down-regulated: Dmp1) transcripts not modulated during mLM myogenesis.

**DISCUSSION**

The breadth of skeletal muscle heterogeneity, beyond that resulting from obvious variations in content of stereotypic fiber types, has received only sparse attention. Yet, knowledge of muscle diversity may be critically important in understanding the frequent and often unexplained targeting of metabolic and neuromuscular diseases to specific muscle groups (63, 70, 71). Skeletal muscle classes, or allotypes, were initially defined by 10.220.32.247 on August 27, 2017 http://physiolgenomics.physiology.org/ Downloaded from muscle allotypes (diaphragm, EOM, limb, and masticatory). In particular, adult EOM fiber types and genome-wide expression profiles represent fundamental departures from traditional skeletal muscle (66, 83). Although cell lines cannot perfectly mirror in vivo development, they offer extraordinary opportunities for studies that are difficult or impossible in vivo and indeed have facilitated rapid progress in understanding myogenic mechanisms. Given the caveats inherent in any myoblast line, properties of EOM and limb muscle were detected in their derivative lines and help to validate them as appropriate models.

The myoblast lines studied here were derived from neonatal skeletal muscle. The source of our cell lines then were not the embryonic muscle precursor cells (primary myoblasts) that originally gave rise to the two muscle groups, but rather myogenic stem cells and satellite cells that populate developing and mature skeletal muscle. It is then necessary to recognize that derivatives of these later-stage myogenic cells may not fully reflect those properties inherent in the original primary myoblasts. Likewise, the morphological and gene expression profiles established here are likely the result of two key factors: 1) inherent differences in the two cell lines based on their origin and 2) differential response to identical culture conditions. In Hoh’s allotype concept (36, 37), precursor cells with differing developmental potentials (i.e., ability to form distinctive allotypes) do so only under the influence of the “correct” innervation. We would suggest that the parallel for our data is that precursor cells of the EOM and limb allotypes form myotubes of distinctive phenotypes based on fundamental differences in both their origin and their responses to identical environmental factors. Hoh’s use of innervation alone as an allotype-specifying factor then may be too restrictive a concept, as a variety of local factors likely influence the emergence of skeletal muscle allotypes during development.

Myoblast morphology and cytoskeletal organization are direct determinants of myotube and myofiber morphology (3). The mLM myoblasts and myotubes were consistently larger in size than those of the mEOM line, and the two lines exhibited substantive differences in expression of the cytoskeletal and sarcomemmal genes that influence cell morphology. In addition, genes and proteins known to promote myofiber hypertrophy (e.g., Dscr111, Fhl1, CBFB, GSK-3b, and PKBa/AKT; Refs. 49, 59, 77) and several in the cell cycle regulation and MAP kinase pathways showed differential expression patterns consistent with the more rapid growth of mLM myotubes. These morphological/cytoskeletal properties of the mEOM and mLM lines directly correlate with in vivo findings that EOM myofibers are considerably smaller in diameter than those of other skeletal muscles. The percentage size differences seen here directly reflect previously reported differences in adult muscle fibers from these two skeletal muscles (83). Myonuclear density is known to relate to myofiber size, but potential differences from prototypical limb muscle in this trait have not yet been determined for EOM. Collectively, data suggest that at least some traits that distinguish the adult muscle allotypes are present in our cell lines.

Several transcription factors and cell signaling genes that are enriched in mLM myoblasts and myotubes have been previously reported in developing or adult limb. Among these, Lbx1h plays an essential role in limb muscle development, and dysgenesis is seen in Lbx1h knockout mice (7, 32, 52, 79). While Lbx1h is expressed in developing chick EOM (5, 53), it
is restricted to only one of the six muscles (i.e., a minority source of myoblasts for the mEOM line) and has not been reported in mammalian EOM. Likewise, Ebfl, Hoxc10, Igbp2, Kremen, Poppd3, Robo1, Six1, and Wnt5a are all established limb development markers (2, 6, 15, 47, 50, 51, 54, 92) that are either expressed at higher levels or differentially regulated only in mLM. Collectively, these expression patterns, and the other patterns in signaling pathways (see below), help validate the mLM line as a model for limb muscle development.

The current level of knowledge of gene expression patterns specific to EOM development provides only a limited basis for direct validation of the mEOM model. Of the transcriptional and signaling genes preferentially expressed by mEOM myoblasts, Arx and Shox2 share a homeodomain motif with other genes involved in craniofacial development (4) and thus are excellent candidates for an important role in EOM. Although Twist1 is widely expressed in muscle development, its enhanced expression during mEOM myogenesis seen here is consistent with an established role in craniofacial muscle development (81) and may explain the association of Twist1 mutations with the inherited EOM disorders of ptosis and exophthalmos (23). Similarly, Arx is closely related to Phox2a/Arx, the disease gene for congenital fibrosis of EOM type 2 (55). Two genes previously linked to EOM development (Pitx2 and Myh13) were not differentially expressed in our data. Reported expression of Pitx2, one of the very few transcription factors known to have an essential role in EOM development (17, 30, 46), in the developing limb (79) might have precluded its detection in our differential expression assays. The EOM-specific myosin (Myh13) was not differentially expressed in the mEOM line; but this finding would be expected, since we previously reported that Myh13 upregulation occurs late in myogenesis (10, 78), requires appropriate neural activation patterns (9, 11), and is not seen even in organotypic cultures of EOM primordia with oculomotor motoneurons (64).

While many features that distinguish the adult allototype are unlikely to emerge during the first 48 h of in vitro myogenesis, several genes and proteins (Atp1b1, Bgn, Car3, Daf1, Islr, Jun, Pdl1m, Pmp22, Rpb1, Sh3bg1, and Xlr; acetylcholine receptor-α, HIC-5/Tgfb1i, MnSOD, ROCK-II/Rokα, and UBCH6) with reported differential expression in EOM vs. hindlimb muscle (44, 66) show the same differences in our cell lines. The differential regulation of Bgn and Daf1, seen both in vivo and in vitro, may play important roles in the novel responses of EOM to neuromuscular disease. Finally, expression patterns of specific transcripts and, more importantly, developmental regulatory pathways (see below) are consistent with the mEOM line being representative of a craniofacial muscle phenotype.

**Signature genes and proteins for mEOM and mLm lines.** Our data show substantial conservation (~85%) of a basic skeletal muscle transcriptome and proteome (expressed genes and proteins) among the mEOM and mLm lines. This finding directly reflects the relatively few allotype-exclusive genes found for adult EOM and hindlimb (17, 18, 29, 44, 66). However, we observed clear allotype-specific patterns during early myogenesis, including both genes that are lineage exclusive and ubiquitous genes that show substantial differences in expression levels. Such myoblast lineage-specific patterns are important, as they may serve to identify putative mechanisms behind the unique functional properties of adult EOM and limb muscle allotypes.

Attempts to identify genes that are signatures of a specific cell type must rely on stringent criteria that, admittedly, require a somewhat arbitrary application of thresholds. Although small differences in expression levels may be critical for transcription factors, while large differences may be insignificant for non-rate-limiting metabolic enzymes, thresholding does identify a subset of highly differentially expressed transcripts that have value for understanding mechanisms and driving further studies. Here, we used several bioinformatics approaches to identify transcripts that were enriched in only one of the two myoblast populations and/or were dynamically regulated during myogenesis in one line only. To identify genes specific to mEOM or mLm myoblasts, we intersected the results from two analytic strategies (present call/SAM analysis with MAS/RMA fold difference analysis); genes recognized by both strategies are shown in Table 1. To identify genes with consistent differences during the 48-h course of myogenesis studied here, we relied on present call/SAM analysis, as this recognizes genes that are at constitutively different levels throughout myogenesis as well as those that are dynamically up- or downregulated in only one of the cell lines (Table 3). This list is, in part, supported by data from CAGED, which determines only the dynamically regulated transcripts. Differentially expressed proteins were those detected in only one of the myoblast or myotube populations (Table 2). Genes and proteins meeting the signature designation are attractive candidates for subsequent in vitro and in vivo deletion studies designed to understand the developmental basis of the EOM and limb muscle allotypes.

**Utilization of developmental regulatory mechanisms in mEOM vs. mLm.** The combined actions of multiple transcription factors and signal transduction pathways are responsible for specification of skeletal, cardiac, and smooth muscle diversity (28). Prior studies demonstrated spatiotemporal heterogeneity in myogenic mechanisms (24, 25, 56, 76, 84) and have shown that lineage-specific cis-regulatory mechanisms drive myogenic regulatory factor expression among muscle groups differing in rostrocaudal position (58, 86). Moreover, several modifications of established myogenic mechanisms are known to exist for craniofacial muscles (5, 16, 34, 48, 53, 86, 87, 89, 91). Our data extend these findings by showing that, given the identical environmental conditions, myoblasts derived from the EOM and limb allotypes make differential use of key developmental regulatory pathways. Specifically, myoblast populations retain some spatially appropriate signaling mechanisms (e.g., Hox and Wnt signaling in mLm > mEOM) while exhibiting surprisingly different patterns for other pathways (e.g., MAP kinase).

Hox family genes are expressed in spatiotemporal patterns suggestive of a role in myogenesis (57), although their precise role is undefined. All of the Hox genes differentially expressed here (Hoxa1, Hoxa9, Hoxa10, Hoxa11, Hoxc10, and Hoxc13) were signature genes for the mLm line. Each is a member of the HoxA or HoxC, but not HoxB or HoxD, clusters, a finding that precisely corresponds to Hox expression patterns in the developing limb (38). Moreover, Hox expression patterns in mLm, a hindlimb derivative, are consistent with Hox expression patterns in the developing limb (38).
regulator of Lbx1h (1), which also is restricted to the mLM lineage and critical to limb muscle development.

Wnt signaling pathways play multiple roles in general limb development, particularly in the positive regulation of myogenesis (2, 19, 20, 26), but are not recruited to the recapitulation of myogenic events during muscle regeneration (95). By contrast, canonical Wnt signaling has the opposite effect of repressing craniofacial myogenesis (91). Consistent with the disparate effects of Wnt on EOM and limb myogenesis, the gene and protein expression patterns seen here suggest that Wnt pathways are more active during myogenesis in mLM, where they are positive regulators, than in mEOM, where they are negative regulators. Several Wnt pathway transcripts are signature genes for mLM (Cend2, Dkk2, Prkcm, Sfrp2, and Wnt5a). Enhanced expression of the promyogenic Wnt5a, upregulation of the Wnt5a signal transducer Camk2d, and downregulation of the Wnt inhibitor Sfrp2 were seen only in the mLM line, thereby supporting the lineage identities and, at least in part, the conservation of in vivo differences in myogenic programs in our cell lines. Negative regulators of Wnt signaling are necessary for myogenesis in embryonic craniofacial muscle primordia due to Wnt expression by surrounding tissues. We did not observe expression of Wnt inhibitors like Frzb in mEOM (although mEOM shows enhanced expression of another negative regulator of canonical Wnt signaling, Dab2), possibly because they too are expressed by nonmuscle tissues associated with developing EOM (91).

In addition, our data show myoblast lineage specificity in regulation of a variety of other pathways involved in myogenesis, including the delta/notch, TGFβ, retinoic acid, and MAP kinase signaling cascades. Some of these differences may ultimately prove to have rather straightforward explanations, such as the differential growth rates of mEOM vs. mLM myotubes (MAP kinase) or the positional signaling intrinsic to limb myoblasts (Wnt and Hox), while others may be more difficult to decipher. Taken together, the analysis of signal transduction pathway differences in the EOM and limb muscle allotypes is complicated by the number of differentially expressed pathways and the numerous and overlapping ligands, antagonists, and receptors for these pathways. However, by using broad gene/protein expression analyses in new allotype-specific cell lines, we have mapped out starting points for subsequent studies to dissect the divergence in early developmental signaling pathways in skeletal muscle allotype specification.

Conceptual and practical value of skeletal muscle allotypes and allotype-specific cell lines. Just as the concept of distinctive myofiber types has been invaluable in guiding research in muscle biology, better utilization of the skeletal muscle allotype concept undoubtedly will advance knowledge of muscle developmental and disease mechanisms. Thirty-five years ago, Brooke and Kaiser (8) reconsidered the concept of skeletal muscle fiber types from the perspective of a single endpoint (myofibrillar ATPase histochemistry), asking, “how many (fiber types) and what kind?”. The identical question must be posed for skeletal muscle allotypes, but from a much broader transcriptome/proteome point of view. In proposing the muscle allotype concept, Hoh et al. (36, 37) relied on the sole endpoint of myosin heavy chain expression and argued for a two-component model: allotype specification requires the interaction of a permissive muscle precursor cell lineage with a facilitating extrinsic environment. Current evidence, from a genome-wide viewpoint, suggests that the basis for muscle allotypes is considerably more complex than myosin expression patterns (present data and Refs. 17, 18, 44, 66, 71).

To date, appropriate models have not been available to dissociate the contributions of intrinsic and extrinsic factors in allotype specification. Here, we derived cell lines for what arguably are two extremes among allotypes, EOM and limb muscle. Through use of the broadest expression profiling tools available, our data show that myoblast lineages, when exposed to the identical environmental conditions, make differential use of transcription factors and myogenic signaling cascades in attaining novel endpoints. These findings suggest that Hoh’s model, that an interaction of intrinsic and extrinsic factors determines allotypes, is correct but on a much broader scale. Myoblast lineage specifies the potential for a broad range of traits, represented among the differentially regulated genes and proteins in mEOM and mLM. While some traits are evident even in the absence of the appropriate in vivo environment, others require input from extrinsic factors for expression. In vivo, EOM and hindlimb are subjected to very different extrinsic influences (circulating hormones, local autocrine/paracrine factors, innervation, and usage patterns), ultimately yielding muscles that are highly adapted to specific functional tasks.

In this study, we performed initial, broad analyses of myogenesis in new allotype-specific cell lines offering the advantage of tight control over the non-cell-autonomous influences. Our results represent one step toward understanding the developmental basis behind fundamentally distinct skeletal muscle groups and, perhaps more importantly, identifying links between muscle group transcriptomes/proteomes/phenotypes and their preferential targeting or sparing in a variety of neuromuscular diseases.

ACKNOWLEDGMENTS

We thank Marty Veigl and Patrick Leahy, of the Case Gene Expression Array Core Facility, for assistance with the microarray studies; Georgiana Cheng, Scott Shnider, and Dan Rischar for assistance with data analysis; and Maziar Assadi for helpful discussions.

GRANTS

These studies were supported by United States Public Health Service (USPHS) Grants R01 EY015306, R01 EY09834, and R01 EY12779 to J. D. Porter. DNA microarray and bioinformatics core module support were provided by USPHS Grants P30 CA43703 and P30 EY11371, respectively.

REFERENCES


63. Porter JD and Baker RS. Muscles of a different ‘color’: the unusual properties of the extraocular muscles may predispose or protect them in neurogenic and myogenic disease. Neurology 46: 30–37, 1996.


