uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Jeff Ishibashi

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Cellular and Molecular Medicine)

GRADE / DEGREE

Department of Cellular and Molecular Medicine

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Transcription Factors Providing Myogenic Identity to Adult Muscle Satellite Cells

TITRE DE LA THÈSE / TITLE OF THESIS

Michael Rudnicki

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

John Bell

Rashmi Kothary

Jeff Dilworth

Fabio Rossi

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Transcription Factors Providing Myogenic Identity to Adult Muscle Satellite Cells

Jeff Ishibashi

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for a Ph.D. degree in
Cellular and Molecular Medicine

Department of Cellular and Molecular Medicine
Faculty of Medicine
University of Ottawa

Your file Votre référence
ISBN: 978-0-494-49360-1
Our file Notre référence
ISBN: 978-0-494-49360-1

Reproduction Authorizations

Permission to reproduce manuscript in Appendix B:

Seale P, Ishibashi J, Scime A, Rudnicki MA. Pax7 Is Necessary and Sufficient for the Myogenic Specification of CD45(+):Sca1(+) Stem Cells from Injured Muscle. PLoS Biol (2004) 2:E130.

Reproduction Authorizations

Permission to reproduce manuscript in Appendix C:

Seale P, Ishibashi J, Holterman C, Rudnicki MA. Muscle satellite cell-specific genes identified by genetic profiling of MyoD-deficient myogenic cells. Dev Biol (2004) 275:287-300.

From: David, Natalie (ELS-OXF) [N.David@elsevier.com]
Sent: Wednesday, June 14, 2006 6:27 AM
To: Ishibashi, Jeff
Subject: Permission Request

Dear Jeff Ishibashi
We hereby grant you permission to reproduce the material detailed below in your thesis at no charge subject to the following conditions:

1. If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies.
2. Suitable acknowledgment to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows: "Reprinted from Publication title, Vol number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier".
3. Reproduction of this material is confined to the purpose for which permission is hereby given.
4. This permission is granted for non-exclusive world English rights only. For other languages please reapply separately for each one required. Permission excludes use in an electronic form. Should you have a specific electronic project in mind please reapply for permission.
5. Should your thesis be published commercially, please reapply for permission.

This includes permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.
Yours sincerely,
Natalie David
Senior Rights Assistant

Abstract

Satellite cells are an adult stem cell population that provide for the regeneration of skeletal muscle. Two major families of transcription factors, the paired-box (Pax) factors and the myogenic regulatory factors (MRFs), have essential roles in specifying that satellite cells become skeletal muscle. It was hypothesized that three of these proteins (Pax7, Myf5, MyoD) have particularly important roles in providing this identity, due to their unique roles in activating the skeletal muscle program. Each was studied using a combination of molecular techniques, including genome-wide microarrays, PCR-based systems, protein assays, and immunostaining, to assess their effects on myoblast gene expression and function. In the absence of Pax7, adult myoblasts are lost through a p53-dependent pathway. Pax7 also regulates the expression of numerous genes in myoblasts, including *Myf5*. Alternative splicing within the Pax7 paired domain alters its ability to activate expression of target genes, including *Myf5*. Although Myf5 and MyoD are both expressed in growing myoblasts and are grossly redundant, MyoD uniquely primes myoblasts for differentiation through its divergent NH_2- and COOH-terminal regions that flank the conserved bHLH domain. Therefore, Pax7, Myf5, and MyoD supply specific myogenic functions in satellite cells that provide for robust repair of injury to skeletal muscle tissue.

Contents

List of Tables

List of Figures

List of Abbreviations

bHLH	basic helix-loop-helix
ChIP	chromatin immunoprecipitation
COOH-	carboxy-
DAPI	4',6-diamidino-2-phenylindole
dblKO	double-knockout (*MyoD*$^{-/-}$;*Myf5*$^{-/-}$)
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
EGFP	enhanced green-fluorescent protein
FACS	fluorescence-activated cell sorting
FGF	fibroblast growth factor
FITC	fluorescein (isothiocyanate)
FKHR	FoxO1
Fluc	firefly luciferase
GAPDH	glyceraldyde-3-phosphate dehydrogenase
GL$^{+/-}$	glycine-leucine +/-
HGF	hepatocyte growth factor
MRF	myogenic regulatory factor
MyHC	myosin heavy chain
NH_2-	amino-
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde

$Q^{+/-}$	glutamine +/-
RDA	representational difference analysis
RIPA	radioimmunoprecipitation assay
Rluc	Renilla luciferase
RMS	rhabdomyosarcoma (alveolar, aRMS; embryonal, eRMS)
RNA	ribonucleic acid
RT	reverse-transcription
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Sp	*Splotch* (*Pax3*-null)
SP cells	side-population cells
UTR	untranslated region

Acknowledgements

Thanks to Dr. Michael Rudnicki for giving me the opportunity in his laboratory to learn to do scientific research. In particular, Rob Perry and Patrick Seale gave generously of their time to teach me. Nonetheless, appreciation is felt for all of the members and friends in the Rudnicki laboratory, present and past, who have contributed to this work in countless ways. Special mention, as well, of Mike Huh and Mark Gillespie for their support both as colleagues and as roommates; and Chet Holterman, who I've worked and played beside since the Guelph years. Fellow Guelph alumnus Jamie Tirone is remembered.

Finally, thanks to my family and friends for their support and encouragement throughout my lengthy education.

Chapter 1

General Introduction

1. Introduction

Skeletal muscle is an essential tissue that, as its most basic function, provides animals with the ability to move. This ability has been extremely important during evolution and has allowed animals to develop the intricate forms and behaviors seen today. So essential has it been that muscle has evolved into a highly specialized and complex tissue with exquisite developmental regulation. However, the lifelong inevitability of injury also required that provisions be made for muscle repair and adaptation. As body plans became larger and more complicated, specialized cells were allocated for this purpose. The molecular mechanisms that produce those cells and orchestrate skeletal muscle regeneration are so far coarsely defined. However, understanding the genetic basis of those functions is an important goal that could have broad medical applications, for the ability to move is now itself an essential function.

1.1 Adult Skeletal Muscle & Satellite Cells

Mature skeletal muscle is a largely post-mitotic tissue composed of bundles of syncytial muscle fibres. During embryogenesis and early post-natal development, successive waves of proliferating myoblasts withdraw from the cell cycle and differentiate into myocytes that express numerous genes involved in muscle function. They align and fuse into multinucleated fibres, synthesizing and organizing the proteins required for contraction. Consequently, fibres develop their typical ultrastructure, with a striated sarcomeric cytoplasm and myonuclei excluded to the fibre periphery by the bulk

of the assembled myofibrils. Individual fibres are separated by their plasma membranes and a basal lamina, as well as the surrounding connective tissue.

Fibres are in turn collected into bundles surrounded by a perimysium layer, and those bundles are bundled into the overall muscle by a myomysium (reviewed in Charge and Rudnicki, 2004). Alternatively, myofibres can be grouped at a logical (rather than spatial) level based on their innervation. Whereas individual fibres are contacted by only a single axon at the motor endplate, one motor neuron may have multiple axon termini to innervate several fibres that are therefore organized as a motor unit. The force of a muscle contraction can thus be modified by the strength of stimulation of a particular motor neuron as well as by the progressive recruitment of increasing numbers of motor units.

Adjacent to skeletal muscle fibres, and sandwiched between the plasma membrane and basal lamina, are small, quiescent, mononuclear cells known as satellite cells (Mauro, 1961) (Figure 1). Whereas microtears and minor damage to the plasma membrane can be repaired directly, greater damage leading to necrosis is corrected in a progressive response involving the removal of damaged tissue and activation of quiescent satellite cells (reviewed in Grounds, 1998). Once activated, satellite cells proliferate, differentiate, and fuse into existing or new fibres in order to rebuild the muscle structure. Although alternative sources of stem cells have been shown to contribute to regenerated skeletal muscle, it is clear that the overwhelming majority of physiological repair is provided by satellite cells (Zammit et al., 2002).

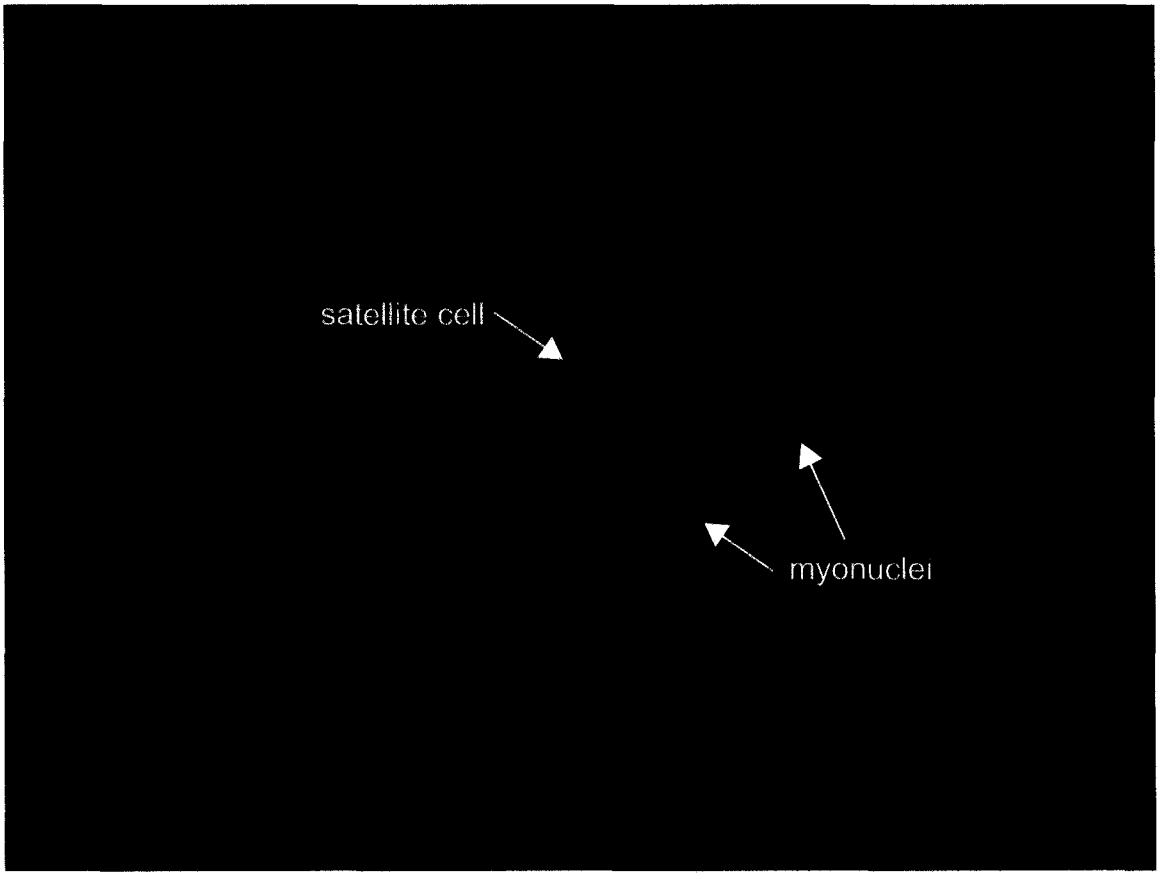


Figure 1. Isolated skeletal muscle fibre and satellite cell. A satellite cell (Pax7+) resides on the surface of the muscle fibre. Many myonuclei (Pax7-) are found along the length of the fibre. Background fluorescence of the fibre highlights the typical striated appearance of skeletal muscle fibres. Pax7+/FITC, green. DAPI, blue (satellite cell nucleus is hidden by overlaid green).

In the absence of trauma, skeletal muscle is a very stable tissue with little nuclear turnover during a lifetime (Decary et al., 1996). However, even when challenged with multiple inflictions of severe injury, satellite cells provide an extraordinary capacity to repair and replace the damaged fibres. Satellite cells constitute only 5% of nuclei in the adult; however, this percentage is maintained through the lifetime of the animal, even in the face of massive regenerative demands (reviewed in Bischoff, 1994; McGeachie and Grounds, 1995). This suggests that they have certain stem cell characteristics, including the ability to self-renew.

The potential practical implications of being able to manage and manipulate satellite cells are diverse. For example, muscle dystrophies are serious inborn diseases. The most critical problem of Duchenne muscular dystrophy, for instance, is the failure to adequately repair skeletal muscle damage, leading to degraded and eventual loss of muscle function. Augmentation of the satellite cell population could be used in treatment to alleviate and delay the loss of function. Satellite cells are also of primary interest for gene therapy-based cures that could replace the missing Dystrophin with a gene encoding a functional substitute. More generally, normal but ageing individuals generally suffer a progressive loss of muscle mass that might eventually be prevented or moderated with therapies to motivate or supplement the resident satellite cells. Similarly, muscle wasting (cachexia) is a common symptom associated with cancer, and a regenerative therapy for muscle would benefit such patients.

As an extensive but largely post-mitotic tissue, skeletal muscle is considered a target for other gene therapies, for which the fibres could act as factories to secrete

therapeutic products. As a stable tissue, this could allow for long-term treatment with a decreased likelihood of causing a hyperproliferative disorder, as can occur with bone marrow-based methods. It is also curious that, despite their many other similarities to striated muscle tissue, cardiac fibres do not possess a satellite cell population. The possibility of implanting satellite cells for cardiac regeneration has also been widely studied. Finally, rhabdomyosarcoma (RMS), a common childhood tumour with a poor prognosis, has many features characteristic of skeletal muscle gene expression. Although it is unclear whether satellite cells are the source of the tumours (Tiffin et al., 2003), the causative chromosomal translocations of the alveolar subtype involve a pair of Pax transcription factors that are unique to and required by satellite cells in the adult animal. These few examples clearly suggest that understanding how satellite cells work will have many applications.

1.2 Embryonic Myogenesis

The basic structure of skeletal muscle is patterned during embryogenesis (reviewed by Buckingham, 2001; Hollway and Currie, 2005; Tajbakhsh, 2003). Paraxial mesoderm is formed bilaterally adjacent to the neural tube in the mouse embryo. This mesoderm becomes progressively segmented along the rostral-to-caudal axis into balls of tissue known as somites. Through the influence of Shh, Wnt, and BMP signals emanating from the notochord, neural tube, surface ectoderm, and lateral plate mesoderm, the somites undergo a transition, partitioning into an underlying mesenchymal schlerotome (bones and cartilage precursors) and an overlying dermomyotome (skin and

muscle precursors). The anterior and posterior regions of the schlerotome give rise to the syndetome (tendon precursors) that will connect bone and muscle. The lateral lips of the dermomyotome migrate to its ventral surface, with the cells undergoing an epithelial-to-mesenchymal transition to form a distinct myotome beneath the dermotome. It is the myotome that eventually forms the skeletal muscles of the body and limbs.

A distinction exists between the dorsomedial and ventrolateral portions of the myotome (reviewed by Christ and Ordahl, 1995). The dorsomedial region forms the epaxial myotome and has responsibility for building the deep back muscles. In contrast, the ventrolateral region, or hypaxial myotome, contains the precursors for the muscles of the body wall. Additionally, hypaxial myoblasts adjacent to the nascent limb buds are induced to delaminate and migrate into the limbs to form the limb muscles. The disparity between these two myoblast pools is also evidenced at the molecular level by differing dependencies upon the important primary myogenic transcription factors Myf5 and MyoD (see Section 1.6; Kablar et al., 1998; Kablar et al., 1997).

Myogenesis proceeds through several waves of differentiation, beginning with an initial group of embryonic myoblasts (reviewed by Hollway and Currie, 2005). Embryonic myoblasts form the primary muscle fibres that serve as a scaffold for later growth. Secondary fetal myoblasts differentiate around the primary fibres and eventually constitute the majority of early post-natal skeletal muscle. Post-natal muscle growth is supported by the proliferation of satellite cells that reside adjacent to existing fibres within muscle tissue. These myoblasts fuse into existing myofibres or differentiate into new myofibres, providing for the considerable hyperplastic muscle growth observed

during the adolescent period. Finally, satellite cells have added responsibility for repairing damage incurred through injury or exercise during the extent of the animal's lifetime. In this, the satellite cell has a singularly important role in post-natal myogenesis.

1.3 Satellite Cells

Satellite cells originate from the central region of the somitic dermomyotome during embryogenesis and migrate into nascent skeletal muscle before taking up their typical positions adjacent to the muscle fibres (Gros et al., 2005; Relaix et al., 2005). In the neonatal mouse, approximately 20-30% of sub-laminar nuclei are satellite cells, although this is reduced to 5% in a neonatal flurry of myogenic differentiation (Bischoff, 1994). Thereafter, satellite cells have predominant responsibility for the regeneration of any moderate to severe damage to the skeletal muscle requiring the addition of nuclei. Normally quiescent in the adult, injury produces activation and mobilization of the satellite cell pool, whereupon they proliferate until sufficient myoblasts are available to repair the compromised muscle fibres (reviewed by Holterman and Rudnicki, 2005).

Although the somite-derived satellite cells constitute the vast majority of adult muscle precursors, several potential alternate sources of myoblasts have been demonstrated in experimental models. Dorsal aorta and embryonic blood vessels have been shown to give rise to myogenic cells strikingly similar to satellite cells (De Angelis et al., 1999). Bone marrow-derived stem cell populations can contribute to the satellite cell population, albeit at a very low frequency (LaBarge and Blau, 2002). Similarly,

resident side-population (SP) stem cells, which lack expression of any known markers of myogenesis, can become myoblasts in some circumstances (Asakura et al., 2002). Muscle-resident $CD45^{+}Sca1^{+}$ cells can also be converted to myoblasts and contribute to fibre regeneration (Seale et al., 2004). However, the relationship between these other populations and satellite cells is unclear. Knowledge of the molecular events controlling myogenic determination will be needed to understand how either somitic or alternative myoblasts arise.

Satellite cells are characterized by the expression of several molecular markers. The cell-surface marker CD34 is expressed by satellite cells, as are the cell-surface receptors Met and Syndecans-3 and -4, and the adhesion molecule M-cadherin (reviewed in Dhawan and Rando, 2005). Particularly specific markers are the transcription factors Pax7 and Pax3, which are uniquely expressed in quiescent satellite cells and proliferating myoblasts but are sharply down-regulated following differentiation (Relaix et al., 2005). Such markers also distinguish satellite cells from other adult stem cells. For instance, the muscle SP population lacks expression of known myogenic markers (Asakura et al., 2002). Similarly, $CD45^{+}Sca1^{+}$ cells are Pax7-negative until they have become myogenic (Seale et al., 2004). Satellite cells are therefore equipped with unique gene expression profiles related to their role in myogenesis.

1.4 Satellite Cell Quiescence, Activation, and Proliferation

Satellite cells are quiescent in undamaged adult skeletal muscle, expressing the molecular markers discussed in the previous section but lacking detectable protein

expression of any myogenic regulatory factors (MRFs; reviewed in Cornelison and Wold, 1997; Dhawan and Rando, 2005). They lurk beneath the basal lamina and next to the muscle fibres, with M-cadherin expression asymmetrically concentrated to the cell surface adjacent to the fibre membrane (Irintchev et al., 1994). Inactive hepatocyte growth factor (HGF) is sequestered in the extracellular matrix surrounding the fibre (Tatsumi et al., 1998). Both *CD34* and *MNF* are uniquely spliced in quiescent satellite cells to produce the truncated and β isoforms, respectively, although the functional consequence of this is unknown (Beauchamp et al., 2000; Garry et al., 2000). Several proteins are implicated in maintaining the quiescent (G_0) state, including the TGFβ super-family member MDF8 (myostatin), the E2F-binding pocket protein p130, and the cyclin-dependent kinase inhibitor p27 (Cao et al., 2003; McCroskery et al., 2003). Thus, satellite cells wait in reserve until called upon in response to damage.

Release of HGF from the matrix after injury stimulates satellite cell activation through the HGF receptor, Met (Tatsumi et al., 1998). Fibroblast growth factors (FGFs) are alternative stimuli, particularly FGF2, possibly acting through FGF-R1 or FGF-R4 (Cornelison et al., 2001; Yablonka-Reuveni et al., 1999). Notably, syndecans-3 and -4 are heparin sulphate proteoglycans that contribute to HGF and FGF signaling in skeletal muscle (Cornelison et al., 2004). Expression of the Delta-1 ligand on local and distant myofibres also triggers Notch receptor signaling that is required for the emergence from quiescence (Conboy et al., 2003; Conboy and Rando, 2002). In response to these signals, satellite cells up-regulate the expression of numerous genes and re-enter the cell cycle.

Changes in gene expression are detected very early in the activation process. Within 24 hours, M-cadherin increases, CD34 is switched to the full-length isoform, and MyoD or Myf5 is induced (Beauchamp et al., 2000; Cornelison and Wold, 1997). By 48 hours, MNF expression changes to the α isoform, most satellite cells co-express Myf5 and MyoD, and there is induction of myogenin and MRF4 (Cornelison and Wold, 1997; Garry et al., 2000). Expression of Myf5 and MyoD is characteristic of proliferating myoblasts, whereas myogenin and MRF4 are confined to myoblasts in the process of differentiating (Cornelison and Wold, 1997). Asymmetric distribution of Numb protein between daughter cells during early satellite cell divisions correlates with a fate choice, with $Numb^{+}$ progeny forming the proliferating group and $Numb^{neg}$ daughter cells returning to repopulate the satellite cell niche (Conboy and Rando, 2002). Following Notch-driven myoblast expansion and activation of MRF expression, the proliferating cells induce Numb to antagonize Notch signaling, which allows for MRF-driven myogenic differentiation (Conboy and Rando, 2002).

A decline in proliferation signals and an increase in Numb in the presence of high levels of MyoD induce expression of cyclin D3, the CDK inhibitor p21, the E2F-binding pocket protein Rb, and the differentiation myogenic factor myogenin (reviewed in Kitzmann and Fernandez, 2001). This leads to the cessation of the cell cycle and the onset of differentiation, in which Myogenin and MRF4 coordinate the expression of numerous muscle genes that act collectively to produce functional, contractile muscle fibres. Myoblast markers including Pax7, Myf5, and MyoD are down-regulated in

mature fibres, having contributed their important functions at the satellite cell and myoblast stages.

1.5 Paired-box (Pax) Genes in Myogenesis

Two of the nine paired-box transcription factors are known to have essential roles in myogenesis. Those two, Pax3 and Pax7, form a closely related subset within the *Pax* gene family, having both related complementary embryonic expression patterns and extremely similar genetic and molecular structures. For many years, the homozygous *Splotch* (*Sp*) mouse mutant was known to lack all appendicular skeletal muscle. The *Sp* mutation (heterozygous equivalent to the human Waardenburg syndrome) was later shown to disrupt the *Pax3* gene. More recently, the adult *Pax7*$^{-/-}$ mouse has been shown to lack histologically defined satellite cells and to have essentially no skeletal muscle repair (Kuang et al., 2006; Seale et al., 2000). The residual regeneration seen in the *Pax7*$^{-/-}$ mouse has been attributed to the minor population of Pax3-expressing cells that reside in the interstitial spaces between muscle fibres (Kuang et al., 2006). A proportion of Pax7/Pax3 co-expressing cells have also been identified, again pointing to heterogeneity within the satellite cell population (Relaix et al., 2005).

Pax7 and Pax3 are transcription factors containing three highly conserved domains. Each has a paired domain and a paired-type homeodomain, as well as an octapeptide motif in the intervening region between the two DNA binding domains (Figure 2). Across these three regions Pax7 and Pax3 are > 95% conserved, suggesting that they may provide related or overlapping functions, including the regulation of similar

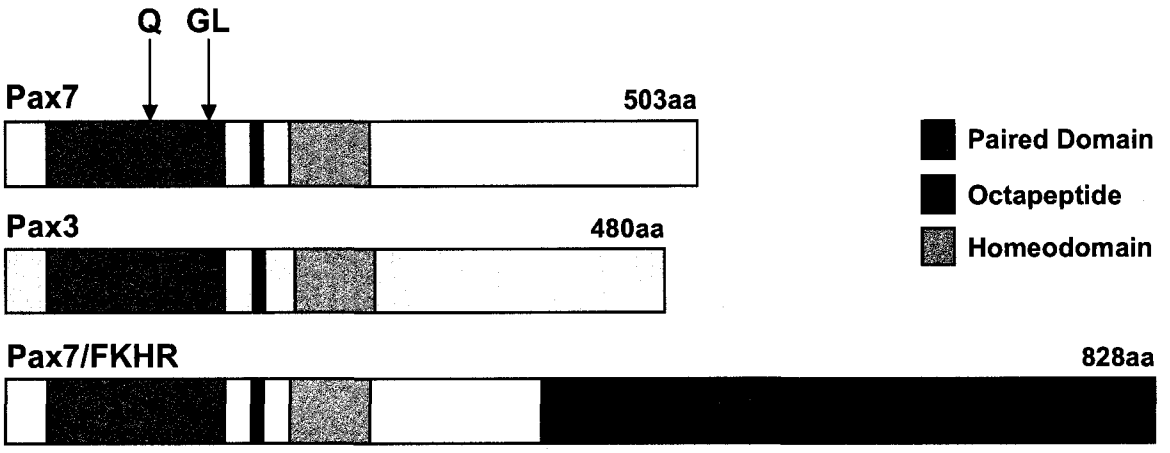


Figure 2. Schematic structures of the myogenic Pax factors. Pax7 and Pax3 exhibit >95% conservation between their paired, octapeptide, and homeodomains. Differential usage of splice acceptor sites at exons 3 and 4 of Pax7 result in the inclusion or exclusion of a glutamine residue or glycine-leucine dipeptide. The glutamine alternative splicing variation also occurs in Pax3. Alveolar rhabdomyosarcoma tumours typically exhibit one of two genomic translocations that produce a fusion of the Pax7 or Pax3 DNA binding domains to the FoxO1 (FKHR) transactivation domain.

target genes involved in myogenesis. This hypothesis has been partially borne out by replacement of *Pax3* with *Pax7*, which compensates for the loss of *Pax3* in neural tube, neural crest, and somite development (Relaix et al., 2004). However, the same experiments have demonstrated that there are also unique functions, as Pax7 cannot rescue the defect in appendicular muscle formation (Relaix et al., 2004). Therefore, those data combined with the distinct phenotypes of the $Pax7^{-/-}$ and $Pax3^{-/-}$ mice indicate that Pax7 and Pax3 are each singularly required.

Pax7 and Pax3 have additional functions in cell survival. Developing myoblasts undergo apoptosis during embryogenesis in $Sp^{-/-}$ mice, a defect that can be rescued by ectopic expression of *Pax3* (Borycki et al., 1999). Splotch-related neural tube defects can also be rescued by concurrent loss of p53 expression (Pani et al., 2002). Similarly, the absence of satellite cells in $Pax7^{-/-}$ skeletal muscle has been attributed to their apoptosis and is not prevented by Pax3 (Relaix et al., 2006). However, it is not known if these anti-apoptotic functions are provided through transcriptional regulation of target genes or through other mechanisms.

Through the selective usage of an alternate splice acceptor site at the beginning of exon 3, the paired DNA binding domains of both Pax7 and Pax3 can be modified to include or exclude a single glutamine residue ($Q^{+/-}$). Specific to Pax7 is a second paired-domain splicing variation at exon 4 that inserts or deletes a glycine-leucine dipeptide ($GL^{+/-}$). While the $Q^{+/-}$ change is known to affect the DNA binding affinity of Pax3 at some sites (Vogan et al., 1996), the *in vivo* functional consequences of these two variations in Pax7 have not been identified.

Pax7 and Pax3 are both associated with the development of rhabdomyosarcomas. In particular, the alveolar subtype generally exhibits one of two highly characteristic chromosomal translocations that fuse the carboxy-terminal transactivation domain of the FoxO1 (FKHR) transcription factor to the amino-terminal DNA binding domains of either Pax7 or Pax3 (Figure 2). *Pax7* expression has been additionally associated with the embryonal subtype of RMS (Tiffin et al., 2003). Notably, these tumours exhibit elements of the skeletal muscle phenotype, including expression of numerous genes (including the myogenic regulatory factors) that are markers for skeletal muscle. The transcriptional and cell survival functions of Pax7 and Pax3 are thus of keen interest for understanding and combating these cancers.

1.6 The Myogenic Regulatory Factors (MRFs)

A family of four basic helix-loop-helix (bHLH) transcription factors function as master regulatory switches for the skeletal muscle lineage (Figure 3). Their expression is unique to this tissue, and ectopic expression of any one of them can cause non-muscle cell types to convert to skeletal muscle. The primary MRFs, Myf5 and MyoD, are expressed in proliferating myoblasts and exhibit partial redundancy of function. Concurrent loss of both genes is a lethal defect resulting from the absence of nearly all skeletal muscle tissue. However, single mutation of *Myf5* or *MyoD* is a viable defect and produces a comparatively mild phenotype.

The secondary MRFs, Myogenin and MRF4, are expressed during the process of myoblast differentiation. Myogenin, in particular, is essential to establishing functional

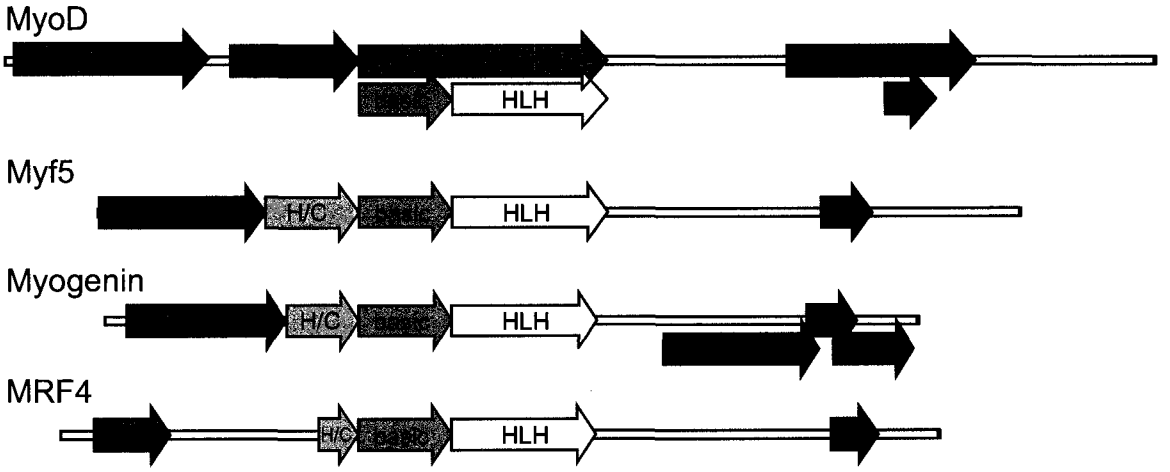


Figure 3. Schematic structures of the myogenic regulatory factors. Using MyoD as the prototype for comparison (Bergstrom & Tapscott, 2001), Myf5, Myogenin, and MRF4 also have highly conserved bHLH domains composed of a basic domain (DNA binding) and a helix-loop-helix motif (HLH; dimerization). The MyoD H/C (His/Cys) region is less conserved, but related sequences can be aligned in the other MRFs (Bergstrom & Tapscott, 2001). The C-terminal domain (CTD) in MyoD contains the important Helix III (H3) that is structurally conserved in the other MRFs, although myogenin's has diverged in function (Bergstrom & Tapscott, 2001). Transcriptional activation domains (AD) have been delineated for MyoD (Weintraub et al., 1991), Myf5 (Braun et al., 1990), Myogenin (Schwart et al., 1992), and MRF4 (Mak et al., 1992).

muscle fibres. MRF4 is expressed in mature fibres, but its specific role is unclear. It has been recently shown that MRF4 can substitute, in a very limited manner, for the loss of MyoD and Myf5 during the earliest stages of myogenesis (Kassar-Duchossoy et al., 2004). Collectively, these four genes activate the skeletal muscle genetic program and define the cells of this lineage.

Myf5 and MyoD are not expressed during quiescence but one or the other is up-regulated within 24 hours of the activation of a satellite cell. Shortly thereafter, both are expressed. It is currently believed that an asymmetric cell division occurs following activation, resulting in a daughter cell that down-regulates MRF expression and returns to the quiescent satellite cell niche. Myoblasts that continue to express Myf5 and MyoD proliferate until there are enough nuclei available to repair the damage, at which point they activate expression of Myogenin, differentiating and fusing into muscle fibres and expressing MRF4. Notably, *Myf5*-null myoblasts proliferate poorly and differentiate prematurely (Montarras et al., 2000). Conversely, *MyoD*-null myoblasts grow well and differentiate poorly, and *MyoD*-null mice are deficient in the repair of muscle injuries (Megeney et al., 1996; Sabourin et al., 1999). Myf5 has therefore been attributed a greater role in the proliferation of myoblasts while MyoD is appears to prepares them for effective differentiation.

Structurally, the four MRFs share a similar genomic organization, and the proteins have highly conserved 65 amino acid bHLH domains of which three specific residues encode myogenic specificity (Davis and Weintraub, 1992). The helix-loop-helix region permits dimerization with bHLH E-proteins such as E12, E47, or HEB, while the

basic domains of the heterodimers recognize E-box sites of the consensus sequence CAnnTG. An additional conserved alpha-helical domain (helix III) located in the C-terminal portion of each MRF is key to the detailed model of MyoD function developed by the Tapscott group. They have proposed that the promoters of inactive MyoD target genes are constitutively bound by Pbx/Meis factors, with which MyoD interacts through helix III (Berkes et al., 2004). This interaction recruits histone acetyl-transferase complexes to acetylate adjacent histones and MyoD itself (Dilworth et al., 2004). Acetylation leads to engagement of SWI/SNF chromatin remodeling complexes that provide access to DNA binding sites for the MRFs (and other transcription factors such as MEF2) and retention of RNA polymerase II for transcription (de la Serna et al., 2005). Notably, the Myogenin helix III acts more like a traditional activation domain and cannot substitute for that of MyoD in this sequence of events, whereas the Myf5 and MRF4 helix III's are more similar to that of MyoD than that of Myogenin (Bergstrom and Tapscott, 2001). In activated satellite cells (which do not express MRF4) this model therefore places Myf5 and MyoD in a key position upstream of Myogenin in providing myogenic specification.

1.7 Hypotheses and Specific Aims

Pax7, MyoD, and Myf5 are key factors in adult myogenesis due to their unique roles in activating the skeletal muscle program. These functions were explored in three segments.

1.7.1 Role of Pax7 in myoblast survival

In addition to its role in myogenic specification, Pax7 appears to have an anti-apoptotic function in myoblast survival, similar to that of Pax3 in embryonic myogenesis. If myoblast apoptosis is p53-dependent, then concurrent loss of *Pax7* and *p53* should prevent their death and allow $Pax7^{-/-}$ myoblasts to be isolated from *Pax7*-null skeletal muscle. Therefore, a $Pax7^{-/-};p53^{-/-}$ mouse line was created and myoblast cultures were derived from the skeletal muscles of those mice.

1.7.2 Pax7 Target Genes in Myoblasts

Paired-box proteins such as Pax7 are DNA binding transcription factors that regulate the expression of downstream genes. Therefore, identification of those target genes should define how Pax7 functions in the myogenic specification of satellite cells. Ectopic expression of Pax7 in a cell line providing the appropriate co-factors for activity allowed for a genome-wide expression comparison between $Pax7^{high}$ and $Pax7^{low}$ myoblasts. Characterization of the differentially expressed genes illuminated a molecular model for Pax7 function in adult myogenesis.

1.7.3 Differences Between Myf5 and MyoD in Adult Myogenesis

Myf5 and MyoD are together essential for myogenesis, although they are grossly redundant if lost individually. Nonetheless, *Myf5*-null and *MyoD*-null myoblasts have distinct and complementary defects that indicate that each factor has unique functions. The exceptional similarity of the two factors within their conserved bHLH dimerization/ DNA-binding domains suggests that the divergent amino- and carboxy-terminal portions of each protein provide for the distinct effects. Therefore, by interchanging the corresponding regions of the two factors, specific functions of Myf5 and MyoD were distinguished and mapped to functional domains in the proteins.

Chapter 2

Loss of *p53* allows the survival of *Pax7*$^{-/-}$ myoblasts

Purpose

To determine if the loss of p53 function rescues the defect in survival and loss of satellite cells observed in *Pax7*$^{-/-}$ mice.

Contributions of Co-Authors

Jeff Ishibashi and Patrick Seale originated this project together.

Patrick Seale did the original cross of the *Pax7*-mutant and *p53*-null mouse lines.

Jeff Ishibashi did all of the experimental work and analysis described herein.

Loss of *p53* allows the survival of *Pax7*$^{-/-}$ myoblasts

Jeff Ishibashi [1,2], Patrick Seale [2,3], and Michael A. Rudnicki [1,2,3]

[1] University of Ottawa, Ottawa, Ontario, Canada
[2] Ottawa Health Research Institute, Ottawa, Ontario, Canada
[3] McMaster University, Hamilton, Ontario, Canada

Corresponding Author: Michael A. Rudnicki
Ottawa Health Research Institute
501 Smyth Road
Ottawa, ON K1H 8L6
CANADA
phone: 613-739-6740
fax: 613-737-8803
email: <mrudnicki@ohri.ca>

Running Title: Isolation of *Pax7-/-;p53-/-* Myoblasts

Abstract

Loss of *p53* rescues the embryonic cell death that occurs in the absence of Pax3, as does expression of Pax7 in place of Pax3. This suggests that Pax7 has an anti-apoptotic function that could manifest in adult skeletal muscle myoblasts, much as its closely related paralog Pax3 has in embryogenesis. Deletion of *p53* was therefore tested for the ability to rescue the apoptosis of *Pax7*-deficient adult myoblasts. In contrast to *Pax7*-null;$p53^{+}$ mice, cultures from $Pax7^{-/-}$;$p53^{-/-}$ hind-limb muscles yielded proliferating myoblasts that expressed desmin and MyoD, and that differentiated with expression of myogenin and myosin heavy-chain. Removal of *p53* therefore allows the survival of *Pax7*-null adult myoblasts.

Introduction

An important aspect of Pax3 function in pre-somitic and newly formed somitic mesoderm is the prevention of apoptosis (Borycki et al., 1999). It was also observed that those somitic regions that survived had compensatory up-regulation of Pax7 expression (Borycki et al., 1999). Similarly, the loss of Pax3 in the neural tube during embryogenesis causes apoptosis, producing neural tube defects (Phelan et al., 1997); however, concurrent loss of *p53* function rescues these cells and prevents the associated defect (Pani et al., 2002). Germline substitution of *Pax7* into the *Pax3* locus demonstrated that Pax7 also provides this anti-apoptotic function amongst neural tube, neural crest, and somitic regions (Relaix et al., 2004).

The primary role of Pax7 in satellite cells is considered to be an essential lineage specification function (Seale et al., 2000), and strong support continues to accumulate for Pax7 having an important role in activating the myogenic program (Chapter 3; Maroto et al., 1997; Zammit et al., 2006). However, Pax7 has an additional role in the continued survival of adult myoblasts, as Relaix et al. (2006) observed widespread apoptosis in postnatal skeletal muscle of *Pax7*-null pups and were able to induce apoptosis in cultured primary myoblasts by expressing a dominant-negative Pax7. We therefore tested whether *Pax7*-null myoblasts enter apoptosis through a p53-dependent pathway, and whether loss of p53 function could compensate for the loss of Pax7. Although there was no rescue of the gross *Pax7*-null phenotype, we were able to culture a novel $Pax7^{-/-};p53^{-/-}$ myoblast population that is absent from $Pax7^{-/-};p53^{+}$ preparations.

Materials & Methods

Mice

Pax7$^{+/-}$;*p53*$^{+/-}$ mice were derived by intercrossing Pax7$^{+/-}$ mice (Mansouri et al., 1996) with *p53*-null mice (Jackson Laboratory). Compound-heterozygous mice were intercrossed to generate combinations of *Pax7* and *p53* genotypes. Mice were maintained at the University of Ottawa Animal Care Facility.

Myoblast Isolation & Culture

Myoblasts were isolated and cultured from hind leg muscles of 6-8 week old mice as previously described (Huh et al., 2004). For immunocytochemistry, cells were seeded onto collagen-coated coverslips and allowed to adhere in growth medium for 24-48 hrs. For anti-myogenin and anti-MyHC staining, cells were rinsed with PBS after 24 hrs and cultured in differentiation medium (5% horse serum + DMEM + penicillin/streptomycin) for 24 hrs or 3 days (respectively).

Immunocytochemistry

Coverslips were rinsed with PBS and fixed with 2%PFA for 10 min at room temperature. Cells were permeabilized for 10 min with 0.5% Triton-X 100, then blocked with 5% goat serum + PBS. Antibodies were diluted into 5% goat serum + PBS for application. Primary antibodies used were: anti-desmin (DAKO; D33; 1:200); anti-MyoD (5.8A; Pharmingen; 1:250); anti-myogenin (hybridoma supernatant; F5D; 1:5); anti-MyHC (hybridoma supernatant; MF20; 1:10). Secondary detection used anti-mouse-FITC or

-rhodamine (Chemicon; 1:200). Final PBS washes include 0.1µg/mL of DAPI for visualizing nuclei. Coverslips were mounted onto slides with fluorescence mounting medium (DAKO).

Microscopy

Images were captured using an AxioCam HRm digital camera mounted on Axioplan2 microscope using 20x/0.50 Plan-NEOFLUAR (Ph2; ∞/0.17) or 40x/0.75 Plan-NEOFLUAR (Ph2; ∞/0.17) objectives (Zeiss). Digital images were captured using Axiovision software (Zeiss) and processed with Photoshop (Adobe). Cell counting was assisted by ImageJ software (http://rsb.info.nih.gov/ij/).

Results & Discussion

Previous literature about the role of Pax3 in cell survival (Borycki et al., 1999; Pani et al., 2002) prompted us to hypothesize that Pax7 has an anti-apoptotic function in adult myogenesis. This cell death could result in the absence of proper satellite cells in the skeletal muscle of adult *Pax7*-null mice (Kuang et al., 2006; Seale et al., 2000). A compound *Pax7/p53* mutant mouse line was therefore derived, by intercrossing mice carrying null *Pax7* or null *p53* alleles. Breeding of the resulting $Pax7^{+/-}$; $p53^{+/-}$ (and eventually also $Pax7^{+/-}$; $p53^{-/-}$) mice demonstrated that there was no overt rescue of the *Pax7*-null phenotype. *Pax7*-null pups remained runted regardless of *p53* status, and very few survived to weaning. As with $Pax7^{-/-}$;$p53^{+/+}$ offspring, the very few that survive to adulthood remain much smaller than their littermates and had a typical *Pax7*-null

appearance (Mansouri et al., 1996). Enumeration of the genotypes at three weeks of age demonstrated that loss of *p53* failed to provide *Pax7*-null pups with improved survival (Table I). Although there is the suggestion that loss of *p53* may have decreased survival (4% of $Pax7^{+/-}$; $p53^{-/-}$ versus 9% of $Pax7^{-/-}$;$p53^{+}$), such a conclusion would require larger numbers for support, as $Pax7^{-/-}$ pups usually survive at rates of 5-10% (Oustanina et al., 2004).

Surprisingly, however, whole preparations from the hind-limb muscles of 6-8 week-old $Pax7^{-/-}$;$p53^{-/-}$ mice revealed the presence of myoblasts representing ~5% of the population (Table II). Immunostaining for myogenic markers desmin (Figure 1) and MyoD (Figure 2) definitively show that myoblasts are present. Additionally, these myoblasts are capable of differentiating in low-mitogen conditions to express the characteristic myogenic differentiation markers myogenin (Figure 3) and myosin heavy-chain (Figure 4). This contrasts dramatically with $Pax7^{-/-}$;$p53^{+}$ isolations, which contain only extremely rare myoblasts (1/150 < 0.7%) (Kuang et al., 2006). Importantly, the $Pax7^{-/-}$;$p53^{-/-}$ myoblasts are capable of proliferating, and maintain their proportion of the culture over several passages. This contrasts with $Pax7^{-/-}$;$p53^{+}$ myoblasts, which do not grow in culture (Kuang et al., 2006).

Post-natal growth and repair of skeletal muscle requires a functional satellite cell population, which becomes heavily depleted prior to adulthood in the absence of *Pax7* (Oustanina et al., 2004). Recently, it was shown that fetal myoblasts do not co-express Pax7 and Caspase3, but that $Myf5^{+}$ myoblasts undergo apoptosis following down-regulation of Pax7 expression (Kassar-Duchossoy et al., 2005). Similarly, the absence of

Table I. Loss of p53 does not rescue Pax7-null lethality (one Pax7+/-;p53+/-, one Pax7+/-;p53-/- parent)

	p53			p53 (Expected)*		
Pax7	+/-	-/-	Total	+/-	-/-	Expected
+/+	57	48	105	49	43	92.5
+/-	89	86	175	99	86	185
-/-	14	6	20	12	11	22.5
Total	160	140	300	150	150	300
	a		b			
P-value (ChiSq)	0.2482		0.0845			

* Expected survival of 5-10% Pax7-/- pups past weaning (Oustanina, 2004). Therefore, used 7.5% to calculate expected number of surviving Pax7-/- pups. Pax7+/+ and Pax7+/- are expected to occur with a 1:2 ratio.

a : no significant difference from expected ratio of p53+/- to -/-
b : no significant difference from expected ratio of Pax7(+/+:+/-:-/-) w.r.t. p53

Table II. Desmin+ cells in myoblast cultures

	Desmin+	Nuclei	%
+/-;+/-	473	534	89%
-/-;+/-	0	570	0%
-/-;-/-	66	1345	5%

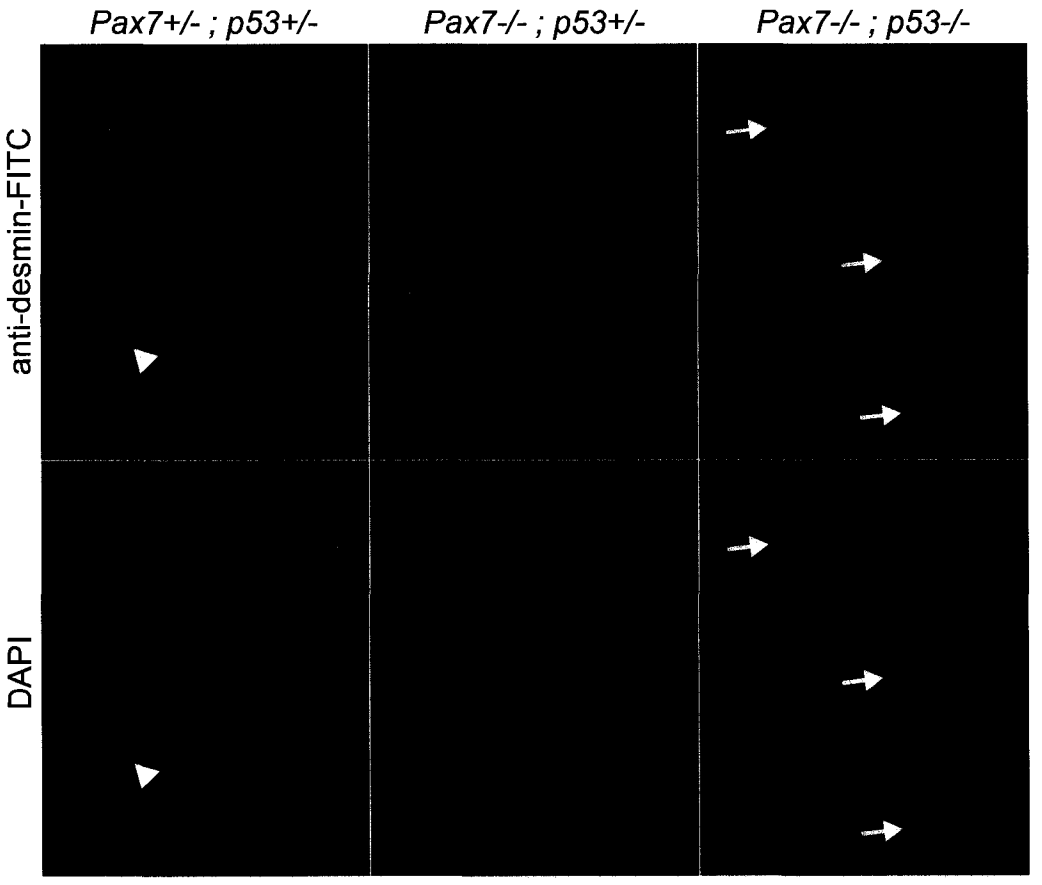


Figure 1: Desmin+ myoblasts are found in *Pax7-/-;p53-/-* cultures. *Pax7-p53* double-heterozygous myoblast cultures are equivalent to wild-type, with more than 95% expressing desmin (arrowhead indicates desmin-negative cell). No desmin-positive cells are found in *Pax7-/-* cultures when a functional *p53* allele is present. However, approximately 5% of cells in *Pax7-p53* double-null myoblast cultures express desmin (arrows).

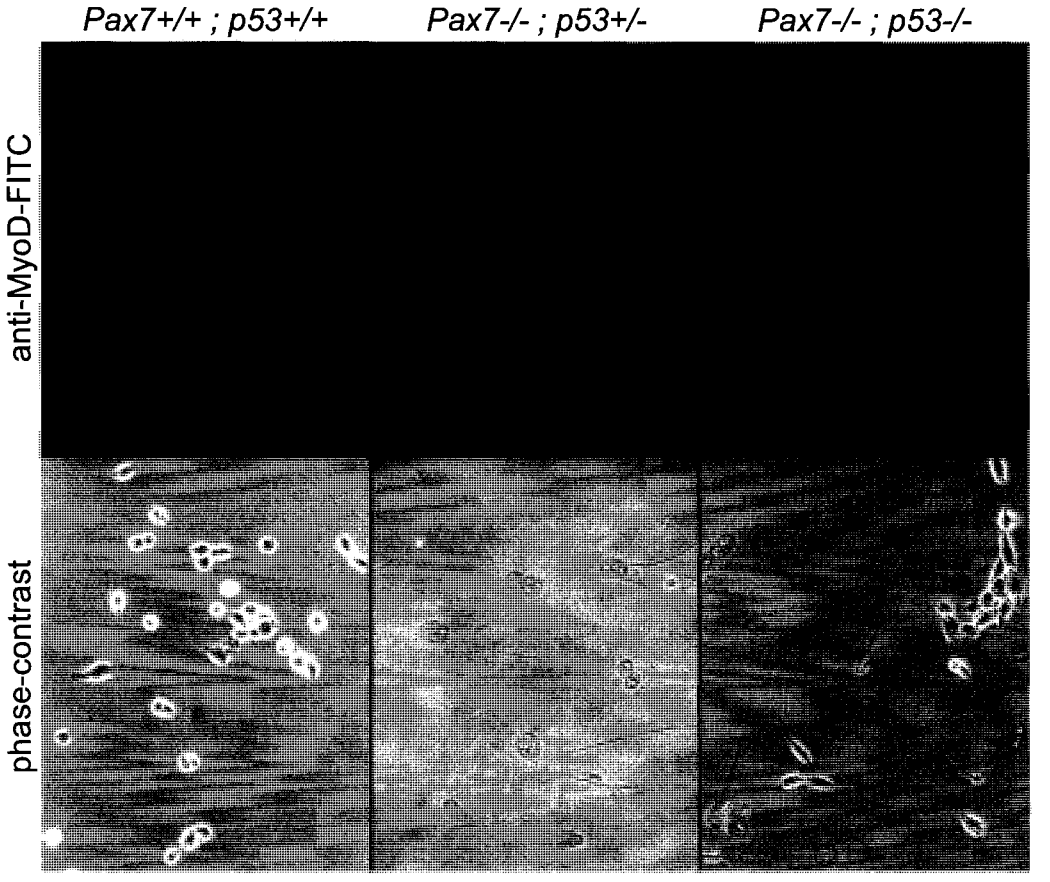


Figure 2: MyoD+ myoblasts are found in *Pax7-/-;p53-/-* cultures. Similarly to desmin, a comparable proportion of *Pax7-/-;p53-/-* cells express MyoD. Wild-type myoblasts are uniformly stained positive whereas *Pax7-/-;p53-/-* isolations are uniformly negative.

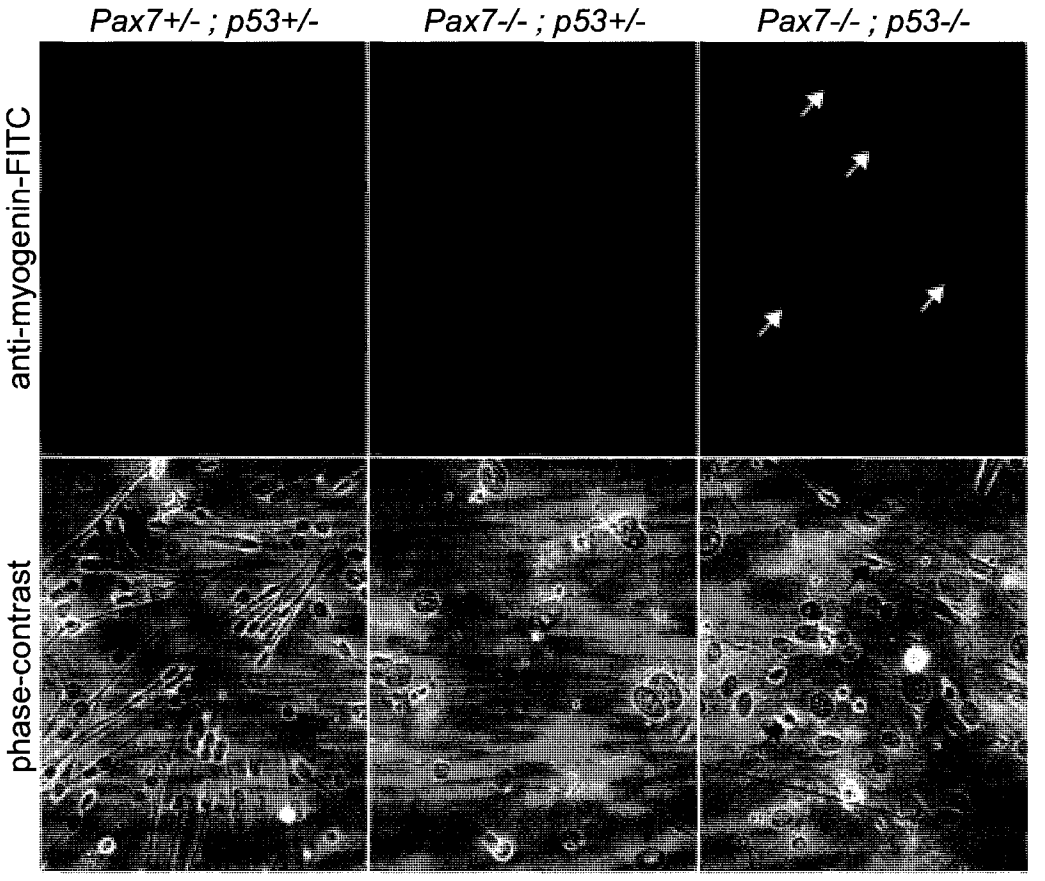


Figure 3: *Pax7-/-;p53-/-* myoblasts express myogenin when asked to differentiate. When cultured in low-mitogen conditions for 24 hours, *Pax7-/-;p53-/-* myoblasts express myogenin and begin to elongate as they differentiate. Wild-type myoblasts differentiate very efficiently and express high levels of myogenin, but *Pax7-/-;p53-/-* cultures are uniformly negative.

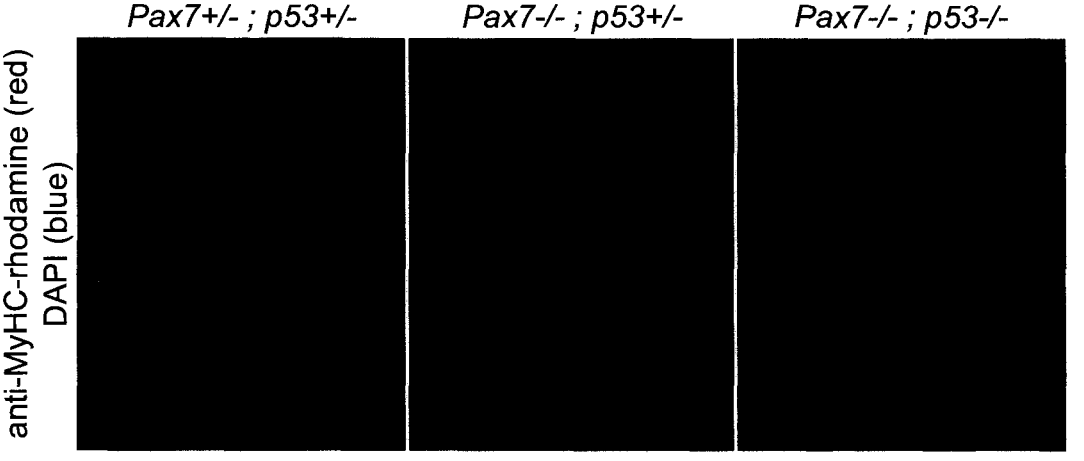


Figure 4: *Pax7-/-;p53-/-* myoblasts express the late marker myosin heavy-chain in differentiation. When cultured in low-mitogen conditions for 3 days, *Pax7-/-;p53-/-* myoblasts express myosin heavy-chain (MyHC). Wild-type myoblasts express high levels of myosin, elongate, and fuse into multinucleated myotubes, whereas *Pax7-/-;p53-/-* cultures do not. MyHC (MF20 antibody), red; nuclei (DAPI), blue.

Pax7 and Pax3 leads to extensive apoptosis and a failure of myogenesis to progress beyond the earliest stages (Relaix et al., 2005). A small number of myogenic precursors have been identified as a $Pax3^+$ population in adult muscle, although their identification as a classical satellite cell or as a muscle-resident myogenic progenitor remains controversial (Kuang et al., 2006; Relaix et al., 2006). However, these $Pax3^+$ myoblasts proliferate very poorly and are inadequate for repairing muscle damage (Kuang et al., 2006). It is a tempting proposition that the *Pax7*$^{-/-}$;*p53*$^{-/-}$ myoblasts are *Pax3*$^+$, with enhanced survival and proliferation.

Nonetheless, loss of *p53* does not rescue the major *Pax7*$^{-/-}$ muscle phenotype, for the mice remain runted and the myoblasts that can be isolated are much less abundant that in wild-type mice. Therefore, Pax7 must have other functions in self-renewal and proliferation that are independent of any cell survival role or are p53-independent. The myoblasts described herein may represent an alternative $Pax3^+/Pax7^+$ population (Relaix et al., 2005) that is otherwise lost in the absence of Pax7.

Further purification and characterization of this population is required. Wild-type bulk myoblast cultures are easily enriched and purified because contaminating fibroblasts arrest and die whereas the myoblasts multiply quickly. However, loss of p53 probably also prevents the apoptosis of the desmin^{neg} cells growing alongside the myoblasts precluding a similar enrichment. A FACS-based approach using markers such as syndecans-3 or -4, CD34, or integrin-α7 could allow the myoblast population to be purified for culture. Such *Pax7*$^{-/-}$ myoblasts would be a valuable reagent for future studies of Pax7 function.

References

Borycki, A.G., J. Li, F. Jin, C.P. Emerson, and J.A. Epstein. 1999. Pax3 functions in cell survival and in pax7 regulation. *Development.* 126:1665-74.

Huh, M.S., M.H. Parker, A. Scime, R. Parks, and M.A. Rudnicki. 2004. Rb is required for progression through myogenic differentiation but not maintenance of terminal differentiation. *J Cell Biol.* 166:865-76.

Kassar-Duchossoy, L., E. Giacone, B. Gayraud-Morel, A. Jory, D. Gomes, and S. Tajbakhsh. 2005. Pax3/Pax7 mark a novel population of primitive myogenic cells during development. *Genes Dev.* 19:1426-31.

Kuang, S., S.B. Charge, P. Seale, M. Huh, and M.A. Rudnicki. 2006. Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. *J Cell Biol.* 172:103-13.

Mansouri, A., A. Stoykova, M. Torres, and P. Gruss. 1996. Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice. *Development.* 122:831-8.

Maroto, M., R. Reshef, A.E. Munsterberg, S. Koester, M. Goulding, and A.B. Lassar. 1997. Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. *Cell.* 89:139-48.

Oustanina, S., G. Hause, and T. Braun. 2004. Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification. *Embo J.* 23:3430-9.

Pani, L., M. Horal, and M.R. Loeken. 2002. Rescue of neural tube defects in Pax-3-deficient embryos by p53 loss of function: implications for Pax-3- dependent development and tumorigenesis. *Genes Dev.* 16:676-80.

Phelan, S.A., M. Ito, and M.R. Loeken. 1997. Neural tube defects in embryos of diabetic mice: role of the Pax-3 gene and apoptosis. *Diabetes*. 46:1189-97.

Relaix, F., D. Montarras, S. Zaffran, B. Gayraud-Morel, D. Rocancourt, S. Tajbakhsh, A. Mansouri, A. Cumano, and M. Buckingham. 2006. Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J Cell Biol.* 172:91-102.

Relaix, F., D. Rocancourt, A. Mansouri, and M. Buckingham. 2004. Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev.* 18:1088-105.

Relaix, F., D. Rocancourt, A. Mansouri, and M. Buckingham. 2005. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature*. 435:948-53.

Seale, P., L.A. Sabourin, A. Girgis-Gabardo, A. Mansouri, P. Gruss, and M.A. Rudnicki. 2000. Pax7 is required for the specification of myogenic satellite cells. *Cell*. 102:777-86.

Zammit, P.S., F. Relaix, Y. Nagata, A.P. Ruiz, C.A. Collins, T.A. Partridge, and J.R. Beauchamp. 2006. Pax7 and myogenic progression in skeletal muscle satellite cells. *J Cell Sci.* 119:1824-32.

Chapter 3

Pax7-Regulated Gene Expression in Adult Myoblasts

Purpose

To identify genes transcriptionally regulated by Pax7d in myoblasts, as well as determining how they are affected by the alternative splice isoforms of Pax7 and their potential roles in myogenic determination.

Contributions of Co-Authors

Jeff Ishibashi did all of the experimental work and analysis.

Myogenic target genes regulated by Pax7 include Myf5 and MyoD

Jeff Ishibashi and Michael A. Rudnicki

University of Ottawa, Ottawa, Ontario, Canada
Ottawa Health Research Institute, Ottawa, Ontario, Canada

Corresponding Author: Michael A. Rudnicki
Ottawa Health Research Institute
501 Smyth Road
Ottawa, ON K1H 8L6
CANADA
phone: 613-739-6740
fax: 613-737-8803
email: <mrudnicki@ohri.ca>

Running Title: Pax7 Regulates Myf5 & MyoD

Abstract

Adult skeletal muscle repair is mediated by satellite cells that require the transcription factor Pax7 for their survival and function. Using comparative microarray analysis, we identified numerous transcripts modulated by Pax7, including one encoding the myogenic regulatory factor Myf5. Pax7 increases Myf5 RNA and protein in myoblasts, and two of four isoforms (Pax7c and Pax7d) can also specifically induce Myf5 expression in non-muscle 10T1/2 fibroblasts. The same functional difference is observed in the regulation of other target genes and corresponds to an alternatively spliced GL^{-} change in the paired DNA binding domain. Chromatin immunoprecipitation (ChIP) data also indicate that Pax7 binds to the *MyoD* -20kb core enhancer and therefore regulates both primary myogenic factors. Accordingly, the alveolar rhabdomyosarcoma fusion protein Pax7d/FKHR is a potent inducer of both Myf5 and MyoD. These data therefore provide a mechanism through which Pax7 can enforce myogenic commitment.

Abbreviations

bHLH	basic-helix-loop-helix
ChIP	chromatin immunoprecipitation
FKHR	forkhead box O1 (FoxO1)
$GL^{+/-}$	glycine-leucine inclusive (+) or exclusive (-) alternative splicing
$Q^{+/-}$	glutamine inclusive (+) or exclusive (-) alternative splicing
MRF	myogenic regulatory factor (Myf5; MyoD; Myogenin, MRF4)
aRMS/eRMS	rhabdomyosarcoma (alveolar or embryonal)

Introduction

Skeletal muscle has an impressive capacity for growth and regeneration that is provided primarily by a resident population of committed skeletal muscle stem cells known as satellite cells (reviewed in Bischoff, 1994; Zammit et al., 2002). The majority of post-natal muscle growth is fueled by these cells, reducing their frequency from 30% of nuclei at birth to a steady-state of approximately 5% in the adult (Bischoff, 1994). Normally quiescent in the adult, satellite cells reside adjacent to the muscle fibres beneath the basal lamina (Mauro, 1961). Traumatic or injury-induced damage to mature muscle is repaired by activating this pool of satellite cells, which proliferate as myoblasts before differentiating and fusing into muscle fibres. A fraction of the myoblasts return to quiescence without differentiating, replenishing the satellite cell population (rev. in Charge and Rudnicki, 2004). At the molecular level, quiescent satellite cells and proliferating myoblasts express transcription factors from the myogenic basic helix-loop-helix (bHLH) and Pax families that commit them to the myogenic lineage and permit them to eventually form functional skeletal muscle fibres (rev. in Berkes and Tapscott, 2005; Holterman and Rudnicki, 2005; Sabourin and Rudnicki, 2000).

The myogenic bHLH genes (*Myf5; MyoD; Myogenin; MRF4*) encode a family of myogenic regulatory factors (MRFs) that are essential for skeletal myogenesis (rev. in Berkes and Tapscott, 2005). In the adult, Myf5 and MyoD provide myogenic identity to growing myoblasts, with Myf5 associated with proliferation (Megeney et al., 1996; Sabourin et al., 1999) and MyoD tending towards differentiation (Ishibashi et al., 2005). Myogenin and MRF4 are expressed through differentiation (rev. in Sabourin and

Rudnicki, 2000). Individual elimination of any of the four MRFs in mice produces unique defects in muscle development. However, the striking exhibition of their importance is seen in compound knockout mice such as the *MyoD*$^{-/-}$;*Myf5*$^{-/-}$ (Kassar-Duchossoy et al., 2004; Rudnicki et al., 1993) or *Myf5*$^{-/-}$;*Pax3*$^{Sp/Sp}$ (*Splotch*~*Pax3*-null) (Tajbakhsh et al., 1997) mice, each of which fails to develop most skeletal muscle tissue. Similarly, ectopic expression of any of the MRFs in non-muscle cells can activate the skeletal muscle program. The MRFs are therefore considered master regulatory genes in the skeletal muscle lineage.

Two additional genes, encoding the paired-box transcription factors Pax7 (Jostes et al., 1990) and Pax3 (Goulding et al., 1991), are important to skeletal muscle development and repair, and function upstream of the MRFs (Relaix et al., 2006; Tajbakhsh et al., 1997). They are very similar, sharing a highly conserved structure with 85% amino acid conservation overall and particularly 94% identity within their DNA binding domains. *Pax3*-mutant *Splotch* mice have substantial defects in embryonic myogenesis and die perinatally, with a complete absence of the muscles of the limbs (Bober et al., 1994). *Pax7*-null pups are severely runted and usually die within 2-3 weeks of birth (Mansouri et al., 1996; Seale et al., 2000). Notably, *Pax7*$^{-/-}$ skeletal muscles are nearly devoid of the satellite cells that are crucial for post-natal muscle growth and injury repair (Kuang et al., 2006; Seale et al., 2000). The residual satellite cells, including those in axial muscles that continue to express Pax3, are unable to give rise to proliferative myogenic precursors and instead are progressively lost, presumably by apoptosis (Kuang et al., 2006; Relaix et al., 2006). As such, Pax7 is uniquely required

by satellite cells for their myogenic function (Relaix et al., 2005; Seale et al., 2000). However, despite the persuasive genetic evidence for the important role of Pax3 and Pax7 in muscle development, the specific mechanisms by which they function in this lineage remain obscure.

Total Pax7 expression in skeletal muscle is a mixture of four isoforms generated by alternative splicing between exons 2 & 3 or exons 3 & 4 (summarized in Table I; Ziman and Kay, 1998). The exon 2/3 event produces an inclusion or exclusion of a single glutamine residue ($Q^{+/-}$ at position #138); the exon 3/4 change results in the presence or absence of a glycine-leucine ($GL^{+/-}$ at position #180 or #181) (Ziman and Kay, 1998). Pax3 uses the same alteration at the $Q^{+/-}$ change (Vogan et al., 1996) but does not undergo $GL^{+/-}$ splicing (Du et al., 2005). Both $Q^{+/-}$ and $GL^{+/-}$ changes occur within the highly conserved paired-box DNA binding domain, suggesting that they could modulate the DNA binding site recognition or affinity of the differing isoforms, which has been confirmed for the $Q^{+/-}$ change in Pax3 (Vogan et al., 1996). While isoform expression has been correlated with neuronal versus myogenic tissue and with mouse strain-dependent differences in skeletal muscle regeneration (Kay and Ziman, 1999; Ziman and Kay, 1998), a specific function for *Pax7* alternative splicing has not been shown.

Pax7 and Pax3, following translocation-induced linkage of their DNA binding domains to the transactivation domain of FoxO1 (FKHR), become Pax7/FKHR and Pax3/FKHR fusion proteins. These fusions are found in the majority of alveolar rhabdomyosarcomas (aRMS), the most common childhood tumour, of which the alveolar

Table I. Alternative splicing in the Pax7 paired-box DNA binding domain.[1]

Isoform	Q (exon 2/3)	GL (exon 3/4)
A	-	+
B	+	+
C	-	-
D	+	-

[1] Ziman and Kay (1998a)

subtype is the most aggressive and carries the worst prognosis (rev. in Xia et al., 2002). Pax3/FKHR has been demonstrated to be a strong but improper regulator of Pax3 target genes (Fredericks et al., 1995), leading to the hypothesis that mis-regulated activation of Pax3 or Pax7 targets by their respective fusions is critical to the development of this form of cancer. This is consistent with the intersection of the prominent roles of Pax3 and Pax7 in myogenesis with the phenotype of RMS tumours, which are characterized the expression of muscle markers including MyoD and Myogenin (Sebire and Malone, 2003). Understanding Pax7 is therefore important for deciphering the role of Pax7/FKHR in aRMS oncogenesis and identifying its therapeutic vulnerabilities.

To gain insight into Pax7's functions as a myogenic transcription factor, we ectopically expressed the four Pax7 isoforms in 10T1/2 fibroblasts, C2C12 myoblasts, and primary myoblasts. Our analysis has revealed that Pax7 regulates the expression of numerous novel target genes in an isoform-dependent manner. Strikingly, Pax7 generally is able to increase *Myf5* expression in myogenic cells, whereas only the GL$^-$ isoforms Pax7c and Pax7d can initiate *Myf5* expression in non-muscle cells. The Pax7d/FKHR fusion oncoprotein is a strong inducer of both *Myf5* and *MyoD*. Regulation of *MyoD* is potentially an effect of DNA binding within the well-characterized *MyoD* -20kb core enhancer. Together, these results provide important insight into the role of Pax7 in enforcing myogenic determination.

Results

Identification of candidate Pax7 target genes in myoblasts.

The C2C12 myoblast cell line expresses low but detectable levels of Pax7, suggesting that it would provide a responsive environment in which ectopic Pax7 would be functional. Pools of C2C12 myoblasts were therefore infected with retrovirus expressing mouse Pax7d (FLAG epitope-tagged on the carboxy-terminus) and a puromycin selectable marker. No difference in function between Pax7d and Pax7d-FLAG tag has been observed (unpublished data). Empty virus (puromycin-resistance alone) was used as a control. Infected pools were stably selected with puromycin, after which changes in gene expression were examined. Whereas Pax7 expression was very low in the control C2C12 pool, high levels of Pax7d were observed in the infected population (Fig. S1, available at http://). Expression of Pax7d in C2C12 robustly inhibited their normal differentiation into myotubes upon serum withdrawal (Fig. 1), similar to Pax3 (Epstein et al., 1995), demonstrating that C2C12 are responsive to Pax7d.

Total RNA was harvested from triplicate Pax7d+ and control pools to generate probes for hybridization to Affymetrix GeneChip microarrays. C2C12-Pax7d samples were stringently compared against C2C12-puro samples to derive sets of candidate Pax7-regulated genes (see Materials & Methods for criteria). There were 26 genes down-regulated by Pax7d, although only one changed by greater than 4-fold (*IGFBP2*, 6.5-fold; Table II). In contrast, 43 genes were up-regulated (Table III), with a number of them exhibiting large changes in expression. *PlagL1*, for instance, was undetected in control C2C12 but highly expressed in Pax7d-expressing C2C12. Others that were considerably

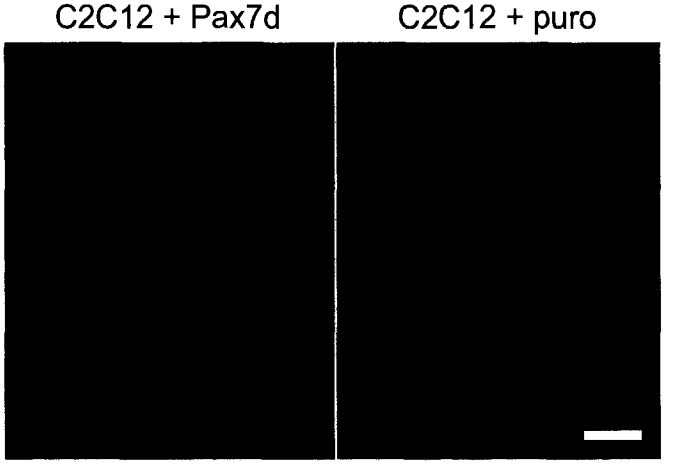


Figure 1. Pax7 inhibits C2C12 differentiation. Stable over-expression of Pax7d strongly inhibited myogenic differentiation compared to empty-vector controls, as assayed by fusion and expression of myosin heavy-chain (MyHC = red; DAPI = blue). Bar = 100 μm.

Table II. Pax7d-induced decreases in expression in C2C12 myoblasts

Symbol	GID	Fold	Name
Igfbp2	NM_008342	**-6.4**	insulin-like growth factor binding protein 2
Fbln2	NM_007992	**-3.7**	fibulin 2
Tnnt2	NM_011619	**-3.3** [2]	troponin T2, cardiac
Ctgf	NM_010217	**-3.2**	connective tissue growth factor
Timp3	NM_011595	**-2.6** [3]	tissue inhibitor of metalloproteinase 3
Serpine1	NM_008871	**-2.6**	serine (or cysteine) proteinase inhibitor, E1
Csrp2	NM_007792	**-2.6**	cysteine and glycine-rich protein 2
4732435N03Rik	AV371987	**-2.5**	RIKEN cDNA 4732435N03 gene
Lox	NM_010728	**-2.5** [2]	lysyl oxidase
Aqp1	NM_007472	**-2.5**	aquaporin 1
Serpinb9b	NM_011452	**-2.4**	serine (or cysteine) proteinase inhibitor, B9b
Thbs2	NM_011581	**-2.4**	thrombospondin 2
Col6a3	NM_009935	**-2.4**	procollagen, type VI, alpha 3
Tubb3	NM_023279	**-2.3**	tubulin, beta 3
Tspan9	NM_175414	**-2.3**	tetraspanin 9
Hbegf	NM_010415	**-2.3**	heparin-binding EGF-like growth factor
Hspb1	NM_013560	**-2.2** [2]	heat shock protein 1
Sgpp1	NM_030750	**-2.2**	sphingosine-1-phosphate phosphatase 1
Ankrd1	NM_013468	**-2.1**	ankyrin repeat domain 1 (cardiac muscle)
Ccl2	NM_011333	**-2.1**	chemokine (C-C motif) ligand 2
Fhl1	NM_010211	**-2.1**	four and a half LIM domains 1
Anxa3	NM_013470	**-2.0**	annexin A3
Akap12	NM_031185	**-2.0**	A kinase (PRKA) anchor protein (gravin) 12
Actb	NM_007393	**-2.0**	actin, beta, cytoplasmic
Cst6	NM_028623	**-2.0**	cystatin E/M
Tm4sf1	NM_008536	**-2.0**	transmembrane 4 superfamily member 1

(2) mean fold change for 2 distinct probesets in Table II directed at the same transcript

(3) mean fold change for 3 distinct probesets in Table II directed at the same transcript

Table III. Pax7d-induced increases in expression in C2C12 myoblasts

Symbol	GID	Fold	Name
Plagl1	**NM_009538**	**385.0** [1]	**pleiomorphic adenoma gene-like 1**
Lix1	**NM_025681**	**12.3**	**limb expression 1 homolog (chicken)**
Syne2	**BF582734**	**11.3**	**synaptic nuclear envelope 2**
Cipar1	**AK008716**	**8.1**	**castration-induced prostatic apoptosis-related 1**
Trim54	**NM_021447**	**7.2**	**tripartite motif-containing 54**
Il13ra1	NM_133990	**5.4** [2]	interleukin 13 receptor, alpha 1
Mest	**NM_008590**	**5.0**	**mesoderm specific transcript**
Peg3	NM_008817	**4.6** [2]	paternally expressed 3
3110001A13Rik	NM_025626	**4.6** [3]	RIKEN cDNA 3110001A13 gene
Ahr	NM_013464	**4.4**	aryl-hydrocarbon receptor
Npnt	NM_033525	**4.1** [2]	nephronectin
Ltb4dh	NM_025968	**4.1**	leukotriene B4 12-hydroxydehydrogenase
Msln	NM_018857	**4.1**	mesothelin
C1qtnf3	NM_030888	**3.8**	C1q and tumor necrosis factor related protein 3
-	BB369191	**3.7**	sim. to mouse pentylenetetrazol-related mRNA
Pparg	NM_011146	**3.4**	peroxisome proliferator activated receptor gamma
Cd24a	NM_009846	**3.1** [2]	CD24a antigen
Pkia	NM_008862	**3.0** [2]	protein kinase inhibitor, alpha
Nqo1	NM_008706	**2.9**	NAD(P)H dehydrogenase, quinone 1
Marcks	NM_008538	**2.6** [3]	Myristoylated alanine rich protein kinase C substrate
Gch1	NM_008102	**2.6**	GTP cyclohydrolase 1
Cdh11	NM_009866	**2.6**	cadherin 11
Myo1d	NM_177390	**2.5**	myosin ID
Cxcr4	NM_009911	**2.5**	chemokine (C-X-C motif) receptor 4
Crlf1	NM_018827	**2.5**	cytokine receptor-like factor 1
Sema3e	NM_011348	**2.4**	semaphorin 3E
Mdfic	NM_175088	**2.4**	MyoD family inhibitor domain-containing
Ass1	NM_007494	**2.4**	argininosuccinate synthetase 1
Id3	NM_008321	**2.4**	inhibitor of DNA binding 3
Ppap2a2	NM_008903	**2.3** [2]	phosphatidic acid phosphatase 2a isoform 2
Myf5	**NM_008656**	**2.2**	**myogenic factor 5**
Cdh2	NM_007664	**2.2**	cadherin 2 (N-cadherin)
Depdc6	NM_145470	**2.2**	DEP domain containing 6
Prss23	NM_029614	**2.2** [2]	protease, serine, 23
Id2	NM_010496	**2.1**	inhibitor of DNA binding 2
Emb	NM_010330	**2.1**	embigin
Rnf128	NM_023270	**2.1** [2]	ring finger protein 128
Rnh1	NM_145135	**2.1**	ribonuclease/angiogenin inhibitor 1
Olfm1	NM_019498	**2.1**	olfactomedin 1
-	BG069607	**2.1**	-
Tob1	NM_009427	**2.1**	transducer of ErbB-2.1
Atp11a	NM_015804	**2.1**	ATPase, class VI, type 11A
2210409B22Rik	BM207133	**2.0**	RIKEN cDNA 2210409B22 gene

(1) signal was very low and called absent in control

(2) mean fold change for 2 distinct probesets in Table II directed at the same transcript

(3) mean fold change for 3 distinct probesets in Table II directed at the same transcript

Bold lines indicate those used for real-time PCR validation

increased included *Lix1* (12-fold); *Syne2* (11-fold); *Cipar1* (8-fold); *Trim54* (7-fold); and *Mest* (5-fold). Interestingly, *Myf5* expression was consistently increased by Pax7d (2.2-fold). PCR primers specific to these transcripts were therefore designed in order to validate them by real-time PCR as candidate Pax7d targets genes.

Real-time RT-PCR of the same RNA samples used for the GeneChip analysis confirmed that all of these transcripts were increased (unpublished data). Five transcripts exhibited increases of similar magnitude or exceeding that suggested by GeneChip analysis, whereas one (*Syne2*) showed a minimal (2-fold) change (unpublished data). More recent sequence data has allowed the design of primers with greater specificity for the *PlagL1* and *Cipar1* transcripts. Real-time PCR validation confirmed that *PlagL1* was strongly up-regulated by Pax7d in C2C12 (135-fold). Similarly, *Cipar1* was up-regulated ~150-fold, suggesting that the GeneChip probesets are sub-optimal and substantially underestimate changes in abundance of that transcript. Six validated candidates were used for our analysis of several additional, independently derived samples.

Pools of C2C12 cells were derived as before, each infected with retrovirus expressing one of the four Pax7 isoforms (Table I), or a control (empty) retrovirus. The abilities of each of the Pax7 isoforms to increase the expression of our candidate targets were then assayed by real-time PCR. Strikingly, Pax7c and Pax7d both produced substantial up-regulation of *Lix1*, *Cipar1*, *Trim54*, and *Mest* transcripts, whereas Pax7a and Pax7b were only able to moderately increase the *Cipar1* and *Mest* transcripts (Fig. 2 A). This immediately pointed to a functional importance for the $GL^{+/-}$ alternative splicing of Pax7. Similar ectopic expression of Pax3(Q^{-}) in C2C12 cells had little or no effect on

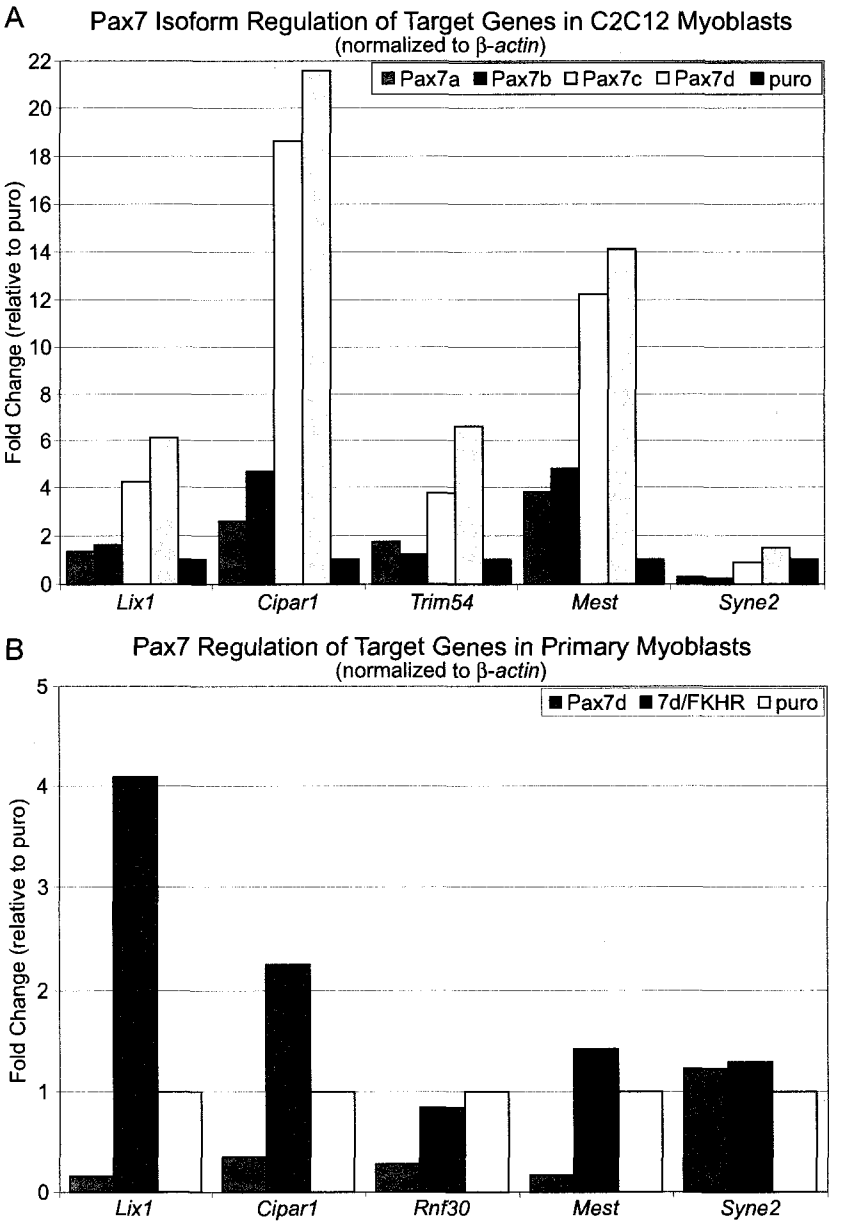


Figure 2. Regulation of candidate target genes by Pax7 isoforms and in primary myoblasts. (A) C2C12 myoblasts were stably infected with retrovirus expressing one of the Pax7 isoforms (a/b/c/d), and the expression levels of candidate target genes were examined by real-time RT-PCR. Pax7c and Pax7c produced substantial increases in target expression, whereas Pax7a and Pax7b had comparatively little or moderate effect. (B) Primary myoblasts (which express high levels of endogenous Pax7) were stably infected with retrovirus expressing Pax7d or the Pax7d/FKHR fusion. Ectopic Pax7d impaired the expression of targets, whereas Pax7d/FKHR maintained or increased target gene expression levels. (A,B) Whereas *Lix1*, *Cipar1*, *Trim54*, and *Mest* are robustly affected, *Syne2* showed little change.

the levels of these transcripts, suggesting a unique role for Pax7 in their regulation (unpublished data).

In contrast to the C2C12 myoblast cell line, primary skeletal myoblasts express high levels of Pax7. We therefore assessed the effects of Pax7d on gene expression in the presence of high endogenous Pax7 expression. As shown in Fig. 2 B, over-expression of Pax7d reduced the expression of the candidate targets in primary myoblasts (2.5- to 6-fold decreased), consistent with high levels of ectopic Pax7d interfering with target activation by endogenous Pax7. The Pax7d/FKHR aRMS fusion was separately introduced in order to find out if it was also capable of regulating these genes. Pax7d/FKHR increased the expression of *Cipar1* and *Lix1* (2- to 4-fold) (Fig. 2 B) and maintained the expression of the others, suggesting that Pax7d/FKHR is a strong transactivator that can offset any interference with the endogenous Pax7.

Pax7 regulates Myf5 and MyoD

The observation that *Myf5* was up-regulated following ectopic expression of Pax7 was consistent with the hypothesis that *Myf5* is a Pax7 target. The microarray result was validated by real-time PCR and by traditional RT-PCR and agarose gel electrophoresis, demonstrating that the quantity of *Myf5* transcript was indeed increased by Pax7d (~3-fold) while *MyoD* mRNA levels were unaffected (Fig. 3). Importantly, Myf5 protein levels changed coordinately with the observed change in mRNA levels. Expressing Pax7d in C2C12 cells resulted in a dramatic increase in Myf5 protein compared to the empty vector control (Fig. 4). An exaggerated change in Myf5 protein levels versus

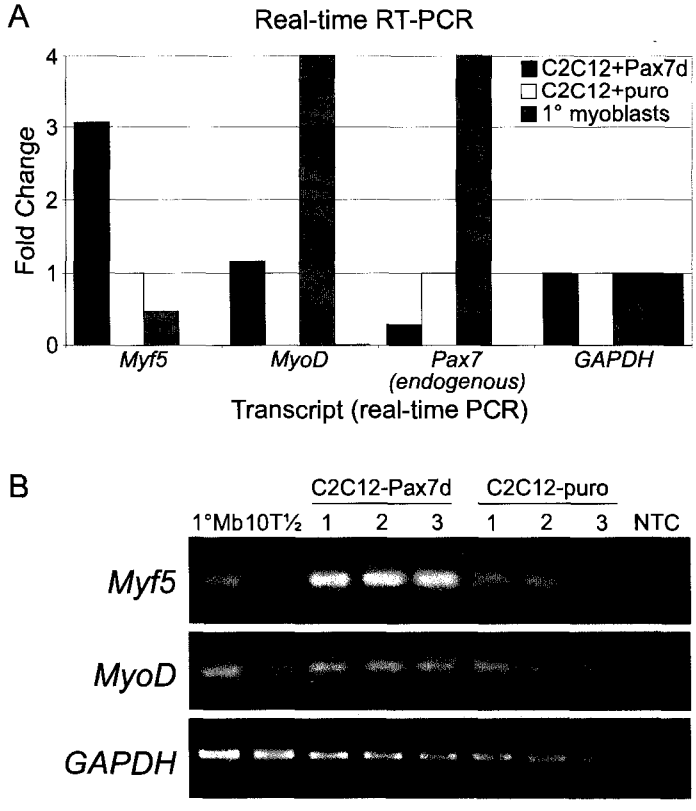


Figure 3. Pax7d increases expression of *Myf5* in C2C12 myoblasts. (A) SYBR green real-time PCR showing increased *Myf5* transcripts, but no change in *MyoD* transcripts; normalized to *GAPDH.* (B) Conventional RT-PCR for *Myf5*, *MyoD*, and *GAPDH* transcripts. Primary myoblasts (1°Mb) are positive for *Myf5*, *MyoD*, and *Pax7*; 10T1/2 are negative.

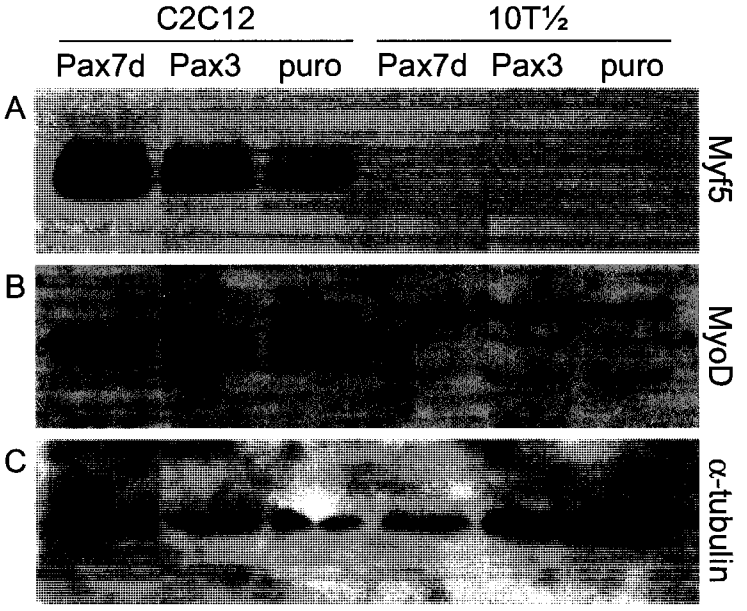


Figure 4. Pax7d and Pax3 up-regulate Myf5 protein. (A) Myf5 protein levels are increased by Pax7d-FLAG and Pax3-FLAG expression in C2C12 myoblasts, while (B) MyoD protein levels are unchanged. (C) anti-α-Tubulin.

mRNA has been previously reported in *MyoD-/-* myoblasts and attributed to a post-transcriptional mechanism (Sabourin et al., 1999). While Pax3 also produced slightly higher levels of Myf5 protein, they remained substantially lower than those induced by Pax7d (Fig. 4). MyoD protein levels remained unchanged, in agreement with the previous RT-PCR results. When the effects of Pax7d on myogenic factor gene expression were re-examined in primary myoblasts, a moderate increase in Myf5 protein was observed (Fig. 5 A). Additionally, however, Pax7d decreased the levels of MyoD and Myogenin protein, and possibly also endogenous Pax7 (Fig. 5 B and C). In primary myoblasts, Pax3 acted very similarly to Pax7d (Fig. 5).

More generally, all four isoforms of Pax7 showed the same ability to up-regulate Myf5 protein in C2C12 cells without effect on MyoD protein levels (Fig. 6). Unexpectedly, Pax7 was also able to induce Myf5 protein expression in the non-myogenic 10T1/2 fibroblast cell line. This occurred without induction of MyoD (Fig. 6; unpublished data). Western blotting of protein from Pax7-expressing 10T1/2 cells demonstrated that Pax7c and Pax7d were uniquely capable of inducing *de novo* Myf5 expression in non-muscle cells (Fig. 6). In contrast, ectopic expression of Pax7a or Pax7b was insufficient to induce Myf5 (Fig. 6).

A pertinent question was whether the induction of Myf5 was occurring at a low level across the entire pool of fibroblasts or if a small sub-population was expressing higher levels. Immunostaining of fixed cells with anti-Myf5 antibody could not detect any positive cells amongst the pools of 10T1/2+Pax7d (Fig. 7 A). However, ectopic expression of the alveolar rhabdomyosarcoma fusion protein Pax7d/FKHR, which is a

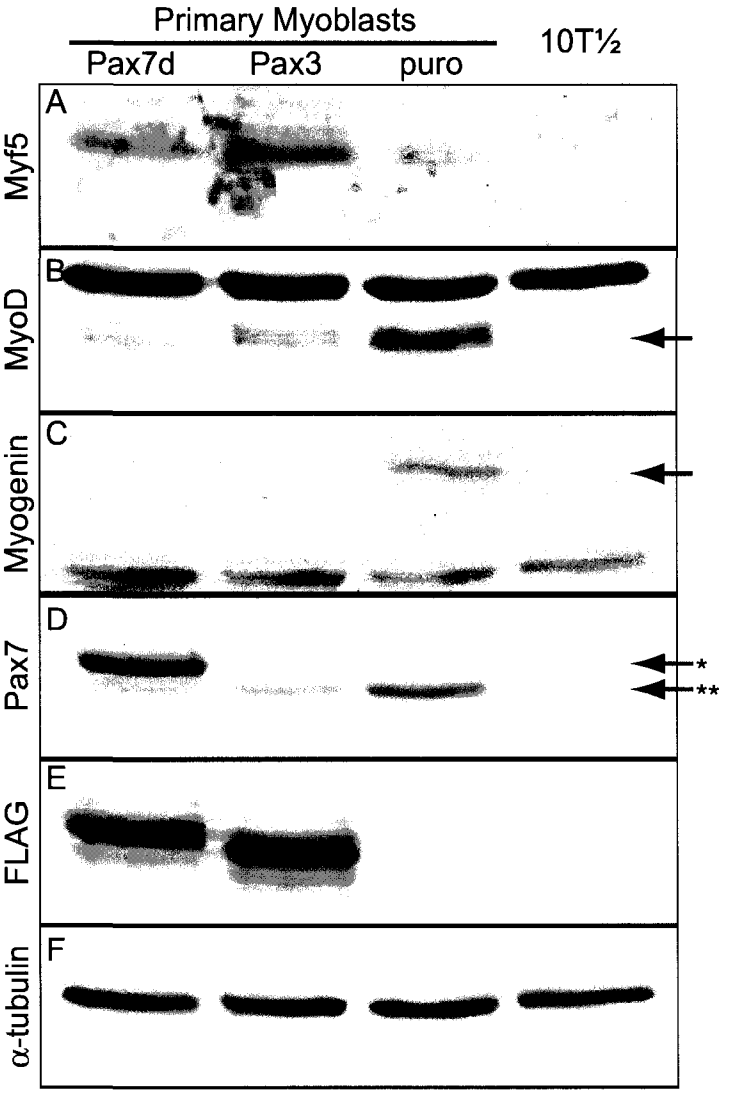


Figure 5. Pax7d and Pax3 regulate MRF expression in primary myoblasts. Pax7d and Pax3 increase (A) Myf5 and decrease (B) MyoD and (C) myogenin protein. (D) Pax7 protein, ectopic Pax7d-FLAG (*) and endogenous (**); (E) Pax7d- and Pax3-FLAG; (F) anti-α-Tubulin.

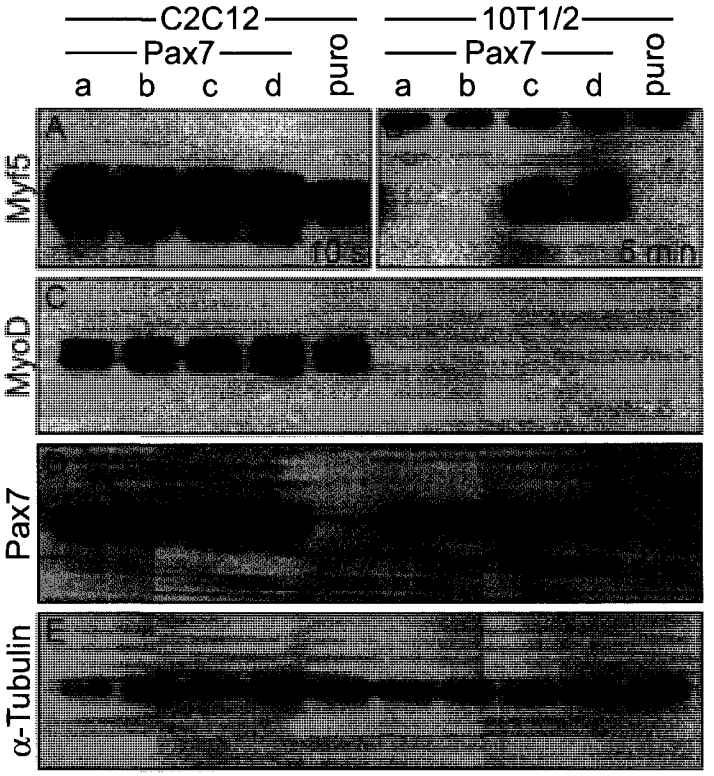


Figure 6. Pax7 activates Myf5 expression. (A,B) All Pax7 isoforms upregulate Myf5 protein in C2C12 myoblasts (A) whereas only Pax7c and Pax7d can activate Myf5 expression in 10T1/2 fibroblasts (B). Note that (A,B) were the same Western blot, but that (B) was exposed 30x longer than (A). (C) MyoD levels are unchanged by Pax7 overexpression in C2C12 myoblasts, and MyoD is not induced in 10T1/2 fibroblasts. (D) Each isoform of Pax7 is robustly expressed from the retrovirus. (E) α-Tubulin loading control. Negative controls (C2C12-puro; 10T1/2-puro) were infected with empty retrovirus.

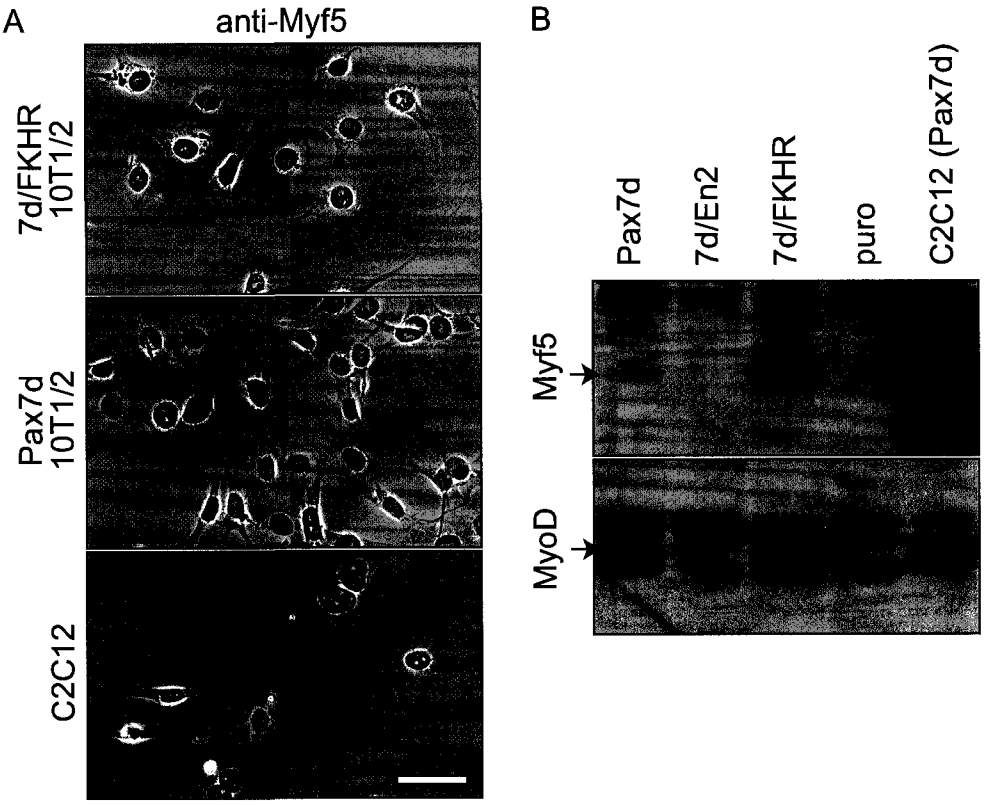


Figure 7. Pax7d/FKHR activates Myf5 and MyoD. (A) In a fraction of 10T1/2 fibroblasts (<20%), Myf5 protein was detectable by immuno-staining with anti-Myf5 antibody, whereas none were visible in Pax7d-overexpressing cells (n>1000). C2C12 myoblasts were >90% positive. Phase/contrast and FITC greyscale images were overlaid and false-coloured, with linear adjustments. Bar, 50 μm. (B) Pax7d induces a low level of Myf5 expression in 10T1/2, whereas Pax7d/FKHR induces moderate levels. Unlike Pax7d, Pax7d/FKHR is capable of inducing MyoD expression in 10T1/2 fibroblasts at levels comparable to that of C2C12 myoblasts.

stronger regulator of Myf5 than wild-type Pax7d (Fig. 7 B), showed that 10-20% of infected 10T1/2 cells were positive for Myf5 expression (Fig. 7 A). The activity of the Pax7d/FKHR fusion protein was also demonstrated by its ability to induce the expression of MyoD (Fig. 7 B). Whereas Myf5 protein levels induced by Pax7d/FKHR were still well below those seen in C2C12 myoblasts, the levels of MyoD induced in 10T1/2 were comparable to those expressed by C2C12 cells.

To investigate the hypothesis that *Myf5* was a direct transcriptional target of Pax7, we synthesized primers to amplify a variety of genomic regions for use in chromatin immunoprecipitation (ChIP)-PCR. In particular, sites in the *Myf5* promoter (-270bp) and at -38kb of the *Myf5* locus were tested by ChIP. The two well-characterized *MyoD* enhancers (the core -20kb enhancer (Goldhamer et al., 1995) and the -5kb enhancer (Asakura et al., 1995; Tapscott et al., 1992)) were also assessed. The carboxy-terminal FLAG-epitope tags on each of Pax7d and Pax3 were used to immunoprecipitate them and their bound DNA sequences from primary myoblasts, with empty-virus control pools of cells used as a baseline.

Control loci *p107*, *Pgc1*, and *GAPDH* were unchanged, and no enrichment was observed for either of the loci examined within the *Myf5* regulatory region (-270bp and -38kb) or in the *MyoD* (-5kb) enhancer region. However, the *MyoD* (-20kb) enhancer region implicated in myogenic specification (Kablar et al., 1999) was greatly enriched by Pax7- and Pax3-ChIP using two different primer pairs for the PCR (Fig. 8). This, in combination with the inhibitory effect of Pax7 and Pax3 in primary myoblasts on MyoD expression (Fig. 5) and the induction of high levels of MyoD in 10T1/2 fibroblasts by

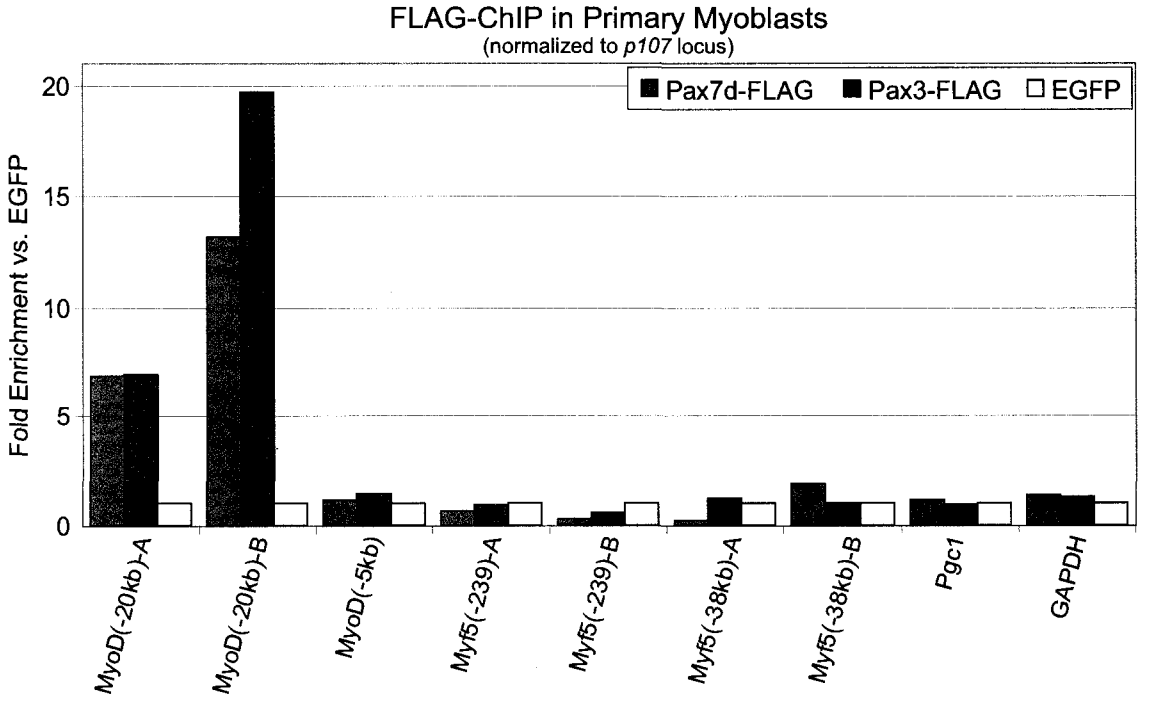


Figure 8. Chromatin immunoprecipitation (ChIP) of *MyoD* core enhancer region bound by Pax7d and Pax3 in primary myoblasts. Real-time PCR showed that a region of the *MyoD* core (-20kb) enhancer flanked by two different pairs of PCR primers (-A, -B) is greatly enriched by chromatin immuno-precipitation when an anti-FLAG antibody is used to recognize Pax7d- or Pax3-FLAG expressed in primary myoblasts. In contrast, the quantities of genomic DNA for the *MyoD* (-5kb) enhancer, *Pgc1*, *GAPDH*, and *p107* loci was unchanged. Similarly, there was no indication of enrichment with two primer pairs at the *Myf5* (-239bp) locus, or at two closely linked sites near *Myf5* (-38kb). Data was normalized against the *p107* locus, which was in close agreement with two other control loci, *Pgc1* and *GAPDH*.

Pax7d/FKHR (Fig. 7 B), strongly supports the assertion that *MyoD* is a direct transcriptional target of Pax7 and Pax3. Together, these data provide important insight into the mechanism by which Pax7 and Pax3 enforce myogenic commitment.

Discussion

Pax7 is a critical regulator of adult skeletal myogenesis and is required for the generation of functional myogenic precursor cells from quiescent satellite cells (Kuang et al., 2006; Seale et al., 2000). Without it, *Pax7*-null mice have essentially no ability to repair damage to their skeletal muscle (Kuang et al., 2006). Whereas its paralog Pax3 has been heavily studied due to its essential role in embryogenesis, Pax7 has only received attention more recently following the discovery of its role in adult myogenesis (Seale et al., 2000). Although their homology and patterns of expression suggest that Pax7 and Pax3 are likely to share overlapping functions, their direct substitution has demonstrated that this overlap is incomplete and that each has unique responsibilities (Relaix et al., 2004). This suggests that further study of Pax7 may provide insight into the role of Pax3, but might just as easily reveal novel functions specifically provided by Pax7.

We therefore undertook to define the target genes regulated by Pax7 in myogenic cells. By stably expressing Pax7 in the myogenic C2C12 cell line, a set of genes was defined by microarray analysis as candidates for regulation. A subset was confirmed by real-time PCR as being differentially expressed in response to Pax7. Notably, these targets were affected quite differently by alternate isoforms of Pax7, pointing to a functional importance for the $GL^{+/-}$ splicing change. The myogenic factor *Myf5* was

identified as one of the candidates, and further work showed that the changes in mRNA expression were reflected at the level of protein. Strikingly, the same $GL^{+/-}$ splicing change that was important for the regulation of the other target genes was also essential in determining the ability of Pax7 to induce Myf5 expression in a non-muscle fibroblast cell line. The aRMS oncogenic fusion protein Pax7d/FKHR produced high levels of Myf5 expression but was also able to induce MyoD. This, as well as the finding that Pax7 is physically associated with one of two enhancers that regulate *MyoD* expression, provides molecular support to its position in myogenesis upstream from the MRFs.

Based on our microarray analysis, we chose several genes with large putative changes for validation. *PlagL1* (*pleomorphic adenoma-like 1*; *Zac1*; *Lot1*) expression showed the largest increase. *PlagL1* is normally expressed in brain, bone, and muscle (Valente et al., 2005) and is a tumour suppressor gene with anti-proliferative and pro-apoptotic effects (Spengler et al., 1997), possibly by enhancing p53-dependent gene activation (Huang et al., 2001). *Lix1* (*limb expression 1*), expressed in developing limb and brain (Moeller et al., 2002) but otherwise poorly described, was also changed substantially. *Trim54* (*tripartite motif-containing 54*; *MURF-3*; *Rnf30*) is expressed specifically in cardiac and skeletal muscle tissue to stabilize the microtubule network, and is required for skeletal muscle differentiation and fusion (Spencer et al., 2000). Its 7-fold increase in expression with Pax7 is thus quite interesting. *Mest* (*mesodermal-specific transcript*; *Peg1*) (Kaneko-Ishino et al., 1995) is an imprinted gene that was up-regulated; it has been observed in developing kidney (Kanwar et al., 2002) and heart (King et al., 2002), and its expression in 3T3-L1 cells can increase expression of *PPARγ*

(Takahashi et al., 2005). Interestingly, *PPARγ* is also increased 3.4-fold in our microarray data (Table III). Rat *Cipar1* (*castration-induced prostatic apoptosis-related protein 1; PARM1*) is induced in castrated rat prostate and may be involved in apoptosis and prostate tumour immortalization (Bruyninx et al., 1999; Cornet et al., 2003). *Syne2* (*synaptic nuclear envelope 2*) was increased in the microarray analysis, but this change could not verified in later trials. These genes (and others in Tables II & III) will need to be further characterized in order to detail their roles in myogenesis and their relationships to Pax7.

Defining a set of uniquely Pax7-regulated target genes creates an important toolkit for investigating Pax7. It is noteworthy that they are at best poorly affected by Pax3, and therefore determination of their roles in myogenesis will be of great interest. They will also provide a useful baseline for comparison of Pax7's responsibilities in the adult to those in the developing somites or brain. This will assist in discovering how Pax7 is providing functions that are useful in both myogenesis and neurogenesis or that are uniquely reserved for skeletal muscle. Finally, current assays for Pax7 function have been limited to those available for Pax3, such as *in vitro* binding assays to consensus Pax3 sites (Vogan et al., 1996), Pax3-responsive luciferase reporter genes *in vitro* (Epstein et al., 1996), or a Pax3-responsive transgenic reporter *in vivo* (Relaix et al., 2003). Having highly responsive targets allows for specific study of Pax7 at the structural level, a task for which they have already been enlisted (J. Ishibashi, unpublished data; I. McKinnell, personal communication).

The validated targets were up-regulated by Pax7d in C2C12 cells but decreased in primary myoblasts (Fig. 2). As primary myoblasts express much higher endogenous levels of Pax7 compared to C2C12, ectopic Pax7 may interfere with its function. Strong activation by Pax7d/FKHR offset or overcome this effect. These were specific effects of Pax7 and Pax7d/FKHR because *Syne2* (which failed validation in C2C12) was not substantially changed by either Pax7 or Pax7d/FKHR in primary myoblasts. The different responses of C2C12 and primary myoblasts are therefore consistent for the candidate Pax7 targets genes.

The differential effects of Pax7c and -d versus -a and -b on target gene expression (including *Myf5* in 10T1/2 cells) are intriguing, as it has been previously shown that splicing changes within the DNA binding domains modify the selection of DNA binding sites (Vogan et al., 1996). In this case, increased activation correlates with the absence of the GL dipeptide, which is present in the -a and -b isoforms (Table I). Pax3 is effectively GL^-, as this alternative splicing event does not occur with *Pax3* (Du et al., 2005). Therefore, expression of multiple Pax7 isoforms in myoblasts may provide developmental versatility in the regulation of global Pax7 target gene expression. Given that our initial microarray work has concentrated on Pax7d, it will be interesting to learn if the GL^+ isoforms preferentially regulate a distinct set of target genes. Similarly, the $Q^{+/-}$ change can be anticipated to affect distinct subsets of targets. The possibility that the proportions of the four isoforms are dynamically regulated during development is also being examined.

The effect of Pax7d on Myf5 expression is consistent across three cell types with a range of myogenic potentials and endogenous Pax7 expression (Fig. 4, 5, and 6). Particularly interesting is the induction of Myf5 in 10T1/2 cells, which have no existing expression of the gene (Fig. 6). This occurs without MyoD, negating the possibility that cross-activation amongst the MRFs leads to the appearance of Myf5. It also counters the argument that increased Myf5 expression is a compensatory response to decreased MyoD levels, similar to the pattern seen in the *MyoD*-null mouse (Rudnicki et al., 1992) and its myoblasts (Sabourin et al., 1999). Therefore, activation of *Myf5* as a direct transcriptional target of Pax7 survives as a leading hypothesis. Using bacterial artificial chromosome transgenics, Zammit et al. (2004) showed that 140 kb of *Myf5* upstream sequence was required to recapitulate all aspects of *Myf5* expression, while a region between -88 kb and -140 kb was necessary for consistent expression of *Myf5* in satellite cells (Fig. S2 B, available at http://). Scanning of this 140 kb region with the Pax3 paired-box consensus site (Chalepakis and Gruss, 1995) and homeobox site ('taat', Hayashi and Scott, 1990) yields eleven close matches (Fig. S2 C, available at http://), with one site at -96 kb closely resembles the optimized Pax3 binding site described by Vogan & Gros (1997). However, the limitations of such an *in silico* approach are evidenced by the absence of such a consensus site at the *MyoD* core enhancer despite our ChIP experimental evidence (Fig. 8). Further work will be required to determine if, and where, Pax7 binds to DNA in the *Myf5* gene.

Regulation of *MyoD* expression by Pax7 is consistent with existing data. Pax3 induction of *MyoD* was shown during embryogenesis (Maroto et al., 1997) and the

concurrent loss of *Pax3* and *Myf5* prevents *MyoD* expression (Tajbakhsh et al., 1997). More recently, *MyoD* was shown to be down-regulated in response to a dominant-negative Pax7 (Relaix et al., 2006) and increased with Pax7 expression in a *Pax7*-negative C2C12 clone (Zammit et al., 2006). The latter result contrasts with our lack of effect on MyoD in a pool of C2C12 cells (Fig. 3, 4, and 6), which could also relate to our high level of ectopic Pax7d expression (Fig. S1, available at http://). However, direct association of Pax7d with the *MyoD* gene (Fig. 8) in primary myoblasts with decreased MyoD expression suggests that Pax7 can directly modulate *MyoD* expression, perhaps also as a repressor. Thus, by increasing Myf5 but limiting MyoD, Pax7 may act to maintain a proliferating myoblast population by lowering its propensity for differentiation (Ishibashi et al., 2005). In such a model, the high level of MyoD activation by Pax7d/FKHR (Fig. 7) would therefore represent mis-expression due to the FKHR domain, as has been seen with the Pax3/FKHR fusion at other loci (Fredericks et al., 1995).

Pax7 has a critical role in the satellite lineage, but many questions remain as to how this is accomplished. An amenable environment (of co-factors, DNA structure, signaling, and other influences) is a probable prerequisite for Pax7 function, and may underlie the variable activity of Pax7 in different cell types. The combination of isoforms introduces still greater complexity. Identification of Pax7-regulated genes is an important entry point for addressing all of these issues, and further application to study of the Pax7/FKHR fusion in aRMS will yield insight into the tumour phenotype. However, the induction of Myf5 by Pax7 remains most notable, for it points to a critical convergence at

the "nodal point" of myogenic factor induction (Weintraub et al., 1991) and thus provides a molecular connection between Pax7 and myogenic commitment.

Materials & Methods

Cell Culture: C2C12 myoblasts were cultured in DMEM with 10% fetal calf serum and 1% penicillin/streptomycin. Primary myoblasts were isolated from adult C57BL/6 mice as described previously (Huh et al., 2004) and cultured on rat-tail collagen-coated plates (rat-tail collagen; Roche-Boehringer) in Ham's F-10 medium with 20% FCS, 1% penicillin/streptomycin, and 2.5ng/mL human recombinant bFGF (Invitrogen).

Stable pools were produced by infecting C2C12 cells or primary myoblasts with retrovirus made using the 3-plasmid HIT system (Soneoka et al., 1995) as described (Ishibashi et al., 2005). Expression plasmids were based on the pHAN backbone (with puromycin resistance driven from a distinct SV40 promoter) into which Pax7, Pax7d/FKHR, or Pax3 was cloned along with sequence for a carboxy-terminal 3xFLAG epitope tag. Control virus expressed EGFP and puromycin-resistance or puromycin-resistance alone. Infected pools were drug selected with 1μg/mL puromycin (Sigma) in growth medium for at least five days, by which time uninfected control cells had been uniformly obliterated.

Microarray Analysis: Triplicate pools of C2C12 cells stably infected with retrovirus expressing Pax7d, or with empty-control retrovirus, were produced. Total RNA was purified from cultured cells using the RNeasy Mini kit (Qiagen) according to the

manufacturer's instructions and quantified by spectrophotometry (OD_{260}). Samples hybridized to MOE430A GeneChips (Affymetrix, Inc.) at the Ottawa Genome Centre and analyzed as previously described (Ishibashi et al., 2005). Criteria used to derive Tables II and III included: log-fold change of greater than one (i.e. 2-fold cut-off); significant change (Increase or Marginal Increase call); detectable expression above background (non-Absent call) for the higher signal, and consistency between replicate samples. Raw microarray data is available from StemBase (http://www.scgp.ca:8080/StemBase/; Ontario Genomics Innovation Centre; Perez-Iratxeta et al., 2005) as experiment E102 (samples S137, S310) and from the National Center for Biotechnology Information's Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/; Barrett et al., 2005; Edgar et al., 2002) under series accession no. GSE3224 (GSM72628, -30, -32, -34...6).

Primers: PCR primers for RT- and real-time PCR were designed using the online Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi; Whitehead Institute) (Rozen and Skaletsky, 2000). Wherever possible, primers were designed to span at least one intron. Primer sequences specifications are provided in Supplementary Table S1.

RT-PCR: Total RNA was isolated as with microarray samples. For each set of samples, equal masses of total RNA (500-750ng) were used as template for first-strand reverse transcription (RT) using the RNA PCR Core kit (Perkin Elmer) with random hexamer

primers. RT products were diluted 2-5 fold with TE to produce a common pool for analysis with multiple primer pairs.

SYBR green real-time PCR reactions were performed in duplicate using an MX4000 PCR machine (Stratagene), gathering fluorescence data on FAM (SYBR green/experimental) and ROX (loading normalization) channels. Reactions included 1x iQ SYBR green Supermix (Stratagene), 30nM ROX passive reference dye, 50nM of each forward and reverse PCR primer, and 2µl of diluted template in a 20µl volume. Fluorescence data was gathered for 40 cycles of {94°C - 30 s; 58°C - 60 s; 72°C - 30 s}. Primer specificity was validated by denaturation curve analysis (55-94°C) and direct sequencing of the PCR products. Amplification curve plotting and C_t value calculation were performed using the MX4000 software (v4.20; Stratagene), with further calculations performed using Excel (Microsoft). Traditional agarose gel RT-PCR was performed as described above except that component reagents, including standard Taq DNA polymerase (New England Biolabs), replaced the SYBR green Supermix. The number of cycles to be used was estimated from the real-time PCR plot for *GAPDH* (24 cycles), *Myf5* (27 cycles), *MyoD* (34 cycles), and endogenous *Pax7* (37 cycles).

Western Blots: Total protein was harvested in RIPA lysis buffer fortified with protease inhibitors (Complete-Mini; Roche-Boehringer). Protein concentration was determined by Bradford assay (Biorad), after which samples (10-25µg each) were subjected to SDS-PAGE and electroblotted onto Immobilon-P membrane (Millipore). Membranes were blocked in 5% non-fat milk in PBS, prior to sequential probing with primary antibody

and HRP-conjugated secondary antibody in blocking solution. Target proteins were visualized by ECL (Amersham-Pharmacia) on Biomax XAR film (Kodak). Primary antibodies used were: anti-FLAG (M2; 1:x; Sigma); anti-Myf5 (C-20; 1:1250; Santa Cruz); anti-MyoD (C-20; 1:1250; Santa-Cruz); anti-Pax7 (hybridoma supernatant; 1:5-1:10; DSHB); anti-α-tubulin (1:x; Sigma). Secondary antibodies were HRP-conjugated anti-mouse and anti-rabbit (1:4000; Biorad).

Chromatin Immunoprecipitation (ChIP)-PCR: ChIP was performed as generally described and with the solutions listed in the ChIP kit instructions (Upstate Biotechnology), except as follows. Formaldehyde cross-linked (1.0% for 10 minutes at room temperature) protein/DNA complexes were isolated from $4x10^6$ growing primary myoblasts by lysis in 400µl of lysis buffer. Genomic DNA was fragmented to an average size of 500-600bp by sonication (Vibra-cell; Sonics and Materials Inc.) with 2 x 10 s pulses at 20% power. After centrifugation (10 minutes at 14,000 rpm at 4°C), the supernatant was split into 2x200µl aliquots and diluted with 1200µl of ChIP dilution buffer each. A 40µl aliquot (~1.3%) was removed as "Input". Following pre-clearing with protein-A/agarose for 30 min at 4°C, FLAG-tagged protein complexes were immunoprecipitated with 2.5µg per 10^6 cells (5µg/tube) of monoclonal mouse anti-FLAG (M2) antibody (Sigma) overnight at 4°C. Immune complexes were captured with protein-A/agarose beads (Upstate Biotechnology), washed through a series of buffers, and eluted. Input and eluted ChIP DNA were purified by phenol-chloroform extraction

and ethanol precipitated, with 2µl GlycoBlue (Ambion) added to visualize the pellet. DNA was resuspended in 50µl of 10mM Tris pH8.0.

The degree of DNA sonication was analysed by agarose gel electrophoresis of 20µl of the resuspended Input sample followed by ethidium bromide staining. The quantities of specific genomic fragments in the ChIP samples were assayed by SYBR green real-time PCR. Enrichment of specific fragments was calculated between ChIP samples from FLAG versus non-FLAG controls, with normalization against the non-specific background observed for irrelevant loci *p107*, *Pgc1*, and *GAPDH* to verify that changes were specific. The widespread agreement (no enrichment) between samples at the majority of the loci validates the normalization. The identities of the PCR targets are shown in Supplementary Table S2.

Immunocytochemistry: Cells were fixed with 4% paraformaldehyde, permeabilized with 0.5% Triton X-100, and blocked with 5% normal goat serum in PBS. Primary and secondary antibodies were applied in blocking solution. Primary antibodies used included anti-Pax7 (hybridoma supernatant; Developmental Studies Hybridoma Bank) and anti-Myf5 (C-20; Santa-Cruz). Secondary antibodies were anti-mouse and anti-rabbit conjugated to fluorescein (Chemicon). 4',6-diamidino-2-phenylindole (DAPI) (0.25µg/mL in PBS) was used to stain nuclei. Samples were mounted in fluorescence mounting medium (DAKO), coverslipped, and imaged using an Axiophot 2 microscope, Plan-NEOFLUAR objective (10x/0.30; Ph1; ∞/0.17) and AxioCam digital camera

(Zeiss). Digital images were captured using Axiovision (Zeiss) and processed with Photoshop (Adobe).

Binding Site Search: Genomic sequences from LocusLink (NCBI) were imported into a custom database and searched with the Pax3 paired-box consensus sequence 'tcgtcacrchyha' (Chalepakis and Gruss, 1995) and homeobox consensus sequence 'taat' (Hayashi and Scott, 1990) in both directions. The separation between paired-box and homeobox was limited to 10bp or less (Chalepakis et al., 1994).

Acknowledgements

We thank Dr. V. Sartorelli for providing retroviral expression plasmids; and Fabien Le Grand, Shihuan Kuang, Kazuki Kuroda, and Mark Gillespie for careful reading of the manuscript. M.A. Rudnicki holds the Canada Research Chair (CRC) in Molecular Genetics and is a Howard Hughes Medical Institute (HHMI) International Scholar. This work was supported by grants to M.A. Rudnicki from the Muscular Dystrophy Association, the National Institutes of Health, the Canadian Institutes of Health Research, the HHMI, and the CRC Program.

References

Asakura, A., G.E. Lyons, and S.J. Tapscott. 1995. The regulation of MyoD gene expression: conserved elements mediate expression in embryonic axial muscle. *Dev Biol.* 171:386-98.

Barrett, T., T.O. Suzek, D.B. Troup, S.E. Wilhite, W.C. Ngau, P. Ledoux, D. Rudnev, A.E. Lash, W. Fujibuchi, and R. Edgar. 2005. NCBI GEO: mining millions of expression profiles--database and tools. *Nucleic Acids Res.* 33:D562-6.

Berkes, C.A., and S.J. Tapscott. 2005. MyoD and the transcriptional control of myogenesis. *Semin Cell Dev Biol.* 16:585-95.

Bischoff, R. 1994. The satellite cell and muscle regeneration. *In* Myology. A.G. Engel, Franzini-Armstrong, C., editor. McGraw-Hill, New York. 97-118.

Bober, E., T. Franz, H.H. Arnold, P. Gruss, and P. Tremblay. 1994. Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development.* 120:603-12.

Bruyninx, M., B. Hennuy, A. Cornet, P. Houssa, M. Daukandt, E. Reiter, J. Poncin, J. Closset, and G. Hennen. 1999. A novel gene overexpressed in the prostate of castrated rats: hormonal regulation, relationship to apoptosis and to acquired prostatic cell androgen independence. *Endocrinology.* 140:4789-99.

Chalepakis, G., and P. Gruss. 1995. Identification of DNA recognition sequences for the Pax3 paired domain. *Gene.* 162:267-70.

Chalepakis, G., J. Wijnholds, and P. Gruss. 1994. Pax-3-DNA interaction: flexibility in the DNA binding and induction of DNA conformational changes by paired domains. *Nucleic Acids Res.* 22:3131-7.

Charge, S.B., and M.A. Rudnicki. 2004. Cellular and molecular regulation of muscle regeneration. *Physiol Rev.* 84:209-38.

Cornet, A.M., E. Hanon, E.R. Reiter, M. Bruyninx, V.H. Nguyen, B.R. Hennuy, G.P. Hennen, and J.L. Closset. 2003. Prostatic androgen repressed message-1 (PARM-1) may play a role in prostatic cell immortalisation. *Prostate.* 56:220-30.

Du, S., E.J. Lawrence, D. Strzelecki, P. Rajput, S.J. Xia, D.M. Gottesman, and F.G. Barr. 2005. Co-expression of alternatively spliced forms of PAX3, PAX7, PAX3-FKHR and PAX7-FKHR with distinct DNA binding and transactivation properties in rhabdomyosarcoma. *Int J Cancer.* 115:85-92.

Edgar, R., M. Domrachev, and A.E. Lash. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30:207-10.

Epstein, J.A., P. Lam, L. Jepeal, R.L. Maas, and D.N. Shapiro. 1995. Pax3 inhibits myogenic differentiation of cultured myoblast cells. *J Biol Chem.* 270:11719-22.

Epstein, J.A., D.N. Shapiro, J. Cheng, P.Y. Lam, and R.L. Maas. 1996. Pax3 modulates expression of the c-Met receptor during limb muscle development. *Proc Natl Acad Sci U S A.* 93:4213-8.

Fredericks, W.J., N. Galili, S. Mukhopadhyay, G. Rovera, J. Bennicelli, F.G. Barr, and F.J. Rauscher, 3rd. 1995. The PAX3-FKHR fusion protein created by the t(2;13)

translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3. *Mol Cell Biol.* 15:1522-35.

Goldhamer, D.J., B.P. Brunk, A. Faerman, A. King, M. Shani, and C.P. Emerson, Jr. 1995. Embryonic activation of the myoD gene is regulated by a highly conserved distal control element. *Development.* 121:637-49.

Goulding, M.D., G. Chalepakis, U. Deutsch, J.R. Erselius, and P. Gruss. 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *Embo J.* 10:1135-47.

Hayashi, S., and M.P. Scott. 1990. What determines the specificity of action of Drosophila homeodomain proteins? *Cell.* 63:883-94.

Holterman, C.E., and M.A. Rudnicki. 2005. Molecular regulation of satellite cell function. *Semin Cell Dev Biol.* 16:575-84.

Huang, S.M., A.H. Schonthal, and M.R. Stallcup. 2001. Enhancement of p53-dependent gene activation by the transcriptional coactivator Zac1. *Oncogene.* 20:2134-43.

Huh, M.S., M.H. Parker, A. Scime, R. Parks, and M.A. Rudnicki. 2004. Rb is required for progression through myogenic differentiation but not maintenance of terminal differentiation. *J Cell Biol.* 166:865-76.

Ishibashi, J., R.L. Perry, A. Asakura, and M.A. Rudnicki. 2005. MyoD induces myogenic differentiation through cooperation of its NH2- and COOH-terminal regions. *J Cell Biol.* 171:471-82.

Jostes, B., C. Walther, and P. Gruss. 1990. The murine paired box gene, Pax7, is expressed specifically during the development of the nervous and muscular system. *Mech Dev.* 33:27-37.

Kablar, B., K. Krastel, C. Ying, S.J. Tapscott, D.J. Goldhamer, and M.A. Rudnicki. 1999. Myogenic determination occurs independently in somites and limb buds. *Dev Biol.* 206:219-31.

Kaneko-Ishino, T., Y. Kuroiwa, N. Miyoshi, T. Kohda, R. Suzuki, M. Yokoyama, S. Viville, S.C. Barton, F. Ishino, and M.A. Surani. 1995. Peg1/Mest imprinted gene on chromosome 6 identified by cDNA subtraction hybridization. *Nat Genet.* 11:52-9.

Kanwar, Y.S., A. Kumar, K. Ota, S. Lin, J. Wada, S. Chugh, and E.I. Wallner. 2002. Identification of developmentally regulated mesodermal-specific transcript in mouse embryonic metanephros. *Am J Physiol Renal Physiol.* 282:F953-65.

Kassar-Duchossoy, L., B. Gayraud-Morel, D. Gomes, D. Rocancourt, M. Buckingham, V. Shinin, and S. Tajbakhsh. 2004. Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature.* 431:466-71.

Kay, P.H., and M.R. Ziman. 1999. Alternate Pax7 paired box transcripts which include a trinucleotide or a hexanucleotide are generated by use of alternate 3' intronic splice sites which are not utilized in the ancestral homologue. *Gene.* 230:55-60.

King, T., Y. Bland, S. Webb, S. Barton, and N.A. Brown. 2002. Expression of Peg1 (Mest) in the developing mouse heart: involvement in trabeculation. *Dev Dyn.* 225:212-5.

Kuang, S., S.B. Charge, P. Seale, M. Huh, and M.A. Rudnicki. 2006. Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. *J Cell Biol.* 172:103-13.

Mansouri, A., A. Stoykova, M. Torres, and P. Gruss. 1996. Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice. *Development.* 122:831-8.

Maroto, M., R. Reshef, A.E. Munsterberg, S. Koester, M. Goulding, and A.B. Lassar. 1997. Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. *Cell.* 89:139-48.

Mauro, A. 1961. Satellite cell of skeletal muscle fibers. *J Biophys Biochem Cytol.* 9:493-5.

Megeney, L.A., B. Kablar, K. Garrett, J.E. Anderson, and M.A. Rudnicki. 1996. MyoD is required for myogenic stem cell function in adult skeletal muscle. *Genes Dev.* 10:1173-83.

Moeller, C., M.B. Yaylaoglu, G. Alvarez-Bolado, C. Thaller, and G. Eichele. 2002. Murine Lix1, a novel marker for substantia nigra, cortical layer 5, and hindbrain structures. *Brain Res Gene Expr Patterns.* 1:199-203.

Perez-Iratxeta, C., G. Palidwor, C.J. Porter, N.A. Sanche, M.R. Huska, B.P. Suomela, E.M. Muro, P.M. Krzyzanowski, E. Hughes, P.A. Campbell, M.A. Rudnicki, and M.A. Andrade. 2005. Study of stem cell function using microarray experiments. *FEBS Lett.* 579:1795-801.

Relaix, F., D. Montarras, S. Zaffran, B. Gayraud-Morel, D. Rocancourt, S. Tajbakhsh, A. Mansouri, A. Cumano, and M. Buckingham. 2006. Pax3 and Pax7 have distinct

and overlapping functions in adult muscle progenitor cells. *J Cell Biol.* 172:91-102.

Relaix, F., M. Polimeni, D. Rocancourt, C. Ponzetto, B.W. Schafer, and M. Buckingham. 2003. The transcriptional activator PAX3-FKHR rescues the defects of Pax3 mutant mice but induces a myogenic gain-of-function phenotype with ligand-independent activation of Met signaling in vivo. *Genes Dev.* 17:2950-65.

Relaix, F., D. Rocancourt, A. Mansouri, and M. Buckingham. 2004. Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev.* 18:1088-105.

Relaix, F., D. Rocancourt, A. Mansouri, and M. Buckingham. 2005. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature.* 435:948-53.

Rozen, S., and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* Bioinformatics Methods and Protocols: Methods in Molecular Biology. S. Krawetz and S. Misener, editors. Humana Press, Totowa, NJ. 365-386.

Rudnicki, M.A., T. Braun, S. Hinuma, and R. Jaenisch. 1992. Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell.* 71:383-90.

Rudnicki, M.A., P.N. Schnegelsberg, R.H. Stead, T. Braun, H.H. Arnold, and R. Jaenisch. 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell.* 75:1351-9.

Sabourin, L.A., A. Girgis-Gabardo, P. Seale, A. Asakura, and M.A. Rudnicki. 1999. Reduced differentiation potential of primary MyoD-/- myogenic cells derived from adult skeletal muscle. *J Cell Biol.* 144:631-43.

Sabourin, L.A., and M.A. Rudnicki. 2000. The molecular regulation of myogenesis. *Clin Genet.* 57:16-25.

Seale, P., L.A. Sabourin, A. Girgis-Gabardo, A. Mansouri, P. Gruss, and M.A. Rudnicki. 2000. Pax7 is required for the specification of myogenic satellite cells. *Cell.* 102:777-86.

Sebire, N.J., and M. Malone. 2003. Myogenin and MyoD1 expression in paediatric rhabdomyosarcomas. *J Clin Pathol.* 56:412-6.

Soneoka, Y., P.M. Cannon, E.E. Ramsdale, J.C. Griffiths, G. Romano, S.M. Kingsman, and A.J. Kingsman. 1995. A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res.* 23:628-33.

Spencer, J.A., S. Eliazer, R.L. Ilaria, Jr., J.A. Richardson, and E.N. Olson. 2000. Regulation of microtubule dynamics and myogenic differentiation by MURF, a striated muscle RING-finger protein. *J Cell Biol.* 150:771-84.

Spengler, D., M. Villalba, A. Hoffmann, C. Pantaloni, S. Houssami, J. Bockaert, and L. Journot. 1997. Regulation of apoptosis and cell cycle arrest by Zac1, a novel zinc finger protein expressed in the pituitary gland and the brain. *Embo J.* 16:2814-25.

Tajbakhsh, S., D. Rocancourt, G. Cossu, and M. Buckingham. 1997. Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell.* 89:127-38.

Takahashi, M., Y. Kamei, and O. Ezaki. 2005. Mest/Peg1 imprinted gene enlarges adipocytes and is a marker of adipocyte size. *Am J Physiol Endocrinol Metab.* 288:E117-24.

Tapscott, S.J., A.B. Lassar, and H. Weintraub. 1992. A novel myoblast enhancer element mediates MyoD transcription. *Mol Cell Biol.* 12:4994-5003.

Valente, T., F. Junyent, and C. Auladell. 2005. Zac1 is expressed in progenitor/stem cells of the neuroectoderm and mesoderm during embryogenesis: differential phenotype of the Zac1-expressing cells during development. *Dev Dyn.* 233:667-79.

Vogan, K.J., and P. Gros. 1997. The C-terminal subdomain makes an important contribution to the DNA binding activity of the Pax-3 paired domain. *J Biol Chem.* 272:28289-95.

Vogan, K.J., D.A. Underhill, and P. Gros. 1996. An alternative splicing event in the Pax-3 paired domain identifies the linker region as a key determinant of paired domain DNA-binding activity. *Mol Cell Biol.* 16:6677-86.

Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T.K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, and et al. 1991. The myoD gene family: nodal point during specification of the muscle cell lineage. *Science.* 251:761-6.

Xia, S.J., J.G. Pressey, and F.G. Barr. 2002. Molecular pathogenesis of rhabdomyosarcoma. *Cancer Biol Ther.* 1:97-104.

Zammit, P.S., J.J. Carvajal, J.P. Golding, J.E. Morgan, D. Summerbell, J. Zolnerciks, T.A. Partridge, P.W. Rigby, and J.R. Beauchamp. 2004. Myf5 expression in

satellite cells and spindles in adult muscle is controlled by separate genetic elements. *Dev Biol.* 273:454-65.

Zammit, P.S., L. Heslop, V. Hudon, J.D. Rosenblatt, S. Tajbakhsh, M.E. Buckingham, J.R. Beauchamp, and T.A. Partridge. 2002. Kinetics of myoblast proliferation show that resident satellite cells are competent to fully regenerate skeletal muscle fibers. *Exp Cell Res.* 281:39-49.

Zammit, P.S., F. Relaix, Y. Nagata, A.P. Ruiz, C.A. Collins, T.A. Partridge, and J.R. Beauchamp. 2006. Pax7 and myogenic progression in skeletal muscle satellite cells. *J Cell Sci.* 119:1824-32.

Ziman, M.R., and P.H. Kay. 1998. Differential expression of four alternate Pax7 paired box transcripts is influenced by organ- and strain-specific factors in adult mice. *Gene.* 217:77-81.

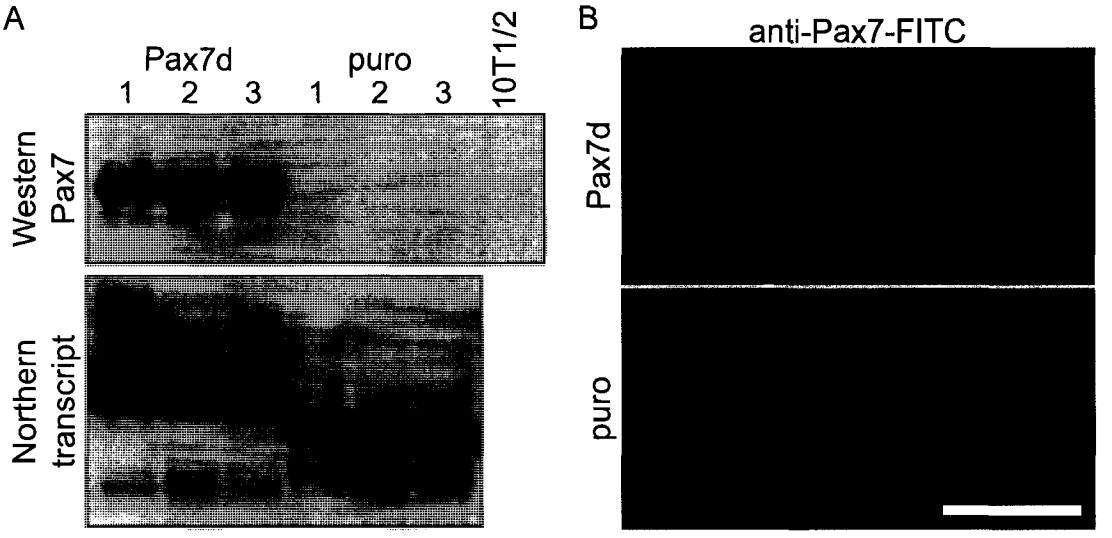
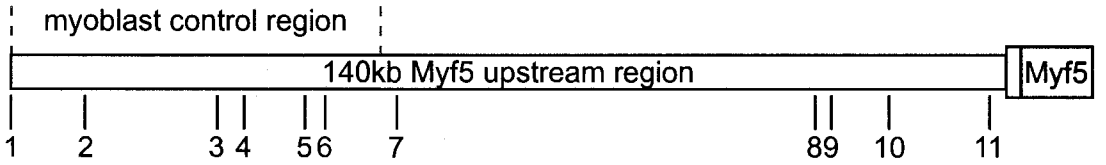


Figure S1. Pax7 expression in C2C12 myoblasts. (A) Pax7d protein was highly expressed in infected C2C12 myoblast pools by Western analysis (a low level of Pax7 was detectable in puro controls with a longer exposure; 10T1/2 fibroblasts were negative). By Northern analysis, both Pax7 and control pools expressed similar levels of the retroviral transcript. (B) Ectopic expression of Pax7d in C2C12 was readily detectable by immunostaining with anti-Pax7 monoclonal antibody compared to the very low level of endogenous Pax7. (Pax7 = green) Bar, 100 μm.

A

	forward	reverse complement
Pax3/PD (Chalepakis et al., 1995)	a a A C tcgTCAC Cc cA G T t t	a a A C Tg gG GTGAcga G T t t
P3-C-OPT (Vogan & Gros, 1997)	tggtgGTCACGCCTCA g TGAATATTA t	

B



C

#	Location	Matches (/13)	Spacing to homeobox	Site
1	-140232	11	5 bp	TAGTGTATGACaAttcatATTA
2	-129618	11	10 bp	ATTAATgttgattgctTCaTCACAGTTCA
3	-111114	11	overlap	ATTATGTGGGAtGA
4	-107226	11	1 bp	TCaTCACATCTCAcTAAT
5	-98660	11	10 bp	TAATttgctacaacTaGTGACACACAA
6	**-95792**	**12**	**3 bp**	TgGTCACACCTTAggaATTA
7	-85664	11	10 bp	TGATGTGTAAaGAttcaaactttTAAT
8	-26777	11	3 bp	TAATTAcagTCaTGACACATTAAT
9	-24505	11	0 bp	ATTATTGGGTTTGAtGA
10	-16445	11	8 bp	ATTATTAaatctgatTGAAGAGTGACtA
11	-239	11	3 bp	ATTAtgtTAAAGTGTGATcA

Figure S2. Pax3 consensus DNA binding sites within the Myf5 regulatory region. (A) Several groups have described closely related sequences bound by Pax3 or other paired-type homeodomains. In particular, Chalepakis et al. (1995) identified a 13bp consensus sequence for Pax3-paired domain binding, while Vogan & Gros (1997) identified an optimized Pax3 paired + homeodomain binding site. (B) A critical 88 kb-140 kb region is required for Myf5 expression in satellite cells and myoblasts (Zammit, 2004). (C) Scanning of the 140 kb of Myf5 upstream sequence required to recapitulate Myf5 expression (Zammit, 2004) identified 11 segments with close matches (11/13 bases) to the Pax3 consensus that were also within 10 bp of a homeodomain consensus site. One site (#6) was very closely matched (12/13 bases) to the consensus as well as to the optimized P3-C-OPT sequence, although most sites had mismatches in the core 'gtcacgc' motif described by Jun & Desplan (1996). paired domain (green); homeodomain (blue); overlapping site (orange); mismatch with core base (red), non-core base (lowercase)

Table S1. Sequences of PCR primers used for SYBR green real-time PCR

			forward		reverse		
Gene	GID	amplicon	sequence	length	sequence	length	Spans Introns?
beta-actin	NM_007393	167 bp	5' accctgtgctgctcacc 3'	17 mer	5' tccggagtccatcacaat 3'	18 mer	#3 (893bp)
Cipar1 (ex4)	AK008716	281 bp	5' ttggggaccacagaggat 3'	18 mer	5' ggtgaatcgacgtgcaaa 3'	18 mer	No
Cipar1 (ex1/2)	NM_145562	141 bp	5' ttggtcagacccaggaaact 3'	20 mer	5' caatgggactgttggtgaac 3'	20 mer	#1 (75860bp)
Cipar1 (ex2/3)	NM_145562	133 bp	5' acctggcaaagtgatctgtg 3'	20 mer	5' aaacaccagcaacactacgg 3'	20 mer	#2 (18496bp)
Syne2	BF582734	269 bp	5' acccgggtcagctctttt 3'	18 mer	5' gccaacagctcaaccaga 3'	18 mer	No
GAPDH	NM_008084	228 bp	5' tcggtgtgaacggatttg 3'	18 mer	5' ggtctcgctcctggaaga 3'	18 mer	No
Lix1	NM_025681	288 bp	5' gcagcagaaagccacctt 3'	18 mer	5' gggtcatccgcatcatct 3'	18 mer	#2 (16329bp), #3 (2180bp)
Mest	NM_008590	370 bp	5' gcaacctggtcatcgaca 3'	18 mer	5' tgatggccaggacctctt 3'	18 mer	No
Myf5	NM_008656	85 bp	5' acagcagctttgacagcatc 3'	20 mer	5' aagcaatccaagctggacac 3'	20 mer	#2 (436bp)
MyoD	NM_010866	86 bp	5' cgctccaactgctctgatg 3'	19 mer	5' tagtaggcggtgtcgtagcc 3'	20 mer	#1 (441bp)
Pax7	NM_011039	328 bp	5' gctaccagtacagccagtatg 3'	21 mer	5' gtcactaagcatgggtagatg 3'	21 mer	#8 + #9 (8030bp)
PlagL1	NM_009538	98 bp	5' ccactaccacagccacagat 3'	20 mer	5' tgctgctgaggttgcagt 3'	18 mer	#8 (33bp)
Trim54	NM_021447	342 bp	5' agcggaacctgctagtgg 3'	18 mer	5' tctgtctgcggctgttgt 3'	18 mer	#3 (1891bp)

Table S2. ChIP PCR Target Loci

Name	Product	Primer Sequence and Location		
MyoD (-20kb)		MyoD core enhancer, ~20kb upstream from start site.		
-A	237 bp	large PCR product covering most of enhancer		
		F - 5' gggcatttatgggtcttcct	3'	20 mer
		R - 5' ccctaggcctgagctagaga	3'	20 mer
-B	103 bp	smaller PCR product towards distal end of enhancer		
		F - 5' gcttctttcggccaagtatc	3'	20 mer
		R - 5' ccaactggctgtgttgtga	3'	19 mer
MyoD (-5kb)	202 bp	MyoD DRR enhancer region, ~5kb upstream from start site.		
		F - 5' acaggtccagactgggtagg	3'	20 mer
		R - 5' tttcagctcccttggctagt	3'	20 mer
Myf5 (-239bp)		Myf5 promoter region, ~239bp upstream from start site.		
-A	322 bp	larger product centred at -270bp		
		F - 5' atccatgaaaatgccacctc	3'	20 mer
		R - 5' ggtccctttgacgctaatga	3'	20 mer
-B	159 bp	smaller product centred at -190bp		
		F - 5' aatgtcttgctaccgtgctg	3'	20 mer
		R - 5' ctggtccctttgacgctaat	3'	20 mer
Myf5 (-38kb)		Myf5 gene, approximately 38kb upstream from start site.		
-A	340 bp	centred at ~ -37kb		
		F - 5' gaacctcagagggatttcttcaaa	3'	24 mer
		R - 5' tgccccaatctactgataaggatt	3'	24 mer
-B	461 bp	centred at ~ -37.5kb		
		F - 5' caggatgtgtctgtatcagacgtg	3'	24 mer
		R - 5' agttgttttctacacacccaccaa	3'	24 mer
p107	290 bp	4kb 3' of p107 locus		
		F - 5' aaaaggccaggtctcatcct	3'	20 mer
		R - 5' tcttcacctccaggagtgct	3'	20 mer
Pgc1	191 bp	2kb 3' of Pgc1 locus		
		F - 5' cttcgtttgggaaactcagc	3'	20 mer
		R - 5' atggcagcgactccatactc	3'	20 mer
GAPDH	228 bp	within GAPDH coding sequence		
		F - 5' tcggtgtgaacggatttg	3'	18 mer
		R - 5' ggtctcgctcctggaaga	3'	18 mer

Supplemental Data - Myf5-Luciferase Reporters

The previous work in Chapter 3 demonstrated that Myf5 expression is regulated by Pax7. However, it was not shown that this regulation was a direct result of Pax7 binding to DNA in the Myf5 gene. Dual-luciferase reporter assays were used to test the functional requirement for specific Myf5 gene regulatory sequences in driving Pax7 responsive gene expression. A candidate region that could be involved in Pax7's effect on Myf5 was identified; however, unexpectedly broad gene expression changes limit the strength of this conclusion.

Results and Discussion

The goal for this reporter assay was the demonstration that a response was due to the interaction of our transfected gene with a putative regulatory sequence in the reporter. It was therefore incumbent that the basic reporter plasmid (lacking the sequence of interest) and the normalization gene were themselves stable and invariant between treatments. These assumptions were tested, in order to assess the practicality of assaying candidate Pax7-responsive regulatory sequences in transfected cells.

Pax7 expression plasmids were co-transfected with firefly luciferase (Fluc) and Renilla luciferase (Rluc) reporters into 10T1/2, C2C12, or primary myoblasts. Pax7/FKHR was used in parallel to Pax7, as it (like Pax3/FKHR) has been shown to be a strong transactivator of Pax target genes. The initial Fluc reporters were highly variable in the presence of Pax7 or Pax7/FKHR (Supplemental Figure S3). The common Fluc reporter

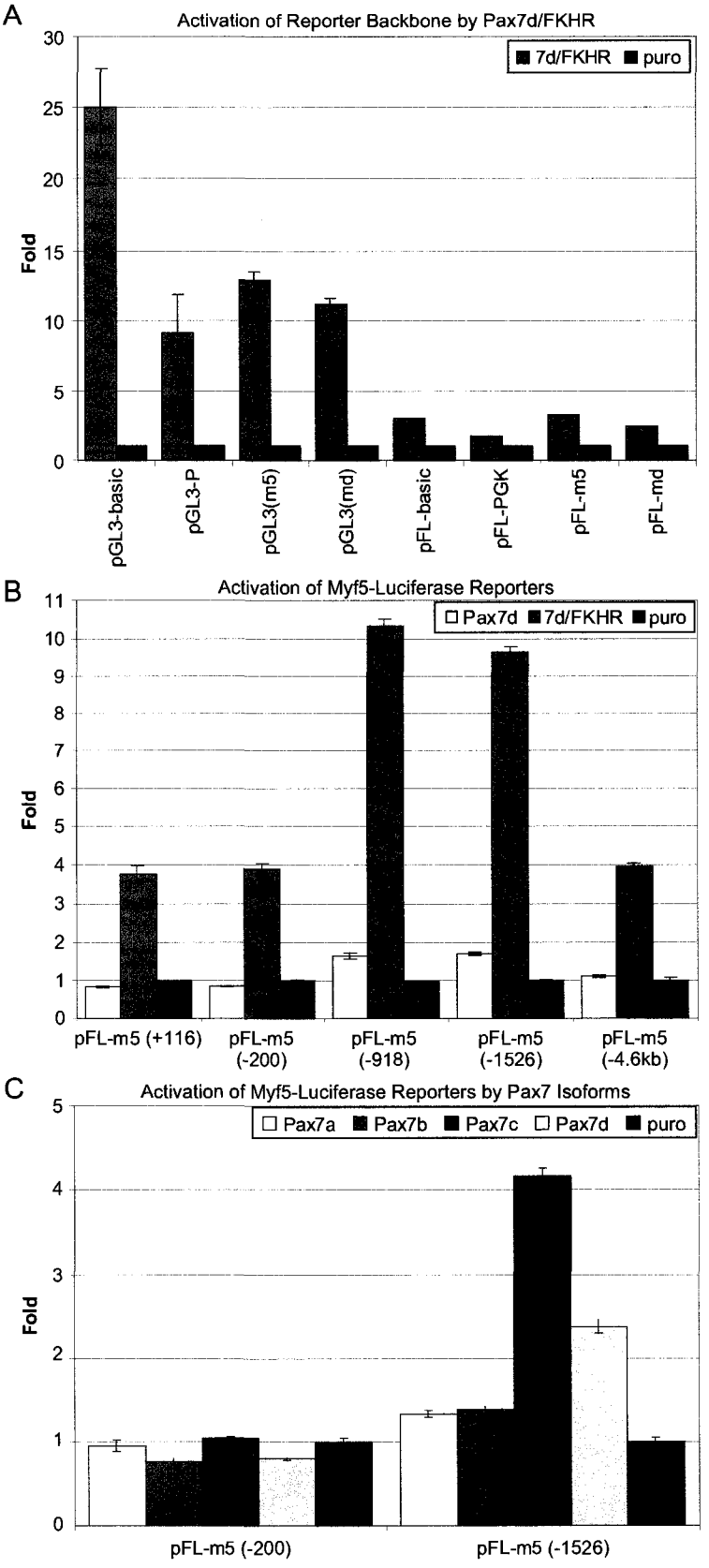


Figure S3. Pax7 activates Myf5-Luciferase reporter plasmids through an element between -200 bp and -918 bp. (A) Firefly luciferase reporter plasmids with no promoter or only minimal promoters are upregulated by Pax7d/FKHR. The pGL3 backbone was particularly responsive, whereas the reconstructed pFL backbone was much more stable. Note that the absolute activities were much lower in the absence of a promoter. (B) Addition of a region of the Myf5 promoter between -200 and -918 correlates with a slight increase in luciferase reporter activity with Pax7d and with Pax7d/FKHR. (C) The extended -1526 region is responsive to Pax7c and Pax7d, whereas the shorter promoter fragment is not.

plasmid pGL3, driven only by a minimal SV40 promoter, was induced up to 25-fold by Pax7/FKHR. The pGL3-basic plasmid, which entirely lacks a promoter, was also up-regulated several-fold. To overcome this problem, the reporter was reconstructed in an alternative backbone. The PGK promoter produced high-level expression and mild (2-fold) responsiveness to Pax7/FKHR. Alternative, minimal promoters of 150-200 bp (plus 5'UTR) from the Myf5 and MyoD genes (pFL-M5 and pFL-MD, respectively) showed moderate basal expression that was unaffected by Pax7 and minimally (2- to 3-fold) increased by Pax7/FKHR.

The Rluc reporters driven by common viral promoters such as SV40 or CMV were also responsive to Pax7 (not shown). This had the unfortunate consequence of creating artifactual changes in the normalized Fluc response. The PGK promoter was the most stable of the set in the presence of Pax7 and Pax7/FKHR, with variation of less than 2-fold.

Fragments of Myf5 and MyoD genomic DNA were cloned into the pFL-M5 vector so that they could be tested for regulation by Pax7. A series of proximal Myf5 gene fragments were initially used, starting from +116 bp, -200 bp, -918 bp, -1526 bp, and from -4.6kb (relatively to the transcriptional start site), in increasing order of length. Transfection of these reporters along with a Pax7d expression plasmid into 10T1/2 produced a 2-fold increase in activity that correlated with the addition of the -918 bp to -200 bp region (Supplemental Figure S3B). In agreement with the specific activation of Myf5 by Pax7 isoforms -C and -D, the (-918) and (-1526) reporters were up-regulated by Pax7C and Pax7D by 2-fold relative to Pax7A, Pax7B, or the control (Supplemental Figure S3C). The magnitude of this activation is similar to that of Myf5 mRNA expression in C2C12. These

data therefore suggest that a Pax7C/D responsive element exists in the Myf5 -918 bp to -200 bp region. One potential Pax3-consensus site is found within this region (Figure S2). However, given the variability encountered with the basic Fluc and Rluc reporter plasmids, closer examination of this fragment will be required to isolate the specific sequence motif involved and to determine if it is physically interacting with Pax7 protein.

Materials and Methods

Cloning

The pFL backbone was created by substituting the firefly luciferase coding sequence from pGL3 for the Renilla luciferase sequence in pRL. Myf5 and MyoD promoter fragments were amplified by PCR, inserting an NcoI restriction sites across the ATG start site at the 3' end of the fragment. These products were cloned into the pGL3 and pFL backbones using that NcoI site to align the fragments with the luciferase ATG. Longer upstream portion of the Myf5 promoter were retrieved from the plasmid pMyf5(5kb)EGFP and cloned seamlessly into pFL-M5.

Cell Culture & Transfections

Fibroblast 10T1/2 cells seeded on 6 cm plates were transfected by calcium phosphate precipitate with 5μg of pFL reporter plasmid, 2-4μg of expression plasmid, and 50ng of pRL-PGK Renilla luciferase plasmid for 48 hours. Cells were collected, lysed with passive lysis buffer and freeze-thawed at -20°C. Lysates were assayed with the dual luciferase assay kit (Promega) using an auto-injecting luminometer (Berthold).

Chapter 4

MyoD induces myogenic differentiation through cooperation of its NH_2- and COOH-terminal regions

Purpose

To establish the differential roles of MyoD and Myf5 in proliferating myoblasts, and to identify the molecular domains involved in providing those distinct functions.

Submission Information

Submitted: 16 February 2005
Accepted: 3 October 2005
Published: 7 November 2005

Reproduced from "**The Journal of Cell Biology, 2005, Volume 171, pp. 471-82**" by copyright permission of The Rockefeller University Press.

Contributions of Co-Authors

Jeff Ishibashi did all of the experimental work and analysis.

Robert Perry provided the *MyoD*$^{-/-}$;*Myf5*$^{-/-}$ double-knockout fibroblast cell lines.

Atsushi Asakura provided the raw microarray data comparing wild-type primary myoblasts with *MyoD*$^{-/-}$ primary myoblasts.

MyoD Induces Myogenic Differentiation Through Cooperation of its NH_2- and COOH-Terminal Regions

Jeff Ishibashi [1,2], Robert L. Perry [3], Atsushi Asakura [2], and Michael A. Rudnicki [1,2,3]

[1] University of Ottawa, Ottawa, Ontario, Canada
[2] Ottawa Health Research Institute, Ottawa, Ontario, Canada
[3] McMaster University, Hamilton, Ontario, Canada

Corresponding Author: Michael A. Rudnicki
Ottawa Health Research Institute
501 Smyth Road
Ottawa, ON K1H 8L6
CANADA
phone: 613-739-6740
fax: 613-737-8803
email: <mrudnicki@ohri.ca>

Running Title: Cooperation between MyoD domains

Abstract

MyoD and Myf5 are basic helix-loop-helix transcription factors that play key but redundant roles in specifying myogenic progenitors during embryogenesis. However, there are functional differences between the two transcription factors that impact myoblast proliferation and differentiation. Target gene activation could be one such difference. We have used microarray and polymerase chain reaction approaches to measure the induction of muscle gene expression by MyoD and Myf5 in an *in vitro* model. In proliferating cells, MyoD and Myf5 function very similarly to activate the expression of likely growth phase target genes such as *L-myc*, *m-cadherin*, *Mcpt8*, *Runx1*, *Spp1*, *Six1*, *IGFBP5*, and *Chrnβ1*. MyoD, however, is strikingly more effective than Myf5 at inducing differentiation-phase target genes. This distinction between MyoD and Myf5 results from a novel and unanticipated cooperation between the MyoD NH_2- and COOH-terminal regions. Together, these results support the notion that Myf5 functions toward myoblast proliferation, whereas MyoD prepares myoblasts for efficient differentiation.

Introduction

The process of skeletal muscle differentiation is orchestrated by a family of four conserved basic helix-loop-helix (bHLH) transcription factors that are collectively known as myogenic regulatory factors (MRFs). Mice harboring single mutations of either *MyoD* or *Myf5* are viable and do not have overt muscle phenotypes, suggesting that MyoD and Myf5 have considerable overlap in their roles (Braun et al., 1992; Rudnicki et al., 1992). However, either MyoD or Myf5 is required for proper myogenesis during embryogenesis because compound *MyoD* / *Myf5*-null mice lack essentially all skeletal muscle tissue at birth (Rudnicki et al., 1993; Kassar-Duchossoy et al., 2004). In contrast, myogenin is important for the terminal differentiation and fusion of myoblasts into mature muscle fibers (Rawls et al., 1998; Vivian et al., 2000). MRF4 appears to have a role as a determination factor in a subset of myocytes in the early somite and as a differentiation factor in later muscle fibers (Kassar-Duchossoy et al., 2004).

Postnatal growth and regeneration of skeletal muscle is mediated primarily by a pool of myogenic stem cells known as satellite cells, which reside adjacent to the fibers. In response to damage through injury or exercise, these satellite cells activate expression of MyoD and Myf5 and undergo numerous rounds of proliferation as myoblasts. A small number of myoblasts return to a quiescent state, thus replenishing the pool of satellite cells; the remainder continue their differentiation, fusing into existing or new myofibres and expressing myogenin and MRF4 while down-regulating Myf5. In contrast to wild-type myoblasts, *MyoD*-null myoblasts grow more quickly, show aberrant expression of

muscle markers, and differentiate inefficiently (Sabourin et al., 1999), a phenotype that is the cause of the regeneration deficit exhibited by compound dystrophic (*mdx*) *MyoD*-null animals (Megeney et al., 1996). Conversely, *Myf5*-null myoblasts proliferate poorly and differentiate precociously (Montarras et al., 2000). The sequence of MRF expression in activated satellite cells, in conjunction with the phenotypes of single-null myoblasts and animals, argue that MyoD and Myf5 do not have identical roles in myoblast proliferation and induction of differentiation.

MyoD and Myf5 target genes have largely been examined after the onset of differentiation and, hence, are involved in producing the enormous phenotypic shift from a proliferating myoblast to a contractile, multinucleated muscle fiber. Although it has been suggested that MyoD and Myf5 transactivation is checked in growing myoblasts (whether by degradation [Thayer et al., 1989], modification [Lindon et al., 1998], signaling [Vaidya et al., 1989; Li et al., 1992], or interfering heterodimerization [Benezra et al., 1990]), the distinct phenotypic differences that are exhibited by growth phase *MyoD*- and *Myf5*-null myoblasts suggest that MyoD and Myf5 do have active roles in myoblasts. Indeed, Wyzykowski et al. (2002) identified *Id3* and *NP1* as target genes of MyoD under growth conditions.

Distinguishing the distinct functions of MyoD and Myf5 is complicated by their abilities to auto- and cross-activate expression from the endogenous loci (Tapscott et al., 1988; Braun et al., 1989; Thayer et al., 1989; Wyzykowski et al., 2002). Such a circular network could account for the stabilization and irreversibility of the commitment of a cell to a myogenic fate (Thayer et al., 1989; Weintraub et al., 1991a). However, gene

expression changes resulting from the introduction of exogenous MyoD or Myf5 could be an indirect effect that is mediated through the other MRF. Expression of MyoD in the absence of Myf5 (and vice versa) permits the comparative evaluation of each factor's functions in myogenic commitment.

To this end, we reintroduced MyoD or Myf5 into a *MyoD / Myf5* double-null (double knockout [dblKO]) fibroblast cell line that was maintained in high serum growth conditions. These cell lines are normally non-myogenic but can be converted to skeletal muscle upon the exogenous expression of MyoD or Myf5. Microarray analysis identified numerous differentially regulated genes, which were further validated by examining specific candidates using real-time PCR. A number of growth phase targets were identified, demonstrating that MyoD and Myf5 are transcriptionally active in proliferating cells. Surprisingly, we did not find unique targets, and both MRFs were able to induce the expression of these genes.

MyoD, however, was strikingly more effective at activating differentiation markers than Myf5. Additional support for functional differences between MyoD and Myf5 were obtained by using chimeric MRFs that interchanged their NH_2-terminal, bHLH, and COOH-terminal domains. The bHLH domains (DNA binding and dimerization) are highly conserved between the two genes, whereas their NH_2- and COOH-terminal regions are more divergent. Whereas Myf5 was inefficient at inducing differentiation gene expression, the activation of a cohort of these markers by the chimeric MRFs provided strong evidence for cooperative gene activation by the NH_2-terminal and bHLH + COOH-terminal regions of MyoD. *MyoD*-null primary myoblasts

have a greatly reduced expression of the same genes relative to wild-type myoblasts. Therefore, these data support the idea that Myf5 is biased toward myoblast proliferation, whereas MyoD promotes myoblast differentiation (Sabourin et al., 1999; Seale et al., 2001).

Results

MyoD and Myf5 activate skeletal muscle genes in growth conditions

To circumvent the potential problem of auto- and cross-activation by the primary myogenic factors (Braun et al., 1989; Thayer et al., 1989; Weintraub et al., 1991b; Hollenberg et al., 1993), we infected clonal double-null *MyoD*$^{-/-}$;*Myf5*$^{-/-}$ (dblKO) embryonic fibroblast lines with retrovirus expressing MyoD or Myf5 or with empty control retrovirus. With this approach, we were able to assess which genes were common targets of MyoD and Myf5 versus those that could be uniquely regulated by one or the other primary MRF. These genes could be directly activated by MyoD or Myf5 or could be indirectly activated through an intermediate transcription factor; we considered both classes to be downstream targets of the primary MRFs. Shortly after infection of the dblKO target cells with retrovirus, positive-expressing cells were purified by FACS based on GFP fluorescence expressed from the bi-cistronic retroviral transcript (Fig. 1, A and B). Pools of > 10^6 cells were then harvested for total RNA after a further 24 h of culture in high serum growth conditions. Fluorescently labeled probes generated from biological– triplicate RNA samples were hybridized to MG-U74Av2 GeneChips, each containing probesets directed at ~6,000 genes and an additional ~6,000 ESTs.

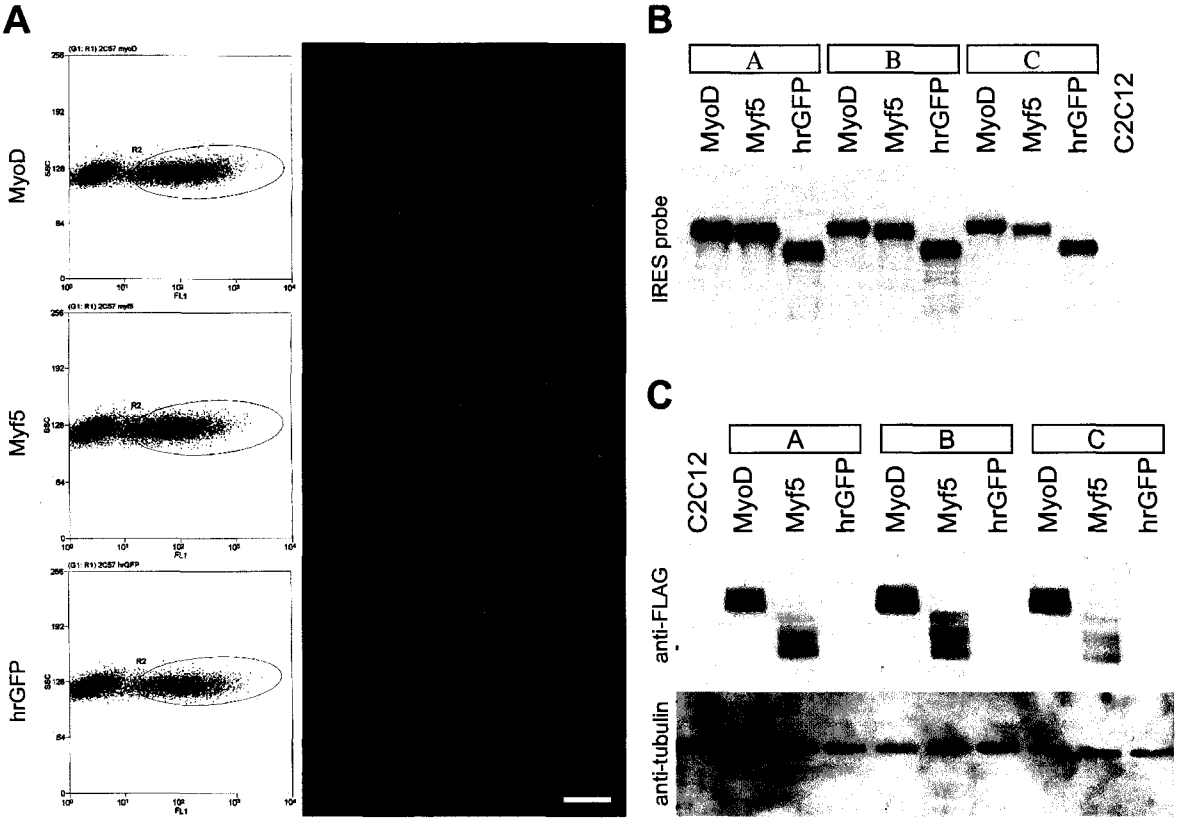


Figure 1. Preparation of RNA for GeneChip analysis. (A) Representative FACS plots of *MyoD-/-;Myf5-/-* fibroblasts infected with retrovirus expressing MyoD, Myf5, or no gene as well as GFP from an internal ribosomal entry site (IRES) within the same transcript. GFP expression amongst sorted cells after 24 h of culture was verified by fluorescence microscopy immediately before harvesting for total RNA. SSC, side scatter; FL1, fluorescence channel 1; G, gate; R, region. The circled regions denote the sorted populations. Bar, 100μm. (B) Northern blot demonstrating equivalent levels of retroviral transcript expression amongst samples. (C) Western blots demonstrating robust MyoD or Myf5 expression in corresponding samples.

Comparison of MyoD or Myf5 arrays with control arrays produced candidate lists that were considered to contain genes potentially regulated by MyoD or Myf5 during growth phase (Table I and Table S1, available at http://www.jcb.org/cgi/content/full/jcb.200502101/DC1).

MyoD expression in a *MyoD / Myf5* double-null background produced increases in 47 genes (including *L-myc* and *cadherin-15*), whereas only 17 genes were increased by Myf5. Of these genes, 11 were common targets of both MyoD and Myf5; eight were activated to similar degrees by both (*Chrnβ 1, Mcpt8, Spp1, Six1, Runx1, Idb2, Ugcg,* and *Kctd12*), whereas the others were more strongly activated by MyoD (*IGFBP5, H19,* and α-*actin*). In contrast, only six genes were down-regulated by either MyoD or Myf5 (none by both), and only four had fold changes of < -2 (*Tcf20,* -4.9; *Dlk1,* -2.9; *TgfβR3,* -2.7; and *S100a13,* -2.0). Relaxing the stringency of the selection criteria had only a moderate effect, producing 70 increases by MyoD and 32 increases by Myf5 versus 5 and 14 decreases, respectively (Tables S1–S4, available at http://www.jcb.org/cgi/content/full/jcb.200502101/DC1).

One goal of our study was to identify targets that are uniquely regulated by Myf5 but not by MyoD (and vice versa). However, very few genes were suggested by the microarray data to be increased by Myf5 and not by MyoD, and real-time PCR directed at several of those transcripts, in turn, did not support them as targets (*Skiip,* Table I; *Refbp1* and *Snk,* unpublished data). In contrast, 36 genes were up-regulated by MyoD but not by Myf5. The majority of these targets (e.g., *myogenin, myosin heavy chain,* and *troponin-T*), however, are associated with skeletal muscle differentiation.

Table I - Candidate MyoD/Myf5 target genes

		GeneChip (U74Av2)				Real-time PCR	
		Fold[1]			Call[2]	Fold[3]	
Gene	GID	MyoD	Myf5	myogenin	GFP	MyoD	Myf5
Transcription Factors							
myogenin	X15784	10.1	1.2	2.8	A	27.0	8.4
sine oculis-related homeobox 1 (Six1)	X80339	3.8	3.5	2.2	P/A	2.6	2.2
inhibitor of DNA binding 2 (Idb2)	AF077861	3.6	2.4	2.7	P	-	-
runt related transcription factor 1 (Runx1)	D26532	3.4	2.9	3.3	P/A	4.3	2.7
hairy and enhancer of split 6 (Hes6)	AW048812	2.7	0.9	1.1	A	1.3	1.0
lung carc. myc related oncogene 1 (L-myc)	X13945	2.3	1.5	0.8	A	[4] 4.7	2.6
Adhesion and Receptors							
cholinergic receptor, nicotinic, beta-1	M14537	24.6	16.0	15.8	A	5.0	6.1
cholinergic receptor, nicotinic, gamma	X03818	9.3	2.0	4.8	A	18.1	16.4
cholinergic receptor, nicotinic, alpha-1	M17640	8.4	2.6	2.5	A	3.2	1.9
cadherin 15 (m-cadherin)	AJ245402	4.2	0.5	1.9	A	[4] 290.6	85.4
transmembrane 4 superfamily member 6	AF053454	2.6	2.3	2.5	P/A	-	-
discoidin domain receptor family, 1	L57509	2.3	1.6	2.6	A	-	-
Secreted Factors							
insulin-like growth factor binding protein 5	L12447	142.6	9.8	7.0	A	123.1	89.0
mast cell protease 8 (Mcpt8)	X78545	11.7	11.1	14.4	A	6.2	15.6
secreted phosphoprotein 1 (Spp1)	X13986	8.3	5.7	4.6	P	11.5	18.5
ESTs							
RIKEN cDNA 1190002N15 gene	AW125453	5.0	1.7	2.9	P/A	-	-
cDNA clone	AW120874	3.2	1.3	1.9	P	-	-
RIKEN cDNA 2610201A13 gene	AA222883	1.4	2.1	1.2	P	-	-
cDNA clone	AA796831	1.1	2.1	1.7	P	-	-
Others							
H19 fetal liver mRNA	X58196	35.6	6.1	13.7	P/M/A	-	-
C1q/tumor necrosis factor related protein 3	AI315647	5.8	1.9	3.1	P	[5] -	-
K+ channel tetramerisation domain 12	AI842065	4.9	6.0	3.2	P/A	[5] 1.1	1.3
protein kinase inhibitor, alpha (Pkia)	AW125442	4.6	1.6	2.1	A	1.8	1.0
UDP-glucose ceramide glucosyltransferase	AI853172	4.3	3.3	2.9	P	-	-
paternally expressed 3 (Peg3)	AF038939	3.6	1.0	1.9	P	-	-
enolase 3, beta muscle	X61600	3.5	1.8	2.0	P	-	-
H2B and H2A histones	X05862	3.3	0.6	1.7	A	-	-
alpha-methylacyl-CoA racemase	U89906	3.0	2.0	1.7	P/A	-	-
ankyrin repeat domain 1 (cardiac muscle)	AF041847	3.0	3.0	3.1	P	-	-
SH3-domain GRB2-like B1 (endophilin)	AI842874	2.2	1.4	2.0	P	-	-
WNT1 inducible signaling pathway 1	AF100777	2.2	2.3	2.6	P	-	-
ADP-ribosylation factor-like 6 interact. prot. 5	AW049647	2.1	1.5	1.6	P	-	-
SKI interacting protein (Skiip)	AW046671	1.2	4.8	3.6	P/A	1.0	0.5
enabled homolog (Drosophila)	D10727	1.9	2.4	2.9	P/A	-	-
Differentiation Markers							
actin, alpha, cardiac	M15501	240.7	3.1	9.6	A	42.9	10.1
troponin C, cardiac/slow skeletal	M29793	172.8	9.3	16.8	A	1.8	1.0
actin, alpha 1, skeletal muscle	M12347	93.3	9.2	14.7	M/A	265.6	130.9
troponin T3, skeletal, fast	L48989	49.3	2.8	10.4	A	-	-
myosin light chain, phosphorylatable, fast sk.	AV290649	43.5	3.0	4.6	A	3850	1540
myosin, light polypeptide 4	M19436	28.5	2.1	6.8	A	-	-
troponin T2, cardiac	L47600	21.4	2.6	3.4	P/A	-	-

myosin, heavy polypeptide 3, skeletal	M74753	17.0	0.8	2.0	A	-	-
troponin T1, skeletal, slow	AV213431	16.4	0.8	0.8	A	-	-
ATPase, Ca++ transporting, cardiac fast 1	X67140	11.0	1.1	1.9	A	-	-
troponin I, skeletal, slow 1	AJ242874	10.5	0.9	2.7	A	-	-
myosin, light polypeptide 1	X12973	9.3	0.7	1.6	A	-	-
retinoblastoma 1	M26391	5.5	1.6	1.5	A	-	-
ryanodine receptor 1, skeletal muscle	D38216	4.7	1.2	1.2	A	-	-
myosin binding protein H	U68267	3.9	0.5	1.5	A	-	-
sarcoglycan, beta	AB024921	3.5	1.0	1.8	A	-	-
myocyte enhancer factor 2A (MEF2A)	AW045443	3.1	1.2	1.4	P/A	0.8	0.9
growth arrest and DNA-damage-inducible 45 alp	U00937	2.5	0.9	1.1	P/A	-	-
cyclin-dependent kinase inhibitor 1A (P21)	AW048937	2.2	1.1	1.0	P	1.4	1.4

[1] Average fold change for pairwise comparisons of MyoD/Myf5/myogenin vs. GFP, from log-fold change in MAS5.0

[2] Present/Marginal/Absent call from Affymetrix MAS5.0

[3] Fold change based on ΔCt between MyoD/Myf5 vs. puro-alone Realtime PCR, normalized to GAPDH expression

[4] Added following manual inspection of dataset

[5] C1q and K+channel identified in current annotations; previously listed as ESTs

At least six of the identified targets are transcription factors (Table I) and may themselves regulate the expression of other genes. Foremost amongst them is myogenin, which is an MRF that is activated by MyoD immediately upon the switch to differentiation conditions (Hollenberg et al., 1993; Bergstrom et al., 2002). Therefore, we examined (by GeneChip analysis) the possibility that our candidates were activated indirectly by myogenin using myogenin retrovirus–infected dblKO cells that were prepared as for MyoD and Myf5. 50% (13/26) of myogenin targets were also downstream from MyoD and Myf5. However, for most non-differentiation class genes, our microarrays indicated that target activation was similar between MyoD, Myf5, and myogenin treatments, which contrasts with the considerable induction of myogenin by MyoD (10-fold) compared with Myf5 (1.2-fold;Table I). This suggests that these genes are common targets of MRFs rather than being strictly dependent on myogenin.

To identify potential growth phase targets of MyoD, probable differentiation markers were removed using previous work by Bergstrom et al. (2002). They used an inducible MyoD–ER fusion system to examine gene expression by microarray analysis during early differentiation in low serum conditions(Bergstrom et al., 2002). They identified nine subsets in their data by using a clustering algorithm to find coordinate patterns of temporal regulation, including those with transient increases, early or delayed increases, and decreases in expression. Of the 571 genes identified by Bergstrom et al. (2002), we mapped 298 to probesets on the MU74Av2 GeneChip on the basis of GenBank, Unigene, or LocusLink IDs. Of these 298, 22 genes were also found to be up-regulated by MyoD in our experiment, and an additional seven were also up-regulated by

Myf5. This provided added support to our data. A portion of these genes fell within Bergstrom clusters 5 and 6 and primarily represented differentiation-specific targets such as structural genes (e.g. *myosin* and *troponin*). However, others within clusters1–4 (early induction) and 7–9 (decreased expression through differentiation) were also seen to have increased expression in our GeneChip data and were considered as possible growth targets.

All of the genes identified in Table I are robustly expressed by proliferating wild-type primary myoblasts (Fig. 2). Most (47/53 = 89%) of these genes are decreased in *MyoD*$^{-/-}$ myoblasts, whereas the remainder (6/53) exhibit a mixture of increase/no change/decrease calls and moderate fold increases. Differentiation markers are vastly decreased (up to 140-fold for *myosin heavy chain-3* or *troponin T1*), whereas most putative growth-phase genes showed only moderate changes (e.g., approximately eight-fold for *m-cadherin* or *L-myc* and ~2.7–1.8- fold for *Six1* or *Runx1*).

Thus, in high serum conditions, MyoD and Myf5 are capable of regulating the transcription of numerous genes. The growth phase regulation of a selection of those candidates was then specifically examined.

Growth phase candidate validation by SYBR Green real-time PCR

The expression of these candidate genes was re-examined using a second set of independently derived RNA samples. This second set was produced by the infection of a distinct dblKO cell line with drug-selectable retrovirus, yielding proliferating puromycin-resistant pools of cells that were expanded for 2 wk under drug selection. SYBR Green

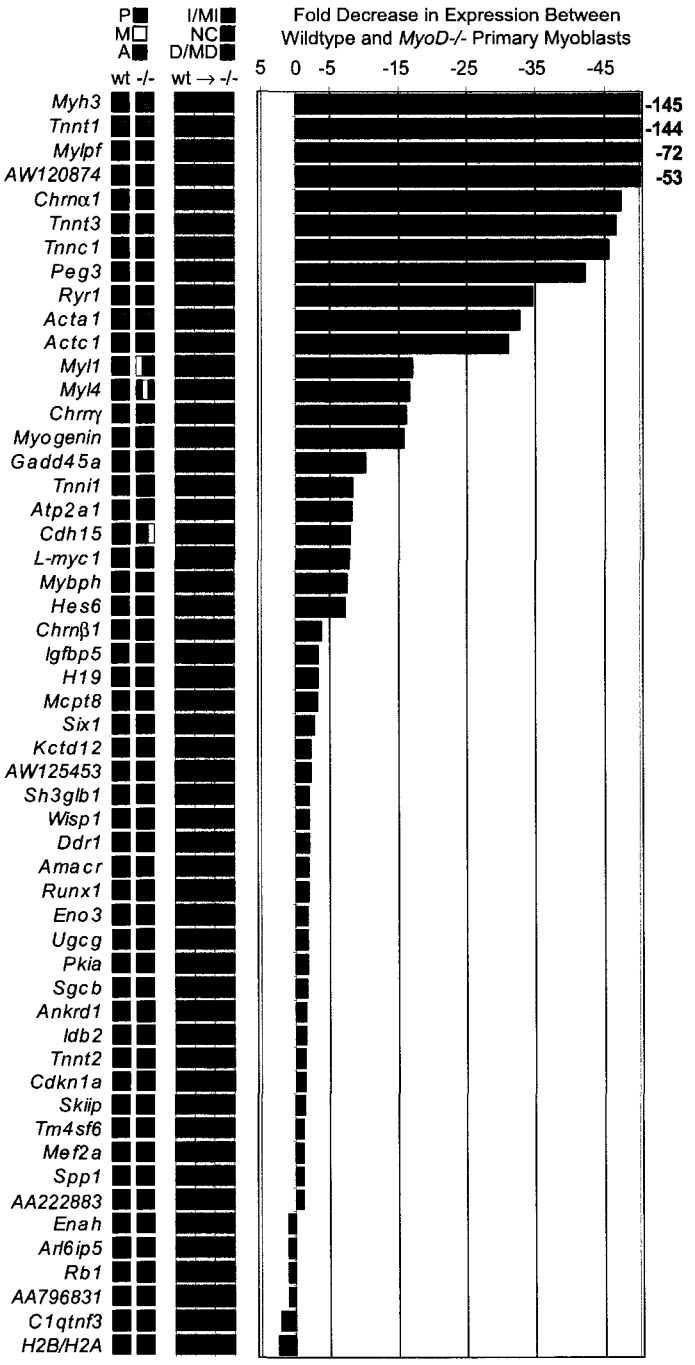


Figure 2. Decreased expression of differentiation markers by *MyoD-/-* primary myoblasts. Expression levels in *MyoD*-null primary myoblasts of the genes in Table I show that the majority of differentiation markers are greatly reduced relative to wild-type myoblasts (e.g., *Myh3* to *myogenin*). In contrast, genes that are regulated in growth phase by MyoD and Myf5 are reduced to a lesser degree, if at all (e.g., *Mcpt8, Six1*, and *Runx1*). Calls are shown for wild-type (n = 3) or *MyoD-/-* (n = 3) myoblasts. P, present; M, marginal; A, absent. Change calls are shown for nine pairwise comparisons between wild-type and *MyoD-/-* myoblasts. I, increase; MI, marginal increase; NC, no change; MD, marginal decrease; D, decrease.

real-time PCR was used to quantitate target gene transcript levels using PCR primers that were chosen to span at least one intron wherever possible. The specificity of the PCR was verified by denaturing curve analysis and direct sequencing of the products.

The majority of the genes that were selected as possible growth phase target genes in the GeneChip data were also found by real-time PCR to be up-regulated (Table I). The estimates of the fold up-regulation of *Six1*, *Runx1*, *L-myc*, *IGFBP5*, and *Mcpt8* were similar by either technique. *Spp1* was increased more, and *Chrnβ1* was increased slightly less as estimated by real-time PCR; nonetheless, each was significantly increased by MyoD and Myf5. The increases in *m-cadherin* and *myogenin* levels by real-time PCR were much greater than by microarray, probably as a result of the lower background of the PCR assay. Again, however, the consistency of the changes in target expression that were produced by MyoD and Myf5 contrast with the variable induction of *myogenin* by MyoD and Myf5, arguing that this is not a strictly indirect effect.

Approximately one-third of the candidates that were selected for real-time PCR validation did not exhibit significant changes in the stable pool samples (Table I). These genes tended to be those that had modest fold changes in the GeneChip data (e.g., *Hes6*, *Pkia*, *Skiip*, and *p21*). Probesets for differentiation markers that exhibited larger changes by microarray (such as *cardiac troponin C* and *MEF2A*) were likely seen as a result of spontaneous differentiation in the original GeneChip samples and were minimized in these proliferating cultures. These observations support the conclusion that the stable pools lacked the spontaneously differentiating cells that were observed in our original samples.

Wyzykowski et al. (2002) used the same inducible MyoD system as Bergstrom et al. (2002) to generate inputs for a representational difference analysis protocol. They identified *Id3* and *NP1* as growth phase targets of MyoD (Wyzykowski et al., 2002). Their cells were maintained in a high serum growth medium during MyoD induction, suggesting that *Id3* and *NP1* are not differentiation targets that are expressed simply as a consequence of serum deprivation. Rather, they are induced in the presence of cyclohexamide, indicating that they are likely to be direct targets of MyoD that do not require intervening protein synthesis for activation. In contrast, our GeneChip experiments did not reveal an induction of *Id3* or *NP1* in dblKO cells by either MyoD or Myf5. To assess whether this was a consequence of low sensitivity to those genes, we used SYBR Green real-time PCR to specifically examine the expression of *Id3* and *NP1* in both the GeneChip RNA samples and in RNA from drug-selected pools; however, no significant changes were detected (unpublished data). It is likely that these inconsistencies are attributable to the numerous differences in our experimental systems, including the method of MyoD expression, the type of host cell used, and possibly the presence/absence of endogenous *MyoD* and *Myf5* genes.

From this combination of GeneChip and real-time PCR analyses, we defined a validated set of growth phase targets. Notably, MyoD and Myf5 were both capable of regulating each of these genes, and none were strictly associated with one MRF. These targets were then applied to examine the differences in function between corresponding regions of MyoD and Myf5.

Association of growth phase gene activation with MyoD and Myf5 domains

Within their bHLH regions, Myf5 and MyoD exhibit 88% identity and > 95% homology at the amino acid level. In contrast, the regions that are NH_2 and COOH terminal to this highly conserved DNA-binding region exhibit considerably more differences in sequence and function (Gerber et al., 1997). Therefore, we hypothesized that functional differences between MyoD and Myf5 would be a consequence of their divergent NH_2- and COOH-terminal domains rather than a result of the bHLH recognition of discrete DNA sequences. To test this, chimeric MRFs were built, interchanging the MyoD and Myf5 NH_2- and COOH-terminal regions around their bHLH domains. This yielded six chimeric MRF genes (termed d5d, d55, 55d, 5d5, 5dd, and dd5, where d denotes the portion derived from MyoD and 5 denotes the portion derived from Myf5) in addition to the full-length wild-type MyoD and Myf5 (Fig. 3 A).

Retrovirus expressing one of the chimeric MRFs was introduced into dblKO fibroblasts, yielding drug-selected pools of $>10^5$ clones from which RNA and protein were extracted for analysis. The growth phase targets identified previously were then assayed within these samples by SYBR Green real-time PCR.

Both *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) transcript (input RNA quantity; reverse transcription) and MRF protein levels (MRF expression) were used to enable equitable comparisons by providing normalization between samples. *GAPDH* transcripts were quantified by real-time PCR. Relative protein expression levels of full-length and chimeric MRFs were derived by direct digital densitometry of Western blots probed with three primary antibodies that each recognized unique but overlapping

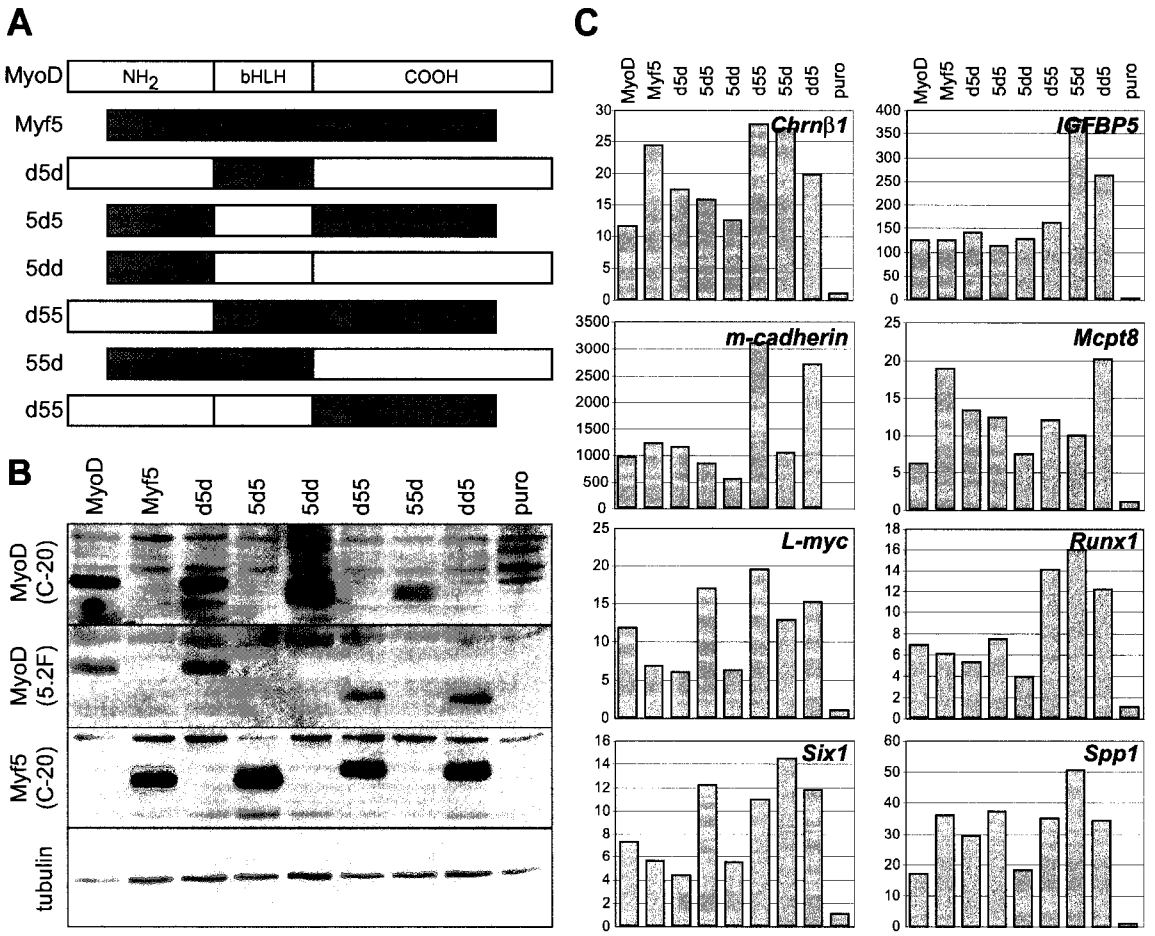


Figure 3. Gene expression induced by chimeras in growing dblKO cells. (A) Schematic of MyoD/Myf5 chimeras. Chimeric MRFs were constructed by interchanging the corresponding NH_2-terminal, bHLH-, and COOH terminal regions of MyoD and Myf5. (B) Levels of MRF and tubulin protein expression in each pool. The combination of three different epitopes that were recognized by MyoD and Myf5 antibodies was used to normalize the expression results in C for relative MRF expression. Puro, puromycin-resistant empty vector negative control. (C) Induction of transcripts for potential growth phase targets by each of MyoD, Myf5, and the chimeric MRFs (normalized to MRF protein and *GAPDH* transcript levels). Numbers on y-axis indicates fold changes.

sets of four MRFs (Fig. 3 B). All of the chimeras were demonstrated to be capable of regulating the genes identified in the previous analysis, producing substantial increases in expression compared with the puromycin-alone negative control pool.

After normalization, 5dd tended to be the poorest relative activator, whereas 55d was (in most cases) the best (Fig. 3 C). However, in contrast to the other samples, these two chimeras were subject to significant normalization corrections (Fig. 3 B), which may have tended to overstate their effects. Nonetheless, no single region of MyoD or Myf5 strictly correlated with enhanced or decreased relative activation. In certain cases (e.g., *Chrnβ1*, *m-cadherin*, and *Spp1*), there is a tantalizing suggestion that the Myf5 bHLH domain could have greater activity than that of MyoD, although this pattern is not borne out amongst the others. Chimeras dd5 and d55 have a greater effect than MyoD or d5d. This might suggest that an interaction between the flanking portions of MyoD acts to suppress transactivating activity, which is an effect that is disrupted when a portion of Myf5 is substituted. A similar effect was observed in deletion studies of MyoD (Weintraub et al., 1991b). Thus, for growth-phase genes, the corresponding domains of MyoD and Myf5 are otherwise interchangeable, and differences emerge only with respect to the degree of target induction.

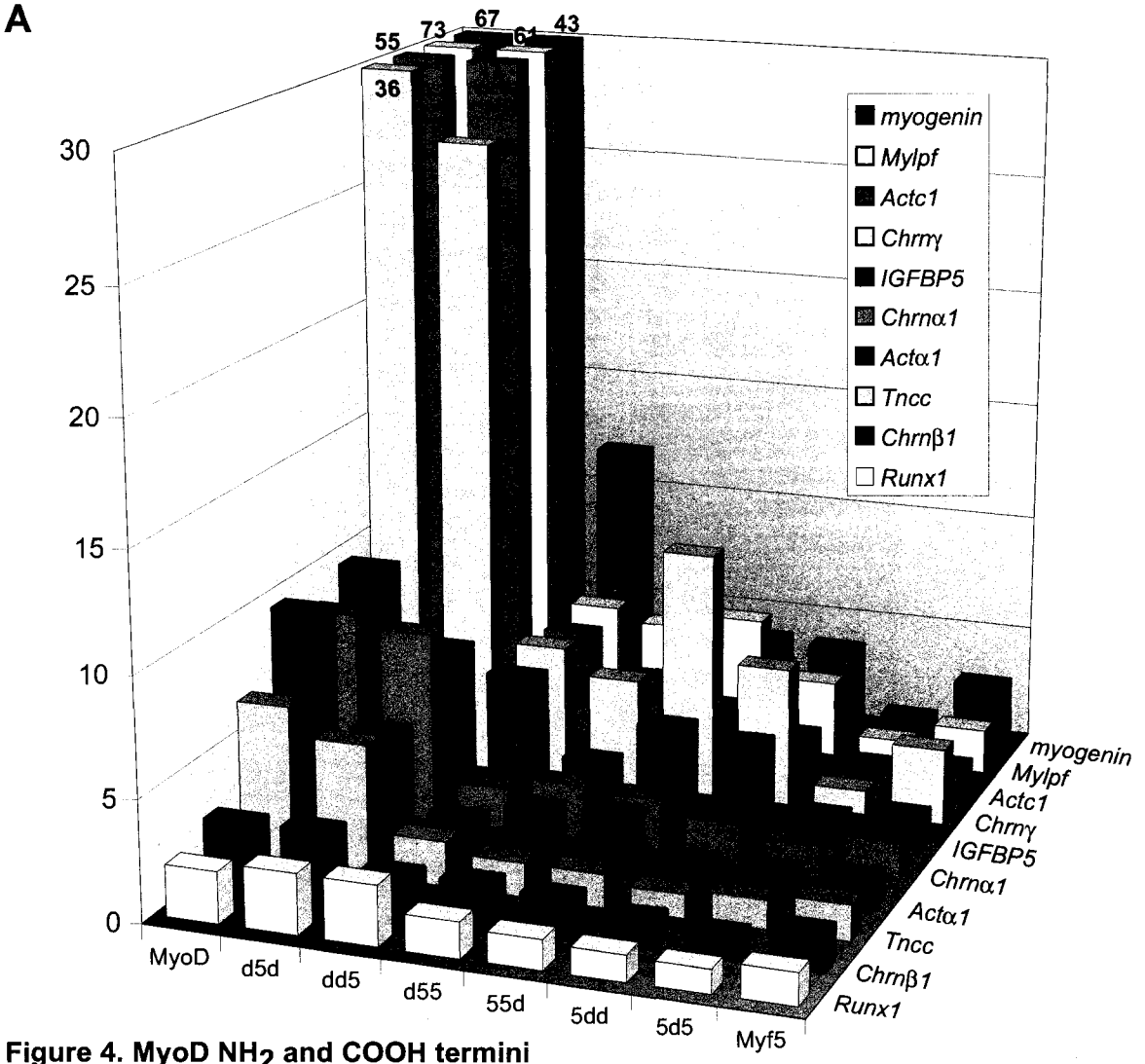
The MyoD NH_2- and COOH-terminal domains cooperate to induce differentiation

A significant number of potential MyoD target genes that were identified by GeneChip analysis were differentiation markers. Wild-type MyoD and Myf5 and chimeras were expressed in dblKO cells using retrovirus and were maintained for 3 d in

growth conditions before harvest. The proportion of infected cells was very similar between pools (see Fig. 5 A). Differentiation marker expression was examined by real-time RT-PCR (Fig. 4 A). This set of vectors included a COOH-terminal FLAG epitope tag that allowed for the normalization of gene expression against MRF protein levels (Fig. 4 B). Two growth phase markers (*Chrnβ1* and *Runx1*, identified in the aforementioned GeneChip experiment) showed little relative difference between MyoD, Myf5, and the chimeras. In concordance with the GeneChip results, however, the expression of MyoD produced a considerable activation of genes such as the *cholinergic receptor* α and γ subunits, *myogenin*, α-*actin*, *myosin*, and *troponin* (Fig. 4; Table II shows un-normalized changes vs. empty vector controls). In comparison, the level of Myf5 induction of these genes was moderate relative to the empty vector control.

Substitution of the MyoD COOH terminus into Myf5 (55d) was not sufficient in growth phase to cause differentiation and could produce only modest increases in the expression of numerous differentiation markers. Replacement with the MyoD NH_2 terminus alone (d55) also resulted in just moderate gene induction relative to full-length Myf5. Importantly, the concurrent presence of MyoD NH_2- and COOH-terminal regions (d5d) resulted in activity approaching that of full-length MyoD.

The same pools of dblKO cells expressing one of the wild-type or chimeric MRFs were challenged to differentiate under reduced serum conditions. When the MyoD COOH terminus was present (MyoD; d5d, 5dd, and 55d), there was enhanced differentiation compared with the corresponding Myf5 region (Myf5; dd5, d55, and 5d5; Fig. 5 B). Deletion of the putative cdk4-interacting domain (Zhang et al., 1999) from the



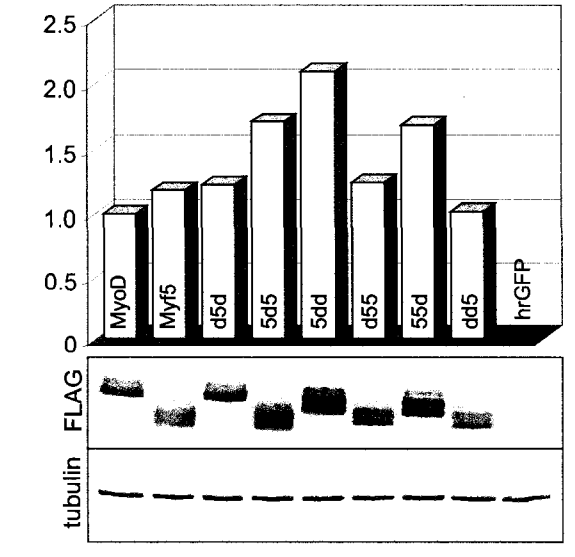


Figure 4. MyoD NH_2 and COOH termini cooperate to activate differentiation marker expression. Amongst a variety of differentiation genes, the d5d chimera had near wild-type activity, whereas the separated NH2 terminus and COOH terminus of MyoD had much less. In contrast, growth-phase genes *Chrnβ1* and *Runx1* were induced similarly by MyoD, Myf5, and the chimeras. (A) Gene expression measured by real-time PCR and normalized to MRF protein expression (B) and *GAPDH* transcript levels. Plotted as relative activity between MRFs. (B) Relative expression levels of FLAG-tagged chimeric MRFs in growth phase (normalized to tubulin). Numbers on y-axis indicate fold changes.

Table II - Fold-Activation of Myogenic Genes by Real-time PCR (relative to GFP; normalized to GAPDH, but NOT to FLAG protein)

	MyoD	d5d	5dd	55d	dd5	d55	5d5	Myf5	GFP
Actc1	89.2	68.5	6.5	7.8	8.9	7.6	2.8	2.5	1.0
myogenin	145.0	114.8	17.4	14.5	26.2	18.4	3.7	7.4	1.0
Mylpf	2645.3	2727.9	249.2	356.0	211.1	234.3	62.4	81.8	1.0
Chrng	48.5	45.3	16.9	24.4	8.0	8.0	2.3	5.2	1.0
Acta1	35.4	21.8	7.7	16.1	5.7	10.8	6.9	5.6	1.0
ChrnA1	8.7	10.0	3.1	3.6	1.8	2.9	1.7	1.9	1.0
IGFBP5	127.4	104.0	104.0	91.4	74.2	35.9	22.8	23.0	1.0
Tncc	5.1	5.1	1.6	1.8	1.4	1.3	1.6	1.3	1.0
Chrnb1	11.7	14.9	10.2	11.2	6.7	8.8	7.1	7.7	1.0
Runx1	2.4	3.4	2.6	2.3	2.8	2.0	1.9	1.8	1.0

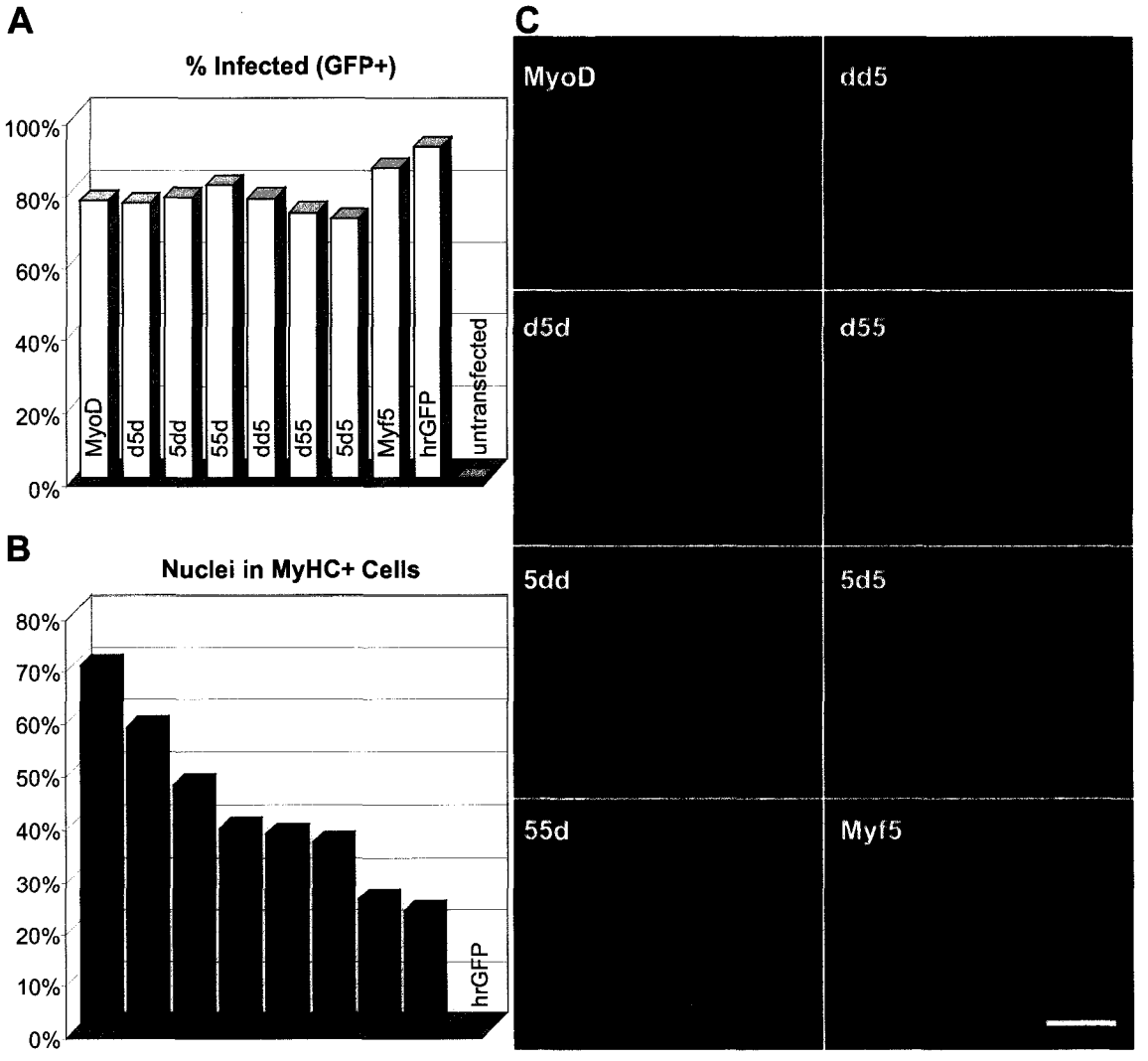


Figure 5. MyoD NH_2 and COOH termini cooperatively promote differentiation. The expression of MyoD/Myf5 chimeric MRFs that included the MyoD NH_2- or COOH terminal regions in *MyoD-/-;Myf5-/-* fibroblasts produced more efficient differentiation in low serum conditions than those with the corresponding Myf5 region. (A) Percentage of infected cells in each pool based on GFP expression immediately before differentiation. (B) Percentage of total nuclei (n > 1,000) found within a differentiated myosin heavy chain cell, normalized to A. (C) Myosin heavy chain immunostaining of differentiated pools (MF20, red; DAPI, blue). Bar, 100 μm.

COOH-terminal region of full-length MyoD or insertion of this region into the corresponding location of Myf5 does not significantly affect these results based on protein (myosin heavy chain immunostaining) or RNA (real-time PCR) markers (Punch, V., personal communication). The MyoD NH_2 terminus also had a noticeable effect on differentiation (Fig. 5 B, compare MyoD with 5dd, d5d with 55d, dd5 with 5d5, and d55 with Myf5). The MyoD bHLH domain, which has previously been connected to cell cycle arrest (Crescenzi et al., 1990; Sorrentino et al., 1990), enhanced differentiation when combined with the MyoD COOH-terminal region (Fig. 5 B, compare MyoD with d5d and 5dd with 55d) but had little effect otherwise. Overall, the combined NH_2- and COOH-terminal portions of MyoD (MyoD and d5d) were most effective in producing robust differentiation, whereas the two MRFs lacking both (Myf5 and 5d5) were the poorest. Thus, in both growth and differentiation conditions, the MyoD NH_2- and COOH-terminal regions cooperate to strongly activate the myogenic differentiation program.

Discussion

Previous work has suggested that there are unique roles for MyoD and Myf5 in adult myogenesis (Megeney et al., 1996; Sabourin et al., 1999; Montarras et al., 2000). To investigate the possibility that these phenotypes are a result of the activation of unique target genes by each of these transcription factors, we conducted genome-wide surveys of gene expression changes in response to MRF expression. The embryonic fibroblast cells that were used are normally non-myogenic but can be converted to myoblasts by ectopic

MRF expression. Importantly, they were derived in a *MyoD*$^{-/-}$;*Myf5*$^{-/-}$ background, precluding cross-activation between MyoD and Myf5.

Our microarray data indicates that MyoD expression induces more myogenic genes than Myf5, often to a greater degree. Comparison of our data with a previous study (Bergstrom et al., 2002), however, demonstrated that the majority of these MyoD-regulated targets are markers of differentiated skeletal muscle. Thus, a unique function of MyoD (vs. Myf5) is a strong ability to induce differentiation. *MyoD*-null primary myoblasts have vastly reduced levels of differentiation marker gene expression (Fig. 2), which is consistent with the phenotypic data on single-null primary myoblast cultures; in the absence of MyoD, the myoblasts proliferate well and differentiate poorly (Sabourin et al., 1999), whereas in the absence of Myf5, the opposite is true (Montarras et al., 2000). The moderate reduction of levels of growth-phase genes in *MyoD*-null myoblasts likely indicates a preference for activation by MyoD that Myf5 compensates for only partially.

A previous study by Seale et al. (2004) used representation difference analysis to identify potential satellite cell markers and MyoD target genes using primary myoblasts from *MyoD*$^{-/-}$ mice as a model for early satellite cell activation. The decreased expression of many genes (including *Chrnγ, Chrnα1, myogenin, troponin T1, H19*, and *Peg3*) was seen (Seale et al., 2004); these genes were up-regulated by MyoD expression in our experiments, supporting the contention that MyoD regulates these genes and promotes differentiation. However, Seale et al. (2004) also showed that levels of numerous genes (e.g., *k-cadherin, integrin-α7, Plgf, VCAM1, Igsf4a, Tcrαv13, laminin α5, neuritin-1*, and *Klra18*) are increased in *MyoD*$^{-/-}$ myoblasts, which express considerable Myf5. None of

these genes was identified in the current study as a Myf5 target. Thus, although MyoD and Myf5 have critical roles in the myogenic program, other transcription factors (e.g., Pax3 or Pax7) are likely needed to produce the full spectrum of normal myoblast gene expression.

We initially focused on the genes that were activated in the proliferating population by pruning our microarray results of differentiation genes. To assist in this, we leveraged the work of Bergstrom et al. (2002), who used an inducible MyoD in differentiating dblKO cells. Their identified target genes were those induced by serum deprivation, growth arrest, and MyoD activity in the context of those conditions. However, their data also included genes that are induced immediately before growth arrest but remain up-regulated either as a result of a failure of their RNA levels to decay to baseline during the early times of the experiment or because expression continues in differentiation. Thus, we identified a set of genes that were growth phase targets of MyoD and Myf5. Furthermore, real-time PCR validation confirmed that they were not unique targets but, rather, that they all could be induced by either of these MRFs.

These similarities in function are perhaps not surprising. Both *MyoD*- and *Myf5*-null mice possess grossly normal skeletal muscle, indicating that there is redundancy and compensation for the loss of either factor during development. Indeed, only with the concurrent loss of MRF4 activity does the defect in skeletal myogenesis become fully penetrant (Kassar-Duchossoy et al., 2004). Therefore, rather than MyoD and Myf5 each having a unique set of target genes that they are responsible for activating, it is reasonable that these two factors are capable of regulating the same downstream targets. Differences

in MyoD and Myf5 function in proliferating cells may instead vary in the strength of their effect on similar sets of genes. Whether MyoD and Myf5 function similarly during differentiation is a distinct question.

Previous work by Wyzykowski et al. (2002) suggested that *Id3* and *NP1* were both activated by MyoD in myoblast growth phase. Although our data does not show this, the differences between our methods and theirs are considerable and are more than adequate to explain the discrepancy. The cell types in which the experiments were conducted are a primary example: Wyzykowski et al. (2002) used 10t1/2 fibroblasts (genetically wild type), whereas our fibroblasts were derived from *MyoD/Myf5* compound-mutant animals (it is also likely that our cells have disrupted MRF4 function; Kassar-Duchossoy et al., 2004). It is quite conceivable that the presence of other myogenic factors could be capable of modulating MyoD activity.

An important conclusion to be drawn from these data is that both MyoD and Myf5 are active transcription factors in proliferating myoblasts (previously shown for MyoD by Wyzykowski et al., 2002). Rather than passively awaiting a differentiation signal, these two factors induce the expression of myoblast growth-phase genes. MyoD and Myf5, therefore, act not only to commit the cell to a myoblast identity but also to prepare the myoblast for skeletal muscle differentiation.

The expression of MyoD or Myf5 led to the down-regulation of only a very small number of genes. This could indicate that they act primarily as transcriptional activators in high serum conditions. However, our experimental system involved converting a

fibroblast cell type to the myogenic lineage and, thus, was incapable of detecting MRF-mediated suppression of any genes that were not initially expressed in the control cells.

Myogenin is not usually found in proliferating myoblasts and is induced at the onset of differentiation (Hollenberg et al., 1993; Weintraub, 1993; Bergstrom et al., 2002). The non-differentiation targets that were assayed (e.g., *Six1*, *Runx1*, *Mcpt8*, and *Spp1*) were induced to similar levels by either MyoD or Myf5, whereas *myogenin* was induced strongly by MyoD (10-fold) but only weakly by Myf5 (1.2-fold). Thus, although myogenin can activate similar targets to MyoD (Wyzykowski et al., 2002) and Myf5 (Table I), indirect regulation through myogenin is not dominant. The over-expression of myogenin alone was much less effective than MyoD at inducing differentiation marker transcripts, also demonstrating that MyoD has distinct functions that are not simply consequences of myogenin induction.

The high level of structural conservation between MyoD and Myf5, particularly in the bHLH domain, provides a strong rationale for our observation that there are few, if any, unique growth phase target genes of MyoD or Myf5 (this might be extended to myogenin - the conserved bHLH domain may allow the induction of growth targets, although myogenin would not usually be found in growing cells). It has been shown previously that just three differing amino acids between the MyoD and E12 bHLH regions encode specificity for the myogenic program (Davis and Weintraub, 1992). However, our results suggest that the three non-conservative and five conservative amino acid substitutions between MyoD and Myf5 bHLH domains are not likely to be involved

in overt target gene selection, although they could potentially provide sufficient variation for more subtle types of regulation.

Chimeric MRFs, in which the NH_2-, bHLH-, and COOH terminal domains of MyoD and Myf5 were interchanged, were used to explore which portions of MyoD might provide for enhanced myogenic differentiation. Strikingly, only the d5d chimera approached MyoD levels of differentiation gene expression, whereas the other chimeras, Myf5, and myogenin were many times less effective. The MyoD NH_2- and COOH-terminal regions had much less activity when separated, thus indicating a functional interaction between them.

In contrast to growth conditions (Fig. 4), differentiation conditions allowed the MyoD NH_2- and COOH-terminal regions to function independently (Fig. 5, B and C), with the concurrent presence of both producing the greatest differentiation. Therefore, each region is effective when growth signals are naturally reduced under low serum conditions. However, cooperation between these two MyoD regions was required to overcome the strong growth signals that were provided by high serum medium. Full-length MyoD, unlike Myf5, could, therefore, bias a cell toward differentiation by sensitizing it to a moderate reduction in growth signals.

Interestingly, the MyoD bHLH domain enhanced differentiation at the molecular and phenotypic levels (Figs. 4 and 5), but only when combined with the MyoD COOH terminus. Despite its exceptional similarity with the Myf5 bHLH, the MyoD bHLH domain functionally differs in its interaction with the MyoD COOH terminus. However,

this MyoD bHLH effect modulates gene expression levels rather than target gene selection (Fig. 4).

The NH_2-terminal activity is likely to be the MyoD transcriptional activation domain, which was previously mapped between amino acids 3–56 (Weintraub et al., 1991b). It is interesting to speculate that the COOH-terminal function is provided by the chromatin remodeling amphipathic α-helix (helix III) domain that was previously described by Bergstrom and Tapscott (2001). Whereas MyoD can induce muscle marker expression at repressed loci, their studies demonstrated that a similar motif in myogenin was ineffective at this task. This α-helix motif is conserved in Myf5 (Gerber et al., 1997), but Myf5 is similar to myogenin in being less effective than MyoD at initiating muscle gene expression. Thus, MyoD might be more efficient than Myf5 for inducing differentiation because of a greater ability to remodel chromatin at lineage-restricted loci. A similar domain appears in MRF4 (Rhodes and Konieczny, 1989; Bergstrom and Tapscott, 2001) and could mediate the residual skeletal myogenesis that is found in the absence of MyoD and Myf5 (Kassar-Duchossoy et al., 2004). Therefore, MyoD may use its COOH-terminal chromatin- remodeling domain to provide access to silenced muscle genes for the NH_2-terminal activation domain. In support of this, the substitution of either portion alone into Myf5 produced only moderate increases in gene expression, whereas together they strongly activated a variety of differentiation markers (Fig. 4).

If MyoD is inclined to promote differentiation, whereas Myf5 activates growth (but not differentiation) targets, then it is consistent with their cell cycle regulation in growing myoblasts. MyoD levels peak at the differentiation checkpoint in G_1 of the cell

cycle, whereas Myf5 levels are high in S/G_2 and G_0 in association with proliferation and a failure to differentiate (Kitzmann et al., 1998). This complementary pattern of expression allows for the maintenance of expression of growth-phase genes and myogenic identity while ensuring that high levels of MyoD occur only at a cell cycle point that is appropriate for differentiation.

That MyoD and Myf5 activate the same downstream target genes, but to differing degrees, does much to explain the phenotypes observed in single knockout mice and cells. Our data reinforce the partitioning of myogenic factors into primary and secondary MRFs and, furthermore, add support to the concept of a role for Myf5 in myoblast proliferation versus MyoD instigating myogenic differentiation. MyoD's ability to activate differentiation marker expression despite the presence of high levels of serum is a distinguishing biochemical function that is not found in Myf5 and involves the cooperation of separate domains of MyoD. These regions may allow MyoD to interact with co-activators for which Myf5 has much less affinity. In the future, studies to further understand the distinct roles played by MyoD and Myf5 in growing and differentiating myoblasts, as well as the structural constraints and intermolecular interactions upon which those roles are built, will be essential to our understanding of the mechanism by which damaged skeletal muscle is efficiently regenerated.

Materials and Methods

Cell culture

MyoD-null/*Myf5*-null dblKO mouse embryo fibroblasts were isolated, and clonal lines were selected and expanded (designated 2C5/7 and 4C5/2). Fibroblasts were cultured in sub-confluent conditions in growth medium of DMEM supplemented with 10% FCS and 1% penicillin/streptomycin. Proliferating primary myoblasts from wild-type and *MyoD*$^{-/-}$ mice were isolated and cultured as described previously (Megeney et al., 1996; Sabourin et al., 1999).

Myf5, *MyoD*, chimeric MRFs, or myogenin were introduced into dblKO cells using retrovirus based on the three-plasmid HIT system (provided by V. Sartorelli, National Institutes of Health, Bethesda, MD; Soneoka et al., 1995), including expression plasmids based on the pHAN backbone (with puromycin resistance driven from a distinct SV40 promoter) or a modified pHAN backbone in which the puromycin cassette was replaced by an internal ribosomal entry site and humanized Renilla GFP (IRES-hrGFP) sequence (Stratagene). MRFs that were expressed using the latter retroviral plasmid were COOH-terminally tagged with a 3x FLAG epitope. Empty control virus expressed only puromycin resistance or hrGFP. Retrovirus was produced by calcium phosphate transient co-transfection of *gag-pol* (pHIT60), *env* (pHIT456), and expression (pHAN) plasmids into 293T cells; virus-containing medium was harvested 48 h from the beginning of the transfection and was filtered through a 0.45-μm syringe filter (Millipore). Cells were infected overnight with diluted, filtered viral supernatant plus 8 μg/ml polybrene

(hexadimethrine bromide; Sigma-Aldrich). Drug selection, where appropriate, was conducted with 1 μg/ml puromycin (Sigma-Aldrich) in growth medium for at least 1 wk; uninfected controls were obliterated after 4–5 d of selection. Differentiation was induced by replacing growth medium with DME + 2% horse serum (GIBCO BRL) and 1% penicillin/streptomycin.

FACS

3 d after infection with hrGFP-expressing retrovirus, pools of cells were trypsinized, centrifuged, resuspended in PBS + 5% FCS, filtered through a MACS filter (Miltenyi Biotec) to remove aggregates, and placed on ice. Cell sorting was performed using a MoFlo sorter (DakoCytomation), with gating on hrGFP+ populations set by a comparison with an uninfected negative control pool. A small portion of the sorted cells was reanalyzed to assess purity, and the remainder was re-plated in growth medium for an additional 24 h before harvesting for total RNA.

RNA and protein isolation

Total RNA was purified using the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions and quantitated by OD260 or by RiboGreen (Invitrogen). RNA samples used for microarray analysis were ethanol precipitated in order to reach a minimum concentration of 2 μg/μl. Protein extracts were made by lysis of pelleted cells in radioimmunoprecipitation assay (RIPA) lysis buffer supplemented with MiniComplete protease inhibitors (Roche).

Microarray analysis

Triplicate RNA samples for microarray analysis were submitted to the Ottawa Genome Centre for hybridization to MG-U74Av2 GeneChips (Affymetrix, Inc.). Manufacturer's quality controls were verified by the Centre. Raw data files were processed with MicroArray Suite (MAS 5.0; Affymetrix, Inc.) using the statistical algorithm (Affymetrix, Inc.) to derive signals and present/marginal/absent calls for each sample; all possible pair-wise comparisons were also performed between experimental and control replicates, producing log(fold) ratio estimates of change and increase/no change/decrease calls. Processed results were then exported from MAS 5.0 and imported into Excel and Access (Microsoft) for further manipulation. Probesets showing consistent statistically significant changes between MRF and control samples were screened for detectable signal values (present/absent calls) and absolute log(fold) changes of at least one. Probeset annotations were obtained from http://www.affymetrix.com/analysis/index.affx. Microarray data is available from StemBase (http://www.scgp.ca:8080/ StemBase/; Ontario Genomics Innovation Centre; Perez-Iratxeta et al., 2005) under experiments E223 (samples S361-4) and E59 (S78-9) and from the National Center for Biotechnology Information's Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/; Edgar et al., 2002; Barrett et al., 2005) under series accession no. GSE3245 (GSM73053...64) and GSE3244 (GSM73026, -29, -32, -35, -38, and -41).

Real-time PCR

RNA samples were reverse transcribed using random hexamer primers that were included in the RNA PCR Core Kit (Perkin-Elmer) according to the manufacturer's instructions. Reverse transcription reactions were diluted (1:10 in Fig. 3 C and 2:5 in Fig. 4 A) with 10 mM Tris, pH 8.0, yielding master samples of reverse-transcribed products from which related PCR reactions were drawn. Real-time PCR reactions included the following: 2 μl of diluted reverse transcription product, 10 μl of 2x iQ SYBR Green Supermix (Bio-Rad Laboratories), 30 nM ROX passive reference dye (Stratagene), and 50 nM of each forward and reverse PCR primer. Real-time data was gathered using a system (MX4000; Stratagene) over 40 cycles (30 s at 94°C, 60 s at 58°C, and 30 s at 72°C) followed by a denaturation curve from 54 to 94°C in 30 s increments of 0.5°C to ensure amplification specificity. C_t values were calculated with the MX4000 software by using moving window averaging and an adaptive baseline. Fold changes, other calculations, and chart plotting were performed in Excel. A PCR efficiency of 85% was assumed. Primer sequences can be found in Table S2.

Chimeric MRF construction

Chimeric *MyoD/Myf5* MRFs were created by using an overlapping PCR approach to seamlessly fuse the NH_2- and COOH-terminal regions to the central bHLH domain (Fig. 3 A). The 318 aa of MyoD (available from GenBank/EMBL/DDBJ under accession no. NM_010866) were divided into aa 1–96 (NH_2 terminus), 97–161 (bHLH), and 162–318 (COOH terminus). The 255 aa of Myf5 (available from GenBank/EMBL/DDBJ under

accession no. NM_008656) were divided into aa 1–70 (NH_2 terminus), 71–135 (bHLH), and 136–255 (COOH terminus). Each PCR product was cloned into an appropriate expression plasmid and verified by sequencing (Applied Biosystems).

Immunocytochemistry and Western blot analysis

Primary antibodies that were used are listed as follows: mouse anti-FLAG (M2 and M5; Sigma-Aldrich), rabbit anti-Myf5 (C-20; Santa Cruz Biotechnology, Inc.), rabbit anti-MyoD (C-20; Santa Cruz Biotechnology, Inc.), mouse anti-MyoD (5.2F; Sigma-Aldrich), mouse anti-myogenin (F5D) and mouse anti-myosin heavy chain (MF20; hybridoma supernatants), mouse anti-desmin (D33; DakoCytomation), and mouse anti-α-tubulin (Sigma-Aldrich). For immunostaining, cells were fixed with 4% PFA, permeabilized with Triton X-100, and blocked with 5% normal goat serum in PBS. Primary and secondary antibodies were applied in 5% goat serum–PBS. Secondary detection used appropriate fluorescein- or rhodamine-conjugated antibodies (Chemicon). 0.25 μg/ml DAPI was included in a final wash step to highlight nuclei. Samples were mounted in fluorescence mounting medium (DakoCytomation), coverslipped, and imaged using a microscope (Axiophot 2; Carl Zeiss MicroImaging, Inc.), a 10X NA 0.30 plan-Neofluar (Ph1; ∞/0.17) or 20X NA 0.75 plan Apochromat (∞/0.17) objective, and a digital camera (AxioCam; Carl Zeiss MicroImaging, Inc.). Digital images were captured by using Axiovision (Carl Zeiss MicroImaging, Inc.) and were processed with Photoshop (Adobe). Enumeration was assisted by ImageJ software (http://rsb.info.nih.gov/ij/).

Western blots were made by electroblotting standard SDS-PAGE gels onto Immobilon-P membranes. Membranes were blocked in 5% skimmed milk in PBS; secondary detection was performed with an appropriate HRP-conjugated antibody (Bio-Rad Laboratories) visualized by ECL (GE Healthcare). For densitometry, blots were digitally imaged with a 16-bit GeneGnome chemiluminescence gel-doc system (Syngene), and bands were quantified with the bundled GeneTools software (Syngene) except parts of Fig. 3 B, for which the MyoD (C-20), Myf5 (C-20), and corresponding tubulin (not depicted) blots were exposed to film, scanned, and analyzed by using ImageJ software.

Online supplemental material

Tables S1–S4 show the expansion of Table I, where candidate gene lists were derived by testing for consistent pair-wise increase/decrease changes (at least six of nine) as well as minimum mean threshold log(fold) changes of ±1 (two-fold). Increases for MyoD, increases for Myf5, decreases for MyoD, and decreases for Myf5 correspond to Tables S1, S2, S3, and S4, respectively. Table S5 shows the sequences and targets for primers that were used for SYBR Green real-time PCR. Online supplemental material is available at http://www.jcb.org/cgi/content/full/ jcb.200502101/DC1.

Acknowledgements

We thank Vince Punch for work on MyoD(cdk4-) and Myf5(cdk4+) constructs, Dr. V. Sartorelli for providing retroviral expression plasmids, and Mark Gillespie and Mike Huh for careful reading of the manuscript. M.A. Rudnicki holds the Canada Research Chair (CRC) in Molecular Genetics and is a Howard Hughes Medical Institute (HHMI) International Scholar. This work was supported by grants to M.A. Rudnicki from the Muscular Dystrophy Association, the National Institutes of Health, the Canadian Institutes of Health Research, the HHMI, and the CRC Program.

References

Barrett, T., T.O. Suzek, D.B. Troup, S.E. Wilhite, W.C. Ngau, P. Ledoux, D. Rudnev, A.E. Lash, W. Fujibuchi, and R. Edgar. 2005. NCBI GEO: mining millions of expression profiles–database and tools. *Nucleic Acids Res.* 33:D562–D566.

Benezra, R., R.L. Davis, D. Lockshon, D.L. Turner, and H. Weintraub. 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell.* 61:49–59.

Bergstrom, D.A., and S.J. Tapscott. 2001. Molecular distinction between specification and differentiation in the myogenic basic helix-loop-helix transcription factor family. *Mol. Cell. Biol.* 21:2404–2412.

Bergstrom, D.A., B.H. Penn, A. Strand, R.L. Perry, M.A. Rudnicki, and S.J. Tapscott. 2002. Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. *Mol. Cell.* 9:587–600.

Braun, T., G. Buschhausen-Denker, E. Bober, E. Tannich, and H.H. Arnold. 1989. A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. *EMBO J.* 8:701–709.

Braun, T., M.A. Rudnicki, H.H. Arnold, and R. Jaenisch. 1992. Targeted inactivation of the muscle regulatory gene Myf-5 results in abnormal rib development and perinatal death. *Cell.* 71:369–382.

Crescenzi, M., T.P. Fleming, A.B. Lassar, H. Weintraub, and S.A. Aaronson. 1990. MyoD induces growth arrest independent of differentiation in normal and transformed cells. *Proc. Natl. Acad. Sci. USA.* 87:8442–8446.

Davis, R.L., and H. Weintraub. 1992. Acquisition of myogenic specificity by replacement of three amino acid residues from MyoD into E12. *Science.* 256:1027–1030.

Edgar, R., M. Domrachev, and A.E. Lash. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30:207–210.

Gerber, A.N., T.R. Klesert, D.A. Bergstrom, and S.J. Tapscott. 1997. Two domains of MyoD mediate transcriptional activation of genes in repressive chromatin: a mechanism for lineage determination in myogenesis. *Genes Dev.* 11:436–450.

Hollenberg, S.M., P.F. Cheng, and H. Weintraub. 1993. Use of a conditional MyoD transcription factor in studies of MyoD trans-activation and muscle determination. *Proc. Natl. Acad. Sci. USA.* 90:8028–8032.

Kassar-Duchossoy, L., B. Gayraud-Morel, D. Gomes, D. Rocancourt, M. Buckingham, V. Shinin, and S. Tajbakhsh. 2004. Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature.* 431:466–471.

Kitzmann, M., G. Carnac, M. Vandromme, M. Primig, N.J. Lamb, and A. Fernandez. 1998. The muscle regulatory factors MyoD and myf-5 undergo distinct cell cycle-specific expression in muscle cells. *J. Cell Biol.* 142:1447–1459.

Li, L., J. Zhou, G. James, R. Heller-Harrison, M.P. Czech, and E.N. Olson. 1992. FGF inactivates myogenic helix-loop-helix proteins through phosphorylation of a conserved protein kinase C site in their DNA-binding domains. *Cell.* 71:1181–1194.

Lindon, C., D. Montarras, and C. Pinset. 1998. Cell cycle-regulated expression of the muscle determination factor Myf5 in proliferating myoblasts. *J. Cell Biol.* 140:111–118.

Megeney, L.A., B. Kablar, K. Garrett, J.E. Anderson, and M.A. Rudnicki. 1996. MyoD is required for myogenic stem cell function in adult skeletal muscle. *Genes Dev.* 10:1173–1183.

Montarras, D., C. Lindon, C. Pinset, and P. Domeyne. 2000. Cultured myf5 null and myoD null muscle precursor cells display distinct growth defects. *Biol. Cell.* 92:565–572.

Perez-Iratxeta, C., G. Palidwor, C.J. Porter, N.A. Sanche, M.R. Huska, B.P. Suomela, E.M. Muro, P.M. Krzyzanowski, E. Hughes, P.A. Campbell, et al. 2005. Study of stem cell function using microarray experiments. *FEBS Lett.* 579:1795–1801.

Rawls, A., M.R. Valdez, W. Zhang, J. Richardson, W.H. Klein, and E.N. Olson. 1998. Overlapping functions of the myogenic bHLH genes MRF4 and MyoD revealed in double mutant mice. *Development.* 125:2349–2358.

Rhodes, S.J., and S.F. Konieczny. 1989. Identification of MRF4: a new member of the muscle regulatory factor gene family. *Genes Dev.* 3:2050–2061.

Rudnicki, M.A., T. Braun, S. Hinuma, and R. Jaenisch. 1992. Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell.* 71:383–390.

Rudnicki, M.A., P.N. Schnegelsberg, R.H. Stead, T. Braun, H.H. Arnold, and R. Jaenisch. 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell.* 75:1351–1359.

Sabourin, L.A., A. Girgis-Gabardo, P. Seale, A. Asakura, and M.A. Rudnicki. 1999. Reduced differentiation potential of primary MyoD-/- myogenic cells derived from adult skeletal muscle. *J. Cell Biol.* 144:631–643.

Seale, P., A. Asakura, and M.A. Rudnicki. 2001. The potential of muscle stem cells. *Dev. Cell.* 1:333–342.

Seale, P., J. Ishibashi, C. Holterman, and M.A. Rudnicki. 2004. Muscle satellite cell-specific genes identified by genetic profiling of MyoD-deficient myogenic cell. *Dev. Biol.* 275:287–300.

Soneoka, Y., P.M. Cannon, E.E. Ramsdale, J.C. Griffiths, G. Romano, S.M. Kingsman, and A.J. Kingsman. 1995. A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res.* 23:628–633.

Sorrentino, V., R. Pepperkok, R.L. Davis, W. Ansorge, and L. Philipson. 1990. Cell proliferation inhibited by MyoD1 independently of myogenic differentiation. *Nature.* 345:813–815.

Tapscott, S.J., R.L. Davis, M.J. Thayer, P.F. Cheng, H. Weintraub, and A.B. Lassar. 1988. MyoD1: a nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. *Science.* 242:405–411.

Thayer, M.J., S.J. Tapscott, R.L. Davis, W.E. Wright, A.B. Lassar, and H. Weintraub. 1989. Positive autoregulation of the myogenic determination gene MyoD1. *Cell.* 58:241–248.

Vaidya, T.B., S.J. Rhodes, E.J. Taparowsky, and S.F. Konieczny. 1989. Fibroblast growth factor and transforming growth factor beta repress transcription of the myogenic regulatory gene MyoD1. *Mol. Cell. Biol.* 9:3576–3579.

Vivian, J.L., E.N. Olson, and W.H. Klein. 2000. Thoracic skeletal defects in myogenin- and MRF4-deficient mice correlate with early defects in myo-

Weintraub, H. 1993. The MyoD family and myogenesis: redundancy, networks, and thresholds. *Cell.* 75:1241–1244.

Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T.K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, et al. 1991a. The myoD gene family: nodal point during specification of the muscle cell lineage. *Science.* 251:761–766.

Weintraub, H., V.J. Dwarki, I. Verma, R. Davis, S. Hollenberg, L. Snider, A. Lassar, and S.J. Tapscott. 1991b. Muscle-specific transcriptional activation by MyoD. *Genes Dev.* 5:1377–1386.

Wyzykowski, J.C., T.I. Winata, N. Mitin, E.J. Taparowsky, and S.F. Konieczny. 2002. Identification of novel MyoD gene targets in proliferating myogenic stem cells. *Mol. Cell. Biol.* 22:6199–6208.

Zhang, J.M., X. Zhao, Q. Wei, and B.M. Paterson. 1999. Direct inhibition of G(1) cdk kinase activity by MyoD promotes myoblast cell cycle withdrawal and terminal differentiation. *EMBO J.* 18:6983–6993.

Table S1 - Increased by MyoD

log2(FoldChange)

		MyoD			Myf5			hrGFP			MyoD					Myf5					MyoD	Myf5
Title	GenbankID	P	M	A	P	M	A	P	M	A	I	MI	NC	MD	D	I	MI	NC	MD	D	SLR	SLR
actin, alpha, cardiac	M15501	3	0	0	0	0	3	0	0	3	9	0	0	0	0	1	0	8	0	0	7.9	1.6
troponin C, cardiac/slow skeletal	M29793	3	0	0	1	0	2	0	0	3	9	0	0	0	0	9	0	0	0	0	7.4	3.2
insulin-like growth factor binding protein 5	L12447	3	0	0	3	0	0	0	0	3	9	0	0	0	0	9	0	0	0	0	7.2	3.3
actin, alpha 1, skeletal muscle	M12347	3	0	0	3	0	0	0	1	2	9	0	0	0	0	9	0	0	0	0	6.5	3.2
troponin T3, skeletal, fast	L48989	3	0	0	2	0	1	0	0	3	9	0	0	0	0	0	0	9	0	0	5.6	1.5
myosin light chain, phosphorylatable, fast skeletal muscle	AV290649	3	0	0	0	0	3	0	0	3	9	0	0	0	0	0	0	9	0	0	5.4	1.6
H19 fetal liver mRNA	X58196	3	0	0	3	0	0	1	1	1	9	0	0	0	0	9	0	0	0	0	5.2	2.6
myosin, light polypeptide 4	M19436	3	0	0	3	0	0	0	0	3	9	0	0	0	0	3	0	6	0	0	4.8	1.1
cholinergic receptor, nicotinic, beta polypeptide 1 (muscle)	M14537	3	0	0	3	0	0	0	0	3	9	0	0	0	0	9	0	0	0	0	4.6	4.0
troponin T2, cardiac	L47600	3	0	0	3	0	0	1	0	2	9	0	0	0	0	8	0	1	0	0	4.4	1.4
myosin, heavy polypeptide 3, skeletal muscle, embryonic	M74753	3	0	0	0	0	3	0	0	3	9	0	0	0	0	0	0	9	0	0	4.1	-0.4
troponin T1, skeletal, slow	AV213431	3	0	0	0	0	3	0	0	3	9	0	0	0	0	1	1	7	0	0	4.0	-0.2
mast cell protease 8	X78545	3	0	0	3	0	0	0	0	3	9	0	0	0	0	9	0	0	0	0	3.5	3.5
ATPase, Ca++ transporting, cardiac muscle, fast twitch 1	X67140	3	0	0	0	0	3	0	0	3	9	0	0	0	0	0	0	9	0	0	3.5	0.2
troponin I, skeletal, slow 1	AJ242874	3	0	0	0	0	3	0	0	3	9	0	0	0	0	0	0	9	0	0	3.4	-0.2
myogenin	X15784	3	0	0	1	0	2	0	0	3	9	0	0	0	0	0	0	9	0	0	3.3	0.3
ATPase, Ca++ transporting, cardiac muscle, fast twitch 1	AV241808	3	0	0	0	0	3	0	0	3	8	0	1	0	0	0	0	9	0	0	3.3	-1.3
myosin, light polypeptide 1	X12973	3	0	0	2	0	1	0	0	3	9	0	0	0	0	0	0	9	0	0	3.2	-0.4
cholinergic receptor, nicotinic, gamma polypeptide	X03818	3	0	0	0	0	3	0	0	3	9	0	0	0	0	0	0	9	0	0	3.2	1.0
cholinergic receptor, nicotinic, gamma polypeptide	AV248455	3	0	0	0	0	3	0	0	3	6	2	1	0	0	0	0	9	0	0	3.1	-0.2
cholinergic receptor, nicotinic, alpha polypeptide 1 (muscle)	M17640	3	0	0	0	0	3	0	0	3	8	1	0	0	0	0	0	9	0	0	3.1	1.4
secreted phosphoprotein 1	X13986	3	0	0	3	0	0	3	0	0	9	0	0	0	0	9	0	0	0	0	3.1	2.5
reticulon 2 (Z-band associated protein)	AF093624	3	0	0	0	0	3	0	0	3	6	0	3	0	0	0	0	9	0	0	2.8	-0.5
C1q and tumor necrosis factor related protein 3	AI315647	3	0	0	3	0	0	3	0	0	9	0	0	0	0	5	0	4	0	0	2.5	0.9
retinoblastoma 1	M26391	3	0	0	2	0	1	0	0	3	9	0	0	0	0	0	0	9	0	0	2.5	0.7
actinin alpha 3	AF093775	2	1	0	0	0	3	0	0	3	7	0	2	0	0	0	0	9	0	0	2.3	-0.8
RIKEN cDNA 1190002N15 gene	AW125453	3	0	0	3	0	0	1	0	2	9	0	0	0	0	9	0	0	0	0	2.3	0.8
myosin light chain, phosphorylatable, fast skeletal muscle	U77943	3	0	0	0	0	3	0	0	3	8	0	1	0	0	1	0	8	0	0	2.3	0.2
potassium channel tetramerisation domain containing 12	AI842065	3	0	0	3	0	0	1	0	2	9	0	0	0	0	9	0	0	0	0	2.3	2.6
v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	AI006228	3	0	0	0	0	3	0	0	3	8	0	1	0	0	0	0	9	0	0	2.2	0.3
ryanodine receptor 1, skeletal muscle	D38216	2	1	0	0	0	3	0	0	3	9	0	0	0	0	0	0	9	0	0	2.2	0.3
protein kinase inhibitor, alpha	AW125442	3	0	0	1	0	2	0	0	3	9	0	0	0	0	4	0	5	0	0	2.2	0.7
UDP-glucose ceramide glucosyltransferase	AI853172	3	0	0	3	0	0	3	0	0	9	0	0	0	0	9	0	0	0	0	2.1	1.7
cadherin 15	AJ245402	3	0	0	0	0	3	0	0	3	6	0	3	0	0	0	0	9	0	0	2.1	-1.0
myosin binding protein H	U68267	3	0	0	0	0	3	0	0	3	9	0	0	0	0	0	0	8	0	1	2.0	-1.1
sine oculis-related homeobox 1 homolog (Drosophila)	X80339	3	0	0	3	0	0	1	0	2	9	0	0	0	0	9	0	0	0	0	1.9	1.8
tropomyosin 2, beta	M81086	3	0	0	3	0	0	3	0	0	6	0	3	0	0	0	0	9	0	0	1.9	0.0
inhibitor of DNA binding 2	AF077861	3	0	0	3	0	0	3	0	0	9	0	0	0	0	9	0	0	0	0	1.9	1.2
paternally expressed 3	AF038939	3	0	0	3	0	0	3	0	0	9	0	0	0	0	0	1	8	0	0	1.8	0.0
11 days embryo head cDNA, RIKEN full-length enriched library, clon	AW212672	2	1	0	0	2	1	0	0	3	6	0	3	0	0	1	0	8	0	0	1.8	1.3
enolase 3, beta muscle	X61600	3	0	0	3	0	0	3	0	0	9	0	0	0	0	8	1	0	0	0	1.8	0.8
sarcoglycan, beta (dystrophin-associated glycoprotein)	AB024921	3	0	0	0	1	2	0	0	3	9	0	0	0	0	0	0	9	0	0	1.8	-0.1
runt related transcription factor 1	D26532	3	0	0	3	0	0	2	0	1	9	0	0	0	0	9	0	0	0	0	1.8	1.6
	X05862	3	0	0	0	0	3	0	0	3	9	0	0	0	0	0	0	8	0	1	1.7	-0.8
	AW120874	3	0	0	3	0	0	3	0	0	9	0	0	0	0	0	1	8	0	0	1.7	0.3
myocyte enhancer factor 2A	AW045443	3	0	0	3	0	0	2	0	1	9	0	0	0	0	5	0	4	0	0	1.6	0.3

Title	GenbankID	MyoD			Myf5			hrGFP			MyoD					Myf5					MyoD SLR	Myf5 SLR
		P	M	A	P	M	A	P	M	A	I	MI	NC	MD	D	I	MI	NC	MD	D		
osteoglycin	D31951	3	0	0	3	0	0	2	0	1	8	0	1	0	0	0	0	9	0	0	1.6	-0.1
alpha-methylacyl-CoA racemase	U89906	3	0	0	3	0	0	1	0	2	9	0	0	0	0	3	0	6	0	0	1.6	1.0
serum amyloid A 3	X03505	3	0	0	3	0	0	1	0	2	7	0	2	0	0	2	0	7	0	0	1.6	0.1
ankyrin repeat domain 1 (cardiac muscle)	AF041847	3	0	0	3	0	0	3	0	0	8	0	1	0	0	9	0	0	0	0	1.6	1.6
sarcoglycan, gamma (dystrophin-associated glycoprotein)	AB024922	3	0	0	0	0	3	0	0	3	6	0	3	0	0	0	0	9	0	0	1.5	-0.5
hairy and enhancer of split 6 (Drosophila)	AW048812	3	0	0	3	0	0	0	0	3	9	0	0	0	0	0	0	9	0	0	1.4	-0.1
transmembrane 4 superfamily member 6	AF053454	3	0	0	3	0	0	1	0	2	9	0	0	0	0	6	0	3	0	0	1.4	1.2
growth arrest and DNA-damage-inducible 45 alpha	U00937	3	0	0	1	0	2	1	0	2	8	1	0	0	0	0	0	9	0	0	1.3	-0.2
protein kinase inhibitor, alpha	M63554	3	0	0	3	0	0	1	0	2	8	0	1	0	0	0	0	9	0	0	1.3	0.3
periostin, osteoblast specific factor	D13664	3	0	0	3	0	0	3	0	0	7	0	2	0	0	6	0	3	0	0	1.3	0.7
lung carcinoma myc related oncogene 1	X13945	3	0	0	1	1	1	0	0	3	6	0	3	0	0	0	0	9	0	0	1.2	0.6
	X65128	3	0	0	3	0	0	3	0	0	7	0	2	0	0	3	0	3	0	3	1.2	-0.1
discoidin domain receptor family, member 1	L57509	3	0	0	3	0	0	0	0	3	9	0	0	0	0	7	2	0	0	0	1.2	0.7
actin-binding LIM protein 1	AI841606	1	2	0	1	0	2	0	0	3	7	0	2	0	0	1	0	8	0	0	1.2	0.3
SH3-domain GRB2-like B1 (endophilin)	AI842874	3	0	0	3	0	0	3	0	0	9	0	0	0	0	0	0	9	0	0	1.1	0.5
UDP-glucose ceramide glucosyltransferase	D89866	3	0	0	3	0	0	2	0	1	6	1	2	0	0	0	0	9	0	0	1.1	0.5
cyclin-dependent kinase inhibitor 1A (P21)	AW048937	3	0	0	3	0	0	3	0	0	9	0	0	0	0	1	0	8	0	0	1.1	0.1
WNT1 inducible signaling pathway protein 1	AF100777	3	0	0	3	0	0	3	0	0	8	0	1	0	0	8	1	0	0	0	1.1	1.2
	AV109962	2	1	0	3	0	0	0	0	3	6	0	3	0	0	8	0	1	0	0	1.1	0.7
activating transcription factor 6	AA592069	3	0	0	3	0	0	2	0	1	8	0	1	0	0	5	1	3	0	0	1.1	0.9
ADP-ribosylation factor-like 6 interacting protein 5	AW049647	3	0	0	3	0	0	3	0	0	9	0	0	0	0	9	0	0	0	0	1.1	0.6
cadherin 11	D21253	3	0	0	3	0	0	2	0	1	5	1	3	0	0	3	0	6	0	0	1.1	0.4
tumor protein p53 inducible nuclear protein 2	AW125508	3	0	0	3	0	0	1	2	0	8	0	1	0	0	0	0	9	0	0	1.0	-0.1
choline phosphotransferase 1	AW212131	3	0	0	3	0	0	1	0	2	7	0	2	0	0	2	1	6	0	0	1.0	0.5

Call:
- P Present
- M Marginal
- A Absent

Change:
- I Increased
- MI Marginally Increased
- NC No Change
- MD Marginally Decreased
- D Decreased

Ratio:
- SLR Signal Log Ratio, ie. logarithm(base2) of the fold-change, averaged

Table S2 - Increased by Myf5

log2(FoldChange)

Title	GenbankID	MyoD			Myf5			hrGFP			MyoD					Myf5					MyoD SLR	Myf5 SLR
		P	M	A	P	M	A	P	M	A	I	MI	NC	MD	D	I	MI	NC	MD	D		
cholinergic receptor, nicotinic, beta polypeptide 1 (muscle)	M14537	3	0	0	3	0	0	0	0	3	9	0	0	0	0	9	0	0	0	0	4.6	4.0
mast cell protease 8	X78545	3	0	0	3	0	0	0	0	3	9	0	0	0	0	9	0	0	0	0	3.5	3.5
insulin-like growth factor binding protein 5	L12447	3	0	0	3	0	0	0	0	3	9	0	0	0	0	9	0	0	0	0	7.2	3.3
selectin, platelet	M72332	1	0	2	3	0	0	0	0	3	3	0	6	0	0	8	0	1	0	0	2.3	3.3
actin, alpha 1, skeletal muscle	M12347	3	0	0	3	0	0	0	1	2	9	0	0	0	0	9	0	0	0	0	6.5	3.2
H19 fetal liver mRNA	X58196	3	0	0	3	0	0	1	1	1	9	0	0	0	0	9	0	0	0	0	5.2	2.6
potassium channel tetramerisation domain containing 12	AI842065	3	0	0	3	0	0	1	0	2	9	0	0	0	0	9	0	0	0	0	2.3	2.6
secreted phosphoprotein 1	X13986	3	0	0	3	0	0	3	0	0	9	0	0	0	0	9	0	0	0	0	3.1	2.5
SKI interacting protein	AW046671	2	0	1	3	0	0	1	0	2	0	0	9	0	0	9	0	0	0	0	0.3	2.3
sine oculis-related homeobox 1 homolog (Drosophila)	X80339	3	0	0	3	0	0	1	0	2	9	0	0	0	0	9	0	0	0	0	1.9	1.8
UDP-glucose ceramide glucosyltransferase	AI853172	3	0	0	3	0	0	3	0	0	9	0	0	0	0	9	0	0	0	0	2.1	1.7
serine (or cysteine) proteinase inhibitor, clade E, member 1	M33960	3	0	0	3	0	0	3	0	0	4	0	2	0	3	6	0	3	0	0	0.3	1.7
connective tissue growth factor	M70642	3	0	0	3	0	0	3	0	0	5	0	1	0	3	6	0	0	0	3	0.6	1.7
ankyrin repeat domain 1 (cardiac muscle)	AF041847	3	0	0	3	0	0	3	0	0	8	0	1	0	0	9	0	0	0	0	1.6	1.6
runt related transcription factor 1	D26532	3	0	0	3	0	0	2	0	1	9	0	0	0	0	9	0	0	0	0	1.8	1.6
sema domain, immunoglobulin domain (Ig), short basic domain, secr	Z80941	2	0	1	3	0	0	0	0	3	8	0	1	0	0	8	0	1	0	0	1.7	1.5
guanine nucleotide binding protein, alpha inhibiting 1	AI153412	3	0	0	3	0	0	1	0	2	5	0	4	0	0	7	0	2	0	0	1.3	1.5
troponin T2, cardiac	L47600	3	0	0	3	0	0	1	0	2	9	0	0	0	0	8	0	1	0	0	4.4	1.4
calmodulin binding protein 1	AF062378	1	1	1	3	0	0	1	0	2	3	0	4	0	2	6	0	3	0	0	0.6	1.3
enabled homolog (Drosophila)	D10727	3	0	0	3	0	0	2	0	1	6	0	3	0	0	9	0	0	0	0	0.9	1.3
inhibitor of DNA binding 2	AF077861	3	0	0	3	0	0	3	0	0	9	0	0	0	0	9	0	0	0	0	1.9	1.2
WNT1 inducible signaling pathway protein 1	AF100777	3	0	0	3	0	0	3	0	0	8	0	1	0	0	8	1	0	0	0	1.1	1.2
transmembrane 4 superfamily member 6	AF053454	3	0	0	3	0	0	1	0	2	9	0	0	0	0	6	0	3	0	0	1.4	1.2
tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation	U56243	0	0	3	1	2	0	0	0	3	2	0	7	0	0	6	0	3	0	0	0.1	1.2
RIKEN cDNA 2610201A13 gene	AA222883	3	0	0	3	0	0	3	0	0	5	0	4	0	0	9	0	0	0	0	0.5	1.1
RIKEN cDNA 1300002F13 gene	AI853531	3	0	0	3	0	0	3	0	0	5	0	2	1	1	6	0	3	0	0	0.3	1.1
	AA796831	3	0	0	3	0	0	3	0	0	2	0	6	0	1	9	0	0	0	0	0.2	1.1
phospholipase A2, group VII (platelet-activating factor acetylhydrolas	U34277	1	2	0	3	0	0	2	0	1	0	0	9	0	0	6	0	3	0	0	0.7	1.0
A kinase (PRKA) anchor protein 8	AI646981	3	0	0	3	0	0	2	0	1	4	0	5	0	0	6	0	3	0	0	0.6	1.0
RIKEN cDNA 2810026P18 gene	AW045664	3	0	0	3	0	0	1	0	2	6	0	0	0	3	6	0	3	0	0	0.6	1.0
RIKEN cDNA 5530600A18 gene	AW125433	2	0	1	3	0	0	1	0	2	5	0	4	0	0	6	0	3	0	0	0.7	1.0
	AA177826	0	0	3	3	0	0	1	0	2	2	1	6	0	0	6	0	3	0	0	0.3	1.0

Call:

P	Present
M	Marginal
A	Absent

Change:

I	Increased
MI	Marginally Increased
NC	No Change
MD	Marginally Decreased
D	Decreased

Ratio:

SLR	Signal Log Ratio, ie. logarithm(base2) of the fold-change, averaged

Table S3 - Decreased by MyoD

log2(FoldChange)

Title	GenbankID	MyoD			Myf5			hrGFP			MyoD					Myf5					MyoD	Myf5
		P	M	A	P	M	A	P	M	A	I	MI	NC	MD	D	I	MI	NC	MD	D	SLR	SLR
hypermethylated in cancer 1	AJ007511	1	0	2	3	0	0	3	0	0	0	0	2	0	7	0	0	8	0	1	-2.4	-0.5
delta-like 1 homolog (Drosophila)	Z12171	3	0	0	3	0	0	3	0	0	0	0	0	0	9	0	0	3	0	6	-1.6	-0.7
quiescin Q6	AW123556	3	0	0	3	0	0	3	0	0	0	0	2	0	7	0	0	6	0	3	-1.1	-0.8
	AA002843	2	1	0	3	0	0	3	0	0	0	0	2	1	6	1	0	6	1	1	-1.1	-0.1
endothelial differentiation, sphingolipid G-protein-coupled receptor, 3	AF108021	3	0	0	3	0	0	3	0	0	0	0	1	0	8	0	0	0	0	9	-1.1	-0.8

Call:

P	Present
M	Marginal
A	Absent

Change:

I	Increased
MI	Marginally Increased
NC	No Change
MD	Marginally Decreased
D	Decreased

Ratio:

SLR	Signal Log Ratio, ie. logarithm(base2) of the fold-change, averaged

Table S4 - Decreased by Myf5

log2(FoldChange)

Title	GenbankID	MyoD			Myf5			hrGFP			MyoD					Myf5					MyoD	Myf5
		P	M	A	P	M	A	P	M	A	I	MI	NC	MD	D	I	MI	NC	MD	D	SLR	SLR
transcription factor 20	AI847906	3	0	0	2	0	1	3	0	0	1	0	4	2	2	0	0	0	0	9	-0.4	-2.3
Solute carrier family 39 (zinc transporter), member 1 (Slc39a1), mRN	AA606878	3	0	0	3	0	0	3	0	0	2	0	2	0	5	0	0	3	0	6	-0.6	-1.5
actin, beta, cytoplasmic	J04181	3	0	0	3	0	0	3	0	0	3	0	2	0	4	0	0	2	1	6	-0.5	-1.5
KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention re	AW123408	3	0	0	3	0	0	3	0	0	2	0	3	0	4	0	0	2	0	7	-0.6	-1.5
transforming growth factor, beta receptor III	AF039601	3	0	0	3	0	0	3	0	0	0	0	5	1	3	0	0	0	0	9	-0.8	-1.4
OTU domain, ubiquitin aldehyde binding 1	AI838318	3	0	0	3	0	0	2	1	0	1	0	3	0	5	0	0	3	0	6	-1.0	-1.4
UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1	M27923	3	0	0	3	0	0	3	0	0	1	0	2	0	6	0	0	3	0	6	-1.0	-1.3
RIKEN cDNA 2300006M17 gene	AV379809	3	0	0	3	0	0	3	0	0	0	0	4	0	5	0	0	3	0	6	-0.9	-1.2
lamin A	D49733	3	0	0	3	0	0	3	0	0	2	0	2	0	5	3	0	0	0	6	-1.2	-1.2
synaptopodin	AW046661	3	0	0	3	0	0	3	0	0	0	0	5	0	4	0	0	3	0	6	-0.8	-1.1
procollagen, type I, alpha 2	AV230631	3	0	0	1	2	0	3	0	0	0	0	6	0	3	0	0	3	0	6	-0.6	-1.0
solute carrier family 23 (nucleobase transporters), member 3	AV222871	3	0	0	3	0	0	3	0	0	0	0	5	1	3	0	0	3	0	6	-0.6	-1.0
dynein, cytoplasmic, intermediate chain 1	AF063229	2	0	1	1	0	2	2	1	0	0	0	5	1	3	0	0	2	0	7	-0.5	-1.0
secretin	X73580	3	0	0	3	0	0	3	0	0	1	0	3	0	5	0	0	2	0	7	-0.7	-1.0

Call:

P	Present
M	Marginal
A	Absent

Change:

I	Increased
MI	Marginally Increased
NC	No Change
MD	Marginally Decreased
D	Decreased

Ratio:

SLR	Signal Log Ratio, ie. logarithm(base2) of the fold-change, averaged

Table S5 - Sequences of PCR primers used for SYBR green real-time PCR

Gene	GID	RLT-PCR Ta	amplicon	forward sequence	forward length	reverse sequence	reverse length	Spans Introns?
Chrna1	NM_007389	58.0 °C	325 bp	5' ccatgaagtcagaccaggag 3'	20 mer	5' atacagccgtgtgagcagag 3'	20 mer	8(395bp)
Chrnb1	NM_009601	58.0 °C	315 bp	5' tcaggagctacgagaggtca 3'	20 mer	5' cacttgaacgtcttcccaga 3'	20 mer	10(289bp)
Chrng	NM_009604	58.0 °C	227 bp	5' ctccagaatggctcttcctc 3'	20 mer	5' aggttacaggcatccacaca 3'	20 mer	8(395bp)
GAPDH	NM_008084	58.0 °C	228 bp	5' tcggtgtgaacggatttg 3'	18 mer	5' ggtctcgctcctggaaga 3'	18 mer	single exon gene
Hes6	NM_019479	58.0 °C	363 bp	5' acccacctcgctgagtagtt 3'	20 mer	5' atgcatgcactggatgtagc 3'	20 mer	1(78bp);2(91bp)
Hmgb3	NM_008253	58.0 °C	258 bp	5' ctaggtggccaagaaggaag 3'	20 mer	5' ctcccgatcatatcggactt 3'	20 mer	3(1863bp);4(226bp)
HnrpL	NM_177301	58.0 °C	307 bp	5' tccagcaccctagcaatgta 3'	20 mer	5' gagtcttccttcccagcatc 3'	20 mer	10(470bp);11(119bp); 12(361bp)
Idb3	NM_008321	58.0 °C	299 bp	5' cctgcagcgtgtcatagact 3'	20 mer	5' aagtgaagagggctgggtta 3'	20 mer	1(116bp);2(504bp)
IGFBP5	NM_010518	58.0 °C	313 bp	5' gcacctgagatgagacagga 3'	20 mer	5' gtaggatagggggaggaagg 3'	20 mer	1(589bp);2(635bp)
L-myc	NM_008506	58.0 °C	248 bp	5' ggcactcctagtctggaagc 3'	20 mer	5' ttatgctgctgttggtggat 3'	20 mer	1(2887bp)
m-cadherin	NM_007662	58.0 °C	317 bp	5' agccctgagttcttcagcat 3'	20 mer	5' ccttcaaggatggtgaacct 3'	20 mer	5(459bp);6(393bp)
Mcpt8	NM_008572	58.0 °C	279 bp	5' ccaaacctacaacgactcca 3'	20 mer	5' aggctcagacaaccataccc 3'	20 mer	4(480bp)
myogenin	NM_031189	58.0 °C	94 bp	5' gcaatgcactggagttcg 3'	18 mer	5' acgatggacgtaagggagtg 3'	20 mer	2 (527)
Osf2	NM_015784	58.0 °C	238 bp	5' ggggttgtcactgtgaactg 3'	20 mer	5' tctcgaaagcttcattggtg 3'	20 mer	unknown
Pkia	NM_008862	58.0 °C	287 bp	5' tgaagtccctgctatgtgga 3'	20 mer	5' gttgatgaccgtcagaatgg 3'	20 mer	2(4496bp)
Runx1	NM_009821	58.0 °C	285 bp	5' tctgcagaactttccagtcg 3'	20 mer	5' ggcgccgtagtatagatggt 3'	20 mer	4(6712bp)
Selp	NM_011347	58.0 °C	273 bp	5' aggacggtcactggtcagat 3'	20 mer	5' cgaatgagacattggcagac 3'	20 mer	11(816bp)
Six1	NM_009189	58.0 °C	290 bp	5' gaaccggaggcaaagaga 3'	18 mer	5' agtcctggagctggtgct 3'	18 mer	1(2038bp)
Spp1	NM_009263	58.0 °C	322 bp	5' gtggtgccaagagtgtgttt 3'	20 mer	5' catggctttcattggaattg 3'	20 mer	1(99bp);2(1079bp); 3(491bp);4(998bp)
Tgfbr3	NM_011578	58.0 °C	281 bp	5' gtggtcaaaaacctcgtcct 3'	20 mer	5' ttgttctcaagccgaagatg 3'	20 mer	unknown
Tncc	NM_009393	58.0 °C	261 bp	5' cacaggtgagaccattacgg 3'	20 mer	5' gacaggaatggggagagaaa 3'	20 mer	5(90bp)

Chapter 5

General Discussion

Overview of Findings

Skeletal muscle satellite cells provide a local reservoir of lineage-committed adult stem cells for repairing exercise- or trauma-induced damage to muscle fibres. The paired-box transcription factor Pax7 is a specific molecular marker of satellite cells, corresponding to the histological criteria of Mauro (1961). Adult *Pax7*$^{-/-}$ mice are severely depleted of satellite cells and have very little capacity to regenerate muscle. The function of Pax7 in myoblasts is therefore of primary interest in deciphering how satellite cell populations are maintained and committed to the skeletal muscle lineage.

Quiescent satellite cells are activated by damage to muscle fibres. An early event in their activation is the expression of the myogenic regulatory factors Myf5 and MyoD, each of which can function as a master regulator of myogenesis. As transcription factors, Myf5 and MyoD induce numerous other muscle genes, eventually culminating in the differentiation and fusion of the myoblasts into contractile multinucleated fibres. Both proteins are expressed in growing myoblasts and each provides unique properties to the growing cells. The relationship of Pax7 to Myf5 and MyoD in the transition of quiescent satellite cells into activated myoblasts, and the differential functions of Myf5 and MyoD in the progression from proliferative myoblasts into differentiated myofibres, are key questions in a molecular model of adult myogenesis.

This work has addressed these questions. The residual myoblasts that persist in *Pax7*$^{-/-}$ muscle are rescued from apoptosis by the loss of *p53*, indicating that Pax7 has a cell survival function (Chapter 2). Importantly, Pax7 also acts in lineage determination

as a transcription factor, increasing the expression of numerous genes when introduced into either myoblasts or fibroblasts (Chapter 3). These target genes are normally expressed in satellite cell-derived myoblasts, but many are poorly characterized and may have roles in maintaining an adult stem cell phenotype. Notably, different isoforms of Pax7 have varying abilities to activate these targets, providing a rationale for the evolutionary conservation of its alternative splicing. In particular, Pax7 has the capacity to regulate the expression of Myf5 and MyoD, providing a functional connection between the paired-box and myogenic factors. In the context of non-myogenic adult $CD45^{+}/Sca1^{+}$ stem cells (Appendix B), expression of Pax7 activates Myf5, MyoD, and ultimately the full myogenic program. As Pax7 is highly expressed in growing myoblasts, the up-regulation of Myf5 by Pax7 fits with the association of Myf5 to proliferation and MyoD to potentiating differentiation (Chapter 4; Appendix C). These data therefore integrate three key myogenic transcription factors into a framework for future studies of satellite and other adult myogenic stem cells (Figure 1).

Biomedical Implications

Translocation fusions of Pax7 and Pax3 to FKHR have a prominent role in alveolar rhabdomyosarcoma. Identifying the normal functions of Pax7 and how they are deregulated in these tumours could point to therapeutic vulnerabilities in these aggressive childhood tumours. On a related issue, knowledge of how Pax7 functions in cell survival and proliferation, as well as in maintaining an undifferentiated and self-renewing satellite cell pool, could revolutionize myoblast-based treatments for muscle degenerative

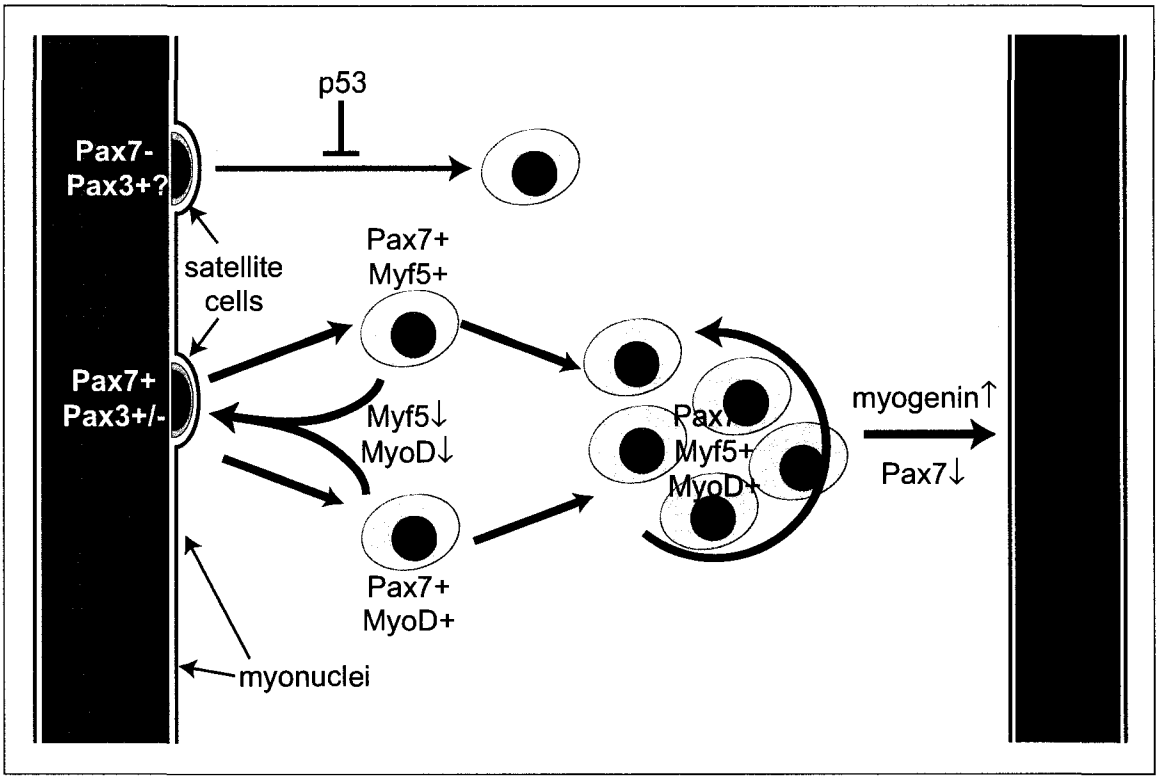


Figure 1. Model of Pax and MRF transcription factor functions in adult satellite cells. Pax7 is expressed in quiescent satellite cell that reside upon muscle fibres until activated by damage. In the absence of Pax7, residual Pax3+ satellite cells are lost through a p53-dependent pathway (Chapter 2). Following activation, satellite cells upregulate *Myf5*, *MyoD*, and other myoblast genes regulated by Pax7 (Chapter 3). A fraction of activated satellite cells down-regulate MRF expression and return to quiescence, replenishing the satellite cell pool. The remaining myoblasts proliferate to provide an expanded pool of cells for fibre repair. High levels of Myf5 expression are associated with this growth phase, whereas increased levels of MyoD prepare the myoblasts for differentiation (Chapter 4). At the onset of differentiation, MyoD efficiently induces myogenin and other genes involved in myoblast fusion and muscle fibre function (Chapter 4). Expression of Pax7 in non-myogenic $CD45^{+}/Sca1^{+}$ stem cells induces a similar cascade of myogenic specification (Appendix B).

diseases such as the muscular dystrophies, which have been hindered by poor survival and engraftment of donor cells.

Adult stem cells are a potential alternative to purified myoblasts. The ability to convert the $CD45^{+}Sca1^{+}$ population by expressing Pax7 is promising. Understanding the mechanism by which Pax7 accomplishes this could allow non-muscle stem cell populations to be employed in future interventions. Similarly, manipulating the balance of Myf5 and MyoD in implanted cells will be important for engineering a proliferative population that differentiates efficiently into working muscle fibres. Gene therapies for other conditions involving secreted products could also benefit if implanted myoblasts capable of manufacturing an important factor could fuse into fibres.

Pax genes are involved in the development of other organs, including the eye (Pax6), kidney (Pax2), pancreas (Pax4), and teeth and skeleton (Pax9). Neurogenesis is also affected in many Pax mutants. Understanding Pax7 function may provide general lessons as to the roles in organogenesis of this important family of genes. Similarly, bHLH transcription factors are widely utilized during development. Myf5 and MyoD are therefore useful models for the functions of master regulatory genes in lineage determination.

Pax3/7 Factors in Myogenesis

The high degree of sequence conservation and nearly identical genomic organizations of Pax7 and Pax3 indicate that Pax7 arose in a duplication event from an ancestral Pax3/7 gene (Relaix et al., 2004). The invertebrate Amphioxus possesses a

single Pax3/7 protein that shares 90%, 100%, and 97% amino acid identity with human PAX7 through the paired, octapeptide, and homeodomain segments, respectively (Holland et al., 1999), suggesting that the ancestral gene had ancient origins. Indeed, Pax7 has been proposed to have the structure most closely resembling the prototypical Pax factor that arose in primitive animals long before the radiation of the Metazoa (Vorobyov and Horst, 2006). Yet despite millions of years of evolutionary separation, Pax3 can functionally substitute for the many of the functions of its Drosophila homologue Paired (Prd; Xue and Noll, 1996). This high degree conservation suggests that their roles are deeply embedded within the development of the animal body plan. Nonetheless, their usage has diverged significantly. For instance, loss of Prd in Drosophila produces a severe pair-rule patterning defect early in the development (Nusslein-Volhard and Wieschaus, 1980). This function is not shared with other arthropods, for which a segmentation role is more general (Davis et al., 2005). Conversely, mutation of Pax7 and Pax3 affected myogenesis and neurogenesis but not overall segmental patterning in mice (Mansouri and Gruss, 1998; Relaix et al., 2005). Furthermore, although they share a degree of overlap in expression and function, Pax7 and Pax3 have differing mouse knockout phenotypes that indicate that each has acquired distinct functions since the original duplication that separated them.

Pax7 clearly has an essential role in satellite cells based on evidence from mouse (Seale et al., 2000) and frog models (Chen et al., 2006). Kuang et al. (2006) demonstrated that Pax3-expressing myoblasts are interstitial and that only Pax7-expressing cells reside in classical satellite cell locations beneath the fibre basal lamina.

Paralogous Pax7 and Pax3 genes are found in mammals, birds, amphibians, and fish, whereas single Pax3/7 genes have been identified in other vertebrates (Chen et al., 2006; Relaix et al., 2004). Interestingly, satellite cells have been found in mammals, birds, and amphibians (rev. Charge and Rudnicki, 2004), and possibly in fish (Stoiber and Sanger, 1996). It is an intriguing (and speculative) hypothesis that the production of a distinct Pax7 by duplication of the ancestral Pax3/7 gene was critical to the appearance of proper satellite cells. In contrast to other regenerative strategies, it may be advantageous to have a reservoir of resident, committed stem cells. Pax7 provides for survival, proliferation, and self-renewal of satellite cells, permitting this population to be maintained through adulthood over many cycles of regeneration. Although Pax3-expressing myoblasts are present, they are inadequate in post-natal and regenerative roles (Kuang et al., 2006). Examination of satellite cells and Pax7 function in many more species will help to define the evolutionary importance of Pax7.

Functions of Pax7

Pax7 mRNA and protein are strongly expressed during myoblast proliferation, before being drastically down-regulated during the earliest stages of differentiation. A role for Pax7 in promoting myoblast survival would be consistent with these observations, as activated satellite cells must enter and multiply within the harsh environment of damaged and necrotic muscle tissue. It was shown in Chapter 4 that whereas Myf5 was functional in growth phase, MyoD was far more efficient in activating the differentiation program. Therefore, by increasing Myf5 expression and limiting

MyoD expression, Pax7-expressing myoblasts would continue to increase in numbers prior to fusion. In this model, myoblasts lacking Pax7 activity would be depleted by apoptosis, resulting in a deficit in proliferation and earlier differentiation that would reduce the number of myoblasts available for repair. This would produce a loss of regenerative capacity, as is observed in *Pax7*$^{-/-}$ muscle.

The strength of Pax7 as a transcription factor is not clear. While there are a number of genes that are well-induced by Pax7d in C2C12, so far only a few respond with high-level changes. Alternatively, Pax7 may have a more prominent function in defining lineage-related sets of genes for transcriptional availability than in direct recruitment of RNA polymerase complexes. Marking those genes by DNA binding or with chromatin modifications could be a necessary preparatory step for their later usage by specific master control factors. For example, such a role has been shown for the homeobox gene Pbx1 in conjunction with MyoD (Berkes et al., 2004). Rather than cooperating with the MRFs, however, Pax7 may help to provide for the ability to activate them. Although the earliest stages of embryonic myogenesis are accomplished without Pax7 or Pax3, they are required for fetal, post-natal, and adult stages (Relaix et al., 2005). Pax7 and Pax3 may therefore be permissive for the myogenic lineage by providing the opportunity for MRF expression in later periods.

Pax7 and Pax3 are prominent during neurogenesis and provide dorsal identity to the neural plate (Relaix et al., 2004). Myogenesis and neurogenesis have many common developmental features, amongst them a prominent role for bHLH factors in lineage determination through the MRFs for muscle, and neurogenic factors such as Neurogenin,

Mash, or NeuroD for neurons. Conceivably, Pax7 could regulate neurogenic bHLH factors (much as for Myf5 and MyoD) with the combinatorial influence of other factors specifying neurons versus muscle. Pax7 will also need to be examined in neuronal systems, where it may regulate a distinct set of targets. Alternatively, the targets that we have identified may provide an underlying function that is useful in the context of both tissues. These comparisons will enlighten the process by which lineages are chosen.

Alternative Splicing of Pax7 and Pax3

The alternative splicing in the paired DNA binding domains of Pax7 and Pax3 is a conserved but strictly vertebrate innovation. Amphioxus Pax3/7 lacks the intron and splicing that produce the $Q^{+/-}$ change (Holland et al., 1999). It was therefore inserted after the divergence of vertebrates but prior to the duplication of ancestral Pax3/7, because the $Q^{+/-}$ option is found in both Pax7 and Pax3. The conservation of this splicing suggests a functional biological importance, perhaps providing an important variation in Pax function without requiring a wholesale gene duplication event. Interestingly, the intron necessary for the $GL^{+/-}$ change is conserved from Amphioxus to Pax3 and to Pax7; however, only Pax7 appears to be alternatively spliced (Holland et al., 1999). Nonetheless, the GL^{+} insertion is the basis for the specific functional change that we have shown to modify Pax7 regulation of Myf5 and other target genes (Chapter 3). The addition of these variations in the paired domains of Pax7 and Pax3 has therefore coincided with the evolutionary development of vertebrate skeletal muscle.

The implications of Pax7 alternative splicing provide many lines of possible inquiry. Unknowns include: how the isoforms are expressed; tissue-to-tissue variation in expression; dynamic regulation through development and ageing; changes in expression during satellite cell activation and regeneration; and differences between muscle groups, fibre types, or satellite cell sub-populations. Limited study has been made of these descriptive issues (Kay and Ziman, 1999; Ziman et al., 1997; Ziman and Kay, 1998), perhaps in part due to difficulties in quantitatively analyzing the transcripts in an efficient manner. Appendix A discusses this problem and provides a real-time PCR-based strategy to quickly assay numerous samples for comparison. This system will be useful for studies describing the *in vivo* expression of alternative transcripts as well as how they are regulated.

Pax7d induced the endogenous Pax7 gene during the conversion of the adult stem cell $CD45^{+}/Sca1^{+}$ population into myoblasts. Therefore, a combination of Pax7 isoforms may provide for myogenesis. However, existing engineered mouse lines have been generated by replacing a genomic fragment (often an exon) with a single contiguous coding sequence for Pax7, Pax3, or Pax3/FKHR, ensuring that the knock-in allele is expressed as only a single isoform. As such, data regarding preferential activation of isoform-specific targets, and even potential interactions between the different isoforms, are lost, as is also the case in expression experiments conducted in much the same way using cultured cell lines. Current conclusions regarding the specific roles of Pax7 or Pax3 may thus be incomplete. Similarly, there is only limited literature on alternative splicing of the Pax/FKHR fusion transcript (Du et al., 2005), and it appears probable that

studying a single fusion isoform will be inadequate to fully explain its role in alveolar rhabdomyosarcoma.

Primary Myogenic Regulatory Factors

In contrast to Pax7, the expression of Myf5 or MyoD is restricted to skeletal muscle, and induction of either primary myogenic factor commits that cell to the muscle lineage. They are, therefore, master regulators of cell fate. Whereas Pax7 may act to restrict lineage choices, Myf5 and MyoD specify the myogenic lineage and stimulate expression of the genes that are specific for muscle. They thus integrate numerous developmental signals into a decision to become muscle, from which the myogenic phenotype is then elaborated.

The duplication of myogenic bHLH factors into a family of four has allowed for specialization of function for each. Myf5 has taken a position as an "early" myogenic marker associated with myoblast proliferation. Most satellite cells have expressed Myf5 but down-regulate it during quiescence (S. Kuang, submitted). This majority proliferates quickly during regeneration, and likely provides the bulk of the acute repair capacity after damage. Myf5 is much less effective at inducing differentiation than MyoD (Chapter 4), and MyoD-negative myoblasts have gene expression profiles intermediate to quiescent satellite cells and fully activated myoblasts (Appendix C). *In vitro*, *MyoD*$^{-/-}$ myoblasts express decreased levels of differentiation markers. However, increased expression of Myf5 in the *MyoD*$^{-/-}$ mouse (Sabourin et al., 1999) may partially compensate for this deficit *in vivo*. When challenged with severe or prolonged damage, the deficit is more

clearly exposed (Megeney et al., 1996). This model would predict that the *Myf5*$^{-/-}$ mouse would also exhibit a regeneration defect, with inadequate proliferation and premature differentiation of *Myf5*$^{-/-}$ satellite cells leading to poor and insufficient repair of severe or repeated injury, and depletion of the satellite cell population.

Activation of myogenin, and the consequent expression of the myogenic differentiation program, is probably a key role for MyoD that is less effectively executed by Myf5 or MRF4. Therefore, Myf5 and MyoD provide myogenesis with a degree of redundancy and robustness, but also with the opportunity to ensure a balance between capacity (proliferation) and capability (differentiation).

Conclusion

Pax7, MyoD, and Myf5 occupy critical positions in the molecular circuitry of adult myogenesis. Their actions specify an adult stem cell population to become myogenic precursors and ultimately skeletal muscle. Expression of these transcription factors thereby provides myogenic identity to skeletal muscle satellite cells.

References

References

Asakura, A., G.E. Lyons, and S.J. Tapscott. 1995. The regulation of MyoD gene expression: conserved elements mediate expression in embryonic axial muscle. *Dev Biol.* 171:386-98.

Asakura, A., P. Seale, A. Girgis-Gabardo, and M.A. Rudnicki. 2002. Myogenic specification of side population cells in skeletal muscle. *J Cell Biol.* 159:123-34.

Barrett, T., T.O. Suzek, D.B. Troup, S.E. Wilhite, W.C. Ngau, P. Ledoux, D. Rudnev, A.E. Lash, W. Fujibuchi, and R. Edgar. 2005. NCBI GEO: mining millions of expression profiles--database and tools. *Nucleic Acids Res.* 33:D562-6.

Beauchamp, J.R., L. Heslop, D.S. Yu, S. Tajbakhsh, R.G. Kelly, A. Wernig, M.E. Buckingham, T.A. Partridge, and P.S. Zammit. 2000. Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *J Cell Biol.* 151:1221-34.

Benezra, R., R.L. Davis, D. Lockshon, D.L. Turner, and H. Weintraub. 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell.* 61:49-59.

Bergstrom, D.A., B.H. Penn, A. Strand, R.L. Perry, M.A. Rudnicki, and S.J. Tapscott. 2002. Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. *Mol Cell.* 9:587-600.

Bergstrom, D.A., and S.J. Tapscott. 2001. Molecular distinction between specification and differentiation in the myogenic basic helix-loop-helix transcription factor family. *Mol Cell Biol.* 21:2404-12.

Berkes, C.A., D.A. Bergstrom, B.H. Penn, K.J. Seaver, P.S. Knoepfler, and S.J. Tapscott. 2004. Pbx marks genes for activation by MyoD indicating a role for a homeodomain protein in establishing myogenic potential. *Mol Cell.* 14:465-77.

Berkes, C.A., and S.J. Tapscott. 2005. MyoD and the transcriptional control of myogenesis. *Semin Cell Dev Biol.* 16:585-95.

Bischoff, R. 1994. The satellite cell and muscle regeneration. *In* Myology. A.G. Engel, Franzini-Armstrong, C., editor. McGraw-Hill, New York. 97-118.

Bober, E., T. Franz, H.H. Arnold, P. Gruss, and P. Tremblay. 1994. Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development.* 120:603-12.

Borycki, A.G., J. Li, F. Jin, C.P. Emerson, and J.A. Epstein. 1999. Pax3 functions in cell survival and in pax7 regulation. *Development.* 126:1665-74.

Braun, T., G. Buschhausen-Denker, E. Bober, E. Tannich, and H.H. Arnold. 1989. A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. *Embo J.* 8:701-9.

Braun, T., M.A. Rudnicki, H.H. Arnold, and R. Jaenisch. 1992. Targeted inactivation of the muscle regulatory gene Myf-5 results in abnormal rib development and perinatal death. *Cell.* 71:369-82.

Braun, T., B. Winter, E. Bober, and H.H. Arnold. 1990. Transcriptional activation domain of the muscle-specific gene-regulatory protein myf5. *Nature.* 346:663-5.

Bruyninx, M., B. Hennuy, A. Cornet, P. Houssa, M. Daukandt, E. Reiter, J. Poncin, J. Closset, and G. Hennen. 1999. A novel gene overexpressed in the prostate of

castrated rats: hormonal regulation, relationship to apoptosis and to acquired prostatic cell androgen independence. *Endocrinology*. 140:4789-99.

Buckingham, M. 2001. Skeletal muscle formation in vertebrates. *Curr Opin Genet Dev*. 11:440-8.

Cao, Y., Z. Zhao, J. Gruszczynska-Biegala, and A. Zolkiewska. 2003. Role of metalloprotease disintegrin ADAM12 in determination of quiescent reserve cells during myogenic differentiation in vitro. *Mol Cell Biol*. 23:6725-38.

Chalepakis, G., and P. Gruss. 1995. Identification of DNA recognition sequences for the Pax3 paired domain. *Gene*. 162:267-70.

Chalepakis, G., J. Wijnholds, and P. Gruss. 1994. Pax-3-DNA interaction: flexibility in the DNA binding and induction of DNA conformational changes by paired domains. *Nucleic Acids Res*. 22:3131-7.

Charge, S.B., and M.A. Rudnicki. 2004. Cellular and molecular regulation of muscle regeneration. *Physiol Rev*. 84:209-38.

Chen, Y., G. Lin, and J.M. Slack. 2006. Control of muscle regeneration in the Xenopus tadpole tail by Pax7. *Development*. 133:2303-13.

Christ, B., and C.P. Ordahl. 1995. Early stages of chick somite development. *Anat Embryol (Berl)*. 191:381-96.

Conboy, I.M., M.J. Conboy, G.M. Smythe, and T.A. Rando. 2003. Notch-mediated restoration of regenerative potential to aged muscle. *Science*. 302:1575-7.

Conboy, I.M., and T.A. Rando. 2002. The regulation of Notch signaling controls satellite cell activation and cell fate determination in postnatal myogenesis. *Dev Cell.* 3:397-409.

Cornelison, D.D., M.S. Filla, H.M. Stanley, A.C. Rapraeger, and B.B. Olwin. 2001. Syndecan-3 and syndecan-4 specifically mark skeletal muscle satellite cells and are implicated in satellite cell maintenance and muscle regeneration. *Dev Biol.* 239:79-94.

Cornelison, D.D., S.A. Wilcox-Adelman, P.F. Goetinck, H. Rauvala, A.C. Rapraeger, and B.B. Olwin. 2004. Essential and separable roles for Syndecan-3 and Syndecan-4 in skeletal muscle development and regeneration. *Genes Dev.* 18:2231-6.

Cornelison, D.D., and B.J. Wold. 1997. Single-cell analysis of regulatory gene expression in quiescent and activated mouse skeletal muscle satellite cells. *Dev Biol.* 191:270-83.

Cornet, A.M., E. Hanon, E.R. Reiter, M. Bruyninx, V.H. Nguyen, B.R. Hennuy, G.P. Hennen, and J.L. Closset. 2003. Prostatic androgen repressed message-1 (PARM-1) may play a role in prostatic cell immortalisation. *Prostate.* 56:220-30.

Crescenzi, M., T.P. Fleming, A.B. Lassar, H. Weintraub, and S.A. Aaronson. 1990. MyoD induces growth arrest independent of differentiation in normal and transformed cells. *Proc Natl Acad Sci U S A.* 87:8442-6.

Davis, G.K., J.A. D'Alessio, and N.H. Patel. 2005. Pax3/7 genes reveal conservation and divergence in the arthropod segmentation hierarchy. *Dev Biol.* 285:169-84.

Davis, R.L., and H. Weintraub. 1992. Acquisition of myogenic specificity by replacement of three amino acid residues from MyoD into E12. *Science.* 256:1027-30.

De Angelis, L., L. Berghella, M. Coletta, L. Lattanzi, M. Zanchi, M.G. Cusella-De Angelis, C. Ponzetto, and G. Cossu. 1999. Skeletal myogenic progenitors originating from embryonic dorsal aorta coexpress endothelial and myogenic markers and contribute to postnatal muscle growth and regeneration. *J Cell Biol.* 147:869-78.

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

Decary, S., V. Mouly, and G.S. Butler-Browne. 1996. Telomere length as a tool to monitor satellite cell amplification for cell-mediated gene therapy. *Hum Gene Ther.* 7:1347-50.

Dhawan, J., and T.A. Rando. 2005. Stem cells in postnatal myogenesis: molecular mechanisms of satellite cell quiescence, activation and replenishment. *Trends Cell Biol.* 15:666-73.

Dilworth, F.J., K.J. Seaver, A.L. Fishburn, S.L. Htet, and S.J. Tapscott. 2004. In vitro transcription system delineates the distinct roles of the coactivators pCAF and p300 during MyoD/E47-dependent transactivation. *Proc Natl Acad Sci U S A.* 101:11593-8.

Du, S., E.J. Lawrence, D. Strzelecki, P. Rajput, S.J. Xia, D.M. Gottesman, and F.G. Barr. 2005. Co-expression of alternatively spliced forms of PAX3, PAX7, PAX3-FKHR and PAX7-FKHR with distinct DNA binding and transactivation properties in rhabdomyosarcoma. *Int J Cancer.* 115:85-92.

Edgar, R., M. Domrachev, and A.E. Lash. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30:207-10.

Epstein, J.A., P. Lam, L. Jepeal, R.L. Maas, and D.N. Shapiro. 1995. Pax3 inhibits myogenic differentiation of cultured myoblast cells. *J Biol Chem.* 270:11719-22.

Epstein, J.A., D.N. Shapiro, J. Cheng, P.Y. Lam, and R.L. Maas. 1996. Pax3 modulates expression of the c-Met receptor during limb muscle development. *Proc Natl Acad Sci U S A.* 93:4213-8.

Fredericks, W.J., N. Galili, S. Mukhopadhyay, G. Rovera, J. Bennicelli, F.G. Barr, and F.J. Rauscher, 3rd. 1995. The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3. *Mol Cell Biol.* 15:1522-35.

Garry, D.J., A. Meeson, J. Elterman, Y. Zhao, P. Yang, R. Bassel-Duby, and R.S. Williams. 2000. Myogenic stem cell function is impaired in mice lacking the forkhead/winged helix protein MNF. *Proc Natl Acad Sci U S A.* 97:5416-21.

Gerber, A.N., T.R. Klesert, D.A. Bergstrom, and S.J. Tapscott. 1997. Two domains of MyoD mediate transcriptional activation of genes in repressive chromatin: a mechanism for lineage determination in myogenesis. *Genes Dev.* 11:436-50.

Goldhamer, D.J., B.P. Brunk, A. Faerman, A. King, M. Shani, and C.P. Emerson, Jr. 1995. Embryonic activation of the myoD gene is regulated by a highly conserved distal control element. *Development*. 121:637-49.

Goulding, M.D., G. Chalepakis, U. Deutsch, J.R. Erselius, and P. Gruss. 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *Embo J.* 10:1135-47.

Gros, J., M. Manceau, V. Thome, and C. Marcelle. 2005. A common somitic origin for embryonic muscle progenitors and satellite cells. *Nature*. 435:954-8.

Grounds, M.D. 1998. Age-associated changes in the response of skeletal muscle cells to exercise and regeneration. *Ann N Y Acad Sci*. 854:78-91.

Hayashi, S., and M.P. Scott. 1990. What determines the specificity of action of Drosophila homeodomain proteins? *Cell*. 63:883-94.

Holland, L.Z., M. Schubert, Z. Kozmik, and N.D. Holland. 1999. AmphiPax3/7, an amphioxus paired box gene: insights into chordate myogenesis, neurogenesis, and the possible evolutionary precursor of definitive vertebrate neural crest. *Evol Dev.* 1:153-65.

Hollenberg, S.M., P.F. Cheng, and H. Weintraub. 1993. Use of a conditional MyoD transcription factor in studies of MyoD trans-activation and muscle determination. *Proc Natl Acad Sci U S A*. 90:8028-32.

Hollway, G., and P. Currie. 2005. Vertebrate myotome development. *Birth Defects Res C Embryo Today*. 75:172-9.

Holterman, C.E., and M.A. Rudnicki. 2005. Molecular regulation of satellite cell function. *Semin Cell Dev Biol.* 16:575-84.

Huang, S.M., A.H. Schonthal, and M.R. Stallcup. 2001. Enhancement of p53-dependent gene activation by the transcriptional coactivator Zac1. *Oncogene.* 20:2134-43.

Huh, M.S., M.H. Parker, A. Scime, R. Parks, and M.A. Rudnicki. 2004. Rb is required for progression through myogenic differentiation but not maintenance of terminal differentiation. *J Cell Biol.* 166:865-76.

Irintchev, A., M. Zeschnigk, A. Starzinski-Powitz, and A. Wernig. 1994. Expression pattern of M-cadherin in normal, denervated, and regenerating mouse muscles. *Dev Dyn.* 199:326-37.

Ishibashi, J., R.L. Perry, A. Asakura, and M.A. Rudnicki. 2005. MyoD induces myogenic differentiation through cooperation of its NH2- and COOH-terminal regions. *J Cell Biol.* 171:471-82.

Jostes, B., C. Walther, and P. Gruss. 1990. The murine paired box gene, Pax7, is expressed specifically during the development of the nervous and muscular system. *Mech Dev.* 33:27-37.

Kablar, B., A. Asakura, K. Krastel, C. Ying, L.L. May, D.J. Goldhamer, and M.A. Rudnicki. 1998. MyoD and Myf-5 define the specification of musculature of distinct embryonic origin. *Biochem Cell Biol.* 76:1079-91.

Kablar, B., K. Krastel, C. Ying, A. Asakura, S.J. Tapscott, and M.A. Rudnicki. 1997. MyoD and Myf-5 differentially regulate the development of limb versus trunk skeletal muscle. *Development.* 124:4729-38.

Kablar, B., K. Krastel, C. Ying, S.J. Tapscott, D.J. Goldhamer, and M.A. Rudnicki. 1999. Myogenic determination occurs independently in somites and limb buds. *Dev Biol.* 206:219-31.

Kaneko-Ishino, T., Y. Kuroiwa, N. Miyoshi, T. Kohda, R. Suzuki, M. Yokoyama, S. Viville, S.C. Barton, F. Ishino, and M.A. Surani. 1995. Peg1/Mest imprinted gene on chromosome 6 identified by cDNA subtraction hybridization. *Nat Genet.* 11:52-9.

Kanwar, Y.S., A. Kumar, K. Ota, S. Lin, J. Wada, S. Chugh, and E.I. Wallner. 2002. Identification of developmentally regulated mesodermal-specific transcript in mouse embryonic metanephros. *Am J Physiol Renal Physiol.* 282:F953-65.

Kassar-Duchossoy, L., B. Gayraud-Morel, D. Gomes, D. Rocancourt, M. Buckingham, V. Shinin, and S. Tajbakhsh. 2004. Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. *Nature.* 431:466-71.

Kassar-Duchossoy, L., E. Giacone, B. Gayraud-Morel, A. Jory, D. Gomes, and S. Tajbakhsh. 2005. Pax3/Pax7 mark a novel population of primitive myogenic cells during development. *Genes Dev.* 19:1426-31.

Kay, P.H., and M.R. Ziman. 1999. Alternate Pax7 paired box transcripts which include a trinucleotide or a hexanucleotide are generated by use of alternate 3' intronic splice sites which are not utilized in the ancestral homologue. *Gene.* 230:55-60.

King, T., Y. Bland, S. Webb, S. Barton, and N.A. Brown. 2002. Expression of Peg1 (Mest) in the developing mouse heart: involvement in trabeculation. *Dev Dyn.* 225:212-5.

Kitzmann, M., G. Carnac, M. Vandromme, M. Primig, N.J. Lamb, and A. Fernandez. 1998. The muscle regulatory factors MyoD and myf-5 undergo distinct cell cycle-specific expression in muscle cells. *J Cell Biol.* 142:1447-59.

Kitzmann, M., and A. Fernandez. 2001. Crosstalk between cell cycle regulators and the myogenic factor MyoD in skeletal myoblasts. *Cell Mol Life Sci.* 58:571-9.

Kuang, S., S.B. Charge, P. Seale, M. Huh, and M.A. Rudnicki. 2006. Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. *J Cell Biol.* 172:103-13.

LaBarge, M.A., and H.M. Blau. 2002. Biological progression from adult bone marrow to mononucleate muscle stem cell to multinucleate muscle fiber in response to injury. *Cell.* 111:589-601.

Li, L., J. Zhou, G. James, R. Heller-Harrison, M.P. Czech, and E.N. Olson. 1992. FGF inactivates myogenic helix-loop-helix proteins through phosphorylation of a conserved protein kinase C site in their DNA-binding domains. *Cell.* 71:1181-94.

Lindon, C., D. Montarras, and C. Pinset. 1998. Cell cycle-regulated expression of the muscle determination factor Myf5 in proliferating myoblasts. *J Cell Biol.* 140:111-8.

Mak, K.L., R.Q. To, Y. Kong, and S.F. Konieczny. 1992. The MRF4 activation domain is required to induce muscle-specific gene expression. *Mol Cell Biol.* 12:4334-46.

Mansouri, A., and P. Gruss. 1998. Pax3 and Pax7 are expressed in commissural neurons and restrict ventral neuronal identity in the spinal cord. *Mech Dev.* 78:171-8.

Mansouri, A., A. Stoykova, M. Torres, and P. Gruss. 1996. Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice. *Development.* 122:831-8.

Maroto, M., R. Reshef, A.E. Munsterberg, S. Koester, M. Goulding, and A.B. Lassar. 1997. Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. *Cell.* 89:139-48.

Mauro, A. 1961. Satellite cell of skeletal muscle fibers. *J Biophys Biochem Cytol.* 9:493-5.

McCroskery, S., M. Thomas, L. Maxwell, M. Sharma, and R. Kambadur. 2003. Myostatin negatively regulates satellite cell activation and self-renewal. *J Cell Biol.* 162:1135-47.

McGeachie, J.K., and M.D. Grounds. 1995. Retarded myogenic cell replication in regenerating skeletal muscles of old mice: an autoradiographic study in young and old BALBc and SJL/J mice. *Cell Tissue Res.* 280:277-82.

Megeney, L.A., B. Kablar, K. Garrett, J.E. Anderson, and M.A. Rudnicki. 1996. MyoD is required for myogenic stem cell function in adult skeletal muscle. *Genes Dev.* 10:1173-83.

Moeller, C., M.B. Yaylaoglu, G. Alvarez-Bolado, C. Thaller, and G. Eichele. 2002. Murine Lix1, a novel marker for substantia nigra, cortical layer 5, and hindbrain structures. *Brain Res Gene Expr Patterns.* 1:199-203.

Montarras, D., C. Lindon, C. Pinset, and P. Domeyne. 2000. Cultured myf5 null and myoD null muscle precursor cells display distinct growth defects. *Biol Cell.* 92:565-72.

Nusslein-Volhard, C., and E. Wieschaus. 1980. Mutations affecting segment number and polarity in Drosophila. *Nature.* 287:795-801.

Oustanina, S., G. Hause, and T. Braun. 2004. Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification. *Embo J.* 23:3430-9.

Pani, L., M. Horal, and M.R. Loeken. 2002. Rescue of neural tube defects in Pax-3-deficient embryos by p53 loss of function: implications for Pax-3- dependent development and tumorigenesis. *Genes Dev.* 16:676-80.

Perez-Iratxeta, C., G. Palidwor, C.J. Porter, N.A. Sanche, M.R. Huska, B.P. Suomela, E.M. Muro, P.M. Krzyzanowski, E. Hughes, P.A. Campbell, M.A. Rudnicki, and M.A. Andrade. 2005. Study of stem cell function using microarray experiments. *FEBS Lett.* 579:1795-801.

Phelan, S.A., M. Ito, and M.R. Loeken. 1997. Neural tube defects in embryos of diabetic mice: role of the Pax-3 gene and apoptosis. *Diabetes.* 46:1189-97.

Rawls, A., M.R. Valdez, W. Zhang, J. Richardson, W.H. Klein, and E.N. Olson. 1998. Overlapping functions of the myogenic bHLH genes MRF4 and MyoD revealed in double mutant mice. *Development.* 125:2349-58.

Relaix, F., D. Montarras, S. Zaffran, B. Gayraud-Morel, D. Rocancourt, S. Tajbakhsh, A. Mansouri, A. Cumano, and M. Buckingham. 2006. Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J Cell Biol.* 172:91-102.

Relaix, F., M. Polimeni, D. Rocancourt, C. Ponzetto, B.W. Schafer, and M. Buckingham. 2003. The transcriptional activator PAX3-FKHR rescues the defects of Pax3 mutant mice but induces a myogenic gain-of-function phenotype with ligand-independent activation of Met signaling in vivo. *Genes Dev.* 17:2950-65.

Relaix, F., D. Rocancourt, A. Mansouri, and M. Buckingham. 2004. Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev.* 18:1088-105.

Relaix, F., D. Rocancourt, A. Mansouri, and M. Buckingham. 2005. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature.* 435:948-53.

Rhodes, S.J., and S.F. Konieczny. 1989. Identification of MRF4: a new member of the muscle regulatory factor gene family. *Genes Dev.* 3:2050-61.

Rozen, S., and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* Bioinformatics Methods and Protocols: Methods in Molecular Biology. S. Krawetz and S. Misener, editors. Humana Press, Totowa, NJ. 365-386.

Rudnicki, M.A., T. Braun, S. Hinuma, and R. Jaenisch. 1992. Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. *Cell.* 71:383-90.

Rudnicki, M.A., P.N. Schnegelsberg, R.H. Stead, T. Braun, H.H. Arnold, and R. Jaenisch. 1993. MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell.* 75:1351-9.

Sabourin, L.A., A. Girgis-Gabardo, P. Seale, A. Asakura, and M.A. Rudnicki. 1999. Reduced differentiation potential of primary MyoD-/- myogenic cells derived from adult skeletal muscle. *J Cell Biol.* 144:631-43.

Sabourin, L.A., and M.A. Rudnicki. 2000. The molecular regulation of myogenesis. *Clin Genet.* 57:16-25.

Seale, P., A. Asakura, and M.A. Rudnicki. 2001. The potential of muscle stem cells. *Dev Cell.* 1:333-42.

Seale, P., J. Ishibashi, C. Holterman, and M.A. Rudnicki. 2004a. Muscle satellite cell-specific genes identified by genetic profiling of MyoD-deficient myogenic cell. *Dev Biol.* 275:287-300.

Seale, P., J. Ishibashi, A. Scime, and M.A. Rudnicki. 2004b. Pax7 is necessary and sufficient for the myogenic specification of CD45+:Sca1+ stem cells from injured muscle. *PLoS Biol.* 2:E130.

Seale, P., L.A. Sabourin, A. Girgis-Gabardo, A. Mansouri, P. Gruss, and M.A. Rudnicki. 2000. Pax7 is required for the specification of myogenic satellite cells. *Cell.* 102:777-86.

Sebire, N.J., and M. Malone. 2003. Myogenin and MyoD1 expression in paediatric rhabdomyosarcomas. *J Clin Pathol.* 56:412-6.

Soneoka, Y., P.M. Cannon, E.E. Ramsdale, J.C. Griffiths, G. Romano, S.M. Kingsman, and A.J. Kingsman. 1995. A transient three-plasmid expression system for the production of high titer retroviral vectors. *Nucleic Acids Res.* 23:628-33.

Sorrentino, V., R. Pepperkok, R.L. Davis, W. Ansorge, and L. Philipson. 1990. Cell proliferation inhibited by MyoD1 independently of myogenic differentiation. *Nature.* 345:813-5.

Spencer, J.A., S. Eliazer, R.L. Ilaria, Jr., J.A. Richardson, and E.N. Olson. 2000. Regulation of microtubule dynamics and myogenic differentiation by MURF, a striated muscle RING-finger protein. *J Cell Biol.* 150:771-84.

Spengler, D., M. Villalba, A. Hoffmann, C. Pantaloni, S. Houssami, J. Bockaert, and L. Journot. 1997. Regulation of apoptosis and cell cycle arrest by Zac1, a novel zinc finger protein expressed in the pituitary gland and the brain. *Embo J.* 16:2814-25.

Stoiber, W., and A.M. Sanger. 1996. An electron microscopic investigation into the possible source of new muscle fibres in teleost fish. *Anat Embryol (Berl).* 194:569-79.

Tajbakhsh, S. 2003. Stem cells to tissue: molecular, cellular and anatomical heterogeneity in skeletal muscle. *Curr Opin Genet Dev.* 13:413-22.

Tajbakhsh, S., D. Rocancourt, G. Cossu, and M. Buckingham. 1997. Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell.* 89:127-38.

Takahashi, M., Y. Kamei, and O. Ezaki. 2005. Mest/Peg1 imprinted gene enlarges adipocytes and is a marker of adipocyte size. *Am J Physiol Endocrinol Metab.* 288:E117-24.

Tapscott, S.J., R.L. Davis, M.J. Thayer, P.F. Cheng, H. Weintraub, and A.B. Lassar. 1988. MyoD1: a nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. *Science.* 242:405-11.

Tapscott, S.J., A.B. Lassar, and H. Weintraub. 1992. A novel myoblast enhancer element mediates MyoD transcription. *Mol Cell Biol.* 12:4994-5003.

Tatsumi, R., J.E. Anderson, C.J. Nevoret, O. Halevy, and R.E. Allen. 1998. HGF/SF is present in normal adult skeletal muscle and is capable of activating satellite cells. *Dev Biol.* 194:114-28.

Thayer, M.J., S.J. Tapscott, R.L. Davis, W.E. Wright, A.B. Lassar, and H. Weintraub. 1989. Positive autoregulation of the myogenic determination gene MyoD1. *Cell.* 58:241-8.

Tiffin, N., R.D. Williams, J. Shipley, and K. Pritchard-Jones. 2003. PAX7 expression in embryonal rhabdomyosarcoma suggests an origin in muscle satellite cells. *Br J Cancer.* 89:327-32.

Vaidya, T.B., S.J. Rhodes, E.J. Taparowsky, and S.F. Konieczny. 1989. Fibroblast growth factor and transforming growth factor beta repress transcription of the myogenic regulatory gene MyoD1. *Mol Cell Biol.* 9:3576-9.

Valente, T., F. Junyent, and C. Auladell. 2005. Zac1 is expressed in progenitor/stem cells of the neuroectoderm and mesoderm during embryogenesis: differential phenotype of the Zac1-expressing cells during development. *Dev Dyn.* 233:667-79.

Vivian, J.L., E.N. Olson, and W.H. Klein. 2000. Thoracic skeletal defects in myogenin- and MRF4-deficient mice correlate with early defects in myotome and intercostal musculature. *Dev Biol.* 224:29-41.

Vogan, K.J., and P. Gros. 1997. The C-terminal subdomain makes an important contribution to the DNA binding activity of the Pax-3 paired domain. *J Biol Chem.* 272:28289-95.

Vogan, K.J., D.A. Underhill, and P. Gros. 1996. An alternative splicing event in the Pax-3 paired domain identifies the linker region as a key determinant of paired domain DNA-binding activity. *Mol Cell Biol.* 16:6677-86.

Vorobyov, E., and J. Horst. 2006. Getting the proto-pax by the tail. *J Mol Evol.* 63:153-64.

Weintraub, H. 1993. The MyoD family and myogenesis: redundancy, networks, and thresholds. *Cell.* 75:1241-4.

Weintraub, H., R. Davis, S. Tapscott, M. Thayer, M. Krause, R. Benezra, T.K. Blackwell, D. Turner, R. Rupp, S. Hollenberg, and et al. 1991a. The myoD gene family: nodal point during specification of the muscle cell lineage. *Science.* 251:761-6.

Weintraub, H., V.J. Dwarki, I. Verma, R. Davis, S. Hollenberg, L. Snider, A. Lassar, and S.J. Tapscott. 1991b. Muscle-specific transcriptional activation by MyoD. *Genes Dev.* 5:1377-86.

Wyzykowski, J.C., T.I. Winata, N. Mitin, E.J. Taparowsky, and S.F. Konieczny. 2002. Identification of novel MyoD gene targets in proliferating myogenic stem cells. *Mol Cell Biol.* 22:6199-208.

Xia, S.J., J.G. Pressey, and F.G. Barr. 2002. Molecular pathogenesis of rhabdomyosarcoma. *Cancer Biol Ther.* 1:97-104.

Xue, L., and M. Noll. 1996. The functional conservation of proteins in evolutionary alleles and the dominant role of enhancers in evolution. *Embo J.* 15:3722-31.

Yablonka-Reuveni, Z., R. Seger, and A.J. Rivera. 1999. Fibroblast growth factor promotes recruitment of skeletal muscle satellite cells in young and old rats. *J Histochem Cytochem.* 47:23-42.

Zammit, P.S., J.J. Carvajal, J.P. Golding, J.E. Morgan, D. Summerbell, J. Zolnerciks, T.A. Partridge, P.W. Rigby, and J.R. Beauchamp. 2004. Myf5 expression in

satellite cells and spindles in adult muscle is controlled by separate genetic elements. *Dev Biol.* 273:454-65.

Zammit, P.S., L. Heslop, V. Hudon, J.D. Rosenblatt, S. Tajbakhsh, M.E. Buckingham, J.R. Beauchamp, and T.A. Partridge. 2002. Kinetics of myoblast proliferation show that resident satellite cells are competent to fully regenerate skeletal muscle fibers. *Exp Cell Res.* 281:39-49.

Zammit, P.S., F. Relaix, Y. Nagata, A.P. Ruiz, C.A. Collins, T.A. Partridge, and J.R. Beauchamp. 2006. Pax7 and myogenic progression in skeletal muscle satellite cells. *J Cell Sci.* 119:1824-32.

Zhang, J.M., X. Zhao, Q. Wei, and B.M. Paterson. 1999. Direct inhibition of G(1) cdk kinase activity by MyoD promotes myoblast cell cycle withdrawal and terminal differentiation. *Embo J.* 18:6983-93.

Ziman, M.R., S. Fletcher, and P.H. Kay. 1997. Alternate Pax7 transcripts are expressed specifically in skeletal muscle, brain and other organs of adult mice. *Int J Biochem Cell Biol.* 29:1029-36.

Ziman, M.R., and P.H. Kay. 1998. Differential expression of four alternate Pax7 paired box transcripts is influenced by organ- and strain-specific factors in adult mice. *Gene.* 217:77-81.

Appendix A

Real-time PCR Analysis of Pax7 Isoform Ratios

Purpose

To develop a fast and efficient method for determining the ratios of expression of the four alternatively spliced Pax7 transcripts.

Abstract

Alternative splicing in the paired domain of Pax7 has a substantial effect on the magnitude of target gene transactivation, probably by altering DNA binding site recognition or affinity. However, the *in vivo* distribution of Pax7 isoforms, as well as the possibility of dynamic regulation of isoform expression, has been under-characterized due to the lack of a rapid and efficient method for distinguishing the relative levels of four transcripts that differ by no more than nine nucleotides. Real-time PCR is well-suited for such an assay, yet the specific structure of the Pax7 $GL^{+/-}$ junction led to non-specific amplification that complicated the quantitation. However, the quantitative nature of the real-time PCR assay allowed for a subtractive correction that provided reasonable estimates of the Pax7 isoform transcript ratios. This method was applied to demonstrate that, although total Pax7 mRNA expression is sharply down-regulated at the onset of differentiation, the proportions of the four Pax7 isoforms are maintained.

Introduction

Pax7 is a paired-box transcription factor with an essential role in the development and maintenance of the myogenic satellite cell population in post-natal skeletal muscle (Kuang et al., 2006; Seale et al., 2000). Satellite cells are the major stem cell population involved in regenerating damaged skeletal muscle in response to injury or exercise (rev. in Bischoff, 1994; Zammit et al., 2002). In the absence of Pax7, satellite cells are missing from adult tissue and the regenerative capacity of the muscles is nearly entirely ablated (Kuang et al., 2006; Seale et al., 2000). The closely related Pax3 gene is expressed in a residual myogenic cell population (<1% of wild-type myoblasts) but is incapable of providing any significant rescue in response to damage (Kassar-Duchossoy et al., 2005; Kuang et al., 2006; Oustanina et al., 2004).

Pax7 and Pax3 share 94% identity within their paired-type homeodomain, octapeptide, and homeodomain regions and 85% similarity overall, and thus form a subgroup amongst the nine mammalian Pax genes. Both are alternatively spliced between their second and third exons, resulting in the inclusion or exclusion of a single glutamine ($Q^{+/-}$) residue at position #138 (Vogan et al., 1996; Ziman and Kay, 1998). Pax7 has an additional alternative splicing change that produces the inclusion or exclusion of a glycine-leucine dipeptide ($GL^{+/-}$) at the exon 3/4 junction for residues #181/2 or #182/3 (Table I) (Ziman and Kay, 1998). Both changes occur within the paired DNA binding domain. In the case of Pax3, this has been shown to result in altered

Table I. Alternative splicing in the Pax7 paired-box DNA binding domain.[1]

Isoform	exon 2/3	exon 3/4
	aga ***CAG*** gtg	tca G***GT TTA g***tg
	Q	GL
A	-	+
B	+	+
C	-	-
D	+	-

[1] Ziman and Kay (1998a)

affinity to its DNA binding site (Vogan et al., 1996), while the GL change in Pax7 modifies downstream target gene activation (see Chapter 3).

Our particular interest is in Pax7 due to its major role in adult myogenesis, in satellite cells (Kuang et al., 2006; Seale et al., 2000) and in driving non-muscle adult stem cells into the skeletal muscle lineage (Seale et al., 2004). As the four isoforms differentially regulate Pax7 target genes (see Chapter 3), it will be important to be able to assess how their expression is regulated within the context of overall Pax7 transcription. Previously, the relative isoform expression levels were characterized amongst different mouse strains using a labour-intensive PCR-cloning-sequencing strategy (Ziman and Kay, 1998). We have used an efficient semi-quantitative real-time PCR-based system to address this question in cultured myoblasts during growth and differentiation. This has demonstrated that the proportions of each isoform are consistent while total Pax7 transcripts are rapidly down-regulated with the onset of skeletal muscle differentiation.

Materials and Methods

Cell Culture

10T1/2 fibroblast and C2C12 myoblast cells were cultured in growth medium of 10% fetal calf serum in DMEM supplemented with penicillin/streptomycin. Myogenic differentiation of the C2C12 cells was induced by changing the culture medium to 2% horse serum in DMEM with penicillin/streptomycin. Pax7-expressing 10T1/2 fibroblasts were made by infecting pools ($> 10^5$ cells) with retrovirus expressing a particular isoform of Pax7 and selecting for successful integration and expression of an included

puromycin-selection cassette with 1-2 μg/μl of puromycin. Retrovirus was made as previously described (Ishibashi et al., 2005).

RNA Isolation and Reverse Transcription

Total RNA was purified using the RNeasy Mini kit (Qiagen) and quantified by spectrophotometry. Equal quantities (1000 ng) of total RNA were reverse transcribed using the RNA PCR Core Kit (Perkin-Elmer) with random hexamer primers according to the manufacturer's instructions. Each 20 μl RT reaction was diluted 1:5 with water to provide a master template for all related PCR reactions.

Primer Design

Primers were designed against the reference mouse Pax7 sequence (NM_011039). Annealing temperatures were calculated using the online Primer3 software (Whitehead Institute; http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Primer sequences are shown in Table II.

Real-time PCR

Real-time PCR reactions included 2 μl of diluted reverse transcription product, 10 μl of 2x iQ SYBR Green Supermix (Bio-Rad Laboratories), 30 nM ROX passive reference dye (Stratagene), and 50 nM of each forward and reverse PCR primer. Data were gathered using the MX4000 real-time system (Stratagene) over 40 cycles of (94°C - 30 s; 58-60°C - 30-60 s; 72°C - 30 s) followed by a 41-cycle denaturation curve from 55° to 95°C to

Table II. Mouse Pax7 isoform primers

Primer	Orientation	Sequence	Targets
Q+	forward	5' cagcaagcccagacag 3'	B, D
Q-	forward	5' cagcaagcccagagtg 3'	A, C
GL+	reverse	5' ggctaatcgaactcactaaac 3'	A, B
GL-	reverse	5' ggctaatcgaactcactga 3'	C, D

ensure amplification specificity. C_t values were calculated by the MX4000 software using moving window averaging and an adaptive baseline. Further calculations and chart plotting were performed in Excel (Microsoft).

PCR Cross-Amplification Compensation

We calculated a compensation for cross-amplification between different isoforms based on the following reasoning:

1. The fluorescence value in a SYBR green real-time PCR reaction is proportional to the total quantity of double-stranded DNA present.
2. Primer efficiency (E) - the probability that a product will be amplified in a cycle - is equivalent to an exponential increase with a base of (1+E) over the course of S cycles, or $(1+E)^S$.
3. The SYBR green fluorescence with a single template will therefore correspond to the starting quantity of template (X) multiplied S cycles of amplification, or $X(1+E)^S$.
4. The total fluorescence in a mixed sample after S cycles will be:

   (1) $F = X_1(1+E_1)^S + X_2(1+E_2)^S + X_2(1+E_2)^S + \ldots = \Sigma_1^n X_n(1+E_n)^S$

   for starting template quantities $X_1, X_2, \ldots, X_n$ and the corresponding amplification efficiencies of a given pair of primers on that template $E_1, E_2, \ldots, E_n$.

C_t values are calculated as the cycle number at which the fluorescence crosses a threshold fluorescence level. By using the identical threshold level for all samples within a 96-well plate, we assume that the fluorescence signal (F_{WC}) for each well W is equivalent, with

cycle number S (= C_t) varying depending on the amount of amplification required to reach the threshold. We have used dRn for this purpose, a value that is background-subtracted and normalized to the ROX passive reference dye that is spiked into reactions. The MX4000 uses a scanning head to read each well, so that the equipment is invariant from well to well. With this in mind, and the appropriate use of master mixes to prepare reactions, the dRn value is an appropriate measure for this analysis.

Thus, the amplification of different primer pairs on related transcripts can be formulated as a system of linear equations, for Pax7 appearing as:

(2)
$$\begin{aligned} F_A &= X_A(1+E_{AA})^{SA} + X_B(1+E_{AB})^{SA} + X_C(1+E_{AC})^{SA} + X_D(1+E_{AD})^{SA} \\ F_B &= X_A(1+E_{BA})^{SB} + X_B(1+E_{BB})^{SB} + X_C(1+E_{BC})^{SB} + X_D(1+E_{BD})^{SB} \\ F_C &= X_A(1+E_{CA})^{SC} + X_B(1+E_{CB})^{SC} + X_C(1+E_{CC})^{SC} + X_D(1+E_{CD})^{SC} \\ F_D &= X_A(1+E_{DA})^{SD} + X_B(1+E_{DB})^{SD} + X_C(1+E_{DC})^{SD} + X_D(1+E_{DD})^{SD} \end{aligned}$$

for fluorescence value F_A with initial Pax7A template quantity X_A, efficiency E_{AA} of the primers for Pax7A on template Pax7A, after S_A cycles of PCR.

The values for E_{xx} were estimated by amplifying the identical quantity of known template with the mismatched primer pairs and comparing them to the matched primer pair. Again, with identical threshold fluorescence values, comparing two primer pairs becomes:

(3)
$$\begin{aligned} X_A(1+E_{AA})^{SA} &= X_B(1+E_{AB})^{SB} \\ (1+E_A)^{SA} &= (1+E_B)^{SB} \qquad \text{(identical template quantity, } X_A=X_B\text{)} \\ (1+E_A) &= (1+E_B)^{SB\,/\,SA} \end{aligned}$$

Thus, based on a known or estimated efficiency of amplification for a perfect-match primer pair and the critical cycle values S_A and S_B for both pairs, the relative efficiencies of the other primers can be estimated.

Using those values of E_x, and assaying each sample with each of the four Pax7 primer pairs to obtain the values for S_x, we therefore have sufficient information to solve the system of equations for the initial isoform template quantities X_x. While the system could be solved analytically, we have instead implemented it as a Solver function in an Excel worksheet.

Results and Discussion

Our objective was to design a PCR-based system that could be easily and quickly be used to assess the relative expression levels of four closely related alternatively spliced mouse Pax7 transcripts. To do this, the designs for our primers were constrained by a number of criteria. Each pair of primers required one member to be located at the exon 2/3 junction and the other at the exon 3/4 junction. This results in a PCR product slightly larger than exon 3 (127 bp, Q^-GL^-), which is acceptably sized for SYBR green real-time PCR. Primers were designed to be at least 16bp in length to provide adequate specificity in a mixture of reverse-transcribed products of genomic complexity. Also, primer sequences needed to be mismatched against the very closely related mouse Pax3 transcript. Primer annealing temperatures were matched in order to allow all of the PCR reactions to be conducted simultaneously within the same run.

Our initial attempts to design PCR primers specific to each isoform were only partially successful. The specificity of each primer pair was evaluated with RT product templates generated from 10T1/2 fibroblasts stably infected with retrovirus expressing a single Pax7 isoform. 10T1/2 cells do not express the endogenous Pax7 gene (Chapter 3).

Primers Q^+ and GL^+ (for Pax7B) were easily designed because the 3' end of each recognizes the included bases. This results in primers with three or six base-pair mismatches against the Q^- or GL^- transcripts. The other two primers cannot be so designed and by necessity recognize sequences that are found within all four different transcripts. Fortunately, overlapping the splice junction with primer Q^- was effective, with the requirement for the formation of a 3 bp bulge in the Q^+ template apparently limiting. Primer GL^- selection is confounded because there is a 'G' base-pair at the splice junction with either the GL^+ and GL^- transcripts. Therefore, there is always at minimum 2 bp match at the 3' end of the primer. Unfortunately, the 6 bp loop related to primer GL^- binding to a GL^+ template was not disruptive, and amplification occurred with surprisingly high efficiency on such transcripts, even when the position of the mismatch was moved to various positions within the primer (not shown). This resulted in considerable cross-amplification by the Pax7C primers on Pax7A templates and the Pax7D primers on Pax7B templates.

Varying both the annealing temperature and time within the PCR cycle demonstrated a trade-off between specificity and sensitivity. For conditions of 58°C / 60 s, the efficiency of amplification (and hence sensitivity) was high (90-95%), but specificity was poor, with primer GL^- improperly amplifying the GL^+ template with equal (and possibly greater) efficiency than the GL^- template. In contrast, at 60°C / 30 s, the same primer was much more specific (10-fold) but the efficiency of amplification was also poor 70-75%. Intermediate values showed that increased specificity always cost considerable sensitivity.

Although highly specific primers are obviously desirable, we noted that the problematic primer (GL^-) resulted in cross-amplification of only one additional transcript. We therefore used equation (2) (see Materials and Methods) to compensate for this effect. A notable aspect of this system is that its quality is highly dependent on the estimation of the individual amplification efficiencies. Empirically, we have found that estimating the relative efficiencies according to (3) produces good results with relatively little dependence on having a highly accurate value for the perfect match efficiency. While standard curves can be used to derive each set of efficiencies, the variability in that process appears to introduce more uncertainty than the simpler process of finding relative values by (3).

Eq. (2), with high-specificity primers, results in a very high perfect-match efficiency (possibly approaching 100%, or theoretically ideal PCR) and a very low efficiency (down to 0%, or no increase in product) for mismatched pairs. In this case (particularly over the standard 20 or more cycles of PCR), the contributions of the mismatch terms rapidly shrink to insignificance. In our system, this is critical because it allows us to compensate for the two non-specific primer pair/template combinations. The highly specific amplification of Pax7A and Pax7B allow their contributions to the Pax7C and Pax7D estimate to be reasonably subtracted. Improvements to the specificities of the primers (the GL^- in particular) would substantially improve the robustness of this calculation.

We applied our system to the described set of known samples (10T1/2 fibroblasts over-expressing Pax7) and tested its ability to estimate the relative levels of the Pax7

isoforms in a 1:1:1:1 mixture of RT products from all four sources. All of the samples were RT products from cells and therefore were complex mixtures of templates. This is noted because we have observed differences in amplification efficiencies between, for example, purified PCR product template and RT products.

Whereas the proportions of each isoform were expected to be 25% each, our estimates ranged from a low value of 20% for Pax7A to a high value of 31% for Pax7D (Figure 1). As noted earlier, estimating the relative match/mismatch efficiencies by (3) produced the most consistent results, with little change when the reference perfect-match efficiency was arbitrarily adjusted up or down (with the mismatch efficiencies obviously following through (3)). In contrast, constructing standard curves for each primer pair led to much greater variation ranging from 15% (Pax7A) to 39% (Pax7D) (not shown). Thus, within a compact 96-well PCR assay we can assay the isoform ratios within ~20% of the expected value in a known mixture. While this leaves substantial room for improvement, this system is well-suited to examining the ratios of Pax7 isoform expression in a series of related samples.

A primary myoblast differentiation time course was used to test whether or not there was a change in the Pax7 isoform (mRNA) distribution in differentiating myoblasts. Pax7 is normally expressed in growing myoblasts and down-regulated as they enter differentiation. As expected, real-time PCR with primers for a common region of the Pax7 transcript shows that total Pax7 mRNA levels are rapidly reduced when differentiation is induced (Figure 2 A). Within this total pool of Pax7 mRNAs, however,

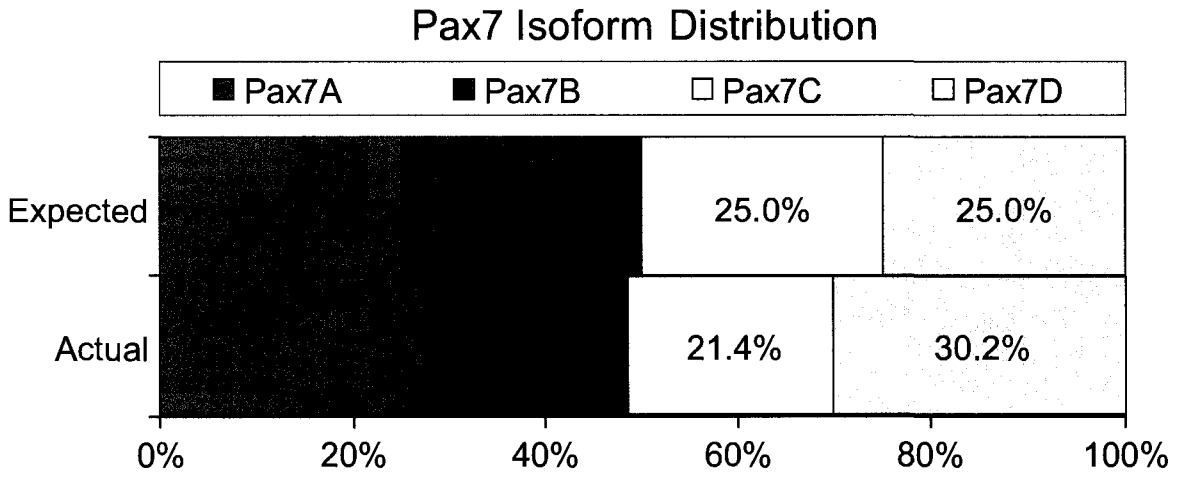


Figure 1: Pax7 isoform distribution in a known mixture. A 1:1:1:1 mixture of RT products derived from four pools of 10T1/2 cells, each expressing a different isoform of Pax7, was used as template to test the accuracy of the relative PCR quantitation.

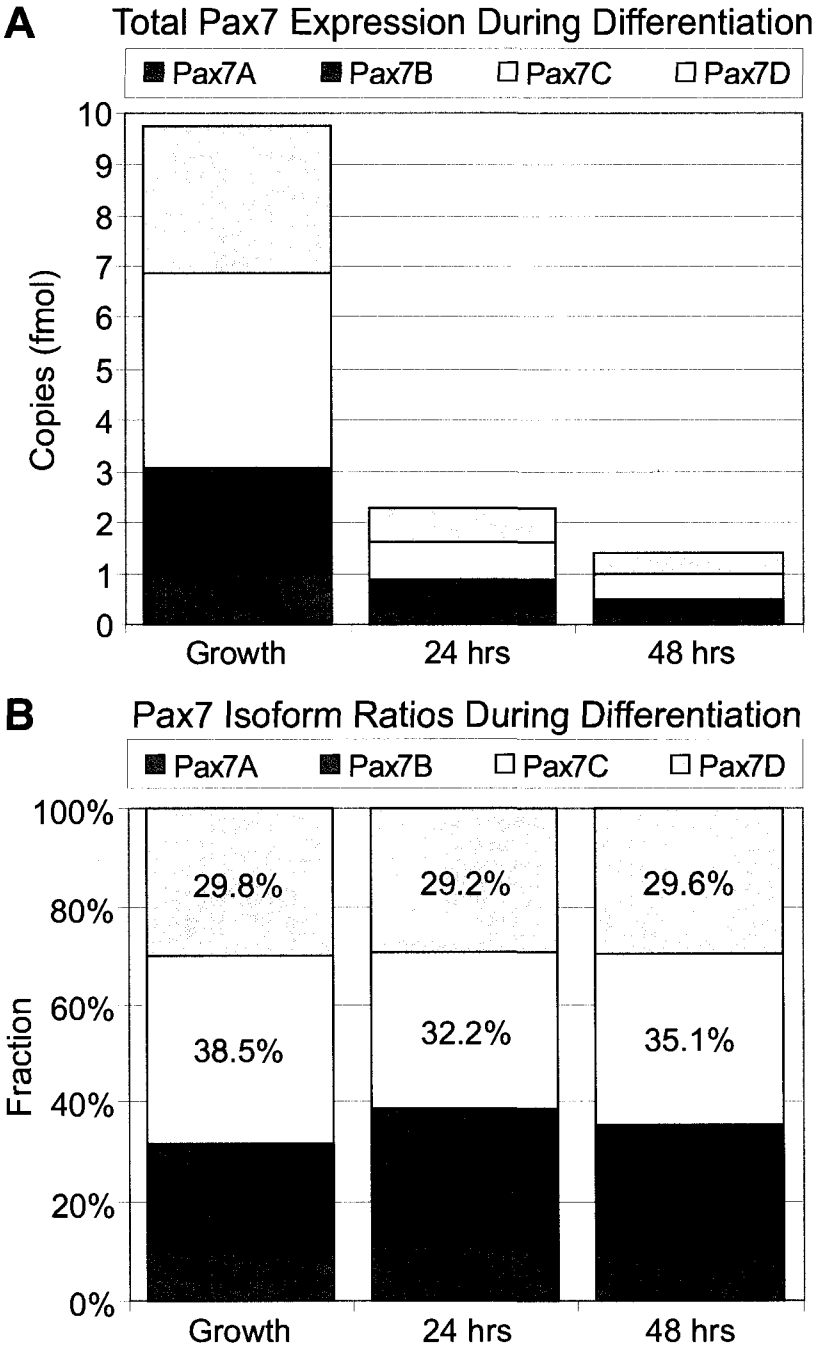


Figure 2: Pax7 isoform distribution during primary myoblast differentiation. Total RNA was harvested from primary myoblast cultures in growth phase and at 24 hrs and 48 hrs after the induction of differentiation, and equal quantities used for RT and Pax7 isoform-specific SYBR green real-time PCR. While the total quantity of Pax7 transcripts drops dramatically in the first 48 hours of differentiation, the proportions of the four isoforms are consistent over the same period.

the relative proportions of the transcripts for each isoform are effectively unchanged (Figure 2 B).

The best available conditions for this PCR occur with an annealing temperature of 60°C, for 30-60 s. For an abundant target, 30 s is more effective because the gain in specificity is valuable when the sensitivity of detection is not a problem. Conversely, amplification is more efficient with a step-duration of 60 s (close to 100% for all primer pairs) for very good sensitivity, but the GL^- primer does not distinguish between $GL^{+/-}$ transcripts under those conditions. Nonetheless, the GL^+ primer is more selective by several orders of magnitude, therefore allowing the GL^+ contribution to be subtracted.

Several alternative strategies are available that fulfill the same purpose as this real-time PCR system, with different costs and benefits. Traditionally, reverse-transcribed fragments have been amplified by PCR, ligated into plasmids, and transformed into bacteria. By picking a large number of transformed colonies and sequencing the inserts of their plasmids, the frequency of recovery of each isoform can be estimated. This method, however, is very time consuming, labour intensive, and expensive. Such a method could be modified into a SAGE-like system, with the RT-PCR products concatemerized for sequencing in order to decrease the number of clones required. However, although streamlined, it would still suffer from the same drawbacks that afflict the original method.

Using flanking PCR primers that amplify all four isoforms simultaneously is another possibility. By radioactively (or non-radioactively) labeling the products during PCR, the different isoforms would be easily separable by native polyacrylamide gel

electrophoresis into four distinct bands, with 3bp of spacing between each (smallest to largest = C, A, D, B). Phosphor imaging and quantitation of the bands would provide the ratios amongst them for a sample. Notably, because all four products amplify in the same tube competing for the same primers and reagents, an endpoint measure after the PCR plateau should remain an accurate gauge of the isoforms' relative proportions. Total Pax7 transcript amounts for comparison between different samples should be obtained quantitatively with a distinct primer pair amplifying a common region in the transcript. This method is more time-consuming than the real-time PCR format and will have less dynamic range for differences in concentrations between isoforms. It is, however, more conceptually straightforward.

The real-time PCR method outlined is an expedient method, requiring only a few hours to go from an RNA sample to analysis. It is very sensitive and has the exceptional dynamic range typical of real-time PCR quantitation. However, the expense of the PCR reagents and the corrections required in analyzing the data may be limiting to the scale and precision of the method. It remains to be seen if there are situations where the isoform ratios are large or change greatly (more than an order of magnitude), both conditions that would benefit this strategy compared to the phosphor imaging approach. In practice, a hybrid method, using phosphor imaging for the isoform ratios and real-time PCR for overall Pax7 transcript quantitation, may be most accurate and effective.

There are numerous problems awaiting such a solution. For example, Pax7 is also expressed in developing neural tube, neuroepithelium, and nasal olfactory epithelium (Jostes et al., 1990; Mansouri et al., 1996), and is required for the specification of neural

crest during gastrulation (Basch et al., 2006). The expression of Pax7 isoforms differs between brain and skeletal muscle tissues (Ziman and Kay, 1998) and could be involved in differing functions for Pax7 in neural versus myogenic development. This seems likely, given the recent finding that alternate Pax7 isoforms have differing abilities to activate target genes (Chapter 1). Additionally, the translocation-induced fusion of the DNA binding domains of Pax7 with the translocation domain of FoxO1 (FKHR) produces Pax7/FKHR and the childhood tumour alveolar rhabdomyosarcoma. Any change in the relative expression levels of the isoforms may skew the activation of target genes and hence contribute to tumour development and metastasis. This fast and scalable method allows such questions to be answered.

References

Basch, M.L., M. Bronner-Fraser, and M.I. Garcia-Castro. 2006. Specification of the neural crest occurs during gastrulation and requires Pax7. *Nature*. 441:218-22.

Bischoff, R. 1994. The satellite cell and muscle regeneration. *In* Myology. A.G. Engel, Franzini-Armstrong, C., editor. McGraw-Hill, New York. 97-118.

Ishibashi, J., R.L. Perry, A. Asakura, and M.A. Rudnicki. 2005. MyoD induces myogenic differentiation through cooperation of its NH2- and COOH-terminal regions. *J Cell Biol*. 171:471-82.

Jostes, B., C. Walther, and P. Gruss. 1990. The murine paired box gene, Pax7, is expressed specifically during the development of the nervous and muscular system. *Mech Dev*. 33:27-37.

Kassar-Duchossoy, L., E. Giacone, B. Gayraud-Morel, A. Jory, D. Gomes, and S. Tajbakhsh. 2005. Pax3/Pax7 mark a novel population of primitive myogenic cells during development. *Genes Dev*. 19:1426-31.

Kuang, S., S.B. Charge, P. Seale, M. Huh, and M.A. Rudnicki. 2006. Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. *J Cell Biol*. 172:103-13.

Mansouri, A., A. Stoykova, M. Torres, and P. Gruss. 1996. Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice. *Development*. 122:831-8.

Oustanina, S., G. Hause, and T. Braun. 2004. Pax7 directs postnatal renewal and propagation of myogenic satellite cells but not their specification. *Embo J*. 23:3430-9.

Seale, P., J. Ishibashi, A. Scime, and M.A. Rudnicki. 2004. Pax7 is necessary and sufficient for the myogenic specification of CD45+:Sca1+ stem cells from injured muscle. *PLoS Biol.* 2:E130.

Seale, P., L.A. Sabourin, A. Girgis-Gabardo, A. Mansouri, P. Gruss, and M.A. Rudnicki. 2000. Pax7 is required for the specification of myogenic satellite cells. *Cell.* 102:777-86.

Vogan, K.J., D.A. Underhill, and P. Gros. 1996. An alternative splicing event in the Pax-3 paired domain identifies the linker region as a key determinant of paired domain DNA-binding activity. *Mol Cell Biol.* 16:6677-86.

Zammit, P.S., L. Heslop, V. Hudon, J.D. Rosenblatt, S. Tajbakhsh, M.E. Buckingham, J.R. Beauchamp, and T.A. Partridge. 2002. Kinetics of myoblast proliferation show that resident satellite cells are competent to fully regenerate skeletal muscle fibers. *Exp Cell Res.* 281:39-49.

Ziman, M.R., and P.H. Kay. 1998. Differential expression of four alternate Pax7 paired box transcripts is influenced by organ- and strain-specific factors in adult mice. *Gene.* 217:77-81.

Appendix B

Endogenous Pax7 Expression Is Induced Following Myogenic Conversion of $CD45^{+}/Sca1^{+}$ Cells by Ectopic Expression of Pax7d

Purpose

To determine if the endogenous Pax7 gene is induced following myogenic conversion of the $CD45^+/Sca1^+$ population by ectopic expression of Pax7d.

Submission Information

Submitted: January 12, 2004
Accepted: February 25, 2004
Published: May 11, 2004

Seale P, Ishibashi J, Scime A, Rudnicki MA. Pax7 is necessary and sufficient for the myogenic specification of $CD45^+$:$Sca1^+$ stem cells from injured muscle. PLoS Biol. 2004 May;2(5):E130.

Contributions of Co-Authors

Jeff Ishibashi provided the RT-PCR data as Figure 4B in the manuscript.

Preface and Discussion of Figure 4B

$CD45^+$:$Sca1^+$ adult stem cells isolated from uninjured muscle do not display any myogenic potential, but become myoblasts if infected with Pax7d-expressing retrovirus. Primers amplifying the 3' end and 3'UTR of the full-length Pax7 transcript (not present in the retrovirus) were used for RT-PCR to assess whether or not the ectopic Pax7d was activating the endogenous Pax7 gene. This result would indicate if Pax7d was capable of myogenic conversion alone, or if this effect might require the expression of all four isoforms.

Pax7d expression in $CD45^+$:$Sca1^+$ cells from uninjured wild-type muscle activated expression of the endogenous Pax7 gene (Figure 4B). Therefore, conversion may be mediated by any one of or a combination of the Pax7 isoforms. The requirement for isoforms other than Pax7d could explain the lack of conversion of the $CD45^+$:$Sca1^+$ population from *Pax7*$^{-/-}$ muscle. However, this is tempered by the successful conversion of the *Pax7*$^{-/-}$;$CD45^-$;$Sca1^-$ cells by Pax7d. Testing of the capabilities of each of Pax7a, Pax7b, and Pax7c to convert the wild-type $CD45^+$:$Sca1^+$ cells is needed. Differences between the $CD45^+$:$Sca1^+$ and $CD45^-$:$Sca1^-$ may also contribute to their receptiveness to Pax7-induced conversion.

Pax7 Is Necessary and Sufficient for the Myogenic Specification of CD45$^+$:Sca1$^+$ Stem Cells from Injured Muscle

Patrick Seale[1,2], Jeff Ishibashi[2], Anthony Scimè[2], Michael A. Rudnicki[1,2*]

1 Department of Biology, McMaster University, Hamilton, Ontario, Canada, 2 Ottawa Health Research Institute, Molecular Medicine Program, Ottawa, Ontario, Canada

CD45$^+$:Sca1$^+$ adult stem cells isolated from uninjured muscle do not display any myogenic potential, whereas those isolated from regenerating muscle give rise to myoblasts expressing the paired-box transcription factor Pax7 and the bHLH factors Myf5 and MyoD. By contrast, CD45$^+$:Sca1$^+$ isolated from injured *Pax7* $^{-/-}$ muscle were incapable of forming myoblasts. Infection of CD45$^+$:Sca1$^+$ cells from uninjured muscle with retrovirus expressing Pax7 efficiently activated the myogenic program. The resulting myoblasts expressed Myf5 and MyoD and differentiated into myotubes that expressed myogenin and myosin heavy chain. Infection of CD45$^-$:Sca1$^-$ cells from *Pax7* $^{-/-}$ muscle similarly gave rise to myoblasts. Notably, infection of *Pax7*-deficient muscle with adenoviral Pax7 resulted in the de novo formation of regenerated myofibers. Taken together, these results indicate that Pax7 is necessary and sufficient to induce the myogenic specification of CD45$^+$ stem cells resident in adult skeletal muscle. Moreover, these experiments suggest that viral transduction of Pax7 is a potential therapeutic approach for the treatment of neuromuscular degenerative diseases.

Introduction

Skeletal muscle regeneration has long been considered to be mediated solely by monopotential skeletal muscle stem cells known as satellite cells (Bischoff 1994; Charge and Rudnicki 2004). However, recent studies have identified novel populations of adult stem cells in skeletal muscle. For example, "side-population" (SP) cells isolated from muscle tissue participate in the regeneration of skeletal muscle and give rise to satellite cells (Gussoni et al. 1999; Asakura et al. 2002). In vitro, muscle SP cells readily form hematopoietic colonies, but do not spontaneously differentiate into muscle cells unless cocultured with satellite-cell-derived myoblasts (Asakura et al. 2002).

Various cell surface markers have been employed to purify adult stem cell populations from skeletal muscle, including c-kit, Sca1, CD34, and CD45 (reviewed by Charge and Rudnicki 2004). Almost all muscle-derived hematopoietic progenitor and blood reconstitution activity is derived from CD45$^+$ cells (Asakura et al. 2002; McKinney-Freeman et al. 2002). Muscle-derived CD45$^+$ cells purified from uninjured muscle are uniformly nonmyogenic in vitro and do not form muscle in vivo (Asakura et al. 2002; McKinney-Freeman et al. 2002). However, coculture and in vivo injection experiments indicate that CD45$^+$ SP, as well as CD45$^-$ SP, cells possess myogenic potential (Asakura et al. 2002; McKinney-Freeman et al. 2002).

Recent experiments have established that CD45$^+$ adult stem cells have a normal physiological role in tissue regeneration (Polesskaya et al. 2003). CD45$^+$:Sca1$^+$ cells display a 30-fold expansion in number following cardiotoxin-induced (ctx-induced) injury. Importantly, a large proportion of CD45$^+$:Sca1$^+$ cells isolated from regenerating muscle acquire myogenic potential and appear to represent a significant source of myogenic progenitors during regenerative myogenesis (Polesskaya et al. 2003). Moreover, the myogenic specification of these adult stem cells during regeneration occurs by a Wnt-signaling-dependent mechanism (Polesskaya et al. 2003).

The paired-box transcription factor Pax7 is specifically expressed in satellite cells and is required for the specification of the satellite cell lineage (Seale et al. 2000). Following Wnt treatment of isolated CD45$^+$ adult stem cells, Pax7 is rapidly induced as an early marker of satellite cell myogenic specification (Polesskaya et al. 2003). Together, these data suggest the hypothesis that Pax7 represents the target of Wnt signaling that directs the myogenic specification of adult stem cells resident in muscle. To investigate this hypothesis, we examined the myogenic potential of adult stem cells from *Pax7* $^{-/-}$ muscle, and employed viral vectors to transduce Pax7 into cells in vivo and in vitro. Our experiments demonstrate that Pax7 induces the myogenic program in specific populations of adult stem cells within muscle tissue and support the conclusion that Pax7 regulates myogenic determination during regenerative myogenesis.

Results

Pax7 Is Required for the Myogenic Commitment of CD45$^+$:Sca1$^+$ Cells

To determine whether Pax7 is required for myogenesis in muscle-derived CD45$^+$ cells, we analyzed the myogenic

Received January 12, 2004; Accepted February 25, 2004; Published May 11, 2004
DOI: 10.1371/journal.pbio.0020130

Abbreviations: Ad-Pax7, adenovirus-Pax7; Ad-LacZ, adenovirus-β-*galactosidase*; CDSC-Pax7, CD45$^+$:Sca1$^+$ cells expressing retroviral Pax7; ctx, cardiotoxin; DSHB, Developmental Studies Hybridoma Bank; HAN-puro, retroviral HAN-puromycin; MRF, myogenic regulatory factor; MyHC, myosin heavy chain; RT-PCR, reverse transcriptase PCR; SP, side population; TA, tibialis anterior

Academic Editor: Jon Epstein, University of Pennsylvania

*To whom correspondence should be addressed. E-mail: mrudnicki@ohri.ca

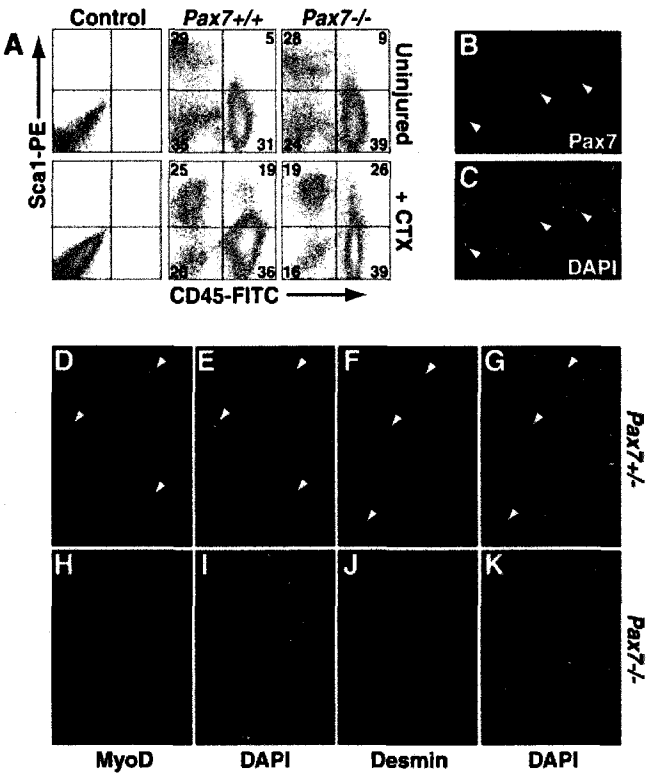


Figure 1. Pax7 Is Required for the Myogenic Specification of $CD45^+$:$Sca1^+$ Cells

(A) Flow cytometric analysis of cell suspensions derived from uninjured and regenerating wild-type and *Pax7* $^{-/-}$ muscle (4 d after ctx injection) showed an increased proportion of $CD45^+$ cells in *Pax7* $^{-/-}$ samples.
(B and C) Pax7 protein was expressed in approximately 6%–10% of $CD45^+$:$Sca1^+$ cells purified from regenerating *Pax7* $^{+/-}$ muscle.
(D–K) MyoD (D and E) and Desmin (F and G) were induced in $CD45^+$:$Sca1^+$ cells from regenerating *Pax7* $^{+/-}$ but were not expressed in $CD45^+$:$Sca1^+$ cells from regenerating *Pax7* $^{-/-}$ muscle (H–K).
DOI: 10.1371/journal.pbio.0020130.g001

differentiation capacity of $CD45^+$:$Sca1^+$ cells from *Pax7* $^{-/-}$ muscle undergoing ctx-induced regeneration. Flow cytometry analysis revealed a higher average proportion of CD45-expressing cells in *Pax7* $^{-/-}$ muscle relative to wild-type (Figure 1A). Specifically, in muscle suspensions from *Pax7* $^{-/-}$ and wild-type littermates, 39% ± 4% versus 31% ± 9% of cells were $CD45^+$:$Sca1^-$ and 9% ± 2% versus 5% ± 2% of cells were $CD45^+$:$Sca1^+$, respectively ($n \geq 6$). Four days following ctx injury, a significantly higher proportion of $CD45^+$:$Sca1^+$ cells (26% ± 3% compared to 19% ± 4% for *Pax7* $^{-/-}$ and wild-type, respectively, $p < 0.05$) and a reduced proportion of $CD45^-$:$Sca1^+$ cells were observed in *Pax7* $^{-/-}$ muscle (19% ± 4% compared to 25% ± 6% for *Pax7* $^{-/-}$ and wild-type, respectively, $p = 0.07$, $n = 3$) (Figure 1A).

By immunohistochemical analysis, Pax7 protein was upregulated in 6%–10% of $CD45^+$:$Sca1^+$ cells from wild-type muscle 4 d after ctx injury (Figure 1B and 1C). Importantly, Pax7 expression was not detected in $CD45^+$:$Sca1^+$ cells purified from uninjured muscles (Polesskaya et al. 2003). Furthermore, MyoD– (Figure 1D and 1E) and Desmin–immunoreactive cells (Figure 1F and 1G) were readily detected in cultured (18 h in growth medium) $CD45^+$:$Sca1^+$ cells purified from regenerating *Pax7* $^{+/-}$ muscle (4 d post-ctx).

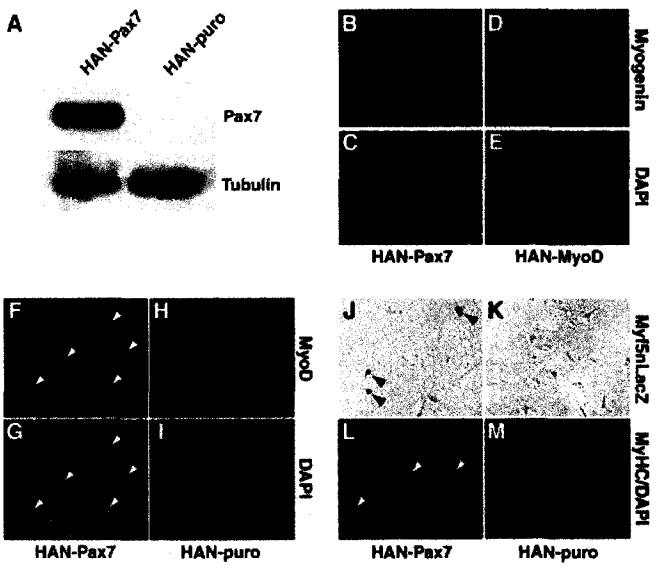


Figure 2. Pax7 Induces Myogenic Commitment in $CD45^+$:$Sca1^+$ Cells

(A) Western blot analysis with anti-Pax7 antibody confirmed high levels of ectopic Pax7 in C3H10T1/2 cells infected with retrovirus-Pax7 (HAN-Pax7) but not with control virus expressing a puromycin-resistance marker (HAN-puro).
(B and C) HAN-Pax7 did not induce expression of myogenin in C3H10T1/2 cells.
(D and E) By contrast, MyoD virus (HAN-MyoD) efficiently converted C3H10T1/2 cells to myogenin-expressing myocytes (green) .
(F–I) HAN-Pax7 (F and G) but not HAN-puro (H and I) activated expression of MyoD (red) in $CD45^+$:$Sca1^+$ cells from uninjured muscle.
(J–M) HAN-Pax7 (J) but not HAN-puro (K) also induced Myf5nLacZ expression in $CD45^+$:$Sca1^+$ cells. Furthermore, HAN-Pax7-infected $CD45^+$:$Sca1^+$ cultures differentiated into MyHC-expressing myocytes (green) under differentiation conditions (L), whereas HAN-puro-infected cells did not undergo myogenic differentiation (M). DAPI staining (blue) was used to visualize all nuclei.
DOI: 10.1371/journal.pbio.0020130.g002

By contrast, *Pax7* $^{-/-}$ $CD45^+$:$Sca1^+$ cells from regenerating muscle did not give rise to any MyoD-expressing (Figure 1H and 1I) or Desmin-expressing (Figure 1J and 1K) myogenic cells ($n = 3$ independent isolations with three mice per experiment). Taken together, these results support a central role for Pax7 in the myogenic specification of $CD45^+$:$Sca1^+$ cells in response to acute muscle damage.

Pax7 Is Sufficient to Induce Myogenesis in $CD45^+$:$Sca1^+$ Cells

Adenoviral and retroviral expression systems were developed to ectopically introduce the Pax7 gene into putative adult stem cell populations. Pax7 was efficiently expressed from retrovirus (HAN-Pax7) in C3H10T1/2 fibroblasts and other cell cultures (Figure 2A). Stable expression of Pax7 did not induce MyoD (data not shown) or Myogenin protein expression (Figure 2B and 2C) in C3H10T1/2 cells. MyoD, as expected, readily converted C3H10T1/2 cells into skeletal myocytes (Figure 2D and 2E). These results show that Pax7 is not sufficient to induce myogenic determination in an established multipotent mesenchymal cell.

To determine whether Pax7 expression was sufficient to activate myogenesis in adult $CD45^+$ progenitors, cells were fractionated from uninjured muscle and infected with Pax7-expressing retrovirus. Strikingly, $CD45^+$:$Sca1^+$ cells expressed Myf5 (data not shown) and MyoD (Figure 2F–2I) protein after infection only with Pax7 (HAN-Pax7), and not with puromy-

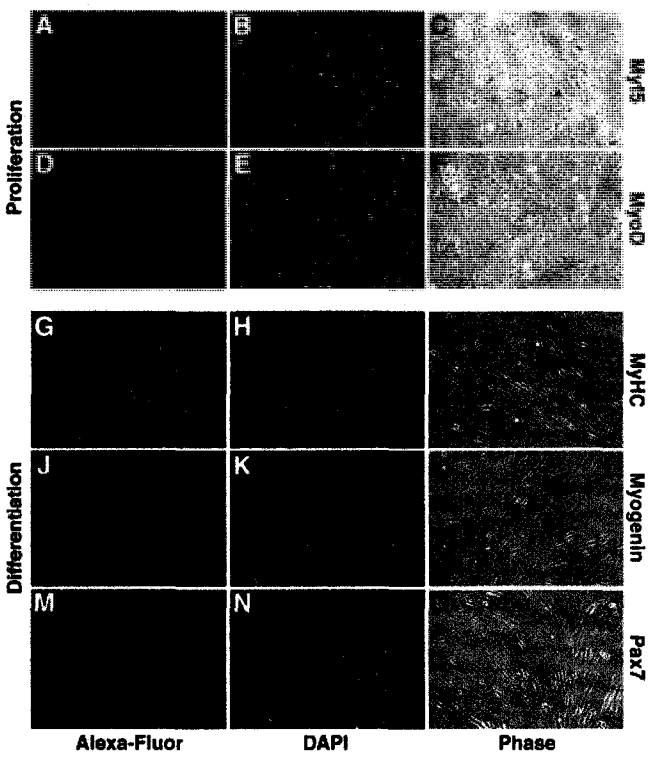


Figure 3. CDSC-Pax7 Cells Become Myogenic Progenitors

Myf5 (A–C) and MyoD (D–F) protein (green) are expressed in proliferating CDSC-Pax7 cells. Exposure of $CD45^+$:$Sca1^+$ cultures to low mitogen medium induced the formation of multinucleated myotubes and expression of myogenic differentiation markers including MyHC (red) (G–I) and myogenin (red) (J–L). Sustained expression of Pax7 (red) (M–O) in these cultures did not interfere with their differentiation. DAPI staining (blue) was used to visualize all nuclei.

DOI: 10.1371/journal.pbio.0020130.g003

cin-alone control virus (HAN-puro), indicating that these cells undergo myogenesis in response to Pax7

Infection of $CD45^+$:$Sca1^+$ cells from *Myf5nLacZ* reporter mice with HAN-Pax7 retrovirus specifically induced Myf5n-LacZ expression and myogenesis in about 50% of infected cells (Figure 2J). The *Myf5nLacZ* allele faithfully recapitulates the expression pattern of the endogenous *Myf5* gene and is rapidly induced following myogenic commitment (Tajbakhsh et al. 1996; Tajbakhsh et al. 1997). Importantly, infection of $CD45^+$:$Sca1^+$ cells with control retrovirus expressing the puromycin-resistance gene (HAN-puro) did not activate Myf5nLacZ expression (Figure 2K). Moreover, exposure of these cultures to differentiation conditions caused Pax7-expressing cells to differentiate into myotubes expressing myosin heavy chain (MyHC) (Figure 2L and 2M). Ectopic expression of Pax7 in $CD45^-$:$Sca1^+$ or $CD45^+$:$Sca1^-$ cells did not result in the generation of myogenic cells. Taken together, these results demonstrate that Pax7 induces the myogenic program selectively in $CD45^+$:$Sca1^+$ adult stem cells from skeletal muscle.

Expression of Pax7 Converted $CD45^+$:$Sca1^+$ Cells into Myogenic Progenitors

$CD45^+$:$Sca1^+$ cells expressing retroviral Pax7 were stably selected using puromycin, hereafter called CDSC-Pax7 cells (n = 4 independent isolates analyzed). CDSC-Pax7 cells displayed a stellate, fibroblastic morphology that was distinct from the round, refractile appearance of primary satellite-cell-derived myoblasts. Proliferating CDSC-Pax7 cells expressed the myogenic determination bHLH factors, Myf5 (Figure 3A–3C), and MyoD (Figure 3D–3F). CDSC-Pax7 cells cycled approximately three times faster than satellite-cell-derived myoblasts isolated simultaneously (data not shown) and maintained their myogenic identity as primary cultures in excess of three months. CDSC-Pax7 cultures also differentiated efficiently into multinucleated myotubes expressing the terminal differentiation markers MyHC (Figure 3G–3I) and myogenin (Figure 3J–3L). These results demonstrate that the constitutive expression of Pax7 (Figure 3M–3O), which is normally downregulated during differentiation (Seale et al. 2000), did not interfere with cell-cycle arrest and normal myotube formation. By contrast, overexpression of Pax7 in C2C12 myoblasts prevented their differentiation into MyHC-positive myocytes (data not shown). These experiments therefore demonstrate that myoblasts derived from Pax7-infected $CD45^+$:$Sca1^+$ stem cells are amenable to ex vivo expansion and subsequent terminal muscle differentiation.

CDSC-Pax7 Cells Express High Levels of Myf5 and Sca1

The expression pattern of myogenic factors in proliferating and differentiating CDSC-Pax7 cell lines was analyzed by Western blot (n = 2). These experiments indicated that Myf5 was expressed at high levels in proliferating CDSC-Pax7 cells (Figure 4A; day 0). Moreover, CDSC-Pax7 cells continued to express Myf5 protein during their differentiation. CDSC-Pax7 cells also expressed MyoD but at low levels relative to primary myoblasts. MyoD was transiently upregulated in CDSC-Pax7 cells as they entered their differentiation program (Figure 4A; days 1 and 2).

The primary myogenic regulatory factor (MRF) expression profile in CDSC-Pax7 cells contrasted with the pattern observed in satellite-cell-derived primary myoblasts (Figure 4; Wt-Mb). Primary myoblasts expressed higher levels of MyoD and lower levels of Myf5 and downregulated Myf5 immediately upon differentiation (Wt-diff). Myogenin (Myg) was upregulated during the differentiation of CDSC-Pax7 cells, albeit at lower levels compared with differentiating satellite-cell-derived myoblasts (Wt-diff). Interestingly, CDSC-Pax7 cells also expressed endogenous *Pax7* mRNA as demonstrated by reverse transcriptase PCR (RT-PCR) using primers that amplify a sequence within the *Pax7* 3′ UTR that is not present in the viral-Pax7 vector (Figure 4B). This result suggests that autoregulatory mechanisms may control *Pax7* gene expression. Taken together, these analyses demonstrate that CDSC-Pax7 cells and primary satellite-cell-derived myoblasts express different levels of MyoD and Myf5 but are similar in their ability to undergo terminal differentiation.

CDSC-Pax7 cells were originally derived from cells expressing cell surface CD45 and Sca1 proteins. Flow cytometry was employed to determine whether expression of these markers was maintained in vitro. Notably, CDSC-Pax7 cells continued to express high levels of Sca1 (approximately 90% of cells showed intense staining), but CD45 expression was extinguished (Figure 4C). Interestingly, approximately 24% of primary satellite-cell-derived myoblasts displayed low levels of Sca1 staining. Sca1 levels were not increased in satellite-cell-derived myoblasts overexpressing Pax7, demonstrating that CDSC-Pax7 cells did not arise from a small

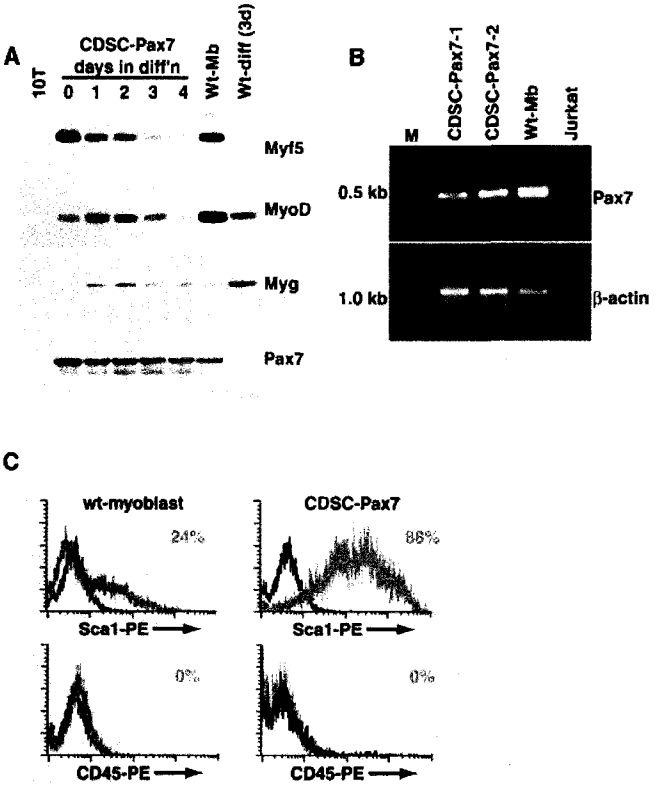


Figure 4. CDSC-Pax7 Cells Express High Levels of Myf5 and Sca1

(A) Western blot analysis of CDSC-Pax7 cells in proliferation conditions (day 0) and during differentiation (days 1–4) revealed high levels of Myf5 expression and low levels of MyoD expression. By contrast, satellite-cell-derived myoblasts (Wt-Mb) displayed the opposite profile of Myf5 and MyoD expression. Myogenin (Myg) was upregulated during the differentiation of CDSC-Pax7 and satellite-cell-derived myoblasts (Wt-diff). Note the sustained expression of Pax7 during the differentiation of CDSC-Pax7 cells. C3H10T1/2 (10T) lysate was used as a negative control.

(B) RT-PCR analysis indicated that CDSC-Pax7 cells (two different lines) upregulated the endogenous *Pax7* mRNA. Satellite-cell-derived myoblasts (Wt-Mb) and Jurkat cells were used as positive and negative controls, respectively.

(C) Flow cytometry indicated that CDSC-Pax7 cells lost expression of CD45 but retained high levels of Sca1. About 24% of satellite-cell-derived myoblasts (wt-myoblasts) expressed low levels of Sca1. (Black graph depicts staining with IgG-PE control antibody; red graph shows target staining using Sca1-PE or CD45-PE.)

DOI: 10.1371/journal.pbio.0020130.g004

number of committed myoblasts fractionated with $CD45^+$:$Sca1^+$ cells (data not shown).

CDSC-Pax7 Cells Differentiate In Vivo

To establish whether CDSC-Pax7 cells were capable of integrating and differentiating as myofibers in vivo, intramuscular transplantation studies were performed in dystrophic (*dystrophin*-deficient) muscle. Specifically, 1×10^5 CDSC-Pax7 cells were injected into the tibialis anterior (TA) muscle of 4- to 6-week-old *mdx:nude* mice. *Mdx* mice carry a point mutation in the *dystophin* gene and are a mouse model of Duchenne muscular dystrophy (Bulfield et al. 1984; Sicinski et al. 1989; Blaveri et al. 1999). As expected, dystrophin was localized at the myofiber sarcolemma in wild-type muscle (Figure 5A) and was absent in *mdx:nude* skeletal muscle (Figure 5B). Two months after transplantation, TA muscles were processed for immunohistochemical detection of dystrophin

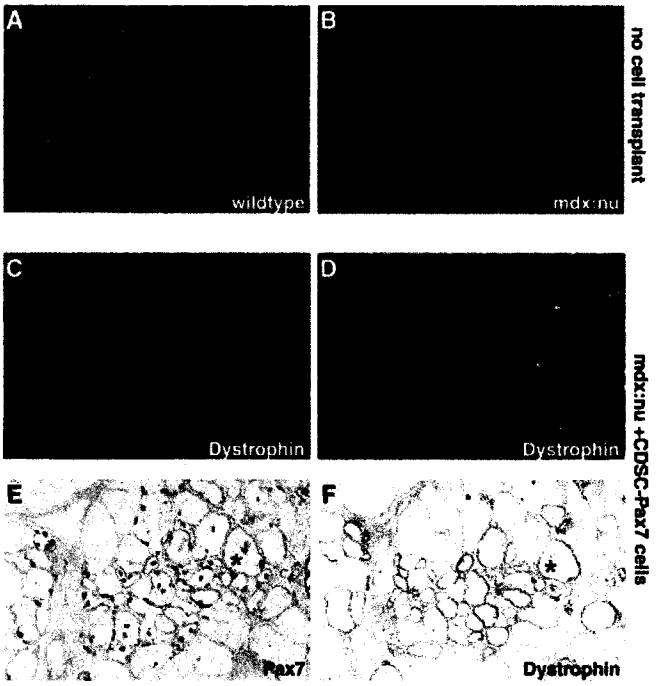


Figure 5. CDSC-Pax7 Cells Efficiently Contribute to the Repair of Dystrophic Muscle

(A) Wild-type muscle expressed dystrophin at the plasmalemma of all myofibers.

(B) Dystrophin protein was not detected in muscle sections from *dystrophin*-deficient *mdx:nude* mice *(mdx:nu)*.

(C–F) CDSC-Pax7 cells differentiated in vivo after transplantation, readily forming large numbers of dystrophin-expressing myofibers (green) in *mdx:nude* muscle (C and D). Serial cross sections showing the viral expression of Pax7 protein in central nuclei of regenerated fibers (red staining in [E]) confirmed the donor origin of dystrophin-positive myofibers (red staining in [F]).

DOI: 10.1371/journal.pbio.0020130.g005

and Pax7. These experiments revealed that CDSC-Pax7 cells differentiated in vivo, readily forming dystrophin-expressing myofibers in the *dystrophin*-deficient recipient muscle (Figure 5C and 5D). Endogenous Pax7 protein expression was not observed in central nuclei within differentiated wild-type myofibers (data not shown). Therefore, the expression of Pax7 protein (from retrovirus) in central nuclei within $dystrophin^+$ fibers established the contribution of CDSC-Pax7 donor cells to recipient muscles (Figure 5E and 5F). These results thus document the capacity for CDSC-Pax7 cells to differentiate in vivo and contribute to the repair of dystrophic muscle.

Pax7 Does Not Induce Myogenesis in $CD45^+$:$Sca1^+$ Cells from $Pax7^{-/-}$ Muscle

The myogenic differentiation of wild-type $CD45^+$:$Sca1^+$ muscle cells suggested that ectopic Pax7 would induce myogenesis in this cell population from $Pax7^{-/-}$ muscle. Infection of $Pax7^{-/-}$ $CD45^+$:$Sca1^+$ cells with Pax7 retrovirus resulted in high levels of retroviral Pax7 transcript but no expression of *Myf5* mRNA by Northern blot hybridization (Figure 6A) or RT-PCR (data not shown). The absence of Myf5 (Figure 6B–6D) or MyoD (data not shown) expression, determined by immunochemical staining of Pax7-transduced cells, ruled out the possibility that a minor subpopulation of $CD45^+$:$Sca1^+$ cells underwent myogenesis. These experiments

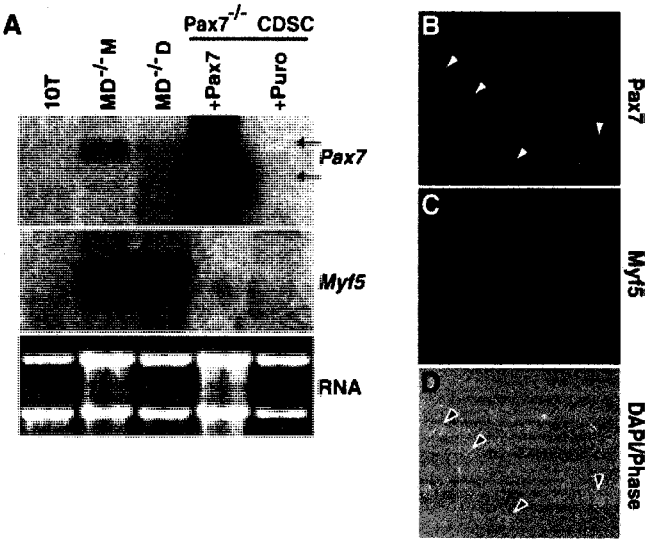


Figure 6. Pax7 Does Not Induce Myogenesis in CD45$^+$:Sca1$^+$ Cells from *Pax7*$^{-/-}$ Muscle

(A) Northern analysis shows that *MyoD*$^{-/-}$ satellite-cell-derived myoblasts (*MD*$^{-/-}$ M) and differentiating cells (*MD*$^{-/-}$ D) express endogenous *Pax7* (upper arrow, Pax7 blot) and *Myf5* transcripts. *Pax7*$^{-/-}$ CD45$^+$:Sca1$^+$ cells (CDSC) transduced with HAN-Pax7 (+Pax7) or HAN-puro (+puro) did not initiate expression of *Myf5* mRNA. The retroviral transcript producing Pax7 (lower arrow) is smaller than the endogenous *Pax7* mRNA (e.g., lower arrow).

(B–D) Ectopic expression of Pax7 (red) (B) in *Pax7*$^{-/-}$ CDSC cells did not induce Myf5 protein expression (C). DAPI staining (blue) was used to visualize nuclei (D).

DOI: 10.1371/journal.pbio.0020130.g006

illustrate that *Pax7*$^{-/-}$ CD45$^+$:Sca1$^+$ cells do not enter the myogenic lineage in response to Pax7, suggesting that intrinsic differences exist between wild-type and *Pax7*-deficient populations of CD45$^+$:Sca1$^+$ cells.

Pax7 Promotes Myogenic Commitment in *Pax7*-Deficient CD45$^-$:Sca1$^-$ Cells

In cell suspensions from uninjured muscle, satellite cells and their daughter myogenic precursors are uniformly CD45$^-$ and Sca1$^-$ (Polesskaya et al. 2003). In *Pax7*$^{-/-}$ mice, the extremely rare myogenic cells in muscle tissue do not express CD45 or Sca1, and do not survive or expand in a variety of culture conditions (S.B.P. Chargé, P. Seale, and M.A. Rudnicki, unpublished data). Interestingly, ectopic expression of Pax7 in CD45$^-$:Sca1$^-$ cells isolated from *Pax7*$^{-/-}$ muscle resulted in the expression of Myf5 protein in more than 50% of infected cells (n = 3) (Figure 7A–7C). Analysis of HAN-puro-infected control cultures did not reveal any myogenic cells (Figure 7D–7F). Importantly, all Myf5-expressing myoblasts (Figure 7G–7I) and MyHC-expressing differentiated myotubes (Figure 7J–7L) in Pax7-infected CD45$^-$:Sca1$^-$ cultures expressed viral Pax7.

In these experiments we cannot formally exclude the possibility that Pax7 promoted the survival and proliferation of committed myoblasts. However, given the extremely low number of myogenic cells recovered in culture (less than 0.7%), the low efficiency of primary myoblast infection (approximately 5%–10%), and the absence of any Myf5- or MyoD-expressing cells in control HAN-puro cultures, our results strongly suggest that Pax7 induces myogenic specification in a nonmyoblast, CD45- and Sca1-negative cell.

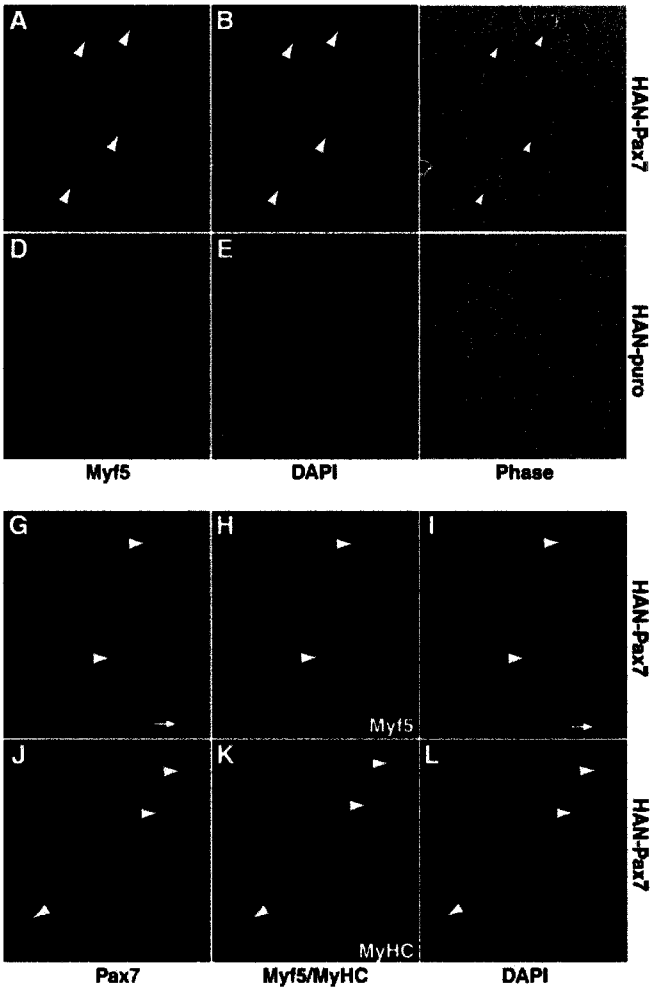


Figure 7. Pax7 Promotes Myogenesis in CD45$^-$:Sca1$^-$ Cells from *Pax7*$^{-/-}$ Muscle

(A–C) Ectopic expression of Pax7 (HAN-Pax7) induced Myf5 expression (green) and myogenic commitment in CD45$^-$:Sca1$^-$ cells from *Pax7*$^{-/-}$ muscle.

(D–F) By contrast, Myf5-expressing cells were completely absent from HAN-puro-infected cultures after selection.

(G–L) CD45$^-$:Sca1$^-$ cells from *Pax7*$^{-/-}$ muscle expressed Myf5 (red) (H) and MyHC (red) (K) only in cells that also coexpressed high levels of Pax7 protein (G and J). Arrowheads indicate cells coexpressing Pax7 and Myf5/MyHC. Arrow in (G) and (I) depicts a Pax7$^+$, Myf5$^-$ cell.

DOI: 10.1371/journal.pbio.0020130.g007

Adenoviral Expression of Pax7 Enhances Regeneration in *Pax7*-Deficient Muscle

To investigate whether Pax7 was sufficient to stimulate myogenesis in vivo, adenovirus was used to ectopically express Pax7 in damaged *Pax7*$^{-/-}$ muscle. Adenoviral particles (1×10^8) expressing either Pax7 (Ad-Pax7) or the bacterial β-*galactosidase* gene *(LacZ)* (Ad-LacZ) were injected directly into injured TA muscles of 4- to 6-week-old *Pax7*$^{-/-}$ animals 2 d after administration of ctx (n = 3). Immunohistochemistry for Pax7 in adenovirus-infected muscles demonstrated widespread Pax7 expression primarily in mononuclear cells within the damaged tissue (data not shown).

To assess the effect of Pax7 expression in damaged tissue, TA muscles were analyzed and scored for regeneration 12 d

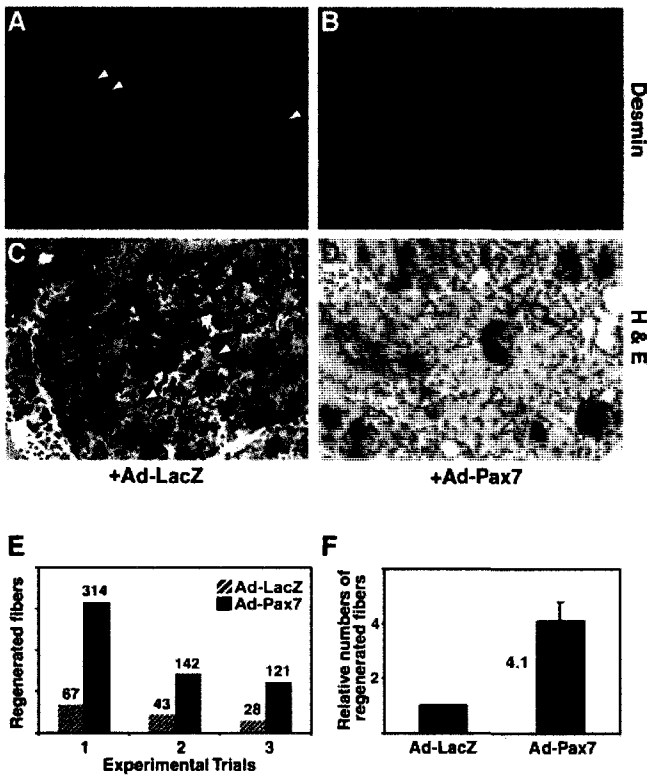


Figure 8. Adenovirus-Pax7 Significantly Improves Regeneration In Vivo

(A and B) Infection of ctx-damaged *Pax7* $^{-/-}$ muscles with Ad-Pax7 resulted in markedly improved muscle integrity and a significantly increased number of Desmin immunoreactive (green) regenerated fibers (B) relative to muscles treated with Ad-LacZ (A).

(C and D) Hematoxylin and Eosin staining similarly showed an increased number of centrally nucleated fibers in Ad-Pax7-treated *Pax7* $^{-/-}$ muscles.

(E) In three separate experimental trials, the number of regenerated fibers was markedly increased in Ad-Pax7-treated muscles relative to Ad-LacZ; however, the response was biologically variable between groups. On average, Ad-Pax7 infection resulted in a 4.1 ± 0.72–fold increase in regenerated *Pax7* $^{-/-}$ myofibers (F).

DOI: 10.1371/journal.pbio.0020130.g008

after infection by enumerating the number of regenerated fibers with centrally located nuclei. The newly regenerated status of centrally nucleated fibers was confirmed by Desmin and embryonic MyHC immunoreactivity. Ad-Pax7 induced a markedly enhanced regenerative response relative to Ad-LacZ in *Pax7* $^{-/-}$ muscle as evidenced by the increased number of Desmin-positive (Figure 8A and 8B) and centrally nucleated fibers (Figure 8C and 8D).

Wild-type TA muscles typically contained in excess of 700 regenerated fibers 14 d after injury (data not shown). In three independent experiments, ctx-damaged TA muscle from *Pax7* $^{-/-}$ mice typically contained an average of 46 surviving or regenerated fibers following regeneration (Figure 8E). By contrast, infection of regenerating *Pax7* $^{-/-}$ TA with Ad-Pax7 resulted in the generation of an average of 192 myofibers (Figure 8E). Therefore, Pax7-infected tissue contained a 4.1 ± 0.72–fold increase in the number of regenerated fibers (Figure 8F). Together, these results demonstrate the ability of virally transduced Pax7 to direct the de novo generation of myogenic progenitors capable of forming new myofibers and participating in regenerative myogenesis.

Discussion

In this article, we demonstrate that expression of Pax7 induces the myogenic specification of CD45$^+$ muscle-derived adult stem cells. First, CD45$^+$:Sca1$^+$ cells isolated from regenerating *Pax7* $^{-/-}$ muscle were incapable of undergoing myogenic specification (see Figure 1). Second, expression of Pax7 with viral vectors in CD45$^+$:Sca1$^+$ cells purified from uninjured muscle promoted the formation of highly proliferative myoblasts that readily differentiated as multinucleated myotubes (see Figures 2 and 3). CD45$^+$:Sca1$^+$ cells engineered to express Pax7 (CDSC-Pax7) also differentiated in vivo, readily contributing to the regeneration of dystrophic muscle (see Figure 5). Lastly, Ad-Pax7 gene delivery into chemically damaged *Pax7* $^{-/-}$ muscle resulted in the efficient de novo generation of myofibers in the absence of endogenous satellite cells. Taken together, these data unequivocally establish that Pax7 plays a key regulatory role for directing myogenic specification in some populations of adult stem cells during regenerative myogenesis. Moreover, these results emphasize the possibility of designing strategies to upregulate or ectopically express Pax7 in stem cells for the treatment of muscle degenerative diseases.

The presence of adult stem cell populations distinct from satellite cells resident in skeletal muscle tissue has been well documented (Gussoni et al. 1999; Jackson et al. 1999; Torrente et al. 2001; Asakura et al. 2002; McKinney-Freeman et al. 2002; Qu-Petersen et al. 2002; Cao et al. 2003). An understanding of the developmental origin of these various cell populations and their physiological relevance in the maintenance of tissue integrity is beginning to emerge. Several lines of evidence argue that skeletal muscle regeneration is normally mediated entirely by stem cells resident in muscle tissue. First, destruction of stem cells resident in muscle with high-dose local irradiation of limbs results in a long-term deficit in muscle growth and regeneration (Wakeford et al. 1991; Pagel and Partridge 1999; Heslop et al. 2000). Second, transplanted muscles do not incorporate host nuclei after injury and regeneration (Schultz et al. 1986). Together, those experiments argue that CD45$^+$ stem cells from marrow do not normally transit in significant numbers through the circulation to sites of muscle damage. Our experiments, however, suggest that a population of specialized CD45$^+$ cells resides in muscle and can efficiently form myogenic progenitors in response to Wnt signaling (Polesskaya et al. 2003). In the current work we demonstrate that induction of Pax7 is required for the myogenic specification of CD45$^+$ stem cells and that retroviral transduction can dominantly induce the myogenic specification of these cells. These observations therefore provide compelling evidence that some adult stem cells participate in regenerative myogenesis by forming myogenic progenitors following Pax7 induction in response to Wnt signaling. These data additionally suggest the hypothesis that Pax7 is a transcriptional target of the β-Catenin complex in Wnt-stimulated adult stem cells.

Interesting parallels exist between the inductive mechanisms and transcriptional networks in embryonic and regenerative myogenesis (Parker et al. 2003). For example, the Pax7-dependent myogenic specification of CD45$^+$ adult stem cells appears analogous to the Pax3-dependent induction of muscle precursors during somitogenesis. In the early embryo, Pax3 is expressed in the presomitic mesoderm and

immature epithelial somites prior to the onset of muscle-specific gene expression (Goulding et al. 1994; Williams and Ordahl 1994). Moreover, Pax3 functions upstream of MyoD in the formation of trunk and body-wall muscle (Tajbakhsh et al. 1997). Consistent with a direct role for Pax3 in myogenic induction, ectopic Pax3 activates MyoD expression in embryonic tissues (Maroto et al. 1997; Bendall et al. 1999; Heanue et al. 1999). However, Pax3 also regulates cell survival in the presomitic mesoderm in areas that do not express Pax7, suggesting an indirect mechanism by which Pax3 may act genetically upstream of MyoD (Borycki et al. 1999). Our experiments do not rule out the possibility that Pax7 promotes the survival of $CD45^+$ progenitors that are already competent to give rise to myogenic cells. Characterization of the downstream targets of Pax7 in $CD45^+$ cells will be required to directly address this issue.

In explanted embryonic tissues, signals from the floor plate and neural tube are required for induction of the MRFs (Munsterberg and Lassar 1995; Pourquie et al. 1995, 1996; Cossu et al. 1996). In particular, Wnt7a activates expression of MyoD in explanted paraxial mesoderm from 10.5-d-old mouse embryos (Tajbakhsh et al. 1998). The requirement for Pax3 expression in somitic precursors prior to the onset of MyoD expression suggests that Wnt signals may activate Pax3 and indirectly promote MRF expression (Borycki et al. 1999). An analogous requirement for Pax7 in the myogenic commitment of adult $CD45^+$ progenitors suggests a conserved hierarchy whereby Wnt signaling activates Pax3 or Pax7 expression upstream of the MRFs in somitic and adult muscle stem cells, respectively. This notion is supported by the observed loss of Pax3 expression in P19 mesodermal precursors engineered to express a dominant negative form of the Wnt effector protein, β-Catenin (Petropoulos and Skerjanc 2002).

A confounding result from our study was the inability of Pax7 to induce myogenesis in $CD45^+$:$Sca1^+$ cells recovered from *Pax7* $^{-/-}$ muscle (see Figure 6). Several possible explanations may account for this observation. First, $CD45^+$:$Sca1^+$ muscle cells represent a heterogeneous cell population, as evidenced by their nonuniform response to stimuli such as myoblast coculture, Wnt proteins, and ectopic expression of Pax7 (results herein and Polesskaya et al. 2003). Analysis of muscle suspensions from young *Pax7*$^{-/-}$ mice revealed a significantly increased number of hematopoietic progenitors and adipogenic cells (Seale et al. 2000). We also observed altered proportions of CD45- and Sca1-expressing cells in uninjured and regenerating muscle (see Figure 1A). The putative stem cell subfraction coexpressing CD45 and Sca1 may have been exhausted prematurely during postnatal *Pax7* $^{-/-}$ muscle development. It is also conceivable that a reduced proportion of stem cells in the *Pax7* $^{-/-}$ $CD45^+$:$Sca1^+$ muscle fractions was not detected in our assay due to a low efficiency of retroviral transduction (approximately 10% of surviving $CD45^+$:$Sca1^+$ cells with GFP virus). The identification of additional markers expressed by adult muscle-derived stem cells is required to more thoroughly explore these issues.

Alternatively, adult stem cells may require additional signals to undergo myogenesis in response to Pax7. The profound growth deficit in *Pax7* $^{-/-}$ muscles is likely to invoke nonspecific and indirect changes to the muscle microenvironment (Seale et al. 2000). Specific cues required to "prime" or activate adult stem cells may thus be absent or ineffective in *Pax7* $^{-/-}$ muscle. Finally, our experiments also revealed that the endogenous *Pax7* gene is upregulated during the myogenic specification of $CD45^+$:$Sca1^+$ cells (Figure 4B). Therefore, endogenous gene activity, possibly through the regulated expression of different isoforms (Kay et al. 1995; Ziman et al. 1997), may be essential to the stability of myogenic commitment. Future experiments addressing the functional differences between $CD45^+$:$Sca1^+$ cells in wild-type and *Pax7*-deficient muscle will provide a unique opportunity to gain a more complete understanding of the role of these cells during postnatal muscle development.

Although $CD45^+$ cells from *Pax7* $^{-/-}$ muscle were apparently unable to undergo myogenesis, ectopic Pax7 induced expression of Myf5 and myogenic specification in *Pax7*-deficient $CD45^-$:$Sca1^-$ cells (see Figure 7). Moreover, Ad-Pax7 significantly increased the in vivo regenerative capacity of *Pax7* $^{-/-}$ muscle (see Figure 8). Skeletal muscle in adult *Pax7* $^{-/-}$ mice displays a profound regeneration deficit with only occasional regenerated fibers observed at the site of injury 30 d after ctx injection (S.B.P. Chargé, P. Seale, and M.A. Rudnicki, unpublished data). Taken together, these results imply the presence of *Pax7* $^{-/-}$ muscle progenitors that require the activity of Pax7 to generate sufficient numbers of myoblasts for effective regeneration. Further studies will be required to molecularly characterize the responsive cells and their developmental relationship to other muscle stem cell populations.

The dominant expression of Myf5 in Pax7-infected $CD45^+$:$Sca1^+$ cells (CDSC-Pax7) (see Figure 4A) suggests a paradigm wherein Pax7 preferentially activates Myf5 compared to MyoD. Interestingly, Pax3 has been implicated in myogenesis specifically upstream of MyoD (Tajbakhsh et al. 1997). Taken together, these observations suggest the hypothesis that Pax3 and Pax7 specify distinct myogenic lineages through the preferential activation of MyoD and Myf5, respectively.

Several experimental observations have noted a role for Myf5 in promoting myoblast proliferation. For example homozygous *Myf5nLacZ*, (e.g., *Myf5*-deficient) embryos display significantly reduced numbers of LacZ-expressing myogenic progenitors (Tajbakhsh et al. 1996). In avian embryos, Myf5 is preferentially expressed in proliferating myoblasts, whereas MyoD appears to be upregulated in differentiating cells (Delfini et al. 2000). Furthermore, *Myf5* $^{-/-}$ satellite-cell-derived myoblasts display a profound proliferation deficit (Montarras et al. 2000). The increased growth rate of CDSC-Pax7 cells is reminiscent of *MyoD*$^{-/-}$ myoblasts that also express elevated levels of Myf5 (Sabourin et al. 1999). These observations raise the possibility that Pax7 activates expression of Myf5 to promote adult myoblast expansion whereas Pax3 preferentially induces MyoD and differentiation.

The requirement for Pax7 in the specification of muscle satellite cells (Seale et al. 2000) and its induction during the myogenic recruitment of $CD45^+$ adult stem cells provide further evidence for a developmental relationship between $CD45^+$ adult muscle stem cells and satellite cells. Together, our experiments suggest the hypothesis that $CD45^+$:$Sca1^+$ cells give rise to satellite cells by a Pax7-dependent mechanism in response to Wnt signals. In conclusion, our work establishes that Pax7 is necessary and sufficient for the myogenic specification of specific populations of adult stem cells resident in muscle tissue. The proliferative myogenic

character of CDSC-Pax7 cells and their efficient engraftment into dystrophic muscle further argue that methods to deliver Pax7 or upregulate its expression in stem cells may be useful in treating degenerative muscle disease.

Materials and Methods

Mice. Mice carrying a targeted null mutation in *Pax7* (hereafter referred to as *Pax7* $^{-/-}$) were generously provided by Drs. A. Mansouri and P. Gruss (Mansouri et al. 1996) and outbred into the SV129 background to increase survival. *Myf5nLacZ* mice were provided by Dr. S. Tajbakhsh (Tajbakhsh et al. 1996). *Mdx* mice were obtained from Jackson Laboratory (Bar Harbor, Maine, United States). *Mdx:nu* mice were provided by Dr. T.A. Partridge (see Blaveri et al. 1999).

Cell sorting. Mononuclear cells were recovered from uninjured hindlimb muscles or from ctx-damaged TA muscles of *Pax7* $^{+/+}$, *Pax7* $^{+/-}$, and *Pax7* $^{-/-}$ mice as described previously (Megeney et al. 1996). Cells were washed twice with ice-cold DMEM supplemented with 5% FBS, passed through 30-μm filters (Miltenyi Biotec, Bergisch Gladbach, Germany) and suspended at a concentration of 2–3 × 10^6 cells/ml. Staining was performed for 30 min on ice using the antibodies CD45-APC (30-F11), CD45.2-FITC (104), Sca1-PE, or FITC (D7), all from BD Biosciences Pharmingen (San Diego, California, United States) and CD45-TC (30-F11) from Caltag Laboratories (Burlingame, California, United States). Primary antibodies were diluted in cell suspensions at 1:200. After two washes with cold PBS supplemented with 2% FBS, cells were separated on a MoFlo cytometer (DakoCytomation, Glostrup, Denmark). Sort gates were strictly defined based on isotype control stained cells and single antibody staining. Dead cells and debris were excluded by gating on forward and side scatter profiles. Sorting was performed using single cell mode to achieve the highest possible purity. The purity of sorted populations was routinely greater than 98%.

Retroviral and adenoviral gene expression. Retrovirus was produced according to the 3-plasmid HIT system with plasmids pHIT60, pHIT456, and pHAN-puro as described elsewhere (Soneoka et al. 1995). pHIT60 encodes the MLV retroviral gag-pol, pHIT456 expresses an amphotrophic envelope protein, and pHAN-puro is an expression vector with a hybrid CMV-5′ LTR promoter driving production of the retroviral transcript. Pax7 expression vectors were generated using the mouse Pax7d isoform containing a single Ala→Thr substitution at the seventh amino acid (the Thr residue is conserved in human, chicken, and zebrafish Pax7 proteins). Pax7d or mouse MyoD are translated from the full retroviral transcript, whereas the puromycin-resistance marker is expressed following integration from a shorter transcript produced by the SV40 early promoter located 3′ to the multiple cloning site. Transient cotransfection of all three plasmids into 293FT cells (Invitrogen, Carlsbad, California, United States) by the calcium phosphate method (Graham and van der Eb 1973) routinely produced viral titres between 10^6 and 10^7 cfu per ml. pHAN-puro was used to produce puromycin-resistant virus for controls.

Purified CD45$^+$:Sca1$^+$ or CD45$^-$:Sca1$^-$ cells were spun down, counted, and 20,00–50,000 cells were then cultured overnight on collagen-coated 4-well chamber slides in HAM's F10 medium (Invitrogen) supplemented with 20% FBS, antibiotics, and 10 ng/ml Stem Cell Factor (R & D Systems, Minneapolis, Minnesota, United States). The following day, cells were incubated for 6 h with retrovirus at a 1:1 ratio (complete medium: retrovirus supernatant) with 8 μg/ml polybrene (hexadimethrine bromide; Sigma, St. Louis, Missouri, United States). After infection, cells were rinsed twice with PBS, and all cells were replated in myoblast growth medium. After 48 h, infected pools were selected in 1 μg/ml puromycin (Sigma) to establish stable CDSC-Pax7 lines. C3H10T1/2 cells were incubated overnight with MyoD, Pax7, or puro virus and 8 μg/ml polybrene.

Adenovirus (type V) was prepared using the Ad-Max adenovirus creation kit (Microbix Biosystems, Toronto, Ontario, Canada). Ad-Pax7d (cDNA as described above) and Ad-LacZ were expressed from the murine CMV promoter. Adenovirus was purified in CsCl gradients by centrifugation, dialyzed against sterile PBS, and frozen down in 15% glycerol at −80 °C. Titres of purified adenovirus were determined by plaque assays on 293 cells and were always above 10^{10} pfu/ml.

Western blot analyses. Cell cultures were lysed in RIPA extraction buffer (50mM Tris-HCl [pH 7.4], 1% Nonidet P-40, 0.5% NaDeoxycholate, 0.1% Sodium-dodecyl-sulphate, 5 mM EDTA, 150 mM NaCl, 50 mM NaF) supplemented with protease inhibitors (Complete; Roche, Basel, Switzerland). The extracts were normalized for protein content using Bio-Rad dye (Hercules, California, United States). Forty micrograms of lysate was separated by sodium-dodecyl-sulfate-polyacrylamide gel electrophoresis and transferred to PVDF filters (ImmobilonP; Millipore, Billerica, Massachusetts, United States). Filters were probed with antibodies to Pax7 (Developmental Studies Hybridoma Bank [DSHB], Iowa City, Iowa, United States); Myf5, 1:1000 (C-20, Santa Cruz Biotechnology, Santa Cruz, California, United States); MyoD, 1:1000 (C-20, Santa Cruz Biotechnology); myogenin (F5D, DSHB); and α-tubulin, 1:2000 (T 9026, Sigma). Secondary detection was performed with horseradish peroxidase-conjugated antibodies (Bio-Rad). Protein expression was visualized using the ECL Plus kit (Amersham Biosciences, Little Chalfont, United Kingdom).

Ctx-induced regeneration and in vivo adenovirus infections. Four- to six-week-old *Pax7* $^{-/-}$ and wild-type littermates were anesthetized with Halothane gas. Twenty-five microliters of 10 μM ctx (Latoxan, Valence, France) was injected into the midbelly of the TA muscle, using a 29½ G insulin syringe. Mice were sacrificed at 4 d or 2 wk after ctx injection. For adenovirus infections, 25 μl of sterile PBS containing 10^8 particles of purified Ad-Pax7 or Ad-LacZ was injected 2 d after ctx injection into damaged TA muscles with a 29½ G insulin syringe.

Cell transplantation. Primary CDSC-Pax7 cells cultured in myoblast conditions were trypsinized, washed twice with PBS, and suspended at 5 × 10^5 cells/25-μl in sterile PBS for cell transplantation. Cells were injected directly into the TA midbelly of 4- to 6-wk-old *mdx:nude* mice. Mice were sacrificed 2 mo after cell injections to analyze the myogenic contribution of transplanted cells.

Cell cultures. Primary satellite-cell-derived myoblasts were established from purified CD45$^-$:Sca1$^-$ fractions of hindlimb muscle of 4- to 6-wk-old *Pax7* $^{+/+}$ or *Pax7* $^{+/-}$ mice. Myoblasts and CDSC-Pax7 cells were maintained in HAM's F-10 medium (Invitrogen) supplemented with 20% FBS and 2.5 ng/ml bFGF (Invitrogen) on collagen-coated dishes. CDSC-Pax7 cells and primary satellite-cell-derived myoblasts were differentiated for 1–3 d in DMEM supplemented with 5% horse serum. C3H10T1/2 and HEK 293 cells were obtained from the ATCC (Manassas, Virginia, United States) and maintained in DMEM supplemented with 10% FBS.

Histology and immunocytochemistry. For analysis of regeneration and enumeration of regenerated myofibers, TA muscles were isolated, embedded in OCT (Tissue-Tek; Sakura Finetek, Torrance, California, United States)/20% sucrose and immediately frozen in liquid nitrogen. Ten-micrometer cryosections (cross sections) from the TA midbelly at the site of ctx injection were stained with Hematoxylin and Eosin. Central myonuclei in regenerating muscles were counted on at least two independent cross sections of the entire TA muscle per mouse analyzed. Fibers were further identified by immunostaining with antibodies specific to Desmin, 1:200 (D33, DakoCytomaton); dystrophin, 1:500 (Sigma); Pax7 (DSHB); or embryonic fast MyHC (F1.652, DSHB) followed by secondary detection with anti-mouse FITC conjugated antibody, 1:200 (Chemicon, Temecula, California, United States). Sections were analyzed on a Zeiss (Oberkochen, Germany) Axioplan 2 microscope.

Cultured cells were fixed with 4% paraformaldehyde, nonspecific antigens were blocked in 5% horse serum/PBS, and cells were reacted with primary antibodies as follows: Desmin, 1:200 (DakoCytomaton); MyoD, 1:200 (5.8A, BD Biosciences Pharmingen); all MyHC (MF-20, DSHB); Myf5, 1:1000 (C-20, Santa Cruz Biotechnology); Pax7 (DSHB); and myogenin (F5D, DSHB). Secondary detection was performed using fluorescein- or rhodamine-conjugated antibodies, 1:200 (Chemicon). *Myf5nLacZ* expression was detected by X-Gal reaction as described previously (Polesskaya et al. 2003).

RT-PCR and Northern blot analysis. Total RNA was extracted using RNeasy kits (Qiagen, Valencia, California, United States), according to manufacturer's instructions. RT-PCR analysis for endogenous *Pax7* mRNA was performed using the GeneAmp PCR Core kit (PerkinElmer, Wellesley, Massachusetts, United States). RT-PCR using 1 μg of total RNA was conducted as per manufacturer's instructions with the following modifications. cDNA synthesis was extended for 1 h at 42 °C, and 5 μl of the first-strand RT product was used for PCR amplification. PCR conditions for endogenous *Pax7* were 94 °C for 5 min; 35 cycles of 94 °C for 45 s, 56 °C for 45 s, 72 °C for 45 s; and finally 72 °C for 7 min. The PCR primers span intron 8 of the *Pax7* gene (Pax7-exon8-fwd 5′ gct acc agt aca gcc agt atg 3′ and Pax7-exon9-rev 5′ gtc act aag cat ggg tag atg 3′) and amplify sequence in the 3′-UTR of the gene that is not contained in the viral Pax7 expression cassette. RT-PCR products were analyzed by electrophoresis through a TAE-ethidium-agarose gel.

Northern blot studies were performed according to standard techniques using random-primed ^{32}P-dCTP radiolabeled cDNA

fragments as probes (Redi-prime, Amersham Biosciences)(Sabourin et al. 1999). Fifteen micrograms of total RNA from various cell cultures was electrophoresed in denaturing formaldehyde gels and transferred to Hybond-N filters (Amersham Biosciences).

Supporting Information

Accession Numbers

The accession numbers for the proteins discussed in this paper are Desmin (LocusLink ID 13346), mouse MyoD (GenBank NM_010866), MyoD (LocusLink ID 17927), Myogenin (LocusLink ID 17928), Pax7 (LocusLink ID 18509), and Pax7d isoform (GenBank AF_254422).

Acknowledgments

The authors would like to thank Dr. Ruedi Braun, Sylvain Gimmig, and Caroline Vergette for flow cytometry, Dr. Robin Parks for assistance with purification of adenovirus, Dr. Terry Partridge for providing *mdx:nude* mice; and Dr. V. Sartorelli for retroviral expression plasmids. PS is supported by a Doctoral Research Award from the Canadian Institutes of Health Research. MAR holds the Canada Research Chair in Molecular Genetics and is a Howard Hughes Medical Institute International Scholar. This work was supported by grants to MAR from the Muscular Dystrophy Association, the National Institutes of Health, the Canadian Institutes of Health Research, the Howard Hughes Medical Institute, and the Canada Research Chair Program.

Conflicts of interest. The authors have declared that no conflicts of interest exist.

Author contributions. PS and MAR conceived and designed the experiments. PS performed the experiments, with the exception of Figure 4B performed by JI. PS and MAR analyzed the data. PS, JI, and AS contributed reagents/materials/analysis tools. PS and MAR wrote the paper. ∎

References

Asakura A, Seale P, Girgis-Gabardo A, Rudnicki MA (2002) Myogenic specification of side population cells in skeletal muscle. J Cell Biol 159: 123–134.

Bendall AJ, Ding J, Hu G, Shen MM, Abate-Shen C (1999) Msx1 antagonizes the myogenic activity of Pax3 in migrating limb muscle precursors. Development 126: 4965–4976.

Bischoff R (1994) The satellite cell and muscle regeneration. In: Myogenesis, AG Engel, C Franszini-Armstrong, editors. New York: McGraw-Hill. pp. 97–118.

Blaveri K, Heslop L, Yu DS, Rosenblatt JD, Gross JG, et al. (1999) Patterns of repair of dystrophic mouse muscle: Studies on isolated fibers. Dev Dyn 216: 244–256.

Borycki AG, Li J, Jin F, Emerson CP, Epstein JA (1999) Pax3 functions in cell survival and in pax7 regulation. Development 126: 1665–1674.

Bulfield G, Siller WG, Wight PA, Moore KJ (1984) X chromosome-linked muscular dystrophy (mdx) in the mouse. Proc Natl Acad Sci U S A 81: 1189–1192.

Cao B, Zheng B, Jankowski RJ, Kimura S, Ikezawa M, et al. (2003) Muscle stem cells differentiate into haematopoietic lineages but retain myogenic potential. Nat Cell Biol 5: 640–646.

Charge SBP, Rudnicki MA (2004) Cellular and molecular regulation of muscle regeneration. Physiol Rev 84: 209–238.

Cossu G, Kelly R, Tajbakhsh S, Di Donna S, Vivarelli E, et al. (1996) Activation of different myogenic pathways: Myf-5 is induced by the neural tube and MyoD by the dorsal ectoderm in mouse paraxial mesoderm. Development 122: 429–437.

Delfini M, Hirsinger E, Pourquie O, Duprez D (2000) Delta 1-activated notch inhibits muscle differentiation without affecting Myf5 and Pax3 expression in chick limb myogenesis. Development 127: 5213–5224.

Goulding M, Lumsden A, Paquette AJ (1994) Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. Development 120: 957–971.

Graham FL, van der Eb AJ (1973) Transformation of rat cells by DNA of human adenovirus 5. Virology 54: 536–539.

Gussoni E, Soneoka Y, Strickland CD, Buzney EA, Khan MK, et al. (1999) Dystrophin expression in the mdx mouse restored by stem cell transplantation. Nature 401: 390–394.

Heanue TA, Reshef R, Davis RJ, Mardon G, Oliver G, et al. (1999) Synergistic regulation of vertebrate muscle development by Dach2, Eya2, and Six1, homologs of genes required for *Drosophila* eye formation. Genes Dev 13: 3231–3243.

Heslop L, Morgan JE, Partridge TA (2000) Evidence for a myogenic stem cell that is exhausted in dystrophic muscle. J Cell Sci 113: 2299–2308.

Jackson KA, Mi T, Goodell MA (1999) Hematopoietic potential of stem cells isolated from murine skeletal muscle. Proc Natl Acad Sci U S A 96: 14482–14486.

Kay PH, Mitchell CA, Akkari A, Papadimitriou JM (1995) Association of an unusual form of a Pax7-like gene with increased efficiency of skeletal muscle regeneration. Gene 163: 171–177.

Mansouri A, Stoykova A, Torres M, Gruss P (1996) Dysgenesis of cephalic neural crest derivatives in *Pax7*$^{-/-}$ mutant mice. Development 122: 831–838.

Maroto M, Reshef R, Munsterberg AE, Koester S, Goulding M, et al. (1997) Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. Cell 89: 139–148.

McKinney-Freeman SL, Jackson KA, Camargo FD, Ferrari G, Mavilio F, et al. (2002) Muscle-derived hematopoietic stem cells are hematopoietic in origin. Proc Natl Acad Sci U S A 99: 1341–1346.

Megeney LA, Kablar B, Garrett K, Anderson JE, Rudnicki MA (1996) MyoD is required for myogenic stem cell function in adult skeletal muscle. Genes Dev 10: 1173–1183.

Montarras D, Lindon C, Pinset C, Domeyne P (2000) Cultured myf5 null and myoD null muscle precursor cells display distinct growth defects. Biol Cell 92: 565–572.

Munsterberg AE, Lassar AB (1995) Combinatorial signals from the neural tube, floor plate and notochord induce myogenic bHLH gene expression in the somite. Development 121: 651–660.

Pagel CN, Partridge TA (1999) Covert persistence of mdx mouse myopathy is revealed by acute and chronic effects of irradiation. J Neurol Sci 164: 103–116.

Parker MH, Seale P, Rudnicki MA (2003) Looking back to the embryo: Defining transcriptional networks in adult myogenesis. Nat Rev Genet 4: 497–507.

Petropoulos H, Skerjanc IS (2002) Beta-catenin is essential and sufficient for skeletal myogenesis in p19 cells. J Biol Chem 277: 15393–15399.

Polesskaya A, Seale P, Rudnicki MA (2003) Wnt Signaling Induces the Myogenic Specification of Resident CD45+ Adult Stem Cells during Muscle Regeneration. Cell 113: 841–852.

Pourquie O, Coltey M, Breant C, Le Douarin NM (1995) Control of somite patterning by signals from the lateral plate. Proc Natl Acad Sci U S A 92: 3219–3223.

Pourquie O, Fan CM, Coltey M, Hirsinger E, Watanabe Y, et al. (1996) Lateral and axial signals involved in avian somite patterning: A role for BMP4. Cell 84: 461–471.

Qu-Petersen Z, Deasy B, Jankowski R, Ikezawa M, Cummins J, et al. (2002) Identification of a novel population of muscle stem cells in mice: Potential for muscle regeneration. J Cell Biol 157: 851–864.

Sabourin LA, Girgis-Gabardo A, Seale P, Asakura A, Rudnicki MA (1999) Reduced differentiation potential of primary MyoD−/− myogenic cells derived from adult skeletal muscle. J Cell Biol 144: 631–643.

Schultz E, Jaryszak DL, Gibson MC, Albright DJ (1986) Absence of exogenous satellite cell contribution to regeneration of frozen skeletal muscle. J Muscle Res Cell Motil 7: 361–367.

Seale P, Sabourin LA, Girgis-Gabardo A, Mansouri A, Gruss P, et al. (2000) Pax7 is required for the specification of myogenic satellite cells. Cell 102: 777–786.

Sicinski P, Geng Y, Ryder-Cook AS, Barnard EA, Darlison MG, et al. (1989) The molecular basis of muscular dystrophy in the mdx mouse: A point mutation. Science 244: 1578–1580.

Soneoka Y, Cannon PM, Ramsdale EE, Griffiths JC, Romano G, et al. (1995) A transient three-plasmid expression system for the production of high titer retroviral vectors. Nucleic Acids Res 23: 628–633.

Tajbakhsh S, Bober E, Babinet C, Pournin S, Arnold H, et al. (1996) Gene targeting the myf-5 locus with nlacZ reveals expression of this myogenic factor in mature skeletal muscle fibres as well as early embryonic muscle. Dev Dyn 206: 291–300.

Tajbakhsh S, Rocancourt D, Cossu G, Buckingham M (1997) Redefining the genetic hierarchies controlling skeletal myogenesis: Pax- 3 and Myf-5 act upstream of MyoD. Cell 89: 127–138.

Tajbakhsh S, Borello U, Vivarelli E, Kelly R, Papkoff J, et al. (1998) Differential activation of Myf5 and MyoD by different Wnts in explants of mouse paraxial mesoderm and the later activation of myogenesis in the absence of Myf5. Development 125: 4155–4162.

Torrente Y, Tremblay JP, Pisati F, Belicchi M, Rossi B, et al. (2001) Intraarterial injection of muscle-derived CD34(+)Sca-1(+) stem cells restores dystrophin in mdx mice. J Cell Biol 152: 335–348.

Wakeford S, Watt DJ, Partridge TA (1991) X-irradiation improves mdx mouse muscle as a model of myofiber loss in DMD. Muscle Nerve 14: 42–50.

Williams BA, Ordahl CP (1994) Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification. Development 120: 785–796.

Ziman MR, Fletcher S, Kay PH (1997) Alternate Pax7 transcripts are expressed specifically in skeletal muscle, brain and other organs of adult mice. Int J Biochem Cell Biol 29: 1029–1036.

Appendix C

Microarray Comparison of Gene Expression Between Wild-type and *MyoD*$^{-/-}$ Primary Myoblasts

Purpose

To identify genes using microarray analysis that are differentially expressed between wild-type and MyoD$^{-/-}$ primary myoblasts for comparison with the set of genes identified using representational difference analysis.

Submission Information

Submitted: 2 December 2003
Accepted: 9 July 2004
Published: 13 July 2004

Contributions of Co-Authors

Jeff Ishibashi conducted the microarray analysis shown as Table 2 in the manuscript and Table S1 online as supplementary material.

Preface and Discussion of Tables 2 and S1

Representational difference analysis (RDA) was used to isolate transcripts that were uniquely expressed in growing wild-type and *MyoD*$^{-/-}$ myoblasts and might represent early satellite cell markers. Microarray analysis of RNA from *MyoD*$^{-/-}$ and wild-type myoblasts was compared with the RDA and Northern blot results for the difference products. Twenty-seven of the forty genes identified by RDA were represented on the Affymetrix Mu11k mouse expression microarray.

Those data are consistent with the established hypothesis [Sabourin, 1999 #196] that *MyoD*$^{-/-}$ myoblasts are representative of an early-activation, undifferentiated myoblast. Genes resolved by the wild-type myoblast RDA (subtracted with fibroblasts and differentiated muscle) were expressed at much higher levels in wild-type myoblasts than in *MyoD*$^{-/-}$ myoblasts (Table 2), and include conventional muscle genes such as cholinergic receptor α1 and troponin T1. Myogenin was expressed at a low but slightly elevated level, indicative of a minor amount of spontaneous differentiation amongst the wild-type myoblasts but lacking in the *MyoD*$^{-/-}$ cells. In contrast, genes identified by the *MyoD*$^{-/-}$ RDA (subtracted with fibroblasts) were expressed in both *MyoD*$^{-/-}$ and wild-type myoblasts (Table 2). Although several (vascular cell adhesion molecule 1; laminin α5; placental growth factor) were moderately higher in *MyoD*$^{-/-}$ cells, they were also general markers of wild-type myoblasts.

Each method has benefits that support the other. Whereas RDA can identify rare and unknown transcripts, microarrays provide a rapid, semi-quantitative, genome-wide survey that can identifies differentially expressed genes that are present in both samples.

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Developmental Biology 275 (2004) 287–300

DEVELOPMENTAL BIOLOGY

www.elsevier.com/locate/ydbio

Muscle satellite cell-specific genes identified by genetic profiling of *MyoD*-deficient myogenic cell

Patrick Seale[a,b], Jeff Ishibashi[b], Chet Holterman[b], Michael A. Rudnicki[a,b,*]

[a]*Department of Biology, McMaster University, Hamilton, Ontario, Canada L8S 4K1*
[b]*Molecular Medicine Program, Ottawa Health Research Institute, Ottawa, Ontario, Canada K1H 8L6*

Received for publication 2 December 2003, revised 9 July 2004, accepted 13 July 2004
Available online 11 September 2004

Abstract

Satellite cells are committed myogenic progenitors that give rise to proliferating myoblasts during postnatal growth and repair of skeletal muscle. To identify genes expressed at different developmental stages in the satellite cell myogenic program, representational difference analysis of cDNAs was employed to identify more than 50 unique mRNAs expressed in wild-type myoblasts and $MyoD^{-/-}$ myogenic cells. Novel expression patterns for several genes, such as *Pax7*, *Asb5*, *IgSF4*, and *Hoxc10*, were identified that were expressed in both quiescent and activated satellite cells. Several previously uncharacterized genes that represent putative MyoD target genes were also identified, including *Pw1*, *Dapk2*, *Sytl2*, and *NLRR1*. Importantly, many genes such as IgSF4, *Neuritin*, and *Klra18* that were expressed exclusively in $MyoD^{-/-}$ myoblasts were also expressed by satellite cells in undamaged muscle in vivo but were not expressed by primary myoblasts. These data are consistent with a biological role for activated satellite cells that induce Myf5 but not MyoD. Lastly, additional endothelial and hematopoietic markers were identified supporting a nonsomitic developmental origin of the satellite cell myogenic lineage.

Keywords: Representational difference analysis; Skeletal muscle; Satellite cell; Regeneration; Pax7; MyoD; Myf5

Introduction

Muscle satellite cells are specialized myogenic progenitors that are activated during postnatal growth and regeneration of skeletal muscle. Consequently, the majority of adult myonuclei are derived from satellite cells following the growth that occurs postnatally (Schultz, 1996). In undamaged adult muscle, most satellite cells are quiescent, contain highly condensed nuclei, and are located beneath the basal lamina of mature muscle fibers (Armand et al., 1983; Mauro, 1961; reviewed by Bischoff, 1994; Hawke and Garry, 2001; Seale and Rudnicki, 2000). In response to a variety of stimuli including exercise, stretching, and injury, satellite cells are activated and give rise to committed myogenic precursor cells (MPCs) that proliferate and differentiate to form new myofibers (Appell et al., 1988; Darr and Schultz, 1987; Grounds and Yablonka-Reuveni, 1993; Rosenblatt et al., 1994; Schultz, 1989; Schultz et al., 1985). Activated satellite cells are also thought to generate progeny that remains undifferentiated, hence restoring the pool of quiescent satellite cells (Bischoff, 1994; Seale and Rudnicki, 2000).

Due to their low abundance in mature muscle (2–5% of sublaminar nuclei), it has been difficult to molecularly investigate early events associated with satellite cell activation. Additionally, there remains a paucity of genetic markers unique to the satellite cell lineage. Many satellite cell markers such as Neural Cell Adhesion Molecule-1 (Ncam1) (Bischoff, 1994), Foxk1 (Garry et al., 1997), c-met (Cornelison and Wold, 1997), and Syndecans 3–4 (Cornelison et al., 2001) are also expressed in other lineages. The characterization of novel genes expressed in satellite cells is essential for elucidating the molecular pathways implicated in their development and function during tissue growth and regeneration.

* Corresponding author. Molecular Medicine Program, Ottawa Health Research Institute, 501 Smyth Road, Ottawa, Ontario, Canada K1H 8L6. Fax: +1 613 737 8803.
E-mail address: mrudnicki@ohri.ca (M.A. Rudnicki).

0012-1606/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2004.07.034

Quiescent satellite cells do not express detectable mRNA or protein for any of the myogenic regulatory factors (MRFs) (Cooper et al., 1999; Cornelison and Wold, 1997; Smith et al., 1994; Yablonka-Reuveni and Rivera, 1994). Following activation, satellite cells upregulate either *MyoD* or *Myf5* mRNA before the initiation of DNA synthesis (Smith et al., 1994). RT-PCR experiments of single cells on isolated muscle fibers showed that activated satellite cells first express either *MyoD* or *Myf5* before coexpressing both factors (Cornelison and Wold, 1997). Analysis of regenerating muscle confirmed that myoblasts express either MyoD or Myf5 alone or coexpress both factors (Cooper et al., 1999). These observations suggest that activated satellite cells possess differential biological properties depending upon whether they initially activate *MyoD* or *Myf5*.

Previous work identified a unique requirement for MyoD in the satellite cell lineage (Megeney et al., 1996). Specifically, $MyoD^{-/-}$ muscles display a severe regeneration deficit following crush-induced damage or on a dystrophic (*mdx*) background. Importantly, $MyoD^{-/-}$ muscle contains an increased number of satellite cells suggesting an increased propensity for $MyoD^{-/-}$ satellite cells to self-renew rather than terminally differentiate. Consistent with these findings, satellite cell-derived myoblasts from *MyoD*-deficient muscle display a profound differentiation deficit and an increased growth rate in vitro (Cornelison et al., 2000; Sabourin et al., 1999; Yablonka-Reuveni et al., 1999). Significantly, $MyoD^{-/-}$ myogenic cells express increased levels of Myf5 mRNA and protein (Sabourin et al., 1999), demonstrating the inability of Myf5 to compensate for the absence of MyoD during differentiation.

The phenotype of *MyoD*-deficient adult myoblasts suggested the hypothesis that $MyoD^{-/-}$ myogenic progenitors are similar to activated satellite cells that express Myf5 alone (i.e., Myf5+:MyoD−) and are developmentally upstream of cells that express MyoD (Sabourin et al., 1999; Seale and Rudnicki, 2000). Therefore, cultured $MyoD^{-/-}$ myogenic cells are a unique resource to identify satellite cell-specific mRNAs. In this study, we employed representational difference analysis of cDNAs (Hubank and Schatz, 1994) to identify markers expressed specifically in wild-type primary myoblasts and *MyoD*-deficient myogenic cells. These experiments provide a collection of novel cDNAs whose expression defines different developmental stages in the satellite cell lineage.

Materials and methods

Cell cultures

Primary myoblast cultures were prepared from adult (6–8 weeks old) $MyoD^{-/-}$ (Rudnicki et al., 1992) and Balb/c (Jackson Laboratories) control animals as described previously (Sabourin et al., 1999). Primary low passage (<p6) myoblast cultures derived from multiple (>3) animals were pooled for use in gene expression studies to control for biological variability and to maintain their primary characteristics. Desmin and Myf5 immunoreactivity confirmed that cell cultures were >98% pure myoblasts. Myoblasts were propagated on collagen-coated dishes in Ham's F-10 medium (Invitrogen) supplemented with 20% FBS and 2.5 ng/ml bFGF (Invitrogen). Myoblast cultures were induced to differentiate in DMEM supplemented with 5% horse serum. Mouse embryonic fibroblasts (MEFs) were obtained from E14.5 Balb/c mouse embryos using standard procedures (Robertson, 1987) and maintained as primary cultures in 10% FBS/DMEM. C3H10T1/2 fibroblasts and C2C12 myoblasts were obtained from ATCC and maintained in 10% FBS/DMEM.

Representational difference analysis

Representational difference analysis of cDNAs was performed as described previously (Hubank and Schatz, 1994). Briefly, double-stranded cDNA was digested with the four-cutter *Dpn*II (New England Biolabs) and ligated with R-Bgl-24 adaptors. PCR was used to amplify the cDNA pools before subtractive hybridizations. R-Bgl-24 adaptors were subsequently removed from the cDNA pools, and J-Bgl-24 adaptors were then ligated only to "tester" cDNA pools. For subtractive hybridizations, wild-type myoblast cDNA "tester" was subtracted against MEF cDNA at 1:100 and 1:400 to yield wtDP1 and wtDP2, respectively. DP3 was generated by subtracting DP2 against cDNA prepared from uninjured whole skeletal muscle at a ratio of 1:400. $MyoD^{-/-}$ cDNA tester was subtracted against C3H10T1/2 fibroblast cDNAs at 1:100 and 1:400 to generate mdDP1 and mdDP2, respectively. Final difference products were cloned directly into pCR2.1 (Invitrogen) for sequence analysis.

Affymetrix array analysis

Total RNA was harvested from two independent isolations of low-passage cultures of wild-type myoblasts and $MyoD^{-/-}$ myogenic cells (as described above for RDA analysis). Hindlimb muscles from three 6- to 8-week-old mice were used for each isolate to control for biological variability between animals. Standard Affymetrix protocols were used to yield fluorescently labeled cRNA fragments, which were hybridized to the Mu11k-SubA and SubB GeneChips (Affymetrix). The hybridized GeneChips were scanned, and the raw image files analyzed using Affymetrix Microarray Suite 4.0's empirical algorithm to generate numeric average difference values and qualitative absent/marginal/present call values. The resulting processed data were further examined using Microsoft Excel and Access. Negative and low average difference values produced by the empirical algorithm were arbitrarily reassigned threshold values of 50 for purposes of calculating fold changes. The average fold change value used to compare the $MyoD^{-/-}$ data against

the wild-type data was calculated as ratios of the means of the replicates.

RNA isolation and synthesis of double-stranded cDNA

Total RNA was prepared from cell cultures and tissues using GIT method as previously described (Birnboim, 1988) or by using Qiagen RNeasy kits for Affymetrix profiling experiments. PolyA+ mRNAs were prepared by two rounds of selection with oligo d(T) cellulose (Amersham Bioscience). Double-stranded cDNAs were generated using the Universal RiboClone cDNA synthesis kit (Promega). The yield of double-stranded cDNA was determined by radioactive monitoring of first and second strand synthesis reactions.

Expression analysis of RDA clones

Northern blot studies were performed according to standard techniques using random-primed ^{32}P-dCTP radiolabeled RDA products as probes (Maniatis et al., 1982). Total RNA (15 μg) from various tissues and cell lines was electrophoresed in denaturing Formaldehyde gels and transferred to Hybond-N filters (Amersham Bioscience). In situ hybridizations were performed on 10-μm cryosections of mouse TA muscles from 8-week-old wild-type or 3-week-old *mdx* mice (Jackson Laboratories) according to previously described procedures (Braissant and Wahli, 1998). Sense and antisense in situ probes were synthesized from RDA products using the DIG labeling mix (Roche) with SP6 or T7 RNA polymerase (Roche). Alkaline phosphatase-conjugated anti-DIG antibody (Roche) followed by reaction with BCIP/NBT (Roche) was used to detect hybridized cRNA probes.

Immunohistochemistry was performed on paraformaldehyde (PFA)-fixed, 10-μm cryosections using goat anti-Vcam1 antibody (Santa Cruz), followed by staining with a biotin-conjugated secondary antibody (Zymed) and streptavidin-HRP (Zymed). Immunoreactive cells were visualized using aminoethyl carbazole (AEC) substrate (Sigma). Immunohistochemistry on cultured cells was performed by fixation with 4% PFA for 5 min, followed by permeabilization with 0.5% Triton X-100 for 5 min. Cells were incubated with primary antibodies against Desmin (DAKO), Pax7 (Developmental Studies Hybridoma Bank), Myf5 (C20, Santa Cruz), β-Gal (Molecular Probes), and Vcam1 (Santa Cruz). Secondary detection was performed with FITC- or HRP-conjugated antibodies (Sigma).

Results

Expression profiling of satellite cell-derived myoblasts

Representational difference analysis of cDNAs (cDNA RDA) was employed to generate libraries of expressed cDNAs in primary myoblasts from wild-type muscle. cDNAs from wild-type primary myoblasts were subtracted sequentially against mouse embryonic fibroblast (MEF) cDNAs at ratios of 1:100 and 1:400 to obtain the first and second difference products (wtDP1 and wtDP2), respectively (Fig. 1A). To refine the products further and eliminate markers of terminal myocyte differentiation, including structural genes (e.g., *myosins*, *dystrophin*, etc.), an additional subtractive step against cDNAs from whole adult skeletal muscle (1:400) was performed to generate wtDP3. As expected, the complexity of the wild-type myoblast cDNA mixture was progressively reduced in wtDP1 and wtDP2, resulting in the appearance of several distinct cDNA products in wtDP3 (Fig. 1B).

A similar strategy was used to identify genes responsible for the behavior of *MyoD*$^{-/-}$ satellite cells including their increased capacity for self-renewal (Megeney et al., 1996; Sabourin et al., 1999). *MyoD*$^{-/-}$ myoblast cDNAs were subtracted twice against cDNAs from C3H10T1/2 fibroblasts (at 1:100 and 1:400) to generate mdDP1 and mdDP2, respectively (Fig. 1C). C3H10T1/2 cDNA was used in this screen to avoid losing genes coexpressed by activated *MyoD*$^{-/-}$ myogenic cells and primary multipotent MEFs. After two rounds of subtraction, mdDP2 contained several distinct cDNA species (Fig. 1D).

Pools of cDNA from wtDP3 and mdDP2 were cloned into pCR2.1 (Invitrogen). To recover lower abundance difference products, 400 individual clones from both subtracted wild-type and *MyoD*$^{-/-}$ libraries were screened. Dot-blots consisting of PCR-amplified RDA products were hybridized with mixtures of labeled clones to identify redundant sequences. After dot-blot and sequence analysis, 18 difference products were identified from wild-type myoblasts and 34 difference products from *MyoD*$^{-/-}$ myogenic cells (Table 1). Comparison of cDNA sequences to database entries in GenBank established the identity of cloned products (accession numbers provided in Table 1).

Expression analyses of RDA products from wild-type primary myoblasts

From reverse Northern blot experiments, 51 of the 52 genes identified were differentially expressed in the starting pools of amplified cDNAs (data not shown), thus validating the efficacy of RDA subtractions. The expression profile of 40 genes during primary adult myoblast differentiation was assessed by Northern blot analysis as outlined in Table 1. The expression patterns for the remaining 11 clones were not determined because of technical difficulties in obtaining nonrepetitive cDNA probes or due to the identity of the cDNA as a gene involved in protein translation, mitochondria, or metabolism that was not the focus of this study (see supplementary material for information).

Northern analysis revealed that *Pax7*, *L-myc*, *Pb99*, and *Asb5* were expressed in proliferating wild-type and *MyoD*$^{-/-}$ myoblasts with no upregulation observed during myotube differentiation. As previously demonstrated, *Pax7* was

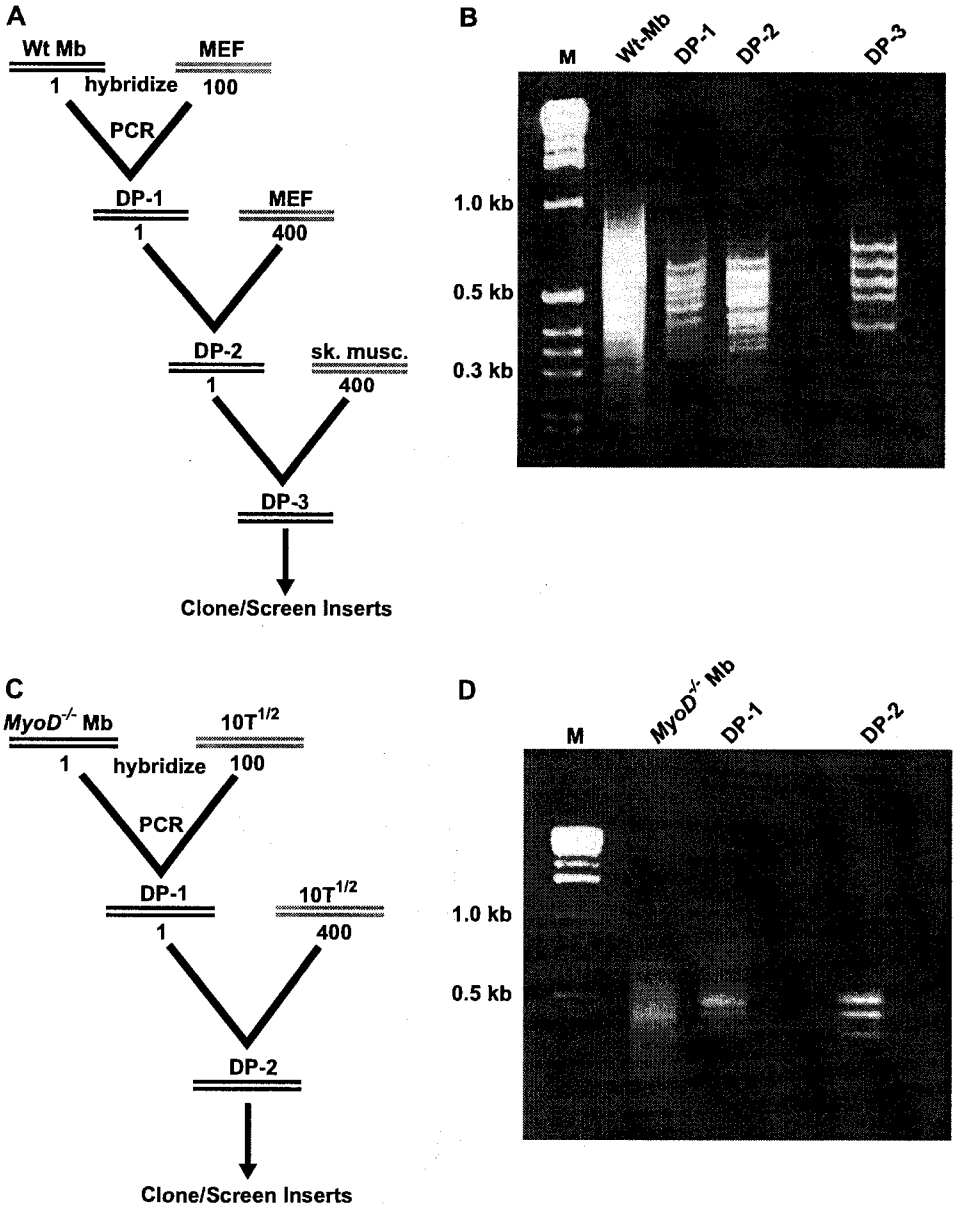


Fig. 1. Experimental strategy for identifying myoblast-specific mRNAs. (A) Wild-type myoblast cDNAs were subtracted against cDNAs prepared from mouse embryonic fibroblasts (MEF) and whole adult skeletal muscle (sk. musc) to generate the third difference product (DP-3). Ratios used for hybridizations are indicated. (B) Agarose gel electrophoresis of the starting pool of wild-type myoblast (Wt-Mb) cDNAs and subtracted DP-1, DP-2, and DP-3 showed a progressive enrichment of specific cDNA molecules. (C) *MyoD*$^{-/-}$ myogenic cell-specific cDNAs were cloned after subtraction against C3H10T1/2 cDNAs. (D) Agarose gel analysis of the starting *MyoD*$^{-/-}$ myogenic cell cDNA pool (*MyoD*$^{-/-}$Mb), DP-1, and DP-2 revealed the appearance of distinct cDNAs in DP-2.

expressed in proliferating wild-type and *MyoD*$^{-/-}$ myoblasts but rapidly downregulated upon differentiation (Seale et al., 2000). *Pb99*, a gene that encodes a seven-pass transmembrane protein was expressed exclusively in undifferentiated myoblast cultures (Fig. 2A). Analogous to *Pax7*, *Pb99* was not detected in total RNA from a panel of cell lines or mouse tissues (Fig. 2A). Furthermore, in C2C12 myoblasts, *Pb99* mRNA was downregulated within the first day of differentiation (data not shown).

Asb5, encoding an Ankyrin-repeat SOCS box containing protein, was expressed at similar levels in proliferating and differentiating myogenic cells including wild-type, *MyoD*$^{-/-}$, and C2C12 myoblasts (Fig. 2A). *Asb5* mRNA was detected in total RNA from adult skeletal muscle but was not expressed in other adult mouse tissues (Fig. 2A). These data therefore suggest previously unrecognized roles for Pb99 and Asb5 in the myogenic satellite cell lineage.

About 10% of cells in cultures of primary growing myoblasts express myosin heavy chain and are presumed to represent a basal level of differentiation (Sabourin et al., 1999). Accordingly, several RDA clones identified from cultures of wild-type myoblasts were defined as differ-

Table 1
Summary and expression of RDA clones

Clone	Title	GenBank	Expression by Northern Blot				Comments/Reference
			Wt-Mb	Wt-D	MD-/- M	MD-/- D	
Transcription Factors							
* MD p286	Hoxc10	NM_017409	++++	++	++++	++++	Expressed in regenerating limbs of Axolotls (Carlson et al., 2001).
dp3-7	Pax7	AF254422	++++	-	++++	-	Paired-box transcription factor required in the development of muscle satellite cells (Seale et al., 2000).
wt-73	Lmyc1	X13945	+	+	+	+	bHLH-Leucine Zipper protein
dp3-9	Myogenin	D90156	++	++++	-	++	Muscle-specific transcription factor that regulates differentiation (Smith et al., 1994; Yablonka-Reuveni and Rivera, 1994).
Receptors and Integral Membrane Proteins							
MD 2	Tcrb-V13	BC034857	+/-	-	+++	-	T cell receptor β, variable 13.
* MD p12	Neuritin-1	BC035531	-	-	++++	-	Expressed exclusively in neural tissue and $Myod^{-/-}$ skeletal myoblasts. GPI- membrane anchored protein that promotes neurite outgrowth (Naeve et al., 1997).
* MD 41	Itm2a	L38971	+	++	++	++++	Integral membrane protein also expressed in chondrogenic progenitors (Deleersnijder et al., 1996).
* MD p259	Klra18	NM_053153	-	-	+++	+++	Killer cell lectin like receptor-18. Cell surface molecule expressed in pre-B cells and T cells (Chan and Takei, 1989).
* wt-30	Pb99	AF249738	++++	-	++++	-	Seven-pass transmembrane protein expressed in pre-B cells (Sleckman et al., 2000).
MD p67	similar to MEGF10	XM_140362	++	-	+++	-	Multiple EGF repeat containing protein.
MD p31	Ptk7	NM_175168	+++	++	+++	++	Protein tyrosine kinase-7 belongs to a novel subclass of receptor tyrosine kinases (Park et al., 1996).
* dp3-1	Lrrn1	D45913	++	+++	+/-	++	Neuronal Leucine Rich Region-1. Expressed in developing nervous system (Taguchi et al., 1996).
dp3-8	Chrnd	BC052153	+	++++	-	++	Cholinergic receptor, nicotinic, delta polypeptide.
* wt-23	Chrna1	NM_007389	+	++++	-	+	Cholinergic receptor, nicotinic, alpha polypeptide-1.
* wt-141	Sytl2	NM_031394	+	+++	-	+	Synaptotagmin-like-2. Contains Slp-homology domain (SHD).
Adhesion/Extracellular Matrix Proteins							
MD 42	Cdh6	NM_007666	+	-	+++	-	Cadherin-6. Cell adhesion molecule expressed in hematopoietic cells and motorneurons (Matthieus et al., 2002; Mbalaviele et al., 1998).
* MD 44	Lama5	XM_203796	+	+	++++	++	Laminin-α5. Adhesive protein for hematopoietic cells (Gu et al., 1999). Expressed in developing and dystrophic muscle (Ringelmann et al., 1999; Sorokin et al., 1997).
* MD p3	Igsf4a	NM_018770	-	-	+++	-	Immunoglobulin superfamily-4. Similar to Synaptic cell adhesion molecule-1 (Biederer et al., 2002).
* MD p168	Vcam1	BC029823	+	-	++++	-	Vascular Cell Adhesion Molecule-1. Muscle satellite cell marker (Rosen et al., 1992).
MD p18	Itgβ4	L04678	++++	++	++	-	Integrin-α6β4 binds components of the extracellular matrix; and is implicated in adhesion and proliferation (Murgia et al., 1998).
MD p16	Itgα7	NM_008398	+	++++	++	++++	Integrin-α7, marks satellite cells (Blanco-Bose et al., 2001; LaBarge and Blau, 2002).

(continued on next page)

Table 1 (*continued*)

Clone	Title	GenBank	Expression by Northern Blot				Comments/Reference
			Wt-Mb	Wt-D	MD-/- M	MD-/- D	
Structural/Cytoplasmic Proteins							
MD p39	Nestin	NM_016701	++	+++	++	++	Intermediate filament protein expressed in myoblasts and regenerating muscle (Carlsson et al., 1999; Vaittinen et al., 2001).
dp3-3	Scga	NM_009161	++	++++	+/-	+/-	α-sarcoglycan, Dystrophin associated glycoprotein. Mutations cause limb girdle muscular dystrophy (Duclos et al., 1998; Liu et al., 1997).
MD p26	Caldesmon	BC019435	+	++++	++	++++	Functions in stabilization of microfilament networks (Matsumara and Yamashiro, 1993).
dp3-13	Tnnt1	NM_011618	+	+++	-	+	Troponin T1 slow. Expressed in differentiated muscle.
Cell Cycle/Apoptosis							
* MD p42	Mcm6	D86726	+	-	+++	+	Cdc21 binding protein expressed in proliferating intestinal crypt cells (Kimura et al., 1996; Sykes and Weiser, 1995).
* MD p123	G0S2	NM_008059	+	+	+++	+	G0/G1 switch gene-2. Target for BMP signals in mesenchymal cells (Buchner et al., 1998).
MD p249	Peg3 (Pw1)	AF038939	+++	n/d	+/-	n/d	Expressed in mesoderm and differentiated muscle (Relaix et al., 1996). Implicated in p53-dependent apoptosis (Coletti et al., 2002).
wt-18	Dapk2	XM_134847	+	++++	-	+	Death-associated protein kinase 2. Calcium/calmodulin regulated kinase implicated in apoptosis (Kawai et al., 1998).
Signaling/Secreted Proteins							
* MD 52	Plgf	BC016567	++	-	++++	+++	Placenta derived growth factor. Closely related to VEGF. Promotes growth of early hematopoietic cells (Luttun et al., 2002).
* wt-17	Asb5	AF398966	++	++	++	++	Ankyrin repeat SOCS-box protein-5. Expressed in satellite cells (Boengler et al., 2003).
MD p158	PAI-2	AJ000386	++	-	++	-	Plasminogen activator inhibitor-2. Plasmin activity is required for muscle regeneration (Lluis et al., 2001; Suelves et al., 2002).
Miscellaneous and Unknown mRNAs							
MD p46	Unknown	XM_129466	-	-	++	-	Putative G-protein coupled receptor.
MD p87	Unknown	BC020031	+/-	+	++	+	Contains CAG repeat region.
9 clones	Neomycin		-	-	+++	+++	Expressed highly from PGK1 promoter in *MyoD*-deficient cells.
MD 62	Unknown		+	++	+	++	Contains S100 calcium binding domain.
MD p35	Unknown	AK008210	+	++	+	++	Putative Androgen-induced protein
MD p11	H19 mRNA	BC025150	++	+++	+	++	Untranslated and imprinted mRNA. Expressed during myoblast and ES cell differentiation (Bartolomei et al., 1993; Leibovitch et al., 1995).
* dp3-2	Unknown	BC024685	+/-	+++	-	+	Hypothetical protein
* dp3-5	Unknown	XM_132832	+	+++	-	-	Hypothetical protein similar to Arachidonate 5-Lipoxygenase. Not expressed in *MyoD*-deficient cells.

*Northern blots are shown in Figures 2 and 3.

mRNAs expressed at elevated levels in *MyoD*$^{-/-}$ cells relative to wildtype cultures

mRNAs upregulated during myoblast differentiation and myotube formation

References for the RDA clones: Bachner et al., 1998; Bartolomei et al., 1993; Carlsson et al., 1999; Chan and Takei, 1989; Coletti et al., 2002; Deleersnijder et al., 1996; Duclos et al., 1998; Gu et al., 1999; Kawai et al., 1998; Kimura et al., 1996; Leibovitch et al., 1995; Liu et al., 1997; Lluis et al., 2001; Luttun et al., 2002; Marthiens et al., 2002; Matsumura and Yamashiro, 1993; Mbalaviele et al., 1998; Murgia et al., 1998; Park et al., 1996; Relaix et al., 1996; Ringelmann et al., 1999; Sorokin et al., 1997; Suelves et al., 2002; Sykes and Weiser, 1995; Taguchi et al., 1996; Vaittinen et al., 2001.

entiation specific (Fig. 2B and hatched rows in Table 1), including *Dapk2*, *Lrrn1*, *Sytl2*, *α-sarcoglycan*, *myogenin*, *troponin T1 slow*, and three unknown genes. Interestingly, these genes were all expressed at lower levels in *MyoD*$^{-/-}$ cells, consistent with a requirement for MyoD during adult myoblast differentiation (Cornelison et al., 2000; Sabourin et al., 1999; Yablonka-Reuveni et al., 1999). *Dapk2*, *Lrrn1*, *Sytl2*, and the unknown genes remain to be studied

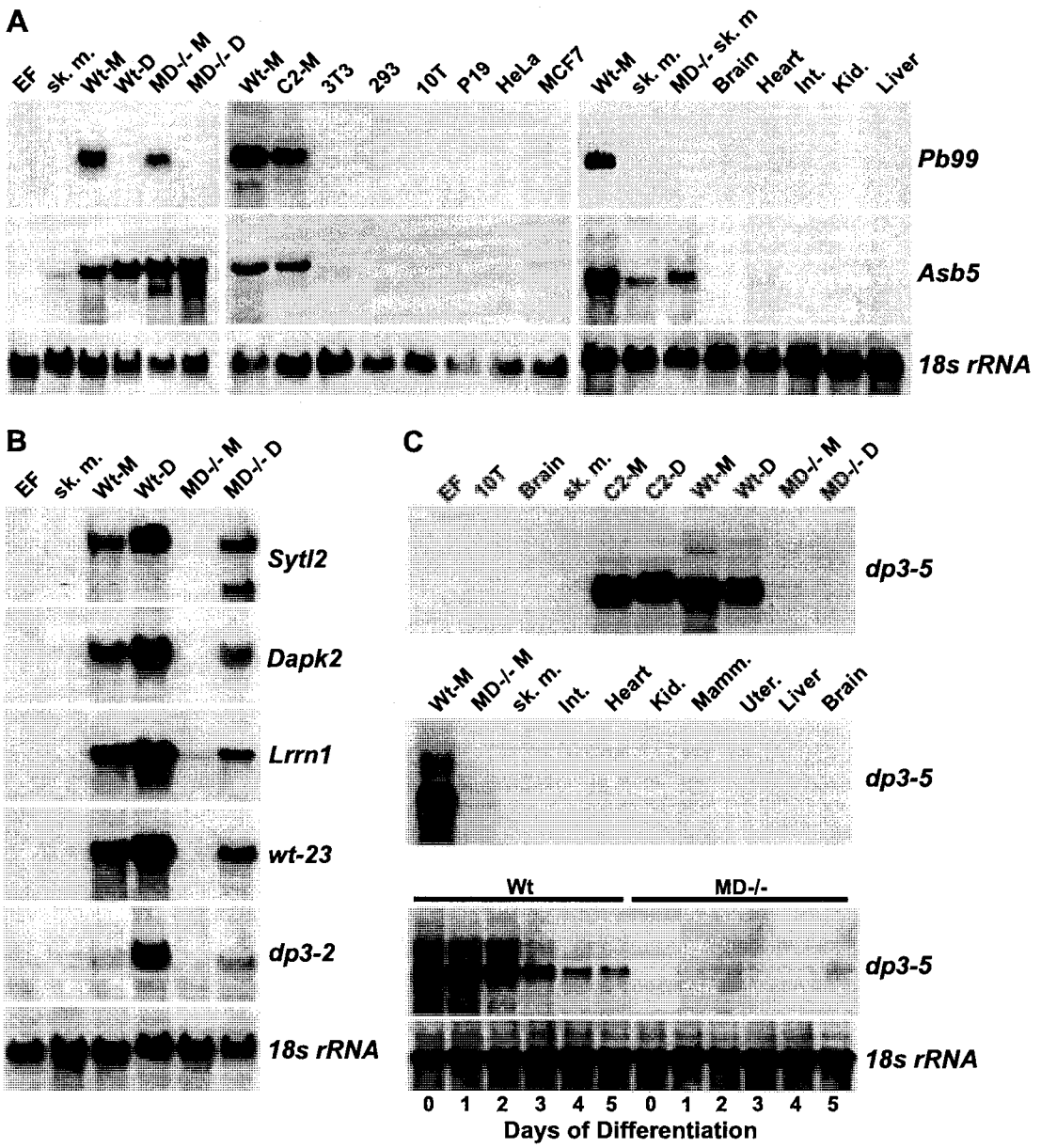


Fig. 2. Expression analysis of wild-type myoblast specific-genes. (A) *Pb99* and *Asb5* were specifically expressed in proliferating wild-type (Wt-M), *MyoD*$^{-/-}$ (MD$^{-/-}$M), and C2C12 (C2-M) myoblasts. *Asb5* was also expressed after 3 days of differentiation (Wt-D and MD$^{-/-}$D). *Asb5* mRNA was also detected in skeletal muscle tissue (sk. m). *Pb99* and *Asb5* mRNAs were not detected in a panel of cell lines. (B) *Sytl2*, *Dapk2*, *Lrrn1*, and two unknown genes (wt-23 and dp3-2) were upregulated during myogenic differentiation in wild-type (Wt) myoblasts. Reduced expression levels were observed in differentiating *MyoD*$^{-/-}$ (MD−/−) cultures. (C) Unknown gene, dp3-5 was specifically expressed in proliferating and differentiating wild-type and C2C12 (C2) myoblasts. However, dp3-5 was not expressed in *MyoD*$^{-/-}$ cultures either during growth conditions (day 0) or throughout differentiation (days 1–5). *18s rRNA* was used to control for loading.

in the context of myogenesis and may play important roles in the differentiation process downstream of MyoD.

Clone *dp3-5*, an unknown gene related to *Arachidonate 5'-lipoxygenase*, was highly expressed in wild-type and C2C12 myoblasts during proliferation and differentiation (Fig. 2C). Furthermore, expression of *dp3-5* was not detected in total RNA isolated from a panel of adult mouse tissues (Fig. 2C). Strikingly, *dp3-5* mRNA was completely absent in proliferating as well as differentiating *MyoD*$^{-/-}$ myoblasts (Fig. 2C). These data therefore suggest that this novel gene is induced downstream of MyoD in skeletal myoblasts.

Expression analyses of MyoD$^{-/-}$ *RDA clones*

Transcriptional profiling of *MyoD*-deficient myogenic cells was employed to identify markers expressed by a more primitive myogenic progenitor. Northern analyses revealed several genes expressed at higher levels in *MyoD*$^{-/-}$ cultures relative to their wild-type counterparts (see gray-shaded rows in Table 1). Importantly, the established satellite cell markers *Vcam1* (Fig. 3A) and *Integrin-α7* (not shown) were expressed at higher levels in *MyoD*$^{-/-}$ cells compared to wild-type primary myoblasts and C2C12

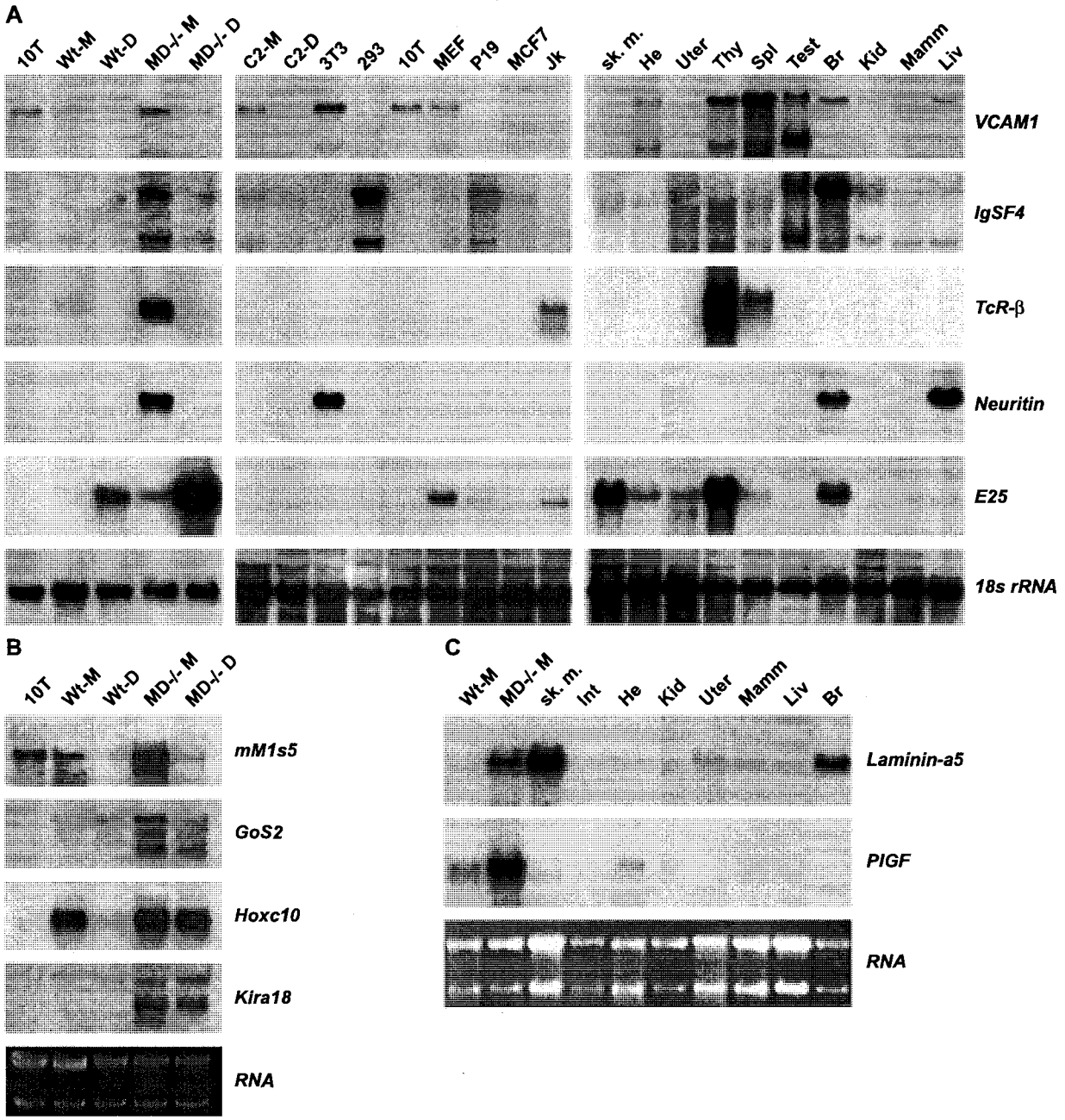


Fig. 3. Expression analysis of $MyoD^{-/-}$-specific genes. (A) *Vcam1* was highly expressed in $MyoD^{-/-}$ (MD−/−M) myogenic cells with low levels detected in wild-type myoblasts. *Vcam1* transcripts were also detected in various cell cultures including MEFs and C2C12 myoblasts (C2-M). *IgSF4* was expressed in $MyoD^{-/-}$ myoblasts, 293 cells, P19 cells, as well as brain (Br) and testis (Test), but not in wild-type myoblasts. *TcR-β* was expressed at surprisingly high levels in $MyoD^{-/-}$ cells as well as the Jurkat T cell line (Jk), Thymus (Thy), and Spleen (Spl). *Neuritin* mRNA was only detected in $MyoD^{-/-}$ myoblasts, NIH 3T3 fibroblasts, brain, and liver. *Itm2a* was expressed at dramatically higher levels in $MyoD^{-/-}$ compared to wild-type myoblasts, with a marked upregulation in differentiating cells. The *Itm2a* transcript was not detected in C2C12 myoblasts. *18s rRNA* was used as loading control. (B) *Mcm6*, *G_0S2*, and *Hoxc10* were all expressed at higher levels in $MyoD^{-/-}$ (MD−/−M) relative to wild-type myoblasts (Wt-M). *Klra18* mRNA was only detected in $MyoD^{-/-}$ myoblasts. Ethidium-stained RNA samples show loading. (C) *Laminin-α5* and *PlGF* were highly expressed in $MyoD^{-/-}$ myoblasts with low levels detected in wild-type cells. *Laminin-α5* mRNA was readily detected in skeletal muscle and brain tissue, while PlGF was expressed at low levels in heart. Ethidium-stained RNA shows loading.

myoblasts (Blanco-Bose et al., 2001; LaBarge and Blau, 2002; Rosen et al., 1992).

RDA facilitated the identification of additional genes expressed in satellite cells. *Immunoglobulin Superfamily-4* (*IgSF4*) was expressed at high levels in $MyoD^{-/-}$ cells, with a complete absence of expression in wild-type primary myoblasts and C2C12 myoblasts (Fig. 3A). *IgSF4* transcripts were also detected in 293 and P19 cells as well as several adult tissues. Similarly, *Neuritin* was expressed in $MyoD^{-/-}$ myogenic cells and not primary or C2C12 myoblasts. *Neuritin* transcripts were also observed in NIH 3T3 fibroblasts, brain, and liver (Fig. 3A).

Surprisingly, *T-cell-Receptor-β* (*TcR-β*) was highly expressed by *MyoD*$^{-/-}$ myogenic cells, with very low levels detected in wild-type myoblasts (Fig. 3A). Similarly, *Placenta Growth Factor* (*PlGF*) was expressed at high levels by *MyoD*$^{-/-}$ myogenic cells relative to wild-type myoblasts (Fig. 3C). Northern analysis of *Laminin-α5* demonstrated expression of the transcript in *MyoD*$^{-/-}$ myogenic cells, adult skeletal muscle, and brain (Fig. 3C). *Hoxc10*, *Killer-cell Lectin Receptor* (*Klra18*), *Mcm6*, *G_0/G_1 switch gene-2* (G_0S2), and *Caldesmon* also displayed elevated expression levels in *MyoD*$^{-/-}$ myogenic cells relative to wild-type cultures (Fig. 3B and gray shaded rows in Table 1).

Several genes cloned by RDA from activated *MyoD*$^{-/-}$ satellite cells were also expressed at similar or higher levels in wild-type myoblasts and throughout differentiation. *Protein tyrosine kinase-7* (*PTK7*), *Plasminogen activator inhibitor-2* (*PAI2*), and a novel EGF repeat containing gene similar to *MEGF10* (MD p67) were expressed at high levels in proliferating wild-type and *MyoD*$^{-/-}$ myogenic cells and downregulated during differentiation (see Table 1). *Integrin β4* exhibited elevated expression in wild-type compared to *MyoD*-deficient cells and was downregulated upon differentiation (see Table 1). *H19* mRNA, *Nestin*, and two unknown genes (MD p35 and MD 62) are examples of genes cloned from *MyoD*$^{-/-}$ cells that were induced upon myogenic differentiation (see Table 1). These results have thus identified a significant number of novel genes that are implicated in various stages of satellite cell myogenesis.

Transcriptional profiling by Affymetrix arrays

GeneChip microarray analysis was used to examine the expression patterns of genes identified by the RDA protocol. Seven of fourteen genes (7/14; 50%) identified in the wild-type myoblast RDA and 20 of 26 genes (14/26; 54%) identified in the *MyoD*$^{-/-}$ RDA experiments were represented in Mu11K probesets. The primitive, undifferentiated state of the *MyoD*$^{-/-}$ myoblasts is well illustrated by the expression patterns of muscle-specific genes identified in the wild-type myoblast RDA screen including *Troponin T1 slow*, *Chrna1*, *Lrrn1*, and *Dapk2* (Table 2). These findings are consistent with the Northern blot data summarized in Table 1.

Similarly, expression of *MyoD*$^{-/-}$ RDA-identified genes corroborates the qualitative Northern blot assessments showing that certain genes such as *Vcam1*, *Laminin-α5*, and *Plgf* are expressed at markedly higher average levels in *MyoD*$^{-/-}$ vs. wild-type myoblasts (see Table 2). A subclass of cDNAs identified in the *MyoD*$^{-/-}$ RDA screen including *H19*, *Integrin-β4* (Itgb4), *Nestin*, and *Peg3* was expressed at elevated levels in wild-type myoblasts relative to *MyoD*-deficient cells (Table 2). These genes are thought to be muscle-specific genes that were not represented in the wild-type myoblast RDA screen because of their expression in skeletal muscle tissue that was used in the subtraction procedure (see Fig. 1).

Table 2
Expression of RDA candidates by Mu11K GeneChip

Gene	GenBank	Avg. Diff./Call[a]		wt/MD
		wt	MD	Fold
Wild-type myoblast RDA candidates				
Chrna1	X03986	19325P	191A	101.4
Tnnt1	W08218	3608P	50A	72.2
Lrrn1	D45913	2919P	50A	58.4
Lmyc1	X13945	257P/A	50A	5.1
Chrnd	L10076	199P/A	50A	4.0
Dapk2	W82116	149P/A	50A	3.0
Myog	D90156	1837P	1319M/A	1.4
MyoD$^{-/-}$ *myoblast RDA candidates*				
Informative				
Vcam1	X67783	1037P	4612P	−4.4
Lama5[b]	U37501	182P/A	731P	−4.0
		596P	1653P	−2.8
Pgf	X80171	799P	1704P	−2.1
Itm2a	L38971	556P	981P	−1.8
Tcrb-V13[b]	X00619	8541P	14345P	−1.7
	AA608090	50A	50A	1.0
Muscle specific				
Mcm6	D86726	171P	230P	−1.3
G0s2[b]	AA036037	450A	593P/M	−1.3
	X95280	762P	610P	1.3
Nes[b]	AA057994	9100P	7427P	1.2
		7453P	4225P	1.8
H19[b]	X58196	12594P	3884P	3.2
		10124P	3062P	3.3
Itgb4	L04678	1244P	212P/A	5.9
Peg3[b]	U48804	2386P	175P/A	13.7
	AA172673	10315P	599P/A	17.2
Uninformative				
Ptk7	W46016	178P	218A	−1.2
Cdh6	D82029	69A	85A	−1.2
Itga7	L23423	50A	50A	1.0

[a] Call = (P)resent, (M)arginal, (A)bsent.
[b] Two distinct probesets were associated with these genes.

Several genes identified by RDA had low average difference values or "Absent" calls on the GeneChips (e.g., *Cadherin6*, *Integrin-α7*), highlighting the utility of the RDA procedure for sensitive detection of differential but low-level expression that may not be effectively detected by the probesets on the microarray. In summary, the Affymetrix array data are consistent with most of the expression data derived from Northern blot studies of clones identified in our RDA screens. We have thus provided a more complete listing of all the genes that were identified in the Affymetrix Mu11K hybridizations as having a consistent >2-fold change after pair-wise comparisons (Supplemental Table S1).

MyoD$^{-/-}$*–specific transcripts are expressed by satellite cells in vivo*

To determine whether the genes expressed predominantly by *MyoD*$^{-/-}$ myogenic cells were also expressed in

uninjured wild-type muscle and regenerating *mdx* muscle in vivo, immunohistochemistry and in situ hybridization studies were employed. Immunohistochemistry revealed expression of Vcam1 in $MyoD^{-/-}$ cells where it was concentrated at specific regions in the plasma membrane and cytoplasm (Fig. 4A). Moreover, Vcam1 was expressed in satellite cells as well as monocuclear interstitial cells in uninjured muscle (Figs. 4B and C).

IgSF4 mRNA was expressed in 2–3% of nuclei associated with uninjured wild-type muscle fibers (Figs. 5A and B), with increased expression in regenerating areas of *mdx* muscle (Figs. 5C and D). This result suggests that *IgSF4* is specifically expressed in quiescent satellite cells and their activated descendents. Transcripts for *Neuritin* were also detected in association with approximately 10% of nuclei in uninjured wild-type muscle inside as well as outside the basal lamina of muscle fibers (Fig. 5E). As with *IgSF4*, expression of *Neuritin* was markedly increased in regenerating *mdx* muscle (Fig. 5F). The satellite cell marker *Integrin-α7* was also detected in uninjured muscle, with upregulated expression during regeneration (Fig. 5H). In addition, *Hoxc10* (Fig. 5I), *TcR-β* (Fig. 5J), *Klra18*, *Itm2a*, and G_0S2 (data not shown) were all expressed in a similar pattern by cells in regenerating muscle. In situ hybridization of muscle sections with nonspecific sense riboprobes for

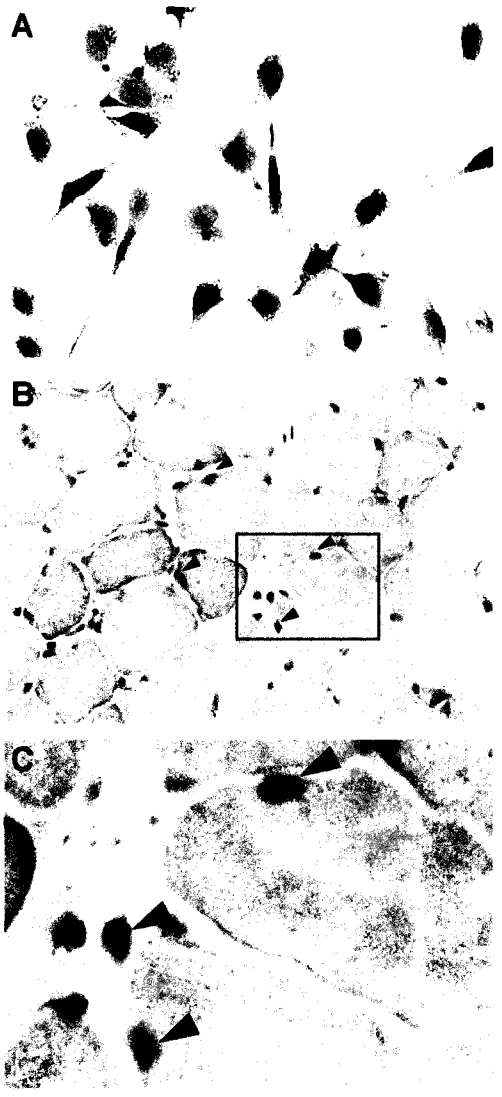


Fig. 4. Vcam1 protein is expressed in $MyoD^{-/-}$ myogenic cells and satellite cells. (A) Immunohistochemistry demonstrates localization of Vcam1 in distinct areas of the plasma membrane in $MyoD^{-/-}$ myoblasts. Hematoxylin (blue) was used to counterstain nuclei. (B) Vcam1 protein was detected in satellite cells in cross-sections of uninjured skeletal muscle (arrowheads). (C) Magnified view of boxed area in B.

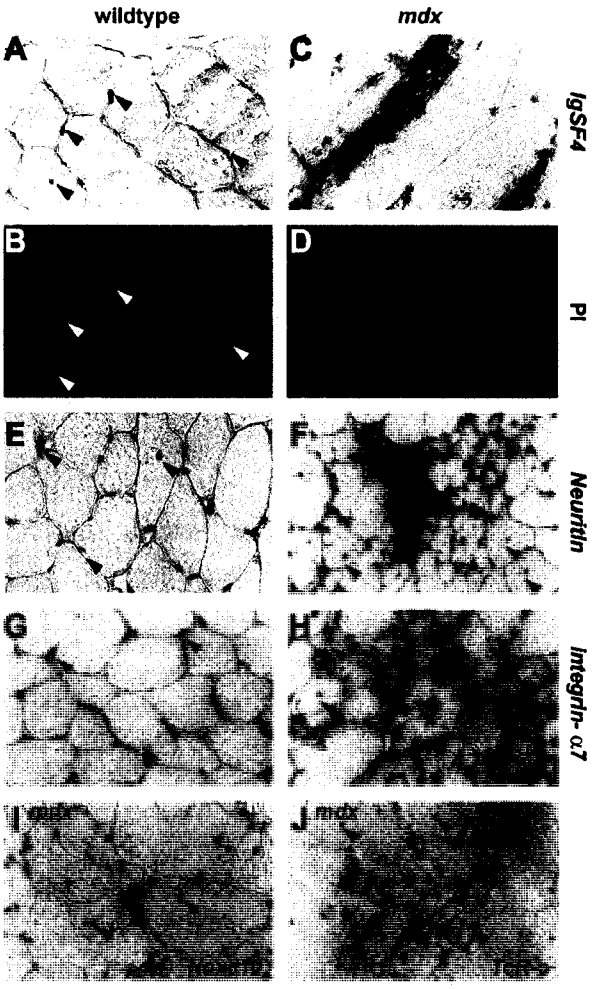


Fig. 5. $MyoD^{-/-}$-specific mRNAs are expressed by satellite cells in vivo. (A and B) *IgSF4* transcripts were detected in uninjured skeletal muscle by in situ hybridization. The expression of *IgSF4* was associated with 2–3% of the sublaminar nuclei visualized by PI staining (B). Arrowheads show *IgSF4* expressing cells and corresponding nuclei. (C and D) *IgSF4* was upregulated in regenerating *mdx* muscles demonstrating its expression in activated satellite cells and myoblasts. (E and F) *Neuritin* mRNA was expressed in uninjured (E) and regenerating skeletal muscle (F). Many *Neuritin*-expressing cells were located beneath the basal lamina of muscle fibers in uninjured muscle (E) (arrowheads) and associated with central nuclei in regenerated fibers (F) (arrowheads). (G and H) *Integrin-α7* was expressed in satellite cells in uninjured muscle (G) and upregulated during regeneration (H). (I) *Hoxc10* was highly expressed in regenerating *mdx* muscle. (J) *TcR-β* transcripts were also readily detected in regenerating muscle.

these genes did not show any nonspecific staining (not shown). Taken together, these studies suggest that many genes expressed specifically in *MyoD*$^{-/-}$ myogenic cells were also expressed by wild-type satellite cells in vivo.

Discussion

Satellite cells play a central role in postnatal muscle growth and regeneration of skeletal muscle. However, the molecular mechanisms responsible for their activation, expansion, and self-renewal remain to be elucidated. A major impediment in the study of satellite cells has been the difficulty in identifying genes expressed specifically in this compartment. In this report we have employed a straightforward experimental strategy to identify novel markers for the satellite cell lineage (Fig. 1).

Transcriptional mRNA profiling of *MyoD*$^{-/-}$ myogenic cells by representational difference analysis revealed that markers expressed by quiescent satellite cells were also expressed in *MyoD*-deficient cells (Fig. 5). This indicates that *MyoD*$^{-/-}$ myogenic cells represent a unique and tractable in vitro system to identify genetic pathways important for the function and maintenance of the satellite stem cell compartment in adult skeletal muscle.

Interestingly, several cDNAs expressed specifically in *MyoD*$^{-/-}$ myogenic cells were also expressed in hematopoietic or endothelial cells including *Vascular cell adhesion molecule-1* (*Vcam1*), *Asb5*, *Immunoglobulin superfamily-4* (*IgSF4*), *T-cell-receptor β-chain* (*TcR-β*), *Killer Lectin Receptor-18* (*Klra18*), *Laminin-α5* (*Lama5*), *Placenta-derived growth factor* (*PlGF*), and *Cadherin-6* (see references given in Table 1). This result supports the suggested developmental relationship between satellite cells and hematopoietic–endothelial lineages that has been proposed previously (De Angelis et al., 1999; Ordahl, 1999).

The developmental origin of satellite cells remains unclear. Quail-chick grafting experiments originally suggested that satellite cells originated from the somitic mesoderm along with embryonic and fetal myogenic cells (Armand et al., 1983). However, De Angelis et al. (1999) described the isolation of satellite cell myogenic clones from explanted dorsal aorta but not from the somites of mouse embryos. Notably, the aorta-derived myogenic cells express markers characteristic of the hematopoietic and endothelial cell lineages. The stem cells associated with blood vessels termed mesangioblasts are multipotential and readily participate in the repair of dystrophic muscle (Minasi et al., 2002). In the present study, we demonstrate that several genes expressed in hematopoietic lineages are also expressed by adult muscle satellite cells in vivo. These findings are consistent with the hypothesis that progenitors associated with blood vessels can give rise to adult muscle satellite cells.

The use of both wild-type myoblasts and *MyoD*-deficient myogenic progenitors in our study identified distinct subsets of genes based on their expression during in vitro satellite cell myogenesis (see Table 1). For example, the paired-box transcription factor *Pax7* was identified as a gene expressed specifically in proliferating myogenic cells. Based on its expression profile defined by this study, we then investigated the role of Pax7 in satellite cells (Seale et al., 2000). Importantly, analyses of *Pax7*-deficient mice revealed a requirement for Pax7 in satellite cell ontogeny. The requirement for Pax7 in the development of satellite cells is the basis of many ongoing studies aimed at identifying the origin of satellite cells and molecular pathways involved in their specification and function. The cloning and functional analysis of *Pax7* in adult progenitors highlights the importance of identifying differentially expressed genes in refined progenitor cell populations.

The Ankyrin repeat SOCS box containing gene *Asb5* was expressed at high levels in wild-type myoblasts and *MyoD*$^{-/-}$ myogenic cells as well as skeletal muscle tissue (Fig. 2A). *Asb5* was independently identified as a gene expressed in smooth muscle and endothelial cells during arteriogenesis (Boengler et al., 2003). Interestingly, immunolocalization of Asb5 protein in skeletal muscle revealed its expression in muscle satellite cells (Boengler et al., 2003). *Pb99* was expressed specifically in proliferating myoblasts and was completely downregulated after 24 h in differentiation conditions (Fig. 2A). Pb99, a putative G-protein-coupled receptor, is also expressed in pre-B cells and thymocytes but not in mature lymphocytes (Sleckman et al., 2000).

Several other genes identified in wild-type satellite cell-derived myoblasts were expressed in a differentiation-dependent manner, including *Synaptotagmin-like 2 (Sytl2)*, *Death-associated kinase 2 (Dapk2)*, and *Neuronal Leucine Rich Region-1* (Lrrn1) (Fig. 2B and Table 1). This subset of genes was expressed at lower levels in *MyoD*-deficient myogenic cells, suggesting specific roles in the differentiation process downstream of MyoD. In particular, an unknown gene related to *Arachidonate 5′-Lipoxygenase* (5-Lox) (dp3-5) was expressed in proliferating and differentiating wild-type and C2C12 myoblasts, but notably absent from *MyoD*$^{-/-}$ cultures (Fig. 2C). The expression pattern of dp3-5 is suggestive of a role for this gene in myogenic differentiation that will require further investigation. Taken together, several genes with previously undescribed roles in myogenesis were identified in wild-type myoblasts.

The increased self-renewal capacity and enhanced proliferation of *MyoD*-deficient myoblasts suggested that their gene expression profile would closely resemble that of satellite cells in vivo. The identification of known satellite cell markers including *Vcam1* (Figs. 3 and 5) and *Integrin-α7* as genes expressed highly in *MyoD*-deficient myoblasts illustrates that *MyoD*$^{-/-}$ cells express genes that are also expressed by satellite cells but that are low or absent in wild-type myoblasts. Moreover, transcripts for *IgSF4* and *Neuritin*, genes not expressed in wild-type myoblasts, were detected in cells within uninjured skeletal muscle (Fig. 5). The frequency of cells expressing *IgSF4* and *Neuritin* and

their overall expression levels were markedly increased in regenerating muscle consistent with their expression in muscle satellite cells. Several other genes expressed at high levels in *MyoD*$^{-/-}$ myogenic cells including *Hoxc10*, *TcRβ*, *Klra18*, and *G_0S2* were similarly activated during muscle regeneration (Figs. 5I and J). These results indicate a specific induction of these markers during muscle regeneration and support the hypothesis that MyoD-:Myf5+ myogenic progenitors represent a distinct developmental stage in the satellite cell program during regenerative myogenesis.

IgSF4 is a recently identified member of the immunoglobulin protein superfamily (IgSF) of proteins related to *Synaptic Cell Adhesion Molecule* (SCAM) that functions at neuronal synapses (Biederer et al., 2002). *IgSF* genes encode a diverse group of proteins characterized by the Ig homology domain that regulate several processes including cell adhesion and signal transduction cascades (reviewed by Rougon and Hobert, 2003). Neuritin is a GPI-anchored protein that is highly expressed in the brain and induced in response to neural activity (Naeve et al., 1997). Nestin is an intermediate filament protein that serves as a marker for neural stem cells and adult pancreatic stem cells (Lendahl et al., 1990; Sawamoto et al., 2001; Zimmerman et al., 1994; Zulewski et al., 2001). The function of these proteins in the myogenic lineage remains to be defined.

Hox genes have been implicated in embryonic growth and pattern formation. Notably, *Hoxc10* is specifically expressed in the developing hindlimbs but not forelimbs of the Axolotl (salamander), mouse, and chick (Carlson et al., 2001; Peterson et al., 1992, 1994). *Hoxc10* is also induced early during the regeneration of Axolotl forelimbs coincident with the appearance of the undifferentiated blastema cells, suggesting a role for Hoxc10 in dedifferentiation (Carlson et al., 2001). The expression of *Hoxc10* in undifferentiated amphibian limb progenitors is consistent with its expression in "more primitive" MyoD-negative myogenic precursors during muscle regeneration. Further studies are required to assess the functional role of *Hoxc10* in skeletal muscle development and regeneration.

Analysis of gene expression in satellite cells by RT-PCR demonstrates that either *Myf5* or *MyoD* is induced upon activation before their coexpression in committed myogenic precursors (Cornelison and Wold, 1997). Analysis of regenerating muscle confirmed these findings, indicating that 50% of activated satellite cells coexpress MyoD and Myf5, 30% express MyoD alone, and 20% express Myf5 alone 3 h postinjury (Cooper et al., 1999). Our study supports these findings and suggests a specific role for myogenic cells expressing Myf5 but not MyoD in satellite cell self-renewal. Taken together, our data indicate that cultured *MyoD*$^{-/-}$ myogenic cells provide a unique opportunity to elucidate genetic networks activated in this adult progenitor compartment.

Methodology to identify differentially expressed genes between RNA preparations has been revolutionized with the advent of spotted cDNA and oligonucleotide array technologies. These procedures are becoming increasingly robust for assessing global changes in gene expression. However, the use of independent methods like RDA remains powerful approaches to identify differentially expressed genes that are novel or are present at low levels. The ability to clone novel genes and tailor cDNA subtractions by altering input cDNA ratios is not possible with microarray methodology. Furthermore, the sensitivity of RDA for identifying genes expressed at low levels (e.g., tissue-specific transcription factors) offers a critical advantage. Finally, inherent to RDA is the "physical" retrieval of differentially expressed cDNA fragments that can then be used in expression studies and screening of cDNA or genomic libraries. In this study, we have successfully identified several specific genes using RDA that will now form the basis of further research into the biology of satellite cells. Future functional and genetic analyses of the genes identified in this survey will help elucidate the mechanisms acting during regenerative myogenesis.

Acknowledgments

The authors thank Adele Girgis-Gabardo for expert technical assistance with myoblast cultures and Luc A. Sabourin for expert technical advice. P.S was supported by a Doctoral Research Award from the Canadian Institutes of Health Research. M.A.R. holds the Canada Research Chair in Molecular Genetics and is a Howard Hughes Medical Institute International Scholar. This work was supported by grants to M.A.R. from the Muscular Dystrophy Association, the National Institutes of Health, the Canadian Institutes of Health Research, the Howard Hughes Medical Institute, and the Canada Research Chair Program.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ydbio.2004.07.034.

References

Appell, H.J., Forsberg, S., Hollmann, W., 1988. Satellite cell activation in human skeletal muscle after training: evidence for muscle fiber neoformation. Int. J. Sports Med. 9, 297–299.

Armand, O., Boutineau, A.M., Mauger, A., Pautou, M.P., Kieny, M., 1983. Origin of satellite cells in avian skeletal muscles. Arch. Anat. Microsc. Morphol. Exp. 72, 163–181.

Bachner, D., Ahrens, M., Schroder, D., Hoffmann, A., Lauber, J., Betat, N., Steinert, P., Flohe, L., Gross, G., 1998. Bmp-2 downstream targets in mesenchymal development identified by subtractive cloning from recombinant mesenchymal progenitors (C3H10T1/2). Dev. Dyn. 213, 398–411.

Bartolomei, M.S., Webber, A.L., Brunkow, M.E., Tilghman, S.M., 1993. Epigenetic mechanisms underlying the imprinting of the mouse H19 gene. Genes Dev. 7, 1663–1673.

Biederer, T., Sara, Y., Mozhayeva, M., Atasoy, D., Liu, X., Kavalali, E.T., Sudhof, T.C., 2002. SynCAM, a synaptic adhesion molecule that drives synapse assembly. Science 297, 1525–1531.

Birnboim, H.C., 1988. Rapid extraction of high molecular weight RNA from cultured cells and granulocytes for Northern analysis. Nucleic Acids Res. 16, 1487–1497.

Bischoff, R., 1994. The satellite cell and muscle regeneration. In: Engel, A.G., Franszini-Armstrong, C. (Eds.), Myogenesis, vol. 2. McGraw-Hill, New York, pp. 97–118.

Blanco-Bose, W.E., Yao, C.C., Kramer, R.H., Blau, H.M., 2001. Purification of mouse primary myoblasts based on alpha 7 integrin expression. Exp. Cell Res. 265, 212–220.

Boengler, K., Pipp, F., Fernandez, B., Richter, A., Schaper, W., Deindl, E., 2003. The ankyrin repeat containing SOCS box protein 5: a novel protein associated with arteriogenesis. Biochem. Biophys. Res. Commun. 302, 17–22.

Braissant, O., Wahli, W., 1998. A simplified in situ hybridization protocol using non-radioactively labelled probes to detect abundant and rare mRNAs on tissue sections. Biochemica 1, 10–16.

Carlsson, L., Li, Z., Paulin, D., Thornell, L.E., 1999. Nestin is expressed during development and in myotendinous and neuromuscular junctions in wild type and desmin knock-out mice. Exp. Cell Res. 251, 213–223.

Carlson, M.R., Komine, Y., Bryant, S.V., Gardiner, D.M., 2001. Expression of Hoxb13 and Hoxc10 in developing and regenerating Axolotl limbs and tails. Dev. Biol. 229, 396–406.

Chan, P.Y., Takei, F., 1989. Molecular cloning and characterization of a novel murine T cell surface antigen, YE1/48. J. Immunol. 142, 1727–1736.

Coletti, D., Yang, E., Marazzi, G., Sassoon, D., 2002. TNFalpha inhibits skeletal myogenesis through a PW1-dependent pathway by recruitment of caspase pathways. EMBO J. 21, 631–642.

Cooper, R.N., Tajbakhsh, S., Mouly, V., Cossu, G., Buckingham, M., Butler-Browne, G.S., 1999. In vivo satellite cell activation via Myf5 and MyoD in regenerating mouse skeletal muscle. J. Cell Sci. 112, 2895–2901.

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.

Cornelison, D.D., Olwin, B.B., Rudnicki, M.A., Wold, B.J., 2000. MyoD(−/−) satellite cells in single-fiber culture are differentiation defective and MRF4 deficient. Dev. Biol. 224, 122–137.

Cornelison, D.D., Filla, M.S., Stanley, H.M., Rapraeger, A.C., Olwin, B.B., 2001. Syndecan-3 and syndecan-4 specifically mark skeletal muscle satellite cells and are implicated in satellite cell maintenance and muscle regeneration. Dev. Biol. 239, 79–94.

Darr, K.C., Schultz, E., 1987. Exercise-induced satellite cell activation in growing and mature skeletal muscle. J. Appl. Physiol. 63, 1816–1821.

De Angelis, L., Berghella, L., Coletta, M., Lattanzi, L., Zanchi, M., Cusella-De Angelis, M.G., Ponzetto, C., Cossu, G., 1999. Skeletal myogenic progenitors originating from embryonic dorsal aorta coexpress endothelial and myogenic markers and contribute to postnatal muscle growth and regeneration [see comments]. J. Cell Biol. 147, 869–878.

Deleersnijder, W., Hong, G., Cortvrindt, R., Poirier, C., Tylzanowski, P., Pittois, K., Van Marck, E., Merregaert, J., 1996. Isolation of markers for chondro-osteogenic differentiation using cDNA library subtraction. Molecular cloning and characterization of a gene belonging to a novel multigene family of integral membrane proteins. J. Biol. Chem. 271, 19475–19482.

Duclos, F., Straub, V., Moore, S.A., Venzke, D.P., Hrstka, R.F., Crosbie, R.H., Durbeej, M., Lebakken, C.S., Ettinger, A.J., van der Meulen, J., Holt, K.H., Lim, L.E., Sanes, J.R., Davidson, B.L., Faulkner, J.A., Williamson, R., Campbell, K.P., 1998. Progressive muscular dystrophy in alpha-sarcoglycan-deficient mice. J. Cell Biol. 142, 1461–1471.

Garry, D.J., Yang, Q., Bassel-Duby, R., Williams, R.S., 1997. Persistent expression of MNF identifies myogenic stem cells in postnatal muscles. Dev. Biol. 188, 280–294.

Grounds, M.D., Yablonka-Reuveni, Z., 1993. Molecular and cell biology of skeletal muscle regeneration. Mol. Cell Biol. Hum. Dis. Ser. 3, 210–256.

Gu, Y., Sorokin, L., Durbeej, M., Hjalt, T., Jonsson, J.I., Ekblom, M., 1999. Characterization of bone marrow laminins and identification of alpha5-containing laminins as adhesive proteins for multipotent hematopoietic FDCP-Mix cells. Blood 93, 2533–2542.

Hawke, T.J., Garry, D.J., 2001. Myogenic satellite cells: physiology to molecular biology. J. Appl. Physiol. 91, 534–551.

Hubank, M., Schatz, D.G., 1994. Identifying differences in mRNA expression by representational difference analysis of cDNA. Nucleic Acids. Res. 22, 5640–5648.

Kawai, T., Matsumoto, M., Takeda, K., Sanjo, H., Akira, S., 1998. ZIP kinase, a novel serine/threonine kinase which mediates apoptosis. Mol. Cell. Biol. 18, 1642–1651.

Kimura, H., Ohtomo, T., Yamaguchi, M., Ishii, A., Sugimoto, K., 1996. Mouse MCM proteins: complex formation and transportation to the nucleus. Genes Cells 1, 977–993.

LaBarge, M.A., Blau, H.M., 2002. Biological progression from adult bone marrow to mononucleate muscle stem cell to multinucleate muscle fiber in response to injury. Cell 111, 589–601.

Leibovitch, M.P., Solhonne, B., Guillier, M., Verrelle, P., Leibovitch, S.A., Verelle, P., 1995. Direct relationship between the expression of tumor suppressor H19 mRNA and c-mos proto-oncogene during myogenesis. Oncogene 10, 251–260.

Lendahl, U., Zimmerman, L.B., McKay, R.D., 1990. CNS stem cells express a new class of intermediate filament protein. Cell 60, 585–595.

Liu, L., Vachon, P.H., Kuang, W., Xu, H., Wewer, U.M., Kylsten, P., Engvall, E., 1997. Mouse adhalin: primary structure and expression during late stages of muscle differentiation in vitro. Biochem. Biophys. Res. Commun. 235, 227–235.

Lluis, F., Roma, J., Suelves, M., Parra, M., Aniorte, G., Gallardo, E., Illa, I., Rodriguez, L., Hughes, S.M., Carmeliet, P., Roig, M., Munoz-Canoves, P., 2001. Urokinase-dependent plasminogen activation is required for efficient skeletal muscle regeneration in vivo. Blood 97, 1703–1711.

Luttun, A., Tjwa, M., Carmeliet, P., 2002. Placental growth factor (PlGF) and its receptor Flt-1 (VEGFR-1): novel therapeutic targets for angiogenic disorders. Ann. N. Y. Acad. Sci. 979, 80–93.

Maniatis, T., Fritsch, E.F., Sambrook, J., 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marthiens, V., Padilla, F., Lambert, M., Mege, R.M., 2002. Complementary expression and regulation of cadherins 6 and 11 during specific steps of motoneuron differentiation. Mol. Cell. Neurosci. 20, 458–475.

Matsumura, F., Yamashiro, S., 1993. Caldesmon. Curr. Opin. Cell Biol. 5, 70–76.

Mauro, A., 1961. Satellite cell of skeletal muscle fibers. J. Biophys. Biochem. Cytol. 9, 493–495.

Mbalaviele, G., Nishimura, R., Myoi, A., Niewolna, M., Reddy, S.V., Chen, D., Feng, J., Roodman, D., Mundy, G.R., Yoneda, T., 1998. Cadherin-6 mediates the heterotypic interactions between the hemopoietic osteoclast cell lineage and stromal cells in a murine model of osteoclast differentiation. J. Cell Biol. 141, 1467–1476.

Megeney, L.A., Kablar, B., Garrett, K., Anderson, J.E., Rudnicki, M.A., 1996. MyoD is required for myogenic stem cell function in adult skeletal muscle. Genes Dev. 10, 1173–1183.

Minasi, M.G., Riminucci, M., De Angelis, L., Borello, U., Berarducci, B., Innocenzi, A., Caprioli, A., Sirabella, D., Baiocchi, M., De Maria, R., Boratto, R., Jaffredo, T., Broccoli, V., Bianco, P., Cossu, G., 2002. The meso-angioblast: a multipotent, self-renewing cell that originates from

the dorsal aorta and differentiates into most mesodermal tissues. Development 129, 2773–2783.

Murgia, C., Blaikie, P., Kim, N., Dans, M., Petrie, H.T., Giancotti, F.G., 1998. Cell cycle and adhesion defects in mice carrying a targeted deletion of the integrin beta4 cytoplasmic domain. EMBO J. 17, 3940–3951.

Naeve, G.S., Ramakrishnan, M., Kramer, R., Hevroni, D., Citri, Y., Theill, L.E., 1997. Neuritin: a gene induced by neural activity and neurotrophins that promotes neuritogenesis. Proc. Natl. Acad. Sci. U. S. A. 94, 2648–2653.

Ordahl, C.P., 1999. Myogenic shape-shifters [comment]. J. Cell Biol. 147, 695–698.

Park, S.K., Lee, H.S., Lee, S.T., 1996. Characterization of the human full-length PTK7 cDNA encoding a receptor protein tyrosine kinase-like molecule closely related to chick KLG. J. Biochem. (Tokyo) 119, 235–239.

Peterson, R.L., Jacobs, D.F., Awgulewitsch, A., 1992. Hox-3.6: isolation and characterization of a new murine homeobox gene located in the 5′ region of the Hox-3 cluster. Mech. Dev. 37, 151–166.

Peterson, R.L., Papenbrock, T., Davda, M.M., Awgulewitsch, A., 1994. The murine Hoxc cluster contains five neighboring AbdB-related Hox genes that show unique spatially coordinated expression in posterior embryonic subregions. Mech. Dev. 47, 253–260.

Relaix, F., Weng, X., Marazzi, G., Yang, E., Copeland, N., Jenkins, N., Spence, S.E., Sassoon, D., 1996. Pw1, a novel zinc finger gene implicated in the myogenic and neuronal lineages. Dev. Biol. 177, 383–396.

Ringelmann, B., Roder, C., Hallmann, R., Maley, M., Davies, M., Grounds, M., Sorokin, L., 1999. Expression of laminin alpha1, alpha2, alpha4, and alpha5 chains, fibronectin, and tenascin-C in skeletal muscle of dystrophic 129ReJ dy/dy mice. Exp. Cell Res. 246, 165–182.

Robertson, E.J., 1987. Embryo-derived stem cells. In: Robertson, E.J. (Ed.), Teratomas and Embryonic Stem Cells: A Practical Approach. IRL Press, Oxford, UK, pp. 71–112.

Rosen, G.D., Sanes, J.R., LaChance, R., Cunningham, J.M., Roman, J., Dean, D.C., 1992. Roles for the integrin VLA-4 and its counter receptor VCAM-1 in myogenesis. Cell 69, 1107–1119.

Rosenblatt, J.D., Yong, D., Parry, D.J., 1994. Satellite cell activity is required for hypertrophy of overloaded adult rat muscle. Muscle Nerve 17, 608–613.

Rougon, G., Hobert, O., 2003. New insights into the diversity and function of neuronal immunoglobulin superfamily molecules. Annu. Rev. Neurosci. 26, 207–238.

Rudnicki, M.A., Braun, T., Hinuma, S., Jaenisch, R., 1992. Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. Cell 71, 383–390.

Sabourin, L.A., Girgis-Gabardo, A., Seale, P., Asakura, A., Rudnicki, M.A., 1999. Reduced differentiation potential of primary *MyoD*$^{-/-}$ myogenic cells derived from adult skeletal muscle. J. Cell Biol. 144, 631–643.

Sawamoto, K., Nakao, N., Kakishita, K., Ogawa, Y., Toyama, Y., Yamamoto, A., Yamaguchi, M., Mori, K., Goldman, S.A., Itakura, T., Okano, H., 2001. Generation of dopaminergic neurons in the adult brain from mesencephalic precursor cells labeled with a nestin-GFP transgene. J. Neurosci. 21, 3895–3903.

Schultz, E., 1989. Satellite cell behavior during skeletal muscle growth and regeneration. Med. Sci. Sports Exerc. 21, S181–S186.

Schultz, E., 1996. Satellite cell proliferative compartments in growing skeletal muscles. Dev. Biol. 175, 84–94.

Schultz, E., Jaryszak, D.L., Valliere, C.R., 1985. Response of satellite cells to focal skeletal muscle injury. Muscle Nerve 8, 217–222.

Seale, P., Rudnicki, M.A., 2000. A new look at the origin, function, and "stem-cell" status of muscle satellite cells. Dev. Biol. 218, 115–124.

Seale, P., Sabourin, L.A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., Rudnicki, M.A., 2000. Pax7 is required for the specification of myogenic satellite cells. Cell 102, 777–786.

Sleckman, B.P., Khan, W.N., Xu, W., Bassing, C.H., Malynn, B.A., Copeland, N.G., Bardon, C.G., Breit, T.M., Davidson, L., Oltz, E.M., Jenkins, N.A., Berman, J.E., Alt, F.W., 2000. Cloning and functional characterization of the early-lymphocyte-specific Pb99 gene. Mol. Cell. Biol. 20, 4405–4410.

Smith II, C.K., Janney, M.J., Allen, R.E., 1994. Temporal expression of myogenic regulatory genes during activation, proliferation, and differentiation of rat skeletal muscle satellite cells. J. Cell. Physiol. 159, 379–385.

Sorokin, L.M., Pausch, F., Frieser, M., Kroger, S., Ohage, E., Deutzmann, R., 1997. Developmental regulation of the laminin alpha5 chain suggests a role in epithelial and endothelial cell maturation. Dev. Biol. 189, 285–300.

Suelves, M., Lopez-Alemany, R., Lluis, F., Aniorte, G., Serrano, E., Parra, M., Carmeliet, P., Munoz-Canoves, P., 2002. Plasmin activity is required for myogenesis in vitro and skeletal muscle regeneration in vivo. Blood 99, 2835–2844.

Sykes, D.E., Weiser, M.M., 1995. Rat intestinal crypt-cell replication factor with homology to early S-phase proteins required for cell division. Gene 163, 243–247.

Taguchi, A., Wanaka, A., Mori, T., Matsumoto, K., Imai, Y., Tagaki, T., Tohyama, M., 1996. Molecular cloning of novel leucine-rich repeat proteins and their expression in the developing mouse nervous system. Brain Res. Mol. Brain Res. 35, 31–40.

Vaittinen, S., Lukka, R., Sahlgren, C., Hurme, T., Rantanen, J., Lendahl, U., Eriksson, J.E., Kalimo, H., 2001. The expression of intermediate filament protein nestin as related to vimentin and desmin in regenerating skeletal muscle. J. Neuropathol. Exp. Neurol. 60, 588–597.

Yablonka-Reuveni, Z., Rivera, A.J., 1994. Temporal expression of regulatory and structural muscle proteins during myogenesis of satellite cells on isolated adult rat fibers. Dev. Biol. 164, 588–603.

Yablonka-Reuveni, Z., Rudnicki, M.A., Rivera, A.J., Primig, M., Anderson, J.E., Natanson, P., 1999. The transition from proliferation to differentiation is delayed in satellite cells from mice lacking MyoD. Dev. Biol. 210, 440–455.

Zimmerman, L., Parr, B., Lendahl, U., Cunningham, M., McKay, R., Gavin, B., Mann, J., Vassileva, G., McMahon, A., 1994. Independent regulatory elements in the nestin gene direct transgene expression to neural stem cells or muscle precursors. Neuron 12, 11–24.

Zulewski, H., Abraham, E.J., Gerlach, M.J., Daniel, P.B., Moritz, W., Muller, B., Vallejo, M., Thomas, M.K., Habener, J.F., 2001. Multipotential nestin-positive stem cells isolated from adult pancreatic islets differentiate ex vivo into pancreatic endocrine, exocrine, and hepatic phenotypes. Diabetes 50, 521–533.

TABLE S1
Genes with elevated expression in MyoD-/- vs wildtype myoblasts (Avg.Diff. / Call^) (153 probesets)

		Avg. Diff. / Call^		wt/MD
Gene	**Genbank**	**wt**	**MD**	**Fold**
Foxg1	U36760	4718 P	50 A	**94.4**
Anxa8	AJ002390	4635 P	50 A	**92.7**
Col1a1	U08020	7161 P	120 P/A	**59.9**
Ptpn8	M90388	1915 P	50 A	**38.3**
Cx3cl1	U92565	1938 P	54 A	**36.2**
Col1a1	U08020	8250 P	244 P	**33.8**
Cx3cl1	U92565	3721 P	127 A	**29.3**
Fbln2	X75285	10794 P	369 P	**29.3**
Ctgf	M70642	6803 P	236 P/A	**28.9**
Lhx8	D49658	1086 P/M	50 A	**21.7**
Bdnf	X55573	2160 P	106 P	**20.5**
Fgf7	Z22703	995 P	50 A	**19.9**
Ndn	C79136	3429 P	210 P	**16.3**
Syk	Z49877	695 P	50 A	**13.9**
?	ET62118	681 P	50 P/A	**13.6**
Rab3b	AA166533	703 P	53 A	**13.4**
Anxa8	AA060106	2061 P	158 P/A	**13.1**
Mgst3	AA518117	5484 P	422 A	**13.0**
Apbb1ip	AF020313	2115 P	183 P	**11.6**
?	X66976	3694 P	333 P	**11.1**
Ltbp2	AF004874	3579 P	326 P	**11.0**
Ly6a	M18184	7606 P	705 P/A	**10.8**
Pold4	AA220798	5210 P	505 P/A	**10.3**
?	AA562685	3678 P	363 A	**10.1**
Kitl	U44725	854 P	85 P	**10.1**
Nfib	U57634	797 P	83 P	**9.6**
Anxa8	AA032354	491 P	52 A	**9.5**
Anxa3	AA003383	2912 P	320 P/A	**9.1**
Pam	U79523	3056 P	340 P/A	**9.0**
Ifi1	U19119	591 P	70 P/A	**8.4**
Ugt1a1	L02333	420 P	50 M/A	**8.4**
Itih2	X70392	414 P/M	50 A	**8.3**
S100a13	X99921	2105 P	256 P/A	**8.2**
1110007C09Rik	AA289353	913 P	113 P	**8.1**

		Avg. Diff. / Call^		wt/MD
Gene	**Genbank**	**wt**	**MD**	**Fold**
Aqp1	L02914	1471 P	184 P	**8.0**
Figf	X99572	441 P	55 A	**8.0**
Gas2	M21828	821 P	103 P	**8.0**
Mglap	D00613	8031 P	1015 P/A	**7.9**
Ly6a	AA097051	2081 P	272 A	**7.7**
Itgb5	W14823	549 P	72 A	**7.6**
Pmp22	M32240	3740 P	515 P	**7.3**
Igfbp7	L75822	12488 P	1800 P	**6.9**
Ctsh	U06119	2399 P	349 P	**6.9**
Mmp24	AA164044	421 P	63 P	**6.7**
Itga6	X69902	3537 P	541 P	**6.5**
?	AA273574	314 P	50 A	**6.3**
Dlx1	U51000	1601 P	259 P/M	**6.2**
Twist1	M63650	1065 P	175 P/A	**6.1**
Timp2	X62622	6689 P	1159 P	**5.8**
Meis1	U33630	594 P	104 A	**5.7**
Rbl2	U36799	304 P	54 A	**5.6**
AI429613	AA267483	281 P	50 P/A	**5.6**
Penk1	M13227	308 P	55 A	**5.6**
Sqrdl	AA245216	1321 P	237 P/A	**5.6**
?	C78076	2881 P	519 P	**5.6**
Cd34	AA064307	275 P	50 A	**5.5**
Cd34	AA000252	802 P	146 A	**5.5**
Bgn	L20276	21379 P	4048 P/A	**5.3**
Gsn	J04953	2407 P	463 P	**5.2**
Sh3d4	U58889	255 P	50 A	**5.1**
Cxcl12	L12030	5959 P	1184 P	**5.0**
Anxa5	U29396	5253 P	1059 P	**5.0**
B130024L21Rik	AA239477	766 P	156 M	**4.9**
1110048B16Rik	W77105	1211 P	249 P/A	**4.9**
Xdh	X75129	360 P	76 A	**4.8**
Sec8	AF022962	552 P	116 P	**4.8**
Nr2f1	X74134	234 P	50 P/A	**4.7**
Vcam1	X67783	4612 P	1037 P	**4.4**
Calm3	M19380	220 P	50 A	**4.4**
Hexb	Y00964	1789 P	407 P	**4.4**
H2-T23	M11284	450 P	104 A	**4.3**

		Avg. Diff. / Call^		wt/MD
Gene	**Genbank**	**wt**	**MD**	**Fold**
Itgb5	W99005	1154 P	271 P	**4.3**
Cd53	X97227	518 P	122 P	**4.2**
Rab14	W82651	1251 P	296 P	**4.2**
Cd53	Z16078	1290 P	313 P	**4.1**
Capg	X54511	4131 P	1004 P	**4.1**
Il4ra	M27960	1586 P	386 P	**4.1**
2210023F24Rik	C76523	1066 P	260 P	**4.1**
Cd53	AA105582	1126 P	276 P	**4.1**
Pea15	AA108330	864 P	213 P/A	**4.1**
Lama5	U37501	731 P	182 P/A	**4.0**
Cd34	W65699	195 P	50 A	**3.9**
Pdgfrb	X04367	789 P	206 P/A	**3.8**
Anxa1	M24554	230 P	60 A	**3.8**
Ifit2	U43085	258 P	68 A	**3.8**
Itgb5	AA142328	720 P	191 P/M	**3.8**
Eps8	L21671	2981 P	792 P	**3.8**
Ifngr2	U69599	233 P	63 P/A	**3.7**
Tm4sf1	L15443	15268 P	4110 P	**3.7**
Pccb	AA237535	208 P	56 A	**3.7**
Dhx16	AA119035	264 P	73 P	**3.6**
Dnajc3	AA265871	1457 P	403 P	**3.6**
Synpo	AA182325	181 P/M	50 A	**3.6**
Neo1	W89821	393 P	110 P	**3.6**
Retsdr2-pending	W20721	965 P	271 P	**3.6**
Anxa4	AA108947	578 P	163 P	**3.6**
Ifi47	M63630	1181 P	336 A	**3.5**
Nucb2	W62646	184 P	53 A	**3.5**
Nid1	X14194	3472 P	1021 P	**3.4**
Atp6v1a1	U13837	584 P	174 P	**3.4**
Klra1	M25812	167 P	50 A	**3.3**
?	W65256	458 P	139 P	**3.3**
Pcolce	X57337	13900 P	4224 P	**3.3**
Lxn	D88769	3214 P	983 P	**3.3**
Tm9sf2	AA117787	796 P	244 P	**3.3**
Rcbtb1	AA689761	210 P	65 P	**3.3**
Tnfsf9	L15435	193 P/M	60 P	**3.2**
Camk2d	W09664	304 P	95 P	**3.2**

		Avg. Diff. / Call^		wt/MD
Gene	**Genbank**	**wt**	**MD**	**Fold**
ank	AA529805	15188 P	4857 P	**3.1**
ank	AA163465	6348 P	2037 P	**3.1**
Crat	X85983	460 P	148 P/A	**3.1**
Cd14	X13333	2749 P	896 P	**3.1**
Nat2	U35886	183 P	60 A	**3.0**
F3	M26071	152 P	50 A	**3.0**
2410019G02Rik	AA684048	1300 P	428 P	**3.0**
Sema3c	X85994	152 P	50 A	**3.0**
Anxa5	W98864	3519 P	1172 P	**3.0**
Cfh	M12660	150 P	50 A	**3.0**
?	AA231147	414 P	139 P/A	**3.0**
?	AA450768	1777 P	597 P	**3.0**
Ltbp3	L40459	3224 P	1084 P	**3.0**
Stk24	AA154321	3251 P	1099 P	**3.0**
Fts	X71978	405 P	138 P	**2.9**
Matn2	U69262	231 P	79 P	**2.9**
Lamc1	J02930	2436 P	836 P	**2.9**
Aes	X73359	4908 P	1698 P	**2.9**
Anxa4	AA142796	486 P	170 P	**2.9**
Iqgap1	AA118739	2942 P	1027 P	**2.9**
MGC47046	AA138866	146 P	52 P	**2.8**
Mpp1	U38196	2270 P	815 P	**2.8**
4931426K16Rik	AA611940	326 P	117 P/A	**2.8**
5730420B22Rik	C77647	499 P	181 P	**2.8**
Kpnb3	AA409333	1427 P	519 P	**2.8**
Kpnb3	AA120563	1552 P	579 P	**2.7**
A430081P20Rik	AA409056	223 P	84 P	**2.7**
Nr2f2	X76653	231 P	89 P	**2.6**
Grn	M86736	2555 P	991 P	**2.6**
Anp32a	W91701	787 P	305 P	**2.6**
Anxa1	X07486	15421 P	6048 P	**2.5**
Dnajc3	U28423	489 P	195 P/A	**2.5**
Myln	AA107052	4941 P	1988 P	**2.5**
Lamp1	M25244	8824 P	3569 P	**2.5**
2210023F24Rik	C76523	881 P	361 P/A	**2.4**
?	D19038	1367 P	565 P/A	**2.4**
Camk1	W30101	124 P	52 P/A	**2.4**

		Avg. Diff. / Call^		wt/MD
Gene	**Genbank**	**wt**	**MD**	**Fold**
Rab7	X89650	9001 P	3744 P	**2.4**
Cd9	C80730	973 P	407 P	**2.4**
9030221M09Rik	AA682033	1007 P	426 P	**2.4**
Jak1	AA561503	508 P	219 P	**2.3**
4921505F14Rik	AA123450	1480 P	644 P	**2.3**
Dpp7	AA122717	435 P	196 P	**2.2**
Gkap42-pending	W98992	121 P	55 P/A	**2.2**
9130415E20Rik	AA289168	1359 P	627 P	**2.2**

^ Call = (P)resent, (M)arginal, (A)bsent

* Two distinct probesets were associated with these genes

An arbitrary floor Average Difference of 50 was used.