Available online at www.sciencedirect.com
ScienceDirect
Developmental Biology 311 (2007) 650–664
DEVELOPMENTAL BIOLOGY
www.elsevier.com/developmentalbiology

Genomes & Developmental Control

Target gene selectivity of the myogenic basic helix–loop–helix transcription factor myogenin in embryonic muscle

Judith K. Davie [1], Jang-Hyeon Cho, Eric Meadows, Jesse M. Flynn, Jennifer R. Knapp [2], William H. Klein *

Department of Biochemistry and Molecular Biology, Unit 1000, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA
Graduate Training Program in Genes and Development, The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX 77030, USA

Received for publication 13 March 2007; revised 4 July 2007; accepted 7 August 2007
Available online 16 August 2007

Abstract

The myogenic regulatory factors MyoD and myogenin are crucial for skeletal muscle development. Despite their importance, the mechanisms by which these factors selectively regulate different target genes are unclear. The purpose of the present investigation was to compare embryonic skeletal muscle from *myogenin*$^{+/+}$ and *myogenin*$^{-/-}$ mice to identify genes whose expression was dependent on the presence of myogenin but not MyoD and to determine whether myogenin-binding sites could be found within regulatory regions of myogenin-dependent genes independent of MyoD. We identified a set of 140 muscle-expressed genes whose expression in embryonic tongue muscle of *myogenin*$^{-/-}$ mice was downregulated in the absence of myogenin, but in the presence of MyoD. Myogenin bound within conserved regulatory regions of several of the downregulated genes, but MyoD bound only to a subset of these same regions, suggesting that many downregulated genes were selective targets of myogenin. The regulatory regions activated gene expression in cultured myoblasts and fibroblasts overexpressing myogenin or MyoD, indicating that expression from exogenously introduced DNA could not recapitulate the selectivity for myogenin observed in vivo. The results identify new target genes for myogenin and show that myogenin's target gene selectivity is not based solely on binding site sequences.

Keywords: Skeletal muscle development; bHLH myogenic regulatory factors; Myogenin; MyoD; Target gene selection; Gene expression profiling; Chromatin immunoprecipitation

Introduction

As transcription factor families evolve, individual members adopt specialized functions for selectively regulating gene expression. Specialization within transcription factor families is considered to be an important driver of developmental diversity (Kirschner and Gerhart, 1998) but how closely related transcription factors exhibit very different molecular functions remains poorly understood. Vertebrate skeletal muscle development is a well-established model for investigating the mechanisms by which related transcription factors selectively regulate the expression of distinct sets of genes (Olson and Klein, 1994; Black and Olson, 1998; Buckingham et al., 2003). In all vertebrates studied to date, two transcription factor families are essential for activating the gene regulatory network that drives skeletal muscle development during embryogenesis — the four closely related myogenic basic helix–loop–helix (bHLH) regulatory factors, MyoD, Myf5, myogenin, and MRF4/Myf6, and the multiple isoforms of the MADS-box factors, Mef2a, Mef2c, and Mef2d (Blais et al., 2005). For full activity, the myogenic bHLH regulatory factors dimerize with E proteins to bind conserved E-box sequences in the regulatory regions of muscle genes (Berkes and Tapscott, 2005).

* Corresponding author. Fax: +1 713 834 6266.
E-mail address: whklein@mdanderson.org (W.H. Klein).
[1] Current address: Department of Biochemistry and Molecular Biology, Southern Illinois University School of Medicine, Carbondale, IL 62901, USA.
[2] Current address: Fred Hutchinson Cancer Research Center, A2-168, Seattle, WA 98109, USA.

0012-1606/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2007.08.014

The functional differences between MyoD and myogenin are particularly relevant to our study. Although these transcription factors have virtually identical bHLH regions and extensive similarity in the remainder of their sequences, they nevertheless play very different roles in myogenesis (Bergstrom and Tapscott, 2001; De la Serna et al., 2005; Berkes and Tapscott, 2005). Specifically, MyoD functions early in myogenesis to confer a myogenic fate on mesodermal progenitor cells, whereas myogenin functions later to cause specified myoblasts to differentiate into functional myofibers (Buckingham et al., 2003; Berkes and Tapscott, 2005).

Recent studies using cultured myoblasts or fibroblasts have revealed functional distinctions between MyoD and myogenin at the molecular level (Berkes and Tapscott, 2005; De la Serna et al., 2005; Tapscott, 2005). MyoD has specialized domains that are involved in chromatin remodeling but myogenin does not have these domains (Bergstrom and Tapscott, 2001; Cao et al., 2006; Ohkawa et al., 2006). Thus, MyoD binds to regulatory regions of skeletal muscle genes, promotes nucleosome remodeling, and recruits transcriptional co-activators (Berkes and Tapscott, 2005; De la Serna et al., 2005; Tapscott, 2005). Furthermore, MyoD binds to regulatory regions of genes expressed early in myogenesis and activates their expression (Cao et al., 2006; Ohkawa et al., 2006, 2007). MyoD also binds to regulatory regions of genes expressed later in myogenesis and initiates nucleosome remodeling but is thought to be insufficient for gene activation (De la Serna et al., 2005; Cao et al., 2006). In contrast, myogenin does not appear to bind to gene regulatory regions early in myogenesis but is thought to bind efficiently within the regulatory regions of genes expressed late in myogenesis following chromatin remodeling by MyoD. Genes activated late in myogenesis may require both MyoD and myogenin binding within the regulatory region (Cao et al., 2006) or require that MyoD remodels chromatin before myogenin can be recruited for binding (De la Serna et al., 2005; Cao et al., 2006; Ohkawa et al., 2006, 2007).

Despite recent progress, many key issues remain unresolved. During embryonic muscle development in vivo, many skeletal muscle genes are not readily classified as early- or late-expressed genes. Whether all late-expressing genes require simultaneous binding of MyoD and myogenin at sites within regulatory regions to activate gene expression is unclear. Additionally, most of the studies attempting to distinguish the differences between MyoD and myogenin have been carried out with C2C12 myoblasts or other tissue culture cells. To be meaningful, the results require validation in vivo.

Myogenin stands out among the myogenic regulatory factors in one important aspect: unlike the phenotypes observed with single knockouts of the three other myogenic regulatory genes, a strong skeletal muscle phenotype is associated with a single knockout of *myogenin* (*myog*) in mice (Olson and Klein, 1994; Buckingham, 2001). Mice homozygous for a *myog*-null allele die at birth because of severe skeletal muscle deficiency (Hasty et al., 1993; Nabeshima et al., 1993). In late-stage embryos, $myog^{-/-}$ myoblasts are specified for a muscle fate throughout the body but myofiber formation is impaired. Despite this strong phenotype, several muscle-specific genes are expressed in $myog^{-/-}$ mice indicating that not all genes involved in muscle differentiation are controlled by myogenin (Venuti et al., 1995). Notably, *myoD* expression is not affected by a loss of myogenin, demonstrating that MyoD cannot compensate for myogenin's absence in vivo (Rawls et al., 1995; Venuti et al., 1995). In addition, $myog^{-/-}$ embryonic stem (ES) cells, when induced to differentiate, form far fewer myotubes than wild-type ES cells do, and forced overexpression of *myoD* in $myog^{-/-}$ ES cells does not make up for myotube deficiencies (Myer et al., 2001).

MyoD and myogenin thus differ fundamentally. Because MyoD and myogenin are DNA-binding proteins and transcriptional activators, the functional differences between them must be accounted for in regulating the expression of distinct gene sets during myogenesis. The question then arises whether any genes targeted by myogenin can be identified that are not also targets of MyoD. The purpose of the present investigation was to compare embryonic skeletal muscle from $myog^{+/+}$ and $myog^{-/-}$ mice to identify genes whose expression was dependent on the presence of myogenin but not MyoD and to determine whether myogenin-binding sites can be found within the regulatory regions of myogenin-dependent genes independent of MyoD.

Using a microarray platform, we performed gene expression profiling to identify a set of genes whose expression was significantly downregulated in the embryonic tongue muscle of $myog^{-/-}$ mice in the absence of myogenin. We identified putative phylogenetically conserved transcriptional regulatory regions associated with several myogenin-dependent genes and found enhancer E-box elements within these regulatory regions that were bound by myogenin but not MyoD. The transcriptional activity of these regulatory regions in cultured myoblasts and fibroblasts suggested that the ability to discriminate between myogenin and MyoD in activating gene transcription requires more than target sequence specificity.

Materials and methods

Affymetrix microarray analysis

$Myog^{+/-}$ mice were intercrossed, and tongue tissue samples from the resulting embryos were harvested at embryonic day 15.5 (E15.5) and stored in RNA*later* (Ambion). In all experiments using $myog^{+/+}$, $myog^{+/-}$, and $myog^{-/-}$ mice, the U.S. Public Health Service Policy on Humane Care and Use of Laboratory Animals was followed and the M. D. Anderson Institutional Animal Care and Use Committee approved all animal protocols for these experiments. $Myog^{+/-}$ mice were maintained on a C57/BL/6J background. Total RNA was extracted from the samples using an RNeasy fibrous tissue kit (Qiagen). Five micrograms of total RNA from each sample was used for preparing hybridization probes. Three independent samples were used in standard Affymetrix protocols to yield fluorescently labeled cRNA fragments, which were hybridized to Mouse Genome 430 2.0 GeneChips (Affymetrix). The hybridized GeneChips were scanned using an Affymetrix GeneArray scanner, and the raw image files were analyzed using ArrayAssist 4.0 (Stratagene) with PLIER (Probe Logarithmic Intensity Error) normalization (Katz et al., 2006).

P-values were calculated using unpaired *t*-tests between $myog^{+/+}$ and $myog^{-/-}$ datasets. Genes were grouped based on their reported Gene Ontology Biological Process descriptions (Hill et al., 2002). All annotations are as reported by the NetAffx Analysis Center (http://www.affymetrix.com/analysis/index.affx). We also used the gene descriptors available at the National Center for Biotechnology (NCBI) Information Mouse Genome Resource web site (http://www.ncbi.nlm.nih.gov/genome/guide/mouse/) to assist in grouping the

genes. The complete microarray dataset is listed in Table S1 and the subset that includes the downregulated genes is listed in Table S2.

Quantitative reverse-transcriptase polymerase chain reaction

Two micrograms of total RNA from E15.5 embryonic tongue tissue of *myog*$^{+/+}$ or *myog*$^{-/-}$ mice was reversed transcribed with Superscript reverse-transcriptase (Invitrogen). cDNA equivalent to 40 ng was used for quantitative polymerase chain reaction (qPCR) amplification (Applied Biosystems) with SYBR green PCR master mix (Applied Biosystems). The relative levels of expression of genes selected from the microarray analysis were normalized according to those of ribosomal protein L7 (Rpl7). qPCR data were calculated using the comparative C_t method (Applied Biosystems). Standard deviations from the mean of the $[\Delta]C_t$ values were calculated from three independent RNA samples. The PCR primers used for detecting the transcripts of the selected genes are listed in Table S3.

In situ hybridization and immunocytochemistry

Twelve-micrometer frozen sections were obtained from 4% paraformaldehyde-fixed tongue and heart samples dissected from E15.5 *myog*$^{+/+}$ and *myog*$^{-/-}$ embryos. To generate antisense and sense probes for in situ hybridization, PCR products from cDNAs derived from tongue tissue samples were cloned into the vector pCRII-TOPO (Invitrogen). The cloned products were linearized using either *Hin*dIII or *Xba*I and subsequently transcribed in vitro using T7 RNA polymerase for the antisense probe and Sp6 RNA polymerase for the sense probe in the presence of digoxigenin (Boehringer-Mannheim, Germany). Non-radioactive in situ hybridization was performed to detect spatial expression as previously described (Lee et al., 2000). Nuclear Fast Red was used for counterstaining of the tissue samples. For double in situ hybridization immunocytochemistry, sections of E15.5 *myog*$^{+/+}$ and *myog*$^{-/-}$ tongue tissue were washed in phosphate-buffered saline (PBS, pH 7.2–7.4) and immunocytochemistry was carried out as described below.

For immunocytochemistry, endogenous peroxidase activity and nonspecific antibody binding in sectioned tissues were blocked using 3% hydrogen peroxide and 3% bovine serum albumin (BSA) in PBS for 1 h. Rabbit anti-MyoD (M-318) and anti-myogenin (M-225) antibodies (Santa Cruz Biotechnology) were used at dilutions of 1:200 and 1:800, respectively. Antibodies were diluted in 3% normal goat serum in PBS containing 3% bovine serum albumin, and sections of tongue tissue were incubated with the antibodies at 4 °C for 16 h. Immunoreactivity in the sections was visualized using an ABC kit (Vector Laboratories) for colorimetric visualization. Methyl green was used for counterstaining of the sectioned tissues in both immunocytochemistry and double in situ hybridization immunocytochemistry labeling experiments.

Chromatin Immunoprecipitation analysis of embryonic tissue

E15.5 tongue tissue from *myog*$^{+/+}$ mice was isolated, cross-linked, and processed for chromatin immunoprecipitation (ChIP) analysis as previously described (Wells and Farnham, 2002). Extracts of tongue tissue were precleared using incubation with preblocked protein-A agarose beads (Upstate Biotechnologies), which were also used to immunoprecipitate the antibody complexes from tongue tissue extracts following antibody addition to the incubation mix. The antibodies used in the ChIP analysis were a ChIP-grade anti-MyoD antibody (C-20; Santa Cruz Biotechnology) and a ChIP-grade anti-myogenin antibody (M-225; Santa Cruz Biotechnology). Two micrograms of the antibodies was used for each immunoprecipitation and with extracts corresponding to approximately 1.5 tongues for each immunoprecipitation. The antibodies showed no cross-reactivity as determined using immunoprecipitation with one antibody followed by Western blot analysis with the second antibody (data not shown). The PCR primers used for detecting protein-bound DNA in embryonic tongue extracts are listed in Table S3. DNA was amplified by qPCR using SYBR green PCR master mix (Applied Biosystems). The data are presented as the fold change between samples with and without antibodies with each sample normalized according to Rpl7. Input samples were tested for amplification and to compare samples across replicates but were not used for normalization. Inputs for all samples were identical, as extracts were obtained from pooled tongue tissue samples and aliquoted for immunoprecipitation. qPCR data were calculated using the comparative C_t method. Values of $[\Delta][\Delta]C_t$ were determined by applying the following formula: $[\Delta]C_{t,\text{template}}$ (experimental) − $[\Delta]C_{t,\text{reference}}$ (Rpl7) = $[\Delta][\Delta]C_t$. Fold enrichments were calculated as $2^{-[\Delta][\Delta]Ct}$ (antibody)/$2^{-[\Delta][\Delta]Ct}$ (no antibody). Standard error from the mean was calculated from replicate $[\Delta][\Delta]C_t$ values.

Transient transfection and luciferase assays

Upstream gene regulatory regions were amplified from genomic DNA using PCR and cloned into the pGL3 luciferase reporter gene vector (Invitrogen). The regulatory regions of candidate myogenin target genes used were AMP deaminase 1 (*Ampd1*; base pairs −8 to −431), LIM domain binding protein 3 (*Ldb3*; base pairs −1 to −1209), Leiomodin2 (*Lmod2*; base pairs −10 to −458), and Calsequestrin2 (*Casq2*; base pairs −3 to −760). Base pair positions are defined with respect to the translational A(TG) start site with A considered to be +1. Fugene 6 and Fugene HD transfection reagents (Roche) were used for DNA transfections and the level of luciferase activity in transfected cells was determined using a dual-luciferase reporter assay system (Promega). The plasmids EMSVmyog (provided by Diane Edmondson, Tanex, Inc.) and pEMCIIs (provided by Andrew Lassar, Harvard Medical School) were used for expressing *myog* and *myoD*, respectively, in transfected NIH-3T3 cells. C2C12 and NIH-3T3 cells were grown in Dulbecco's modified Eagle's medium containing 10% bovine calf serum, 2% L-glutamine, 1% penicillin–streptomycin according to standard tissue culture protocols (Pollard and Walker, 1997). C2C12 cells were induced to differentiate into myotubes by allowing the cells to reach confluence and switching the medium to Dulbecco's modified Eagle's medium in 2% horse serum, 0.2% L-glutamine, 0.5% penicillin–streptomycin. Expression values for each construct were calculated as a percentage of expression of the pGL3 control vector. Standard deviations from the mean was calculated from the mean expression values representing at least three separate experiments.

The *math5* expression construct was generated by cloning the complete *math5* coding sequence (Mu et al., 2005) into the pIRES-hrGFP-1a vector in frame with 3× FLAG sequences. Transactivation experiments were performed at least in duplicate using NIH-3T3 cells and Fugene HD transfection reagent. E-boxes identified as putative myogenin-binding sites through ChIP analysis were deleted from the *Lmod2* and *Ampd1* vectors using the QuickChange II site-directed mutagenesis kit (Stratagene). Primer sequences used to generate the mutations are listed in Table S3. Transactivation experiments were performed at least in duplicate using NIH-3T3 cells and Fugene HD transfection reagent.

Results

Identification of myogenin-dependent genes in E15.5 tongue muscle

From E12.5 to E18.5, committed myoblasts throughout the muscle-forming regions of the body differentiate into myocytes and fuse to form myofibers (Venuti et al., 1995). *Myog* expression begins at E8.5 in the dermamyotome and continues thereafter at all sites of skeletal muscle development (Cheng et al., 1992; Venuti et al., 1995). We elected to perform gene expression profiling at E15.5 with embryonic tongue muscle. *Myog*$^{+/+}$ and *myog*$^{-/-}$ tongues look very much alike at that time, although severe deficiencies in myofibers in *myog*$^{-/-}$ embryonic tongues become apparent later.

We extracted RNA from tongue tissue samples from three independent *myog*$^{+/+}$ and *myog*$^{-/-}$ embryos and used them to generate biotin-labeled cRNA probes for hybridization to Mouse Genome 430 2.0 GeneChips. We used ArrayAssist 4.0 to analyze the microarray hybridization results and used PLIER normalization to identify selected genes that showed changes in expression with P values ≤ 0.05 and empirically chosen fold

changes of ≥1.5. We identified 140 genes whose expression was downregulated and 240 genes whose expression was upregulated that met these criteria.

Myogenin functions mainly as a transcriptional activator although, in certain cases, it may act to repress transcription (Trouche et al., 1995). However, because our objective was to identify direct target genes whose expression was activated by the presence of myogenin, we focused on the downregulated genes. The complete set of upregulated and downregulated genes are listed in Table S1, but in contrast to the downregulated genes (see below), we failed to validate a sampling of the upregulated genes by independent qRT-PCR analysis (data not shown). This suggests that many of the upregulated genes identified by the microarray analysis were not bona fide targets of myogenin.

Of the 140 downregulated genes, 117 were previously characterized, while 23 were identified as genes encoding

Table 1
Fifty representative genes whose expression is downregulated in E15.5 tongue muscle samples obtained from *myog*$^{-/-}$ embryos fold changes (*myog*$^{+/+}$/*myog*$^{-/-}$)

Gene title	Gene symbol	Probe set ID	Array	qRT-PCR	SD
RIKEN cDNA 1100001E04 gene	*1100001E04Rik*	1454660_at	5.7	12.4	3.4
Leiomodin 2 (cardiac)	*Lmod2*	1452345_at	4.2	18.3	4.2
Troponin T3, skeletal, fast	*Tnnt3*	1450118_a_at	4.1	–	–
Myogenin	*Myog*	1419391_at	3.7	–	–
Cysteine and glycine-rich protein 3	*Csrp3*	1460318_at	3.7	8.3	2.8
Cytochrome *c* oxidase muscle isoform 8b	*Cox8b*	1449218_at	3.3	8.7	1.1
Creatine kinase, mitochondrial 2	*Ckmt2*	1428722_at	3.3	2.8	0.2
Troponin I-2	*Tnni2*	1438609_x_at	3.2	6.1	1.1
Integrin beta 1 binding protein 2	*Itgb1bp2*	1423238_at	3.1	6.1	0.9
Myogenic factor 6	*Myf6*	1419150_at	3.0	8.2	0.3
Spastic paraplegia 21 homologue (human)	*Spg21*	1451036_at	2.7	–	–
Myosin, light polypeptide 3	*Myl3*	1427768_s_at	2.7	–	–
Myozenin 2	*Myoz2*	1418769_at	2.6	6.9	0.9
Myozenin 1	*Myoz1*	1460202_at	2.6	14.4	3.4
Dystrophia myotonica-protein kinase	*Dmpk*	1434944_at	2.6	–	–
Myosin heavy chain 1	*Myh1*	1427868_x_at	2.5	4.5	0.7
Adenylosuccinate synthetase like 1	*Adssl1*	1449383_at	2.5	4.7	0.7
Tubulin, beta 6	*Tubb6*	1416431_at	2.4	–	–
Myosin, light polypeptide 1	*Myl1*	1452651_a_at	2.4	–	–
Myomesin 2	*Myom2*	1450917_at	2.4	3.6	0.1
Interleukin 17B	*Il17b*	1431693_a_at	2.4	–	–
Calsequestrin 2	*Casq2*	1422529_s_at	2.4	3.9	0.4
Interferon gamma induced GTPase	*Igtp*	1417141_at	2.3	–	–
Cytochrome *c* oxidase, subunit 6a, polypeptide 2	*Cox6a2*	1417607_at	2.3	6.4	0.3
Serine/Threonine kinase 23	*Stk23*	1447806_s_at	2.2	5.1	0.6
Nephroblastoma overexpressed gene	*Nov*	1426852_x_at	2.2	4.2	1.0
Myosin, light polypeptide 2, regulatory, cardiac, slow	*Myl2*	1448394_at	2.2	–	–
LIM domain protein binding 3	*Ldb3*	1416752_at	2.2	8.2	2.2
Small muscle protein, X-linked	*Smpx*	1418095_at	2.1	3.8	1.0
Double cortin and calcium/calmodulin-dependent protein kinase-like 1	*Dcamkl1*	1436659_at	2.1	4.2	0.7
Zyxin	*Zyx*	1417240_at	2.0	2.7	0.1
Upregulated during skeletal muscle growth 4	*Usmg4*	1417626_at	2.0	–	–
Sarcalumenin	*Srl*	1436867_at	2.0	–	–
Transcription elongation factor A (SII), 3	*Tcea3*	1424531_a_at	1.9	4.9	0.8
PDZ and LIM domain 5	*Pdlim5*	1429783_at	1.8	–	–
Titin	*Ttn*	1431928_at	1.7	–	–
Enabled homologue (*Drosophila*)	*Enah*	1421624_a_at	1.7	1.5	0.3
CDC42 effector protein (Rho GTPase binding) 2	*Cdc42ep2*	1428750_at	1.7	3.3	0.9
Translin-associated factor X	*Tsnax*	1430045_at	1.6	–	–
Synaptic nuclear envelope 1	*Syne1*	1455493_at	1.6	–	–
Fibroblast growth factor receptor-like 1	*Fgfrl1*	1447878_s_at	1.6	7.7	4.4
Calcium channel, voltage-dependent, alpha2/delta subunit 1	*Cacna2d1*	1433643_at	1.6	–	–
AMP deaminase 1 (isoform M)	*Ampd1*	1434722_at	1.6	30.5	9.3
Tropomodulin 4	*Tmod4*	1449969_at	1.5	–	–
Titin-cap	*Tcap*	1423145_a_at	1.5	–	–
Junctophilin 2	*Jph2*	1455404_at	1.5	–	–
Fibroblast growth factor 6	*Fgf6*	1427582_at	1.5	6.0	2.7
Cofilin 2, muscle	*Cfl2*	1418067_at	1.5	–	–
Actin binding LIM protein family, member 3	*Ablim3*	1434013_at	1.5	–	–
Actinin α2	*Actn2*	1456968_at	1.2	3.2	0.5

Fold change calculations for the Affymetrix array and qRT-PCR analyses are described in Materials and methods. One gene, *Actn2*, was included in the dataset although the fold change in the expression of this gene calculated using the array analysis was less than 1.5. SD, standard deviation from the mean for the fold-change values calculated using qRT-PCR.

Table 2
Expression of MRFs in E15.5 tongue muscle samples obtained from *myog*$^{-/-}$ embryos Fold changes (*myog*$^{-/-}$/*myog*$^{+/+}$)

MRF	Affymetrix microarray	qRT-PCR mean±SD
myoD	1.37	1.25±0.12
myf5	0.90	0.83±0.09
mrf4 (myf6)	0.33	0.12±0.07

The fold changes in MRF expression for the Affymetrix microarray and qRT-PCR analyses were calculated as described in Materials and methods. For qRT-PCR, the values were normalized relative to *Rpl7* expression. SD, mean standard deviation.

hypothetical proteins by gene prediction algorithms. Virtually all of the 117 known downregulated genes were represented in various skeletal muscle EST databases and many encoded structural or enzymatic components of skeletal muscle (Table S2). Table 1 presents fifty selected genes that were all downregulated in *myog*$^{-/-}$ tongue muscle. Examples included myosin heavy chain 1 (*Myh1*), Leiomodin2 (*Lmod2*), Troponin I-2 (*Tnni2*), Titin (*Ttn*), Actinin α2 (*Actn2*), cytochrome *c* oxidase muscle isoform 8b (*Cox8b*), muscle-specific serine/threonine kinase 23 (*Stk23*), Calsequestrin2 (*Casq2*), and AMP deaminase1 (*Ampd1*).

Genes encoding subunits of the sarcomeric Z-disks were strongly represented in the downregulated set. These genes included Myozenin1 (*Myoz1*), Myozenin2 (*Myoz2*), muscle-specific integrin β1 binding protein2 (*Itgβ1bp2*), and LIM domain binding protein3 (*Ldb3*). This finding was particularly notable because Z-disks in residual myofibers within *myog*$^{-/-}$ muscle are either severely disrupted or completely absent (Nabeshima et al., 1993). The gene expression profiling results provided a molecular explanation for this phenotype and suggested that myogenin plays a selective role in controlling the expression of Z-disk components.

The microarray analysis also identified Fibroblast growth factor 6 (*Fgf6*) and Fibroblast growth factor receptor-like 1 (*Fgfrl1*) as significantly downregulated genes. FGF6 is known to regulate myogenesis (Armand et al., 2006) and FGFRL1 is expressed in skeletal muscle and is thought to act as a decoy receptor for fibroblast growth factor ligands (Trueb et al., 2003).

Myogenic regulatory factor expression in myog$^{-/-}$ *embryonic tongue muscle*

Our gene profiling analysis included all the myogenic regulatory factors (Table 2). Previous reports showed that the expression of *myoD* and *myf5* was not altered in *myog*$^{-/-}$ embryos (Rawls et al., 1995; Venuti et al., 1995). These genes are genetically upstream of *myog* and have roles in myoblast specification. We confirmed that *myoD* and *myf5* were expressed at wild-type levels in *myog*$^{-/-}$ tongues in our microarray experiments and qRT-PCR analysis (Table 2). In contrast with the expression *myoD* and *myf5*, expression of *mrf4* (*myf6*) was strongly downregulated in *myog*$^{-/-}$ embryonic tongues as had been reported previously in embryonic limb muscle (Table 2; Rawls et al., 1995).

Immunostaining of embryonic tongue tissue with anti-myogenin and anti-MyoD antibodies vividly illustrates the independence of MyoD protein expression from myogenin (Fig. 1). In *myog*$^{+/+}$ tongues, we detected expression of both myogenin (Fig. 1A) and MyoD (Fig. 1C) in well-organized myofibers oriented longitudinally. In *myog*$^{-/-}$ tongues, we could not detect myogenin expression above nonspecific background staining (Fig. 1B) whereas we observed high levels

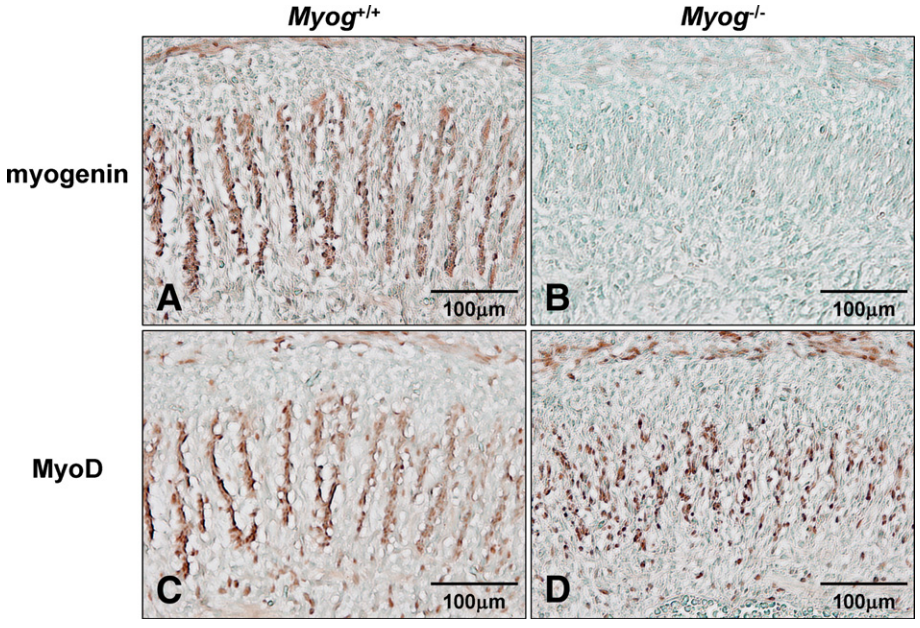

Fig. 1. Myogenin protein expression is absent from E15.5 tongue muscle obtained from *myog*$^{-/-}$ embryos but MyoD protein expression is unaltered. (A, B) Myogenin expression detected using immunostaining with M-225, an anti-myogenin antibody. (C, D) MyoD protein expression detected using immunostaining with M-318, an anti-MyoD antibody.

of MyoD expression in unfused myocytes (Fig. 1D). These results confirmed previous findings (Hasty et al., 1993; Venuti et al., 1995) and further demonstrated that the presence of MyoD is insufficient for myofiber formation in the absence of myogenin.

Functional and expression properties of putative myogenin target genes

To validate our microarray results, we examined the expression of 27 downregulated genes using qRT-PCR. We extracted RNA from E15.5 *myog*$^{+/+}$ and *myog*$^{-/-}$ tongue tissue samples and compared the levels of expression of the putative myogenin target genes using three different samples for each genotype (Table 1; Fig. 2). In all cases, the fold changes in the expression of these genes calculated using qRT-PCR confirmed the results of the microarray experiments. Moreover, we routinely found that the fold changes obtained using qRT-PCR were substantially larger than those obtained in the microarray analysis (Table 1).

We selected several strongly downregulated genes to determine whether they were direct targets of myogenin. Genes encoding structural and enzymatic components of skeletal muscle whose expression was strongly dependent on the presence of myogenin were obvious candidates for target genes. Expression of the representative skeletal muscle structural proteins, *Lmod2*, *Ldb3*, and *Myoz1* decreased significantly in *myog*$^{-/-}$ tongue muscle (Table 1; Fig. 2). Lmod2 is the cardiac isoform of Leiomodin and is expressed in both heart and skeletal muscle during embryonic and adult life (Conley et al., 2001). Leiomodins are a family of proteins highly related to the Z-disk-associated Tropomodulin family of actin filament end-capping proteins (Conley et al., 2001; Fischer and Fowler, 2003). Ldb3, also called Cypher or Zasp, contains an N-terminal PDZ domain and three C-terminal LIM domains. Ldb3 is important for maintaining Z-disk stability by binding to Actinin α2 in its PDZ domain (Faulkner et al., 1999; Zhou et al., 1999; Klaavuniemi and Ylanne, 2006). Myoz1, also called Fatz-1 or calsarcin-1, is a sarcomeric protein implicated in the assembly and stabilization of Z-disks (Takada et al., 2001).

Three other genes downregulated in *myog*$^{-/-}$ tongue muscle had notable features that were suggestive of direct myogenin gene targets. *Fgfrl1*, which had a 7.7-fold change in expression (Table 1), is highly expressed in embryonic tongue and diaphragm muscle but only weakly expressed in embryonic limb muscle (Trueb and Taeschler, 2006). Tongue and diaphragm muscle are the muscle tissues most affected by the absence of myogenin in *myog*$^{-/-}$ mice (Hasty et al., 1993; Venuti et al., 1995). *Ampd1* showed a 30-fold change in expression (Table 1). A previous study identified two distinct regions of the *Ampd1* gene promoter required for expression in myocytes, although it did not determine whether E-box-binding site were contained within these regions (Morisaki and Holmes, 1993). Finally, the transcript encoded by *RIKEN1100001E04* showed one of the greatest changes in expression in both microarray and qRT-PCR analyses (Table 1). The NCBI-specified hypothetical protein for *RIKEN1100001E04* has a RhoGAP motif, which is a motif contained in GTPase activator proteins that act on Rho/Rac/Cdc42-like small GTPases (Zhang et al., 1997).

To confirm that the downregulated genes are in fact expressed in the same myofibers as those expressing myogenin, we performed immunocytochemistry with anti-myogenin M-225 antibody following in situ hybridization with *Ampd1*, *Lmod2*, *Fgfrl1*, and *RIKEN1100001E04* antisense probes (Fig. 3). Because myogenin protein was localized to the nuclei in tongue fibers and the transcripts of the abovementioned genes were localized in the cytoplasm of these fibers, we could readily detect the expression of both myogenin protein and gene transcripts. In all cases, myogenin protein was expressed in the same fibers as were the genes whose expression was downregulated in the absence of myogenin (Fig. 3).

Many of the myogenin-dependent genes that we identified are expressed not only in tongue muscle but also in other skeletal muscle in embryos and in cells other than skeletal muscle cells. If

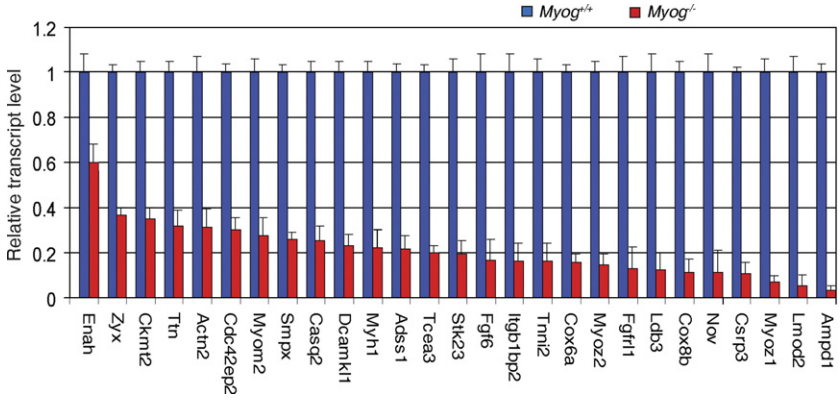

Fig. 2. Expression of genes encoding structural, enzymatic, and secreted factors is downregulated in E15.5 tongue muscle samples obtained from *myog*$^{-/-}$ embryos. The results of qRT-PCR analysis using RNA extracted from E15.5 tongue muscle and the primer pairs described in Table S3 are shown. Calculations of the relative fold changes in gene expression for *myog*$^{+/+}$ and *myog*$^{-/-}$ samples are described in the Materials and methods section. The data are normalized relative to the fold changes for *Rpl7*. Blue bars are *myog*$^{+/+}$ samples in which the average value of three samples is normalized relative to 1 as indicated on the *Y* axis. Red bars are *myog*$^{-/-}$ samples in which the average value of three samples is normalized relative to the *myog*$^{+/+}$ samples. Error bars are standard deviations from the mean. Full gene names are given in Table 1.

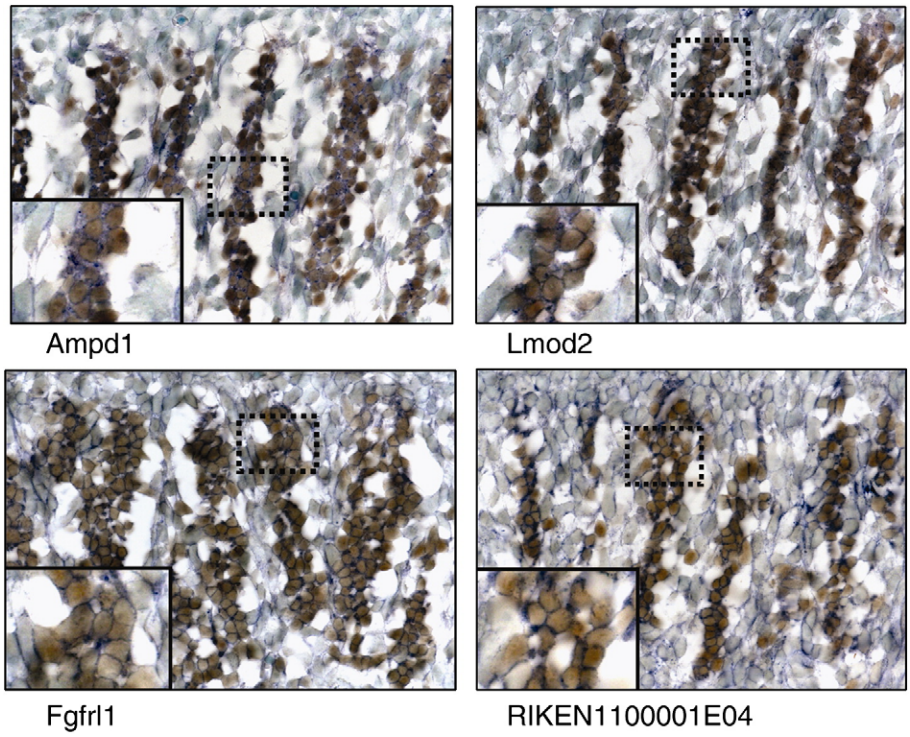

Fig. 3. Myogenin-downregulated genes are expressed in the same tongue muscle fibers as myogenin protein. Digoxigenin-labeled probes for *Ampd1*, *Lmod2*, *Fgfrl1*, and *RIKEN1100001E04* were hybridized in situ to E15.5 wild-type tongue sections (purple, cytoplasmic) followed by immunocytochemical staining with the anti-myogenin antibody M-225 (brown, nuclear). A magnified portion of each field is shown in the lower left corner of each image. The regions selected for magnification are indicated by dashed boxes.

these genes were bona fide targets of myogenin, their expression should be downregulated in tongue and non-tongue skeletal muscle cells of *myog*$^{-/-}$ embryos. Moreover, expression of myogenin target genes in non-skeletal muscle cells should be unaffected by myogenin's absence. We therefore determined the spatial expression patterns of *Ampd1*, *Lmod2*, *Fgfrl1*, *Myoz1*, *Ldb3*, and *RIKEN1100001E04* in *myog*$^{+/+}$ and *myog*$^{-/-}$ embryos on E15.5 using sectioned in situ hybridization. In each case, we observed decreased expression of these genes in *myog*$^{-/-}$ tongues compared with that in *myog*$^{+/+}$ tongues; representative examples using antisense probes for *Ampd1*, *Myoz1*, *Lmod2* and *myog* (serving as a control) are shown in Fig. 4A. Sense probes did not show any significant labeling with comparable tongue tissue samples (data not shown). Some of these genes are also expressed in wild-type embryonic limb muscle and in cardiac muscle where *myog* is never expressed. In *myog*$^{-/-}$ embryos, expression of *Fgfrl1*, *Myoz1*, and *Ldb3* was significantly downregulated in hindlimb muscle but was not downregulated in cardiac muscle (Fig. 4B). These results imply that separate regulatory mechanisms independent of myogenin operate in cardiac muscle to control the expression of *Fgfrl1*, *Myoz1*, and *Ldb3*.

Conserved E-box elements within the regulatory regions of myogenin-dependent genes

To find potential binding sites for myogenin and MyoD within regulatory regions of putative target genes, we searched for the presence of phylogenetically conserved noncoding sequences containing bHLH protein-binding sites (E-boxes) for 10 myogenin-downregulated genes in the NCBI Mouse Genome Resource website. These genes were chosen because they were strongly downregulated in *myog*$^{-/-}$ skeletal muscle, making them likely target genes for myogenin. Prior information on regulatory sequences for these ten genes was limited only to the proximal upstream region of *Ampd1* (Morisaki and Holmes, 1993). Regulatory regions in the nine other candidate genes have not been characterized. We focused on 10-kb regions encompassing DNA sequences upstream and downstream of the transcriptional initiation sites of the candidate genes. To perform this analysis, we used the computational tool regulatory VISTA (rVISTA) to identify clusters of myogenin/MyoD, E-protein and Mef2 consensus binding sites that were conserved between the mouse and human genomes (Dubchak and Ryaboy, 2006). For example, applying these criteria, rVISTA identified putative myogenin/MyoD-binding sites in conserved noncoding sequences in base pairs −1 to −356 upstream of *Casq2*, base pairs −1 to −596 upstream of *Lmod2*, and base pairs −676 to −1417 upstream of *Ldb3* (Fig. 5; Fig. S1). The DNA sequences containing E-box and Mef2-binding sites for these genes are conserved between mouse and human genomes (Fig. 5). The rVISTA analysis identified putative myogenin/MyoD-binding sites upstream of the transcriptional initiation sites for *Myoz1*, *Csrp3*, *Ttn*, *Lmod2*, *Ldb3*, *Fgfrl1*, *Fgf6*, *Casq2*, *Ampd1*, and *RIKEN1100001E04*. In addition, the rVISTA analysis identified potential myogenin/MyoD-binding sites within the first intron of *Ldb3* and *RIKEN1100001E04*. In a few instances, we observed more than one myogenin/MyoD-binding site clustered within a

conserved region, suggesting the possibility of multiple bHLH factor binding sites within one region (Fig. S1).

In vivo binding of myogenin and MyoD to conserved regulatory regions

We used ChIP assays to determine whether myogenin bound to E-box elements within the potential regulatory regions of 9 genes whose expression was downregulated in the absence of myogenin in vivo (Fig. 6A). We prepared chromatin extracts from cells of E15.5 tongue tissue and cross-linked and immunoprecipitated chromatin with anti-myogenin antibody M-225. We designed primer sets to amplify the conserved binding sites identified by rVISTA. In cases where we found more than one conserved region at a gene locus, we generated additional primer sets corresponding to the multiple potential E-box sites. In these experiments, an upstream regulatory region from *mef2c* known to have an E-box that binds to both myogenin

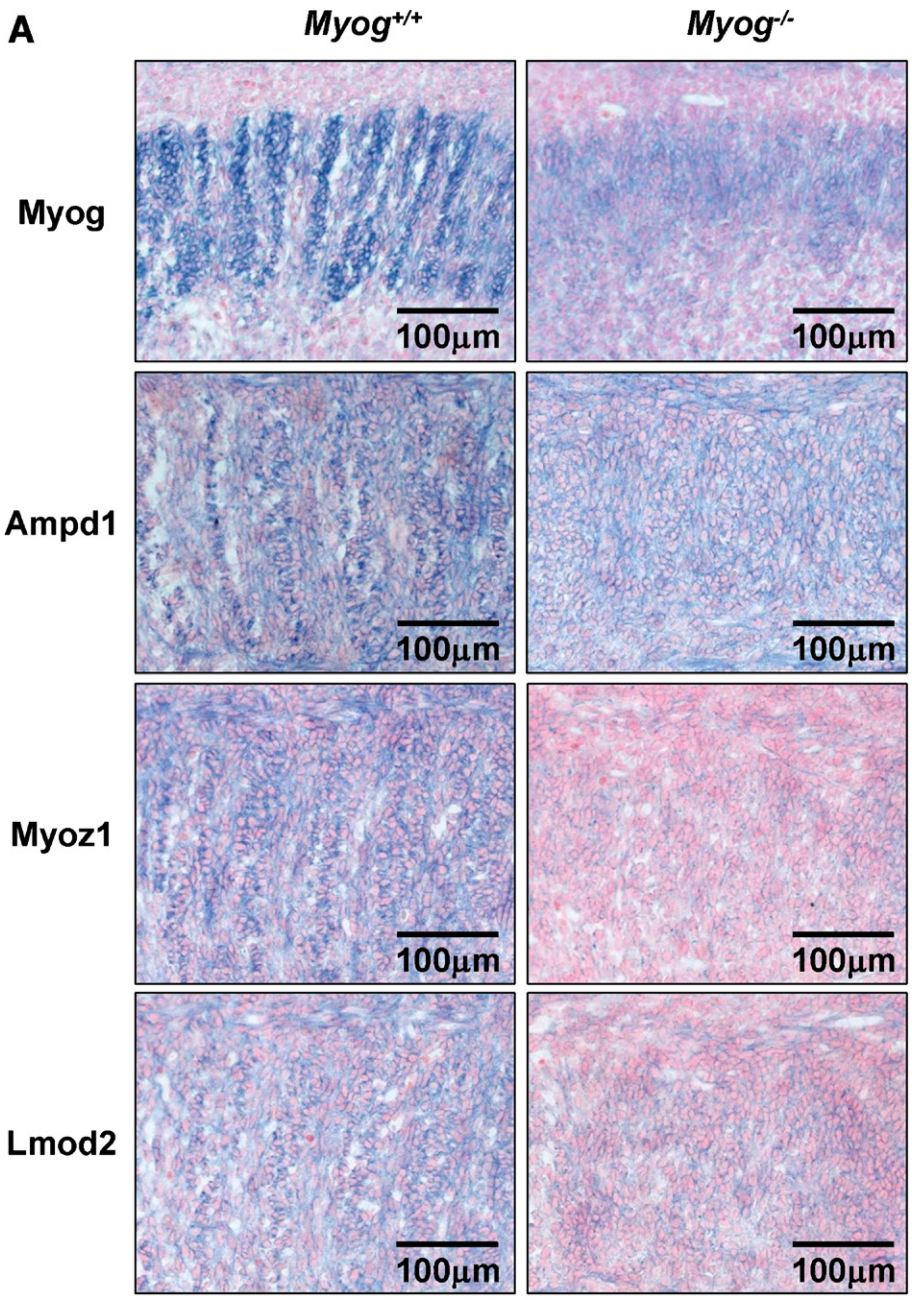

Fig. 4. Expression of myogenin-dependent genes in E15.5 tongue, limb, and heart muscle using in situ hybridization analysis shows skeletal muscle specificity. (A) Expression of *myog*, *Ampd1*, *Myoz1*, and *Lmod2* is downregulated in $myog^{-/-}$ tongue muscle. (B) *Fgfrl1*, *Myoz1*, and *Ldb3* are expressed at wild-type levels in $myog^{-/-}$ heart muscle, whereas their expression is downregulated in $myog^{-/-}$ intercostal muscle (ICM). The left-most panels show low-magnification images of sections from the trunk areas of E15.5 $myog^{+/+}$ (+/+) and $myog^{-/-}$ (−/−) embryos. The sections selected for magnification are indicated by dashed boxes. The upper right dashed boxes are heart muscle and higher magnifications of these sections are shown the middle panels. The lower left dashed boxes are intercostal muscle and higher magnifications of these sections are shown in the right-most panels.

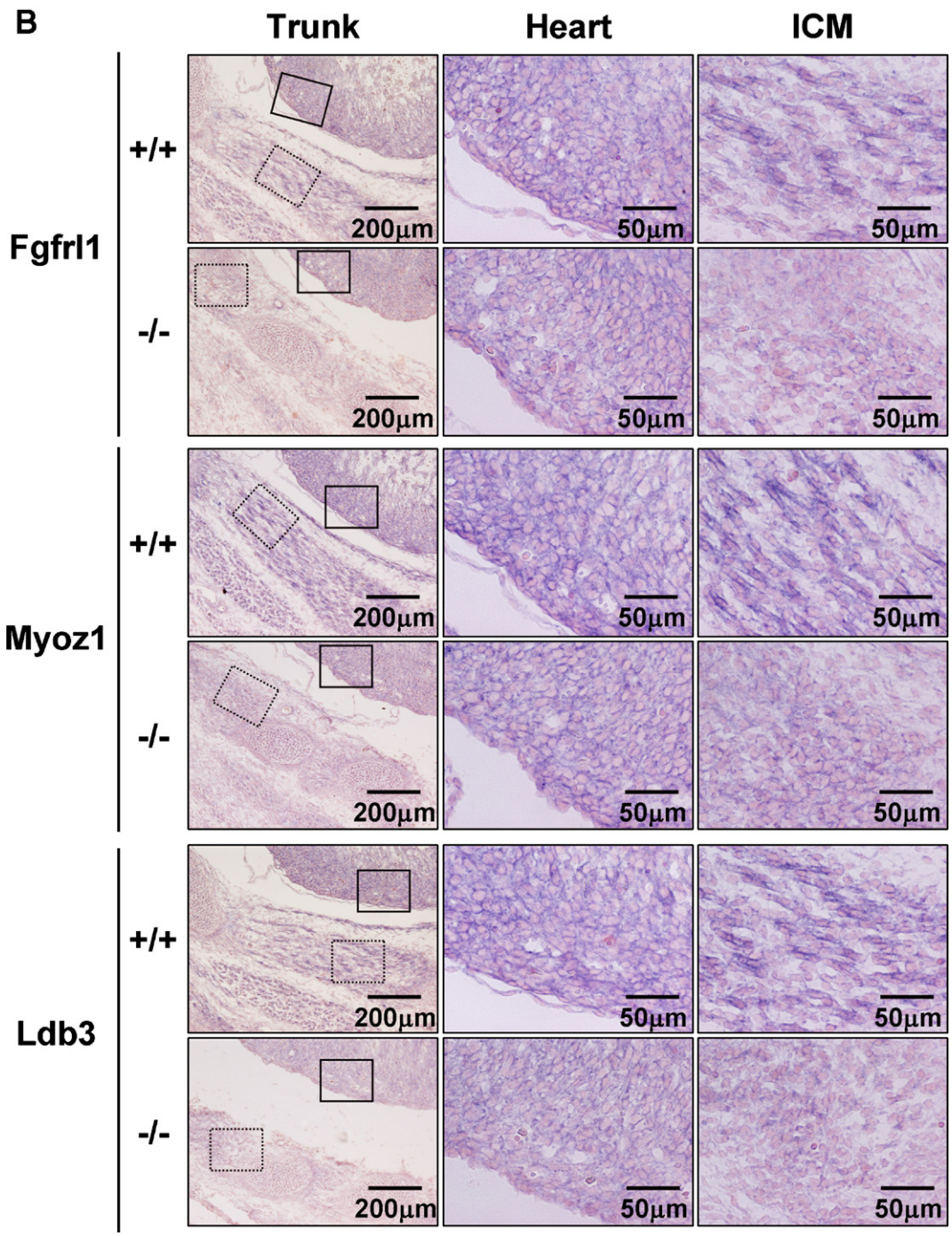

Fig. 4 (*continued*).

and MyoD served as a positive control (Wang et al., 2001), and a region from the ubiquitously expressed gene *Rpl7* encoding ribosomal protein 7 served as a negative control. We considered enrichment of chromatin-bound myogenin of fivefold or greater with respect to the negative control as evidence of meaningful myogenin binding.

The ChIP assay results showed that myogenin binding was associated with regulatory regions of 8 of the 10 genes that we analyzed (Fig. 6A). Myogenin did not bind at upstream regions from *Myoz1* and *Crsp3* and first intron regions from *Ldb3* and *RIKEN1100001E04* (Fig. 6A). In some genes, myogenin was bound to an E-box element within one conserved region but not another. For example, myogenin bound efficiently to the *Ldb3* upstream conserved region but not to the *Ldb3* conserved region within the first intron (Fig. 6B). The ChIP analysis provided direct evidence that *Ampd1*, *Casq2*, *Ttn*, *Fgf6*, *Fgfrl1*, *Ldb3*, *Lmod2*, and *RIKEN1100001E04* are all direct target genes of myogenin.

The identification of conserved regulatory regions with myogenin-binding sites also provided the opportunity to determine whether MyoD played a role in regulating the expression of myogenin-dependent genes within these regions.

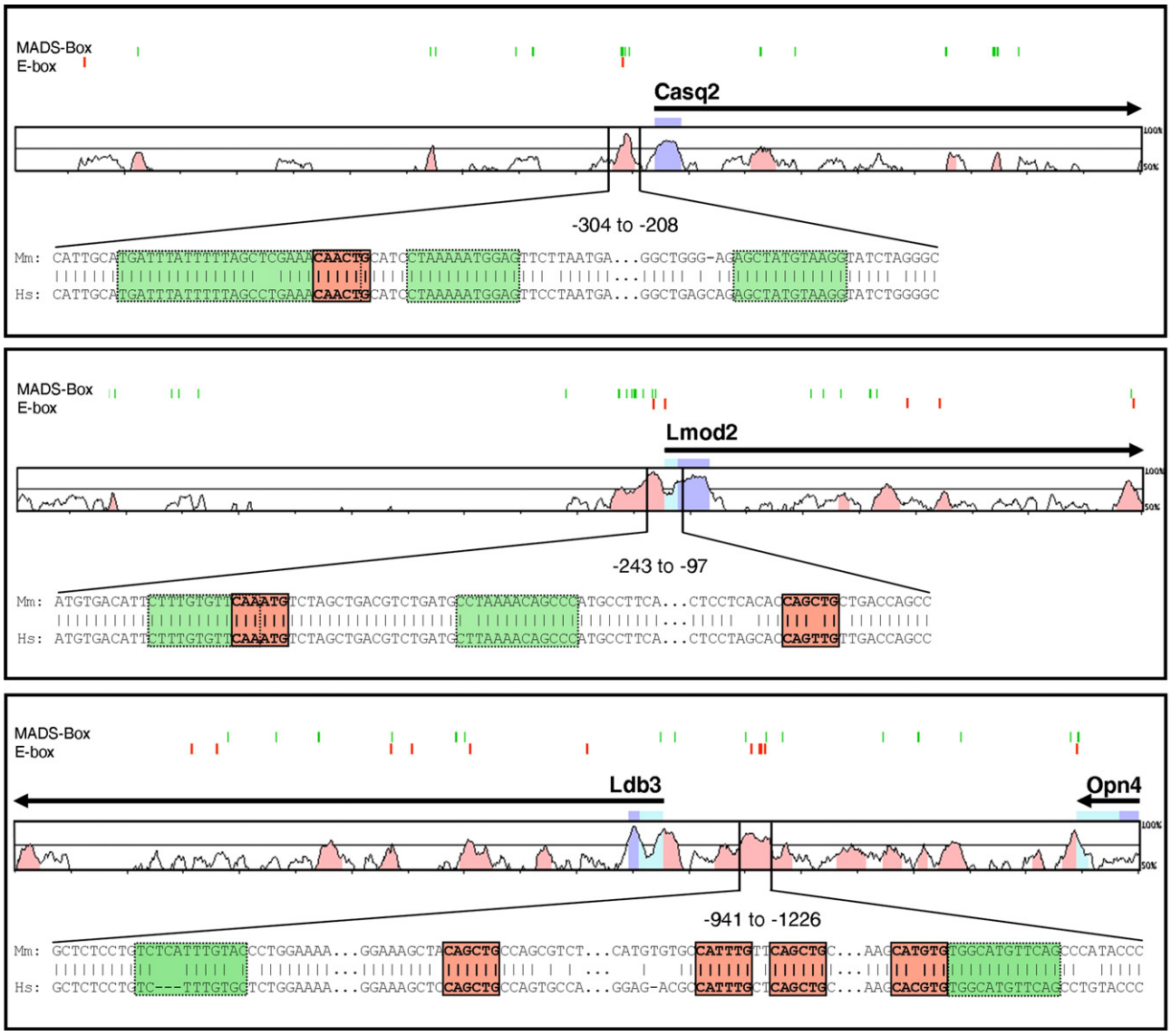

Fig. 5. Myogenin-dependent genes have conserved E-box- and Mef2-binding site motifs in their putative regulatory regions. Examples of rVISTA plots for *Casq2*, *Lmod2*, and *Ldb3* are shown. We identified conserved noncoding sequences and consensus binding sites using the VISTA genome browser and rVISTA, respectively. Phylogenetically conserved noncoding sequences in mouse and human genomes are depicted as pink regions, conserved untranslated regions are depicted as blue regions, and conserved coding regions are depicted as purple regions. The green vertical lines indicate conserved Mef2 (MADS-Box)-binding sites, and the vertical pink lines indicate the myogenin/MyoD E-Box)-binding sites. The gene orientation is indicated by thick arrowed lines. The regions containing clusters of conserved binding sites are positioned immediately upstream of the transcriptional start site for *Casq2*, *Lmod2*, and *Ldb3* and indicated by brackets. The partial sequences of the bracketed regions are shown below the rVISTA plots. The Mef2 MADS-box-binding sites are boxed in green and the myogenin/MyoD E-box-binding sites are boxed in pink. The numbers represent the base pair coordinates with −1 assigned the base pair immediately adjacent to the translational start A(TG) for each gene.

As hypothesized by Cao et al. (2006), MyoD may function in remodeling chromatin rather than functioning directly as a transcriptional activator. Accordingly, we repeated the ChIP experiments with antibodies against both myogenin and MyoD and E15.5 tongue muscle extracts. Control experiments showed that myogenin and MyoD were bound to bHLH E-box elements within the regulatory site in *mef2c* (Fig. 6A). In contrast, the upstream bHLH E-box elements within the conserved regions in *Ampd1*, *Casq2*, *Fgf6*, *Fgfrl1*, *Ldb3*, and *Lmod2* were ineffective binding sites for MyoD. Of the 10 myogenin target genes analyzed using ChIP, only conserved regions in *RIKEN1100001E04* and *Ttn* had significant MyoD binding (Fig. 6A). These results support the hypothesis that myogenin is able to activate transcription from E-boxes within the putative regulatory regions of these target genes without simultaneous input from MyoD.

Transcriptional activity of putative regulatory regions of myogenin target genes

The conserved regulatory regions associated with the myogenin target genes are potentially useful for determining the basis of myogenin's transcriptional selectivity. We sought to determine whether the putative regulatory sites that bound to

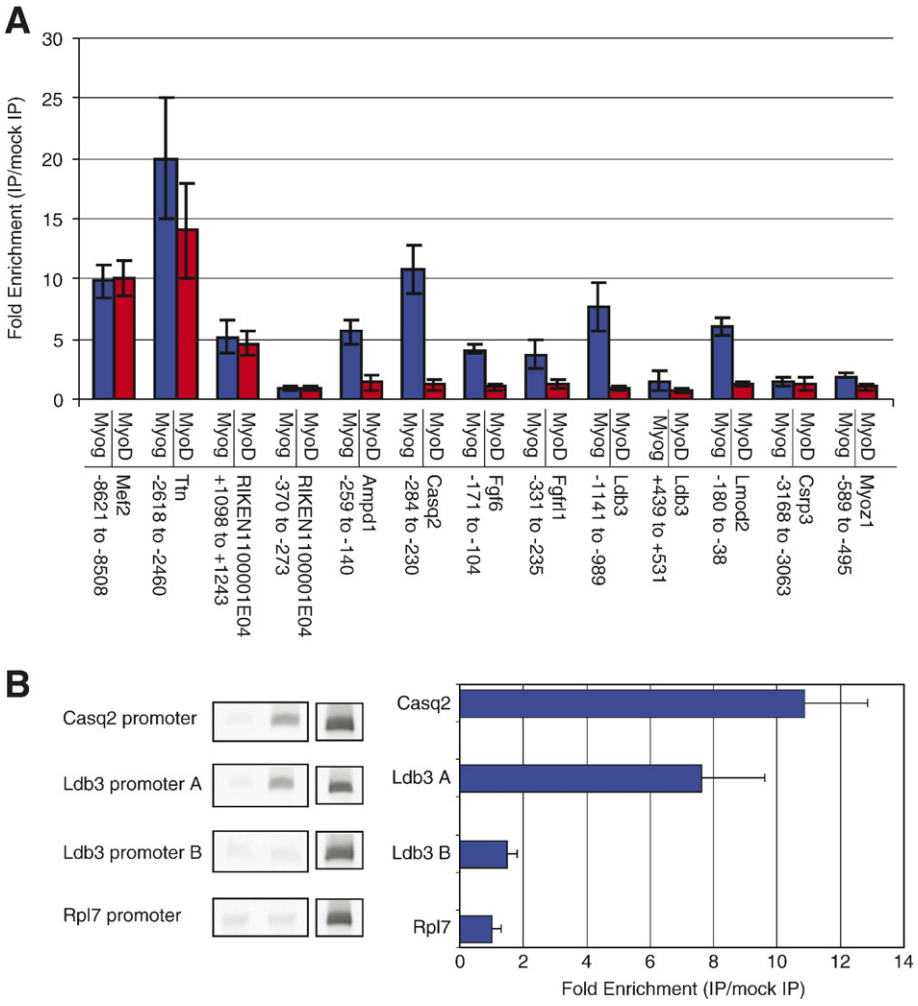

Fig. 6. ChIP analysis of myogenin and Myod using chromatin extracted from E15.5 tongue tissue samples. (A) Myogenin binds to conserved regions containing E-boxes. The ChIP results are graphed as fold enrichment as indicated on the *Y* axis for the antibody samples versus the no-antibody control samples normalized to *Rpl7*. The data represent the mean of three independent experiments, with each PCR performed in duplicate. The base pair positions shown for each gene name represent the primer sites in reference to the translational A(TG) start sites. Blue bars represent the ChIP results for myogenin, and red bars represent the ChIP results for MyoD. Error bars are standard deviations from the mean. (B) ChIP analysis of conserved regulatory regions in *Casq2*, *Ldb3*, and *mef2*. The left panel shows representative agarose gel profiles for the regulatory regions of the indicated genes using the M-225 anti-myogenin antibody, and the right panel shows fold enrichment for the regulatory regions calculated from qPCR values using the same DNA samples.

myogenin were in fact capable of activating transcription of a reporter gene in transient transfection assays. Regulatory regions of *Ampd1*, *Ldb3*, *Lmod2*, and *Casq2* were cloned into pGL3 lacking a eukaryotic promoter.

We first introduced these constructs into C2C12 myoblasts under either proliferating or differentiating conditions. In proliferating C2C12 myoblasts, *myoD* is expressed at high levels but *myog* is not expressed (Cornelison and Wold, 1997; Shen et al., 2003); after differentiation for 48 h, both *myoD* and *myog* are expressed at high levels (Cornelison and Wold, 1997; Shen et al., 2003). In proliferating cells, we detected reporter gene expression using regulatory regions of *Ampd1*, *Casq2*, and *Lmod2*, although the activity of these regulatory regions was only modestly enhanced over that of empty vector controls (Fig. 7A). We found somewhat higher reporter gene activity using the regulatory region of *Ldb3*. In contrast, after differentiation for 48 h, the transcriptional activity of all the regulatory regions was greatly enhanced; reporter gene expression was 3-fold to 31-fold higher in differentiating cells than in proliferating cells for the regulatory regions that we tested (Fig. 7A). Because MyoD is present in both proliferating and differentiating C2C12 cells, the results suggested that myogenin was specifically required for the observed enhancement of transcriptional activity.

In another set of experiments, we used NIH-3T3 fibroblasts, which do not express *myoD* or *myogenin*, in transactivation assays. In these experiments, we co-transfected expression constructs containing either *myoD* or *myog* along with the regulatory region reporter gene constructs to determine whether overexpression of *myoD* or *myog* could stimulate transcription from the regulatory regions of genes we identified by rVISTA. Introducing the regulatory region reporter gene constructs alone into NIH-3T3 cells resulted in low but detectable reporter gene expression in all cases (Fig. 7B). When we induced co-

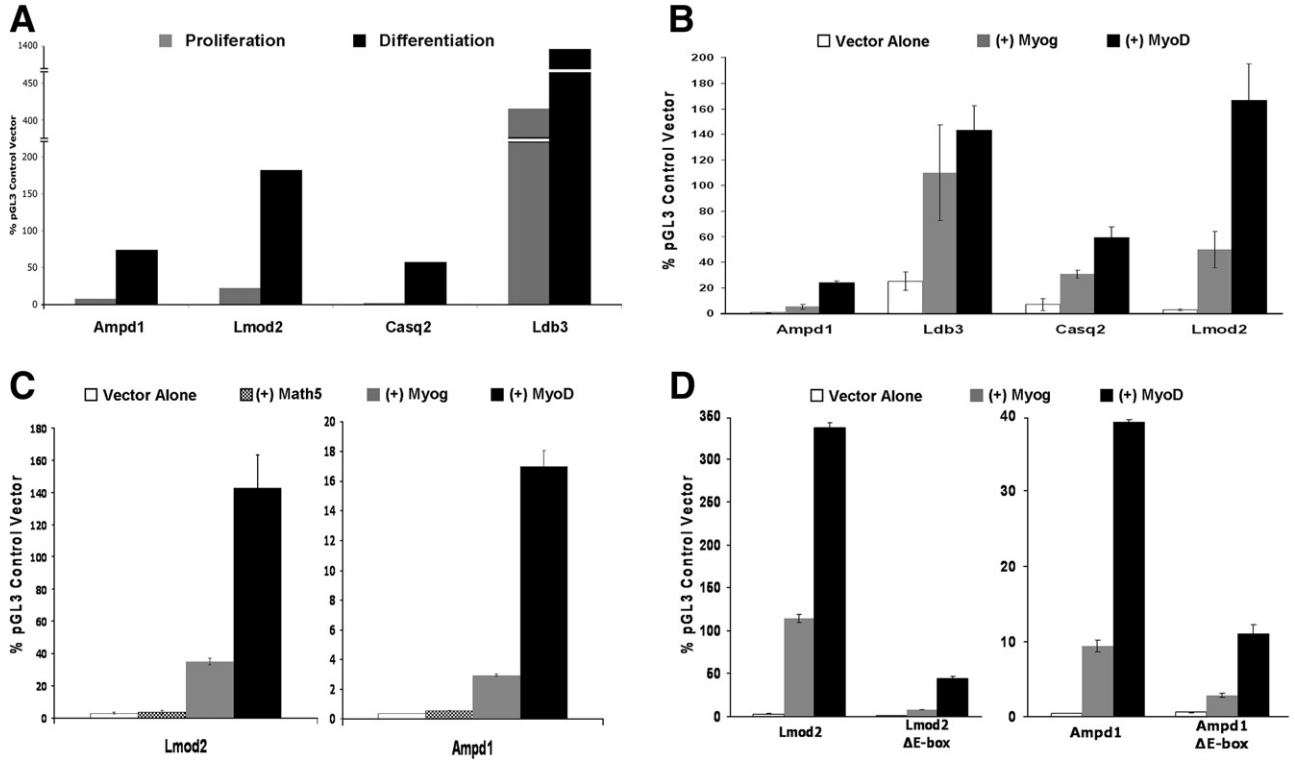

Fig. 7. The putative regulatory regions in the myogenin-dependent genes, *Ampd1*, *Ldb3*, *Casq2*, and *Lmod2* are able to activate expression of the *luciferase* reporter gene. The sequences of these regulatory regions are shown in Fig. S1. (A) Transient transfections of regulatory regions into proliferating and differentiating C2C12 myoblasts. Gray bars show the activity in proliferating myoblasts, and black bars show the activity in differentiating myoblasts. (B–D) Transient transfections of regulatory regions into NIH-3T3 cells co-transfected with no expression vector or expression vectors containing either myogenin or MyoD. The data for both C2C12 and NIH-3T3 cells are expressed as a percentage of the pGL3 control vector expression levels. Gray bars show activity when the *myog* expression construct is co-transfected, black bars show the activity when the *myoD* construct is co-transfected, and white bars show the activity when no expression construct is co-transfected. In panels C and D, activities of the regulatory regions from *Lmod2* and *Ampd1* are shown. In panel C, the dashed bar shows the activity when the *math5* construct is co-transfected. In panel D, ΔE-box refers to the regulatory regions from *Lmod2* and *Ampd1* in which the myogenin/MyoD E-box-binding sites were deleted. Error bars are standard deviations from the mean.

expression of either *myoD* or *myog* with the regulatory region reporter gene constructs, the transcriptional activity was stimulated from 4-fold to 56-fold depending on the regulatory region (Fig. 7B). However, both MyoD and myogenin were able to stimulate transcription in these transactivation experiments; in fact, MyoD was generally more effective than was myogenin. Transfections using concentrations of *myoD* or *myog* expression constructs ranging from 0.05 μg to 1.0 μg showed that MyoD was more effective at stimulating reporter gene transcription at all concentrations (data not shown). Furthermore, when both *myoD* and *myog* constructs were transfected together, we observed reporter gene activity that was approximately the average value of the two constructs transfected separately (data not shown). The results indicated that the transcriptional selectivity that is observed in embryonic tongue muscle is not recapitulated in transiently transfected NIH-3T3 fibroblasts.

The consensus E-box elements in the regulatory regions might be expected to bind to non-myogenic bHLH transcription factors and activate transcription from these E-box sites as effectively as MyoD or myogenin. To determine whether bHLH factors unrelated to MyoD and myogenin could activate transcription, we used a distantly related bHLH transcription factor, Math5 (Mu et al., 2005); a *math5* expression construct was co-transfected with the regulatory regions from *Lmod2* and *Ampd1*. Co-expressing *math5* with either *Lmod2* or *Ampd1* regulatory region reporter gene constructs did not result in significant transcriptional activation above the empty vector control, although co-expression with either *myoD* or *myog* expression constructs stimulated transcription as expected (Fig. 7C). These results indicated that the regulatory regions from *Lmod2* and *Ampd21* were able to selectively respond to myogenic bHLH regulatory factors.

To determine whether the E-box elements in the *Lmod2* or *Ampd1* regulatory regions were required for MyoD-mediated or myogenin-mediated activation, we generated constructs with deleted E-box elements; in *Lmod2* and *Ampd1*, a single E-box element was deleted (see Fig. S1 for E-box sequences). The deleted E-box reporter gene constructs responded only weakly when co-expressed with *myoD* or *myog* expression constructs, although reporter gene activity was not reduced to the background levels of the empty vector control (Fig. 7D). Residual activity might be explained by the presence of Mef2 sites, which are known to bind Mef2 factors in physical association with MyoD and myogenin (Berkes and Tapscott, 2005). These results demonstrate that the E-box elements within the *Lmod2* and *Ampd1* regulatory regions are the major *cis*-regulatory elements

required for MyoD- and myogenin-mediated transcriptional activation.

Discussion

Our results identify new target genes for myogenin and demonstrate that in *myog*$^{-/-}$ embryonic tongue muscle, the presence of normal levels of MyoD cannot compensate for the absence of myogenin in regulating the expression of these target genes. Moreover, binding sites within the upstream regulatory regions of these target genes have myogenin but not MyoD bound to them in vivo. Our transient transfection results with NIH-3T3 cells suggest that the DNA sequences within the regulatory regions of the myogenin target genes cannot be solely responsible for the transcriptional selectivity that we observe in vivo.

The gene expression profiling and ChIP experiments we performed are the first to use embryonic skeletal muscle to identify potential target genes for myogenin. Our experiments used skeletal muscle extracted from *myog*$^{+/+}$ and *myog*$^{-/-}$ embryos. This genetically based approach necessarily identifies genes whose expression strictly depends on the presence of myogenin and should thus identify genes whose expression is directly regulated by myogenin as well as genes whose expression is regulated by indirect mechanisms that require the presence of myogenin. The myogenin-dependent genes that we identified are likely to be bona fide targets of myogenin as evidenced by significant decreases in their expression in *myog*$^{-/-}$ tongue muscle, strong correlations between gene function and the *myog*$^{-/-}$ phenotype, and in vivo binding of myogenin to sites within putative regulatory regions. The target genes represent a wide variety of genes expressed in skeletal muscle with diverse functions and transcriptional regulatory features. They should prove useful in ongoing investigations of the mechanisms by which myogenin exerts its specialized functions as a myogenic regulatory factor.

Recently, several groups have identified potential MyoD and myogenin target genes in cultured myoblasts using gene profiling and ChIP-on-chip approaches (Blais et al., 2005, De la Serna et al., 2005; Ishibashi et al., 2005; Cao et al., 2006). Unexpectedly, the myogenin-dependent genes identified in our gene expression profiling analysis of E15.5 embryonic tongue muscle do not significantly overlap with the genes identified in cultured myoblasts (Table S1) (Blais et al., 2005; De la Serna et al., 2005; Ishibashi et al., 2005; Cao et al., 2006). Moreover, there is little concordance among the cell culture studies on the target genes that were identified. Using C2C12 myoblasts sampled at different stages of myogenesis, Cao et al. (2006) reported a large overlap of target genes between MyoD and myogenin. In contrast, Blais et al. (2005), also using differentiating C2C12 myoblasts, did not find substantial overlap of gene targets among MyoD, myogenin, and Mef2. In our analysis, we specifically selected genes whose expression required the presence of myogenin, irrespective of the presence of MyoD. Furthermore, our ChIP analysis readily detected myogenin binding within the putative regulatory regions of some of these genes but detected MyoD binding only infrequently within these same regulatory regions. These results suggest that the myogenin-dependent genes identified in our analysis are distinct from those identified previously in cultured myoblasts and may therefore be subject to different modes of regulation.

Interestingly, the gene set of 140 downregulated genes that we identified is largely overlapped with that identified in a microarray analysis by Kuninger et al. (2004), who used insulin-like growth factor (IGF)-I to induce myogenesis in C2C12 cells. In particular, Kuninger and colleagues found that expression of many of the genes encoding structural and enzymatic components of muscle that we identified as myogenin-dependent genes was upregulated upon IGF-I induction of myogenesis. These similarities suggest that in most cases, IGF-I signaling is mediated through activation of *myog* and that the subsequent expression of genes is transcriptionally regulated by myogenin. Indeed, the skeletal muscle phenotype of *igf1*-knockout mice is very similar to that of *myog*-knockout mice (Benito et al., 1996).

The myogenin-dependent genes that we identified are only a subset of genes expressed in skeletal muscle; the expression of many genes encoding other structural and enzymatic components of skeletal muscle is not significantly downregulated in *myog*-null tongue muscle. The expression of these genes may not require the action of any of the myogenic regulatory factors, including myogenin. Alternatively, MyoD or Myf5, which are both present in *myog*$^{-/-}$ tongue muscle, may compensate for loss of myogenin. Our results show that myogenin is strictly required in vivo for the normal expression of one set of genes; without the expression of this gene set, myocyte differentiation and myofiber formation are severely compromised. Elucidating the mechanisms that distinguish myogenin-dependent genes from myogenin-independent genes will add substantially to a more precise understanding of gene regulation in skeletal muscle development.

At E15.5, myogenesis in the embryonic tongue is heterogeneous; committed myoblasts are in the process of forming myocytes and some of these myocytes are beginning to fuse to form myofibers (Yamane, 2005). Unlike the case in which C2C12 myoblasts are induced to differentiate into myotubes, myogenesis in embryonic tongue tissue is not synchronous. Therefore, it is difficult to group the myogenin-dependent genes that we identified into early and late expressing genes. In fact, the myogenin-dependent genes identified in our analysis apparently are expressed at various stages of myogenesis. For example, *Fgf6* and *Fgfrl1* are expressed throughout myogenesis (Trueb et al., 2003; Armand et al., 2006), whereas genes encoding sarcomeric Z-disk components like *Myoz1* and *Myoz2* are expressed at late stages of myogenesis (Takada et al., 2001).

Cao et al. (2006) have presented evidence suggesting that myogenin does not bind efficiently to target genes without MyoD. Given the heterogeneity of expression of myogenin-dependent genes in tongue muscle, one would expect to find MyoD and myogenin bound to sites within the same regulatory regions using ChIP analysis. However, our results indicate that this occurs only infrequently. Because myogenin is thought to be unable to efficiently bind to its target site without MyoD-

mediated chromatin remodeling (Berkes and Tapscott, 2005; Tapscott, 2005; Cao et al., 2006), our results suggest that MyoD acts at an earlier embryonic stage in tongue muscle than E15.5. Of course, some of the myogenin-dependent genes that we have identified may never require MyoD for their expression in skeletal muscle, and other factors may be used for chromatin remodeling and myogenin's access to regulatory DNA sequences.

Our transient transfection experiments made use of the putative regulatory regions in myogenin-dependent genes to determine whether exogenously introduced DNA templates would behave in a manner similar to that observed in vivo. That is, would the intracellular environment in cultured C2C12 myoblasts or NIH-3T3 fibroblasts provide a means for distinguishing between MyoD and myogenin? Our tentative hypothesis was that if the DNA sequences within the regulatory regions in these genes were responsible for selective binding of myogenin rather than MyoD, then the transient transfection analysis would reflect that selectivity. We would therefore expect MyoD to be incapable of binding and activating transcription but expect myogenin to perform these functions. Because *myog* is expressed at high levels in differentiating myoblasts but is not expressed in proliferating myoblasts, our results imply that myogenin is responsible for the enhanced activity of myogenin-dependent genes. However, because *myoD* is also expressed in differentiating myoblasts, we are not certain whether the regulatory regions are responding only to myogenin and not to MyoD. In fact, in NIH-3T3 cells, in which *myog* and *myoD* were overexpressed, we found no evidence of sequence selectivity for myogenin. Our results suggest that the cellular environment is a crucial factor in determining myogenin's target gene selectivity. This environment could reflect a muscle cell-specific chromatin organization, in which myogenic cells use factors like MyoD to remodel chromatin. Alternatively, transcriptional co-factors may be essential in myogenesis to promote myogenin's selective activity after it binds to its target site.

Acknowledgments

We thank Xiuqian Mu and Girard Courteau Jr. for assisting with the microarray analysis and Roberto Reynaga for technical assistance. This work was supported by a grant from the Muscular Dystrophy Association (MDA-4289 to W.H.K.) and by the Robert A. Welch Foundation (G-0010 to W.H.K.). It was also supported by the National Cancer Institute Comprehensive Cancer Center Support Grant (CA16672), supporting the DNA Analysis, Genomics, and Research Animal Support Facilities.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2007.08.014.

References

Armand, A.S., Laziz, I., Chanoine, C., 2006. FGF6 in myogenesis. Biochim. Biophys. Acta 1763, 773–778.

Benito, M., Valverde, A.M., Lorenzo, M., 1996. IGF-I: a mitogen also involved in differentiation processes in mammalian cells. Int. J. Biochem. Cell Biol. 28, 499–510.

Bergstrom, D.A., Tapscott, S.J., 2001. Molecular distinction between specification and differentiation in the myogenic basic helix–loop–helix transcription factor family. Mol. Cell. Biol. 21, 404–412.

Berkes, C.A., Tapscott, S.J., 2005. MyoD and the transcriptional control of myogenesis. Semin. Cell Dev. Biol. 16, 585–595.

Black, B.L., Olson, E.N., 1998. Transcriptional control of muscle development by myocyte enhancer factor-2 (MEF2) proteins. Annu. Rev. Cell Dev. Biol. 14, 167–196.

Blais, A., Tsikitis, M., Acosta-Alvear, D., Sharan, R., Kinger, Y., Dynlacht, B.D., 2005. An initial blueprint for myogenic differentiation. Genes Dev. 19, 553–569.

Buckingham, M., 2001. Skeletal muscle formation in vertebrates. Curr. Opin. Genet. Dev. 11, 440–448.

Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meilhac, S., Montarras, D., Rocancourt, D., Relaix, F., 2003. The formation of skeletal muscle from somite to limb. J. Anat. 202, 59–68.

Cao, Y., Kumar, R.M., Penn, B.H., Berkes, C.A., Kooperberg, C., Boyer, L.A., Young, R.A., Tapscott, S.J., 2006. Global and gene-specific analyses show distinct roles for Myod and Myog at a common set of promoters. EMBO J. 25, 502–511.

Cheng, T.C., Hanley, T.A., Mudd, J., Merlie, J.P., Olson, E.N., 1992. Mapping of myogenin transcription during embryogenesis using transgenes linked to the myogenin control region. J. Cell Biol. 119, 1649–1656.

Conley, C.A., Fritz-Six, K.L., Almenar-Queralt, A., Fowler, V.M., 2001. Leiomodins: larger members of the tropomodulin (*Tmod*) gene family. Genomics 15, 127–139.

Cornelison, D.D., Wold, B.J., 1997. Single-cell analysis of regulatory gene expression in quiescent and activated mouse skeletal muscle satellite cells. Dev. Biol. 191, 270–283.

de la Serna, I.L., Ohkawa, Y., Berkes, C.A., Bergstrom, D.A., Dacwag, C.S., Tapscott, S.J., Imbalzano, A.N., 2005. MyoD targets chromatin remodeling complexes to the myogenin locus prior to forming a stable DNA-bound complex. Mol. Cell. Biol. 25, 3997–4009.

Dubchak, I., Ryaboy, D.V., 2006. VISTA family of computational tools for comparative analysis of DNA sequences and whole genomes. Methods Mol. Biol. 338, 69–89.

Faulkner, G., Pallavicini, A., Formentin, E., Comelli, A., Levolella, C., Trevisan, S., Bortoletto, G., Scannapicco, P., Salamon, M., Mouly, V., Valle, G., Lanfranchi, G., 1999. ZASP: a new Z-hand alternatively spliced PDZ-motif protein. J. Cell Biol. 146, 465–475.

Fischer, R.S., Fowler, V.M., 2003. Tropomodulins: life at the slow end. Trends Cell Biol. 13, 593–601.

Hasty, P., Bradley, A., Morris, J.H., Edmondson, D.G., Venuti, J.M., Olson, E.N., Klein, W.H., 1993. Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. Nature 364, 501–506.

Hill, D.P., Blake, J.A., Richardson, J.E., Ringwald, M., 2002. Extension and integration of the gene ontology (GO): combining GO vocabularies with external vocabularies. Genome Res. 12, 1982–1991.

Ishibashi, J., Perry, R.L., Asakura, A., Rudnicki, M.A., 2005. MyoD induces myogenic differentiation through cooperation of its NH2- and COOH-terminal regions. J. Cell Biol. 171, 471–483.

Katz, S., Irizarry, R.A., Lin, X., Tripputi, M., Porter, M.W., 2006. A summarization approach for Affymetrix GeneChip data using a reference training set from a large, biologically diverse database. BMC Bioinformatics 7, 464.

Kirschner, M., Gerhart, J., 1998. Evolvability. Proc. Natl. Acad. Sci. U. S. A. 95, 8420–8427.

Klaavuniemi, T., Ylanne, J., 2006. Zasp/Cypher internal ZM-motif containing fragments are sufficient to co-localize with alpha-actinin—Analysis of patient mutations. Exp. Cell Res. 312, 1299–1311.

Kuninger, D., Kuzmichas, R., Peng, B., Pintar, J.E., Rotwein, P., 2004. Gene discovery by microarray: identification of novel genes induced during growth factor-mediated muscle cell survival and differentiation. Genomics 84, 876–889.

Lee, J.K., Cho, J.H., Hwang, W.S., Lee, Y.D., Reu, D.S., Suh-Kim, H., 2000.

Expression of neuroD/BETA2 in mitotic and postmitotic neuronal cells during the development of nervous system. Dev. Dyn. 217, 361–367.

Morisaki, T., Holmes, E.W., 1993. Functionally distinct elements are required for expression of the *AMPD1* gene in myocytes. Mol. Cell. Biol. 13, 5854–5860.

Mu, X., Fu, X., Sun, H., Beremand, P.D., Thomas, T.L., Klein, W.H., 2005. A gene network downstream of transcription factor Math5 regulates retinal progenitor cell competence and ganglion cell fate. Dev. Biol. 280, 467–481.

Myer, A., Olson, E.N., Klein, W.H., 2001. MyoD cannot compensate for the absence of myogenin during skeletal muscle differentiation in murine embryonic stem cells. Dev. Biol. 229, 340–350.

Nabeshima, Y., Hanaoka, K., Havasaka, M., Esumi, E., Li, S., Nonaka, I., Nabeshima, Y., 1993. Myogenin gene disruption results in perinatal lethality because of severe muscle defect. Nature 364, 532–535.

Ohkawa, Y., Marfella, C.G., Imbalzano, A.N., 2006. Skeletal muscle specification by myogenin and Mef2D via the SWI/SNF ATPase Brg1. EMBO J. 25, 490–501.

Ohkawa, Y., Yoshimura, S., Higashi, C., Marfella, C.G., Dacwag, C.S., Tachibana, T., Imbalzano, A.N., 2007. Myogenin and the SWI/SNF ATPase Brg1 maintain myogenic gene expression at different stages of skeletal myogenesis. J. Biol. Chem. 282, 6564–6570.

Olson, E.N., Klein, W.H., 1994. bHLH factors in muscle development: dead lines and commitments, what to leave in and what to leave out. Genes Dev. 8, 1–8.

Pollard, J.W., Walker, J.M., 1997. Basic Cell Culture Protocols, Second Edition. Humana Press, Totawa, NJ.

Rawls, A., Morris, J.M., Rudnicki, M., Braun, T., Arnold, H.H., Klein, W.H., Olson, E.N., 1995. Myogenin's functions do not overlap with those of MyoD or Myf-5 during mouse embryogenesis. Dev. Biol. 172, 37–50.

Shen, X., Collier, J.M., Hlaing, M., Zhang, L., Delshad, E.H., Bristow, J., Bernstein, H.S., 2003. Genomic-wide examination of myoblast cell cycle withdrawal during differentiation. Dev. Dyn. 226, 128–138.

Takada, F., Vander Woude, D.L., Tong, H.Q., Thompson, T.G., Watkins, S.C., Kunkel, L.M., Beggs, A.H., 2001. Myozenin: an alpha-actinin- and gamma-filamin-binding protein of skeletal muscle Z lines. Proc. Natl. Acad. Sci. U. S. A. 98, 1595–1600.

Tapscott, S.J., 2005. The circuitry of a master switch: Myod and the regulation of skeletal muscle gene transcription. Development 132, 2685–2695.

Trouche, D., Masutani, H., Groisman, R., Robin, P., Lenormand, H., Harei-Bellan, A., 1995. Myogenin binds to and represses c-fos promoter. FEBS Lett. 361, 140–144.

Trueb, B., Taeschler, S., 2006. Expression of FGFRL1, a novel fibroblast growth factor receptor, during embryonic development. Int. J. Mol. Med. 17, 617–620.

Trueb, B., Zhuang, L., Taeschler, S., Wiedemann, M., 2003. Characterization of FGFRL1, a novel fibroblast growth factor (FGF) receptor preferentially expressed in skeletal tissues. J. Biol. Chem. 278, 33857–33865.

Venuti, J.M., Morris, J.M., Vivian, J.L., Olson, E.N., Klein, W.H., 1995. Myogenin is required for late but not early aspects of myogenesis during mouse development. J. Cell Biol. 128, 563–576.

Wang, D.Z., Valdez, M.R., McAnually, J., Richardson, J., Olson, E.N., 2001. The Mef2c gene is a direct transcriptional target of myogenic bHLH and MEF2 proteins during skeletal muscle development. Development 128, 4623–4633.

Wells, J., Farnham, P.J., 2002. Characterizing transcription factor binding sites using formaldehyde crosslinking and immunoprecipitation. Methods 26, 48–56.

Yamane, A., 2005. Embryonic and postnatal development of masticatory and tongue muscles. Cell Tissue Res. 322, 183–189.

Zhang, B., Wang, Z.X., Zheng, Y., 1997. Characterization of the interactions between the small GTPase Cdc42 and its GTPase-activating proteins and putative effectors. Comparison of kinetic properties of Cdc42 binding to the Cdc42-interactive domains. J. Biol. Chem. 272, 21999–22007.

Zhou, Q., Ruiz-Lozano, P., Martone, M.E., Chen, J., 1999. Cypher, a striated muscle-restricted PDZ and LIM domain-containing protein, binds to alpha-actinin-2 and protein kinase C. J. Biol. Chem. 274, 19807–19813.